

Active-site studies of enzymes by X-ray diffraction methods

HILDA CID

Laboratorio de Biofísica Molecular, Facultad de Ciencias Biológicas,
Universidad de Concepción, Concepción, Chile

X-Ray diffraction is the only method currently available for the determination of the detailed three-dimensional structure of macromolecules. In this article, the principles behind the use of X-ray crystallography in the study of protein structure are presented. The results obtained in its application to several enzymes are discussed, as well as the limitations of the method.

Key terms: carbonic anhydrase; lysozyme; serine proteases; triose phosphate isomerase; X-ray crystallography.

STATEMENT OF THE PROBLEM

An accurate knowledge of the mechanism of action of an enzyme could be obtained directly if the following two experiments could be performed:

1. A direct observation of the enzyme with enough resolution to appreciate all the details of the active site.
2. A 'freezing' of the enzymatic reaction in order that all the intermediate enzyme-substrate and enzyme-product complexes last long enough to be identified.

The first experiment implies the determination of the tertiary structure of the enzyme, that is, the spatial location of every one of the thousands of atoms of its hundreds of amino acid residues. To discern an atom from its neighbors means to resolve distances of the order of $1 \text{ \AA} (= 10^{-7} \text{ mm})$. Since the limit of resolution of the naked eye is 10^{-1} mm , a microscope capable of a magnification of $10^6 \times$ would be needed. Therefore, the determination of the tertiary structure of a protein (in fact, the structure of any molecule, large or small) can be reduced to design an adequate microscope.

Since enzymes are usually very efficient catalysts, the mean life of the intermediate complexes formed in the reaction process are too short to allow a direct observation, even with the most powerful microscope. Therefore, the second experiment should be performed with poor substrates and under unfavorable conditions for catalysis. A useful alternative is to work with inhibitors that bind to the active site and remain there; in fact, several catalytic mechanisms are now better understood, because the exact location of inhibitors in the active site has been determined.

OPTICAL BASIS OF THE X-RAY DIFFRACTION METHOD

The resolution of a microscope cannot be better than the dimensions of the wave length of the radiation used to 'illuminate' the object; therefore, in order to obtain a resolving power of 10^{-7} mm a visible light microscope must be discarded. Wavelengths of 10^{-7} mm or lower in the electromagnetic spectrum belong to X-rays; also, according to De Broglie's theory, waves of this order of magnitude are associated with electrons and protons. The most powerful microscopes

actually in use are electron microscopes, but the best resolution obtained so far is not better than 10 to 20 Å, one order of magnitude larger than the resolution needed. The strong interaction between the electrons and the sample observed plays against the possibility of improving this limit of resolution. A proton microscope shows even more disadvantages.

X-rays can easily penetrate samples, but have other drawbacks: the propagation of X-ray waves is not appreciably disturbed by any transparent media, nor by electric or magnetic fields. There are no lenses that can appreciably deviate this radiation, a necessary condition for image formation. The most recent progress in this direction is the focalization of X-rays by the use of Fresnel's zone plates; these are circular diffraction plates formed by concentric rings, alternatively transparent and opaque with a spacing decreasing inversely with their radii. A resolution of the order of 300-1000 Å has been claimed for an X-ray microscope of this type, which is equal to the lowest distance attained between two adjacent rings in the zonal plate (Duke & Michette, 1990).

The current solution to this problem, and with it the discovery of the most powerful method to determine the tertiary structure of a protein, is based on optical principles established at the end of the last century by Abbe, and reformulated by Zernike in 1946 (Lipson, 1972). Abbe's theory states that the process of image formation takes place in

two stages of diffraction. Here diffraction is understood in the more general form, that is, as a process that occurs naturally, every time the propagation of a wave is limited in some way.

In the microscope schematized in Figure 1, the object is illuminated by light coming from the source at the left, becoming a source of wavelets scattered in every direction of space. All wavelets scattered in a particular direction, originated from different points of the object will add to each other. The amplitude and the phase of the resulting wave will be determined by the amplitude and phases of the individual wavelets; thus, each resulting wave will depend on the structure of the object. The new light distribution produced by the diffraction phenomena is called the 'diffraction spectrum' of the object, and, to a first approximation, and for a planar object, it will be focused at the second focal plane of lens I. It must be emphasized that the presence of this lens is not necessary to observe the diffraction spectrum, if the dimensions of the object are small enough to allow the proximity required for the combination of the wavelets originated on different points of the object.

In the second stage, all wavelets diverging from the diffraction spectrum are brought together by lens II in such a way that all waves originated at a single point of the object are focused at a single point in the image plane. This stage can be interpreted as a second diffraction, caused by the limi-

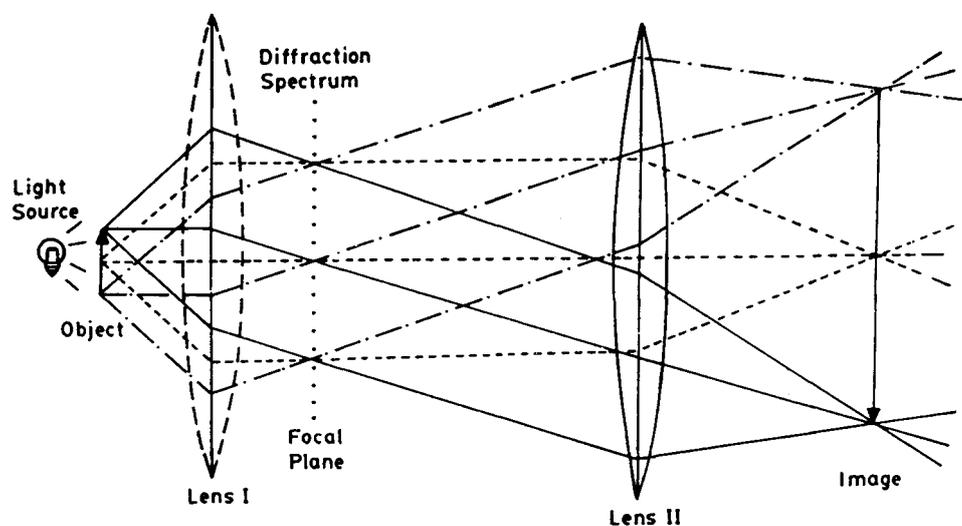


Fig 1. Abbe's theory of image formation in a microscope.

tation of the wave front by lens II. The light distribution at the image plane can be calculated mathematically with the aid of a computer, provided the amplitudes and phases of all waves diverging from the diffraction spectrum are known.

When the light source is replaced by an X-ray source, the diffraction spectrum is produced without the presence of lens I, and, if we can determine the amplitudes and phases of all the waves coming out from this diffraction spectrum, a computer can calculate the 'image' of our object. In this way one can have an 'X-ray microscope' without lenses and with it a resolution power of 10^{-7} mm.

MATHEMATICAL RELATIONSHIPS BETWEEN OBJECT, IMAGE AND DIFFRACTION SPECTRUM

The amplitudes and phases of a scattered wave in a certain direction will be the result of the sum of all the wavelets scattered in that direction by all the points of the object. If we represent by $F(\alpha)$ the wave scattered in a direction making an angle α with the optical axis of the microscope (Fig. 2), for a one-dimensional object we can write:

$$F(\alpha) = \int_{-\infty}^{+\infty} f(x) \exp \left[\frac{2\pi i x \sin \alpha}{\lambda} \right] dx \quad [1]$$

Where $f(x)$ represents the amplitude of the wavelet scattered in the direction α by a point located at the distance x from the axis.

Making $\sin \alpha / \lambda = X$, the expression simplifies to:

$$F(X) = \int_{-\infty}^{+\infty} f(x) \exp [2\pi i X x] dx \quad [1]$$

which is the expression of a Fourier Transform. This formula states that $F(X)$ is the Fourier Transform of $f(x)$. $F(X)$ represents the "light" distribution at the diffraction spectrum, and, since it is a complex number, the intensity of the diffracted wave is proportional to $|F|^2 = FF^*$, where F^* represents the complex conjugate of F . The meaning of $f(x)$ can be better understood if we identify our linear object with a linear molecule of several atoms located at distances x_j from the optical axis. To a first approximation, we can consider discrete centers of scattered wavelets, and, since the electrons are responsible for the scattering of the X-rays, the intensity of each scattered wavelet will be a direct function of the number of electrons of the scattering atom. The continuous function $f(x)$ becomes the 'scattering power' of each atom j , a function of the electron distribution in the object.

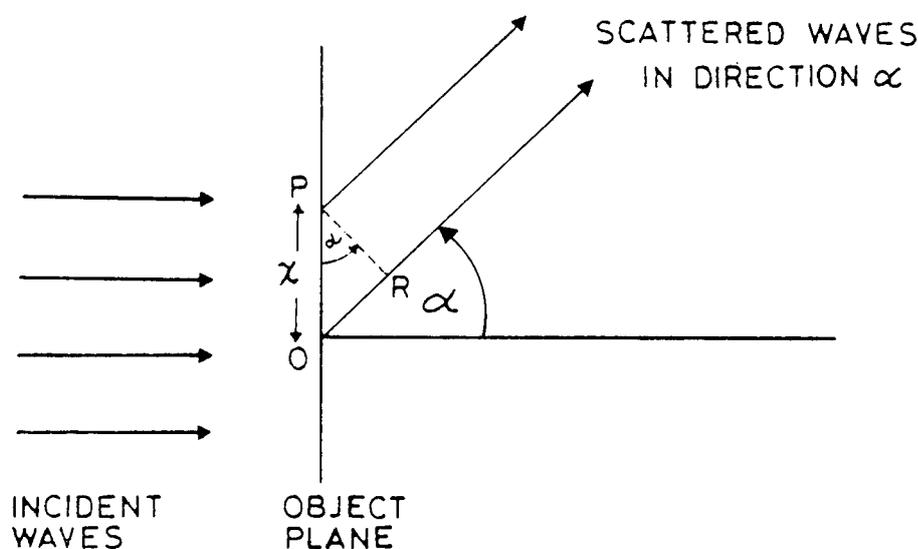


Fig 2. Calculation of the phase difference for two wavelets diffracted in the direction α , emerging from the origin and from a point P located at a distance x from the origin. OP is the optical path difference.

The expression of the Fourier Transform, considering point atoms, is changed to:

$$F(X) = \sum_{j=1}^n f_j \exp [2 \pi i X x_j] \quad [3]$$

$F(X)$ is called the 'structure factor' since it depends on the location (x_j) and on the nature (f_j) of the atoms, that is, the structure of the object. The structure factor is now given as a function of a coordinate X in the plane of the diffraction spectrum, instead of the variable α .

By an analogous deduction, it can be shown that the mathematical relationship between the image (I) and the diffraction spectrum is:

$$I(x) = \int_{-\infty}^{+\infty} F(X) \exp [-2\pi i X x] dX \quad [4]$$

In words, the image is the Fourier Transform of the Fourier Transform of the object.

For a three-dimensional object, one must deal with a three-dimensional image and a three-dimensional diffraction spectrum (equations 5 and 6).

Even though object and image could be assumed to be discontinuous functions, the diffraction spectrum of a single object is a continuous function:

$$I(xyz) = \iiint_{-\infty}^{+\infty} F(XYZ) \exp [-2\pi i (Xx + Yy + Zz)] dX dY dZ \quad [5]$$

$$F(XYZ) = \sum_j f_j \exp [2 \pi i (Xx_j + Yy_j + Zz_j)] \quad [6]$$

For a detailed treatment of the use of Fourier methods in crystallography, see Ramachandran and Srinivasan (1970).

DIFFRACTION SPECTRUM OF A REPEATING OBJECT

In order to calculate the image, it is necessary to know as precisely as possible, the amplitudes and phases of all the waves emerging from the diffraction spectrum. The

amplitude of any wave is proportional to the square root of its intensity. The first problem is solved by a careful measurement of the intensity distribution at the diffraction spectrum. However, a single molecule, even a protein molecule with thousands of atoms, is unable to produce diffracted X-rays strong enough to be detectable. It is necessary to obtain the cooperative diffraction of several thousands of such molecules to be able to measure, on a film or by means of a radiation detector, the diffracted X-ray beams. This cooperative diffraction is obtained with an ordered arrangement of the molecules we want to observe. A perfect three-dimensional periodic repetition of a molecule is a crystal, and the three vectors defining the repetition periods form the 'unit cell' of the crystal. Therefore, the first problem to be solved in the determination of the tertiary structure of a protein is to obtain protein crystals.

What is the relationship between the diffraction spectra produced by a single molecule and that produced by a three-dimensional periodic arrangement of thousands of identical molecules?

