The Structure of Structure
Some Problems in Solid State Chemistry

By A. L. Mackay

Department of Crystallography, Birkbeck College, University of London, Malet Street, London, W.C.1

Dedicated to Professor Walter Feitknecht

Summary

The problem of selecting useful parameters for the description of the complexity of structures at a level higher than that of the arrangement of atoms in the unit cell is discussed. Parameters equivalent to those of pressure, temperature, entropy used in describing disordered gases are sought and concepts such as hierarchic structures, information content, dislocation density and self-assembly are briefly surveyed.

The concept of geometrical structure and the techniques for determining it have been fundamental in the development of chemistry. As regards the structure of molecules—the arrangement of single atoms to form a larger unit—we might take Kekulé's conception of the benzene ring and the tetrahedrally bonded carbon atom as marking the crucial step. (The previous attempts, from Plato and Democritos onwards, were not so successful in predicting and explaining the properties of substances.) The actual X-ray demonstrations, by Kathleen Lonsdale and later by J. Monteath Robertson, of the reality of the picture came somewhat as an anticlimax, so fruitful was the original concept.

Even today, for molecular structures, the interest of chemists is primarily in the configuration of the molecules rather than in how molecules themselves combine to give more complex aggregates. We wish now to scan the whole scale of complexity, both for molecular and for ionic or metallic structures. Professor Walter Feitknecht is one of the few who took seriously the chemical implications, first of crystal structure (the arrangement of atoms in the unit cell) and secondly of higher kinds of structuring involving texture, dislocation patterns, two-phase materials, etc. He has regarded X-ray structure analysis, and later electron microscopy, not simply as methods of confirming the configurations of complex molecules, but as key techniques for developing his speciality—topochemistry—how the chemistry of the solid state operates at all scales from microscopic to macroscopic, atomic to geological. I wish, therefore, to make some speculations on structure, using the word "structure" over a wider range of phenomena than the arrangement of atoms in the unit cell of a perfect crystal.

In thinking about structures I consider that the concept of *hierarchy* is paramount, not just as a device for interpreting or apprehending complex structures, but as a physical reality. Structure itself has structure and hierarchy is perhaps the most important of its non-trivial characteristics and occurs because it increases the configurational entropy. In this respect it is a generalisation of symmetry which does the same by introducing degeneracy, both by the occurence of point elements and of translational elements of symmetry. Another very general property of structure is *dimensionality* but we will here confine ourselves to the three dimensions of length and to the topological laws appropriate to this number. There are perhaps three approaches which give useful *a priori* information about systems. These are *thermodynamics, symmetry* and *dimensional analysis. Information theory* might be considered as a kind of restricted thermodynamics. There are also limitations, derived from experiment or "thought-experiments", such as those imposed by the *uncertainty principle,* the finite *velocities of light and sound,* the *Brownian movement,* etc.

Euclidean geometry, including topology, imposes very stringent conditions on structures. Starting from the simple assumption of spherical, non-overlapping atoms, a great deal of structural chemistry can be deduced. A. F. Wells, for example, has set these out clearly. Euler's rule for simply connected polyhedra (faces + vertices = edges + 2)—more generally, $n_0 - n_1 + n_2 - n_3 \ldots = 1$ (where n_0 is the number of features with dimensionality 0 — vertices; n_1 features with dimensionality 1 — edges; n_2 — faces; n_3 — volumes) is a vital step in relating coordination number to the way in which coordination polyhedra are linked. (As a peripheral question: is it really only coincidence that the Phase Rule takes the same form?)

Under symmetry we include such considerations as Neumann's *Principle* (really an extension of the case of Buridan's Ass) and the symmetry of the bonding to single atoms occasioned by the solutions to the wave equation. It is remarkable that the *a priori* geometrical assumptions that the Platonic solids must figure in the structure of the universe should have failed in the Ptolemaic system of the heavens carried on crystal spheres, and in Kepler's Mysterium Cosmographicum, but should have at last occurred as the spherical harmonics" which are the solutions of Schroedinger's wave equation.

