Decoding Hydrogen-bond Patterns. The Case of Iminodiacetic Acid

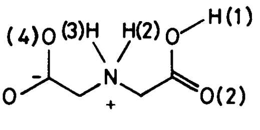
Joel Bernstein *
Department of Chemistry, Ben-Gurion University of the Negev, Beer Sheva 84120, Israel
Margaret C. Etter *† **and John C. MacDonald**
Department of Chemistry, University of Minnesota, Minneapolis, MN 55455

The hydrogen-bond patterns of three polymorphs of iminodiacetic acid are evaluated and compared using a recently developed form of analysis based on graph theory. The results of the analysis are compared with several previous analyses of these structures which are noted for their hydrogen-bond complexities. Graph-set analysis provides a way of abstracting hydrogen-bond patterns as recognizable motifs which are easily compared and contrasted among related structures. It is shown that there are five hydrogen-bond motifs present in different combinations in these polymorphs. Two structures are indistinguishable on the basis of these characteristic motifs alone, and higher-order networks involving multiple hydrogen-bond types are used to define their differences. The two metastable polymorphs are shown to lack a particular cyclic pattern involving hydrogen bonds between two carboxylate and two iminium ions found only in the stable form.

Since intermolecular hydrogen bonds are selective[1,2] and directional[3,4] they are useful tools for designing molecular crystals. To that end we have recently outlined a general scheme for classifying hydrogen-bond connectivity patterns using concepts derived from graph theory.[5] The purpose of this paper is to show how graph sets can be used effectively as a basis for the comparison of complex and intertwined hydrogen-bonded networks. Graph sets will be assigned to the hydrogen-bond patterns of three polymorphs of iminodiacetic acid, HO_2C-$CH_2NHCH_2CO_2H$ (IMDA), and will be used for comparing and contrasting these patterns in a chemically useful manner. The graph sets assigned here also reveal a possible driving force for the solid-state thermal polymorphic transformations that take place. Some graph sets are found to recur in all three of the polymorphs, suggesting that they might be useful for designing hydrogen-bond patterns of other iminodiacetic acids or of related molecules.

Hydrogen Bonding

Earlier Models for Comparison.—Iminodiacetic acid is at least trimorphic[6,7] and three crystal structures have been reported.[7,8] In all three structures the molecule is present in the zwitterionic form, rather than the neutral form. The various combinations of proton donors and acceptors coupled with the conformational freedom of the molecule lead to a large variety of hydrogen-bond possibilities, many of which are present in the three polymorphic structures. Comparison of the hydrogen-bond patterns of these structures is extremely difficult because of their complexity. The difficulties are compounded by a lack of appropriate basis for comparison. An initial survey of the three IMDA polymorphs (**1**)–(**3**) showed three patterns that appeared to be different, and were nearly indecipherable.[8] In that paper several attempts were made to deconvolute the structures, using different criteria for comparison.



IMDA zwitterion.

One criterion used was the presence of different intermolecular hydrogen-bonded ring structures. Boman *et al.*[7] initially cited two rings comprised of ten atoms and containing hydrogen bonds as the 'most striking feature' of the structure of the first published polymorph, (**1**). Following this lead Bernstein[8] also identified rings in the second two polymorphs. In (**2**), which has two molecules in the asymmetric unit, rings with eight and ten members were found. There were also two rings of 22 atoms each. In the third polymorph, (**3**), the smallest rings that were found contained 20 and 22 atoms. Using rings as a criterion for comparison, little chemical insight was gained about similarities between these three structures.

A second possible criterion for comparison of hydrogen-bond patterns is based on the type of hydrogen bond formed, as defined by which atoms participate in the bond. To facilitate these comparisons a consistent system of atom numbering was chosen for IMDA.[8] In the four different molecules present in these structures, each of the three hydrogens available for hydrogen bonding participates in a single hydrogen bond, but the role of the oxygen atoms varies among the polymorphs. The hydroxy oxygen does not act as an acceptor in any of the structures. Carbonyl and carboxylate oxygens are expected to be better acceptors than the hydroxy group, but only in (**3**) does the carbonyl oxygen participate in a hydrogen bond. One of the carboxylate oxygens participates in two hydrogen bonds in (**1**), and the other carboxylate oxygen has two hydrogen bonds in (**2**). The carbonyl oxygen and both carboxylate oxygens act as single acceptors in (**3**).