Figure 3 shows the diffraction spectrum of a circular hole and that produced by two identical holes. It is clear that the diffraction spectrum of the two holes has the same intensity distribution as the spectrum produced by a single aperture, except that it

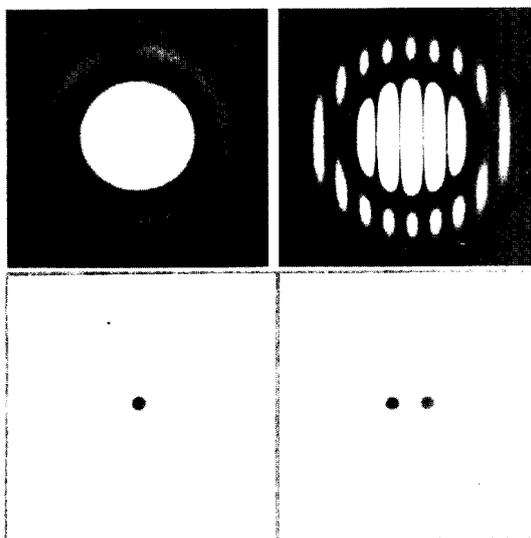


Fig 3. The effect of repeating the object on the diffraction spectrum. (From Harburn *et al*, 1975; reprinted with permission).

is crossed by dark, equally spaced, straight interference lines, perpendicular to the translation vector relating both circular apertures. The repetition of the diffracting object did not generate new bright zones; instead, it 'selected' some regions within the bright zones produced by the diffraction of a single object. Since the amplitudes of two identical waves are added in the bright zone, the intensity is four times that obtained with a single aperture. This is the classical Young experiment; it can be shown that the spacing between the dark interference lines is inversely related to the length of the vector relating both circular holes.

If the object presents a two-dimensional periodicity, the diffraction spectrum will be crossed by two sets of interference lines, with orientations and spacings depending on the direction and magnitudes of the translation vectors. A three-dimensional repetition of the object will produce a more strict selection of the remaining bright zones, reducing them to a periodic distribution of bright spots with different intensities.

Good examples of the effects of periodicity on the diffraction spectra can be found in the book by Haburn *et al* (1975).

Mathematically, a crystal can be considered as a convolution (or product) of a molecule (or group of molecules) and a lattice (the 'real' lattice) formed by the three vectors that describe the crystal periodicity. The diffraction spectrum of the crystal, or Fourier transform, is the convolution of the diffraction spectrum of the single molecule and the interference pattern of the real lattice. This interference pattern is also a lattice, called the 'reciprocal lattice'. Since a lattice is a function with non-zero values only at the lattice nodes, the product of the diffraction spectrum of the single molecule and the reciprocal lattice will have non-zero values only at the reciprocal lattice nodes.

The diffraction spectrum of a crystal is no longer a continuous function of three coordinates XYZ, but it will depend on three integers (hkl) that define the nodes of the reciprocal lattice:

$$F(hkl) = \sum_j f_j \exp [2 \pi i (hx_j + ky_j + lz_j)] \quad [7]$$

This formula allows a calculation of the diffraction spectrum of a molecule, provided the nature and position of all its j atoms are known. It is very useful to test the reliability of the structure determined: if correct, the observed structure factors must be similar to the calculated ones.

Figure 4 shows a plane of the diffraction spectrum of a protein crystal. As expected, it consists of discrete points very close to each other. Since the molecular dimensions require large repetition vectors, the reciprocal vectors are small.

THE ELECTRON DENSITY FUNCTION

Up to now we have considered a protein molecule as a collection of point atoms with all the electrons concentrated at points defining the location of these atoms. An alternative and more realistic way to define the structure of a protein is in terms of the electrons contained in the unit cell, assuming a continuous distribution of electrons. A continuous function called 'electron density' is defined and represented by $\rho(xyz)$, where $\rho(xyz) dx dy dz$ represents the number of electrons included in the volume $dx dy dz$ at the point xyz of the unit cell. As we have pointed out before, the electrons are responsible for the scattering of the X-rays,

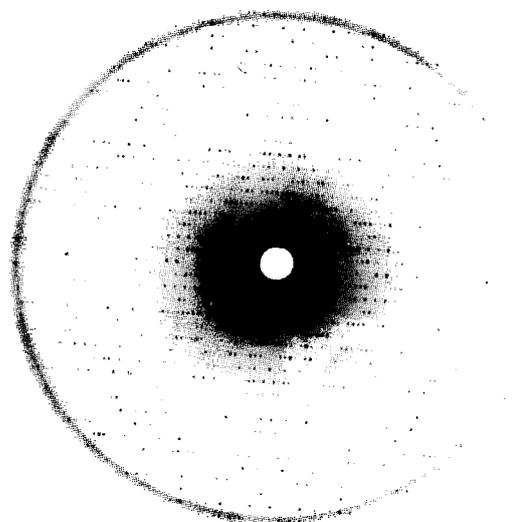


Fig 4. A typical diffraction spectrum from a protein crystal. Precession photograph of the (hk0) plane of the actin-profilin complex.

therefore the X-ray diffraction spectrum is a function of the electron distribution.

The mathematical expression [5] used to calculate the three-dimensional image, considering that hkl are integers and not continuous variables, becomes:

$$\rho(xyz) = (1/V) \sum_h \sum_k \sum_l F(hkl) \exp [-2 \pi i (hx + ky + lz)] \quad [8]$$

The structure factor $F(hkl)$ represents the diffracted wave at the point (hkl) of the diffraction spectrum. This wave has an amplitude and a phase, therefore can be represented as a vector or as a complex number. The amplitude $|F(hkl)|$ is the modulus of the complex quantity $F(hkl)$, and the phase can be represented as $\alpha(hkl)$; this can be written as:

$$F(hkl) = |F(hkl)| \exp [i\alpha(hkl)] \quad [9]$$

and the electron density function becomes:

$$\rho(xyz) = (1/V) \sum_h \sum_k \sum_l F(hkl) \exp [-2 \pi i (hx + ky + lz) + i\alpha(hkl)] \quad [10]$$

The amplitude $|F(hkl)|$ can be measured as the square root of the intensity of each diffracted spot (hkl) in a diffraction spectrum, such as that shown in Figure 4.

Determination of the phase angle $\alpha(hkl)$

This is the second key problem that protein crystallographers need to solve. Several solutions have been found. The first method, still widely used, is the Isomorphous Replacement Method, that determines the phase angle of each reflection by 'marking' the protein with heavy atoms, that do not disturb the packing of the native protein molecule (see for example Blundell & Johnson, 1976). It is based on the vector equation:

$$F_{PH} = F_P + F_H \quad [11]$$

where F_{PH} , F_P and F_H represent the structure factors of the protein-heavy atom complex, the native protein and the heavy atom alone, respectively. In this vector relation the amplitudes of the three vectors are known, and the phase angle of the heavy atom can be

determined; therefore the phases α_{PH} and α_P are the unknown variables and two different solutions are obtained from equation [11].

Theoretically, a minimum of two different heavy-atom derivatives need to be obtained and measured, in the same way as for the native protein, to attain a unique solution. In fact, every protein structure solved by this method utilizes four or more such derivatives, unless one of the heavy atoms used presents anomalous dispersion for the X-ray wavelength used, thus introducing new information. The solution of the crambin structure (Hendrikson & Teeter, 1981) and of a derivative of insulin (Stuart *et al*, 1984) are examples of structures solved by utilizing heavy atoms with anomalous dispersion properties.

When structural information of other proteins, structurally related to the unknown protein is available, the Molecular Replacement Method is preferred to calculate the phases (Rossmann, 1972). There are, at least, two different ways of utilizing the Molecular Replacement method; the first one requires that the asymmetric unit (the minimum motif needed to build the whole crystal, after all the translation and crystallographic symmetry operations have been applied to it) has more than one copy of the protein molecule, related by non-crystallographic symmetry elements. This is the case of the crystalline structure of several virus-coat proteins, and this procedure has played a fundamental role in their solution (See for example Unge *et al*, 1980; Bricogne, 1974, 1976).

The second case implies obtaining a molecular model of the unknown structure based on a similar protein whose structure is known. The known molecule must be correctly oriented and located in the unit cell of the unknown protein in the position that the unknown molecule should occupy. The phase determination would then consists of three stages:

i) The determination of the relative orientation of the model and the unknown protein molecules, that is the determination of the "Rotation Function" [C], defined as an orientation matrix.

ii) The second step is the correct location of the model molecule in the unit cell of the

unknown structure, which is equivalent to determine a "Translation Function" (Rossman, 1972). At the end of these two steps the exact relation between a point R' in the unknown structure and the equivalent point R in the model structure will be given by the equation:

$$R' = [C] R + D \quad [12]$$

where D is the translation vector determined in step ii).

iii) The third step consists in the use of the structural information of the model to calculate a set of phases for the unknown structure, by equations [7] and [9]. These phases, when combined with the amplitudes of the observed intensities of the unknown structure, will allow the calculation of the first electron density map of the unknown molecule.

The Molecular Modeling Method, described in the article by Cid and Bunster (1996), can also be used as a variation of the Molecular Replacement Method to calculate an initial set of phases, and utilize them to build the first electron density map of the unknown protein.

Only when the phase determination is successfully accomplished, can one be sure that the three-dimensional structure of the protein molecule can be solved.

Calculation of the electron density function

Formula [10] is used by the computer to calculate the 'image', in every point (xyz) of the unit cell, as requested by the programmer. Usually one divides the unit cell axes in 50 (or 100) points, and asks the computer to calculate $\rho(xyz)$ for all values of x from 0/50 to 50/50, and the same for the variables y and z . The computer calculates the value and plots a number in the corresponding point of the unit cell. For a three-dimensional electron density function, all calculations are performed in sections, with one coordinate held constant on each section.

With the current technology, the computer can not only plot the electron density function, but it can also fit to these maps a selected strand of the polypeptide chain, by

using superpositions, rotations and displacements. When the best fit is achieved, the computer can produce the 3-dimensional set of coordinates (xyz) for each one of the atoms included in that part of the polypeptide chain.

Figure 5 shows the calculated electron density function for a small planar molecule, 'hand plotted'; the curves joining points of equal electron density define the position of the atoms. Each atom is clearly resolved from its neighbors; the 'X-ray microscope' has reached a magnification power of $10^6 X$.

Other forms of the electron density function: Difference maps and double difference maps. The Patterson Function

Two modifications of the electron density function are used in the X-ray diffraction studies of complexes of enzymes with substrates or inhibitors. These are the 'difference electron density' and the 'double difference electron density' functions. These two functions differ from the electron density function only in the coefficients that sub-

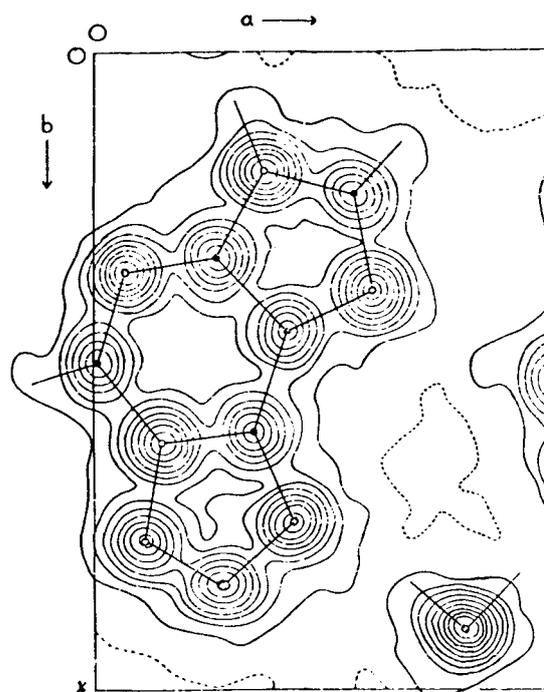


Fig 5. 'Hand drawn' electron density map of a small planar molecule. The height of the peaks, given by the number of contours, is proportional to the number of electrons of the atom defined by that peak.

stitute the absolute value of the structure factor in equation [10], as shown in Table I.

The meaning of these two functions is shown schematically in Figure 6 for the special case of a linear molecule. It is assumed that $\rho_1(x)$ represents the structure (or electron density) of the native 'enzyme' which has three atoms at locations x_1 , x_2 and x_3 ; $\rho_2(x)$ is the structure of the 'enzyme-inhibitor complex' with the inhibitor located at position x_4 . The difference electron density function shows only one peak at position x_4 , thus indicating the location of the inhibitor. The double difference function shows all the maxima corresponding to the native 'protein', plus a double weighted peak indicating the position of the inhibitor. The importance of the double difference map is that it displays the relative position of the inhibitor with respect to other residues involved in the active site.

The use of these two functions is restricted to the condition that there are no drastic changes in the structure or in the packing of the enzyme on binding of the ligand. The calculation is then based on the intensity differences obtained from the diffraction spectra of the complexed and native enzymes and on the phase angles obtained from the native enzyme.

Figure 7 shows electron density maps together with the corresponding difference and double difference maps for the enzyme carbonic anhydrase I (CA I), native and complexed with Diamox, a sulfonamide non-competitive inhibitor. Since in this case all the Fourier summations are three-dimensional functions, the maps have been calculated in sections. Figure 7 shows several sections drawn in transparent plates and stacked together.