Maxwell's kinetic theory of gases achieved very considerable success in connecting microscopic parameters with macroscopic quantities such as temperature, pressure, viscosity and thermal conductivity. However, gases have only molecular structure and beyond that

Table 1. Hierarchy of levels of organisation of matter

Typical dimension	*Name of assembly*	*Interaction with radiation* Diffraction characteristics	*Physical properties* originating at this level	*Symmetry characteristics*
10^{-13} cm	Fundamental particles	Mass/energy transition reactions	Mass	Non-geometrical (parity, strangeness, etc.)
10^{-12} cm	Nuclei	*n*-scattering	Nuclear reactions of the elements	
1 Å $\equiv 10^{-8}$ cm	Atoms, with electronic structure	X-ray scattering. Fluorescence, etc. *e*-scattering	*p*-magnetism atomic dia-magnetism density, temperature	Atomic orbitals, spins aligned in magnetic field
	Valency, chemical bonding (covalent, ionic, van der Waals, metallic, repulsive)			
5 Å $\equiv 5 \times 10^{-8}$ cm	Immediate neighbours, coordination shell or molecule. Complex ions. Molecules	Peaks on scattering curves. Molecular transform	Chemical properties molecular diamagnetic dipole moments	Point groups, cheirality
	Crystal chemistry—systems and combinations of chemical bonding			
10 Å $\equiv 10^{-7}$ cm	Unit of pattern in crystal Crystal structure	Unit cell transform F_{hkl}	Chemical composition	Orientation
1000 Å $\equiv 10^{-5}$ cm	Crystallite	Shape transform	*f*-magnetic domains anti-*f*-magnetics	Space groups; anisotropy of properties
1 m.m. $\equiv 10^{-1}$ cm	Large crystal	∞ grating (extinction)	Optical properties conduction	
1 cm	Texture	Disordering dislocations	Mechanical properties	Texture groups
1 cm	Two- (or more) phase textures	Optical microscopy. Bulk properties. Contrast by R.I. and by absorption	Complex mechanical properties	Twinning, epitaxy, topotaxy
10 cm	Suspensions, lumps, etc.			
$1–10^5$ cm	Engineering-scale structures	Radiowaves		

complete randomness. In gases there are only three pertinent levels of spatial organisation: atoms; molecules; gas. "Gas" is a word coined by VAN HELMONT to resemble the word "chaos". The connection of microscopic parameters with macroscopic is not so satisfactory in the cases of liquids and solids which have varying amounts of order or structure. My main thesis now is that we must try to set up more macroscopic descriptions which are understood in terms of structures at the atomic level. As an example, until relatively recently the mechanical strength of crystals did not fit with the known interatomic forces. The introduction of a parameter "dislocation density" (based on experimental studies) has restored the correspondence by enabling us to speak concisely about the structure at a level above that of the unit cell arrangement derived by X-ray crystallography.

Crystallography is basically the exploitation of the knowledge that a crystal is made up of identical units arranged in an ordered way to obtain information about the individual units which may be individually too small for direct observation. Crystallography works primarily at the levels (marked in Table 1) of interatomic bonding. For liquids and meso-phases we can assume neither complete order nor complete disorder and the difficulties of connecting macroscopic diffraction data with microscopic parameters is correspondingly great. There are too many microscopic parameters (atomic positions) to list them separately but yet they are not independent of each other as in a gas. The problem is to find suitable parameters for the structure—one widely used is radial density distribution but this is not enough.

The method of carrying out a structure determination, where finding the unit cell and space group is an essential and easy preliminary, tends to give the impression that crystals consist of atoms placed artfully in a three-dimensional frame of symmetry elements. This is the exact opposite of the true situation. The unit cell and the symmetry elements are implicit in the individual molecules and their interactions, just as the chicken is implicit in the fresh egg. But who can say into what creature a single cell will develop? Molecules or other units assemble themselves into crystals in accordance with general principles such as the minimisation of free energy. Structures assemble themselves, generally with the aid of a catalyst or substrate to take away the energy released and to present the components to each other in favourable orientations. *Figure 1* shows jigsaw puzzle pieces of 17 types. When many examples of any one of these are taken and fitted together so that no spaces are left (this could in principle be done by shaking, if the joined pieces stayed fixed) a pattern representing one of