Finally, the symmetry relationships between hydrogen-bonded molecules were used as a basis for comparison. In (**2**) and (**3**) the presence of a 5.3 Å axis is compatible with a hydrogen bond created *via* translation.[1] Both (**1**) and (**3**) are centrosymmetric, but hydrogen bonding across the inversion centre is present only in the former. Glide symmetry is used to generate an O–H···O hydrogen bond in all three polymorphs. Glide-generated N–H···O hydrogen bonds are also present in both (**1**) and (**3**).

Even though these procedures involved systematic analyses for comparison of the structures, it was still difficult to disentangle the hydrogen-bond patterns sufficiently to make useful chemical conclusions about similarities and differences

---

† Alfred P. Sloan Foundation Fellow, 1989–1991.

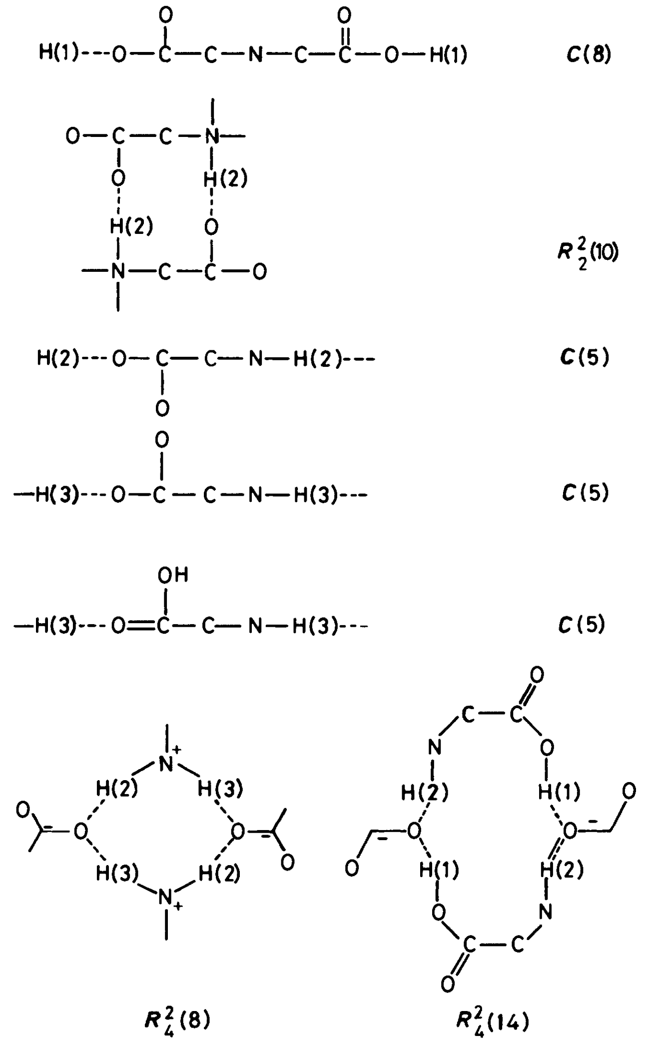


Figure 1. IMDA hydrogen-bond patterns and their graph sets. The top five patterns are the characteristic motifs that constitute first-order networks. The lower two patterns contain two different kinds of hydrogen bonds each, and are thus higher-order networks.

