Russian Chemical Reviews **68** (1) 1 – 18 (1999) © 1999 Russian Academy of Sciences and Turpion Ltd

UDC 548.31

Cambridge Structural Database as a tool for studies of general structural features of organic molecular crystals

L N Kuleshova, M Yu Antipin

Contents

I.	Introduction	1
II.	Organisation and software of the Cambridge Structural Database	3
III.	Van der Waals atomic radii and nonbonded intermolecular distances in organic crystals	4
IV.	Principle of close packing in crystals and distribution of molecular centres in crystallographic unit cells	6
V.	Relative frequencies of occurrence of space groups	8
VI.	Why do organic crystals prefer centrosymmetrical groups?	10
VII.	Number of molecules per asymmetric unit and pseudosymmetry	14
VIII.	Supramolecular synthons and the possibility of prediction of crystal structures	16
IX.	Conclusion	17

Abstract. The review surveys and generalises data on the use of the Cambridge Structural Database (CSD) for studying and revealing general structural features of organic molecular crystals. It is demonstrated that software and facilities of the CSD allow one to test the applicability of a number of known concepts of organic crystal chemistry (the principle of close packing, the frequency of occurrence of space groups, the preferred formation of centrosym metrical molecular crystals, *etc*.) on the basis of abundant statistical data. Examples of the use of the Cambridge Structural Database in engineering of molecular crystals and in the systematic search for compounds with specified properties are given. The bibliography includes 122 references.

I. Introduction

Presently, the Cambridge Structural Database (CSD) is one of the most powerful information databases, which contains exhaustive data (atomic coordinates, molecular geometry and crystal symmetry) on crystal and molecular structures of organic, bioorganic, organoelement and organometallic compounds and to a large extent coordination compounds which have been studied by X-ray diffraction and/or neutron diffraction analysis. (The term organic crystals is used to denote crystals containing at least one 'organic' carbon atom, *i.e*., that involved in formation of C – C or C – H bonds.) Data on crystals which can hardly be called true organic molecular crystals (for example, data on inorganic alkylammonium salts, *etc*.) are also included in the CSD.

Therefore, the CSD involves information on diversified organic crystals. Hence, the correct choice and rejection of compounds in the search for structural regularities in series of related and structurally similar crystals is one of the most important problems in the use of the CSD. Presently, the CSD contains data on 181 000 crystal structures and the body of information in the CSD increases annually by 7000 – 8000 new structures. Clearly, the availability of the sizable database gave impetus to a search for various special structural features of particular classes of compounds as well as to a search for general regularities of organic crystal structures with the use of the CSD on the basis of statistically reliable analysis of large samples. The first line of investigation began actually after the first CSD release has appeared. Studies of the general regularities of organic crystal structures using statistical analysis became possible only with the appearance of computer versions of the CSD and with the development of special programmes for statistical treatment of a large body of data.

There are two approaches to the elucidation of the general rules and regularities of organisation of crystalline solids from isolated structural elements (atoms, molecules and their associates and ions). The first approach involves the use of computational methods, in particular, those used for *a priori* prediction of new crystal structures. In recent years, much progress in this field has been achieved.[1–4] The second approach is based on the correct statistical analysis of numerous crystal-structural data accumulated in information databases. We dwell on the results and prospects of the second approach with the use of modern tools and facilities of the CSD.

Useful new crystal-chemical information can be obtained by performing correct systematic statistical analysis of the large available body of experimental structural data. This idea appeared rather long ago. Thus many chemists and crystal chemists still use the values of the van der Waals radii which were determined by Pauling[5] and Bondi[6] based on the analysis of scarce crystal-structural and other (in particular, gas-kinetic) data available at that time. The fundamental principles of structures of organic molecular crystals (among them, the principle of close packing), which have been explicitly stated for the first time by Kitaigorodskii,[7] were also established on the basis of relatively scarce statistical data. In the 1970's and in the early 1980's, these systematic studies were believed to be very promising. However, as the experimental published data on crystal structures were accumulated, these investigations became more complicated. The major obstacle (which is believed to be insurmountable) lay in the fact that non-computerised search and processing of abundant numerical data were extremely tedious procedures.

L N Kuleshova, M Yu Antipin A N Nesmeyanov Institute of Organoelement Compounds, Russian Academy of Sciences, ul. Vavilova 28, 117813 Moscow, Russian Federation. Fax (7-095) 135 50 85.
Tel. (7-095) 135 93 43 (L N Kuleshova), Tel. (7-095) 135 92 15.
E-mail: mishan@xray.ineos.ac.ru (M Yu Antipin)

Received 23 April 1998
Uspekhi Khimii **68** (1) 3 – 22 (1999); translated by T N Safonova

It is worth noting that since the early days of crystallography (crystal chemistry), the researchers involved in this field have realised that data obtained in their studies are of fundamental importance and can have a wide use. Consequently, they were the first to systematically amass results of their studies. Thus as soon as 16 years after the development of the X-ray diffraction method, a beginning has been made in systematically reviewing X-ray structural data.[8] In this country, results of X-ray diffraction studies of organic crystals were first surveyed by Struchkov and were reported by Kitaigorodskii[7] in 1955. More recently, the well-known handbooks by Kitaigorodskii, Zorky and Belsky were issued.[9,10] The complete bibliography concerning these sources was published by Watson.[11,12]

First computer databases were compiled based on the available publications. Two of them, *viz*., 'Powder Diffraction File'[13] and 'NBS Crystal Data: Database Description and Applications'[14] became available totally as a computer version. These banks contain exhaustive (from the chemical standpoint) information but do not include atomic coordinates and were primarily used (and are still used) for identification and comparative studies of unit cell parameters and crystal symmetry.

The compilation of structure databases which contain atomic coordinates along with crystallographic data, details of chemical structures and bibliographic description dates back to the mid-1960's. When creating these databases, one would have to process information from current publications and to collect data on structures which have been studied previously and which are scattered over numerous literature sources. Therefore, in the initial stage, priority was given to works on the development of software for collecting, testing, processing and depositing data. Subsequently, programmes were developed for the search, selection and statistical analysis of information as well as for the distribution of databases. Systematic statistical studies with the use of all the data accumulated in databases appeared only in the last decade.

Presently, information on virtually all types of chemical compounds can be found in four structure databases: the Cambridge Structural Database (University of Cambridge, UK) contains data on 181 000 structures, the Inorganic Crystal Structure Database (University of Bonn, Germany) contains data on 41 500 structures, the Metals Data File (National Research Council of Canada, Ottawa, Canada) contains data on more than 11 000 structures and the Protein Data Bank (Brookhaven National Laboratories, USA) contains data on more than 1500 structures.

Presently, the CSD[15] is the biggest database. Using the information bases of the CSD, one can perform statistically reliable systematic studies of structural data. The knowledge of precise molecular geometry is necessary for the understanding of the nature of chemical bonds and chemical structures. Therefore, early studies were devoted to the determination of mean ('typical') values of bond lengths and bond angles in various classes of chemical compounds[16,17] and the acquisition of data on structural features of particular fragments or classes of chemical compounds as a whole.[18–20]

Systematic investigations into the mutual arrangement of molecules in crystals were begun with studies of geometrical characteristics of intermolecular interactions. In particular, statistical values of van der Waals atomic radii[21,22] and geometrical characteristics of specific interactions (hydrogen bonds,[23–27] and the CH···O(N), Hal···Hal and some other contacts[28–33]) were determined.

More recently, with the increase in the size of the CSD and the perfection of software for correct selection and statistical analysis of results, some researchers formulated problems based on the belief that the CSD contains, in some form, information on general regularities of molecular crystal structures. In this respect, noteworthy is the paper by Motherwell[34] in which distributions of positions of geometrical centres in molecular crystals were analysed and the preference for the realisation of the principle of close packing was statistically confirmed. Recently, some researchers have cast doubt on the firmness of this principle.[35] The doubts were substantiated by the fact that the principle of close packing should contradict a tendency of molecules to form strong intermolecular hydrogen bonds and by the fact that experimental crystal structures do not necessarily correspond to the calculated global energy minima.

The question of the frequency of occurrence of space groups has received the most study based on the CSD data. Numerous statistical studies were devoted to this problem.[36–44] Admittedly, the most detailed investigations were carried out by Belsky[42] and by Brock and Dunitz.[45] The considerable attention given to this problem is quite justified because many physical properties of crystals depend on the symmetry of the molecular arrangement and hence, the *a priori* estimation of the crystal symmetry is an important practical problem. For example, this is of importance in the design of materials with nonlinear-optical properties, which are constructed using organic chromophore molecules.[46,47]

Often, in studies using the CSD researchers attempt to establish the reasons for which compounds crystallise preferentially in centrosymmetrical (achiral) space groups rather than restrict themselves to simple statement of the facts (for example, about the considerable predominance of racemic crystals.[48]) Thus efforts to estimate the effect of molecular dipoles on the crystal packing were applied,[49] a comparative analysis of densities and stabilities of compounds that crystallise in both racemic and chiral space groups was carried out[50] and the correlations between densities, calculated packing energies and some thermodynamic properties of crystals were considered for pairs of polymorphic modifications existing at room temperature.[3]

In recent works devoted to analysis of the frequency of occurrence of space groups,[45] much attention was given to the number of formula units per asymmetric unit (Z'). Belsky and Zorky were the first to note the importance of the consideration of this parameter.[51,52] Studies of crystals with $Z' > 1$ brought up many questions. For example: are conformations of independent molecules in the unit cell identical or not? Are independent molecules related by any symmetry transformations? What is meant by pseudosymmetry? How do pseudosymmetry elements and true symmetry of real crystals relate to each other? Do crystals with $Z' > 1$ possess any characteristic properties? Abrahams attempted to give the answer to the last question.[53] He suggested that if pseudosymmetry is present, phase transitions of the second kind can occur in crystals of a number of inorganic compounds as the temperature is increased.[53] This hypothesis provided the basis for the development of a procedure for the search for new potential ferroelectrics,[54] which is successfully used for predicting inorganic compounds with high-temperature phase transitions.[55]

Interestingly, unit cells with $Z' > 1$ are more frequent in crystals containing molecules that can form stable associates through hydrogen bonds.[56] The occurrence of unit cells with $Z' > 1$ and the existence of pseudosymmetry are highly probable for compounds exhibiting liquid-crystal properties on melting.[57–59]

The results of studies of characteristic types of supramolecular fragments (supramolecular synthons) that formed in crystals through the strongest nonbonded intermolecular interactions for molecules of similar types or for molecules containing similar chemical functional groups may be very useful in revealing general regularities of organisation of crystal structures.[60–62] These studies fall within the realms of crystal engineering (a new intensively developing branch of science that appeared at the borderline between materials technology and crystal chemistry), which involves the development of procedures for the design of organic crystalline materials with predetermined properties. The crystal symmetry describes the topology of molecules in crystals and determines many physical properties of solids, namely, electrical, magnetic, linear-optical and nonlinear-optical properties, electronic and ionic conductivity, *etc*. Therefore, the current prime objective of crystal engineering is to reveal factors that

affect the formation of crystal structures with a particular molecular architecture, which causes the manifestation of a desirable physical property.

An important role of databases, in particular of the CSD, in the design of new materials was discussed in the review by Aakeroy.[63] It was noted that the development of software of the CSD (new algorithms and statistical and cluster analyses) and the incorporation (wherever possible) of a number of thermodynamic and/or other physicochemical characteristics of compounds into the CSD will favour the directed search and design of new compounds with preset properties.

Therefore, the use of the CSD has promise in the search for fundamental structural features of organic molecular crystals. However, it should be borne in mind that even very large statistical samples cannot assure automatically that statistically reliable characteristics and properties will be obtained without appropriate rejection and processing of the data. Therefore, when considering examples of the use of the CSD, we concentrated most attention on the analysis of schemes applied in the original publications cited for creating representative statistically significant samples. Modern versions of the CSD allow one to compile and reject particular data on the basis of a series of criteria and indications, each being a subject of special discussions in the literature.

The present review is the first attempt to consider the CSD as a tool in studying general structural regularities of organic molecular crystals.

II. Organisation and software of the Cambridge Structural Database

As has been mentioned above, the CSD contains results of complete three-dimensional X-ray and neutron diffraction studies of organic, bioorganic, organometallic and coordination compounds. The CSD is compiled by reviewing original publications in 1025 scientific journals and other sources (data as of 1988). Presently, the majority of crystal chemists and crystallographers as well as many chemists consistently use the above-mentioned databases in their routine work. For example, the 1998 version of the CSD is employed by researchers at about 40 scientific institutes of Russia and Newly Independent States of the Former Soviet Union. About 150 articles were devoted to the use of the CSD. The major principles of selection and systematic numerical analysis of information in crystallographic databases were considered and analysed in detail in one of the latest papers, namely, in the review by Allen[64] 'Crystallographic Data Bases: Samples and Analysis of Precise Structure Information from the Cambridge Structural Database'. Therefore, we mention only briefly the essentials of organisation and software of the CSD.

The use of the CSD requires special software, namely, the programme package developed for the search, selection of required data, analysis of geometrical parameters of molecules and crystal structures and their representation with the use of computer graphics tools. The search for the required information in the CSD is performed using the QUEST programme, which specifies the query. This query can be formulated in the form of three major packages, *viz*., 1D, 2D and 3D.

A query specified by the QUEST1D option allows one to obtain bibliography, the name of a compound, the molecular formula or the amino acid sequence, unit cell parameters, the *R* factor and required comments (the presence of disorder, errors, *etc*.).

A query specified by the QUEST2D option allows one to obtain information on the atom and bond properties, the formal chemical type of the bond, tables of atom connectivities and the structural formula.

A query specified by the QUEST3D option allows one to obtain the description of a crystal structure, *i.e*., the space group, symmetry operations, three-dimensional atomic coordinates, crystallographic connectivity determined by the covalent radii and data on the correspondence between the 'crystallographic' and 'chemical' atoms from tables of connectivities of 2D diagrams.

Queries are specified with the use of special menus (BILD, SEARCH and QUEST). Information can be rejected by specifying a set of BITSCREEN indications.

The QUEST3D option compiles a data file FDAT from information selected according to the query. This file serves as an input file for the GSTAT(VISTA) and(or) PLUTO programmes. The GSTAT programme was originally developed for calculations of intra- and intermolecular geometry and geometry of molecular fragments.[65] Presently, GSTAT is a programme package which allows one to analyse geometrical characteristics with the use of statistical and numerical methods.[66] The VISTA programme (graphical supplement to GSTAT) is destined mainly for numerical, statistical and graphical analysis of the geometrical information extracted from the CSD.

The PLUTO programme is applied not only by crystallographers but also by chemists. This programme is destined for the preparation of various graphical illustrations, including visualisation of both molecular structure and crystal packing based on three-dimensional coordinate data sets.

The most precise data in the CSD are used in the analysis of geometrical characteristics. In the simplest case, entries containing errors and disordered structures are rejected. The CSD contains two major parameters characterising the accuracy of the crystal structure, namely, the *R* factor and the *AS* parameter. The former characterises the overall discrepancy between the diffraction data and the crystal structure solved on the basis of these data. The latter characterises the average standard deviation of bond lengths. These parameters can be used both in QUEST (for preliminary rejection in the course of the search) and GSTAT (in statistical averaging).

