Fundamentals and Applications of Supramolecular Chemistry

Deepak Chopra
Department of Chemistry
IISER Bhopal

Week 05
Lecture 25

W5L25_Solid State Reactivity and Topology of H-bonded solids

So, hello everybody. So, now let us continue our discussion on crystal engineering. So, in the last lecture, we looked at solid-state reactivity in organic crystals. And the field of crystal engineering was pioneered by G. Smith, who actually explored solid-state reactivity in organic crystals. And it is important to note that reactions in the crystalline matrix provide extraordinary spatial control.

Ok, because we already have a very rigid matrix and the reacting double bonds are in the vicinity of each other. So, the structure provides the necessary spatial arrangement of the molecules that come into the required position for reacting with each other, and it also leads to a complete conversion into the final products.

Unlike reactions in solutions, which are diffusion-controlled and involve the reactant molecules coming close to each other, and where there is also an activation barrier involved in the process, in solids, the crystalline matrix is very rigid.

The molecules have functional groups that are already pre-organized; they are in the vicinity to react with each other, and then, in the presence of suitable stimuli, these double bonds can react with each other to form cyclobutane derivatives.

So, in this regard, they provide extraordinary spatial control, and, number two, these are topochemical in nature; that is, those reactions that occur with minimal movement of the reacting species.

So, this is the minimum movement at the atomic or molecular level when there is a topochemical reaction. So, keeping this in mind, Gerhard-Smith first performed the solid-state photodimerization reaction on trans-cinnamic acid.

Trans-cinnamic acid crystallizes initially in three different polymorphs; this means the molecule is the same, but the arrangement of the molecules in the crystal structure is different.

So, there are three different polymorphs: we designate them as alpha, beta, and gamma, and it has been observed that the alpha polymorph is present.

So, we represent this with an open circle. They have the arrangement of Ph and CO2H trans to each other, whereas here there is a CO2H group on the top and the Ph at the bottom, and now the alpha form undergoes a 2+2 cycloaddition reaction.

So, here you see that there is inversion symmetry in the product, and the resulting compound is an acid. On the other hand, when you have the corresponding beta form, the arrangement of the molecules is as follows: And the beta form undergoes a [2 + 2] dimerization reaction to give the beta dimer, which has mirror symmetry present in the product.

And in the third case, we have the double bonds where the separation between the double bonds is greater than 4.6 angstroms in the gamma form.

And therefore, no reaction takes place in the gamma form, and no products are obtained. So, this discusses a very, very specific product formation occurring depending on the arrangement of the molecules in the solid state.

So, if you have the CO2H groups that are trans to each other, then we have inversion symmetry; but if the CO2H groups are on the same side, then that is the cis orientation, and we have mirror symmetry that gives us the beta dimer. The prerequisite for this solid-state photodimerization reaction is that the reacting double bonds must be in the range of 3.5 to 4 angstroms.

And this is within the sum of the van der Waals radii plus a bit more, but still, they are positioned and pre-organized so that they can undergo solid-state dimerization to give the final product.

So, this kind of interesting chemistry occurs in the solid state, and now it is relevant to see whether we can actually bring about reactive double bonds in the solid state into a position where they can dimerize with each other.

So, the second strategy was to make structures such that the reactive double bonds are pre-organized to undergo dimerization reactions in solid. So, in this regard, what was considered is that we have this particular compound, and in the presence of resorcinol, it undergoes a solid-state dimerization at this particular double bond site to form the dimer.

And what is done essentially is we take a stoichiometric amount of A and a resource in B, we perform the process of mechanochemical grinding where we mix the reactants in a mortar and pestle, and we grind them for an hour and a half.

The resulting product, which is obtained, or the resulting solid-state phase, which is obtained, is then crystallized, and after the determination of the crystal structure, it was

observed that resorcinol forms hydrogen bonds via O-H...N hydrogen bonds with the pyridine nitrogen on both sides.

