Pure & Appl. Chem., Vol. 70, No. 2, pp. 319–324, 1998.
Printed in Great Britain.

Quantum crystallography, a technique for extending the concept of structure

J. Karle[a], L. Huang[a*] and L. Massa[b]

[a]Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, D. C. 20375-5341

[b]Hunter College and the Graduate Center of City University, New York, New York 10021-5024

Abstract: A developing area of research concerns a technique for extracting quantum mechanically valid properties from X-ray diffraction experiments. Quantum mechanics and crystallography are joined through the fact that the electron distributions around atoms are the source of X-ray diffraction and electron density distributions are observables that lend themselves readily to quantum mechanical description. Direct contact with the X-ray diffraction data is made by equating the structure factor magnitudes, which are readily obtained from the measured X-ray diffraction intensities with the magnitudes of Fourier transforms of the quantum mechanical description of the electron distribution. The article concerns a further discussion of quantum crystallography and its implications. To create a quantum mechanical model, crystallographic information in the form of atomic positions is used. Since quantum crystallography is applicable to very large structures, methods for handling very large structures are described.

INTRODUCTION

This article concerns some new and developing approaches to the investigation of molecular structure and properties by both theoretical and experimental means. The subject to be discussed is the developing field of quantum crystallography (QCr). QCr has as its objectives insights into the structure and physical properties of crystalline substances that extend those of current practice. The methodology is in place and it has as its basis the intimate combining of crystallographic data and quantum mechanical models in such a way that wave functions can be obtained that are consistent with the crystallographic data (ref. 1). This lends itself readily to the calculation of electron density distributions, atomic charges, molecular energies and many other properties.

In the course of developing these aspects of QCr, it was discovered that it would be rigorously possible to make *ab initio* quantum mechanical calculations on arbitrarily complex molecules with an essentially linear increase in computing time as a function of complexity (ref. 2). In fact, when *ab initio* programs become available on high performance parallel computers, the increase in computing time should hardly be noticed if the number of parallel nodes is sufficiently large to handle the individual fragments. The possibility of making *ab initio* quantum mechanical calculations on molecules large and small, and parts of molecules, greatly facilitates the applications of QCr (Note a).

*Permanent Address: Geo-Centers, Inc., 10903 Indian Head Highway, Fort Washington, MD 20744
Note a: Research supported by the Office of Naval Research and Public Health Grant GM30902.

Some of the aspects of these new opportunities will be illustrated with the use of a naturally occurring antibiotic, the hexadecapeptide, leu^1-zervamicin. In the next section, structural characteristics of this zervamicin will be described, as well as some special properties that derive from its structure and packing that may have relevance to the mechanism of ion-transport through membranes.

BACKGROUND

Some sample calculations have been made on leu^1-zervamicin, Ac-Leu-Ile-Gln-Iva-Ile-Thr-Aib-Leu-Aib-Hyp-Gln-Aib-Hyp-Aib-Pro-Phol (Aib: α-aminoisobutyric acid; Iva: isovaline; Hyp: 4-hydroxyproline; Phol: phenylalininol) (ref. 3). Zervamicin is known to transport potassium ions across cell membranes. Its crystal structure indicates it to be a helical peptide. This is consistent with experimental observations that helical peptides which are ionophores, assemble to form channels so that ions can pass through. Questions arise concerning the number of helical peptides required to form a channel, the size and shape of the pores and other structural features concerning the packing of the helical peptides in a membrane. Clues concerning answers to such questions may come from an examination of the packing of the helical molecules in a crystal structure.

Crystal structure investigations of zervamicin show (refs. 3,5,6), for the type of packing found in the crystalline state, that it requires three molecules to form the walls of an ion channel. The same molecules form the walls of additional channels. The bent helix that forms the structure of the zervamicin molecule has hydrophobic side chains extending from the peptide residues on one side of the molecule and on the other side there are several polar side chains, e.g., at residues 3, 6, 10 and 13. This is called an amphiphilic helix, polar on one side and hydrophobic on the other. One curious feature is the side chain of residue 11, glutamine. This side chain is attached to the backbone on the hydrophobic side of the molecule, but wraps around in such a fashion that its polar end projects into the hydrophilic side of the peptide.