Calculations of average molecular geometrical parameters used in statistical analysis were discussed in many papers (see, for example, Refs 67–69). Generally, either an average or a weighted average is calculated. A weighted average is better suited for cases where environmental effects are insignificant (for example, for rigid parameters, such as bond lengths). If environmental effects are substantial (torsion angles and bond angles), simple unweighted averages should preferably be determined. It is desirable to use an unweighted average also in statistical verifications of hypotheses. The absence of standard deviations of parameters for the majority of structures in the CSD does not allow one to calculate weighted averages without recourse to the original literature. However, Taylor and Kennard[69] demonstrated that an unweighted average can be sufficient even for rigid parameters provided the scatter of the experimental values is limited by the *AS* criterion. The authors believe that when the value of *AS* is 1.2, estimates of, for example, bond lengths in organic structures are quite correct. An unweighted average is used in the GSTAT programme for calculations of standard statistical characteristics.

It should be noted that only selected crystal-structural data, which are obtained by X-ray diffraction analysis of single crystals, are included in the CSD. In particular, atomic anisotropic displacement parameters (previously, they were called thermal ellipsoids) cannot be analysed. These parameters contain unique information on the dynamics of crystal structures and characteristics of atomic movements and/or disorder in crystals. Unfortunately, these data are absent in the CSD. Trueblood and Dunitz[70] were the first to call attention to this disappointing fact. They have demonstrated how information on the dynamics of crystal structures can be used for estimating energy characteristics (for example, barriers) of slow intramolecular movements in crystals, which is important in studies of properties of organic crystals.

III. Van der Waals atomic radii and nonbonded intermolecular distances in organic crystals

The atoms of elements can be, in a rather rough approximation, represented by rigid spheres with radii which are geometrical characteristics of the corresponding elements.[7] The effective values of these characteristics are called van der Waals radii (WR) of elements. Needless to say these concepts are very simplified because atoms in molecules (the more so, in molecules in crystals) cannot be rigid, and their effective sizes depend also on the nearest environment as well as on the hybrid state of the atom. The first studies [71, 72] carried out with the use of the CSD have demonstrated that some bound atoms are flattened rather than spherical.[73] However, the simple 'rigid' spherical model appears to be a very persistent and extremely useful formalism.

Frequent application of WR stems from the necessity of comparing lengths of expected nonbonded intermolecular atom–atom contacts in crystals, which are determined as the sum of the corresponding WR, with the corresponding distances observed in crystal structures. Besides, in early work [7] WR were used for choosing force-field parameters in calculations of the energy of intermolecular interactions in crystals. Presently, the CSD is finding increasing use for choosing these parameters,[1–4] and the parameters of pairwise X···Y interactions are chosen based on the analysis of statistical distributions of bond lengths of the corresponding contacts retrieved from a large body of data in the CSD.

Several systems of WR are known. They were suggested by Pauling,[5] Kitaigorodskii,[7] Bondi,[6] Zefirov and Zorky,[74] Nyburg and Faerman,[75] and Batsanov.[76] Bondi's system (table) of WR [6] is most commonly used in spite of the fact that it has been repeatedly criticised in more recent works.[71, 74, 77] In this country, the system of radii suggested by Zefirov and Zorky [74] has gained acceptance.

Rowland and Taylor performed systematic analysis of nonbonded contacts in organic molecular crystals.[23–26, 78] In one of their recent works, the authors decided to test the compatibility between the commonly accepted systems of WR of nonmetals H, C, N, O, F, P, S, Cl, Br and I and the observed nonbonded interatomic distances in crystals studied over a period of the last 30 years.[22] Studies were carried out with the use of the 1995 version of the CSD containing 126 353 crystal structures. Only precise (R factors < 5%) and ordered structures that contained exclusively the above-listed atoms were considered. Structures containing charged groups were also rejected to exclude the effect of electrostatic interactions on interatomic distances. A total of 28 403 structures satisfying all the above-mentioned requirements were selected. For a pair of atoms A and B under consideration, all intermolecular distances falling within the interval $(R_A + R_B) \pm 1.5$ Å, where R_A and R_B are the corresponding WR according to Bondi, were taken into account. The broad range ensured that the sample thus obtained would not lead to results expected in advance.

Distribution histograms of nonbonded distances for various atom pairs were constructed [22] with a bin size of 0.1 Å. Different types of histograms were obtained. Some of them showed a monotonic dependence of the number of atom pairs (n) on the distance (H···H). Other histograms had a well-defined peak (Cl···Cl). In some cases, a maximum was poorly defined (H···F) or absent [histograms had only a shoulder (C···N)] (Fig. 1).

Histograms of the first type were not analysed because they gave no numerical characteristics of the distribution.† Rowland and Taylor [22] suggested the use of the parameter d (the distance corresponding to the histogram half peak or shoulder height; Fig. 2) as a characteristic of histograms of all the other types. The stability of this parameter is the major argument in favour of its application as a numerical characteristic. Other possible (and most commonly used) characteristics, such as the shortest contacts‡ or the position of the distribution maximum, appeared to depend strongly on the choice of the bin size and the boundaries of the distribution. Because of this, the value of d can be estimated much more reliably than the position of the maximum of the histogram or the shortest contacts.

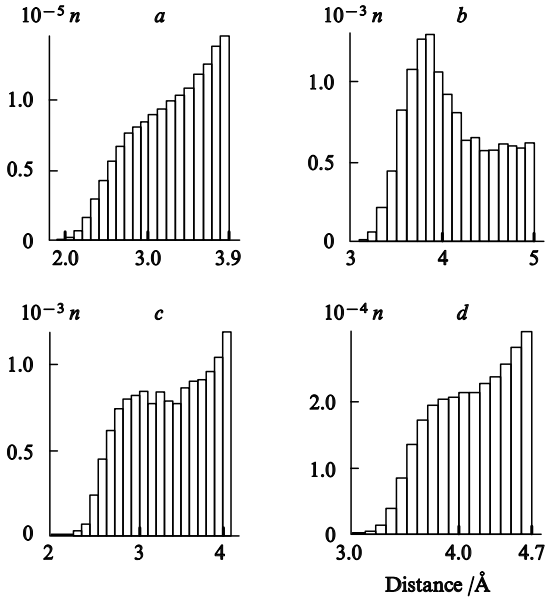


Figure 1. Distribution histograms of nonbonded intermolecular distances for selected atomic pairs.[22] H···H (*a*); Cl···Cl (*b*); H···F (*c*); C···N (*d*).

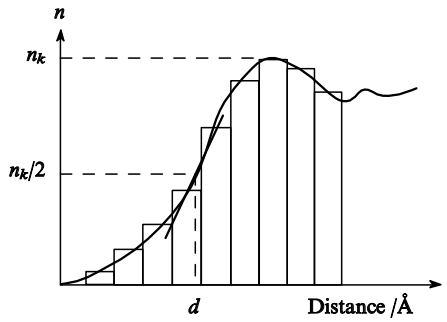


Figure 2. Scheme for the determination of the parameter d;[22] n_k is the height of the maximum of the histogram.

The distributions of the lengths of O···O, N···O and N···N nonbonded distances have a pronounced peak at short distances corresponding to typical lengths of hydrogen bonds (Fig. 3). Since only van der Waals contacts were considered in the work cited,[22] all contacts of the NH and OH groups were also excluded. This restriction resulted in the removal of maxima corresponding to H bonds from all three distributions.

† These histogram patterns are due to the fact that all intermolecular distances (in particular, H···H) are taken into consideration, among them distances that do not correspond to the real van der Waals ('reference' [7]) contacts in crystals. However, the last-mentioned contacts serve as a basis for the determination of WR of atoms in crystals. Consequently, an approach used by the authors of the work [22] is of limited utility and cannot be used for estimating values of typical van der Waals H···H contacts.

‡ It is self-evident that the expression 'interatomic contacts' makes sense only for the left portion of the histogram where the distances are close to the sum of WR.

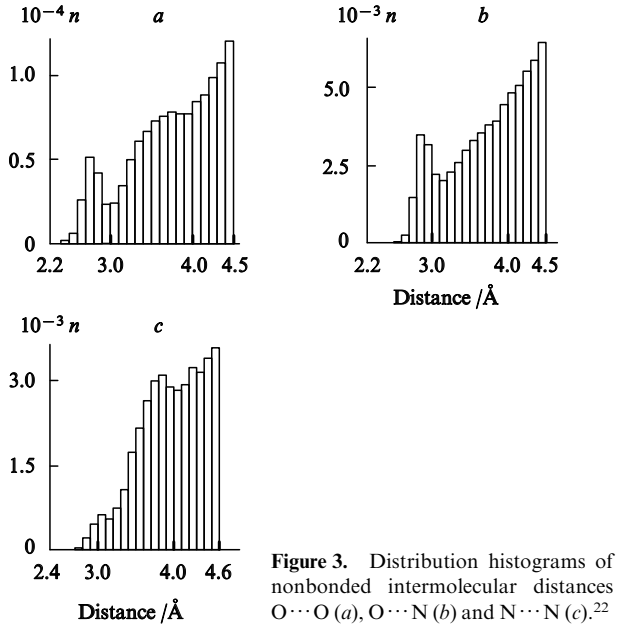


Figure 3. Distribution histograms of nonbonded intermolecular distances O···O (*a*), O···N (*b*) and N···N (*c*).[22]

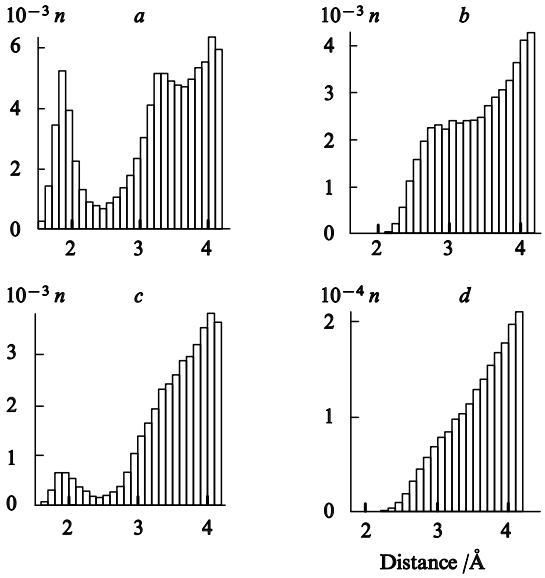


Figure 4. Distribution histograms of nonbonded distances (O,N)H···O (*a*), (C)H···O (*b*), (O,N)H···N (*c*) and (C)H···N (*d*).[22]

As expected, the values of *d* (Table 1) for the (O,N)H···N(O) distances are markedly smaller than those for the (C)H···O(N) distances (Fig. 4). On the contrary, the values of *d* for (O,N)H···Cl(Br) are larger than those for (C)H···Cl(Br). This can be explained by the fact that the OH and NH groups are almost always involved in strong hydrogen bonds with oxygen or nitrogen atoms, while bonds with halogen atoms appear to be 'secondary' and hence, are substantially weaker. If this assumption is true, the concept of WR for the oxygen and nitrogen atoms (at least, for the OH and NH groups) in organic crystal structures containing hydrogen bonds is meaningless. Taking into account the aforesaid, Rowland and Taylor [22] estimated the values of *d* for the H···O, H···N and H···S distributions considering only hydrogen atoms at carbon atoms.

Table 1. Observed values of *d* (Å).[22]

Atom	H	C	N	O	F	S	Cl	Br	I
H		3.02	2.59		2.54	2.91	2.88	2.99	3.14
C	3.02	3.59	3.31	3.24		3.61			
N	2.59	3.31	3.43	3.28	3.13	3.42	3.41		
O		3.24	3.28		3.00	3.45	3.33	3.48	
F	2.54		3.13	3.00	2.90				
S	2.91	3.61	3.42	3.45		3.60	3.55		
Cl	2.88		3.41	3.33		3.55	3.50		
Br	2.99			3.48				3.70	
I	3.14							4.05	

The authors calculated the values (r_i) that characterise WR of the atom of the *i*th type using the values for all kinds of pairs of atom-atomic contacts (see Table 1) by minimising the following function:

$$f = \sum_i \sum_{j<i} w_{ij}[d_{ij} - (r_i + r_j)]^2 ,$$

where d_{ij} is the value of *d* for the pair of atoms *i* and *j*, w_{ij} is the weighting coefficient, which is equal to 1 and 0.25 for bin sizes of 0.1 and 0.2 Å, respectively. The summation was made over all types of atoms.

We recall that Rowland and Taylor [22] proposed to use *d* as a characteristic of the distribution of intermolecular distances based on the numerical constancy of this value regardless of the choice of the bin size and boundaries of the distribution rather than on the expectation that *d* can be equal to the sum of WR of the corresponding atom pair. The authors only assumed that the values of r_i thus determined would correlate with the values of WR (R_w): $r_i = cR_{wi}$, where *c* is the correlation coefficient. The coefficient *c* appeared to be close to unity.

The values of r_i determined as described above as well as the values of WR determined independently are given in Table 2. It can be seen that the values of r_i agree most closely with Bondi's system of radii. The authors [22] believed that this is most likely a random coincidence. It was noted that, on the whole, the radii which were determined in the work cited from a large number of interatomic distances in crystals agree well with WR determined by Bondi, yet some exceptions were observed. In particular, WR of hydrogen atoms in Bondi's system are 0.1 Å larger. The calculation taking account of the distribution of the H···H contacts determined only by the neutron diffraction method gave a radius of 1.19 Å. This value is close to that determined by Bondi. When H···O and H···N contacts were included in the consideration, the value of the radius under consideration decreased. The values of WR in Bondi's system were determined primarily from the H···H contacts in adamantane. It is known that covalently bound hydrogen atoms are partially positively charged in the majority of organic compounds. Therefore, H···H distances would be expected to be somewhat elongated due to

Table 2. Van der Waals radii of selected atoms (Å).

Atom	I	II	III	IV	V	VI	VII
H	1.2	1.17	1.16	–	1.2	1.2	1.10
C	–	1.72	1.71	–	1.7	1.7	1.77
N	1.5	1.57	1.50	1.60	1.6	1.55	1.64
O	1.40	1.36	1.29	1.54	1.5	1.52	1.58
F	1.35	–	–	1.38	1.4	1.47	1.46
S	1.85	2.02	1.84	2.03	1.85	1.80	1.81
Cl	1.80	1.78	1.90	1.78	1.8	1.75	1.76
Br	1.95	–	–	1.84	1.9	1.85	1.87
I	2.12	–	–	2.13	2.1	1.98	2.03

Note. The values of R_w were determined by: I, Pauling;[5] II, Kitaigorodskii;[7] III, Zefirov and Zorky;[74] IV, Nyburg and Faerman;[75] V, Batsanov;[76] VI, Bondi;[6] VII, values of r_i calculated by Rowland and Taylor.[22]

electrostatic repulsions. On the contrary, the H···O distances can be somewhat shortened due to electrostatic attractions. Therefore, the radii which were determined [22] with the use of all types of H···X contacts based on a large body of statistical data are presumably more accurate than those determined by Bondi. Yet another noticeable difference in the values of WR was observed for the nitrogen atoms. A larger (compared to Bondi's radius) value of r_N was obtained due to the unexpectedly large value of d for the N···N distance distribution. The authors [22] believe that the value of WR for the nitrogen atom suggested by them is more appropriate for interpreting N···N contacts in crystals.