So, this kind of stereochemically locks the double bond. Now, these two have come into reactive proximity, and they react with each other to form the dimer. And these are hydrogen-bonded as well. You can see that there is a 2:2 ratio of the reactant materials to give you this product, and then the dimerization takes place in the solid state as well. So, this points to the fact that these kinds of interesting reactions can take place in solid materials.

This essentially provided the much-needed driving force for people to pursue more active research in the field of crystal engineering, and today there are a large number of reports on co-crystals and polymorphism, which we will discuss more in the lectures following next week.

Now, the next thing we would like to discuss is the topology of the hydrogen bond network once these hydrogen bonds have been formed in the solid state. The topology of the hydrogen bond network is very important, and we would like to utilize certain descriptors to describe the hydrogen bonding pattern obtained in the solid state; the language that we use in this regard is referred to as graph set analysis.

The graph-set analysis is the language used to describe the way in which hydrogen-bonded networks are formed in crystals. And this language can help in the identification of related families of crystal geometries, topologies, and connectivity in chemically different structures.

So, once you have the hydrogen-bonded network, you can now identify these hydrogen-bonded networks in related families of different crystal geometries, topologies, and connectivity in different structures. So, we would like to discuss this graph set analysis, which was developed by Margaret C. Etter at the University of Minnesota, followed by Joel Bernstein and Raymond Davis.

So, to start with, all hydrogen-bonded crystal structures can be reduced to combinations of four simple, designated patterns. Number 1 is a chain that represents a C.

Number 2 is rings, which we represent as R; number 3 is an intramolecular hydrogen bond, which we represent as S, representing self; and number 4 is discrete D for other finite patterns.

To these fundamental descriptors, a subscript is added, denoting the number of hydrogen bond donors in the pattern, symbolically represented as d, and a superscript denoting the number of acceptors, which is denoted as a.

The total number of atoms, including hydrogen, in the pattern is termed the degree of the pattern and is symbolically written as n. This is also given in brackets after the pattern descriptor. So, the total graph set descriptor is symbolically written as G_a_d(n).

These are the fundamental rules on which the graph set analysis is based. The structure containing only one type of hydrogen bond is called a motif and can have one or more graph-set descriptors associated with it.

And when you have more complex crystal structures where more than one motif is present in the structure, you have to assign a graph set descriptor for every motif separately.

So, the graph set descriptors have to be assigned independently; that is, individually for each motif, as if the others were not present, and this is called the first-level graph set, written as N1. So, now let us look at these different graph set descriptors with appropriate examples.

For example, if we are to first look at the chain that is being formed, we have already discussed its formation. For example, we have a benzamide molecule. It forms a chain, and between one molecule and another, there is a number of atoms that separate the chain.

So, between this hydrogen bond and this hydrogen bond, the number of atoms that separate the chain is the longest possible chain, which contains the maximum number of atoms; in this case, it is four.

So, this is referred to as a C4 graph set descriptor, with C representing the chain and 4 representing the number of atoms that are repeated in the pattern.

So, when the pattern is being formed, this is the number of atoms in the pattern that are repeating, which is C4. The next one is to look at the classical carboxylic acid dimer, which is represented by the graph set descriptor R; "a" represents the number of acceptor atoms.

So, there are 2 acceptor atoms, "d" represents the number of donor atoms, there are 2 donor atoms, and the number of atoms in the motif or the graph set descriptor is 1, 2, 3, 4, 5, 6, 7, 8, so R22(8). So, the cyclic motif, the dimeric centrosymmetric motif, is represented by $R2^2(8)$. We can consider the case of the intramolecular hydrogen bond.

For example, we can take ortho-hydroxybenzoic acid, where this intramolecular hydrogen bond is present. So, it is a self-cyclic motif, it is an intramolecular hydrogen-bonded system, and the number of atoms present in this descriptor is 1, 2, 3, 4,

5, 6; normally, when the number of donors and acceptors is equal to 1, that is not mentioned.

