Molecular Crystals and Liquid Crystals Science and Technology. Section A. Molecular Crystals and Liquid Crystals

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gmcl19

Structural Aspects of Intermolecular Interactions

Jenny P. Glusker [a]

[a] The Institute for Cancer Research, The Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA, 19111, U.S.A.

Version of record first published: 27 Oct 2006.

To cite this article: Jenny P. Glusker (1992): Structural Aspects of Intermolecular Interactions, Molecular Crystals and Liquid Crystals Science and Technology. Section A. Molecular Crystals and Liquid Crystals, 211:1, 75-88

To link to this article: http://dx.doi.org/10.1080/10587259208025807

Mol. Cryst. Liq. Cryst. 1992, Vol. 211, pp. 75–88
Reprints available directly from the publisher
Photocopying permitted by license only

Printed in the United States of America

STRUCTURAL ASPECTS OF INTERMOLECULAR INTERACTIONS

JENNY P. GLUSKER
The Institute for Cancer Research, The Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, U.S.A.

(Received September 28, 1991)

Abstract The surroundings of functional groups in molecules within a crystal can be analyzed from data in crystallographic databases. A method is described here for preparing contoured maps of the probabilities of interactions of functional groups in various directions. The results of such analyses can provide probability plots of orientational preferences of binding which can be used for model building involving the interaction of two different molecular species. Examples chosen for description here include hydrogen bonding to oxygen-containing functional groups, the surrounding of C-F bonds and the manner by which carboxyl groups bind metal ions.

Keywords: intermolecular interactions, hydrogen bonding, metal ions, carboxylates, metal binding, fluorine

Crystals contain not only information on the stereochemistry of the molecules that they are composed of, but, in addition, on intermolecular interactions between these molecules. How the results of an X-ray diffraction analysis of a crystal structure can give information on the stereochemistry of intermolecular packing is the subject of this article.

Molecules pack closely together in crystals to give an ordered arrangement. The arrangement of atoms throughout a crystal can therefore, as shown in Figure 1, be described in terms of two components: the atomic arrangement in one unit cell and the crystal lattice. Atomic parameters for crystal structures are derived by X-ray diffraction techniques. The crystal lattice is a device for illustrating the periodicity of the structure in three dimensions in which the entire arrangement of atoms in one unit cell is replaced by a point that is repeated with the dimensions of the unit cell, as shown in Figure 1. In the particular example shown in this figure there is only one molecule in the unit cell but, more generally, in most crystal structures, there are two, four or six asymmetric units (depending on the space group) within the unit cell and the asymmetric unit itself may contain one or more molecules, the same or different. Once atomic coordinates are available, the distances between atoms can be determined whether or not these atoms are in the same molecule. Thus if molecules in different