It should be noted that with only a few exceptions, the dispersion of the values of WR for atoms determined by different authors is small (~ 0.1 Å). It is with this (or even narrower) accuracy that WR are generally determined in experiments.[76] Therefore, the results discussed in this section can be considered as the demonstration of a new statistical approach to the estimation of these values rather than as yet another attempt to refine WR of atoms. In our opinion, a possible line of further investigation in this field with the use of the CSD can involve analysis of typical intermolecular distances in crystals at different temperatures. It is known that cooling of molecular crystals results in contraction of intermolecular contacts, which is more substantial (up to 0.2 Å) than those of other contacts. Therefore, systematic analysis of these changes in crystals with a decrease in temperature (generally, from room temperature to -100 °C and below) is of special interest.

IV. Principle of close packing in crystals and distribution of molecular centres in crystallographic unit cells

The major principle of organic crystal structures, namely, the principle of close packing, was stated by Kitaigorodskii in 1955.[7] In recent years, several researchers concerned with computational modelling of crystal packings noted that if calculations were based on the molecular geometry of compounds, several close-packed structures were generally observed, all with packing energies within 5 kJ mol^{-1} of the global minimum.[71] The observed crystal structure is often the one found to coincide with the global minimum, but not always.[1–4]

Some researchers believe that the principle of close packing contradicts the tendency of molecules to form stable associates in crystals through hydrogen bonds or other specific intermolecular interactions, at least in the case of polyfunctional or bulky and branched molecules.[35] On the whole, the problem of *a priori* calculations of a crystal structure of a particular compound is far from an unambiguous solution. However, considerable progress has been observed in this field in recent years. In the recent review,[4] Gavezzotti noted that nowadays the answer to the question 'Are crystal structures predictable?' would have to be 'Yes, sometimes'. This question was also discussed in detail in another paper.[63]

The major aim of *a priori* calculations of crystal structures is to search for optimum packings of molecules and/or their associates. In this case, the principle of close packing can provide a basis for these calculations. Therefore, the examination of its applicability with the use of abundant statistical data in the CSD is an urgent problem.

Motherwell [34] attempted to find regularities of molecular arrangements in crystalline compounds. He expected to reveal preferred molecular arrangements for the purpose of using the rules found in computational modelling of crystal structures.

First, the most common space groups were revealed. In the 1996 version of the CSD (160 000 structures), 82% of all organic structures encompassed by seven space groups. These are: $P\bar{1}$ (19%), $P2_1$ (6%), $P2_1/c$ (35%), $C2/c$ (7%), $P2_12_12_1$ (9%), *Pbca* (4%) and *Pnma* (2%). Crystal structures belonging to the above-listed groups were retrieved from the CSD with the use of the QUEST search programme. The selection was restricted to structures with one molecule per asymmetric unit ($Z' = 1$), Disordered and polymeric structures were rejected. An exception was made for the space group *Pnma*, where all molecules have an inherent mirror plane coinciding with the crystallographic plane m. For this study, a selection was made with $Z' = 1/2$.

It was decided to use the coordinates of the molecular geometrical centre as the characteristics of the molecular position. These coordinates were calculated as the average of all atomic coordinates (including hydrogen atoms if present). The molecular centre coordinates were transformed as to place the molecule into the first quarter of the unit cell. The calculations were performed with the use of a modification of the standard programme PLUTO CSD. Then the data obtained were used as input parameters to the statistical display programme VISTA.

The distribution patterns of the molecular centres were displayed as three-dimensional maps, which readily illustrate the cluster character of the distributions. The maps were presented as planar projections onto the coordinate planes and distribution histograms along the coordinate axes.

1. Group $P\bar{1}$

For the group $P\bar{1}$, 967 structures that satisfy the above-mentioned requirements were selected. The results of calculations of the molecular geometrical centres are shown in Fig. 5. In the normalised distribution (*i.e.*, with the coordinates brought into the first quarter of the unit cell; Fig. 5*b*), the points corresponding to clusters form a distinct cross with the following coordinates of the centre: $x = 0.25$, $y = 0.25$, $z = 0.25$. As expected, in this space group the patterns are very similar for all three projections. The regions of the increased density of the distribution have the coordinates $(x,y) = (0.25, 0.25)$. The regions of the decreased density of the distribution have the coordinates $(x,y) = (x,z) = (y,z) = (0, 0), (0, 0.5), (0.5, 0), (0.5, 0.5)$. The

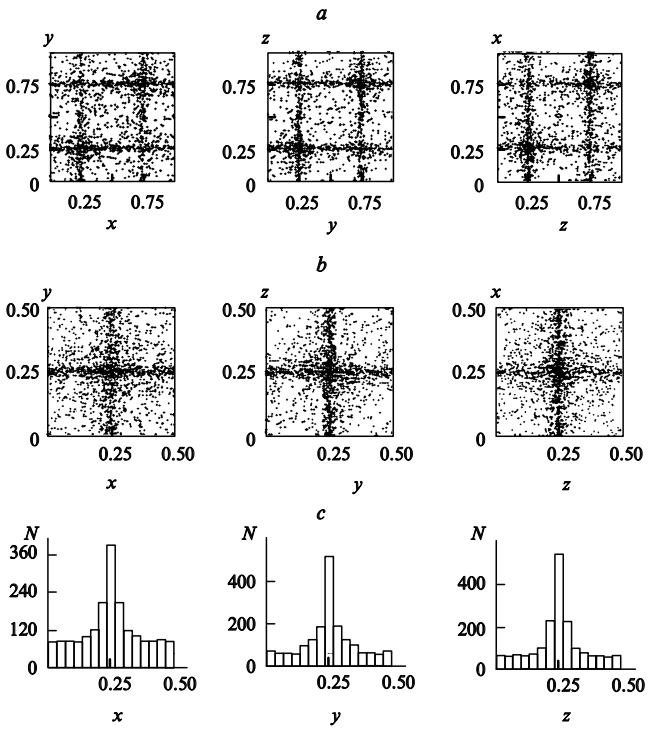


Figure 5. Distributions of geometrical molecular centres in the group $P\bar{1}$;[34] scattergrams for the total unit cell (*a*); scattergrams adjusted to bring the reference molecule to the first quarter of the unit cell (*b*); distribution histograms of positions of the geometrical molecular centres along the axes of coordinates of the unit cell (*c*). Hereinater N is the number of points.

areas of low population correspond to the positions of the symmetry elements in this space group, *i.e.*, to the centres of symmetry.

Motherwell [34] suggested that this distribution pattern of the molecular centres is a graphical illustration of the principle of close packing. Clearly, the molecules cannot occupy inversion centres since asymmetrical molecules, which have no inherent centre of symmetry, were chosen. The reasons why the molecules will tend to 'avoid' regions about the inversion centres are obvious if one considers the location of this molecule between two centres of symmetry at a distance d from each of them. Evidently, the best location of a 'spherical' molecule is halfway between these centres. However, the shapes of real molecules are far from spherical ones and, in addition, the molecules have functional groups which can be involved in specific directional interactions (for example, hydrogen bonds). Nevertheless, judging from the results obtained, a general tendency exists according to which the molecules prefer to occupy the midpoints between centres of symmetry at a distance equal to one-quarter along each coordinate axis.

2. Group $P2_1$

For the group $P2_1$, 1000 structures were selected. To simplify consideration, the projection along the y axis was used as an informative projection. It was established that the molecular centres cluster predominantly halfway between the screw axes $(x,z) = (0.25, 0.25)$ (Fig. 6). In the regions $(x,z) = (0, 0), (1.0, 0.5), (0.5, 0), (0.5, 0.5)$, *i.e.*, on the screw axes, areas of the minimum density of the distribution are observed. This confirms the above-mentioned tendency of molecules to occupy the midpoints between symmetry elements (in this case, between screw axes).

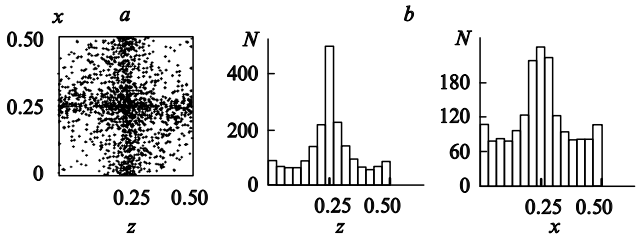


Figure 6. Scattergram (*a*) and distribution histograms (*b*) of molecular centres in the group $P2_1$.[34]

3. Group $P2_1/c$

The group $P2_1/c$ contained 994 structures after elimination of structures with $Z' > 1$. The distribution histogram (Fig. 7) shows one narrow peak on the axis x. The histograms along the axes y and z have substantially less pronounced peaks with the coordinates (0.125, 0.375, 0.5) and (0.125, 0.25, 0.375). The characteristic feature of the distribution pattern in this group is a broad band of points in the plane defined by $x = 0.25$, which corresponds to positions of molecules along the glide plane. The areas of low population in the plane (x,z) are located at the positions corresponding to the centres of symmetry: (0, 0), (0, 0.5), (0.5, 0), (0.5, 0.5).

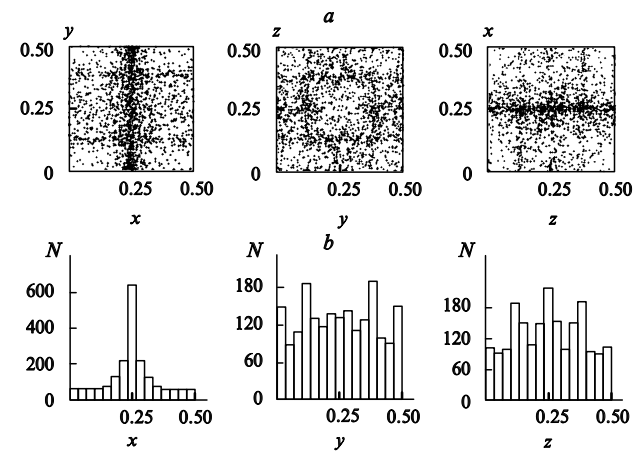


Figure 7. Scattergrams (*a*) and distribution histograms (*b*) of molecular centres in the group $P2_1/c$.[34]

4. Group $C2/c$

The sample for the group $C2/c$ consisted of 984 structures. The distribution histograms show pronounced narrow peaks at $x = 0.125$ and 0.375 (Fig. 8). The distribution pattern along the axis y is less distinct but the histogram has a narrow peak at $y = 0.25$. The distribution histogram along the direction z, as well as the histogram along x, have distinct peaks at 0.125 and 0.375. The distribution patterns consist of narrow bands in the plane xy ($x = 0.125$ and 0.375). It may be noted that there is an obvious increase in the density extending in the direction of the glide plane c and large voids in the regions $(x,z) = (0, 0.25)$ and (0.5, 0.25), which correspond to the positions of rotation twofold axes. This is an expected result because it is obvious that molecules related by rotation axes cannot be closely packed.

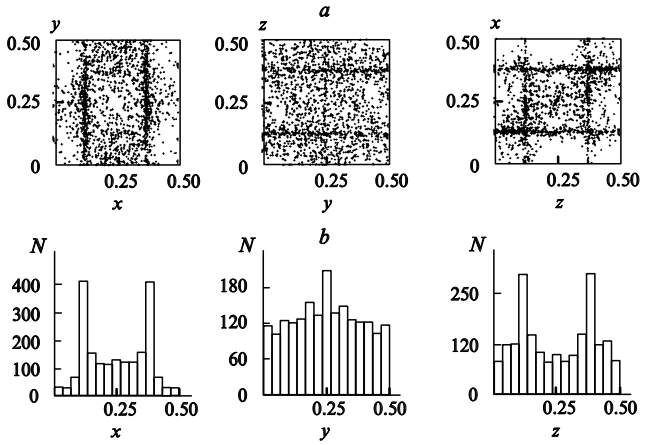


Figure 8. Scattergrams (*a*) and distribution histograms (*b*) of molecular centres in the group $C2_1/c$.[34]

The regions of low population are also observed at the inversion centres, like in other space groups.

5. Group $P2_12_12_1$

The sample for the group $P2_12_12_1$ consisted of 999 structures. The histograms and the distribution densities of molecular centres in the crystal unit cell appeared to be most uniform. The only characteristic feature is the low density along the screw axis parallel to the axis z.

6. Group *Pbca*

The distribution of molecular centres for the group *Pbca* (999 structures) shows clear maxima on each axis at the positions 0.125 and 0.375. The points are distributed as distinct bands in all three planes with the coordinates (0.125, 0.375) along each crystal axis. The regions of low population correspond to the positions of centres of symmetry and positions of screw axes. As in the groups $P\bar{1}$ and $P2_12_12_1$, all three directions are equivalent.

7. Group *Pnma*

The distributions of molecular centres for the group *Pnma* were constructed based on the data on 731 compounds. Since all molecules in this group have a mirror plane, the only informative

distribution histogram is the projection (x,z). This projection has clusters of points in the regions about (0.25, 0) and (0.25, 0.5). However, these clusters are very diffuse. Regions of low population at (0, 0), (0.5, 0) and (0.5, 0.5), which correspond to the centres of symmetry, are better defined.

Let us emphasise again that all the structures considered above were selected regardless of chemical atom types. The distributions of molecular centres were obtained without taking into account hydrogen bonds, electrostatic interactions and the fact that approximately 45% of all the structures under consideration belonged to complex organometallic molecules whose shapes are far from spherical. To exclude possible objections, the distribution patterns of molecular centres were considered for a subset of the $P2_1/c$ sample consisting of hydrocarbons only (126 structures) and no significant difference could be seen in the distribution of molecular centres.

With the aim of analysing the molecular environment in crystals in more detail, an extension was written for the programme PLUTO to display the molecular coordination sphere about the reference molecule according to the approaches of Kitaigorodskii [7] and Gavezzotti.[79] The coordination sphere was determined based on calculations of the energy of intermolecular interactions. Only molecules whose energy of interactions with the reference molecule is not lower than 4.18 kJ mol^{-1} were included in the calculations. The typical coordination numbers of molecules in the CSD were found to be 12–14, which confirmed the suggestion of Kitaigorodskii.[7] The potential energy was calculated by summing energies of pairwise atom-atomic interactions using Gavezzotti's empirical potential.[79] The electrostatic term was ignored. Analysis of several hundreds of coordination spheres determined as described above demonstrated that, almost without exception, a close-packed layer of six adjacent molecules can be observed in this coordination sphere. Needless to say that the real situation is far from simple. Many examples can be presented where the molecules form strongly bound dimers or chains. However, in these cases the molecules also tend to form a coordination sphere, which is the nearest to the closest packing of spheres.