The Patterson Function, defined in equations [13] and [14], is also a Fourier Transform, that, unlike the electron density function, does not require a phase knowledge to be calculated (Table I). However, the interpretation of the results of the Patterson Function is not straightforward: the positions of its peaks determine a set of interatomic vectors, not atomic positions, and the peak heights are proportional to the product of the number of electrons of the two atoms involved. As it is shown in Figure 6, for a

Table I

Different Fourier Transforms used in protein tertiary structure determination

Function	Coefficient	Calculation performed
Electron Density	$ F(hkl) \exp i\alpha(hkl)$	$\rho(xyz)$
Difference Electron Density Function	$(F_2(hkl) - F_1(hkl)) \exp i\alpha_1(hkl)$	$\rho_2(xyz) - \rho_1(xyz)$
Double Difference Electron Density	$(2 F_2(hkl) - F_1(hkl)) \exp i\alpha_1(hkl)$	$2\rho_2(xyz) - \rho_1(xyz)$
Patterson Function	$ F(hkl) ^2$	$\rho(xyz) * \rho(x+u, y+v, z+w)$
Difference Patterson Function	$(F_2(hkl) ^2 - F_1(hkl) ^2)$	$P_2(uvw) - P_1(uvw)$

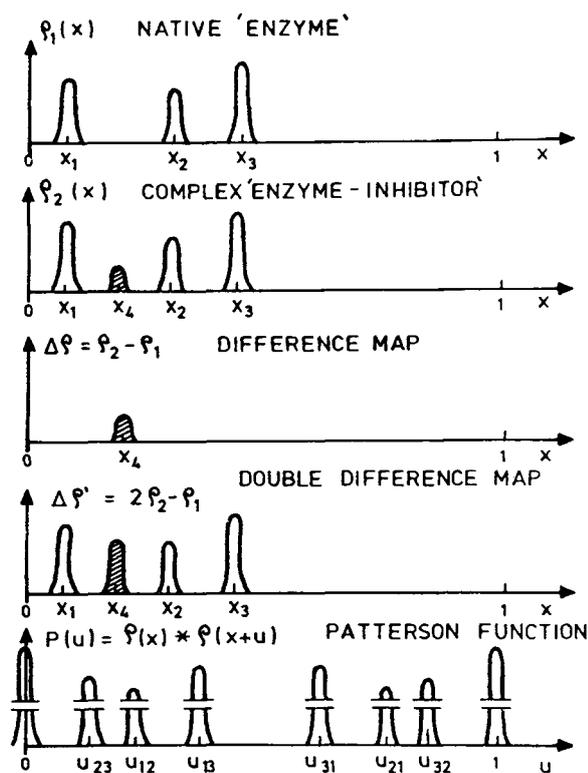


Fig 6. Electron density, difference and double difference electron density maps, illustrated for a one-dimensional 'molecule'. The Patterson function $P(u)$, corresponding to the electron density $\rho_1(x)$ is also shown.

linear structure with 'n' atoms, the Patterson Function will show n^2 peaks, n of them corresponding to vectors relating each atom with itself, and therefore will be located at the origin of the unit cell.

$$P(uvw) = \rho(xyz) * \rho(x+u, y+v, z+w) \quad [13]$$

$$P(uvw) = (1/V) \sum \sum \sum |F(hkl)|^2 \exp [2\pi i (hu + kv + lw)] \quad [14]$$

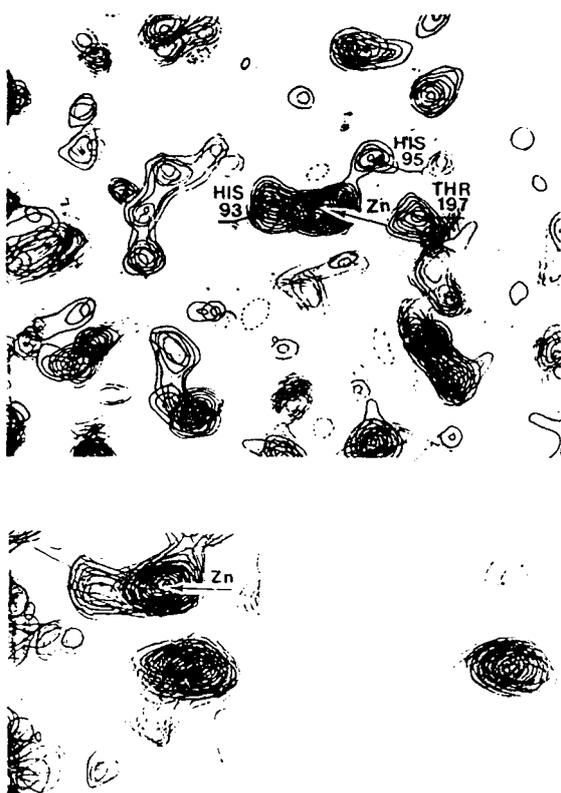


Fig 7. (a) Electron density map of the active site region in native HCAI. (b) Double difference map for the complex HCAI-Diamox ; note that the shape of the maxima including the Zn ion is the same as in (a). (c) Difference map between the complex HCAI-Diamox and the native enzyme. Only the peaks corresponding to the Diamox inhibitor are present. (From Nostrand, 1974; reprinted with permission).

The interpretation of the Patterson Function for a protein with its thousands of atoms is an impossible task. However, the Difference Patterson Function is fundamental in the process of phase determination, since it allows the location of the heavy atoms used to 'mark' the protein in the Isomorphous Replacement method. The difference between the number of electrons of the 'heavy' atoms currently used (Iodine, Gold, Mercury...) compared to the carbon, oxygen or nitrogen found in a protein, makes easy to identify the peaks that correspond to interatomic vectors between heavy atoms, and their location within the unit cell. Also, Patterson Functions are used in the definition of both the Rotation and the Translation Functions in the Molecular Replacement Method.

ACTIVE SITE AND CATALYTIC MECHANISM

The most remarkable characteristic of enzyme catalysis is that the enzymes bind their substrates at the active site in a unique orientation that allows the desired proximity of all the necessary elements that take part in the reaction. To understand the mechanism of an enzyme it is necessary to know the structure of the native enzyme, and that of the complexes of the enzyme with its substrates, intermediates and products as well (Fersht, 1985). As pointed out before, usually the mean life of enzyme-substrate complexes can be of the order of seconds or less, and the recording of the X-ray data takes several hours.

To overcome these difficulties several strategies are used. Some of them are:

1. The determination of the structure of an enzyme-inhibitor, enzyme-product or enzyme-substrate analog complex.
2. The determination of the structure of the enzyme-substrate complex under unfavorable reaction conditions, such as very low temperature, a pH where the substrate is in the wrong ionic state, or the use of very poor substrates.
3. The determination of the structure of the native enzyme and, separately, of the substrate and then, by model building, find the correct location of the substrate in the active site.
4. The examination of a productive enzyme-substrate complex under conditions of rapid reaction, where it is possible to set an equilibrium between substrates and products, with an equilibrium constant favoring the substrate by one order of magnitude or more. In preparing the complexes, use is made of the fact that protein crystals contain about 50% (and never less than 30%) solvent. It is thus possible to soak the native enzyme crystals in a solution containing the substrate, inhibitor or product to be complexed with the enzyme. The ligand will then diffuse into the active site. This method is easier than co-crystallizing the enzyme and the substrate.
5. Use of Site Directed Mutagenesis, in connection with X-ray diffraction methods. It

is possible to test the role of certain amino acid residues in the catalytic process if we have the possibility to change them by others. This can be achieved when the gene of the protein has been cloned and it has been expressed in an appropriate vector. The stability of DNA duplexes containing mismatches and the ability of bacterial DNA polymerases to extend primers hybridized to single stranded DNA templates are also required.

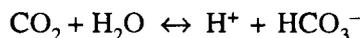
The use of this method would allow the production of mutant proteins with very specific changes in any amino acid residue 'suspicious' of being involved in the protein function, and in this way, different ideas about its role can be tested.

6. Use of Molecular Dynamics calculations to simulate the whole catalytic process, starting by substrate docking.

In the first edition of a book on this subject (Cid, 1987), four examples, illustrating the use of approaches 1 to 4 were presented. In the last years, Site Directed Mutagenesis and Molecular Dynamics calculations have proved to be very efficient, when applied by themselves or in combination with the more classical methods, to deduce the catalytic process. Therefore we will present, together with the original solutions, the contribution made by these new techniques. The basis of the Molecular Dynamics method is presented in the article by Cid and Bunster (1996), and a complete description of the Site-Directed Mutagenesis method is given by Moreno-Hagelsieb and Soberón (1996), also in this volume.

EXAMPLE 1.
A CATALYTIC MECHANISM FOR CARBONIC ANHYDRASE OBTAINED BY X-RAY DIFFRACTION STUDIES OF ENZYME-INHIBITOR COMPLEXES

Carbonic anhydrase is an enzyme that catalyses a fundamental reaction in the respiratory process:



Two varieties of this enzyme have been isolated from human red cells, and have been designated as HCA I and HCA II,

respectively. HCA II has a turnover number of 10^6 mole enz^{-1} sec^{-1} , and HCA I, which is 'slower', has a rate of 1.600.000 mole enz^{-1} sec^{-1} . Both enzymes show great similarities in primary structure, molecular weight, and number of amino acids. The three-dimensional structure of both enzymes were determined by X-ray diffraction methods to a nominal resolution of 2 Å almost twenty years ago. The similarity of these enzymes is maintained at the secondary and tertiary structure level so that it is possible to give a unique description for both of them (Kannan *et al.*, 1975, Nostrand *et al.*, 1975).

Owing to the very complete X-ray study of the complexes of HCA I and HCA II with inhibitors, this enzyme is a good example of the information that X-ray diffraction methods can provide on the catalytic mechanism of an enzyme.

The common secondary structure of carbonic anhydrases HCA I and HCA II is shown in Figure 8 in the 'cylinder and arrow' representation. The Zn^{+2} ion, a cofactor of the enzyme, is coordinated to three histidine residues. Both the location of the Zn ion and its distorted tetrahedral coordination to the histidines and to a solvent molecule, were determined from the electron density maps. The Zn peak, with its 30 electrons, gave by far the highest maximum density in the map. The non-histidine fourth ligand is a water molecule or an OH^- .

Once the tertiary structure of these enzymes was solved, it was possible to locate

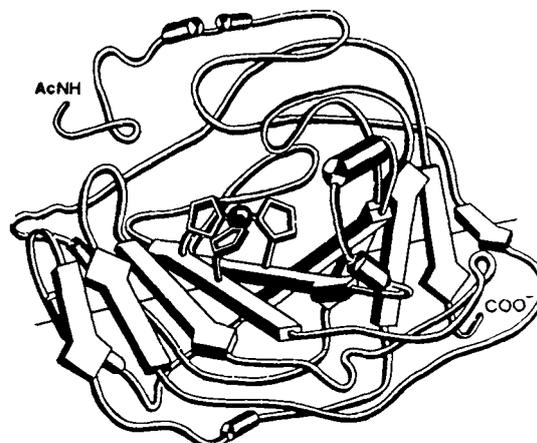


Fig 8. Secondary structure of HCA I and HCA II in the cylinder-arrow representation. The Zn atom is shown attached to the three histidyl residues. (From Nostrand, 1974; reprinted with permission).

the active site and to determine the role of the metal. As part of the active site studies, the enzyme was prepared without the Zn ion (apoenzyme), and with the Zn²⁺ replaced by Mn²⁺, Co²⁺, Cu²⁺ or Mg²⁺ (Lövgren *et al*, 1971). In each case the metal position was obtained from difference Fourier maps between the complex and the apoenzyme. These studies showed that although all these metal ions occupy the same location and bind to the same ligands as Zn²⁺, only the enzyme-Co complex retains catalytic activity (about 50% of the activity of the native Zn-enzyme).

Monovalent anions (I⁻, Br⁻, Cl⁻), as well as sulfonamides, are non-competitive inhibitors of the reaction of both enzymes. Complexes of HCAB and HCAC have been made with each of the three ions and with several sulfonamides such as 3-acetoxymethyl-4-aminobenzene sulfonamide (AMSulf), acetazolamide, Diamox and salamide (Bergstén *et al*, 1972). All of these complexes were studied by X-ray diffraction methods, and the position of the inhibitors were located by difference and double difference Fourier maps, such as those shown in Figure 7.

The results of these studies have clearly shown that:

1. The monovalent anions bind to the Zn ion by replacing the fourth ligand. The electron density peak corresponding to

the anion was found in the location occupied by water, or OH⁻ in the native enzyme. Figure 9 is a computer drawing of the I⁻ coordination at the active site of HCAC.

2. All sulfonamide complexes bind to the Zn ion through the nitrogen of the sulfonamid group. The nitrogen atom also replaces the fourth Zn ligand. Some of the amino acid residues present in the active site also bind to the sulfonamide inhibitors as seen in Figures 10 and 11.

The only known competitive inhibitor for the hydration reaction of CO₂ catalyzed by carbonic anhydrases is imidazole. It acts only on HCA I. Even though this inhibitor is a small molecule, its shape makes possible its location within the active site. Electron density maps of the enzyme-imidazole complex, and difference Fourier maps between the complex and the native enzyme, showed that this inhibitor binds to the Zn ion without disturbing the fourth ligand. It is attached to Zn at a fifth coordination position, with the nitrogen of the ring 2.7 Å from the Zn ion (Kannan *et al*, 1977).