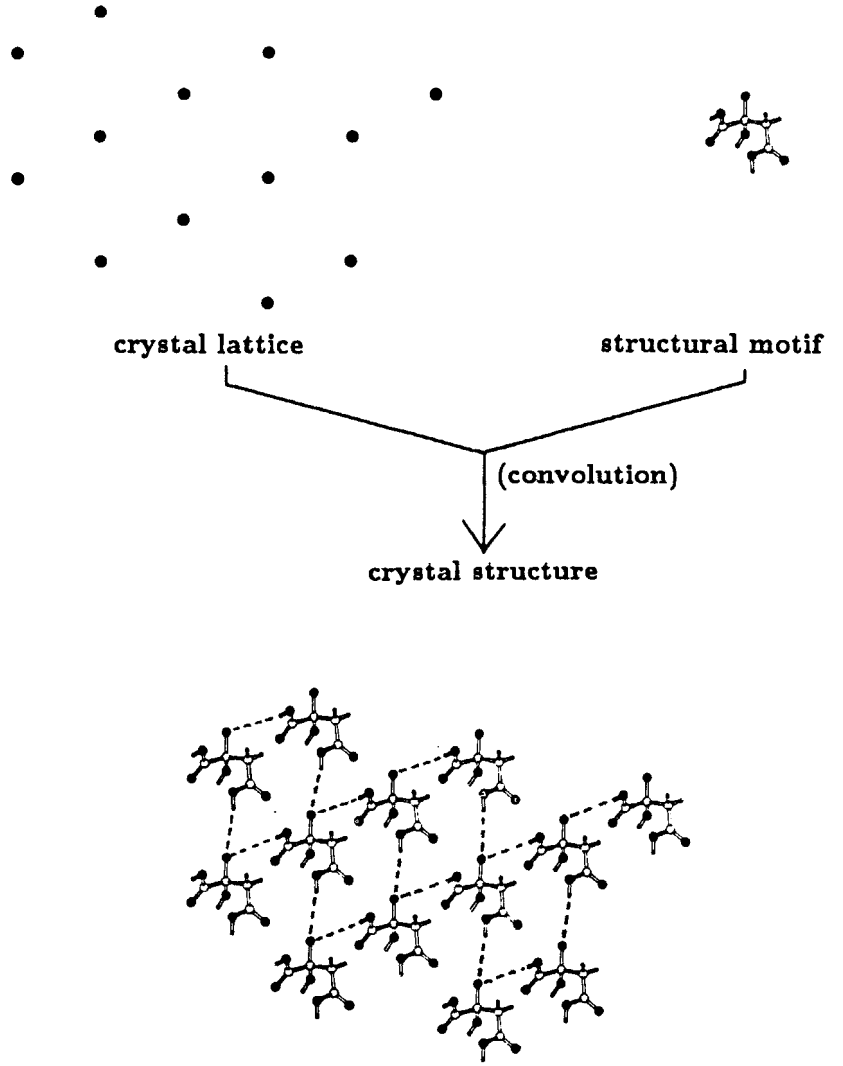

Figure 1. A crystal structure is a regular array of molecules. While the reported information is coordinates of one molecule, it is also possible to determine the directionalities of intermolecular interactions (marked by broken lines).

unit cells (or different asymmetric units) are considered, information on intermolecular interactions results.

There have been tens of thousands of crystal structures determined since the first X-ray diffraction experiment[1] in 1912. The majority of these have been elucidated in the last 20 years because methods for measuring the intensities of the diffraction pattern and analyzing the resulting data have been greatly facilitated by the growth of computer capabilities. Most reported crystal structures were studied in order to find the unknown chemical formula of a compound, such as an intermediate in a chemical synthesis or a compound isolated from natural sources, or to establish details of the stereochemistry and electron distribution in a compound. What is often not described in reported crystal structure determinations, apart from any hydrogen bonding or metal coordination, is how the molecules interact with each other. Pleas have been made for the reporting of fractional coordinates of molecules within the unit cell[2] because otherwise, if the atomic parameters are given in Cartesian coordinates, the packing information is lost.

Since data from tens of thousands of crystal structures are now available, looking for specific intermolecular interactions would at first seem to be a daunting process. This was recognized early on when, for example, R.W.G. Wyckoff initiated volumes describing each crystal structure, an endeavor still continued as "Structure Reports".[3] A quotation from J.D. Bernal in "Science and History"[4] in 1965 sums up the positive attitudes of the times: "However large an array of facts, however rapidly they accumulate, it is possible to keep them in order and to extract from time to time digests containing the most significant information while indicating how to find those items of specialized interest. To do so, however, requires the will and the means." Bernal had already been responsible for the initiation of computer-based data bases for the storage and retrieval of data from crystal structure determinations.

There are several requirements for efficient and "user-friendly" computer-based data storage of crystal structure information. There is a need for the information to be accurate, for all available information to be included in the database and for this information to be readily accessible for use in analyzing general or specific problems. There is also a need for additional computer programs for handling and analysis of the

data in a data base. These requirements have all been addressed by the X-ray crystallographer[5] and five databases are currently available to those interested in the analysis of crystal structure results. These are the Cambridge Crystallographic Database (which contains all data on organic compounds),[6] the Inorganic Crystal Structure Database,[7] the Protein Data Bank,[8] the NRCC Metals Crystallographic Data File[9] and the Powder Diffraction File.[10] These each contain atomic coordinates, unit cell dimensions and space group information except for the last which contains more detailed information on the actual diffraction pattern than do the other databases.

The Cambridge Crystallographic Database, for example, can be accessed by listing on the computer an arrangement of atoms (that is, their connectivities) that one wishes to search for. Each crystal structure in the Database is searched for groups of atoms that have this arrangement. The output from such a search is a listing of the name of the compound, its journal reference and the atomic parameters for each crystal structure containing the fragment that was used as input. This output file can then be searched for additional information. For example, one might initially search the data file for all ketones containing the C-CO-C grouping. Then, additionally, it is possible to search the output file for those with a hydrogen bond to the oxygen atom and to obtain coordinates not only of the ketone but also of the group that forms a hydrogen bond to it. This is the basis of the searches for intermolecular interactions to be described here.