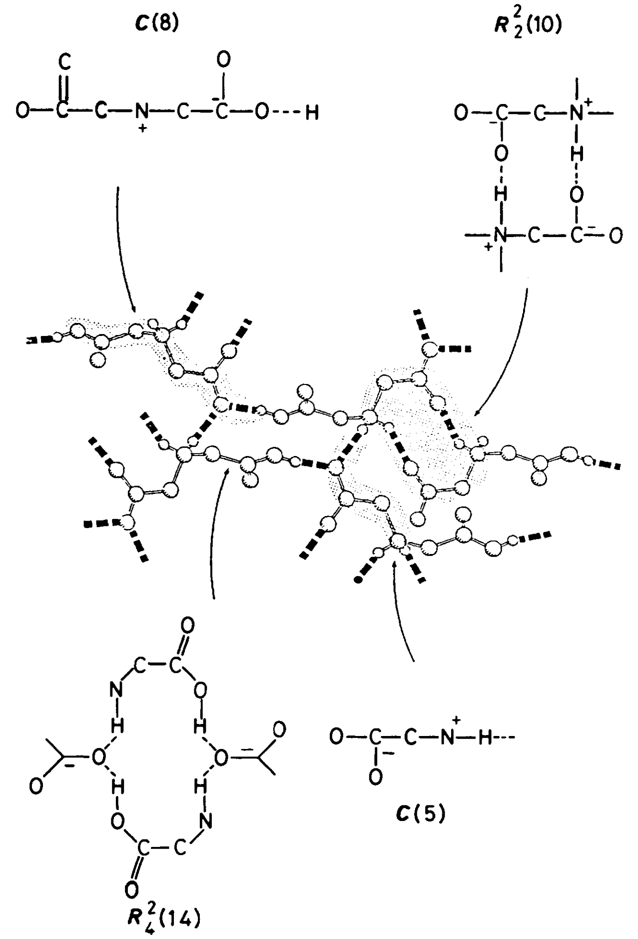


Figure 2. Map of the hydrogen-bond patterns in the crystal structure of (1). Hydrogen bonds are shown as dark dashed lines. There are three motifs in this structure, shown as independent shaded regions and illustrated schematically with their graph set assignments. The three motifs together constitute the first order hydrogen-bond network for (1), given as $N_1 = C(5)R_2^2(10)C(8)$.

between the three polymorphs. Graph-set analysis was developed expressly for this purpose, *i.e.* to decode hydrogen-bond patterns into sets of interpretable and comparable motifs. The IMDA polymorphs will be used to demonstrate the method and to give examples of how to assign and compare graph sets.

Application of Graph Sets.—In accordance with the procedures outlined previously[5] the different kinds of hydrogen bonds (based on the atoms used as donors and acceptors) are identified in the structures of (**1**)–(**3**). Each hydrogen-bond type is isolated from the others and a map is drawn showing all the repetitive occurrences of that single type of hydrogen bond. The set of molecules that is connected by this kind of hydrogen bond is the characteristic motif, which is assigned to one of four graph types determined by whether the motif is intramolecular (***S***), infinite (***C***), cyclic (***R***), or dimeric (***D***).* The number of proton donors and acceptors involved in the motif are assigned as sub- and superscripts, respectively, and the degree or size of the motif is included in the notation in parentheses. Motifs are always composed of only one hydrogen-bond type, as opposed to networks, which contain multiple types of hydrogen bonds. In structure (**1**), for instance, the hydroxy hydrogen is in a chain motif ***C*(8)**, one of the imino hydrogens is in a ring motif comprising two molecules $\mathbf{R_2^2(10)}$, and the second imino hydrogen is in a chain motif ***C*(5)** (Figures 1 and 2).

In (**2**), the presence of two molecules in the asymmetric unit could provide an additional complication. However, here the use of graph sets simplifies the interpretation of the hydrogen-bond patterns since symmetry relations are not a constraint. In Figure 3(*a*) it is seen that the hydroxy hydrogen participates in the same kind of *C*(8) chain regardless of which crystallographically independent molecules are considered. Two *C*(5) chains using the two independent imino hydrogen donors [(Figure 3(*b*)] are also formed. Finally, the *C*(5) chains of molecules are linked by the formation of an $R_2^2(10)$ ring, similar to that found in (**1**). It is possible to distinguish between the motifs of the two crystallographically independent molecules by additional notation in the graph set symbols, but this is not necessary here just to describe the hydrogen-bond patterns. For (**3**), H(1) is again involved in a *C*(8) chain; H(2) participates

---

* ***D*** refers to finite non-cyclic intermolecular hydrogen-bond patterns that involve only one type of hydrogen bond. Dimers are the most frequent ***D*** pattern, but trimers or higher sets of molecules held together by a single repeated hydrogen-bond type also fit the definition of ***D***. An example of a trimer ***D*** pattern would be hydroquinone hydrogen-bonded to two acetone molecules.