One of the crystal structure investigations of zervamicin (ref. 3) showed the side chain of glutamine 11 in two different orientations, Fig. 1, i.e., both conformers co-crystallize. One orientation closes the channel and the second opens it. The closed form occurs to

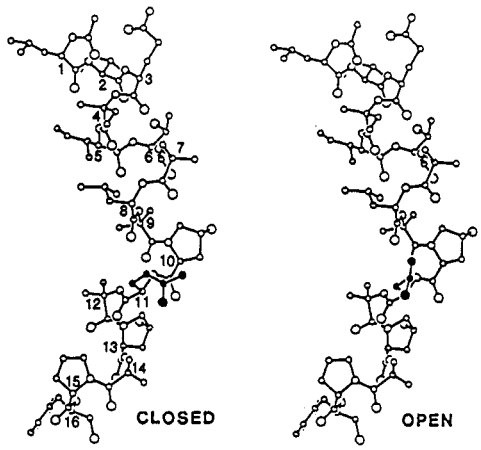

Fig. 1 Closed and open forms of leu^1-zervamicin. The dark circles indicate the atomic positions of the side chains of glutamine 11 in both forms (ref. 3).

about 80% in the crystal. It was possible to simulate the passage of the K^+ ion through the channel. In the course of the simulation, it was found that the gate formed by the side-chain has to open and then close in order for the K^+ ion to proceed through the channel. It appears that the packing configuration formed in the crystal requires the gate to open and

close each time an ion goes through. Experiment shows (ref. 7) that a potential is required for the transport of the K^+ ions. If the packing in a cell membrane resembles that in the crystal, it is conceivable that the zervamicin crystal is suggesting a gating mechanism on the level of atomic resolution.

To illustrate some aspects of quantum crystallography, the nature of the calculations that may be applied to the region involving the side chain of glutamine 11 will be described.

QUANTUM CRYSTALLOGRAPHY

Quantum crystallography (QCr) concerns the combining of X-ray diffraction data for crystals with quantum mechanics with the the objective of obtaining accurate wave functions that are consistent with the X-ray data (ref. 1). Quantum mechanics and crystallography are readily combined through the fact that the electron distributions around atoms are the sources of X-ray scattering and electron density distributions are readily described by quantum mechanical models. The manner in which experiment and theory are connected is through structure factor magnitudes, quantities that are readily obtained from the measured X-ray diffraction intensities, and also obtained from certain Fourier transforms of the quantum mechanical description of the electron density distribution. If the model and the experiment are in agreement, the structure factor magnitudes should be the same. The applications of QCr can be carried out in a fashion that assures that this agreement will initially be rather close. Adjustments can then be readily carried out.

One form of the quantum mechanical description of the electron density distribution involves molecular orbitals and an associated matrix. This matrix is required to be a projector with a normalized trace, which imposes strong constraints on the manner in which the final adjustment between theory and experiment may be carried out. The method for making the fit to the X-ray data makes use of least-squares calculations for which the defining equations are the structure factor equations and the equations that arise from the conditions that define a projector matrix with a normalized trace. The variables are the elements of the projector matrix. Other parameters such as atomic positions can also be refined. The resulting wave functions should provide reliable information concerning electron densities, charge distributions, electrostatic potentials and other quantities of interest.

QCr applications are greatly facilitated by having good projector matrices initially available. This may be achieved by making purely quantum mechanical calculations and extracting projector matrices from them. The fact that computer programs may be limited in the complexity of the structures that can be computed at one time does not limit the size of a structure for which a projector matrix may be obtained. This is so because it is possible to carry out *ab initio* quantum mechanical calculations for very large structures with only a linear increase in time with increase in complexity. As noted previously, when computer programs based on molecular orbital methods in quantum mechanics are arranged for high performance parallel computers with sufficient nodes, the time increase will be essentially unnoticeable.

The key to performing *ab initio* calculations for large structures is the concept of the fragment calculation. It is based, rigorously, on the observation that overlap integrals which enter into the definition of a projector matrix, are very small when atoms are far apart. These integrals measure the extent of the overlap of atomic orbitals and at approximately a distance of 5 Å or somewhat more for first row atoms, for example, they are negligible in value. The practical aspect is that structures that are excessively large can be divided into fragments for which *ab initio* calculations are readily feasible. The special way in which fragments may be formed and projector matrices may be extracted is described in references 1 and 2.

Fragment and QCr calculations in the glutamine 11 region

It is of interest to examine comparable fragments in the vicinity of glutamine 11 for both forms of leu^1-zervamicin. One of the virtues of appropriate fragment calculations, applied to regions of interest, is that they make it unnecessary to apply the calculations to the entire molecule. This is of special value when the molecule is large. If a molecule is small enough, then the only valid fragment, based on the diminishing of overlap of atomic orbitals with increasing distance, is the moelcule itself.

As noted earlier, the behavior of the side chain of glutamine 11 in leu^1-zervamicin, having both the "open" and "closed" conformations in a co-crystallized sample, suggests that it may be of interest to examine the glutamine 11 region with the use of quantum crystallography. A discussion of the procedure to be followed in making this calculation would illustrate a number of new features involved. One feature concerns the fact that matrices called kernel matrices $\mathbf{R}_{i\kappa}$ may not only be summed to form **R** matrices for the entire structure, but also the kernel structure factors $F_{\mathbf{h}i\kappa}$ may be summed to form the structure factors for the total structure $F_{\mathbf{h}t}$. A structure factor is a quantity that may be computed by taking a certain Fourier transform of the quantum mechanical description of the electron density distribution. It is generally a complex number whose magnitude squared is related to the intensity of diffraction of X-rays.