The strategy for analysis is exemplified by an analysis of the manner by which metal ions interact with carboxyl groups.[11] The Cambridge Structural Database was searched for crystal structures containing both metal ions and carboxyl groups. The aim was to determine whether or not the metal ion lay in the plane of the carboxyl group or not and whether it lay in an orientation described as *syn*-, *anti*- or direct as shown in Figure 2. The designation "direct" is given to orientations in which the metal ion is shared equally by the two oxygen atoms of the carboxyl group. Since the initial aim was to find the interaction with an isolated carboxyl group, it was necessary to eliminate crystal structures in which a given molecule or ion formed more than one interaction with the metal ion. In the remaining list of metal carboxylates the carboxyl group was moved to a fixed position, the same for each crystal structure and the metal

ion locations and identities were then recorded with respect to this fixed location. The result was a scatterplot of the type shown in Figure 3 (item 3).

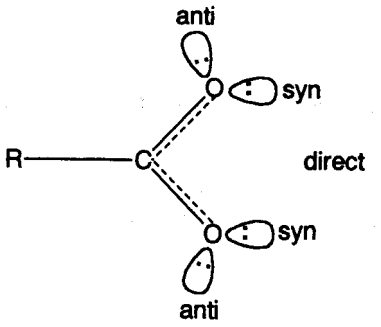

Figure 2. Designations of lone pairs and directions of binding for a carboxyl group.

However, scatterplots are often hard to interpret by eye because one has a tendency to concentrate too heavily on outliers. Therefore the scatterplot was simplified by placing a Gaussian function on each point of the scatterplot and the resulting set of overlapping peaks was contoured, as for a crystallographic electron density map,[12] as shown in Figure 3 (item 6). Peaks in the contour plot indicate areas in which the metal ion is most likely to lie, that is, the directional preferences of binding of the carboxyl group for that particular metal ion. Some examples now follow.

The first case in which intermolecular interactions in crystal structures were studied was an analysis of the surroundings of divalent sulfur, illustrated in Figure 4. It was shown by Parthasarathy and co-workers[13] that electrophiles, such as metal cations or hydrogen bond donors, approach the sulfide group (C-S-C) in a direction that is nearly perpendicular to the sulfide plane. On the other hand, nucleophiles, such as negatively charged groups or ions, approach the sulfur along a line that is almost in the sulfide plane. Therefore electrophilic attack of divalent sulfur would be expected to occur in a direction perpendicular to the sulfide plane while nucleophilic attack would be expected to occur in the direction of an S-C bond.