The above observation suggests that one of the oxygens of the substrate CO₂ can also occupy this fifth ligand position, thus leaving the substrate in a highly favorable orientation for a reaction with the fourth Zn ligand, water.

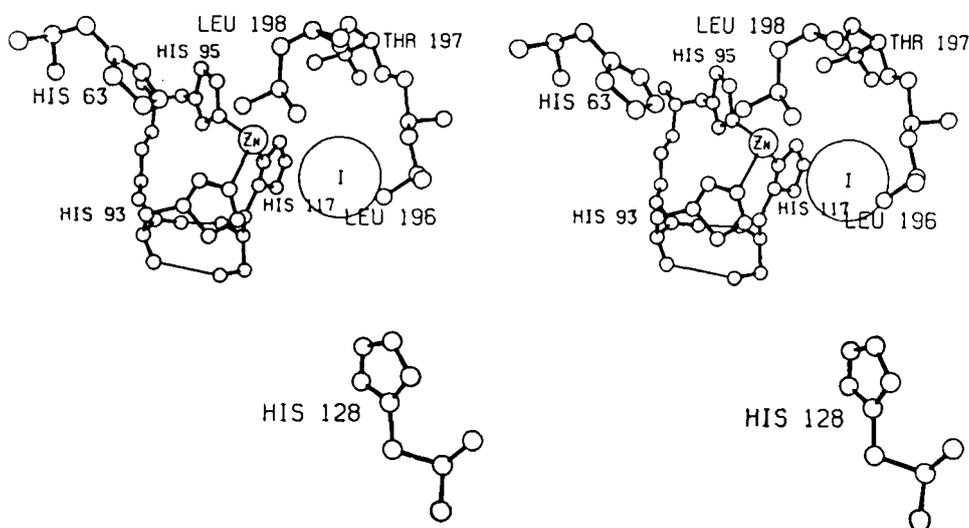


Fig 9. Location of I⁻ in the active site of HCAII. Stereoscopic (ORTEP) drawing. (From Bergstén *et al*, 1972; reprinted with permission).

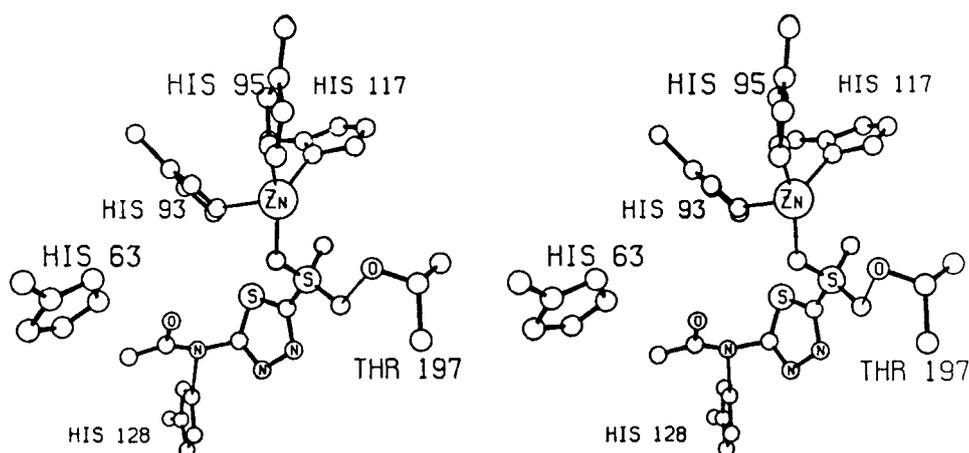


Fig 10. ORTEP stereoscopic drawing of acetazolamide binding to the active site of HCAII (From Bergstén *et al*, 1972; reprinted with permission).

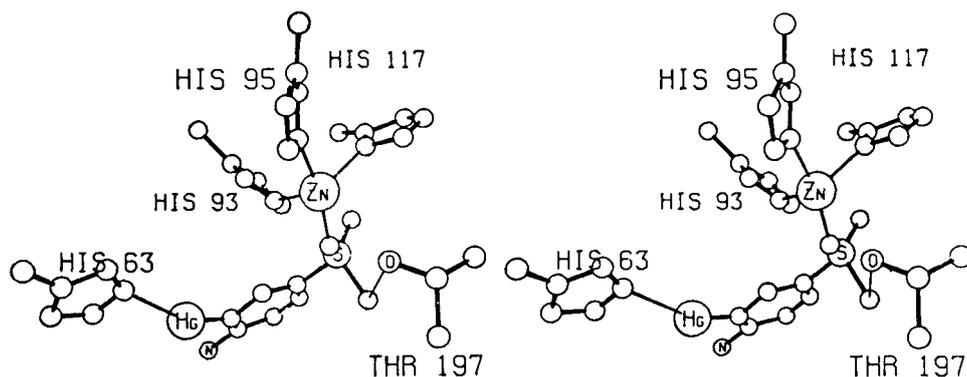


Fig 11. ORTEP stereoscopic drawing of the AMSulf molecule binding to the active site of HCAII (From Bergstén *et al*, 1972; reprinted with permission).

Once the existence of a fifth coordination position for the Zn had been established, the rest of the experimental results began to fit. This explains, for example, why the Co^{+2} -enzyme retains activity, since Co is the only other divalent cation with 5-fold coordination.

It was also shown that, although the N atom in sulfonamide replaces the fourth ligand, one of the oxygens of the sulfonamide group occupies the fifth coordination position of the Zn ion. This fact agrees with reports that sulfonamides act as competitive inhibitors of the hydration reaction of CO_2 by HCAB (Lindskog *et al*, 1971).

The catalytic mechanism for carbonic anhydrases proposed in 1977 is shown in Figures 12 and 13. Two proposals were made, since the data available at that time were not sufficient to decide if the fourth Zn ligand was a water molecule or an OH^- ion. The amino acid residues Glu 106 and Thr

199 participate in both mechanisms. Both residues are conserved in all primate and ruminant carbonic anhydrases.

New data on the structure and function of HCA II based on X-ray studies and site-directed mutagenesis

A careful refinement of the crystal structure of HCA II (Eriksson *et al*, 1988a, Håkansson *et al*, 1992), has been followed by a revision of the structures of several enzyme-inhibitors complexes, some of them already published, unrefined, in the seventies, like those with sulfonamides and anions (Eriksson *et al*, 1988b, Xue *et al*, 1993a, b; Lindahl *et al*, 1993, Liljas *et al*, 1994). The refined X-ray data confirmed the overall correctness of the structure previously determined, allowed the location of 167 solvent molecules in the electron density maps, and allowed to establish the

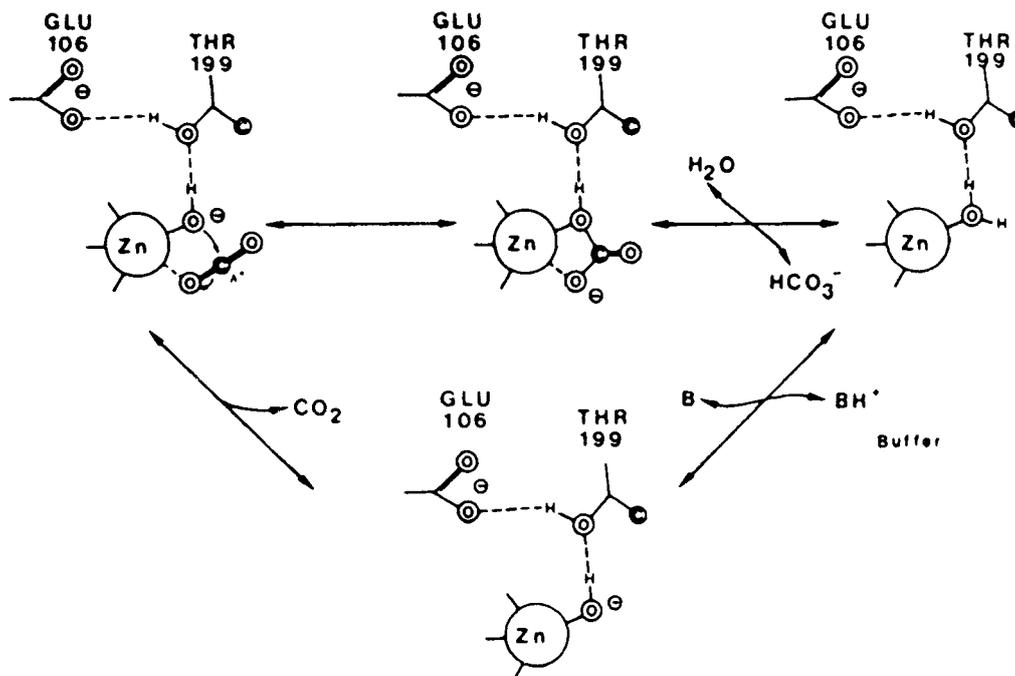


Fig 12. Proposed catalytic mechanism for HCAI, assuming that the fourth Zn ligand is OH (From Kannan *et al.*, 1977; reprinted with permission).

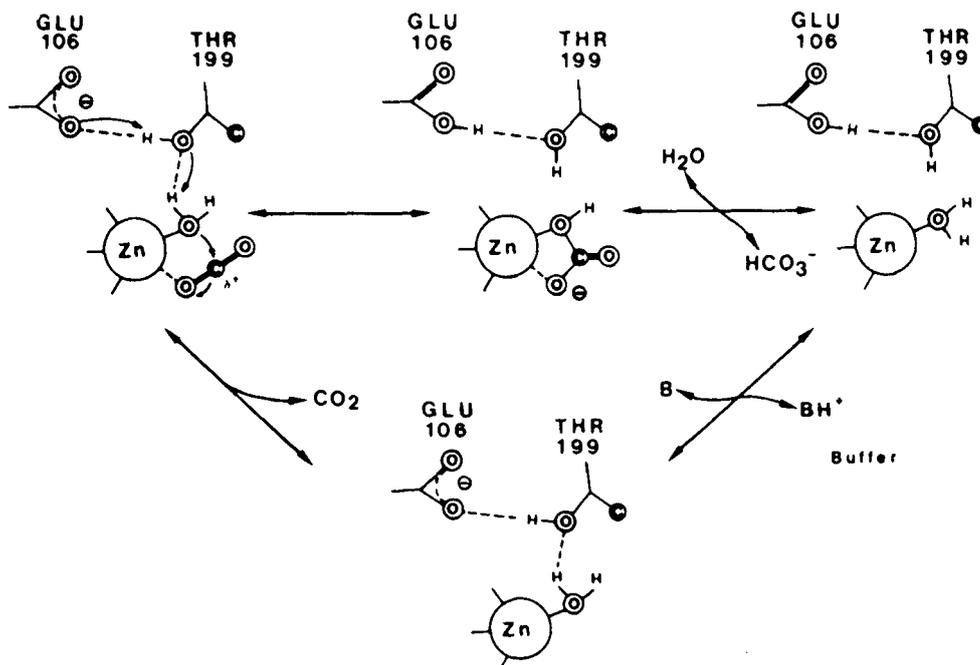


Fig 13. Proposed catalytic mechanism for HCAI, assuming that the fourth Zn ligand is a water molecule (From Kannan *et al.*, 1977; reprinted with permission).

exact hydrogen-bond net within the active site. Two, out of the nine water molecules found within the active site, OHH 63 (the 'Zinc water') and OHH 338 (the 'deep water'), participate of this net and play a fun-

damental role in the catalytic process of the enzyme.

Several site-directed mutagenesis have been made to elucidate the role of specific amino acids. His64 was tested as a proton

transfer amino acid; Glu106 and Thr199 as the 'door keeper' residues, Val143 and Leu198 to examine the role of the hydrophobic pocket residues in the catalytic mechanism, and Thr200 to explain the difference in efficiency between HCAI and HCAII. X-ray diffraction studies were also made for each one of these mutant to detect structural changes. Some of the results obtained are summarized below and on Table II.