It is now our intention to illustrate how to compute kernel $\mathbf{R}_{i\kappa}$ matrices, obtain from them their corresponding partial structure factors $F_{\mathbf{h}i\kappa}$ with and without thermal effects and form from them total structure factors with and without thermal effects. With this accomplished, it is possible to put the experimental and theoretically calculated data on the same scale, remove the thermal effects from the experimental data, correct the experimental data for systematic errors, estimate the experimental values for the partial structure factors associated with glutamine 11 and, finally, perform quantum crystallographic calculations on the residue, glutamine 11.

Calculation of kernel matrices, $\mathbf{R}_{i\kappa}$

R matrices for entire molecules are composed of elements that are linear combinations of the products of the coefficients of pairs of orbitals. There are sensible ways of dividing up the complete matrix into kernel matrices, $\mathbf{R}_{i\kappa}$, in such a way that the kernel matrices add up to the matrix for the entire molecule. As noted, structure factors may be computed for the $\mathbf{R}_{i\kappa}$. Typical forms for two $\mathbf{R}_{i\kappa}$ are seen in Fig. 2. The dark squares have as elements linear combinations of products of orbital coefficients associated only with the atoms forming the kernels. The gray squares have linear combinations of products of coefficients in which one coefficient always comes from a neighborhood orbital and one from a kernel orbital. The formula used for the calculation of the partial structure factor, $F_{\mathbf{h}i\kappa}$, from theory is

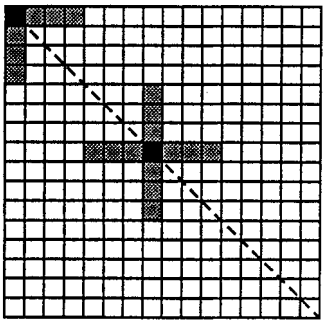

Fig. 2 Schematic diagram of two kernel projector matrices.

$$F_{\mathbf{h}i\kappa T} = 2\mathrm{tr}\left[\mathbf{R}_{i\kappa} \sum_{j=1}^{n} \widehat{R}_j \mathbf{f}_{hi\kappa} \square T\right] \tag{1}$$

where

$$\mathbf{R}_{i\kappa} = (\mathbf{C}^{\dagger}\mathbf{C})_{i\kappa} \tag{2}$$

and the matrix **C** has only those elements that belong to the kernel for glutamine 11, i.e. i = 11. All elements of $\mathbf{C}^{\dagger}\mathbf{C}$ have a weight of one when both orbitals belong to the kernel and a weight of 0.5 otherwise. Therefore, the elements in the dark areas of Fig. 2 have a weight of one, and the ones in the gray areas have a weight of 0.5. $\widehat{R}_j$ is the jth space group operator among n operators and the matrix $\mathbf{f}_{hi\kappa}$ has as its elements the individual Fourier transforms of the elements of $(\psi\psi^{\dagger})_i$. The dimensions for the matrices $\mathbf{R}_{i\kappa}$ and $\mathbf{f}_{hi\kappa}$ in the case that i=11, referring to the matrices for glutamine 11, are 344 x 344. The symbol □ implies element by element multiplication of the matrix by thermal effects derived from experiment (ref. 1). $F_{hi\kappa}$ is obtained if T is omitted.

Total structure factors with and without thermal effects

Having performed the calculation indicated in Eq. (1), it is now possible to calculate readily $F_{hTQ.M.}$ and $F_{hQ.M.}$, the structure factors for the entire molecule with and without thermal effects, respectively, i.e.,

$$F_{hTQ.M.} = \sum_{i=1}^{m} F_{hi\kappa TQ.M.} \tag{3}$$

where m is the total number of kernels, Q.M. indicates use of quantum mechanical theory and $F_{hQ.M.}$ is obtained by omitting the thermal effects from Eq. (1). The experimental structure factor magnitudes $|F_{hexp.}|$ are obtained from the experiment only for the entire structure and with the thermal effects inherent. For scaling purposes, the $|F_{hexp}|$ must therefore be compared with the theoretical data computed with thermal effects. The experimental structure factor magnitudes are summed, as are the theoretical structure factor magnitudes for the same set of **h**. Scaling is achieved by multiplying each of the experimental data by the ratio

$$\sum_{h} |F_{htheor}| / \sum_{h} |F_{hexp}|$$

Removal of thermal effects and correction for systematic errors

Thermal effects may be removed from the experimental data by use of

$$\frac{|F_{hexp}|}{|F_{hTexp}|} = \frac{|F_{hQ.M.}|}{|F_{hTQ.M.}|} \tag{4}$$

Since the other three quantities are known, Eq. (4) may be used to obtain values for the F_{hexp}. Once the $|F_{hexp}|$ have been obtained, it is possible to correct them for some types of systematic errors that occur as a function of the magnitudes of the $|F_{hexp}|$ and also as a function of the scattering angle.