Therefore, the distribution histograms of molecular centres in the symmetry groups under consideration show peaks or clusters corresponding to the preferred positions of molecular centres in unit cells. These centres are located halfway between the inversion centres or screw axes in unit cells. The calculations of the energy of intermolecular interactions with the adjacent molecules in the coordination sphere confirmed that these assumptions are energetically justified. It is useful to approximate molecules by spheres to visualise the preferential formation of close-packed layers of molecules and tendencies to close packings of spheres.

V. Relative frequencies of occurrence of space groups

Analysis of relative frequencies of space-group occurrence for the known crystal structures is one of the most evident areas of application of the CSD in studies of general regularities of crystal structures. This analysis is of interest not only from the theoretical, but also from the practical standpoint, because crystal symmetry determines a number of important properties manifested by many crystals. Evidently, the problem of the choice of a particular space group by a molecule is thus currently the most studied.

Kitaigorodskii [7] was the first to demonstrate that only a few of 230 theoretically plausible groups are of rather frequent occurrence. Examples of crystal structures belonging to other groups are few in number, while certain space groups are 'prohibited'. Among the most populated groups, centrosymmetrical space groups are by far 'favourites'. Kitaigorodskii ascribed this selectivity to the difference in the packing energy of symmetric arrangements of molecules in crystals. On the whole, it was concluded that organic molecules crystallise preferentially in space groups (SG) that can ensure the close packing of spheres or tri-axial ellipsoids (which are most commonly used as approximations for organic molecules). In this respect, the groups $P2_1/c$ and $P\bar{1}$ were assigned to most convenient.

More recently, Belsky and Zorky [40, 51, 52, 80] published a series of papers in which the most populated space groups were established based on a rather large body of data.[9, 10] It appeared that there are only 9 such groups, *viz*., $P2_1/c$, $P\bar{1}$, $P1$, $P2_1$, $P2_12_12_1$, $Pbca$, $C2/c$, $Pna2_1$ and $Pnma$. They account for 79.2% of all the structures known at that time. The results obtained demonstrated that organic molecules actually crystallise preferentially in centrosymmetrical SG. Over 50% of all known organic structures belong to only two space groups, *viz*., $P2_1/c$ and $P\bar{1}$. In addition, it was noted that the inherent molecular symmetry and the number of crystallographic orbits occupied by molecules must be taken into account.

More than a dozen of recent publications were devoted to studies of the relative frequencies of occurrence of SG for both organic (Wilson,[36–39] Padmaja *et al*.,[41] Belsky *et al*.,[42, 81] Mighell *et al*.[44, 82] and Brock and Dunitz [45]) and inorganic (Mackay,[43] Mighell *et al*.[44, 82] and Baur and Kassner [83]) crystals based on statistically reliable data using computer databases.

All authors noted the prevailing occurrence of centrosymmetrical SG. However, the frequency of occurrence of these groups for organic compounds reported in different works varies from 66% to 75%. Among inorganic compounds (Inorganic Crystal Structure Database, 1992), this frequency is, on the average, even higher (78%). Proteins and other natural polymers consisting of pure enantiomeric molecules are the exceptions, where chiral SG prevail.[41]

Note that the preference of a particular SG (centrosymmetrical or noncentrosymmetrical) for a molecule is of exceptional importance from the practical standpoint to researchers involved in the design of new organic materials. In particular, compounds that can possess nonlinear-optical properties and are potentially able to generate the second harmonic of laser radiation crystallise most often in centrosymmetrical SG excluding the manifestation of the above-mentioned properties in crystals. Below we dwell on works aimed at establishing the reasons why the majority of organic molecules tend to crystallise in centrosymmetrical SG.

The preferential occurrence of centrosymmetrical SG may be only indicative of the fact that organic compounds can exist as racemic mixtures of enantiomers rather than reflects the fact that these SG are energetically more favourable. Strictly speaking, to determine correctly the actual frequency of occurrence of SG it is necessary to know whether the crystals were obtained from a solution of a pure enantiomer, from a mixture of enantiomers or from a solution of an achiral compound. The great majority of crystals were grown from mixtures of enantiomers. Therefore, it is not surprising that they belong to centrosymmetrical SG. If this is the case, the problem of the preparation of noncentrosymmetrical materials can be reduced to the problem of spontaneous resolution of enantiomers or to the development of procedures for their forced resolution. It was noted [52] that in recent years the number of noncentrosymmetrical groups in the CSD increased somewhat, which may be indicative of an increase in the number of the natural compounds and products of the stereoselective synthesis studied.§

For the crystalline compounds studied to date, virtually all 230 possible space groups are observed, except for 7 groups for inorganic compounds and 28 groups for organic and organometallic compounds.[83] This fact can be considered as the experimental corroboration of the Kitaigorodskiii theory. However, as for their frequency of occurrence, only certain of space groups are

§ This example provides further evidence that care is required in the data rejection and sample creation. Even a very large body of structural data does not necessarily assure that the statistical characteristics or property is reliable.

Table 3. Distribution of organic crystals among space groups.[42]

Number	SG	*m*	*n* (%)
1	$P2_1/c$	7634	37.97
2	$P2_12_12_1$	3052	15.18
3	$P\bar{1}$	2658	13.22
4	$P2_1$	1684	8.38
5	$C2/c$	1149	5.71
6	*Pbca*	998	4.96
7	$Pna2_1$	414	2.06
8	*Pnma*	257	1.28
9	$Pca2_1$	199	1.00

Note. The following notations are used: SG is the space group; *m* is the number of structures belonging to a particular SG; *n* is the ratio between structures belonging to a particular SG and the total number of the structures under consideration.

rather populated.[7,80] Only a few examples of the remaining groups are available.

Thus only 18 most 'popular' (according to Baur and Kassner [83]) space groups account for 92.71% of all known structures of organic compounds. Belsky [42] has distinguished 9 most frequently occurring groups, which account for 89.8% of all compounds. The nine groups mentioned by Belsky (Tables 3 and 4) enter into 18 groups of Baur (Fig. 9).

Table 4. Distribution of organic crystals among crystal systems.[42]

System	*m*	*n* (%)
Triclinic	2823	14.04
Monoclinic	11061	55.01
Orthorhombic	5563	27.67
Tetragonal	360	1.79
Hexagonal	268	1.33
Cubic	32	0.16

Note. The following notations are used: *m* is the number of structures belonging to a particular crystal system; *n* is the ratio between structures belonging to a particular system and the total number of structures under consideration.

For inorganic compounds, Baur has also distinguished 18 most frequently occurring groups. However, these groups account for only 56.86% of all structures. The distribution histogram of space groups for inorganic compounds is more shallow than that observed for organic compounds (see Fig. 9). In the case of inorganic compounds, it is more difficult to reveal 'leading' group. The dominant space groups for organic and inorganic compounds are also substantially different in spite of the above-mentioned common predominance of centrosymmetrical SG. The most substantial difference is that inorganic compounds crystallise in higher-symmetry space groups, namely, in trigonal, tetragonal, hexagonal and cubic systems. Thus 31.58% of the total number of inorganic compounds crystallise in groups belonging to the symmetry classes 3*m*, 4/*mmm*, 6/*mmm* and *m*3*m*. Only 0.63% of organic compounds belong to the above-listed groups.

According to the data reported by Brock and Dunitz,[45] in the case of protein and polymeric structures, the tetragonal and trigonal systems are also more abundant (13% and 16%, respectively) than in chiral crystals of organic compounds (3% and 2%, respectively). The groups $C2$ and $C222_1$ in the Protein Data Bank account for 13% and 5% of structures, respectively, while these groups in the CSD (for chiral crystals) account for only 4% and 1% of structures, respectively. The populated SG of

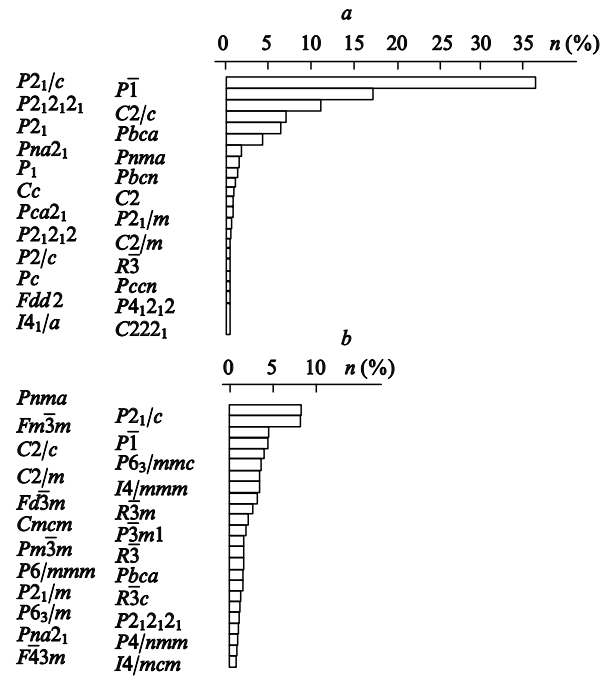


Figure 9. Most populated space groups for organic (*a*) and inorganic (*b*) crystal structures (in percent of the total number of structures).[83]

organic compounds belong exclusively (95.38%) to triclinic, monoclinic and orthorhombic systems (Fig. 10). These crystal systems account for only 48.92% of inorganic compounds. Thus the distribution of inorganic compounds is more diversified.

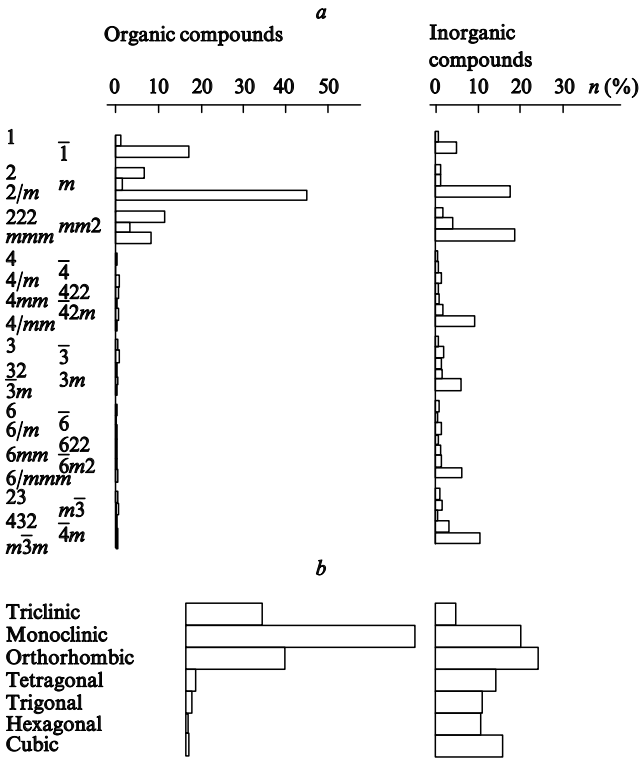


Figure 10. Distributions of frequencies of abundance of crystals of organic and inorganic compounds over classes of point symmetry (*a*) and crystal systems (*b*).[83]

Among the above-listed 18 most frequently occurring SG of inorganic compounds, all crystal systems are present (see Fig. 10).

The majority of organic molecules crystallise in low-symmetry space groups, while the majority of elements, intermetallic compounds and inorganic salts crystallise in high-symmetry groups. However, this does not imply that the packing rules for these compounds are different. The difference in structure is most likely determined by the shape of the initial structural unit as well as by the presence or absence of strong long-range electrostatic interactions. For inorganic compounds, the structural units are most often nearly spherical atoms or ions as well as octahedral or tetrahedral fragments. That is the reason that a large number of inorganic compounds crystallise in high-symmetry groups. The vast majority of organic compounds archived in the CSD consist of low-symmetry electroneutral molecules. Therefore, it is not surprising that organic compounds crystallise predominantly in low-symmetry groups.

Presently, it is generally agreed that the molecular shape affects the symmetry of the crystal structure formed.[38, 39, 45, 84] When estimating the frequency of occurrence of SG, this fact must be taken into account. Brock and Dunitz,[45] like Kitaigorodskii,[7] exemplified this fact by crystals of molecules possessing inherent symmetry elements, namely, a twofold rotation axis (2) and a mirror plane (m). In spite of the obvious fact that these elements are unfavourable,[7] space groups containing the plane m are rather frequent, which is exclusively due to molecules possessing the inherent symmetry m. The plane m in the corresponding SG is virtually always 'occupied'. If molecules possessing the inherent mirror symmetry are excluded from consideration, the distribution pattern of space groups is changed (the above-mentioned group *Pnma* no longer belongs to abundant groups). In a half of the cases, a twofold rotation axis is occupied and in the remaining cases, it is free. In the latter case (for chiral molecules), this axis plays the same role as an inversion centre in the case of a racemic pair, namely, stable dimers are formed due to a twofold axis. Concave-shaped protein molecules often consist of such dimers, which allows them to be closely packed. Apparently, for this reason space groups containing a twofold axis occur more frequently in the Protein Data Bank than in the CSD.[41]

From the analysis of the frequency of occurrence of space groups for organic crystals in the CSD, Brock and Dunitz [45] found what symmetry elements are relatively more favourable. Thus it appeared that the screw axis 2_1 is much more favourable than the simple translation or the screw axes of higher orders. It is generally believed [7] that the axis 2_1 is favoured over the glide plane, which, in turn, is favoured over the translation. However, Brock and Dunitz believe that glide planes play the same role as translations in organisation of crystal structures. This conclusion is not quite without doubt, because the translation is a symmetry element of the first kind under the action of which the chirality of the molecule is retained, while the glide plane is a symmetry element of the second kind, which inverts the chirality. Therefore, it may be that comparisons were carried out based on different classes of compounds, which is not quite correct.

The number of formula units per asymmetric unit (Z') (the number of occupied orbits) is yet another important characteristics necessary for analysing the frequency of occurrence of space groups. Belsky and Zorky, who developed the concept of structural classes,[51,80] were the first to note the importance of the consideration of Z'. Recently, other authors also pointed to the importance of account of this parameter. Padmaja *et al.*[41] found that structures with $Z' > 1$ belong predominantly to low-symmetry crystal systems (particularly often, to the group $P1$). Analysis of the results obtained by Brock and Dunitz with the use of the 1991 version of the CSD [45] confirmed the above conclusions. It appeared that the overall number of structures with $Z' > 1$ in the most populated chiral groups is larger than in the corresponding centrosymmetrical groups:

Group	$P1$	$P2_1$	$C2$	$P2_12_12_1$	$P\bar{1}$	$P2_1/c$	$C2/c$	$Pbca$
$n_{Z'} > 1$, %	45.0	14.0	11.2	4.5	11.3	5.7	3.0	3.0

Brock and Dunitz suggested that many acentric crystals with $Z' > 1$ are actually pseudosymmetrical structures. First sampling tests demonstrated that of 55 of a total of 256 structures with $Z' > 2$ belonging to the group $P2_1$ can be described by the pseudogroup $P2_1/c$. The pseudosymmetry elements $\bar{1}$, 2, 2_1 and t are most often observed in the above structural type. The phenomenon of pseudosymmetry and problems associated with this effect are discussed below in more detail.

