Methods for calculating the enthalpies of sublimation of organic molecular crystals

E. A. Arnautova, M. V. Zakharova, T. S. Pivina, E. A. Smolenskii, D. V. Sukhachev, and V. V. Shcherbukhin*

N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences,
47 Leninsky prosp., 117913 Moscow, Russian Federation.
Fax: 007 (095) 135 5328

Four methods for calculating the enthalpies of sublimation of organic molecular crystals based on different methodological approaches are proposed. Comparative analysis of these methods was carried out and their good predictive ability with respect to various classes of compounds (aliphatic nitro compounds, acyclic and cyclic amines, amides and amino acids, benzene derivatives, heterocyclic compounds, *etc.*) was demonstrated.

Key words: enthalpy of sublimation, additivity approach; QSPR approach; descriptors; regression analysis, trigonometric polynomials, atom—atom potential functions.

The enthalpy of sublimation is one of the important characteristics of molecular crystals. Existing methods for the calculation of this enthalpy are relatively few in number and refer mostly to hydrocarbons.

Previously, theoretical estimations of the enthalpies of sublimation ΔH_{subl} of molecular crystals have been carried out most frequently using the additivity method, based on the assumption that the magnitude of a physicochemical characteristic Y can be represented as the sum of the partial contributions of individual fragments:

$$Y = \Sigma n_i X_i, \qquad (1)$$

where n_i is the number of fragments of an ith sort in a molecule, and X_i is the contribution of the corresponding fragment.

To carry out calculations using the additivity approach, first of all, X_i values must be found from the available experimental data using the least-squares method. Then these values are used to evaluate the enthalpies of sublimation of compounds, for which this value has not been found experimentally. The best known scheme of this type was proposed by A. Bondi.[1]

According to the principle of corresponding states, the standard heat of sublimation is equal to the heat of sublimation at the lowest temperature of a first-order phase transition. At this temperature, the crystalline state of a substance is closest to its liquid state, and the enthalpy of sublimation can be estimated as the sum of the contributions of individual fragments:

$$\Delta H_{\mathrm{subl}} = \Sigma n_i \Delta H_{\mathrm{subl},i},$$

where $\Delta H_{\mathrm{subl},i}$ is the contribution of the ith fragment; n_i is the number of fragments of the ith type in the molecule.

Various classes of hydrocarbons (branched and nonbranched alkanes, cycloalkanes, aromatic compounds, and their alkyl derivatives) have been considered in detail.[1] The group increments that have been proposed[1] for hydrocarbons are fairly accurate, because they are based on the selection and analysis of a great amount of reliable experimental data. The group increments for other classes of compounds are less reliable, since they have been calculated using data for a small number of substances, and sometimes, for only one substance.

The additivity approach suggested by A. Bondi has been developed further.[2—4] Enthalpy of sublimation is known to be defined as the difference between the crystal lattice energy (E) at 0 K and the energy of the substance in the gaseous state. In the general case

$$-\Delta H_{\mathrm{subl}} = E + 2RT + 9/8 \cdot R\theta_D, \qquad (2)$$

where $2RT$ is difference between the energies of a molecule of a substance in the solid and gaseous states, and $9/8 \cdot R\theta_D$ is the energy of zero-point vibrations in the Debye approximation. When the vapor pressure is low (10^{-3}—10^{-5} Torr), the enthalpy of sublimation can be considered to be equal to the lattice energy taken with opposite sign.

The method outlined in Refs. 2—4 is based on the idea that the energy of the crystal lattice is related to the free surface area of the molecule S_{m}. For S_{m}, in turn, an empirical linear relation has been suggested:

$$S_{\mathrm{m}} = 2.601 Z_n + 28.13 \ (\text{Å}^2),$$

where Z_n is the number of valence electrons in the molecule. S_{m} can also be estimated using systems of atomic and group increments for free surface area S_{ai}.[2—4]

Translated from *Izvestiya Akademii Nauk. Seriya Khimicheskaya*, No. 12, pp. 2872—2881, December, 1996.

To calculate the crystal lattice energy E, the following formulas have been used:

$$E = 0.0767S_m + 1.448 \text{ (kcal mol}^{-1}\text{)}, \quad (3)$$

$$E = 0.202Z_n + 3.472 \text{ (kcal mol}^{-1}\text{)}. \quad (4)$$

Systems of atomic increments for lattice energy[3] R_a have also been employed. According to the analysis made in Ref. 3, each atom incorporated in various compounds makes approximately equal contributions R_a, irrespective of its valence state, to the lattice energy of a molecular crystal. In close packing, molecules are arranged in the crystal so that the contribution of each atom to the energy E is maximally close to R_a.

Like the method proposed by A. Bondi,[1] the approach based on dependences (3) and (4)[2–4] for the evaluation of the enthalpy of sublimation shows good agreement with experimental results only in the case of low-polarity molecules. When strong electrostatic intermolecular interactions occur in a crystal, which is often the case for compounds containing nitrogen and oxygen atoms, the additivity approach leads to substantial errors in the estimation of ΔH_{subl}.

Another method for calculating the enthalpy of sublimation involves the use of relationships that are valid at $T = T_{\text{melt}}$:

$$\Delta H_{\text{subl}} = \Delta H_{\text{vapor}} + \Delta H_{\text{melt}}, \quad (5)$$

where ΔH_{subl} is the heat of sublimation at T_{melt}; ΔH_{vapor} is the heat of vaporization at T_{melt}; ΔH_{melt} is the heat of melting. If we pass to T = 298 K using Kirchhoff's equation, we obtain the following expressions for the heats of phase transitions:

$$\Delta H_{\text{vapor},298} = \Delta H_{\text{vapor},T_{\text{melt}}} + \int_{T_{\text{melt}}}^{298} [c_{p,\text{g}} - c_{p,\text{liq}}]dT, \quad (6)$$

$$\Delta H_{\text{subl},298} = \Delta H_{\text{subl},T_{\text{melt}}} + \int_{T_{\text{melt}}}^{298} [c_{p,\text{g}} - c_{p,\text{sol}}]dT. \quad (7)$$