The corrections for those specific types of scaled systematic errors are made by listing the scaled structure factor magnitudes from experiment, corrected for thermal effects, in descending order of magnitude and also listing them in the order of increasing scattering angle. Both listings are handled similarly. Corresponding to each list of experimental structure factor magnitudes is a second list of structure factor magnitudes obtained from quantum mechanical theory. These second lists are put in the same order as the corresponding lists of the experimental structure factor magnitudes. The adjustments for systematic errors are made on the assumption that over a range of a large number of structure factor magnitudes, perhaps 50 to 100, the average values of the theoretical structure factor magnitudes are correct. The correction procedure then becomes an adjustment of the experimental data so that corresponding averages become close in value. Simple mathematical techniques are available to effect these corrections.

Evaluation of experimental partial structure factors for glutamine 11

In order to perform quantum crystallographic calculations for glutamine 11, it is necessary to have an estimate of the values of the structure factor magnitudes associated with glutamine 11. To achieve this, we use

$$\frac{|F_{h(11)exp}|}{|F_{hexp}|} = \frac{|F_{h(11)Q.M.}|}{|F_{hQ.M.}|} \tag{5}$$

where all the quantities are known except the $|F_{h(11)exp}|$. With the values of the $|F_{h(11)exp}|$ available, it is possible to proceed with the quantum crystallographic calculations for glutamine 11.

Quantum crystallographic calculations for the residue glutamine 11

We use

$$F_{h(11)exp} = \sum_{j=1}^{n} \hat{R}_j [2\mathrm{tr}\mathbf{S}^{-1/2}_{(11)}\mathbf{P}_{(11)}\mathbf{S}^{-1/2}_{(11)}\mathbf{f}_{h(11)}] \tag{6}$$

where

$$\mathbf{S}_{(11)} = \int (\psi\psi^{\dagger})_{(11)} \tag{7}$$

is the matrix formed by integrating orbital products over all space. The **S** matrix is obtained in practice from the $\mathbf{f}_{hi}$ matrix when **h** = 0,0,0. When orbitals are far enough apart, the integrals are negligible in value and provide the rational basis for the fragment calculations. Comparing Eq. (6) with Eq. (1), we see that

$$\mathbf{S}^{-1/2}_{(11)}\mathbf{P}_{(11)}\mathbf{S}^{-1/2}_{(11)} = (\mathbf{C}^{\dagger}\mathbf{C})_{(11)} \tag{8}$$

The virtue of the form in Eq. (6) is that it allows improvements to be made in the quantum mechanical model by adjusting the elements of the projector **P** while preserving the projector property, namely $\mathbf{P}^2 = \mathbf{P}$. The details for performing the computations of quantum crystallography are described in a QCr study of maleic anhydride (ref. 8).

REFERENCES

1. L. Massa, L. Huang and J. Karle. *Int. J. Quantum Chem., Quantum Chem. Symp. No.* 29, 371-384 (1995).
2. L. Huang, L. Massa and J. Karle. *Int. J. Quantum Chem., Quantum Chem. Symp. No.* 30. [In press].
3. I. L. Karle, J. L. Flippen-Anderson, S. Agarwalla and P. Balaram. *Biopolymers* **34**, 721-735 (1994).
4. G. Boheim and H. A. Kolb. *J. Membrane Biol.* **38**, 99-150 (1978).
5. I. L. Karle, J. L. Flippen-Anderson, S. Agarwalla and P. Balaram. *Proc. Natl. Acad. Sci. USA* **88**, 5307-5311 (1991).
6. I. L. Karle, J. L. Flippen-Anderson, S. Agarwalla and P. Balaram. In *Structure and Function: Volume 2: Proteins - Proceedings of the Seventh Conversation in Biomolecular Stereodynamics*, (R. H. Sarma and M. H. Sarma, eds.) pp. 97-111. Adenine Press, Schenectady (1992).
7. P. Balaram, K. Krishna, M. Sukumar, I. R. Mellor and M. S. P. Sansom. *Eur. Biophys. J.* **21**, 1178-128 (1992).
8. L. Huang, L. Massa and J. Karle. *Application of Quantum Crystallography to Maleic Anhydride.* [In preparation, to be submitted to Acta Crystallographica A].