The active site of HCAII (as well as that of HCAI) is clearly divided into two regions: one hydrophilic and the other hydrophobic (Fig 14). At the bottom of the hydrophobic pocket the 'deep water' molecule preserves the CO₂ binding site. This site was deduced by analogy with the binding of the isoelectronic NCO⁻ inhibitor, and coincides with the results from molecular dynamics calculations which indicate that the main CO₂ binding site is the hydrophobic pocket where the deep water is located (Liang & Lipscomb, 1990). Amino acid substitutions at Val143, located at the center of the hydrophobic pocket, and Leu198, at its entrance, showed that the enzyme catalytic behavior was more sensible to the presence of bulky

Table II
Kinetic parameters for the catalysis of CO₂ hydration by wild-type and mutant HCA II, at 25°C

Enzyme	k_{cat} (s ⁻¹)	k_{cat}/K_M (μM ⁻¹ s ⁻¹)	Reference
Wild-type	1000	100	Engstrand <i>et al.</i> , 1992
H64A	186	92	Tu & Silverman, 1989
E106Q	1.1	10	Liang <i>et al.</i> , 1993
V143G	742	43	Fierke <i>et al.</i> , 1991
V143C	1075	83	<i>Id.</i>
V143L	667	17	<i>Id.</i>
V143I	1075	12	<i>Id.</i>
V143N	1398	14.6	<i>Id.</i>
V143H	36.6	0.9	<i>Id.</i>
V143F	>8.6	0.04	<i>Id.</i>
V143Y	>0.03	0.003	<i>Id.</i>
L198A	400	37.3	Krebs <i>et al.</i> , 1993
L198H	120	7.9	<i>Id.</i>
L198E	300	13	<i>Id.</i>
T199A	10	1.5	Liang <i>et al.</i> , 1993
T200H	270	44.4	Behravan <i>et al.</i> , 1990

amino acid residues (Phe, His, Tyr) than to their hydrophobic character, probably because the blocking of the pocket does not allow the substrate to approach the Zn with the correct orientation to react (Table II). The hydrogen-bond net of the active site includes the Zinc water bonded to Ogl 1 of

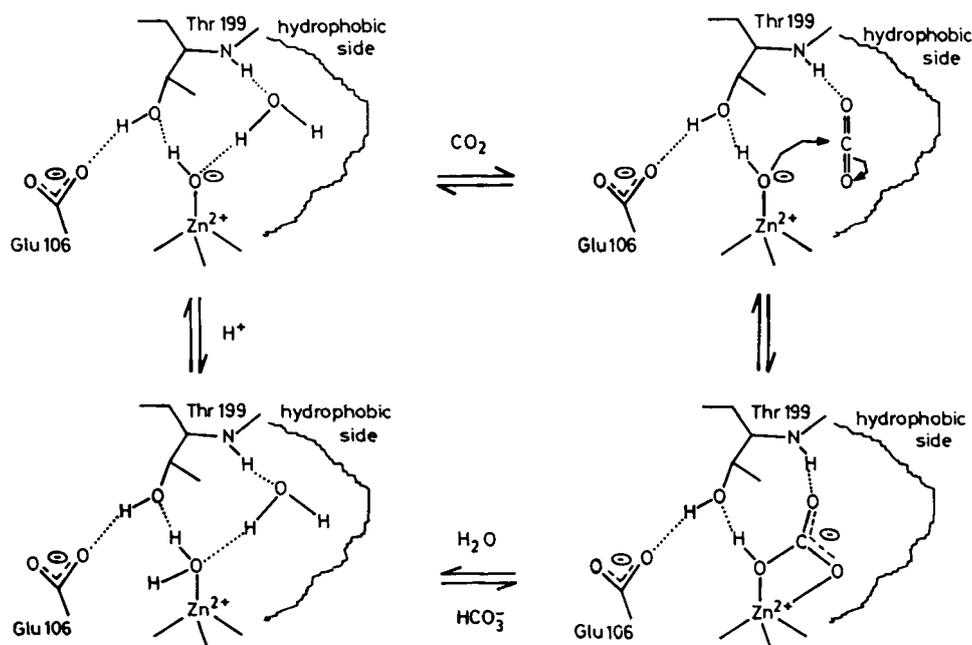


Fig 14. Proposed catalytic mechanism for carbonic anhydrases. The interconversion of CO₂ and HCO₃⁻ is emphasized; the details of the reactions involved in dissociation and proton transfer between the Zn-bound water molecule and the reaction medium has been left out. The enzyme-bicarbonate complex has been drawn as observed in the crystal structure of the mutant T200H of CAII. (From Xue *et al.*, 1993). ©Wiley-Liss Inc.1993. Reprinted with permission from John Wiley & Sons, Inc.

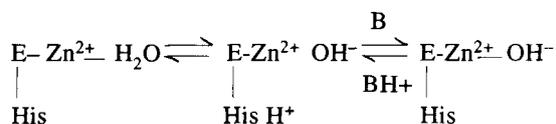
Thr199, Glu106 Oe, also bonded to Thr 199, and the 'deep water' bonded to the nitrogen of Thr199 and to the Zn water. The role of the 'door keeper' residues Glu106 and Thr199 is to control that only protonated ligands, that can *donate* a hydrogen bond to Thr199 hydroxyl group, bind at the tetrahedral Zn position, replacing the Zn water. The mutation of Glu106 by a Gln (E106Q) made possible that unprotonated ligands could bind to the Zinc water position, since the hydrogen of Thr199 Ogl becomes accessible in the active site.

The mutation T199A drastically reduces the enzyme activity, due to the fact that Glu106 exposes its negative charge at the active site; a water molecule 'replaces' the removed Ogl of Thr199.

All these results indicate that the function of the 'door keepers' is not only to orient the HCO_3^- correctly for the dehydration reaction, but also the hydroxide orbitals in the hydration of CO_2 . The T199A mutant showed an increased affinity by HCO_3^- in penta coordination to the metal ion (Liljas *et al.*, 1994). Proton transfer from the Zn water to the buffer is efficiently performed by His 64 in the HCAII enzyme. However, Lys64, Glu64, and imidazole can also perform this role, but less efficiently, as it has been shown by site directed mutagenesis experiments. The role of imidazole was tested by measuring the activity of the H64A mutant adding imidazole to the buffer (Engstrand *et al.*, 1992).

The mutation T200H, resulted in a higher affinity of the enzyme for the HCO_3^- ion, thus reducing the enzyme activity to values similar to those of HCAI (Xue *et al.*, 1993a). A proposed catalytic mechanism which accounts for all these results is schematized on Figures 9 and 14, and in it the Zn-bound OH^- reacts with CO_2 to form HCO_3^- which is then displaced from the metal ion by a water molecule. The OH^- regeneration is a rate-limiting process, and, as schematized below, it is performed via a proton-transfer process involving His64 as the proton shuttle between the Zn ion and buffer molecules (B/BH+) for HCAI and HCAII. The comparison of Figures 12 and 14 indicate that the catalytic mechanism of HCAII is essentially coincident with that proposed for HCAI, except

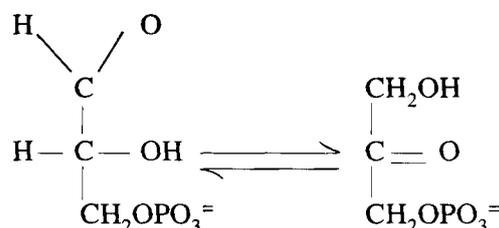
that the initial binding of the substrate replacing the 'deep water' molecule bound to the nitrogen of Thr199 was not considered.



EXAMPLE 2.

X-RAY DIFFRACTION STUDIES OF PRODUCTIVE ENZYME-SUBSTRATE COMPLEXES UNDER APPROPRIATE EQUILIBRIUM CONDITIONS. ACTIVE SITE STUDIES OF TRIOSE PHOSPHATE ISOMERASE

Triose phosphate isomerase (TIM), a dimeric enzyme with 247 amino acid residues in each of the identical chains, catalyses the following reaction:



The crystal structure of the enzyme obtained from chicken muscle has been solved to a resolution of 2.5 Å (Banner *et al.*, 1975). The polypeptide chain folds in a regular manner with eight strands of a β -pleated sheet forming an inner barrel surrounded by eight helices (Fig 15). This enzyme is one of the few that affords the opportunity of a direct observation of a productive enzyme-substrate complex. The reaction catalyzed has an equilibrium constant favoring the dihydroxyacetone phosphate (DHAP) with a value of about 20 (Fersht, 1985). The complex enzyme-DHAP was obtained by diffusing the substrate into the enzyme crystals. Difference Fourier methods allowed the location of the substrate in the enzyme active site even at a resolution of 6 Å (Banner *et al.*, 1971). It is shown that the carboxylate of Glu165 is equidistant from C_3 and C_4 of the substrate. The imidazole ring of His 195 is also equidistant from the carbonyl and hydroxyl oxygens. These findings agree with results obtained from solution studies (Rose, 1975) indicating that the reaction occurs via

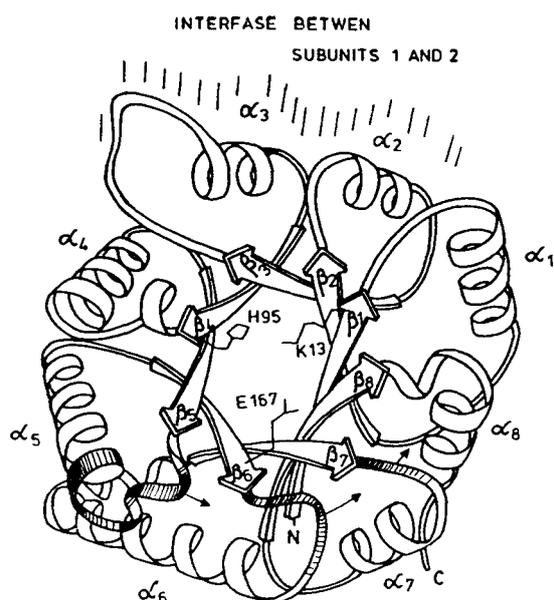


Fig 15. Secondary structure representation of the three-dimensional structure of triose phosphate isomerase displaying the characteristic β -stranded fl-barrel surrounded by 8 helices. The interface between both subunits is indicated, and loops 5, 6 and 7, involved in conformational changes, are shaded (loop 'j' is located between β_5 and β_6). (Redrawn from Richardson, 1985).

a cis-enediol intermediate (Fig 16), with a proton transferred by a single base: Glu165 (Glu167 in the Trypanosomal enzyme). The proposed active site lies at the carboxyl end of the β -barrel, and it is formed by residues from β -strands β_1 , β_5 , β_6 , β_7 , β_8 and from helices α_4 , α_5 and α_8 , with participation of a few residues of the adjacent subunit (Fig 15).

New X-ray data, combined with molecular dynamics calculations

In addition to the chicken enzyme, the refined structures of yeast TIM, to 1.9Å resolution (Lolis *et al*, 1990) and of *Trypanosoma brucei* TIM, to 1.83Å resolution (Wierenga *et al*, 1991a), are now known. Comparative studies among these three enzymes have shown that their functional properties are very well conserved, even though the conservation in the sequences is not better than 53%, when two of them are compared. The superposition of the α -carbons of the three-dimensional structures of these three enzymes give r.m.s differences between 0.7-0.9Å (Wierenga & Noble,

1992). In addition to the high resolution data, X-ray studies of complexes of the enzyme and catalytic relevant ligands, such as substrate or transition state analogs (Lolis & Petsko, 1990; Noble *et al*, 1991), have been performed. All these new data have allowed to answer several questions that the structures of these enzymes had raised, such as:

Why only the dimer formed by two identical subunits is active? It has been shown that the contact surface between both subunits (Fig 15) is formed by 32 residues, belonging to loops 1, 2 and 3. Of these, at least 3 residues (Cys14, Val 46, and Thr75), that are fully exposed in the monomer, are completely buried in the dimer. Thr75 of subunit 1, located at the tip of loop 3, penetrates into a pocket close to the active site of subunit 2, and its hydroxyl group participates in a hydrogen network important for the positioning of the side chains of some catalytic residues of subunit 2: The OH of Thr75 (subunit 1) makes an H-bond to ND₂ of Asn311 and to OE₁ of Glu397 which is hydrogen bonded to N_Z of Lys 313. (Glu397 and Lys313 are the active site residues of subunit 2, equivalent to Lys 13 and Glu97 of subunit 1). This network involving Lys13 (Lys12 in the yeast enzyme), may explain why only the dimer is competent: each subunit provides a "wall" upon which the other active site rests (Lolis *et al*, 1990).

The behavior of loop 6, the 'flexible loop'. This 14-amino acid region undergoes an important conformational change, which involves atomic displacements as large as 7Å, by 'closing' the active site upon substrate fixation. Loops 5 and 7 also experiment conformational changes associated with those of loop 6. The rigid-body movement of loop 6 had been observed for both subunits in the yeast enzyme, when the substrate binding site was occupied by ammonium sulfate (in the crystallization process), or by the substrate or intermediate analogs in the chicken and yeast enzymes. In the Trypanosomal enzyme, the rotation of loop 6 was observed only in subunit 2, since for subunit 1 is inhibited by close crystal contacts, not found in the other enzymes (Wierenga *et al*, 1991a)

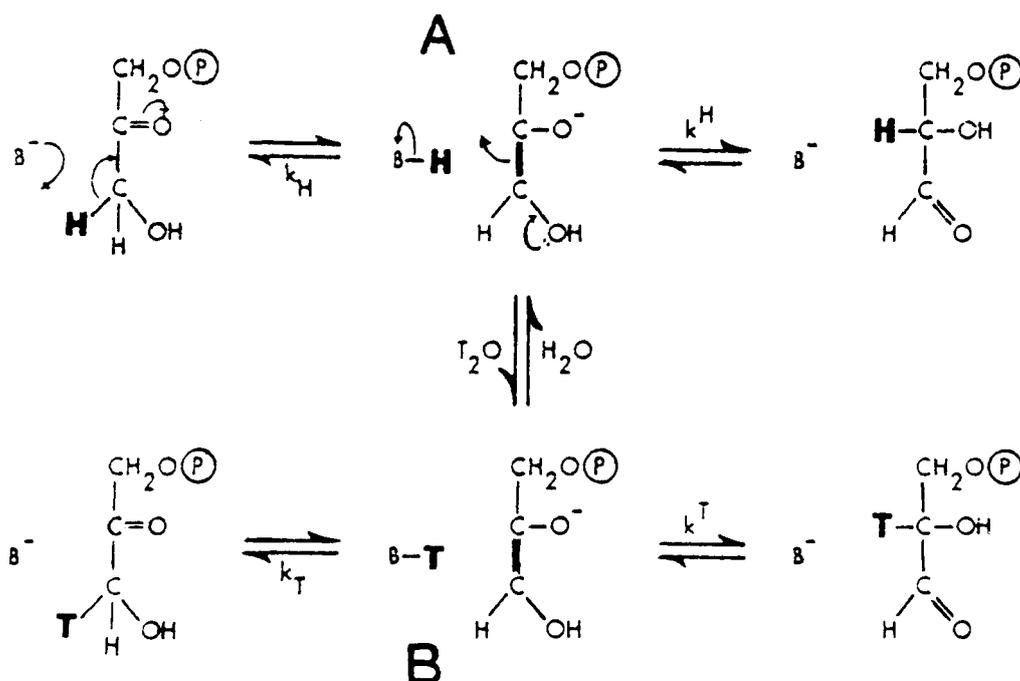


Fig 16. The enediol mechanism proposed by Rose for the reaction catalyzed by triosephosphate isomerase. A base on the enzyme (B) removes the α -proton to give an enzyme-bound enediolate. The intermediate picks up a proton to form the product. The lower half of the figure illustrates the experiment made with tritiated water that led to the proposed mechanism. (From Knowles *et al*, 1972; reprinted with permission).