By combining Eqs. (6) and (7), we obtain the following relation:

$$\Delta H_{\text{subl},298} = \Delta H_{\text{vapor},298} + \Delta H_{\text{melt}} + \int_{T_{\text{melt}}}^{298} [c_{p,\text{sol}} - c_{p,\text{liq}}]dT. \quad (8)$$

The last term in Eq. (8) makes a relatively small contribution to the heat of sublimation; therefore, it can be neglected. The second term, *viz.*, the heat of melting, amounts to 20—30% of the vaporization heat. In actual practice, approximate empirical methods are used to estimate the heats of melting. For example, rough estimates are sometimes carried out using the Walden formula:

$$\Delta H_{\text{melt}}/T_{\text{melt}} = 13\pm3 \text{ (kcal mol}^{-1}\text{ K}^{-1}\text{)}.$$

Heats of melting for some classes of compounds can be estimated using formulas derived in a number of studies.[5–7] However, at present the theory of melting of molecular crystals is still in its developmental stage. This makes wide use of the method for calculation of ΔH_{subl} based on relations (5)—(8) impossible.

In general, the additivity methods described above cannot ensure high accuracy of prediction, since ΔH_{subl} is not by nature an additive quantity. This is why the calculation schemes presented above can be used successfully only for hydrocarbons. Therefore, the purpose of the present study was to develop new methods for calculating the enthalpies of sublimation of organic molecular crystals containing N and O atoms based on known theoretical approaches.

Calculation of enthalpies of sublimation of molecular crystals

Of the existing theoretical approaches, we have chosen the four approaches that seem to be the most promising to fulfill this task: (1) the additivity approach; (2) the method of approximation of physicochemical characteristics by trigonometric polynomials; (3) the method of "structure—property" relationships (QSPR); (4) the method of atom—atom potential functions (AAPF). While we were developing methods based on these approaches for the calculation of ΔH_{subl}, we used a database composed by us and comprising experimental data for the enthalpies of sublimation of more than 280 H-, C-, N-, and O-containing compounds of various classes (aliphatic nitro compounds, cyclic and acyclic amines, amides, amino acids, benzene derivatives, heterocyclic compounds, *etc.*[8]).

1. Modification of the additivity approach. Previously,[9–11] a fairly accurate method for the calculation of additive properties of hydrocarbons and their derivatives (alcohols, ketones, benzene derivatives, halo-derivatives, *etc.*) has been proposed. However, this method is not very good for describing properties that depend not only on the structures of molecules but also on the interactions between them. Without question, the enthalpy of sublimation is one of the latter type of properties.

In this study, we attempted to take these interactions into account by considering not only the contributions of individual fragments (such as hydroxy, nitro, carbonyl, oxo, and other groups) to the enthalpy of sublimation, but also the contributions caused by their interaction. If some group X_i is able to interact with a group X_j, then, in the case where the interacting molecules are identical, the number of X_i—X_j interactions can be defined as

$$n(X_iX_j) = \min([X_i],[X_j]), \quad (9)$$

where $[X_i]$ and $[X_j]$ are the numbers of the corresponding fragments in a molecule. It was assumed that an X_i

fragment can interact with only one X_j fragment of the neighboring molecule. In addition, we took into account not only the total numbers of fragments and interactions in each structure but also their relative weights over the whole experimental database. Since the database under consideration was quite diversified with regard to fragments incorporated in molecular structures, we were unable to select a large number of fragments that were encountered in the vast majority of molecules. Therefore, the greater part of the fragments chosen had relatively small coefficients of occurrence.

Then we constructed a linear function in which the numbers of various fragments $n(X_k)$ and of their pair interactions $n(X_iX_j)$ were used as independent parameters, while the enthalpy of sublimation acted as the dependent variable:

$$\Delta H_{\mathrm{subl}} = \sum_k n(X_k)\Delta H_{\mathrm{subl}}(X_k) + \sum_{i,j} n(X_iX_j)\Delta H_{\mathrm{subl}}(X_iX_j) \ . \qquad (9a)$$

The enthalpies of sublimation obtained using formula (9a) for 205 compounds from the database mentioned above are presented in Table 1. The largest absolute error of the calculation was ~3 kcal mol^{-1}.

It should be noted that although this method is relatively accurate, it does not allow the estimation of enthalpies of sublimation of structural isomers. For example, for all dinitrobenzene isomers, the same value, ΔH_{subl} = 22.3 kcal mol^{-1}, was obtained, whereas experimental values of ΔH_{subl} for *o*-, *m*-, and *p*-dinitrobenzenes are 21.8, 20.8, and 22.2 kcal mol^{-1}, respectively. To increase the accuracy of the calculation, finer characteristics of the interactions, taking into account the differences in the structures of isomers, must be considered. In addition, the fragments chosen have small relative weights in the database used; therefore, this model cannot claim to have a high predictive ability with respect to ΔH_{subl} of compounds of other classes.

2. Approximation of the enthalpy of sublimation by trigonometric polynomials. The fundamentals of this

Table 1. Experimental and calculated enthalpies of sublimation (ΔH_{subl}/kcal mol^{-1}) of organic molecular crystals