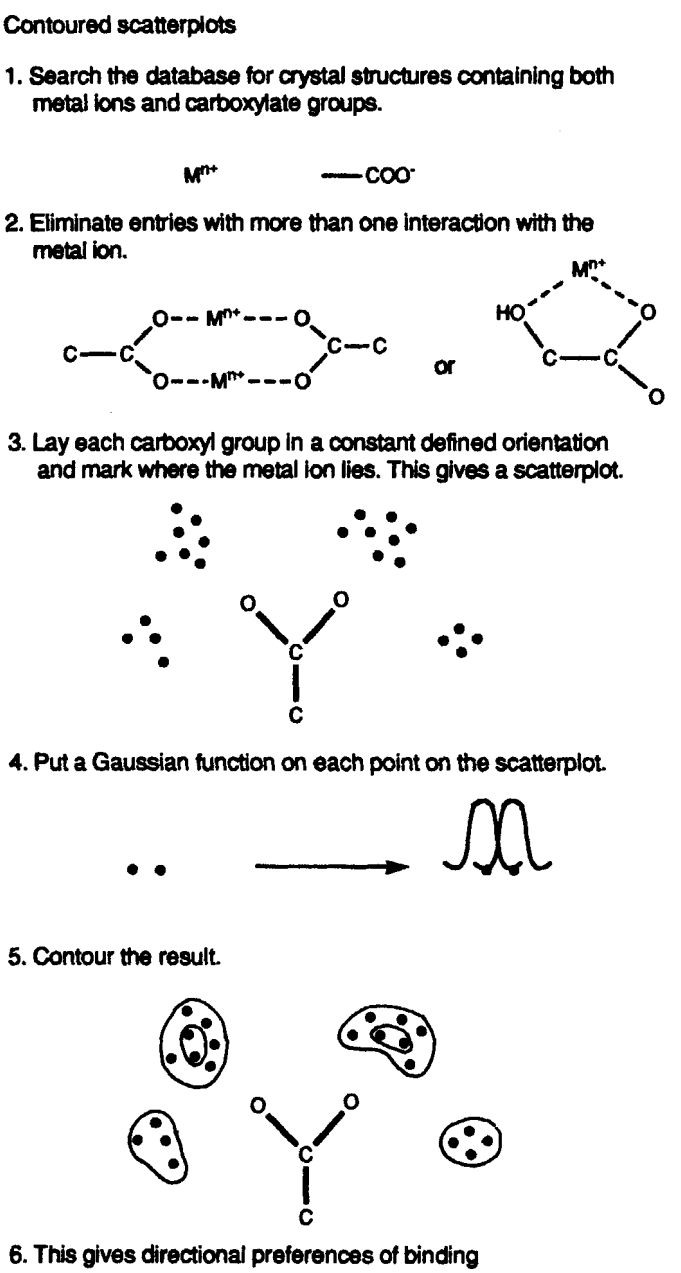

Figure 3. Strategy for producing contoured scatterplots

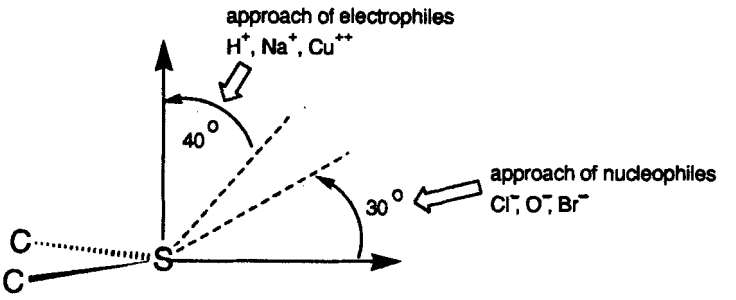

Figure 4. Directions of approach of electrophiles and nucleophiles to a C-S-C group.[13]

The directionality of hydrogen bonding to sp^2 and sp^3-hybridized oxygen atoms in ethers, ketones, esters and epoxides was investigated by the methods just described.[14] The largest concentration of hydrogen-bonded H-O or H-N was found to lie in the directions conventionally considered to be the directions in which the lone-pairs of electrons on the oxygen atoms point, that is, at about 109° to the C-O bond. For the C-O-C oxygen atom in epoxides and ethers the scatterplots, shown in Figure 5, contained two large peaks that lay in a plane perpendicular to that of the C-O-C group. There seemed to be a tighter distribution of hydrogen bond donors towards the oxygen atom in ethers, than towards the oxygen atom in ketones or epoxides. The cause of this is under further investigation. An analysis of metal cation···ether oxygen interactions has also been reported.[15] Most cations are found to approach the ether group along a tetrahedral lone pair direction; the exception was Li^+ which lies along the trigonal lone pair direction.

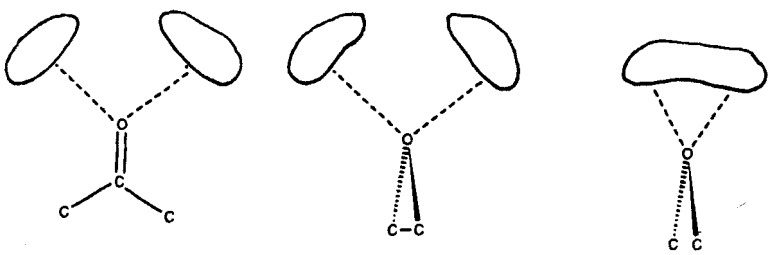

Figure 5. Directionality of hydrogen binding to ketones, epoxides and ethers.[14]