VI. Why do organic crystals prefer centrosymmetrical groups?

1. The effect of molecular dipoles on the crystal-packing organisation

Whitsell and coworkers [49] were among the first to perform studies aimed at elucidating the reasons for the preferential occurrence of centrosymmetrical packings with the use of the CSD. The authors attempted to verify the thesis that molecules with large dipole moments crystallise preferentially in centrosymmetrical space groups. This concept is commonly accepted by researchers involved in studies of nonlinear optics of molecular crystals.[46, 47] It is based on simple inferences that the role of the electrostatic component in stabilisation of the crystal structure increases in the case of the antiparallel arrangement of molecular dipoles, which ultimately leads to the formation of centrosymmetrical crystals.

As early as 1955, Kitaigorodskii [7] noted that electrostatic interactions (in the dipole – dipole approximation) must be taken into account for crystals of strongly polar molecules. However, recent calculations performed by Gavezzotti and Filippini [85, 86] for small molecules containing C = O and N ≡ C groups demonstrated that dipole-dipole interactions make only an insignificant contribution to the total energy of crystal stabilisation.

In the estimation of the role of the molecular dipole moment in the crystal symmetry, it was suggested [49] that if the antiparallel orientation of molecules with large dipole moments is preferable, the number of crystals containing such molecules should be substantially larger among centrosymmetrical crystal structures than among noncentrosymmetrical structures.

This suggestion was verified when crystal structures belonging to three space groups ($P\bar{1}$, $P1$ and $P2_1$) were retrieved from the CSD (1994 version). Typical schemes of the molecular arrangements in these groups are shown in Fig. 11. In the group $P1$, all molecular dipoles are parallel to each other due to the absence of symmetry operations other than translations t. Therefore, the average dipole moment of the crystal is determined by the sum of all molecular dipole moments. In the group $P\bar{1}$, an antiparallel neighbour corresponds to each molecular dipole and the total dipole moment is equal to zero. In the group $P2_1$, pairs of molecules are related by a combination of rotation and translation operations and the value of the structural dipole should depend on the orientation of the molecular dipole relative to the screw axis (angle θ). Since the experimental data on molecular dipole moments (μ) are scarce, these values for a large series of

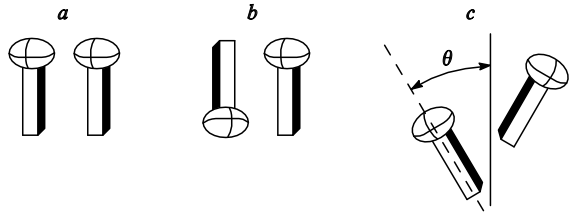


Figure 11. Mutual arrangement of molecular dipole moments in crystals belonging to the space groups $P1$ (*a*), $P\bar{1}$ (*b*) and $P2_1$ (*c*).

compounds were calculated using the AMPAC programme with the AM1 parameterisation.[87]

Crystal structures in the CSD were chosen so that the values of μ could be correctly calculated within the framework of the above-mentioned approximation. Therefore, when statistical samples for the three above-listed space groups were created, the following objects were rejected: (1) compounds in which *d*-orbital interactions can occur; (2) organometallic ionic and polymeric structures; (3) structures containing molecules of solvation; (4) structures in which molecules occupy several systems of equivalent positions and (5) structures in which molecules form strong hydrogen bonds (*i.e.*, compounds containing OH or NH groups). As a result, a total of 28, 161 (out of 1229) and 179 (out of 722) examples were selected in the groups $P1$, $P\bar{1}$ and $P2_1$, respectively.

Table 5. Values of the average dipole moments of organic molecules.[49]

Space group	n	μ /μB	σ	$\bar{\mu}$ /μB
$P1$	28	3.41	0.37	3.36
$P\bar{1}$	161	3.33	0.15	3.22
$P2_1$	179	3.14	0.15	3.04

Note. The following notations are used: n is the number of structures under consideration; σ is the error of calculation of μ; $\bar{\mu}$ is the value of the dipole moment for the midpoint of the distribution.

The results obtained in the work [49] are presented in Table 5 and in the histograms (Fig. 12). It can be seen that the average molecular dipole moments in centrosymmetrical and noncentrosymmetrical SG are virtually identical within the rather small dispersion. No statistically reliable correlation was observed between the dipole moments of molecules and their orientations relative to the polar direction in the group $P2_1$. Hence, the molecular dipole moment is not the governing factor in choosing the space group and does not account for the observed predominance of racemic space groups over chiral groups. However, some effects, such as hydrogen bonds, specific interactions of polar groups and the shape of the molecule, can affect the crystal packing. Apparently, interactions of molecular dipoles play an important role at large distances. In crystals, adjacent molecules are located at distances comparable with their dimensions. Moreover, dimensions of the majority of molecules are substantially larger than the shortest distances between them. Therefore, it is hardly probable that attempts to construct the molecular packing based primarily on the consideration of total molecular dipole moments will succeed.

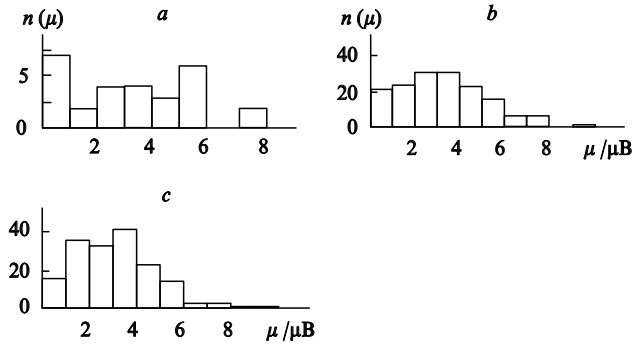


Figure 12. Distribution histograms for molecular dipole moments in the space groups $P1$ (*a*), $P\bar{1}$ (*b*), and $P2_1$ (*c*).[49]

2. Density and stability of racemic crystals and their chiral counterparts

The commonly accepted postulate that explains the preferrence of centrosymmetrical crystals is known as Wallach's rule.[88] According to this rule, racemic crystals tend to be denser than their chiral counterparts.

The question of whether molecules in racemic crystals are packed more closely than in the corresponding chiral crystals has been a subject of speculations since 1895.[88] This question has in essence two aspects:[89] are racemic crystals generally more stable than their chiral counterparts and is this greater stability reflected in a larger density of racemic crystals?

Wallach has deduced his rule by considering 8 chiral – racemic pairs, and only one of these pairs was an exception. More recently, Jacques, Collet and Wilen [89] considered 12 corresponding pairs and found 4 exceptions to Wallach's rule. The improved accuracy of determination of X-ray densities ($\sim 0.15\%$) made it possible more correctly to verify the validity of Wallach's rule. Mason [90] was the first to perform this verification. He analysed 14 pairs of compounds and found 9 exceptions to Wallach's rule. Apparently, only statistical analysis of a large sample can provide a reliable answer to the above question.

For this purpose, Brock and coworkers [50] used the CSD (1989 version). The authors created two files. The first file contained data on only noncentrosymmetrical groups ($\sim 25\%$ instances). The second file contained data on only racemic groups ($\sim 75\%$ instances). The data of the files were sorted according to chemical compositions of compounds and compared. The sets of structures having the same chemical composition were considered as potential objects of investigation. In several instances, compounds filed in the CSD under the same code turned out to be diastereomers. In this case, they were excluded from consideration. Particular attention was given to the problem of ambiguities of space-group determination. The space groups differing only by the presence or absence of an inversion centre were subjected to special scrutiny. For each pair, the relative differences between the densities of racemic (D_r) and chiral (D_c) crystals were calculated as follows:

$$\Delta D = \frac{100(D_r - D_c)}{0.5(D_r + D_c)}.$$

The final list contained 129 pairs. This number of examples was insufficient to obtain the smooth distribution (the shape of the histogram depends on the choice of the bin size and boundaries). Thus, the authors presented two variants of distributions of the values of ΔD, which differ by the choice of the initial boundaries (Fig. 13). In spite of the differences, both distributions exhibit a pronounced bimodal character. The average value of ΔD for 129

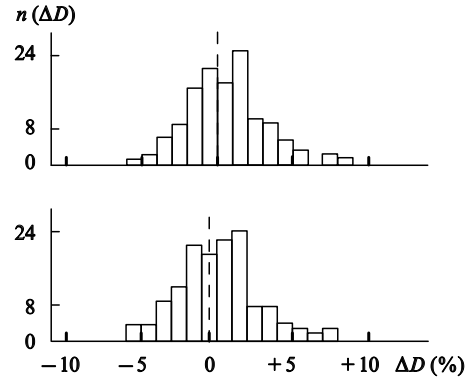


Figure 13. Distribution histogram for the values ΔD for 129 chiral – racemic pairs of the structures under study (two variants of the choice of the initial bin size are given).[50]

pairs was 0.56%. However, this value is too small to confirm the rule, but it is too large to demolish it.

Taking into account the bimodality of the ΔD distribution, the authors attempted to divide the selected objects into two groups. One way of dividing objects is to separate pairs which are polymorphic modifications. An alternative way is to separate pairs which are not polymorphs. This separation was carried out according to the definition of polymorphs suggested by McCrone.[91] According to McCrone, polymorphs are different solids that give the same liquid upon melting or dissolution. Therefore, crystal pairs of the first group (polymorphic modifications) contain either achiral molecules or chiral molecules that can rapidly undergo racemisation in solutions. Crystal pairs of the second group consist of enantiomers that can be readily resolved

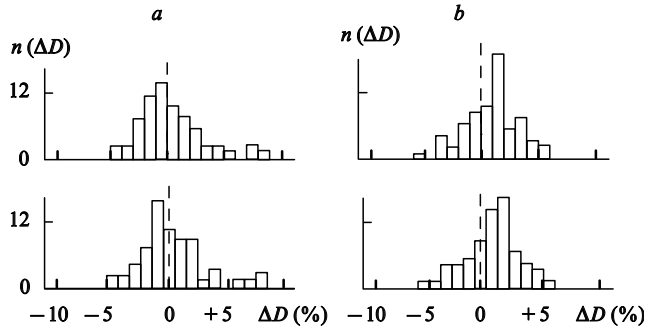


Figure 14. Distribution histograms for the values ΔD for 64 pairs of structures of the first group (enantiomers that can be rapidly interconverted, achiral compounds) (*a*) and for 65 pairs of the structures of the second group (enantiomers that can be resolved) (*b*).[50]

(consequently, they are difficultly interconvertible).

For 64 pairs of the objects of the first group, the average value of ΔD was *ca.* +0.20% (Fig. 14*a*). Consequently, a comparison of polymorphic pairs showed no noticeable difference between the packing densities of racemic and chiral crystals.

The average ΔD value for the second group of objects consisting of 65 pairs was +0.92% (Fig. 14*b*). The packing densities of racemic crystals of the second group are, on the average, $\sim$ 1% higher than those of their chiral counterparts. Therefore, it was concluded that Wallach's rule is apparently valid only for enantiomers that can be resolved.[50]

It is not inconceivable that the difference in the average values of ΔD for the groups of crystals under consideration is due to different initial conditions of crystallisation. Compounds of the first group are always crystallised from solutions of achiral compounds or mixtures of enantiomers, while compounds of the second group are crystallised from solutions of both enantiomeric mixtures and pure enantiomers.

Crystallisation of compounds of the first group from solutions depends on the relative stability of possible solid phases and sometimes affords several crystal forms. Evidently, the stabilities of these forms should be approximately equal. Enthalpies of polymorphic modifications rarely differ by more than 1 kcal $\cdot$ mol^{-1}. Therefore, it is hardly probable that less stable modifications of compounds of the first group will be obtained.

For compounds of the second group, the situation is quite different. Racemic solutions may yield racemic crystals, racemic conglomerates of chiral crystals or mixtures of both types depending on their relative stability. In contrast, crystals obtained from enatiomerically pure solutions must be chiral. Therefore, chiral crystals can be obtained even if they are thermodynamically much less stable than the racemic crystals. This implies that the second group of compounds is statistically biased. Thus, pairs in which racemic crystals are energetically more favourable can be more frequent in the second group than in the first one. This is reflected in different distributions of ΔD.

Jacques and coworkers[89] compared thermodynamic characteristics of the melting process for 36 pairs of chiral and racemic crystals belonging to the second group. The average values of such characteristics as ΔH, ΔS and ΔT are 7.54 kcal mol^{-1}, 18.5 kcal mol^{-1} K^{-1} and 405 K for racemic crystals and 6.28 kcal mol^{-1}, 15.7 kcal mol^{-1} K^{-1} and 395 K for chiral crystals. Based on these data, the authors concluded that racemic crystals are thermodynamically more stable than chiral crystals.

However, it is hardly reasonable to consider this conclusion to be the general rule because the statistical data (in particular, the thermodynamic characteristics) are still scarce. It is quite probable that racemic crystals appear to be more stable because pairs in which they are less stable are absent only among compounds that have already been studied rather than among naturally occurring compounds.

Therefore, statistically reliable evidence for the higher density and thermodynamic stability of racemic crystals compared to their chiral counterparts has not yet been found.

3. Comparison of thermodynamic characteristics of polymorphic modification of organic crystals

The phenomenon of polymorphism in organic crystal structures has long been known. Such aspects as thermodynamic stability, pharmaceutical properties of polymorphs, conformational polymorphism, *etc.* have been studied in considerable detail.[4,92–98] Unfortunately, the data available in the literature are primarily descriptive and each instance of polymorphism was considered by itself. Accumulation of a large body of structural data on polymorphic modifications (PM) in the CSD brought systematic statistical studies in this field to reality. The existence of polymorphs is still a headache problem for scientists involved in computational modelling of molecular packings. However, it is evident that due to the constancy of the chemical composition in polymorphic pairs (triads, *etc.*) these systems can serve as a potential source of detailed information on the structure–property relationship in organic solids because in this case the difference in the physical properties can be explained by the difference in their structural organisations.

The first investigation devoted to this problem was performed by Gavezzotti and Filippini.[3] The authors compared the thermodynamic and structural characteristics of the polymorphs available in the 1993 version of the CSD. They selected only objects which were studied at room temperature, had more than one polymorphic modification and consisted of C, H, N, O, Cl and S atoms. A total of 163 compounds which satisfy these requirements were found, 147 of them with two, 13 with three and three with four polymorphic modifications. A total of 345 crystal structures were found.

For these polymorphic modifications, thermodynamic characteristics of crystals were calculated according to procedures reported in the Refs 3, 86, 99 and 100. The parameters for the 6-exp potential function for intermolecular interactions were chosen[98,100] so as to reproduce the values of thermodynamic characteristics at room temperature. Generally, the calculated values of thermodynamic characteristics agree, at best, only qualitatively with the experimental values (due to an inaccuracy of both calculated and experimental values). Nevertheless, the regularities found by the authors can be considered as justified because possible inaccuracies affect in the same manner the results of calculations for both forms, while the differences between the calculated values used by the authors are free from errors.

The polymorphic pairs of compounds were placed in order of decreasing density. Therefore, the differences between the densities ΔD for polymorphic modifications are positive by definition. The differences (%) in the other properties (P) were determined as follows:

$$\Delta P = \frac{100(P_i - P_j)}{P_j}, \quad j > i.$$

The total number of points was 204. The following parameters P were calculated: D, the crystal density; V, the molecular volume; K, Kitaigorodskii's packing coefficient; E, the packing energy; S, the lattice-vibrational entropy; G, the lattice free energy, and Z', the number of molecules per asymmetric unit.