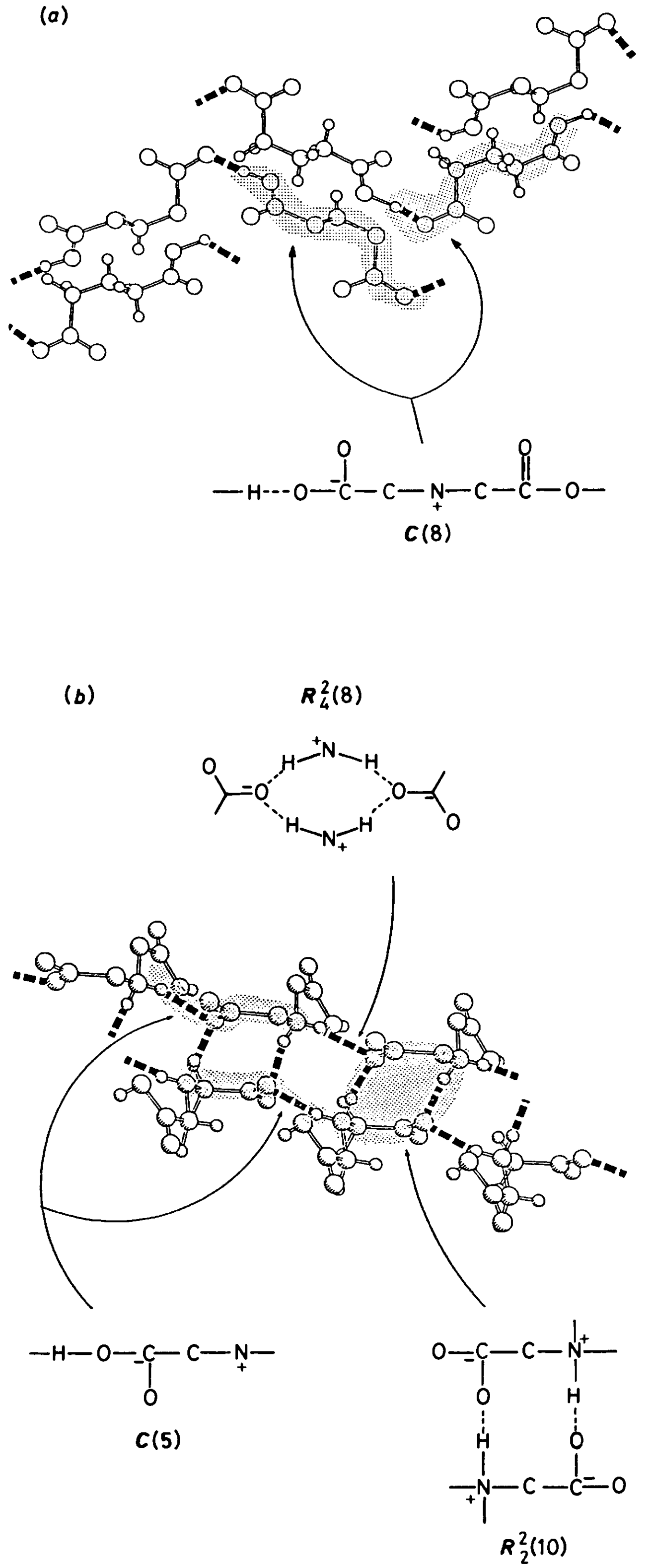


Figure 3. Maps of the hydrogen-bond patterns in the crystal structure of (2). The hydrogen-bond pattern is best understood from two separate figures. Hydrogen bonds are shown as dark dashed lines. The independent shaded regions indicate the motifs: (*a*) The *C*(8) motif for each of the two independent molecules; (*b*) The *C*(5) motif for each of the two independent molecules and the $R_2^2(10)$ which joins them. These three motifs comprise the first-order network specified as N_1 = $C(5)R_2^2(10)C(8)$. The second-order N_2 = $R_4^2(8)$ results from the use of H(2) and H(3) (see the text and Figure 1).