The intermolecular interactions of the C-F bond have been analyzed in a similar way.[16] The C-F bond is often considered rather inert, like the C-H bond, and this conception is reinforced by the inertness of teflon, used in "no-stick" frying pans. In the crystal structure of ethyl fluorocitrate (Figure 6) there is a short C-F···H-N^+ interaction, although not as strong as the C-OH···N interaction also formed.[16] An analysis, by use of the Cambridge Structural Database, shows that the fluorine atom of a C-F bond approaches a hydrogen atom attached to an oxygen or nitrogen atom at an F···O distance that has a minimum value around 2.85 Å. This may be compared with a van der Waals radius sum of 2.75 Å and a minimum O···O distance of around 2.38 Å. The minimum F···H distance was 2.3 Å, much longer than the minimum O···H distance of 1.2 Å in a symmetrical hydrogen bond. In these crystal structures, if there is any possibility of hydrogen bond formation that involves only oxygen atoms, these will form, rather than hydrogen bonds involving fluorine. These findings suggest that C-F···H-O,N hydrogen bonds can occur, with C-F as a weak proton acceptor. They may be evident if there is an excess of proton donors over proton acceptors in the crystal structure.

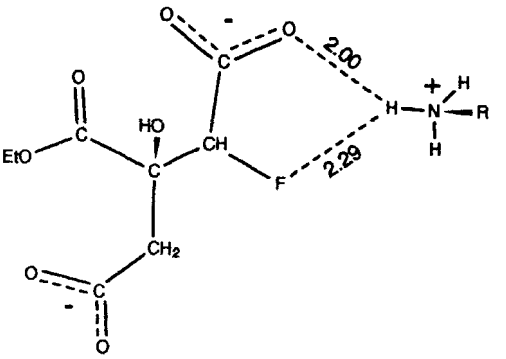

Figure 6. Hydrogen bonds in an alkylammonium salt of ethyl monoethylfluorocitrate.[16]

Why does the C-F bond, where fluorine is the most electronegative atom known, have such weak hydrogen-bonding power when H-F forms the shortest hydrogen bonds known. The strength of a hydrogen bond mainly depends on the

electronegativity of the donor, that is, how readily the proton is lost from, say, -OH, rather than how readily it is taken up by the acceptor (the fluorine of a C-F bond). The nature of the acceptor atom is also of some importance since we have shown here that oxygen works better than fluorine. A fluorine atom in a C-F bond can only act as a proton acceptor, in contrast to the case of an oxygen atom in the C-OH bond which can act both as a proton donor and acceptor. The result for oxygen is a cooperative effect - oxygen can give and take while fluorine can only give.

The interaction of the C-F bond with metal ions in crystal structures was also examined.[16] The C-F bond does not generally appear to form short F···M^{n+} contacts unless M^{n+} is an alkali metal cation. The metal-fluorine contacts were analyzed in terms of the requirement that, in the crystalline state, the sum of the bond valences of the anions around the cation are near 1.0 for a monovalent cation, as expected. This was indeed the case indicating that the fluorine atom contributed to the local neutralization of the charge of the cation. This contribution should be about 1/6 electrons for fluorine around an octahedrally-coordinated alkali metal cation. These analyses have helped in an understanding of the role of certain inhibitory fluorine compounds in enzyme systems. For example, only one isomer of fluorocitrate can inhibit the enzyme aconitase. We have determined the absolute configuration of this stereoisomer[17] and it is found that the fluorine atom resides on the portion of the citrate molecule that is not involved in the enzymatic reaction which converts it to *cis*-aconitate or isocitrate. However, if the fluorocitrate binds to the enzyme in the manner indicated by the crystal structures of metal complexes, shown in Figure 7, and binds near a positively-charged site, then it will bind the wrong way and the stereochemistry of the inhibition can then be explained.[18]

Another interaction studied in this way is the manner by which a carboxyl group recognizes a metal ion.[11] For this analysis, as mentioned in the earlier description of the methods used, care was taken to eliminate cases where an additional group, such as an adjacent hydrogen group in the same molecule took part in this interaction so that results were for the interaction of an isolated carboxyl group with the metal ions. Various possible locations of binding were shown in Figure 2. These include in-plane and out-of-plane possibilities and *syn, anti* and "direct" orientations. The Cambridge

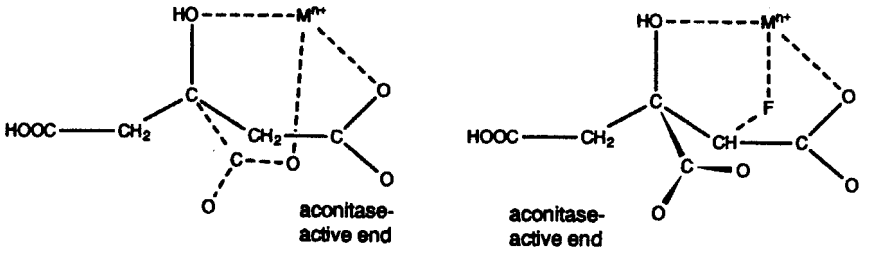

Figure 7. Binding of metals to citrates and fluorocitrates and implications for the action of aconitase for which citrate is a substrate and fluorocitrate is an inhibitor.