Compound	Experimental	1[a]	2[b]	3[c]	4[d]
9-Diacetylaminoanthracene	25.4	25.1	23.4	29.6	28.4
Benzofurazan	17.1	20.6	15.8	16.5	17.2
2,4,6-Trinitroresorcinol	28.9	25.9	29.9	30.4	28.9
4-Amino-4′-nitroazobenzene	32.9	30.3	30.1	32.6	28.9
o-Hydroxybenzalaniline	19.4	21.4	23.4	17.7	21.2
Oxamide	27.0	24.1	21.6	24.8	22.7
Phthalamide	13.7	14.2	20.5	13.5	13.3
Isophthalamide	13.0	14.2	20.5	15.7	15.6
2,4,6-Trinitro-*m*-cresol	26.8	30.1	32.6	27.9	26.6
2,4,6-Trinitromesytilene	24.8	25.3	26.2	23.7	24.4
1-Azabicyclo[2.2.2]octane	12.14	14.5	12.4	13.2	13.1
Tetryl	32.0	28.9	32.7	33.1	36.8
4-(*N*,*N*-Dimethyl-4′-nitramino)azobenzene	32.2	31.8	30.3	30.3	29.3
4-(*N*,*N*-Dimethyl-3′-nitramino)azobenzene	31.9	31.8	30.3	29.4	29.5
5-Methyl-1-phenyltetrazole	25.6	25.9	25.6	23.4	23.9
Phthalonitrile	21.0	20.6	19.2	19.6	21.5
Isocytosine	33.4	30.4	31.7	32.4	35.3
2,4,6-Trinitrophenetole	28.8	27.8	25.9	27.8	29.3
Methyl anthranylate	18.7	17.9	25.8	15.4	14.4
m-Dinitrobenzene	20.8	22.3	22.9	21.7	20.7
o-Dinitrobenzene	21.0	22.3	22.9	19.6	20.7
m-Dinitrobenzene	22.8	22.7	19.5	20.3	19.8
p-Nitroaniline	24.1	25.4	22.6	24.6	21.4
m-Nitroaniline	23.1	25.4	21.6	23.7	21.8
o-Nitrophenol	20.3	20.6	19.2	21.1	20.6

(*to be continued*)

Table 1. (*continued*)

Compound	Experimental	1[a]	2[b]	3[c]	4[d]
2,4,6-Trinitroaniline	29.9	29.0	33.1	28.9	28.8
Guanine	44.5	41.4	52.9	43.9	42.9
1,3-Diamino-2,4,6-trinitrobenzene	33.5	32.4	34.7	35.1	33.3
1,3,5-Trinitro-1,3,5-triazacyclohexane	32.1	32.9	31.2	28.6	30.5
1,3,5-Triamino-2,4,6-trinitrobenzene	40.2	40.7	38.9	39.8	38.7
1,4-Diazabicyclo[2.2.2]octane	14.8	16.5	14.9	14.4	15.7
N-Methyltetrazole	21.7	20.0	19.7	21.1	20.0
2-Nitrofuran	18.0	15.4	13.6	17.6	17.3
3-Amino-4,5-dimethylisoxazole	21.0	21.9	20.2	19.3	18.9
3-Amino-5-methylisoxazole	19.5	19.5	18.1	20.9	19.8
Dimethylnitramine	16.7	16.5	17.3	16.6	17.6
Uracil	28.8	29.5	28.2	27.5	25.9
2-Nitropyrrole	19.1	20.6	20.9	22.7	18.2
3,5,7-Trinitro-1-azaadamantane	29.4	28.7	23.3	30.5	26.3
3,6-Dinitro-1-nitroso-1,3,6-triazacycloheptane	32.1	29.8	30.1	30.2	29.1
1,3,6-Trinitro-1,3,6-triazacyclohexane	29.8	32.9	31.2	31.3	27.3
2,4,6-Trimethyl-1,3,5-trinitro-1,3,5-triazacyclohexane	22.0	23.9	27.9	29.9	28.4
1,3,5-Trinitro-2-nitrosobenzene	28.6	26.4	29.9	28.9	29.5
3,5-Dinitro-1-oxa-3,5-diazacyclohexane	24.5	26.1	25.6	25.3	27.7
p-Dinitrobenzene	22.2	22.3	22.9	21.8	21.4
2,5-Dinitrotoluene	23.5	23.2	21.9	23.1	22.0
2,3-Dinitrotoluene	22.32	23.2	21.9	20.9	23.1
3,5-Dinitrotoluene	21.5	23.4	21.9	22.3	21.7
3,4-Dinitrotoluene	22.3	23.4	21.9	21.7	23.4
2,4-Dinitrophenol	25.0	22.2	24.9	24.8	24.8
4-Nitropyrazole	23.6	22.5	25.3	23.1	19.8
1,3,5-Trinitrobenzene	23.3	24.4	27.4	22.9	23.9
1,2,3-Trinitrobenzene	27.15	24.4	27.4	24.6	27.0
2,4,6-Trinitrotoluene	25.0	25.3	24.9	25.5	24.0
2,4,5-Trinitrotoluene	26.3	25.3	24.9	26.4	26.9
2,3,4-Trinitrotoluene	25.4	25.3	24.9	25.5	27.7
2,4,6-Trinitrophenol	25.1	24.3	29.7	27.2	26.8
Nitroguanidine	34.1	34.6	31.4	35.5	35.9
3-Nitrotriazole	24.4	24.4	24.0	21.2	23.5
1,3,5,7-Tetranitro-1,3,5,7-tetraazacyclooctane	39.3	39.0	33.2	34.7	34.4
cis-Azobenzene	22.2	21.7	22.8	22.2	26.4
1,2,4-Triazole	19.9	22.4	21.5	18.3	22.4
p-Nitrotoluene	18.9	21.8	18.2	19.8	19.7
L-Valine	38.9	38.3	31.9	35.6	35.4
n-Propylcarbamate	19.4	21.9	22.5	20.9	21.1
2-Aminobutyric acid	38.9	35.7	31.9	34.4	34.7
D-Alanine	34.6	35.8	28.8	33.4	34.4

(*to be continued*)