Molecular dynamics calculations performed for the Trypanosomal enzyme, starting with the refined coordinates of a structure obtained with a crystal transferred to a sulfate-free mother liquid, were able to show the movement of the loop from a closed to an open conformation, involving atomic displacements as large as 7\AA . This process began with 8 steps, 50 cycles each, of energy minimization, followed by 15 steps of molecular dynamics minimization at a constant simulated temperature of 300°K , followed by 5 more steps of energy minimization. At this stage the temperature was raised to 1000°K , and then was slowly reduced to 100°K , by increasing four times the temperature relaxation time of 0.1psec used up to that point. The molecular dynamics process was considered complete, and a final step of potential energy minimization gave an R factor of 20.3%. It was then possible to trace the 'open' position of a loop of 14 amino acid residues which had been impossible to locate in the electron density function, without disturbing the generally correct conformation of the rest of the protein (Wierenga *et al*, 1991b).

The first residue of loop 6 in the trypanosomal TIM is the catalytic residue Glu167; atomic movements as large as 2\AA have been detected for it upon inhibitors' binding. The role that the movement of loop 6 plays in the catalytic process has not been fully established yet. However, it has been suggested that the "closure" of loop 6 may play a role in preventing phosphate elimination after the formation of intermediate in the catalytic reaction. The elimination of phosphate might be hindered by removing the electrostatic shielding powers of the solvent in the active site, thus making the formation of a negative charge necessary for this reaction, unfavorable. In addition, the interaction of the loop in the "closed" position with the phosphate group will prevent it from adopting a conformation favoring phosphate elimination (Lolis & Petsko, 1990).

EXAMPLE 3. A COMBINATION OF X-RAY DIFFRACTION METHODS AND MODEL BUILDING IN THE ACTIVE SITE STUDIES OF LYSOZYME

Lysozyme catalyses the hydrolysis of a polysaccharide that is a major constituent of

the cell wall of some bacteria. The substrate is formed by alternating units of N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) (Fig 17). The three-dimensional structure of this enzyme was totally solved by 1965 (Blake *et al*, 1965) and it has been further refined to a resolution of 1.5 Å, becoming one of the protein structures known with higher detail (Artymiuk & Blake, 1981).

The enzyme has a molecular weight of 14600 with 120 amino acid residues in a single polypeptide chain and four disulfide bridges. Lysozyme cleaves the polysaccharide chain in a specific position, between a C₁ of NAM and the oxygen attached to the C₄ of NAG (Fig 17).

The enzyme has a cleft on one side. This cleft presents regions with amino acid residues which facilitate the binding of the non-polar regions of the substrate and also hydrogen bonding sites for the acylamino and hydroxyl groups. The crevice is divided into six sites designated as ABCDEF. NAM residues can bind to B, D, and F only, while NAG residues can occupy any site. The bond to be cleaved must lie between sites D and E. Figure 18 illustrates the location of several saccharide inhibitors of lysozyme, as obtained from low-resolution difference electron density maps.

The structure of the enzyme complexed with the trimer NAG-NAG-NAG has also been solved (Phillips, 1967). This trimer behaves as a stable inhibitor, whereas longer NAG polymers are substrates, the velocity of cleavage being a direct function of the number of sugar rings up to hexasaccharides. Figure 19 shows the α -carbon representation of the structure of lysozyme with the trisaccharide (NAG)₃, located at the active site.

(NAG)₃, proved to be a very poor substrate. Fourier maps showed that it binds to sites ABC and avoids the sites where the cleavage takes place. The structure of a productive substrate-enzyme complex was obtained by building a wire model of the complex enzyme-(NAG)₃, and extending the polysaccharide chain by adding more NAG units. Chemical intuition was used to select appropriate contacts between the enzyme's three-dimensional model and the polysac-

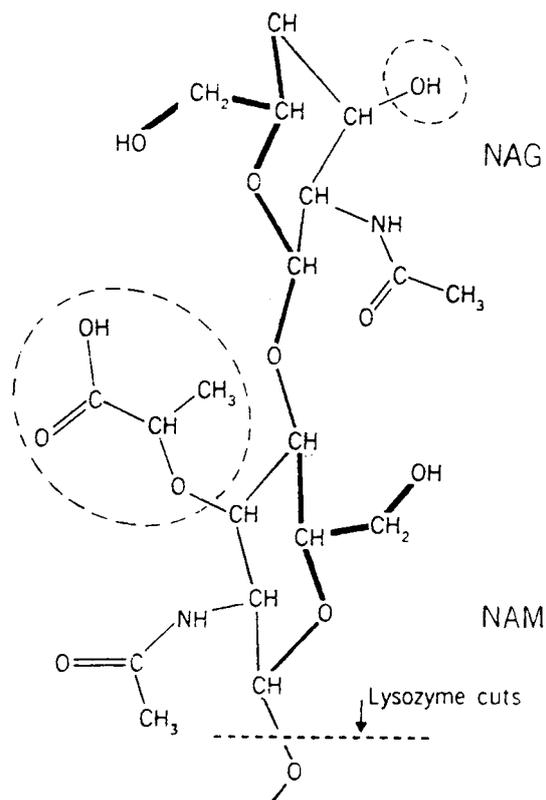


Fig 17. One unit of the (NAG-NAM)_n substrate of lysozyme, showing the cleaved bond. (From Dickerson & Geis, 1969; reprinted with permission).

charide chain. These studies determined that the bond to be cleaved is placed between the carboxyl groups of Glu35 and Asp52 (Phillips, 1967).

Model building is currently done using a computer program to optimize the fitting between the enzyme and the substrate. Interactive graphic systems display the atomic coordinates of the enzyme and of the substrate in a television screen and the program allows selective rotations or displacements in order to obtain the best fit. This procedure has been used in a combination of X-ray studies and model building of the complex of lysozyme and the trisaccharide NAM-NAG-NAM (Fig 20a). The crystals of the enzyme and the trisaccharide were grown by cocrystallization (Kelly *et al*, 1979). The coordinates of the trisaccharide were obtained by docking the atomic coordinates of two crystal structures, β -(1-4)-N,N'-diacetylchitobiose (monohydrate) and N- α -D-muramic acid. The fitting of the substrate

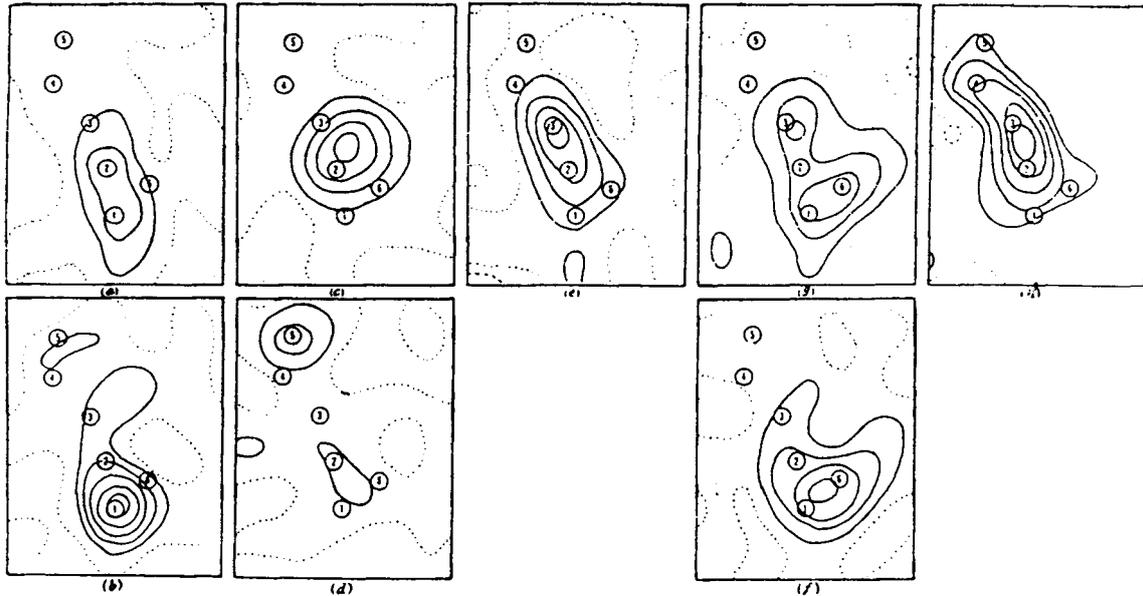


Fig 18. The section $y = 39/120$ in difference electron density maps at 6-Å resolution calculated with several lysozyme-inhibitor complexes and with the native enzyme. These show the binding of (a) NAG; (b) NAM; (c) 6-iodo- α -methyl-NAG; (d) α -benzyl-NAM; (e) and (g) di-NAG; (f) NAG-NAM; (h) $(\text{NAG})_3$ to the enzyme. Apparent binding sites of amino-sugar residues are marked 1 to 6 (2-Å resolution studies showed later that sites 4, 3 and 2 correspond to binding sites A, B and C, respectively). (From Blake *et al*, 1967; reprinted with permission).

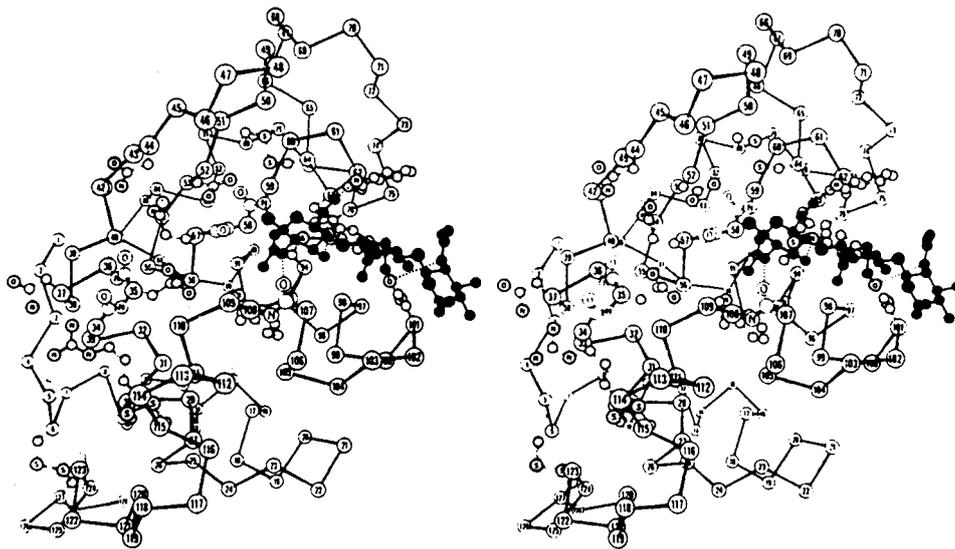


Fig 19. Stereo view of the α -carbon representation of the tertiary structure of lysozyme with the inhibitor $(\text{NAG})_3$ positioned at the active site. (From Dickerson & Geis, 1969; reprinted with permission).

in the difference electron density maps was made with an interactive graphics system and it is shown in Figure 20b. A list of close contact distances (less than 5 Å between the trisaccharide and the native enzyme) clearly show the participation of Asn103 and Asp101 in the B site; Ala107, Trp108, Gln57, Trp63

and Asn59 in the C site, and Asp52, Val109, Asn46, and Glu35 in the D site.