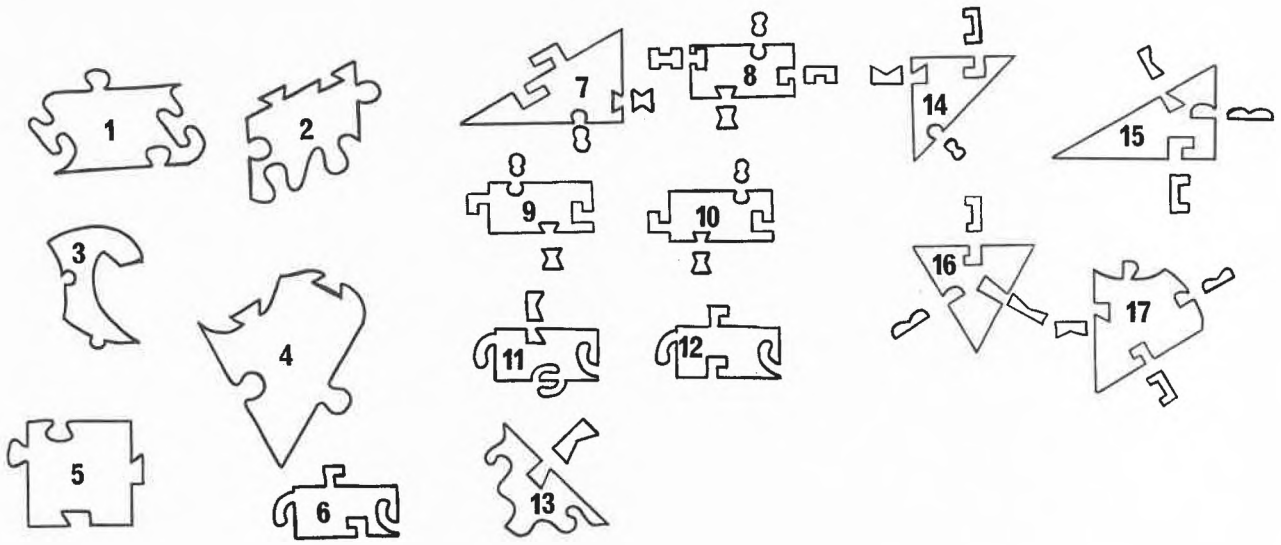

Fig. 1. Specimes of pieces of jigsaw puzzles which automatically assemble themselves to give patterns of a specified symmetry

1. $p1$	4. $p6$	7. cmm	10. pm	13. $p4g$	16. $p3m1$
2. $p2$	5. $p4$	8. pmm	11. pmg	14. $p4m$	17. $p31m$
3. $p3$	6. pgg	9. cm	12. pg	15. $p6m$	

When made of plywood in the conventional style small connecting pieces are required to give bonding across mirror planes

the 17 plane groups is obtained automatically. This system was suggested by N.M. Bashkirov (1959). (In the jigsaw puzzle style there is some difficulty in making an interlocking joint with a mirror plane of symmetry and extra joining piece is necessary.) However, although this is the simplest possible example, it is not immediately obvious which piece gives which symmetry group. It is clear that planes are the most restrictive symmetry elements and glides the least restrictive. Such considerations lead me to formulate the general question:

Is it possible, even in principle, to know enough of the properties of a fundamental unit in a hierarchic structure be able to predict how a complex assembly of these units will behave? If the answer is "no" (and such is my own tentative answer), then it follows that chemistry is not entirely reducible to physics, and that biology is not reducible to chemistry. It implies too that a subject must be studied at each level of complexity.

We may look at Bernal's definition of life: "Life is a partial, continuous, progressive, multiform and conditionally interactive, self-realisation of the potentialities of atomic electron states" (1967). All possibilities for life are written into the atomic energy levels if only we could read them. But it is clear that we cannot, and even if we could, it would often happen that the branching point between two courses of development would be finely balanced and that the energies would have to be known very accurately. The situation is clearly beyond our present, and even our imaginable capabilities. At present we still have great difficulty in predicting which metals will be cubic and which hexagonal. Topochemistry is not reducible to solution chemistry.

Even in a most restricted field like the mathematical properties of the integers, rich and unexpected properties continue to emerge as we pass from one integer to the next. Indeed, it has been shown, parodying Euclid, that there is no greatest interesting number N, for if so, then $N + 1$ would be of interest as the least integer of no interest, thus contradicting the assumption. Similarly with geometrical and atomic structures—ever more unexpected properties emerge with complexity. Table 1 shows how different properties are associated with different levels of organisation.

