Graphical Enumeration of Hydrogen-Bonded Structures

BY L. N. KULESHOVA AND P. M. ZORKY

Chemistry Department, Moscow State University, Moscow 117234, USSR

(Received 2 November 1979; accepted 25 April 1980)

Abstract

A classification of hydrogen-bonded structures (H structures) which is based on the representation of these structures as graphs has been developed. Statistical data on the distribution of 776 homomolecular H structures are given in the form of graphs. The most common modes of H-structure formation and some anomalous H structures are described.

When molecules in crystals are joined by H bonds the aggregates so formed may be (1) dimers, trimers and some other finite clusters (islands), (2) infinite chains, (3) layers, or (4) frameworks. We designate such molecular aggregates as H aggregates and hydrogen-bonded structures as H structures. The H-aggregate type present in the structure has an obvious influence on the physicochemical properties of the crystals. This is why this aspect of H structures is of special interest.

A similar scheme has been given by Hamilton & Ibers (1968). They used a symbol \((N,M)\), where \(N = n/2\) is the number of H bonds per molecule and \(M = m\). In describing the system of intermolecular H bonds, they considered each molecule as a single point with lines representing hydrogen bonds extending from it. This meant that an H aggregate was represented as a graph (Harary, 1967; Harary & Palmer, 1973) (though Hamilton & Ibers did not use this term). It should be remembered that the term 'graph' is applied to a system of points, some of which (or all of which) are connected by lines. The number of lines extending from a point is called the degree of this point.

In the present work the classification of H structures based on the representation of H aggregates as graphs is continued. We have taken into consideration only homonuclear crystals built up from symmetrically related molecules. Such crystals are described as graphs with topologically equivalent points. Like other authors, we have used the parameters \(n\) and \(m\). The first is the degree of the point. We also call it the firmness of the H aggregate, because it shows how many H bonds must be broken to release a molecule. It is clear that for the structures considered here, \(n\) is always even. \(m\) is the number of points with which every point of the graph is directly joined. We also take into account the dimension \((k)\) of the rings present in the graph, considering only those rings which do not overlap.

In general, a graph symbol may be written as \(G_{nm}^k\). In particular cases we write \(I, C, L\) or \(F\). These letters correspond to islands, chains, layers and frameworks respectively. For instance, the hydrogen-bonded layer

\[
\begin{array}{c}
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\n\end{array}
\]

which is present in the meso-tartaric acid structure (Butsma & Schoone, 1967) is described by the symbol \(L_{48}^2(2,4)\). It should be noted from Table 1 that these symbols are not completely adequate. For example, there are two different graphs represented by the symbols \(C_{48}^2(2,3)\). Nevertheless, these symbols are sufficiently informative to classify H structures.

In order to determine the most common modes of H-structure formation and to investigate the variety of these modes (including rare and unique variants), we have statistically processed the data in the literature. Only those H bonds which were indicated by the authors were taken into account. Bifurcated H bonds were considered as two bonds.

As the source for the data we used Vols 1 and 2 of Structure of Organic Substances (Kitaigorodsky, Zorky & Belsky, 1980a,b). The data cover the period from 1929 to 1974 (inclusive).

The number of hydrogen bonds formed by different molecules for the 960 structures derived from this source is distributed as follows:

\[
\begin{array}{c}
n \quad 2 \quad 4 \quad 6 \quad 8 \quad 10 \quad 12 \quad 14 \quad 16 \quad 24 \\
\text{Number of structures} \quad 445 \quad 255 \quad 91 \quad 113 \quad 29 \quad 21 \quad 4 \quad 1 \quad 1
\end{array}
\]
Table 1. The distribution of H aggregates in the form of graphs

<table>
<thead>
<tr>
<th>n = 8 (28 H structures)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of H aggregate</td>
<td>Symbol</td>
</tr>
<tr>
<td></td>
<td>C (2)</td>
</tr>
<tr>
<td></td>
<td>L (2,3)</td>
</tr>
<tr>
<td></td>
<td>L (2,4)</td>
</tr>
<tr>
<td></td>
<td>F (2,6)</td>
</tr>
</tbody>
</table>

It is of interest to point out those structures with very large n values: cyclobutanoctol (Bock, 1968) where n = 16, and 1-kestose (Jeffrey & Park, 1972) where n = 24.