A different approach to the model building process has been presented by Pincus *et al* (1976, 1977). Successive NAG residues were added to the rigid active site of lysozyme, and the dihedral angles ϕ and φ , together

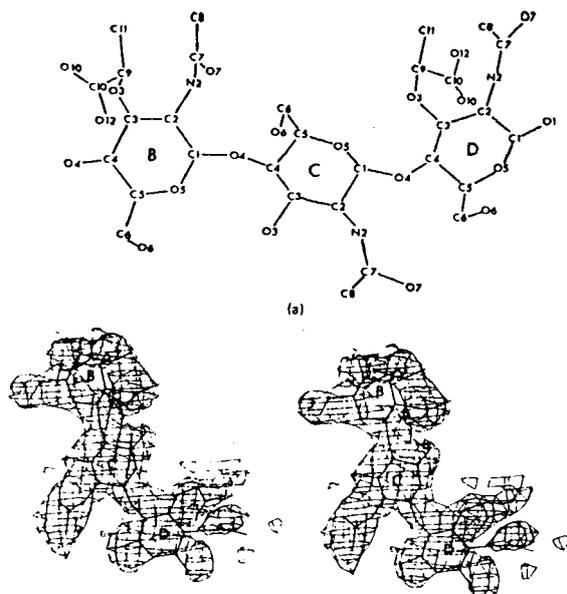


Fig 20. (a) The structure of NAM-NAG-NAM. The B, C and D labels designate the binding sites occupied by the sugars in lysozyme. (b) Stereo view of the difference electron density function of the enzyme-(NAM-NAG-NAM) complex and the native enzyme, superimposed with the fitted model of the substrate. This drawing was produced with an MMS-X interactive graphic system. (From Kelly *et al*, 1979; reprinted with permission).

with six rigid body variables that describe the orientation of the oligosaccharide substrates were allowed to vary to give unique low energy structures. This method was used with the complexes lysozyme-(NAG)₄, lysozyme-(NAG)₅ and lysozyme-(NAG)₆. Even though the results were promising, it was clear to the authors that the restriction of a rigid active site can introduce important distortions in the calculated energies. In any case, the model building performed after the energy minimization process allowed the removal of several unfavorable contacts between the polysaccharide and the enzyme (Fig 21).

Use of site-directed mutagenesis to clarify the role of two active-site residues in lysozyme

Lysozyme is one of the best studied proteins, and several site-directed mutagenesis have been successfully performed in this enzyme to identify the role of specific amino acid residues in the catalytic process and/ or in the

enzyme stability. We will present here, as an example, a summary of the conclusions obtained from 4 lysozyme mutants, where Trp108 was replaced by Tyr or Gln (W108Y and W108Q) and Asp52 was substituted by Asn or Glu (D52N and D52E) (Inohue *et al*, 1992). In the catalytic mechanism proposed for this enzyme, two carboxyl groups (Glu35 and Asp52) have been identified: it is accepted that, the carboxyl group of Glu35, in a non-ionized form, would serve as a general-acid catalyst, whereas the ionized form of the Asp52 carboxyl group stabilizes, via electrostatic interactions, the oxocarbenium ion intermediate. In the reaction, the pK_a of Glu35 has a value of 6.1, abnormally high as compared to the normal value of 4.4. It has been proposed that this high pK_a value may be due to the hydrophobic influence of Trp108, located in the vicinity of Glu35 (Rupley *et al*, 1967) or, to the electrostatic interactions between Glu35 and Asp52 (Imoto *et al*, 1972).

The study of the mutants W108Y and W108Q showed that the hydrophobic character of Trp108 contributed to raise the pK_a value of Glu35 in 0.6 units, it played a role in the stabilization of the folding structure of lysozyme and it increased the binding ability of the enzyme. The studies also indicated that it may exist an interaction between Trp108 and a saccharide unit of the substrate at the C binding site, as it was suggested by the model building studies mentioned above (Kelly *et al*, 1979).

Mutant D52N, which lacks the electrostatic interaction with Glu35, presented a pK_a value of 5.5 for this last residue. X-ray studies of this mutant showed that, even though Asp52 is 7 Å apart from Glu35, both amino acids are connected with hydrogen bonds, via two tightly-bound water molecules. Therefore Trp108 and Asp52 are almost equally responsible for maintaining the high pK_a value for Glu35; nevertheless 0.4 units of pK_a still needs to be explained by other factors (Inoue *et al*, 1992). Both mutants D52E and D52N showed a drastic decrease of the enzyme's lytic activity, thus confirming that Asp52 is a catalytic residue, and it is related to the D binding site, as proposed by model-building studies.

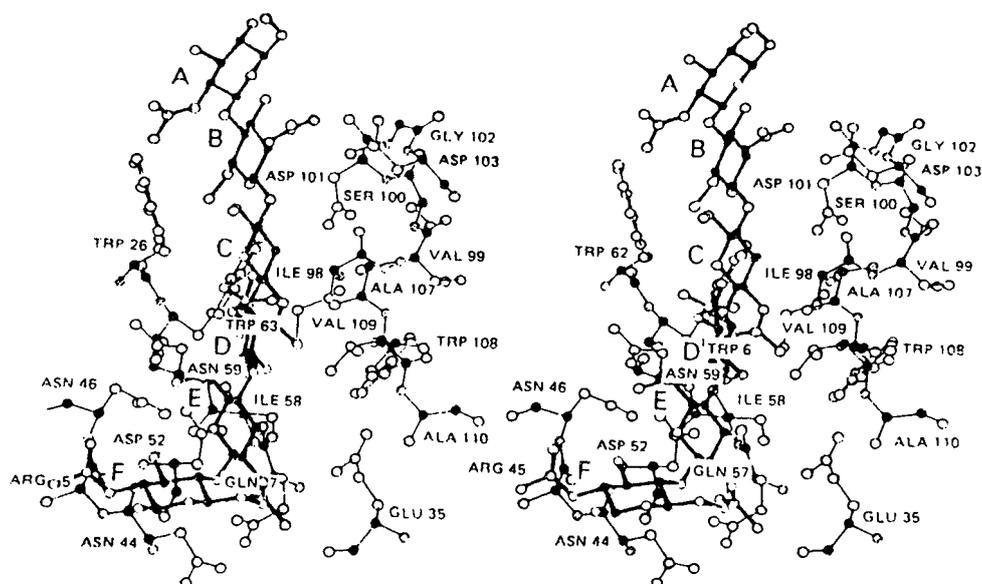


Fig 21. Stereo view of the calculated minimum energy structure of bound (NAG)6 to the rigid active site of lysozyme. The substrate bonds are indicated in heavy lines. (From Pincus *et al*, 1977; reprinted with permission).

EXAMPLE 4.
ACTIVE SITE STUDIES IN SERINE-PROTEASES.
EXAMPLES OF X-RAY DIFFRACTION STUDIES ON AN ACYL-ENZYME INTERMEDIATE, AN ENZYME-PRODUCT COMPLEX AND A COMPLEX WITH A NATURALLY OCCURRING POLYPEPTIDE INHIBITOR

Chymotrypsin, trypsin, and elastase are three pancreatic enzymes that are structurally and kinetically very similar, hydrolyzing peptides and synthetic ester substrates. Their primary structures are composed of approximately 50% identical or similar amino acid residues, and the similarity in their tertiary structures allows a unique structural description (Fersht, 1985). Specificity is the major difference between these enzymes. Trypsin is specific for the positively charged amino acids lysine and arginine; chymotrypsin acts on the large hydrophobic chains of phenylalanine, tyrosine, and tryptophan, whereas elastase is effective with small hydrophobic amino acids such as alanine (Fig 22).

When the tertiary structures of α -chymotrypsin (Matthews *et al*, 1967), elastase (Shotton & Watson, 1970) and bovine trypsin (Huber *et al*, 1974, Stroud *et al*, 1974, Bode & Schwager, 1975) were determined by X-ray diffraction methods, it was shown that the polypeptide backbones of these three enzymes were superimposable,

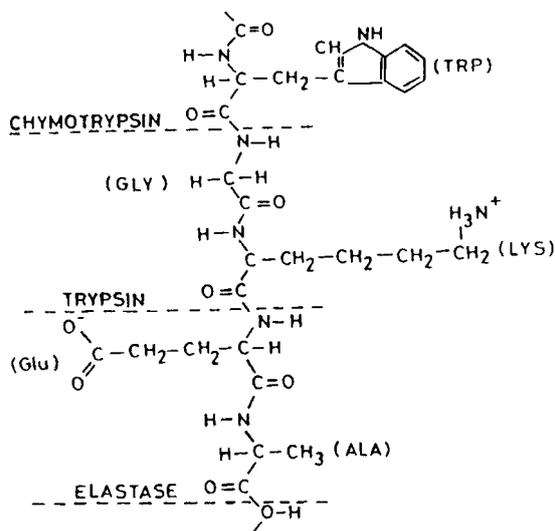


Fig 22. Sites of substrate cleavage for three serine proteases.

except for some small changes in the active sites. These changes explain their specificities: the binding pocket for the aromatic side-chains of the chymotrypsin substrates is a slit 10-12Å deep and 4 by 6 Å in cross section that can perfectly accommodate an aromatic ring 6 Å wide and 3.5 Å thick. A methylene group with 4 Å diameter can fit into the same slit in trypsin, the main difference being that trypsin has replaced

Ser189 by Asp189 to bind the positively charged side-chains of lysine or arginine. The main changes in the active site of elastase with respect to chymotrypsin are the substitution of Gly216 and Gly226 by the bulkier Val216 and Thr226 groups. These changes make the active site accessible only to the small hydrophobic amino acid side chain of alanine (Fig 23).

The three enzymes belong to the group of the 'serine proteases', since it has been shown that a specific serine residue (Ser195) takes part in the catalytic process of each of these enzymes. The tertiary structures show that the imidazole ring of His57 is hydrogen bonded to the carboxylate of Asp102, to form a catalytic triad common to all serine proteases. The mechanism of hydrolysis for all serine proteases is shown schematically in Figure 24 (Fastrez & Fersht, 1973). A non-covalent intermediate is first formed, followed by attack of the hydroxyl group of Ser195 on the substrate to give a tetrahedral intermediate. This intermediate collapses to give an acyl-enzyme, releasing the amino or alcohol. The acyl-enzyme then hydrolyzes to form the enzyme-product complex.

The tertiary structure of both the chymotrypsin product complex and the acyl-enzyme intermediate have been solved by X-ray diffraction methods. The enzyme-product complex was formed by diffusing into α -chymotrypsin crystals the acylated amino acids formyl-L-tryptophan and formyl-L-phenylalanine. The interpretation of the difference Fourier maps at 2.5 Å resolution allowed the location of the acylated amino acid in the active site, as shown for formyl-L-tryptophan in Figure 23 (Steitz *et al.*, 1969).

The lifetime of an acyl-enzyme intermediate in the serine-proteases is of the order of 0.01 seconds at neutral pH, too short for crystallographic observations even with the most modern equipment. This limitation has been solved by using an abnormally poor substrate and a non-physiological pH to produce the acyl-enzyme indoleacryloyl- α -chymotrypsin complex. The crystals of the complex were obtained by incubation of α -chymotrypsin crystals in a saturated solution of indoleacryloyl-imidazole in 3% dioxane

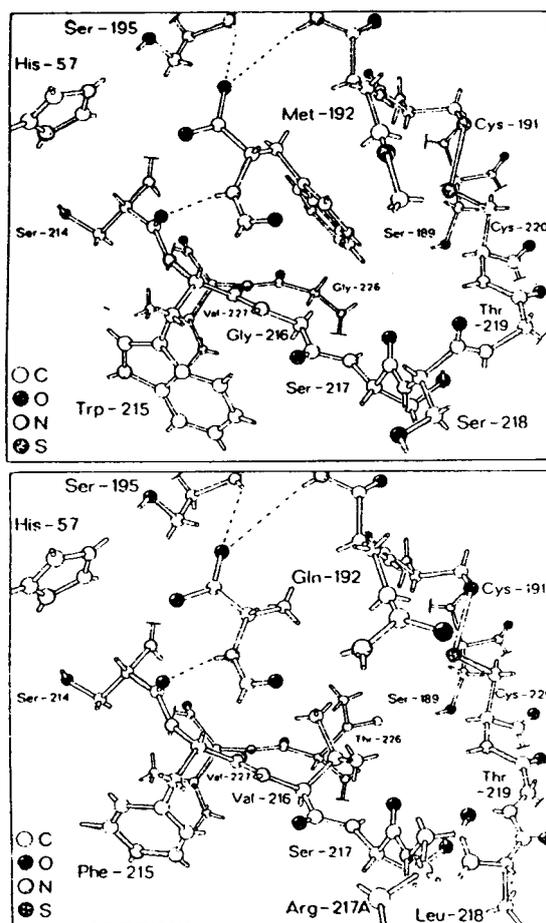


Fig 23. The active sites of chymotrypsin (top) and elastase (bottom). Formyl-trp and formyl-ala are shown bound to the active sites of chymotrypsin and elastase, respectively. The trypsin active site resembles that of chymotrypsin, except that Ser189 becomes Asp189. This difference allows trypsin to bind the positively charged Lys or Arg residues. (From Fersht, 1985; reprinted with permission).

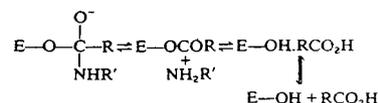
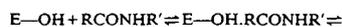


Fig 24. The mechanism of action of serine-proteases. (From Fersht, 1985; reprinted with permission).

and 65% ammonium sulfate at pH 4.0 (Henderson, 1970). The results obtained were extrapolated to the structure of the enzyme complex at neutral pH and lead to the proposal of an acylation mechanism for chymotrypsin.