So, when there is only one type of hydrogen bond and there is only one donor and one acceptor, you do not have to mention the number 1; you can omit that, and that is a case of a discrete motif.

We can now take a case where we have triphenyl phosphine oxide, hydrogen-bonded with phenol via an O-H...O hydrogen bond, and this forms a discrete motif represented as D, and the number of atoms in a discrete motif is 1, 2, 3, or 4. Therefore, the graph set descriptor for such a kind of motif is D4.

So, in an actual crystal structure, you can have more than one type of motif, and as mentioned, you will have to consider every motif using a different graph set descriptor, and the combination, the product of all these descriptors, will represent the first-level graph set.

And in this regard, Etter's rules have been proposed, which are used to predict and forecast solid-state arrangements. And what does the Etter rule say? Number 1: all good proton donors and acceptors are used in hydrogen bonding.

Hydrogen bonds are preferred over intermolecular forces. Thus, these two are the fundamental rules on which Etter's rules, the foundational principles for predicting solid-state arrangements, are based.

Obviously, these are just a set of rules; they can be applied, and sometimes they work, but it is not necessary for the Etter rules to work in all cases. However, in cases where you have strong donor and acceptor groups, such as acids, amides, ureas, anilines, imides, and phenols.

So, the first rule works very well, and the justification for the second rule is similar to that of the chelate effect. Just like the chelate effect stabilizes the ring-like structure, in this case, the cyclic structure, wherein a proton is essentially chelated between the two donor atoms, provides extra stability to this hydrogen bond; hence, it is preferred in comparison to the intermolecular hydrogen bonds.

And let us take up some more examples of the applications of graph-set analyses. For example, we can consider N-H interacting with the motif we looked at. So, this is a ring motif that has 2 acceptors, 1 donor, and the number of atoms in the ring is 1, 2, 3, or 4.

So, when you have an N-H…O hydrogen bond involving the nitro group, this kind of motif is formed, and this is also a representation of a three-center hydrogen bond. We can have amino functionality that interacts with the nitro group, and here it will be

R22123456.

So, this is the graph set descriptor, and this is a case of a hydrogen bond and an N-H...O hydrogen bond. Similarly, we can look at diaryl ureas.

So, in this case, we have two sets of hydrogen bonds, and we can first describe the cyclic one, which is the R1 that has one acceptor and two donors: 1, 2, 3, 4, 5, and 6. So, this is one: the acceptor is 1, and the donors are 2. Yes, we have another one such that the chain repeats itself; that is, this particular C4 unit. So, this is also present. So, both of these graph-set descriptors are present in diarylureas.

So, this is interesting in the sense that we now have two donors and one acceptor. We can also look at other types of motifs. For example, we can look at N-H; we looked at the structure. So here we have got 1, 2, 3, 4, 5.

So we can have this N-H...O 1, 2, 3 and 4, 5, 6, 7, 8. This is a 1,8-membered ring, and then we can also have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12. So this is the first one: it is the primary graph set analysis.

Here we have 1, 2, 3, 4, 5, 6, 7, 8; we have 1 donor and 1 acceptor, whereas we can also consider another graph set analysis, where we have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12, and there are 2 donors and 2 acceptors. This is the secondary graph set analysis.

So instead of talking about two independent primary sets, we can talk about one independent primary set and a much larger graph set descriptor, which essentially covers the larger-sized motifs, represented as R 2 2 (12).

So, in this way, graph set analysis is very useful for understanding the arrangements of different hydrogen bonds in various crystal structures.

Particularly in the case of polymorphs, this is very useful for differentiating between the different hydrogen bonding arrangements that govern the formation of polymorphs.

So, with this, we have come to the end of this particular week's lectures, and next week, we will go over the other set of topics related to polymorphism, co-crystallization, and explore other relevant phenomena in the solid state.

Thank you.