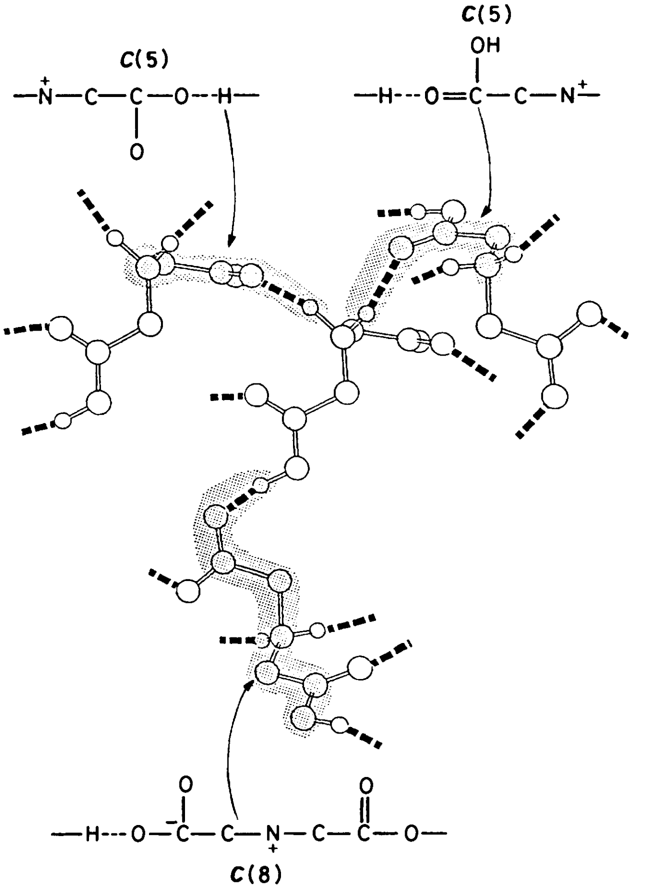


Figure 4. Map of the hydrogen-bond patterns in the crystal structure of (3). Hydrogen bonds are shown as dark dashed lines. Three motifs which define the first order hydrogen-bonding network are indicated by independent shaded regions and illustrated schematically by their graph-set assignments. The first order network is N_1 = *C*(5)*C*(5)*C*(8).

in a *C*(5) chain; and H(3) is part of an additional *C*(5) chain of molecules (Figure 4).

These are the minimum number of graph set assignments (one for each hydrogen-bond type) needed to characterize completely the three polymorphs. As a set they comprise the first-order networks, N_1, of the hydrogen-bond patterns.[5] These assignments are summarized in the Table. The 10-membered ring described by Boman *et al.*[7] for (1) is the $R_2^2(10)$ ring described above. It is now seen that there are several common features among the graph sets of the three forms. The first-order networks for (1) and (2) are identical, for example, forcing us to look to the second-order networks to distinguish between these two polymorphic forms. Polymorph (3) can be differentiated from (1) and (2) in terms of its first-order network only, which contains only relatively short *C*(5) and *C*(8) chains. The 20- and 22-membered rings which were used to characterize this form earlier[8] clearly belong to higher-order networks that are not as useful in the present context as the characteristic first-order networks given here.

In the course of this study we have found a simple protocol for identifying the graph sets of higher-order networks. As we noted above, the first-order networks are comprised of motifs containing only one type of hydrogen bond. What distinguishes a higher-order network from a motif is that networks of N_2 or higher describe patterns containing more than one type of hydrogen bond. In each specific case we searched for that pattern (of any graph set type) that had the smallest number of atoms in its repeat. This procedure is carried out in a

Table. Summary of first- and second-order graph-set assignments for IMDA polymorphs.

Polymorph (**1**)	$N_1 = C(5)R^2_2(10)C(8)$	$N_2 = R^2_2(14)$
Polymorph (**2**)	$N_1 = C(5)R^2_2(10)C(8)$	$N_2 = R^2_4(8)$
Polymorph (**3**)	$N_1 = C(5)C(5)C(8)$	

straightforward way: *e.g.* in searching for chains, we start with H(1) and proceed along a hydrogen bond and along the continuing atom chain until reaching H(2). Then a pathway is found that returns to H(1) in another molecule without including any other H-atoms. If this sequence propagates itself then it is accepted as a chain. The procedure is applied to all the pairwise combinations of the individual hydrogen atoms (in this particular case there are three pairs) in order to identify second-order chains.