Table 1. *(continued)*

Compound	Experimental	1[a]	2[b]	3[c]	4[d]
Glycine	32.6	31.3	25.1	32.9	33.9
Butanamide	20.6	18.8	21.9	21.8	19.3
Propanamide	18.9	18.9	15.9	21.2	20.7
Acetamide	18.8	19.0	17.9	21.8	20.0
Melamine	29.5	29.9	26.5	32.1	28.3
6,9-Dimethyladenine	26.4	25.7	27.0	29.9	28.5
Adenine	30.2	30.1	26.7	29.6	30.4
Tetrazole	23.3	24.3	27.4	22.7	19.0
Phenazine	21.6	25.1	20.4	18.9	18.8
Imidazole	19.5	20.3	22.3	21.6	18.5
2,5-Dinitro-2,5-diazahexane	26.7	26.7	28.2	23.4	26.5
2,4-Dinitro-2,4-diazapentane	26.0	24.0	27.9	23.1	24.2
1-Phenyltetrazole	27.4	24.5	25.9	30.1	26.5
1,5-Diphenyltetrazole	28.7	28.9	31.9	27.5	29.3
5-Phenyltetrazole	24.8	25.9	21.9	28.3	27.8
2,3-Dinitrophenol	22.7	22.2	24.9	23.5	28.0
2,5-Dinitrophenol	21.6	22.2	20.4	24.8	26.0
1,4-Dinitropiperazine	26.6	25.9	24.5	22.9	25.9
1-Nitroadamantane	17.1	17.5	18.3	17.9	18.8
1,3-Dinitroadamantane	24.8	21.9	21.9	25.8	23.3
2,2-Dinitroadamantane	21.3	21.9	21.9	22.1	20.4
N-Methyltetrazole	21.7	20.0	24.1	21.1	20.0
Hexamethylenetetramine	17.9	20.3	18.3	18.1	18.6
γ,γ-Dinitropropylnitramine	22.9	25.5	21.9	24.6	24.7
Carbamide	21.0	23.1	23.5	22.7	19.4
1-Adamantanyl nitrate	19.1	17.5	18.9	21.4	22.9
2,5-Diphenyltetrazole	28.6	28.9	27.9	27.6	28.9
N-Methyl-*N*,*N*′-dinitroethylenediamine	21.7	21.9	26.6	25.8	27.5
Diaminofurazan	24.7	22.7	25.5	26.9	23.9
Methylnitrofuroxan	20.8	17.6	18.7	20.5	17.3
Benzofuroxan	19.6	20.6	21.7	18.2	18.4
4,6-Dinitrobenzofuroxan	27.5	26.2	29.7	30.1	27.4
5-Nitrobenzofuroxan	23.0	24.6	22.8	24.6	23.6

[a] A modified version of the additivity approach (R = 0.913). [b] The method of approximation by trigonometric polynomials (R = 0.843). [c] The method of "structure—property" correlations (R = 0.929). [d] The method of "structure—property" correlations with the use of three-dimensional descriptors (R = 0.933).

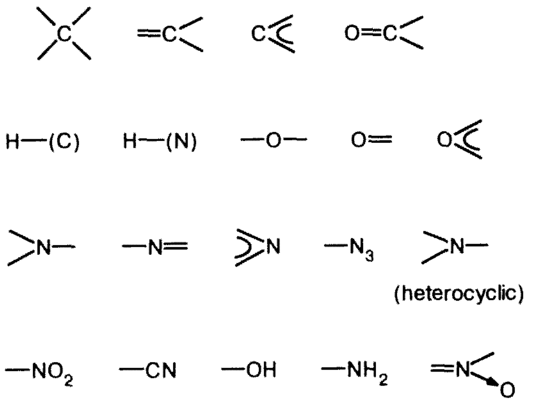
method have been published previously.[12—14] The method is based on the following idea. In terms of the additivity approach, function Y describing some physicochemical characteristic is expressed as Eq. (1), which can be regarded as a representation of Y by a linear polynomial. To increase the accuracy of the estimation of the property Y, it would be resonable to take into account also terms of higher degrees. Therefore, in calculations of the densities of nitrogen- and oxygen-containing compounds, power[12] and trigonometric[13] polynomials have been used. In addition, the approximation by trigonometric polynomials has also been used for the calculation of the enthalpies of formation of nitrogen- and oxygen-containing compounds.[14]

In this work, we attempted to estimate the enthalpies of sublimation of H-, C-, N-, and O-containing organic molecular crystals using this method. The a_0, $a^n{}_k$, and $b^n{}_k$ parameters for the required trigonometric polynomial

$$Y = a_0/2 + \sum_k \sum_n (a_k^n \cos nX_k + b_k^n \sin nX_k)$$

were calculated using a computer program (written by A. E. Evtushenko, N. D. Zelinsky Institute of Organic

Chemistry of the RAS) based on the techniques of regression analysis and successive inclusion of variables. Several sets of descriptors were considered. The best correlation coefficients for series of both reference and check compounds were obtained with the set of descriptors that included the types of atoms and atomic fragments presented below.



The reference series included 235 compounds of various classes. The best correlation coefficient in the reference series corresponded to the first harmonic (the number of parameters was 54) and was equal to 0.843 (the average absolute error was 2.4 kcal mol^{-1}). When the resulting formula was verified in relation to the check series (35 compounds), the correlation coefficient was equal to 0.797, and the average absolute error was 3.8 kcal mol^{-1}. The enthalpies of sublimation calculated for the compounds under consideration are listed in Table 1. The largest error in the calculations of ΔH_{subl} by this method was observed in the case of ethylenedinitramine and was equal to 11 kcal mol^{-1}.

It should be noted that the estimation of enthalpy of sublimation carried out using this approach is substantially less accurate than that using the additivity methods. The use of polynomials with large numbers of parameters (the fourth and higher harmonics) increases the accuracy of the calculation for compounds of the reference series; however, the number of polynomial terms can approach the number of compounds in the reference series. Consequently, estimation of ΔH_{subl} for any compound not included in this series becomes absolutely unreliable.[15] In our case, this already occurs when the third harmonic is considered.

3.1. Estimation of the enthalpy of sublimation based on "structure—property" relationships. Another method for the calculation of ΔH_{subl} of organic molecular crystals is based on the search for "structure—property" relationships, *i.e.*, on the construction of linear correlation models

$$A = K_0 + \Sigma K_i D_i + \varepsilon \qquad (10)$$

for a specified group of compounds, for which experimental values of the given property are known. In Eq. (10), A is a function of a dependent property, determined for all the compounds of the sample; ε is an unknown function of the error. All coefficients K_i of the equation are chosen by minimization of the absolute error.