Crystals are perhaps unique in that they may be of enormous extent at the same level of complexity. A crystal may contain 10^{25} identical unit cells. Structures usually evince new properties with increasing size. For example, in most biological structures a clear hierarchy is to be seen. We may have, atoms: amino acid: α-helix: twisted helices: fibril: fibre: muscle. Analogously, in a Nylon hawser about 10 cm in circumference, five orders of fibres, from extruded filament to complete rope, can be seen. Each twisted bundle is a helix of hand opposite to that of the order below.

As an independent scientific subject, crystallography is at a critical stage of development because the determination of crystal structures is likely to become a standard technology. Interest is thus tending to move up the chain of complexity towards dealing with some of the consequences of crystal structure. Much work centres on the defects of crystals, dislocation structure, etc. and its importance for chemistry and electronics.

However, an important class of problem in crystallography, which is overdue for solution, is that of organis-

ing the results of structure analysis and formulating general principles and concepts. Here the guiding principles of local coordination into polyhedra and the linking of molecules or polyhedra into networks were early formulated by LINUS PAULING but, in spite of ligand field theory, there has been no sweeping advance. For molecular crystals A.I. KITAIGORODSKII, aided by the computer, has confirmed that, given plausible intermolecular forces, the actual packings correspond to potential minima. Nevertheless it is not yet possible with any confidence to predict a complex structure. We see the first signs that some progress in this direction may be possible. Firstly, R.G.J. STRENS (1967) has been quantifying concepts of symmetry and using it to calculate configurational entropy. Secondly, computer science has made progress in handling multi-dimensional data. In the 10,000 or so known structures we have a great reserve of data which should be used in discussing any individual case. The problem is how to organise it. Here again there are possibilities of semi-automatic numerical taxonomy which can be realised on a computer. If crystal structures were classified automatically, as if they were beetles or fossils would the emerging system look like our present one, or would it afford new insights? In short, the computer simulation, classification and storage of structures and associated thermodynamic information seems a promising field.

Table 2. Number of distinguishable sequences of N digits

Direction and starting point not given

N	Binary digits	Ternary digits	Sexagenary digits
1	1	1	1
2	1	1	3
3	1	2	6
4	2	4	26
5	3	8	83
6	5	18	402
7	8	39	
8	14	94	9212
9	21	222	
10	39	572	
11	62		
12	112		
13	189		
14	352		
15	607		
16	1144		
17	2055		
18	3885		
19	7154 ($=2^{12\cdot8}$)		

Ternary and sexagenary cases by courtesy of H. TAKEDA.

We have remarked above, with reference to the jigsaws of Figure 1, that in a sense simple structures contain their own assembly instructions in the atomic energy levels. These are not capable of arbitrary variation or evolution, except by the addition of different atoms or other pre-assembled units. There are, however, structures which can contain almost arbitrary information because they have metastable states of depth large compared with KT. The differences between individual potential wells and the absolute energy minimum are small compared with the depths of the individual metastable wells. W.A. ASHBY's "homeostat" machine is a model illustrating this principle. SiC is an example. Layers can succeed each other in either cubic or hexagonal sequence and growth conditions, screw dislocations, etc. can give rise to an almost arbitrary sequence such as "*ccchhcch*". This sequence repeats after a number of layers N by the screw mechanism. "*c*" and "*h*" may be considered as binary digits (like 0 and 1). We can ask then how many different sequences of N layers are possible (assuming the two directions of reading are indistinguishable and that reading can begin at any layer. We have calculated these up to $N = 19$ (the International Tables list values up to $N = 12$) and H. TAKEDA has done this also for the ternary and sexagenary digits which can be represented in the layer stackings of certain micas. These are shown in Table 2. A 19 layer polytype of SiC could thus store 12 binary digits (bits of information and this could be reproduced by the growth spiral mechanism almost as in DNA. Several typically hierarchic structures are summarised together in table 3.

There is, however, a bigger class of materials which contain their own assembly instructions written in another code, a meta-language, which is capable of evolution in the Darwinian sense. These are living materials, for example the simplest virus structure which consists of an icosahedral protein cage structure, just sufficiently big to contain a length of DNA or (RNA) which codes for the formation of the enzymes necessary to catalyse the production of the coat protein. The inclusion of assembly instructions in a meta-language

Table 3. Hierarchic systems analogous to those of "aperiodic" crystals

SiC Polytypes	Genetic Code	Written English	Written Chinese
SiC sheet	4 nucleotides	binary code on punched tape	12 strokes
Stacking into unit of N layers	20 amino amino acids 10,000 enzymes	26 letters 50,000 words	214 radicals (+ other graphic elements) 8,000 characters 100,000 compounds
Repetition of unit cell	Indefinitely large number of structures	Indefinitely large number of sentences	Indefinitely great number of sentences

capable of evolution might be taken as an alternative definition of living structures.