Structural Database was searched for such interactions. Only when there were at least ten cases for a cation in a particular oxidation state were the results used. It was found that for low-valent metal ions, especially sodium, potassium, rubidium and cesium, the metal ion could lie either in-plane or out-of-plane of the carboxyl group, as shown in Figure 8(a). This may be why it is easy to form crystalline salts of alkali metal polycarboxylates such as the citrates, because the requirements for the location of metal ions with respect to the carboxyl group is not strict. On the other hand, transition metal ions were generally found to lie in the plane of the carboxyl group. Overall the preference was for the *syn* orientation, but individual metal ions had somewhat different preferences (Figure 8(b)). Hydrogen-bonded carboxyl dimers are formed in this orientation. Some metal ions, notably divalent calcium and other metal ions with M···O distances near 2.6 Å, tended to lie in the plane and shared both oxygen atoms of the carboxyl group. In proteins this can be a way in which the higher coordination number of calcium versus the other ions can be satisfied in the same environment of carboxyl groups.[19]

When additional groups are near the carboxyl groups, metal chelation results. The α-hydroxycarboxylate group is found in many important biochemical compounds such as citrates and malates. It is an excellent metal-chelating group and its geometry in several crystal structures has been investigated. If the cation is of a suitable size the entire chelate group is approximately planar. If no cation is present, as in the free acid,

the hydrogen atom of the hydroxyl group (but not that of the carboxyl group in free acids) then forms an internal hydrogen bond across the group, as shown in Figure 9.

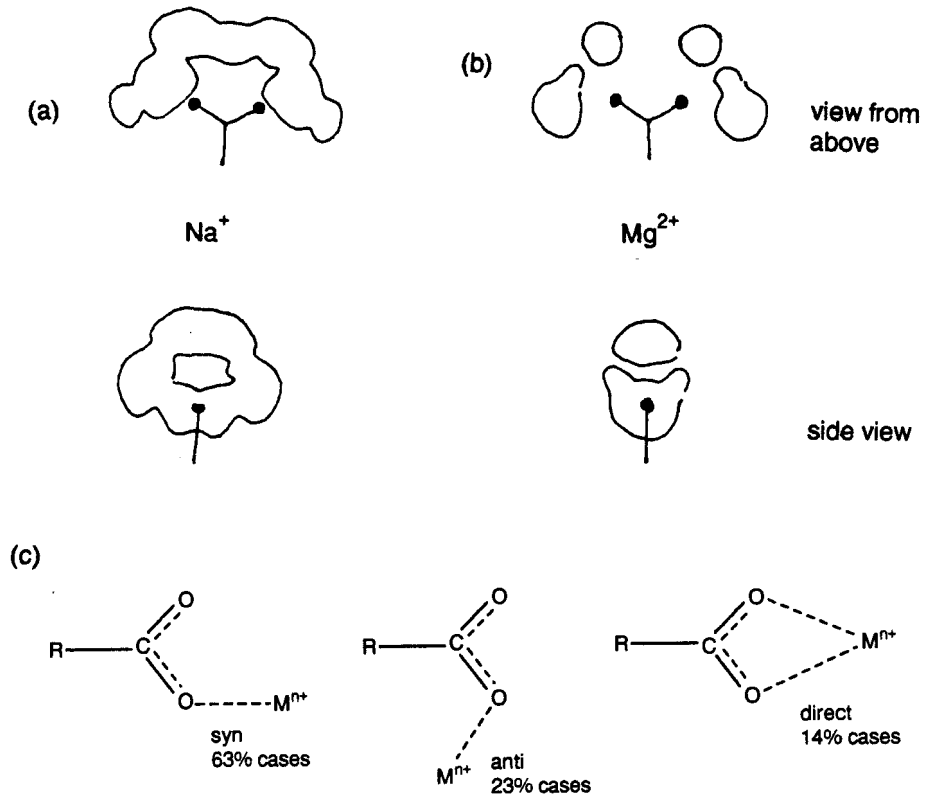

Figure 8. Contoured scatterplots for (a) Na^+ and (b) Mg^{2+} bound to c carboxylate ion. (c) The directionalities of binding for all types of cations.