It can be seen from Eq. (10) that the "structure—property" correlation is not established directly; it is expressed as a relationship between the properties and functions D_i of the structures of compounds called descriptors and found by a known algorithm. Various molecular characteristics can be used as descriptors D_i, for example, connectivity indices, solvation indices, molecular shape indices, information indices, indices based on atomic charges and electronegativities, substructural indices, *etc.*[16]

To develop a model with good predictive ability, it is desirable to use sets of descriptors that reflect as fully as possible the characteristic features of the structures of the compounds included in the available database that are responsible for the manifestation of the property under consideration. Therefore, the first stage in the construction of a linear model consists of the calculation of the set of descriptors (~3000) of various classes, possible in models of this type. The number of the resulting descriptors can exceed substantially the number of the available compounds. Therefore, in order to reduce the number of descriptors by excluding the least significant descriptors from this set, in the next stage, they are analyzed for being correlated in pairs and for being nondegenerate. Since the information contents of descriptors that are strongly correlated with one another are nearly identical, it is sufficient to include in the model only one of them, in particular, the most meaningful from the chemical viewpoint. In addition, those descriptors that are poorly correlated (singly) with the property are also eliminated from consideration. After the preliminary analysis, only a small number of descriptors (150—200) remain for the construction of the models.

Using the set of meaningful descriptors, linear regression equations are constructed based on the recursive method and on the least squares method. After that, among the models obtained, the best models are chosen. The following values are used as the criteria for this selection: coefficient of correlation R, standard deviation σ, the Fischer criterion F, the maximum residue δ_{max}, arithmetic mean of the residues δ_{mean} over the training sample, and the average absolute error Δ.

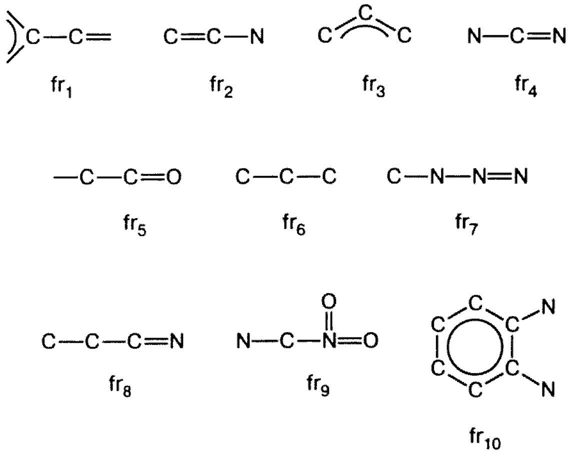
In the estimation of enthalpies of sublimation by this method, we used the EMMA (effective modeling of molecular activity) computer package.[17] The effective atomic charges were found by a known method.[18] The sample that served as the base for the construction of the models included 146 compounds from the database under consideration.

Of the models that we constructed, one of the most stable models is based on 13 descriptors and is expressed by the following equation:

$$\begin{aligned}\Delta H_{\text{subl}} = {} & 10.277 + 1.683(HB_{\text{don}} + HB_{\text{acc}}) + 0.111\,{}^0k + \\ & + 3.475\min_{\text{fr}_1}\sum \text{S}_i - 66.068\sum_{\text{fr}_1}\max q - \\ & - 36.388\min_{\text{fr}_2}(\min q) + 0.873\min_{\text{fr}_3}(\sum \text{S}_i) + \\ & + 55.139\sum_{\text{fr}_4}\max q_{\text{H}} - 11.717\sum_{\text{fr}_5}\min \text{S}_i + \\ & + 6.561\sum_{\text{fr}_6}\max q - 1.480\sum_{\text{fr}_7}\min \text{S}_i + \\ & + 59.662\sum_{\text{fr}_8}\min q - 0.799n(\text{fr}_9) - 0.201\sum_{\text{fr}_{10}}\text{S}_i, \qquad (11)\end{aligned}$$

where HB_{don} is the number of possible donors of hydrogen bonds[19] (HB_{don} = [the number of H atoms at O atoms] + [the number of H atoms at trivalent N atoms] + [the number of ≡CH fragments]), HB_{acc} is the number of possible hydrogen bond acceptors (HB_{acc} = [the number of O atoms] + [the number of trivalent N atoms]), ${}^0k$ is the zero-order molecular connectivity index,[20] ΣS_i is the sum of electrotopological states S_i [21] of atoms in a fragment fr_i, max$q$ is the maximum charge on the nonhydrogen atoms in a fragment fr_i, min$q$ is the minimum charge on the nonhydrogen atoms in a fragment fr_i, maxq_H is the maximum charge on the hydrogen atoms attached to the atoms of a fragment fr_i, minS_i is the minimum number of electrotopological states S_i of the atoms in a fragment fr_i, and $n(\text{fr}_i)$ is the number of fragments fr_i in the molecule.

For the model described by Eq. (11), n = 146, R = 0.929, σ = 2.661, F = 64.62, Δ = 1.905. The fragments used in Eq. (11) are shown below.



The enthalpies of sublimation obtained by this method for a number of compounds are presented in Table 1. Of the compounds included in the database under consideration, the largest error in the estimation of ΔH_{subl} was obtained in the case of 2,4,6-trimethyl-1,3,5-trinitro-1,3,5-triazacyclohexane (7.9 kcal mol^{-1}) (see Table 1). For 1,3,5,7-tetranitro-1,3,5,7-tetraazacyclooctane, 2-aminobutanoic acid, 9-diacetylaminoanthracene, and N-methyl-N,N'-dinitroethylenediamine, the errors were somewhat smaller and amounted to 4.6, 4.5, 4.2, and 4.1 kcal mol^{-1}, respectively.

As an evaluation of the potentialities of the method of "structure—property" correlations for the elucidation of the relationship between the structures and enthalpies of sublimation of the compounds considered, one can note that this model involves descriptors associated with the electrotopological state of atoms, atomic charges, and the numbers of hydrogen bonds. Thus, this model makes it possible to take into account the main types of intermolecular interactions in the crystal and to select molecular descriptors that determine the relationship between the structure of molecular crystals and the enthalpy of sublimation.