776 of these structures have been represented as graphs (all structures with n = 2 and 4 and some of those structures with n = 6 and 8). The results are summarized in Table 1.

It should be noted that the most common H aggregate is C \(2(0), the frequency of its occurrence being 28%. The second most common mode of H structure is I \(2(2) (19%), and the third is C \(2(2) (11%). The list of the more common H aggregates is completed by layers L \(4(4) (6%) and L \(2,6) (5%). These five modes of H bonding cover ~69% of H structures. All other H aggregates are encountered far less frequently (less than 2%). Examples of the most common H aggregates are: C \(2(0), 2,5-dimethylphenol (Neuman & Gillier-Pandraud, 1973); I \(2(2), propionic acid (Streiter, Templeton, Scheuerman & Sass, 1962); C \(2(2), dihydrouracil (Rohrer & Sundaralingam, 1970); L \(4(4), chloranilic acid (Andersen, 1967); L \(2,6), o-dihydroxybenzene (Brown, 1966).

We now consider some examples of the unusual modes of H aggregates. In some cases H bonding leads to the cyclic aggregates I \(3(3), I \(3(4) and I \(3(6) present in the structures of acetone oxime (Bierlein & Lingafelter, 1951), anti-\(\alpha\)-bromoacetophenone oxime (Wetherington & Moncrief, 1973), and (\(\pm\))-\(\beta\)-promedol alcohol (De Camp & Ahmed, 1972). The unique H-bonded layer L \(2,8 occurs in the structure of methyl 5-thio-\(\beta\)-D-ribopyranoside (Girling & Jeffrey, 1973). Sometimes a system of H bonds in a crystal is similar to a diamond structure, its graph being F \(2(6). Such a system occurs in 6-azauracil (Singh & Hodgson, 1974). The symbol F \(2(6) can also be applied
to the structure of hexagonal diamond (Ergun & Alexander, 1962). However, so far we have not detected any example of such a graph. It is interesting to consider structures of the type $F^2_{3}(2,3,4)$; these contain a diamond-like system of H bonds, but some lines of the graph are doubled by additional H bonds, and the six-membered rings are divided into three- and four-membered ones. Such a structure occurs, for example, in l-tyrosine (Mostad, Nissen & Romming, 1972).

The method of H-bonded structure representation described here does not take into account the fact that one of the two molecules connected by a H bond is a donor and that the other is an acceptor. This shortcoming can be eliminated by using so-called digraphs which consist of points and directed lines.

We plan to apply the digraph concept to the further systematization of H structures.

References


The Structure of 7-Chloro-1,3-dihydro-1-(N-methylacetamido)-5-phenyl-2H-1,4-benzodiazepin-2-one

By Pongchan Chananont and T. A. Hamor

Department of Chemistry, University of Birmingham, Birmingham B15 2TT, England

and I. L. Martin

MRC Neuropharmacology Unit, Department of Pharmacology, The Medical School, Birmingham B15 2TJ, England

(Received 4 March 1980; accepted 24 March 1980)

Abstract

$C_{16}H_{16}ClN_3O_2$ is monoclinic, space group $P2_1/c$, with $a = 4.775$ (5), $b = 11.74$ (1), $c = 30.54$ (2) Å, $\beta = 91.60$ (5)°, $Z = 4$. Final $R = 6.0$% for 1300 observed counter amplitudes [$I > 2.5\sigma(I)$]. E.s.d.'s average 0.008 Å for bond lengths and 0.6° for bond angles not involving H atoms. The N-methylacetamido group is planar and oriented at an angle of 75±5° to the C(6)–(11) phenyl residue of the 1,4-benzodiazepine system. The angle between the two phenyl rings is 61±6°. Comparison of molecular parameters with those of diazepam indicates that the geometries of the 5-phenyl-1,4-benzodiazepine residues of the two molecules differ only slightly.