The ΔD, ΔE, ΔS and $\Delta Z'$ distribution histograms are shown in Fig. 15. It appeared that the difference in the density between polymorphs is small and seldom exceeds a few percent. Thus 93%

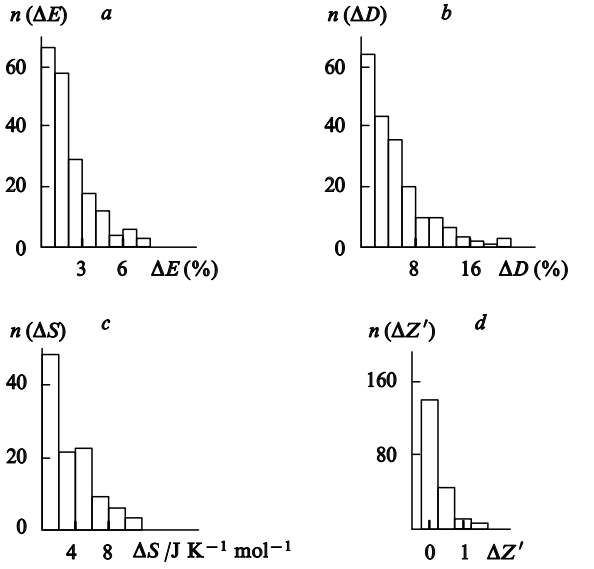


Figure 15. Distribution histograms for the packing energy ΔE (*a*), ΔD (*b*), ΔS (*c*), and $\Delta Z'$ (*d*).[3]

of the values of ΔD are smaller than 5%, which confirms the results reported in the work.[50] This implies that molecules can be packed in different fashions, but, apparently, not at the expense of a decrease in compactness. Cases with $\Delta D > 7\%$ occur either due to inaccuracy in determination of atomic coordinates or to large conformational differences in the molecules.

Similar conclusions are drawn by examination of the ΔE histograms: the packing energies of PM are generally close. A few exceptions can be attributed to the structural disorder or to inaccurate positioning of some key atoms of the system. The differences in the lattice entropies ΔS between polymorphs in pairs are also very small (see Fig. 15).

Thus the density, packing energy and lattice entropy for a polymorphic pair have close values. The results on the number of molecules per asymmetric unit (Z') are unexpected: 62 structures (18% of the pairs selected) have $Z' > 1$ [this value is substantially higher than the overall percentage in the CSD (8.3%)] and 46 compounds (28%) of 163 compounds selected have modifications with different values of Z'.

The authors considered also possible relationships between the calculated values (bivariate statistics). The dependence of ΔD on ΔE (Fig. 16*a*) demonstrates that in most cases higher packing energy is accompanied by higher density. The existence of crystal structures with high energies and low packing densities can be attributed to the presence of exceptionally strong and directional hydrogen bonds in a crystal. In all cases under consideration, the differences in the free energies ΔG between polymorphs in pairs have the same sign as the differences in the packing energies ΔE (Fig. 16*b*). As expected, the higher the density the lower the entropy (Fig. 16*c*). These results demonstrate that higher crystal

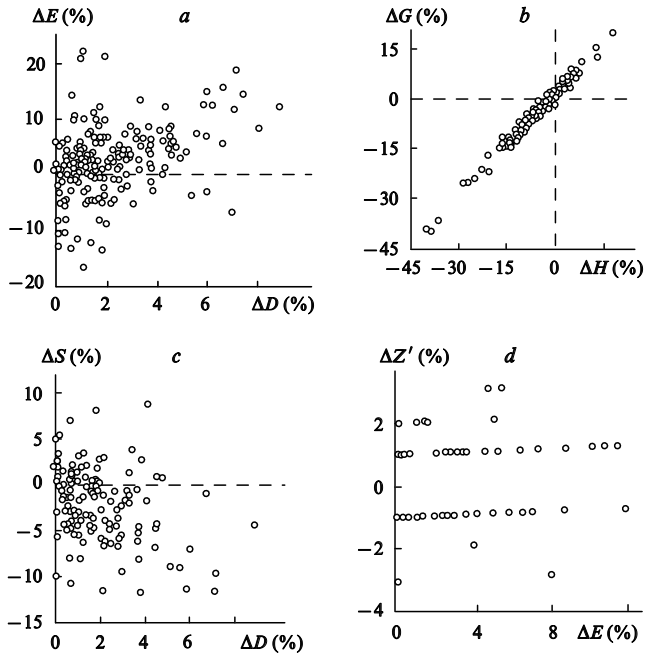


Figure 16. Relationships between ΔE and densities ΔD (*a*), ΔG and enthalpies ΔH (*b*), ΔS and ΔD (*c*), and $\Delta Z'$ and ΔE (*d*).[3]

density corresponds to higher packing energy and lower vibrational entropy.

The dependence of $\Delta Z'$ on ΔE is shown in Fig. 16*c*. The structures with $Z' > 1$ are more stable in $\sim 50\%$ of the occurrences. Hence, the presence of more than one molecule per asymmetric unit dose not lead to a decrease in the stability of a crystal.

Therefore, the knowledge of thermodynamic characteristics of polymorphs existing at room temperature is insufficient for explanation of their relative stability because the differences in these characteristics are generally small. Thus, it is useful to analyse molecular-packing symmetry to gain insight into the ways of formation of crystal structures in different polymorphic modifications.

The space group remains the same for 28% of the polymorphic pairs under consideration; 82% of all structures belong to the group $P2_1/c$ and the remaining 18% of the structures are distributed among the groups $P\bar{1}$, $Pbca$, $C2/c$ and $P2_12_12_1$. The greatest number of plausible arrangements of the same molecule was observed for organic crystals belonging to the group $P2_1/c$.

As mentioned above, one of polymorphic forms often has $Z' > 1$. Although a detailed analysis of such cases has not been performed, preliminary consideration of the data revealed instances in which molecules in the asymmetric unit are related by pseudosymmetry operations. A detailed examination of pseudosymmetry relations in crystals will allow one to 'visualise' possible ways of formation of crystal structures, which have not been realised.

The authors also noted that in 20 polymorphic pairs (24%) the molecules decide between centrosymmetrical and noncentrosymmetrical space groups. Ten pairs $P2_1/c - P2_12_12_1$, six pairs $P2_1/c - P2_1$, four pairs $P2_1/c - Pc$ and four pairs $P2_1/c - Pna2_1$ were found.

Comparison of thermodynamic characteristics of the polymorphs did not allow the authors of the works cited (Refs 3, 49 and 50) to reach definite conclusions about regularities of organisation of crystal structures. However, although the analysis of structural regularities in polymorphic modifications did not provide answers to the problems under considerations, it doubtedlessly made it possible to mark the route and develop a strategy of further investigations in this field.

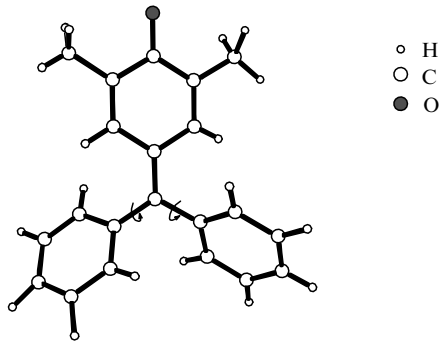


Figure 17. Molecular structure of 2,6-dimethyl-4-(diphenylmethylene)-cyclohexa-2,5-dienone (DMFUSC).[101, 102]

We attempted to search for PM containing chiral–racemic partners in the 1996 version of the CSD. A total of 102 such pairs were found. The same distribution pattern of chiral and racemic modifications was observed: 45 pairs (44.1%) $P2_1/c-P2_12_12_1$, 17 pairs (16.7%) $P2_1/c-P2_1$, 14 pairs (13.6%) $P2_1/c-Pna2_1$, 10 pairs (9.8%) $P2_1/c-Pc$, 3 pairs (2.9%) $P2_1/c-P\bar{1}$, 7 pairs (6.9%) $P\bar{1}-P2_12_12_1$, 2 pairs (2.0%) $P\bar{1}-P2_1$, 2 pairs (2.0%) $P\bar{1}-Pc$ and 2 pairs (2.0%) $P\bar{1}-Pna2_1$. We hope to reveal reasons for the preferential occurrence of these chiral–racemic pairs from a detailed analysis of structures of these polymorphic modifications.

Let us offer an example of a typical situation for the $P2_1/c-P2_12_12_1$ pair. In the case of 2,6-dimethyl-4-(diphenylmethylene)cyclohexa-2,5-dienone[101, 102] (DMFUSC),¶ the molecules adopt identical conformations in both polymorphic modifications. The chirality in the molecule (Fig. 17) is determined by the difference in the rotation of the phenyl rings. In solutions, both enantiomeric forms are present due to rotation of the rings. In both crystals, molecules of one enantiomeric form are packed in identical layers (Fig. 18). The adjacent layers in the β form ($P2_12_12_1$) are stacked in a congruent fashion, while in the α form ($P2_1/c$) the adjacent layers are related by centres of symmetry. The α–β phase transition was observed at 383 K. Therefore, in this case two polymorphic modifications exist due to different modes of packing of layers and the enantiomeric forms readily undergo interconversion with a small consumption of energy. Analogous

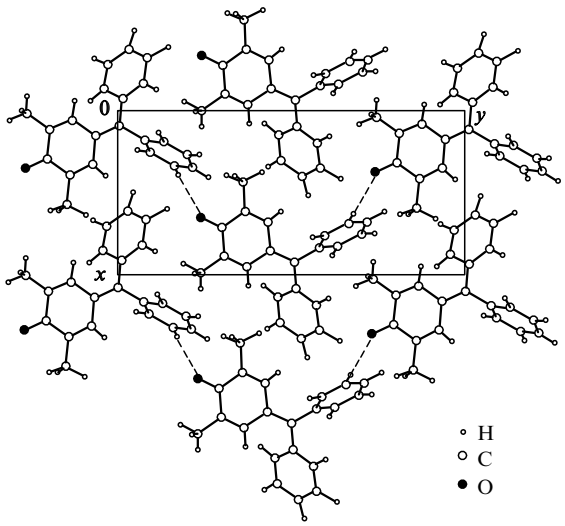


Figure 18. Structure of the layer in the orthorhombic modification of DMFUSC (projection onto the *xy*0 plane). The layer in the monoclinic modification has an analogous structure.

situations were observed in di(1-piperidyl) disulfide[103] (DPIPDS),¶ *o*-chlorobenzamide[104] (CLBZAM),¶ *m*-nitrophenol[105, 106] (MNPHOL),¶ and bis(trimethylbenzylammonium) tetrachlorocuprate(II)[107] (MBAMCC).¶ These compounds can be assigned to the first group, *i.e.*, they are enantiomers that interconvert rapidly.

VII. Number of molecules per asymmetric unit and pseudosymmetry

As mentioned above in the section devoted to the relative frequency of occurrence of space groups, the increasing attention is paid to crystal-chemical analysis of structures with $Z' > 1$. Several computer programmes have already been developed for revealing pseudosymmetry in such crystal structures.[108, 109] These programmes were devised originally for revealing improper symmetry in structural studies and therefore they treat pseudosymmetry as a 'negative factor'. The main purpose of these programmes is to determine deviations of atoms from certain ideal symmetry positions, which imply that some symmetry operations were missing in the course of determination of crystal symmetry.

Initially, when only a few instances of crystals containing two independent molecules per asymmetric unit were known, researchers pinned hopes on comparison of their conformations and molecular geometry. This approach seemed to be very promising for studying the effect of crystal environment on molecular structure. However, expectations were not realised completely because such structures remained for long few in number.

Studies of noncrystallographic (pseudosymmetrical) symmetry operations that related independent molecules in crystals have attracted growing interest as the corresponding structural data has been accumulated.[45, 53, 54] Analysis of the data in the 1991 version of the CSD demonstrated that in 27% of structures with $Z' > 1$. the molecules are related by approximate pseudosymmetry elements $\bar{1}$, 2, 2_1 and *t* (translation), *i.e.*, by elements which are also present most often in space groups of crystals.† Undoubtedly, a pseudocentre of symmetry is the most frequently occurring element.

In the next stage, relationships between space groups describing the crystal symmetry and pseudosymmetry transformations were searched for. The symmetry of real crystals often appeared to be a symmetry subgroup of a certain 'hypothetical' crystal. It is derived by 'eliminating' the corresponding symmetry element (which becomes a pseudoelement) from the symmetry group of the hypothetical crystal. Thus, Brock and Dunitz[45] found 256 crystals belonging to the group $P2_1$ with $Z' = 2$ in the 1989 version of the CSD of which 55 crystal structures contain an inversion pseudocentre and are described by the symmetry pseudogroup $P2_1/c$. It was suggested that the great majority of structures with $Z' > 1$ belonging to the group $P1$ are actually pseudosymmetrical structures.

The application of the CSD made it possible to reveal also some other statistical regularities. Padmaja *et al.*[41] noted that structures with $Z' > 1$ are substantially more frequent in low-symmetry crystal systems (especially in the group $P1$). These results were confirmed.[45] It was demonstrated that structures with $Z' > 1$ are, on the whole, more abundant in chiral groups than in the corresponding centrosymmetrical groups (see Section IV). Certain preliminary conclusions were made about classes of chemical compounds for which structures with $Z' > 1$ are most often realised. In particular, ~ 40% of crystals of aliphatic alcohols in the CSD have $Z' > 1$. Craven[110] noted that in ~ 50% of crystals of cholesterol derivatives, $Z' > 1$.

We performed sampling of structures of 3- and 4-hydroxy derivatives of hydropyridine in the 1996 version of the CSD. A total of 320 structures were found of which 21% instances have

¶ The reference code of the compound in the CSD.

† R Davies and A Willer (private communication; see Ref. 45).