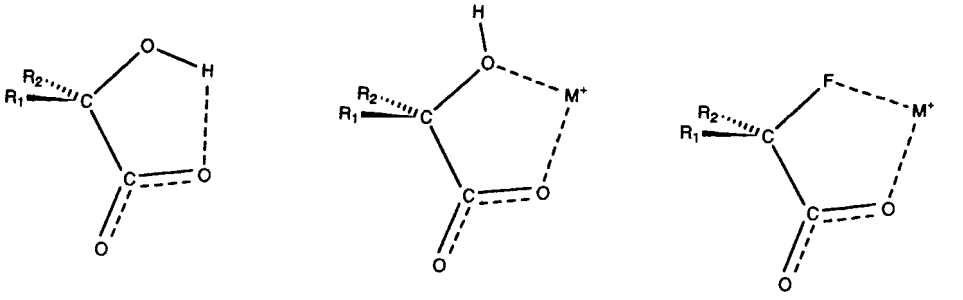

Figure 9. Binding to α-substituted carboxylates (citrate and fluorodeoxycitrate).[20]

A study of one form of potassium citrate[20] revealed, surprisingly, that the potassium ion did not chelate the α-hydroxycarboxylate group. However, the hydroxyl hydrogen atom formed an internal hydrogen bond, as shown in Figure 9. The location of the hydrogen atom was verified by a subsequent neutron diffraction study. When the central hydroxyl group of citric acid is replaced by fluorine, the α-fluorocarboxylate group is still a good chelating group, but no hydrogen atom is available to form an internal hydrogen bond in salts such as the potassium salt. In the case of dipotassium 3-fluorodeoxycitrate, a potassium cation is chelated by the α-fluorocarboxylate group. Thus the α-hydroxycarboxylate group will either chelate a cation or form an internal hydrogen bond via the hydroxyl hydrogen atom. When no such hydroxyl-type hydrogen atom is available, the group will chelate a metal cation as is the case in dipotassium 3-fluorodeoxycitrate (Figure 9).

The information that we obtained on the mode of binding of carboxyl groups, together with that also found for imidazole groups, was used to locate the metal-binding sites in the crystal structure of the enzyme D-xylose isomerase[21] that we are presently working on. The locations of metal ions in the probe, shown in Figure 10, and the arrangement and binding of the metal ions in the active site of this enzyme, is shown in Figure 11. The method gave correct locations for the metal ions.

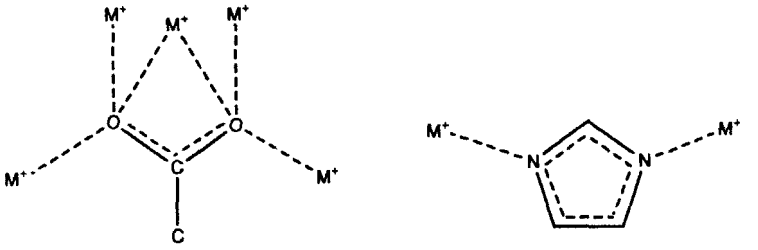

Figure 10. Probes used to find metal-binding positions in a protein.

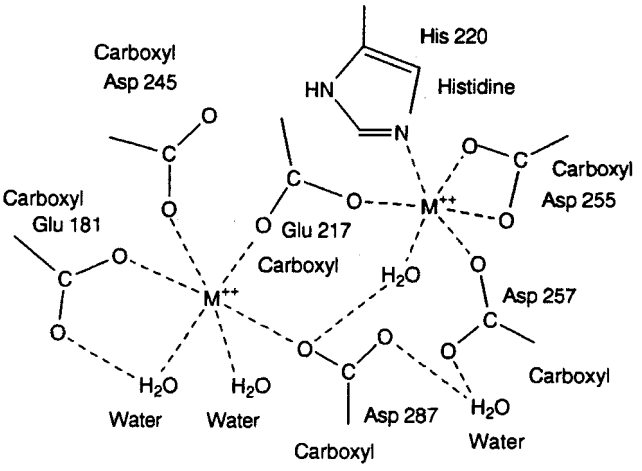

Figure 11. Metal ion binding in D-xylose isomerase.[19]

I wish to thank the following people who have made significant contributions to the studies described here: H.L. Carrell, Peter Murray-Rust, Christopher Carrell, Andrew Carrell, Amy Kaufman, Daiqing Liao and Gautam Desiraju. I also thank the National Institutes of Health for support by grants GM-44360 and CA-10925. This work was also supported by grants CA-06927 and by an appropriation from the Commonwealth of Pennsylvania.