Rings are readily identified by a similar procedure. A ring of second order or above must contain at least two types of hydrogen bonds. Hence we again start with H(1), proceed along its hydrogen bond and along its intramolecular atomic chain to H(2). The smallest ring possible would then include a second H(1), and a second H(2), returning in a cycle to the starting H(1). Again, proceeding through the three pairwise combinatorial possibilities, the higher-order rings can be identified. Rings with more than two types of hydrogen bonds in them are assigned as even higher networks, with ring size (or degree) determining their respective priorities.

Applying this procedure we have identified the second-order graph sets of (**1**) and (**2**). In (**1**), N_2 is the fourteen-membered ring $R^2_4(14)$, identified earlier by Bomans *et al.*,[7] while in (**2**) it is the eight-membered ring $R^2_4(8)$. Note that in keeping with the definition of higher-order graph sets, both contain two different kinds of hydrogen bonds. Interestingly, for (**2**) the ring in the second-order set is smaller than the ring found in the first-order network, resulting from combination of parts of two $C(5)$ chains and the sides of two $R^2_2(10)$ rings. The $R^2_4(8)$ pattern is unique to (**2**) which is the stable crystal form. This pattern is isographic with a common pattern found for co-crystals of nucleotide bases and carboxylic acids,[5] and it may be the feature that gives (**2**) its extra stability.

It has been shown[6] that (**2**) is the most stable of the three forms and that both (**1**) and (**3**) can be converted into (**2**) in the temperature range 453–457 K. The structural implications of the hydrogen-bonding changes taking place in these phase transformations can be readily understood in terms of graph-set analysis. For the transformation from (**1**) into (**2**), $R^2_2(14)$ is converted into $R^2_4(8)$, while that from (**3**) into (**2**) involves the loss of a $C(5)$ with the concurrent generation of $R^2_2(10)$. While it is not possible at this point to propose a detailed structural mechanism for these transformations, the graph-set descriptions provide an analytical description of the net change involved in the transformation.

Conclusions

We have demonstrated the use of graph sets for classifying the hydrogen-bonding schemes of the three rather complex polymorphs of iminodiacetic acid. The graph-set procedure provides a topological method for compariing the three structures and considerably simplifies the understanding of the similarities and differences among them. It also provides a basis for studying the mechanism of the polymorphic transformations that are known to take place in this system. We believe that further application of this graph set method to other hydrogen-bonded systems will usefully reveal the extents and limitations of this method for classifying and comparing crystal structures dominated by hydrogen bonds.

Acknowledgements

We wish to acknowledge financial support for this work from NIH (GM42148-01) and the Petroleum Research Foundation of the American Chemical Society.

References

1 L. Leiserowitz, *Acta Crystallogr., Sect. B*, 1976, **32**, 775.
2 L. Leiserowitz and G. M. J. Schmidt, *J. Chem. Soc. A*, 1969, 2372.
3 O. Kennard and R. Taylor, *Acc. Chem. Res.*, 1984, **17**, 320.
4 J. P. Glusker and P. Murray-Rust, *J. Am. Chem. Soc.*, 1984, **106**, 1018.
5 (*a*) M. C. Etter, J. Bernstein, and J. C. MacDonald, *Acta Crystallogr., Sect. B*, submitted August 1989; (*b*) F. Harary, *Graph Theory and Theoretical Physics*, Academic Press, New York, 1967.
6 Y. Tomita, Y. Ando, and K. Ueno, *Bull. Chem. Soc. Jpn.*, 1965, **38**, 138.
7 C.-E. Boman, H. Herbertsson, and A. Oskarsson, *Acta Crystallogr., Sect. B*, 1974, **30**, 378.
8 J. Bernstein, *Acta Crystallogr., Sect. B*, 1979, **35**, 360.

Paper 9/03557E
Received 21*st August* 1989
Accepted 31*st October* 1989