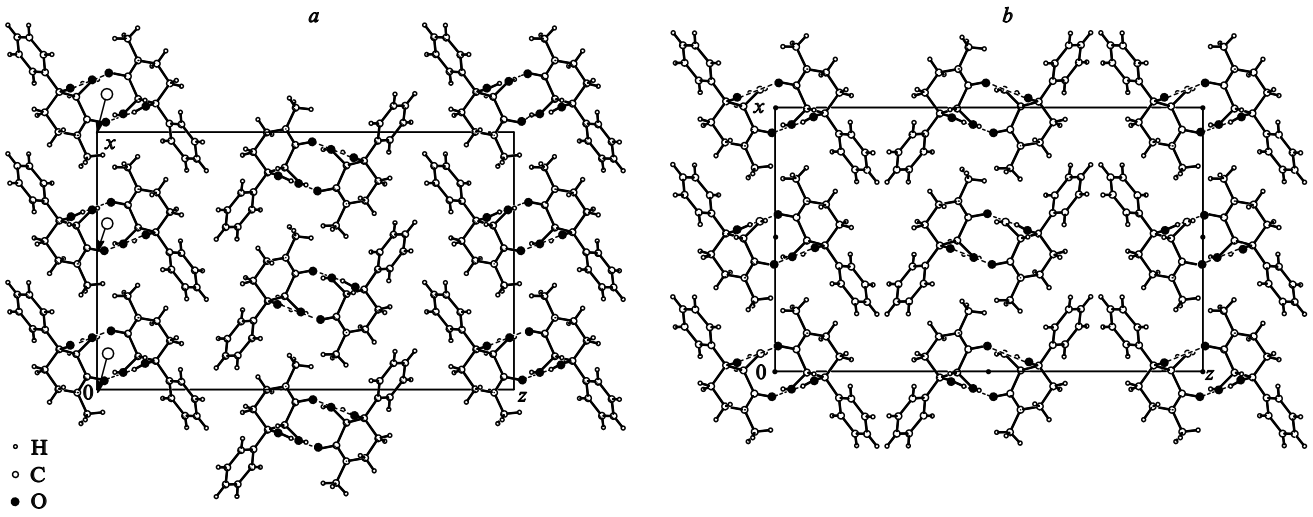


Figure 19. Projection of the crystal structure of 3,4-dihydroxy-1-methyl-2-oxo-4-phenylpiperidine onto the $x0z$ plane (*a*) and the hypothetical crystal structure obtained by shifting pseudosymmetrical chains in the crystal (*b*).[111]

$Z' > 1$. Apparently, for all classes of compounds containing hydroxy groups there is a high probability of formation of crystals with $Z' > 1$. It is known that compounds containing hydroxy groups form stable associates in crystals through hydrogen bonds. It can be assumed that these associates are present even in the initial mother liquor from which crystals are grown. Thus single crystals can be formed from associates rather than from isolated molecules. The packing of associates does not necessarily lead to the optimum symmetry due to insignificant shifts required for closer packing.

Let us exemplify the aforesaid by the crystal structure of 3,4-dihydroxy-1-methyl-2-oxo-4-phenylpiperidine,[111] which crystallised in the space group $Pca2_1$, $Z' = 2$. Two independent molecules are linked in a pseudocentrosymmetrical dimer through hydrogen bonds. These dimers are linked in infinite chains through hydrogen bonds, and the crystal structure as a whole is formed from these layers (Fig. 19*a*). This noncentrosymmetrical structure can be transformed into the centrosymmetrical one (symmetry *Pbcn*) by slightly shifting the chains with respect to each other. The shift is equal to ~ 1.5 Å. The pseudocentrosymmetrical dimers can be seen in Fig. 19*a*. Chains formed in the structure through hydrogen bonds are located perpendicular to the plane of the paper. The hypothetical crystal structure and the molecular arrangement are shown in Fig. 19*b*. It can be seen that the shift of chains in the real crystal with respect to each other leads to a closer packing.

The probability of the detection of $Z' > 1$ and consequently, of the presence of pseudosymmetry is also very high for crystals of mesogenic compounds, which are solid-crystalline precursors of liquid crystals (among them are cholesterol derivatives).[110] It is known that mesogenic compounds are prone to aggregation in the liquid state, which does not necessarily occur through hydrogen bonding. For example, in four of eight crystal structures of cholesteryl benzoate derivatives, which do not form intermolecular hydrogen bonds,[58] $Z' > 1$. Calculations of the energy of intermolecular interactions demonstrated that molecules are packed in crystals in a layered-stacked fashion and the strongest interactions occur between molecules related either by screw axes or by translations along the stack axis. The existence of this energy-distinguishable direction is very important because weaker interactions between layers and stacks are primarily weakened when passing to the mesophase. In this case, shifts of stacks with respect to each other as well as slight mutual rotations of adjacent molecules in stacks about stack axes become possible. Rotations result in the appearance of a supramolecular structure twisted about the stack direction. This structural feature emphasises the role of screw axes in cholesterogenic crystals. The results reported in the work [58] suggested that a genetic relationship exists between crystalline and liquid-crystalline structures and can provide the basis for further systematic studies with the use of the CSD data aimed at the design of new compounds potentially exhibiting liquid-crystalline properties.

Interestingly, at least one polymorphic form often has $Z' > 1$. The overall percentage of structures with $Z' > 1$ in the CSD is 8.3%, while twice as many such structures are observed among polymorphic modifications (16%).[3] Thus, the presence of more than one independent molecule per asymmetric unit may indicate that several different crystal packings can exist.

The existence of pseudosymmetry in a particular crystal structure is an indication that a distorted higher-symmetry structure can exist. Based on this fact, Abrahams [53, 54] suggested that if this distortion is rather small, the crystal structure would be expected to become more symmetrical at higher temperature as a result of phase transition. Recall that in the cases of phase transitions of the first and second kinds, the crystal symmetry changes abruptly at the transition point. In the case of phase transitions of the first kind, a correlation between the structures of the initial and newly formed phases can be either present or absent. The characteristic feature of phase transitions of the second kind is that the symmetry group of the low-symmetry phase is a subgroup of the symmetry group of the other phase, because only some symmetry elements disappear as a result of displacements of atoms, while other elements persist. It is the latter that form a subgroup. A higher-symmetry phase corresponds, as a rule, to the high-temperature modification. These reasonings provided the basis for a procedure proposed by Abrahams for a search for new crystals possessing ferroelectric properties in the crystallographic databases. This method is based on the analysis of the available structural data. The ferroelectric phase transition is accompanied by loss of an inversion centre, a mirror plane m or a rotation axis 2. If H is a polar space group of the structure under study, the minimum supergroup G of the paraelectric phase can be written as follows:

$$G = H + \left(\frac{g}{t}\right)H,$$

where the operator $g(\bar{1}, m, 2)$ represents the rotational component. In the general case, the orientation of a plane m or an axis 2 is limited by the space group H, but the translation term (t) of the equation can take any value within limits determined by the group G. To put it differently, an inversion centre, a mirror plane or a

rotation axis that are lost can be located in any position with respect to the original one chosen for the description of the group H. Abrahams restricted the consideration to pseudocentres and mirror pseudoplanes. Polar crystal structures in which deviations (shifts) of atoms Δr with respect to hypothetical nonpolar configurations lie in the range between 0.1 and 1 Å were considered as materials that will, most probably, exhibit ferroelectric – paraelectric phase transitions. It was suggested that the maximum deviation of atoms Δr along the polar axis from the totally symmetrical position be considered as a quantitative parameter for estimating the degree of pseudosymmetry of the structure. On the basis of this parameter, the following empirical equation was derived for predicting the transition temperature:

$$T_c = C(\Delta r)^2 ,$$

where $C \approx 2.0 \times 10^4$ K Å^{-2}. Abrahams has cast doubt on space groups determined for structures with $\Delta r < 0.1$ Å. Compounds satisfying the above conditions [54] were selected in the Inorganic Crystal Structure Database (1987). For these compounds, more than 50 new ferroelectric crystals were predicted.

Based on the approach developed by Abrahams, Igartua and coworkers [55] suggested a procedure for systematic studies of materials exhibiting both kinds of high-temperature phase transitions. The authors suggested that compounds possessing pseudosymmetry under ordinary conditions would possess structural phase transitions at elevated temperatures. This was exemplified with the space group $P2_12_12_1$. A total of 442 structures were retrieved from the Inorganic Crystal Structure Database and for 407 of them the values Δr were calculated. It appeared that in all 14 compounds, for which high-temperature phase transitions and pseudosymmetry were reported, $\Delta r < 1.5$ Å. About 20 structures with pseudosymmetry and $\Delta r < 1$ Å were additionally found for which phase transitions are highly probable.

Therefore, it is evident that systematic studies of pseudosymmetrical crystals are promising for establishing structural regularities and for design of materials possessing desired physical properties.

An analogous search for organic crystals in the CSD was not carried out. However, it is possible to trace the route to a search for compounds with the desired physical properties in the database based on certain crystal-structural 'indications' (for example, values of $Z' > 1$). In particular, systematic analysis of structures of long-chain molecules with $Z' > 1$ will apparently allow one to reveal crystals of new mesogenic compounds. A successful search for new nonlinear-optical noncentrosymmetrical crystals in the CSD was carried out based on a search for molecules similar in geometry to dicyanovinylanisole (a known nonlinear-optical material).[112] It appeared that the CSD already contained data on the crystal structure of *p*-chlorodicyanovinylbenzene,[113] which possesses high ability to generate the second harmonic of laser radiation in the crystal.

VIII. Supramolecular synthons and the possibility of prediction of crystal structures

A governing role of intermolecular interactions, in particular of hydrogen bonds, in crystal structures of organic molecules became apparent about 40 years back.[114, 115] At that time, it was already noticed that hydrogen bonds favour the formation of stable molecular associates in crystals. A series of works by Leiserowitz and coworkers in 1969 – 1978 [116 – 118] can be considered as the first systematic study in this field. In these works, the authors attempted to derive possible molecular associates that are realised in some classes of chemical compounds, namely, in carboxylic acids and in primary and secondary amides. The geometrical characteristics of hydrogen bonds in crystals were determined for the first time [119] based on statistical data on the published crystal structures and the theory of graphs was applied for deducing theoretically possible fragments formed through hydrogen bonds in crystals. In addition, typical and anomalous modes of formation of such molecular associates were determined based on the analysis of more than 2000 structures.[120]

The determination of the geometrical parameters of different types of hydrogen bonds based on more detailed statistical data became possible as the CSD data have become available. Taylor and Kennard [23] were the first to determine the geometrical parameters of CH···O interactions. These interactions are commonly assigned to weak hydrogen bonds. Berkovitch-Yellin and Leiserowitz [121] noted that these weak CH···O interactions in crystals along with hydrogen bonds can lead to the formation of stable molecular associates. The works devoted to the consideration of the role of CH···O hydrogen bonds in the formation of crystal structures were most completely surveyed in the review.[29]

More recently, it was shown that the majority of functional groups form a limited number of graph-sets of associates, which are found repeatedly in different structures. Moreover, even chemically different functional groups often form associates described by the same graph-sets.[122] Each type of interaction is characterised by an inherent ideal geometry of the contact. Therefore, the knowledge of functional groups of a molecule makes it possible to predict, with a high probability, associates that will be present in crystals.

Etter [122] derived rules for formation of hydrogen bonds in crystals. She has supplemented Donohue's principle, according to which all 'acid' hydrogen atoms in solids participate in hydrogen bonds, by the addition of the following concepts: (1) the hydrogen-bond acceptors are used to the maximum extent if H-bond donors are available and (2) in crystals containing several potential donors and acceptors, H bonds are preferentially formed between the best donor and the best acceptor. The latter principle is very helpful in the analysis of crystal structures of polyfunctional molecules because it makes it possible to reveal a 'hierarchy' of formation of hydrogen bonds. Based on this principle in the case of branched systems of H bonds in crystals, Bernstein *et al.*[61] suggested that systems of H bonds be divided into separate subfragments with account of intermolecular distances and symmetry operations of space groups. This approach allows one to unify a procedure for revealing molecular associates formed through H bonds in all crystals listed in the CSD. Etter and Bernstein were the first to perform studies in the field of crystal engineering.[61, 122]

Presently, investigations in this field follow two paths. The first aim is to reveal principles of molecular aggregations in crystals based on studies of the diversity of the available crystal structures. The second line of investigation is aimed at the development of calculation procedures for crystal structures design on the basis of only the molecular geometry.[1 – 3] These two lines are 'two sides of a coin' and are aimed at solving the direct and inverse problems of crystal engineering.

In studies devoted to the first problem, the nature of forces owing to which molecules in crystal structures are linked in associates (hydrogen bonds, other weaker specific interactions or even usual van der Waals interactions) is first established and then one tries to reveal structural fragments consisting of most strongly bound molecules (supramolecular synthons). Works devoted to the second line of investigation are based on the molecular geometry. One tries to calculate the most favourable mutual arrangement of two molecules (dimers), of chains or planes formed by molecules and finally of a three-dimensional structure.[1] Alternative calculation algorithms are also available (see Ref. 31).

Note that chemists involved in computation of crystal structures address themselves more and more often to structural databases (most often to the CSD) primarily for choosing potential-function parameters and also for testing calculation procedures with the use of a large number of the published crystal structures. The algorithm (FlexCryst) which was developed [1] for predicting organic crystal structures was used for calculating 131 crystal structures belonging to the space group $P1$ and 95 crystal structures belonging to the space group $P\bar{1}$ retrieved from the

CSD. The agreement between the calculated and the real structures was 98% and 85% for the groups $P1$ and $P\bar{1}$, respectively. Needless to say that these results are very promising. However, this algorithm does not presently allow calculations for other space groups. The algorithm MOLPAK [2] allows one to treat also the space groups $P2_1$, $P2_1/c$ and $P2_12_12_1$. A total of 242 crystal structures were calculated with the use of this program. The programme PROMET allows one to make rather reliable predictions about realisation of a particular space group. This programme made it possible to calculate polymorphic modifications of sulfathiazole, probucol and aspirin.[3]

IX. Conclusion

Investigations with the use of the CSD as a tool in studying general regularities of organic crystal structures demonstrated that this approach produces good results. Thus the principle of close packing of organic and organometallic compounds in crystals was confirmed by statistical analysis of positions of molecular geometrical centres in crystal unit cells. The statistical values of the van der Waals radii of nonmetals were determined based on the analysis of a large sample of nonbonded interatomic distances in the published structures. The preferential types of molecular packings of organic molecules were revealed and an overwhelming predominance of centrosymmetrical space groups was noted. However, statistical analysis demonstrated no evidence that centrosymmetrical packings are energetically preferable to noncentrosymmetrical structures. Neither gain in density nor noticeable differences in the thermodynamic characteristics were revealed.

Conceivably, a predominance of centrosymmetrical space groups reflects the fact that compounds can exist as racemic mixtures of enantiomers rather than the fact that they are energetically more favourable. If this is the case, then crystals belonging to noncentrosymmetrical space groups can be obtained if one establishes the reasons for spontaneous resolution of enantiomers upon crystallisation or finds a procedure for their forced resolution. It is also probable that noncentrosymmetrical crystals just have received lesser attention and, in spite of a large body of data in the CSD, a sample of compounds in this case is not yet representative. Statistical studies confirmed that molecular association in crystals through hydrogen bonds plays a governing role in the formation of crystal packings. Some structural features of crystals whose packings contain stable synthons (dimers, chains or layers) were revealed. These are, first of all, a high probability of occurrence of polymorphic modifications, structures with $Z' > 1$ and with pseudosymmetry and cases of spontaneous resolution of isomers. These structures often possess (or can possess) useful physical properties. A procedure was developed for the search for new ferroelectrics and other materials exhibiting phase transitions.

Analysis of the results of the first systematic statistical studies of organic crystal structures with the use of the CSD allows one to draw up a plan of further investigations in this field. Thus the first stage of studies of the reasons for spontaneous resolution of enantiomers upon crystallisation should apparently involve the establishment of types and forms of molecules that can undergo spontaneous resolution and analysis of the role of molecular association in crystals (apparently, from the initial solution). In this respect, it is very advantageous to compare crystal packings of compounds having centro- and noncentrosymmetrical polymorphic modifications. A search for compounds exhibiting polymorphic transformations and compounds potentially possessing liquid-crystalline and nonlinear-optical properties in the CSD can be rather simple. These are only some applications of the CSD.