REFERENCES

1. W. Friedrich, P. Knipping and M. Laue, *Sitzungsberichte der mathematisch-physikalischen Klasse der Königlichen Bayerischen Akademie der Wissenschaften zu München*, pp. 303-322 (1912). English translation: J.J. Stezowski, in *Structural Crystallography in Chemistry and Biology*, edited by J.P. Glusker (Hutchinson & Ross, Stroudsburg, PA, 1981), pp. 23-39.
2. J.P. Glusker, *J. Accts. Chem. Res., 18,* 95 (1985)
3. *Structure Reports. A. Metals and Inorganic. B. Organic (including organometallic compounds)* (Oosthoek (IUCr), Utrecht, yearly to date).
4. J.D. Bernal, *Science in History*, 3rd edn. (Hawthorn, New York, 1965), p. 943.
5. *Crystallographic Databases. Information Content. Software Applications. Scientific Applications.* Published by the Data Commission of the International Union of Crystallography (Bonn, Cambridge, Chester, 1987).

6. F.H. Allen, O. Kennard and R. Taylor, *Acc. Chem. Res., 16,* 146-153 (1983).
7. G. Bergerhoff, R. Hundt, R. Sievers and I.D. Brown, *J. Chem. Inf. Comput. Sci., 23,* 66-69 (1983).
8. F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer Jr., M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi, *J. Mol. Biol. 112,* 535-542 (1977).
9. J.R. Rodgers and G.H. Wood, in *Crystallographic Databases. Information Content, Software Applications, Scientific Applications* (International Union of Crystallography, Bonn, Cambridge, Chester, 1987), Section 2.3, pp. 96-106.
10. J.D. Hanawalt, in *Crystallography in North America,* Eds. D. McLachlan Jr. and J.P. Glusker (American Crystallographic Association, New York, 1983), Section D, Ch. 2, pp. 215-219.
11. C.J. Carrell, H.L. Carrell, J. Erlebacher and J.P. Glusker, *J. Amer. Chem. Soc., 110,* 8651-8656 (1988).
12. R.E. Rosenfield, Jr., S.M. Swenson, E.F. Meyer, Jr., H.L. Carrell and P. Murray-Rust, *J. Mol. Graphics, 2,* 43-46 (1984).
13. R.E. Rosenfield, R. Parthasarathy and J.D. Dunitz, *J. Amer. Chem. Soc., 99,* 4860-4862 (1977).
14. P. Murray-Rust and J.P. Glusker, *J. Amer. Chem. Soc., 106,* 1018-1025 (1984).
15. P. Chakrabarti and J.D. Dunitz, *Helv. Chim. Acta, 65,* 1482-1488 (1982).
16. P. Murray-Rust, W.C. Stallings, C.T. Monti, R.M. Preston and J.P. Glusker, *J. Amer. Chem. Soc., 105,* 3206-3214 (1983).
17. W.C. Stallings, C.T. Monti, J.F. Belvedere, R.K. Preston and J.P. Glusker, *Arch. Biochem. Biophys., 203,* 65-72 (1980).
18. H.L. Carrell, J.P. Glusker, J.J. Villafranca, A.S. Mildvan, R.J. Dummel and E. Kun, *Science, 170,* 1412-1414 (1970).
19. N.C.J. Strynadka and M.N.G. James, *Annu. Rev. Biochem., 58,* 951-958 (1989).
20. H.L. Carrell, J.P. Glusker, E.A. Piercy, W.C. Stallings, D.E. Zacharias, R.L. Davis, C. Astbury and C.H.L. Kennard, *J. Amer. Chem. Soc., 109,* 8067-8071 (1987).
21. H.L. Carrell, J.P. Glusker, V. Burger, F. Manfre, D. Tritsch and J-F. Biellmann, *Proc. Natl. Acad. Sci. USA, 86,* 4440-4444 (1989).