3.2. Estimation of enthalpy of sublimation based on "structure—property" correlations using three-dimensional descriptors. This method is a modified version of the conventional method of "structure—property" correlations involving descriptors that take into account three-dimensional structures of molecules. The three-dimensional descriptors used in this version can formally be divided into two groups. The first group includes geometric descriptors sensitive to changes in the three-dimensional molecular structure, *i.e.*, conformations, interatomic distances, bond lengths, and bond angles. For the estimation of enthalpy of sublimation, we chose the following descriptors from this group: shadow indices,[22] parameters of shape,[23] molecular volume,[24] and solvation indices.[25] To calculate these descriptors, the molecule is represented as a set of overlapping rigid spheres, whose radii are equal to the van der Waals radii of the corresponding atoms. Then the molecular volume, the surface area, and other descriptors are calculated for the resulting geometrical structure.

The second group comprises three-dimensional descriptors that reflect the combination of electronic and geometric characteristic features of a molecule. The area of the charged surface of the molecule[26] and the solvation indices belong to this group. The effective charges on the atoms used in the calculations can be evaluated by empirical and quantum-chemical methods.[18,27]

The search for "structure—property" correlations involving three-dimensional descriptors carried out in relation to ΔH_{subl} for a test sample consisting of 146 compounds (their possible conformations were calculated using the MOLGEO program[28]) led to the following regression model:

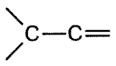
$$\begin{aligned}\Delta H_{\text{subl}} = {} & -5.8146 - 15.6399(\max(i' - i))^2/N_{\text{at}} - \\ & - 7.8448\ln(\text{PNSA-2}) + 599.7542(\sum_{\text{fr}_1}\max q_{\text{H}})^2/N_{\text{at}} + \\ & + 3.4219\min_{\text{fr}_2}\sum \text{S}_i - 66.8991\sum_{\text{fr}_2}\max q + \\ & + 55.5239\sum_{\text{fr}_3}\max q_{\text{H}} - 10.1101\sum_{\text{fr}_4}\min \text{S}_i - \\ & - 0.7469n(\text{fr}_5) + 0.7879\sum_{\text{fr}_6}\text{S}_i - 1.4411\sum_{\text{fr}_7}\min \text{S}_i + \\ & + 67.3580\sum_{\text{fr}_8}\min q - 2.2458n(\text{fr}_9) - 1.5047\sum_{\text{fr}_{10}}\min \text{S}_i, \qquad (12)\end{aligned}$$

where N_{at} is the number of nonhydrogen atoms in a molecule, max($i' - i$) is the change in the inductive parameter i [29] upon transfer of inductive effect (the largest value of all the differences between i' and i for each nonhydrogen atom in the molecule[17]), PNSA-2 is the area of the negatively charged surface. In turn,

$$\text{PNSA-2} = Q_t^{-}\Sigma(-SA_i),$$

where $(-SA_i)$ is the contribution of the ith negatively charged atom to the molecular surface area and Q_t^- is the sum of all the negative charges in the molecule.

For this particular model, n = 146, R = 0.933, σ = 2.599, F = 68.23, and Δ = 1.950. The fragments used in Eq. (12) are presented below (• is an arbitrary nonhydrogen atom).



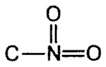
fr$_1$

—N—C=N

fr$_2$

C—C=O

fr$_3$

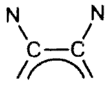


$C{-}\overset{\displaystyle O \atop \|}{N}{=}O$

fr$_4$

C—N=C—N

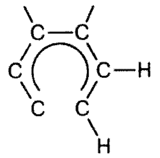
fr$_5$



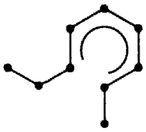
fr$_6$

C—C—C=N

fr$_7$



fr$_8$



fr$_9$

The values of ΔH_{subl} for a number of compounds from the database considered are listed in Table 1. The largest error, equal to 6.4 kcal mol^{-1}, was obtained in the case of 2,4,6-trimethyl-1,3,5-trinitro-1,3,5-triazacyclohexane.

Comparison of the results obtained in terms of the two "structure—enthalpy of sublimation" correlation models presented above shows that the use of three-dimensional descriptors in this method improves only slightly the quality of prediction (the correlation coefficient for model (11) derived without three-dimensional descriptors was 0.929, while that for model (12) involving three-dimensional descriptors was 0.933). In general, the method of "structure—property" correlations ensures fairly high accuracy of the evaluation of the enthalpies of sublimation of organic molecular crystals.

4. Evaluation of the enthalpies of sublimation of nitroaromatic compounds by atom-atom potential functions. In terms of the method of atom—atom potential functions (AAPF),[30] the interaction of molecules in the crystal is represented as the sum of the interactions between the atoms constituting them:

$$U(r) = 1/2\Sigma U_{ij}(r), \quad (13)$$

where $U_{ij}(r)$ is the potential of nonvalent atom-atom interaction (the summation involving i is taken over all the atoms in one molecule, while that involving j is taken over all the atoms in the rest of the molecules in the crystal). It is known that there are no direct experimental methods for determining the exact form of the $U_{ij}(r)$ potential. However, now we speak only about evaluation of the parameters of the $U_{ij}(r)$ function, the general form of which is specified beforehand using theoretical dependences for a particular model.

The $U_{ij}(r)$ potential is normally assumed to be the sum of attraction and repulsion potentials:

$$U = U_{rep} + U_{attr}. \quad (14)$$

The following functions are used most frequently as model potentials:

$$U(r) = -A/r^6 + B/r^{12} \text{ (Lennard—Jones potential)}, \quad (15)$$

$$U(r) = -A/r^6 + B\exp(-Cr) \text{ ("6—exp" potential)}, \quad (16)$$

where A, B, and C are empirical or semiempirical parameters characterizing the interaction of a particular pair of atoms.

Along with van der Waals interactions, electrostatic interactions also make a substantial contribution to the energy of a crystal lattice. The energy of the electrostatic interaction is found from the Coulomb formula:

$$E_{coul} = \Sigma(-e_ie_j/r_{ij}), \quad (17)$$

where e_i and e_j are effective charges on atoms, which are calculated using quantum-chemical methods.