References

1. D W M Hofmann, T Lengauer *Acta Crystallogr., Sect A* **53** 225 (1997)
2. J R Holden, Z Du, H L Ammon *J. Comput. Chem.* **14** 422 (1993)
3. A Gavezzotti, G Filippini *J. Am. Chem. Soc.* **117** 12299 (1995)
4. A Gavezzotti *Acc. Chem. Res.* **27** 309 (1994)
5. L Pauling, in *The Nature of the Chemical Bond* (Ithaca, New York: Cornell University Press, 1948) p. 187
6. A J Bondi *J. Phys. Chem.* **68** 441 (1964)
7. A I Kitaigorodskii *Organicheskaya Kristallokhimiya* (Organic Crystal Chemistry) (Moscow: Izd. Akad. Nauk SSSR, 1955); *Molekulyarnye Kristally* (Molecular Crystals) (Moscow: Nauka, 1971)
8. P P Ewald, C Hermann *Strukturberichte 1913–1928* (Leipzig: Academische Verlagsgesellschaft, 1929)
9. A I Kitaigorodskii, P M Zorky, V K Belsky *Stroenie Organicheskogo Veshchestva* (The Structure of Organic Substance) (Moscow: Nauka, 1980) Vol. 1
10. A I Kitaigorodskii, P M Zorky, V K Belsky *Stroenie Organicheskogo Veshchestva* (The Structure of Organic Substances) (Moscow: Nauka, 1982) Vol. 2
11. D G Watson, in *CODATA Directory of Data Sources for Science and Technology* (Paris: CJLFNF, 1977) p. 15
12. D G Watson, in *Crystallographic Data Bases* (Eds F H Allen, G Bergerhoff, R Sievers) (Chester: International Union of Crystallography, 1987) p. 25
13. R Jenkins, D K Smith, in *Crystallographic Data Bases* (Eds F H Allen, G Bergerhoff, R Sievers) (Chester: International Union of Crystallography, 1987) p. 158
14. A D Mighell, J K Stalick, V L Himes, in *Crystallographic Data Bases* (Eds F H Allen, G Bergerhoff, R Sievers) (Chester: International Union of Crystallography, 1987) p. 134
15. F H Allen, O Kennard *Chem. Design Autom. News* **8** (1) 31 (1993)
16. F H Allen, O Kennard, D G Watson, L Brammer, A G Orpen, R Taylor *J. Chem. Soc., Perkin Trans 2* S1 (1987)
17. A G Orpen, L Bramer, F H Allen, O Kennard, D G Watson, R Taylor *J. Chem. Soc., Dalton Trans. 1* S1 (1989)
18. F H Allen, in *Modelling of Structure and Properties of Molecules* (Ed. Z B Maksic) (Chichester: Horwood-Wiley, 1987) p. 51
19. A I Yanovskii, Yu T Struchkov, in *Problemy Kristallokhimii* (The Problems of Crystal Chemistry) (Moscow: Nauka, 1984) p. 161
20. Yu L Slovokhotov, I V Moskaleva, V I Shilnikov, E F Valeev, Yu N Novikov, A I Yanovsky, Yu T Struchkov *Mol. Mater.* **8** 117 (1996)
21. S C Nyburg, C H Faerman *Acta Crystallogr., Sect. B* **41** 274 (1985)
22. R S Rowland, R Taylor *J. Phys. Chem.* **100** 7384 (1996)
23. R Taylor, O Kennard *J. Am. Chem. Soc.* **104** 5063 (1982)
24. R Taylor, O Kennard *Acta Crystallogr., Sect. B* **39** 517 (1983)
25. R Taylor, O Kennard *Acta Crystallogr., Sect. B* **39** 133 (1983)
26. R Taylor, O Kennard *Acta Crystallogr., Sect A* **41** 85 (1985)
27. R Taylor, O Kennard, W Versichel *J. Am. Chem. Soc.* **106** 244 (1984)
28. J A R Sarma, G R Desiraju *Acc. Chem. Res.* **19** 222 (1986)
29. G R Desiraju *Acc. Chem. Res.* **29** 441 (1996)
30. F H Allen, O Kennard, R Taylor *Acc. Chem. Res.* **16** 146 (1983)
31. G R Desiraju *Crystal Engineering* (Amsterdam: Elsevier, 1989)
32. T Dahl *Acta Chem. Scand., Ser. B* **48** 95 (1994)
33. A Gavezzotti *Chem. Phys. Lett.* **161** 67 (1989)
34. W D S Motherwell *Acta Crystallogr., Sect. B* **53** 726 (1997)
35. P M Zorky *Zh. Fiz. Khim.* **68** 9667 (1994) [a]
36. A J C Wilson *Acta Crystallogr., Sect A* **44** 715 (1988)
37. A J C Wilson *Acta Crystallogr., Sect A* **46** 742 (1990)
38. A J C Wilson *Z. Kristallogr.* **197b** 85 (1991)
39. A J C Wilson *Acta Crystallogr., Sect A* **49** 795 (1993)
40. V K Belsky, P M Zorky *Acta Crystallogr., Sect A* **33** 1004 (1977)
41. N Padmaja, S Ramakumar, M A Viswamitra *Acta Crystallogr., Sect A* **46** 725 (1990)
42. V K Belsky, in *The XVII Congress of the International Union of Crystallography (Abstracts of Reports), Seattle, KY, 1996* N T 12, p. 1
43. A Mackay *Acta Crystallogr.* **22** 329 (1967)
44. A D Mighell, J R Rodgers *Acta Crystallogr., Sect A* **36** 321 (1980)
45. C P Brock, J D Dunitz *Chem. Mater.* **6** 1118 (1994)

46. J F Nicoud, R J Twieg, in *Non-Linear Optical Properties of Organic Molecules and Crystals* Vol. 1 (Eds D S Chemla, J Zyss) (New York: Academic Press, 1987) p. 253
47. G R Meredith *ACS Symp. Ser.* **233** 29 (1983)
48. J Jacgues, A Collet, S H Wilen *Enantiomers, Racemates and Resolutions* (New York: Wiley, 1981) p. 81
49. J K Whitsell, R E Davis, L L Saunders, R J Wilson, J P Feagins *J. Am. Chem. Soc.* **113** 3267 (1991)
50. C P Brock, W B Schwezer, J D Dunitz *J. Am. Chem. Soc.* **113** 9811 (1991)
51. V K Belsky, P M Zorky *Kristallografiya* **15** 704 (1970) [b]
52. N Yu Chernikova, V K Belsky, P M Zorky *Zh. Strukt. Khim.* **31** 661 (1991) [c]
53. S C Abrahams *Acta Crystallogr., Sect. B* **45** 228 (1989)
54. S C Abrahams *Ferroelectrics* **104** 35 (1990)
55. J M Igartua, M I Aroyo, J M Perez-Mato *Phys. Rev. B, Condens. Matter* **45** 12744 (1996)
56. W Nowacki *Helv. Chim. Acta* **34** 1957 (1951)
57. T V Timofeeva, A P Polischuk, M Yu Antipin, V I Kulishov, Yu T Struchkov, in *The VIIIth European Crystallography Meeting (Abstracts of Reports), Liege, 1983* p. 1627
58. A P Polishchuk, Candidate Thesis in Chemical Sciences, Institute of Organoelement Compounds, Academy of Sciences of the USSR, Moscow, 1983
59. B K Vainshtein, V M Fridkin, V L Indenbom *Sovremennaya Kristallografiya* (Modern Crystallography) (Moscow: Nauka, 1979) Vol. 2, p. 185
60. G R Desiraju *Angew. Chem., Int. Ed. Engl.* **34** 2311 (1995)
61. J Bernstein, R Davis, L Shimoni, Ning-Leh Chang *Angew. Chem., Int. Ed. Engl.* **34** 1555 (1995)
62. J J Wolff *Angew. Chem., Int. Ed. Engl.* **35** 2115 (1996)
63. C B Aakeroy *Acta Crystallogr., Sect. B* **53** 569 (1997)
64. F H Allen, in *Molecular Structure. Their Determination and Importance* (Eds A Domenicano, I Hargittai) (Oxford: Oxford University Press, 1992)
65. F H Allen, S Bellard, M D Brice, B A Cartwright, A Doubleday, A Higgs, T Hummelink, B G Hummelink-Peters, O Kennard, W D S Matherwell, J R Rodgers, D G Watson *Acta Crystallogr., Sect. B* **35** 2331 (1979)
66. P Murray-Rust, J Raftery *J. Mol. Graphics* **3** 50 (1985)
67. R Taylor, O Kennard *Acta Crystallogr., Sect. B* **39** 517 (1983)
68. R Taylor, O Kennard *Acta Crystallogr., Sect. A* **41** 85 (1985)
69. R Taylor, O Kennard *J. Chem. Inform. Comput. Sci.* **26** 28 (1986)
70. K N Trueblood, J Dunitz *Acta Crystallogr., Sect. B* **39** 120 (1983)
71. S C Nyburg, W Wong-Ng *Proc. R. Soc. London, A Math. Phys. Sci.* **367** 29 (1979)
72. F H Allen *Acta Crystallogr., Sect. B* **42** 515 (1986)
73. S L Price, A J Stone, J Lucas, R S Rowland, A E Thornley *J. Am. Chem. Soc.* **116** 4910 (1994)
74. Yu V Zefirov, P M Zorky *Zh. Strukt. Khim.* **15** 118 (1974) [c]
75. S C Nyburg, C H Faerman *Acta Crystallogr., Sect. B* **41** 274 (1985)
76. S S Batsanov *Zh. Neorg. Khim.* **36** 3015 (1991) [d]
77. Yu V Zefirov *Kristallografiya* **42** 122 (1997) [b]
78. R O Gould, A M Gray, R Taylor, M D Walkinshaw *J. Am. Chem. Soc.* **107** 5921 (1985)
79. A Gavezzotti *Acta Crystallogr., Sect. B* **46** 275 (1990)
80. V K Belsky *Zh. Strukt. Khim.* **15** 726 (1974) [c]
81. V K Belsky, O N Zorkaya, P M Zorky *Acta Crystallogr., Sect A* **51** 473 (1995)
82. A D Mighell, V L Himes, J R Rodgers *Acta Crystallogr., Sect. A* **39** 737 (1983)
83. W H Baur, D Kassner *Acta Crystallogr., Sect. B* **48** 356 (1992)
84. R P Scaringe, in *Electron Crystallography of Organic Molecules* Vol. 328 (Eds J R Fryer, D L Dorset) (Dordrecht, Netherlands: Kluwer Academic, 1991) p. 85
85. A Gavezzotti *J. Phys. Chem.* **94** 4319 (1990)
86. A Gavezzotti, G Fiilipini *J. Phys. Chem.* **98** 4831 (1994)
87. M J S Dewar, E G Zoebisch, E F Healy, J P Stewart *J. Am. Chem. Soc.* **107** 3902 (1985)
88. O Wallach *Liebigs Ann. Chem.* **286** 90 (1895)
89. J Jacgues, A Collet, S H Wilen *Enantiomers, Racemates and Resolutions* (New York: Wiley, 1981) p. 4, 23, 94
90. S F Mason *Molecular Optical Activity and the Chiral Discriminations* (Cambridge: Cambridge University Press, 1982) p. 171
91. W C McCrone, in *Physics and Chemistry of the Organic Solid State* Vol. II (Eds D Fox, M M Labes, A Weissberger) (New York: Interscience, 1965) p. 725
92. A Burger *Pharm. Int.* **3** 158 (1982)
93. A Ellern, J Bernstein, J Y Becker, S Zamir, L Shahal *Chem. Mater.* **6** 1378 (1994)
94. J D Dunitz, J Bernstein *Acc. Chem. Res.* **28** 193 (1995)
95. J Bernstein *J. Phys. D, Appl. Phys.* **26** B66 (1993)
96. K Sato *J. Phys. D, Appl. Phys.* **26** B77 (1993)
97. L Borka J K Haleblian *Acta Pharm. Yugosl.* **40** 71 (1990)
98. G Filippini, A Gavezzotti *Chem. Phys. Lett.* **231** 86 (1994)
99. A Gavezzotti, G Filippini *J. Chem. Soc., Perkin Trans. 2* 1399 (1995)
100. G Filippini, A Gavezzotti *Acta Crystallogr., Sect. B* **49** 868 (1993)
101. W Lewis, I C Paul, D Y Curtin *Acta Crystallogr., Sect. B* **36** 70 (1980)
102. P C Minshall, G M Sheldrick *Acta Crystallogr., Sect. B* **33** 160 (1977)
103. R Kivekas, T Laitalainen *Acta Chem. Scand., Ser. B* **41** 213 (1987)
104. Y Kato, Y Takaki, K Sakurai *Acta Crystallogr., Sect. B* **30** 2683 (1974)
105. F Pandarese, L Ungaretti, A Coda *Acta Crystallogr., Sect. B* **31** 2671 (1975)
106. F Hamzaoui, F Baert, G Wojcik *Acta Crystallogr., Sect. B* **52** 159 (1996)
107. M Bonamico, G Dessy, A Vaciago *Theor. Chim. Acta* **7** 367 (1967)
108. Y Le Page *J. Appl. Crystallogr.* **20** 264 (1987)
109. Y Le Page *J. Appl. Crystallogr.* **21** 983 (1988)
110. B M Craven *Acta Crystallogr., Sect. B* **35** 1123 (1979)
111. L N Kuleshova, V N Khrustalev, Yu T Struchkov, A T Soldatenkov, I A Bekro, Zh A Mamyrbekova, S A Soldatova *Kristallografiya* **41** 673 (1996) [b]
112. T V Timofeeva, V N Nesterov, M Yu Antipin, R D Clark, M Sanghadasa, B Cardelino, C Moore, D Frazier *J. Phys. Chem.* (1999) (in the press)
113. Y Delugeard *Cryst. Struct. Commun.* **4** 289 (1975)
114. A F Wells *Structural Inorganic Chemistry* (Oxford: Clarendon Press, 1962) p. 294
115. W C Hamilton, J A Ibers *Hydrogen Bonding in Solids* (New York: Benjamin, 1968)
116. L Leiserowitz *Acta Crystallogr., Sect. B* **32** 775 (1976)
117. L Leiserowitz, G M J Schmidt *J. Chem. Soc., A* 2372 (1969)
118. L Leiserowitz, M Tuval *Acta Crystallogr., Sect. B* **34** 1230 (1978)
119. L N Kuleshova, P M Zorky *Acta Crystallogr., Sect. B* **37** 1363 (1981)
120. L N Kuleshova, P M Zorky *Acta Crystallogr., Sect. B* **36** 2113 (1980)
121. Z Berkovitch-Yellin, L Leiserowitz *Acta Crystallogr., Sect. B* **40** 159 (1984)
122. M C Etter *Acc. Chem. Res.* **23** 120 (1990)

[a] — *Russ. J. Phys. Chem. (Engl. Transl.)*
[b] — *Russ. Crystallogr. (Engl. Transl.)*
[c] — *J. Struct. Chem. (Engl. Transl.)*
[d] — *Russ. J. Inorg. Chem. (Engl. Transl.)*