It is logically necessary that the assembly instructions should be written in a meta-language since, as the whole system of structure plus instructions has to be written in atoms, the plan cannot be a model of the structure on a reduced scale—like a sauce bottle with a picture of a sauce bottle on the label, or as the spermatozoon regarded as a homunculus—this involves an infinite regress.

As SCHROEDINGER pointed out in 1948, there is a serious problem in miniaturising the instructions, as these must be stable against the Brownian motion, and must be capable of being reproduced and read with only very small possibilities of error. The beautiful mechanism by which this is achieved in the coding of protein sequences by DNA is now familiar, but we should recall SCHROEDINGER's prescient statement that an "aperiodic crystal" is involved: "In physics we have dealt hitherto with periodic crystals. To the humble physicist's mind these are very interesting and complicated objects; the constitute one of the most fascinating and complex material structures by which inanimate nature puzzles his wits. Yet compared with the aperiodic crystal, they are rather plain and dull ..." Crystallography is now moving on to more general, aperiodic crystals, both organic and inorganic. BERNAL has made a preliminary statement of the principles of "generalised crystallography" (1967). Professor FEITKNECHT has specialised in experimentally elucidating the chemical properties of real crystals which are more than the naive structure analyst might expect. Table 3 is meant to suggest that writing and laguage also exhibit some of the features of an "aperiodic crystal".

We see that in DNA relatively heavy units, stable against KT, are used and that many atoms (a nucleotide plus a sugar and a phosphate group) are needed for each "bit" of stored information. Since in a completely arbitrary structure each atom requires three positional parameters and each group or molecule six, it is clear that a structure cannot describe itself unless it is largely degenerate (has large regions which are described in bulk by a general rule). This implies a hierarchic structure of standard components arranged so as to reduce their information content.

BERNAL's restatement of CHARLIER's Principle of hierarchic organisation: "The probability of the formation of a highly complex structure from its elements is increased, or the number of possible ways of doing it is diminished, if the structure in question can be broken down into a finite series of successively smaller substructures" is here apposite and can easily be translated into thermodynamic (or information theory) quantities (by BOLTZMANN's Law, $S = k \log W$). In either case it corresponds to the calculation of the configurational entropy.

Here we may conjecture as to whether there is a complexity limit for crystal structures. When the number of parameters in the unit cell rises does there come a time at which only a hierarchic structure is stable? There are, I believe, no very large inorganic structures with atoms of 10 different elements, each lying in a distinct position. The principal inorganic structure which have large cells are *polytypes*, which are modulations of a smaller cell, or hierarchic structures (such as silicates with strongly expressed substructural units). In organic structures almost all are formed by the crystallisation of pre-formed molecular units and thus already include much information. Can such considerations of ordered structure be quantified?

Directly from this consideration of organic structures we can formulate an important question for solid state chemistry. How densely can information be stored in a solid integrated circuit, considering the mentioned above of stability against KT? There are two cases. The orientation of magnetic domains, used in computer cores and the chemical differences used in micro-circuits. This is the reverse problem to the study of topochemical reactions. Namely how far can reaction (and diffusion) be prevented when they are not required?

Acknowledgments

In producing this essay for the intellectual bouquet presented to Professor WALTER FEITKNECHT I am happy to acknowledge my obvious debt to Professor J.D. BERNAL who would wish to be associated with the congratulations, if not with the opinions expressed. I am also indebted to Dr. KHUDU MAMEDOV and Dr. HIROSHI TAKEDA for stimulating discussions on the wider aspects of crystal science.

References

ASHBY W.R., *An Introduction to Cybernetics*, 1956, p. 83.
BASHKIREV N.M., *Fedorov Session on Crystallography* (abstracts), Leningrad 1959.
BERNAL J.D., *The Origin of Life*, London 1967, p. xv and Appendix 3.
SCHROEDINGER E., *What is life?* Cambridge 1948, p. 3.
STRENS R.G.J., *Miner. Mag. 36* (1967) 565–77.
WELLS A.F., *Structural Inorganic Chemistry*, 3rd edition, Oxford 1962.