On the other hand, the energy of interaction between a pair of atoms not linked by a chemical bond is a function of the distance r_{ij} between their centers. The set of these distances in a crystal lattice of known symmetry depends on the lattice parameters and on the mutual arrangement of the molecules that constitute the crystal. Therefore, the energy of the lattice can be regarded as a function of the unit cell periods a, b, and c; unit cell angles α, β, and γ; Euler angles φ, θ, and ϕ describing the orientation of the basis molecule with respect to the crystal axes; and the coordinates of the center of gravity of the molecule x, y, z:

$$E = f(a,b,c,\alpha,\beta,\gamma,x,y,z,\phi,\theta,\varphi). \quad (18)$$

By minimizing E as a function of the above parameters, one finds the global minimum in the potential energy surface of the crystal and the corresponding crystal lattice parameters.

Thus, knowing the structure of a crystal (for example, from the Cambridge Crystallographic Database), atomic charges, and enthalpies of sublimation of a number of compounds, one can find parameters A, B, and C from formulas (13)—(18) and use them subsequently for

Table 2. Parameters of atom—atom potential functions ("6—12") for the calculation of ΔH_{subl} of nitroaromatic compounds

Atoms	A /kcal Å^6 · · mol^{-1}	B /kcal Å^{12} · · mol^{-1}	r_0 /Å	ε_0 /kcal mol^{-1}
H...H	45.434	14372	2.93	−0.0359
H...C	125.81	83480	3.32	−0.0474
H...N	40.26	30015	3.38	−0.0135
H...O	222.05	85069	3.03	−0.1449
H...H*	45.49	11079	2.81	−0.0467
C...C	370.49	475289	3.70	−0.0722
C...N	366.43	592024	3.85	−0.0567
C...O	599.03	470921	3.41	−0.1905
C...H*	125.82	66292	3.19	−0.0597
N...N	358.30	722866	3.99	−0.0450
N...O	698.82	705298	3.56	−0.1731
N...H*	122.45	84238	3.34	−0.0445
O...O	369.15	170256	3.12	−0.2001
O...H*	104.44	2457	1.90	−1.1100
H*...H*	45.50	8429	2.68	−0.0614

the estimation of ΔH_{subl} of other related compounds, for which the enthalpies of sublimation have not been determined experimentally.[31—33]

However, only a few studies devoted to the selection of AAPF parameters specially meant for the calculation of enthalpy of sublimation have been published.[32] They refer mostly to aliphatic and aromatic hydrocarbons and to a quite small number of heterocyclic compounds. No AAPF parameters for nitro compounds have been reported. Therefore, using this approach, we selected parameters for the atom—atom potential functions of an aromatic series of nitro compounds (this work will be continued for other classes of compounds). From the database under consideration, 14 nitroaromatic compounds were selected, for which experimental enthalpies of sublimation and crystal structure parameters have been reported. The enthalpies of sublimation of these compounds were evaluated by the AAPF method using sets of parameters published in the literature (electrostatic interactions were not taken into account). The crystal lattice energies were calculated using the PMC (Packing of Molecules in Crystals) program developed by A. V. Dzyabchenko (L. Ya. Karpov Scientific Research Physicochemical Institute). Analysis of the results has shown that none of the sets of parameters considered ensures high accuracy (the average error varied from 20 to 80%).

To increase the accuracy of the calculation of ΔH_{subl} of aromatic nitro compounds, a new set of parameters for "6—12" type potential functions (15) was obtained by the least squares method. The calculation was based on a published set of parameters.[33] The optimized parameters of potential functions are listed in Table 2, and the crystal lattice energies obtained using these parameters for the 14 nitro compounds considered are presented in Table 3, together with the experimental ΔH_{subl} values.

Table 3. Calculated and experimental enthalpies of sublimation (kcal mol^{-1}) of nitroaromatic compounds

Compound	ΔH_{subl}		Ref.
	calc.	exp.	
Nitrobenzene	19.0	—	—
m-Dinitrobenzene	22.6	20.8	33
o-Dinitrobenzene	23.3	21.0	33
p-Dinitrobenzene	24.3	22.2	33
1,3,5-Trinitrobenzene	27.9	23.3	34
m-Nitroaniline	19.2	23.1	35
o-Nitroaniline	18.7	21.5	34
p-Nitroaniline	20.2	24.1	35
2,4,6-Trinitroaniline	31.8	29.9	36
2,4,6-Trinitrotoluene	29.3	28.3	34
1,3-Diamino-2,4,6-trinitrobenzene	33.6	33.5	37
2,3,4,6-Tetranitroaniline	36.2	—	—
1,3,5-Triamino-2,4,6-trinitrobenzene	39.5	40.2	37
Hexanitrobenzene	45.5	—	—

The accuracy of the evaluation of the enthalpy of sublimation increases substantially when intermolecular hydrogen bonds are taken into account. This can be seen most clearly in the case of 2,4,6-trinitroaniline. In fact, we were able to decrease the error of the calculation for this compound from 17.7 to 6.5% by introducing additional parameters characterizing hydrogen bonds. In general, for the set of AAPF parameters obtained, the average error amounted to 9.7%, and the average absolute error was 2.5 kcal mol^{-1}. The error can be decreased by taking into account electrostatic interactions and by using three-parameter potential functions of the "6—exp" type.

Thus, as a result of our study, several methods for the evaluation of the enthalpies of sublimation of various organic molecular crystals were developed. It can be seen from Table 1 that all these methods, except approximation by trigonometric polynomials, lead to similar correlation coefficients and average absolute errors. The method of atom—atom potential functions should be specially distinguished, since it permits more than just an estimation of the enthalpies of sublimation of compounds of various classes including structural isomers and polymorphic modifications; this method also makes it possible to take into account specific intermolecular interactions (in particular, hydrogen bonds).

The methods discussed above possess satisfactory predictive abilities and can be used successfully to evaluate enthalpies of sublimation both of compounds that have already been synthesized and of hypothetical compounds.

References

1. A. Bondi, *J. Chem. and Eng. Data*, 1963, **8**, 371.
2. A. Gavezzotti, *J. Am. Chem. Soc.*, 1985, **107**, 962.
3. A. Gavezzotti, *Nouv. J. Chim.*, 1982, **6**, 443.
4. A. Gavezzotti, *J. Am. Chem. Soc.*, 1983, **105**, 5220.
5. W. R. Cambill, *Chem. Eng.*, 1958, **65**, No. 5, 147.
6. G. S. Parks and S. S. Todd, *Ind. and Eng. Chem.*, 1929, **21**, 1235.
7. E. A. Moelwyn-Hyghes, *Physical Chemistry*, Pergamon, London, 1961.
8. Yu. A. Lebedev and E. A. Miroshnichenko, *Termokhimiya paroobrazovaniya organicheskikh veshchestv* [*Thermochemistry of Vapor Formation for Organic Compounds*], Nauka, Moscow, 1981, 215 pp. (in Russian).
9. E. A. Smolenskii, *Zh. Fiz. Khim.*, 1964, **38**, 1288 [*Russ. J. Phys. Chem.*, 1964, **38** (Engl. Transl.)].
10. E. A. Smolenskii, *Zh. Fiz. Khim.*, 1966, **40**, 2982 [*Russ. J. Phys. Chem.*, 1966, **40** (Engl. Transl.)].
11. E. A. Smolenskii, *Dokl. Akad. Nauk SSSR*, 1976, **230**, 373 [*Dokl. Chem.*, 1976, **230** (Engl. Transl.)].
12. A. V. Evtushenko, B. B. Smirnov, and V. A. Shlyapochnikov, *Izv. Akad. Nauk, Ser. Khim.*, 1993, 685 [*Russ. Chem. Bull.*, 1993, **42**, 645 (Engl. Transl.)].
13. A. V. Evtushenko and B. B. Smirnov, *Izv. Akad. Nauk, Ser. Khim.*, 1994, 22 [*Russ. Chem. Bull.*, 1994, **43**, 13 (Engl. Transl.)].
14. A. V. Evtushenko and B. B. Smirnov, *Izv. Akad. Nauk, Ser. Khim.*, 1994, 1389 [*Russ. Chem. Bull.*, 1994, **43**, 1315 (Engl. Transl.)].
15. E. N. L'vovskii, *Statisticheskie Metody Postroeniya Empiricheskikh Formul* [*Statistical Methods for Construction of Empirical Formulas*], Vysshaya shkola, Moscow, 1982, 323 pp. (in Russian).
16. S. M. Kavun, D. G. Petelin, V. A. Palyulin, I. I. Baskin, Yu. M. Genkina, I. G. Tsibko, and N. S. Zefirov, *Dokl. Akad. Nauk*, 1993, **333**, 189 [*Dokl. Chem.*, 1993, **333** (Engl. Transl.)].
17. D. V. Sukhachev, T. S. Pivina, V. A. Shlyapochnikov, E. A. Petrov, V. A. Palyulin, and N. S. Zefirov, *Dokl. Akad. Nauk*, 1993, **333**, 188 [*Dokl. Chem.*, 1993, **333** (Engl. Transl.)].
18. J. Gasteiger and M. Marsili, *Tetrahedron*, 1980, **36**, 3219.
19. Guang-Zhong Yang, Eric T. Lien, and Zong-Yu Guo, *Quant. Struct.-Act. Relat.*, 1986, **5**, 12.
20. L. B. Kier, *Medicinal Res. Rev.*, 1987, **7**, 417.
21. C. H. Hall, B. Mohney, and L. B. Kier, *J. Chem. Inf. Comput. Sci.*, 1991, **31**, 76.
22. R. H. Rohrbaugh and P. C. Jurs, *Anal. Chim. Acta*, 1987, **199**, 99.
23. A. Radecki, H. Lamparczyk, and R. Kaliszan, *Chromatographia*, 1979, **12**, 595.
24. J. Higo and N. Go, *J. Comput. Chem.*, 1989, **10**, 376.
25. K. Jinno and K. Kawasaki, *Chromatographia*, 1983, **17**, 445.
26. D. T. Stanton and P. C. Jurs, *Anal. Chem.*, 1990, **62**, 2323.
27. N. S. Zefirov, M. A. Kirpichenok, F. F. Izmailov, and M. I. Trofimov, *Dokl. Akad. Nauk SSSR*, 1987, **296**, 883 [*Dokl. Chem.*, 1987, **296**, 440 (Engl. Transl.)].
28. A. R. Katritzky, E. V. Gordeeva, V. V. Shcherbukhin, and N. S. Zefirov, *J. Chem. Inf. Comput. Sci.*, 1993, **33**, 102.
29. A. N. Vereshchagin, *Induktivnyi Effekt* [*Inductive Effect*], Nauka, Moscow, 1988, 110 pp. (in Russian).
30. A. I. Kitaigorodskii, *Molekulyarnye Kristally* [*Molecular Crystals*], Nauka, Moscow, 1971, 424 pp. (in Russian).
31. D. E. Williams and T. L. Starr, *Computers and Chemistry*, 1977, **1**, 173; D. E. Williams and D. J. Houpt, *Acta Cryst.*, 1986, **B42**, 286.
32. V. I. Poltev, *Kristallografiya* [*Crystallography*], 1977, **22**, 3, 453 (in Russian); V. I. Poltev and B. I. Sukhorukov, *Zh. Strukt. Khim.* [*J. Struct. Chem.*], 1968, **9**, 298 (in Russian).
33. F. A. Momany, L. M. Carruthers, R. F. McGuire, and H. A. Scheraga, *J. Phys. Chem.*, 1974, **78**, 16, 1595.
34. D. Ferro, V. Piacente, R. Gigli, and G. D'Ascenzo, *J. Chem. Thermodyn.*, 1976, **8**, 1137.
35. J. D. Cox and G. Pilcher, *Thermochemistry of Organic and Organometallic Compounds*, Acad. Press, London—New York, 1970, 643 pp.
36. L. Malaspina, R. Gigli, G. Bardi, and G. de Maria, *J. Chem. Thermodyn.*, 1973, **5**, 699.
37. R. B. Cundall, *J. Chem. Soc., Faraday Trans. 1*, 1978, **74**, 1339.
38. J. M. Rosen and C. Dickenson, *J. Chem. and Eng. Data*, 1969, **14**, 120.

Received May 28, 1996;
in revised form June 24, 1996