There are many naturally occurring polypeptide inhibitors which bind to chymotrypsin and to trypsin producing stable

enzyme-substrate complexes. They bind as true substrates, but their own structure does not allow the flexibility that a normal substrate shows on binding. These inhibitors lock the active site and do not allow the diffusion of the amino group released on the cleavage of the peptide from the enzyme. One of these inhibitors is the pancreatic trypsin inhibitor (PTI), a polypeptide with 58 amino acid residues and a molecular weight of 6,500. The tertiary structure of PTI (Huber *et al*, 1970, 1971) and of the complex trypsin-PTI (Ruhlmann *et al*, 1973) have been solved by X-ray diffraction methods. The electron density maps were interpreted with the aid of a detailed model of the complex α -chymotrypsin and those of PTI as determined from their individual structures.

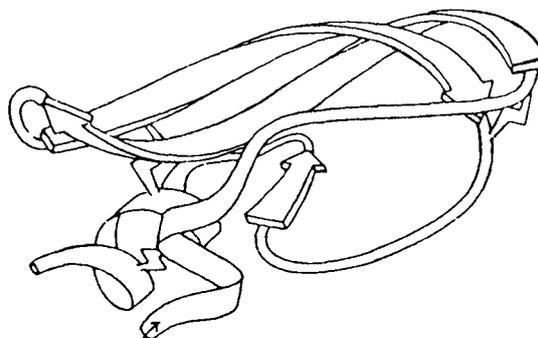
In addition to the active site, the polypeptide substrates bind to a series of subsites across the enzyme surface. Model building studies of the complex trypsin-PTI have shown the existence of subsites S_1 , (active site), S_2 , S_3 on one side of the polypeptide chain and S'_1 on the other side (Fersht, 1985). Figure 25 shows the secondary structure representation of PTI and the regions where this inhibitor binds to trypsin.

New findings on the structure and function of serine proteases based on a combination of X-ray diffraction methods and site-directed mutagenesis

To assess the role of Asp102 in the catalytic process, mutant D102N of trypsin was obtained, crystallized at pH 6.0 and 8.0 and analyzed by X-ray diffraction. The high pH crystal produced a structure very similar to the wild type enzyme, whereas the structure obtained from crystals grown at pH 6.0 presented the side chain of the catalytic residue His57 statistically disordered between two positions, one coincident to that found in the wild type enzyme, and the other with the imidazole ring displaced from the active site towards the solvent (Sprang *et al*, 1987). X-ray studies of the D102N trypsin suggested that the 'gauche' conformation of the imidazole ring of His57 found in the native enzyme is stabilized by a hydrogen bond between the Nd1 atom of His57 and the

carboxylate oxygen of Asp102. In D102N, the conformation of His57 seems to be determined by its protonation state: above neutral pH, the monoprotonated imidazole predominates, and the Nd1 atom of His can accept a hydrogen bond from Nd2 of Asn102. At lower pH this hydrogen bond is lost and the imidazole ring is free to rotate to the 'trans' position.

These findings were complemented with activity measurements of the mutant trypsin at neutral (7.15-7.95) and alkaline (8.77-10.18) pHs values (Craik *et al*, 1987). As shown on Table III, the catalytic efficiency of D102N trypsin was measured as its ability to hydrolyze the ester substrate N-benzyloxycarbonyl-L-lysine thiobenzylester (Z-Lys-S-Bzl). Also the reactivity of Ser195 and His57 were measured with their specific inhibitors diisopropylfluorophosphate (DFP) and tosyl-L-lysine chloromethyl ketone (TLCK), respectively.



Pancreatic Trypsin Inhibitor

Fig 25. Secondary structure representation of the three-dimensional structure of the pancreatic trypsin inhibitor. The N-terminal strand and the large loop joining one of the large and the small β -strands, make contact with trypsin. The disulfide bridges are represented by little zigzag strands. (From Richardson, 1985; reprinted with permission).

Table III

Ratios of activity for native trypsin and D102N trypsin*

Ligand	Kinetic Constant	Relative activity at neutral pH	Relative activity at alkaline pH
Z-Lys-S-Bzl	k_{cat}	4400	18
Z-Lys-S-Bzl	k_{cat}/K_m	11300	152
DFP	k_{obs}	11300	10000
TLCK	k_{obs}	5.1	1.4

* From Craik *et al* (1987).

The conclusions obtained from these results are:

- The relative activity of the mutant enzyme increases with increasing pH values, in agreement with X-ray diffraction results.
- The nucleophilicity of Ser195 is dependent on the negative charge of Asp102.
- The imidazole ring of His57 in the D102N trypsin is not in the correct tautomeric state for removal of the Ser195 proton and thus it reduces the reactivity of the enzyme to DFP, but His 57 can still react with TLCK.

CONCLUSIONS

At the beginning of this article we were faced with two problems: to observe a protein molecule with a resolution power of 10^{-7} mm, a million times larger than that of our naked eye, and, the necessity to 'freeze' an enzymatic reaction in order to discern all the intermediate complexes formed.

The solution to the first problem is given by the X-ray diffraction methods. The aim of this article is to give to non-crystallographers the basis of these methods together with some selected examples of their application to active site studies.

The examples were selected to illustrate different solutions to the problem of 'freezing' the enzymatic reaction, but these examples cover by no means all possible solutions to the study of enzyme-intermediate complexes. A powerful method, used very frequently today, is site-directed mutagenesis. By replacing selected amino acid residues, it is also possible to change the affinity of substrates or intermediate analogs for the active site, and thus to be able to record their positions by X-ray methods.

This article also intends to show the power of the X-ray diffraction methods. The possibility of knowing the exact location of every atom in a molecule, as well as those from the attached solvent molecules, has drastically changed the trends of the biological sciences since the first tertiary structure of a protein was determined. However, the X-ray diffraction methods also have some limitations, such as the following:

1. It used to be a difficult technique which required many years of patient and hard work to obtain and measure hundreds of diffraction spectra with thousands of diffraction spots each. Today, the production of very intense X-ray sources, like the synchrotron radiation, and the automatization of data collection, have solved most of these problems.
2. The crystallization of a protein, the first step in the use of these methods, is not a straightforward task even today. Also, there are proteins that fail to form stable enzyme-heavy atom complexes, thus making the phase determination, necessary to calculate the electron density function, impossible. The knowledge of structurally related proteins can provide the information necessary to overcome the phase problem, either by using the Molecular Replacement method, or by trying to build directly a 3-dimensional model by Molecular Modelling Methods.
3. A very serious objection to the X-ray diffraction method is that it is a 'static' method, and enzymatic catalysis is essentially a dynamic process. However, the combination of the exact knowledge of the enzyme structure, provided by X-ray diffraction, with Molecular Dynamics calculations, and with Site Directed Mutagenesis, can certainly provide most of the information needed to understand these dynamic processes.

The above limitations explain why it is not always possible to use these methods; this has led to the development of a host of alternative methods for the analysis of enzyme active sites. Some of them are discussed elsewhere in this volume.

REFERENCES

- ARTYMIUK PJ, BLAKE CCF (1981) Refinement of human lysozyme at 1.5 Å resolution. Analysis of non-bonded and hydrogen-bond interactions. *J Mol Biol* 52: 737-762
- BANNER DW, BLOOMER AC, PETSKO GA, PHILLIPS DC, POGSON CI (1971) Crystallographic studies of chicken triose phosphate isomerase. *Cold Spring Harbor Symp Quantit Biol* 36: 151-155
- BANNER DW, BLOOMER AC, PETSKO GA, PHILLIPS DC, POGSON CI, WILSON IA, CORRAN PH, FURTH AJ, MILMAN JD, OFFORD R, PRIDDLE JD, WALEY SG (1975) Structure of the chicken muscle

- triosephosphate isomerase determined crystallographically at 2.5 Å resolution. *Nature*, London 255: 608-614
- BEHAVAN G, JONSSON B-H, LINDSKOG S (1990) Fine tuning of the catalytic properties of carbonic anhydrase. Studies of the Thr200-His variant of human isoenzyme II. *Eur J Biochem* 190: 351-357
- BERGSTEN PC, WAARA I, LOVGREN S, LILJAS A, KANNAN KK, BENGTSSON U (1972) Crystal structure of human erythrocyte carbonic anhydrase C.V. Complexes with some anion and sulphonamide inhibitors. In: RØRTH M, ASTRUP P (eds) *Proceedings of Alfred Benzon Symposium IV*. Copenhagen: Munksgaard. pp 363-383
- BLAKE CCF, KOENIG DF, MAIR GA, NORTH ACT, PHILLIPS DC, SARMA VR (1965) Structure of hen egg-white lysozyme. *Nature*, London 206: 757-761
- BLAKE CCF, JOHNSON LN, MAIR GA, NORTH ACT, PHILLIPS DC, SARMA VR (1967) Crystallographic studies of hen egg-white lysozyme. *Proc Royal Soc, London*, B167: 378-388
- BODE W, SCHWAGER P (1975) The refined crystal structure of bovine β-trypsin at 1.8 Å resolution. *J Mol Biol* 98: 693-717
- BRICOGNE G (1974) Geometrical sources of redundancy in intensity data and their use for phase determination. *Acta Crystallogr A* 30: 395-405
- BRICOGNE G (1976) Methods and programs for direct space exploitation of geometrical redundancies. *Acta Crystallogr A* 32: 832-847
- CID H (1987) Active-site studies by X-ray diffraction methods. In: EYZAGUIRRE J (ed) *Chemical Modification of Enzymes: Active site-studies*. England: Ellis Horwood. pp 119-149
- CID H, BUNSTER M (1996) Active-site studies of enzymes by secondary structure prediction and by molecular modeling. *Biol Res* 29: 77-100
- CRAIK CS, ROCZNIAK S, LARGMAN C, RUTTER WJ (1987) The catalytic role of Aspartic Acid in serine proteases. *Science* 237: 909-913
- ENGSTRAND C, FORSMAN C, LIANG Z, LINDSKOG S (1992) Proton transfer roles of Lysine 64 and Glutamic Acid 64 replacing Histidine 64 in the active site of human carbonic anhydrase II. *Biochim Biophys Acta* 1122: 321-326
- ERIKSSON AE, JONES TA, LILJAS A (1988a) Refined structure of human carbonic anhydrase II at 2.0 Å resolution. *Proteins* 4: 274-282
- ERIKSSON AE, KYLSTEN PM, JONES TA, LILJAS A (1988b) Crystallographic studies of inhibitor binding sites in human carbonic anhydrase II: A pentacoordinated binding of SCN⁻ ion to the zinc at high pH. *Proteins* 4: 283-293
- FASTREZ J, FERSHT AR (1973) Demonstration of the acyl-enzyme mechanism for the hydrolysis of peptides and amides by chymotrypsin. *Biochemistry* 12: 2025-2041
- FIERKE CA, CALDERONE TL, KREBS JF (1991) Functional consequences of engineering the hydrophobic pocket of carbonic anhydrase II. *Biochemistry* 30: 11054-11063
- HENDERSON R (1970) Structure of crystalline α-chymotrypsin. IV. The structure of indoleacryloyl-α-chymotrypsin and its relevance to the hydrolytic mechanism of the enzyme. *J Mol Biol* 54: 341-354
- HENDRICKSON WA, TEETER M (1981) Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. *Nature*, London 290: 107-113
- HOWELLS MR (1990) Soft X-ray Imaging for the Life Sciences. In: SS HASNAIN (ed) *Synchrotron Radiation and Biophysics*. England: Ellis Horwood
- HUBER R, KUKLA D, RÜHLMANN A, EPP O, FORMANSK H (1970) The basic trypsin inhibitor of bovine pancreas. *Naturwissenschaften* 57: 389-392
- HUBER R, KUKLA D, RÜHLMANN A, STEIGEMANN W (1971) Pancreatic trypsin inhibitor (Kunitz). Part I. Structure and function. *Cold Spring Harbor Symp Quantit Biol* 36: 141-150
- HUBER R, KUKLA D, BODE W, SCHWAGER P, BARTELS K, DEISENHOFER J, STEIGEMANN W (1974) Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. *J Mol Biol* 89: 73-101
- HÅKANSSON K, CARLSSON M, SVENSSON LA, LILJAS A (1992) Structure of native and apo carbonic anhydrase II and structure of some of its anion-ligand complexes. *J Mol Biol* 227: 1192-1204
- INOHUE M, YAMADA H, YASUKOCHI T, KUROKI R, MIKI T, HORIUCHI T, IMOTO T (1992) Multiple role of hydrophobicity of Tryptophan 108 in chicken lysozyme. Structural stability, saccharide binding ability and abnormal pKa of Glutamic Acid 35. *Biochemistry* 31: 5545-5553
- IMOTO T, JOHNSON LN, NORTH ACT, PHILLIPS DC, RUPLEY JA (1972) Vertebrate lysozymes. *The Enzymes* (3rd Ed.) 7: 665-868
- KANNAN K, NOSTRAND B, FRIDBORG K, LÖVGREN S, OHLSSON A, PETEF M (1976) Crystal structure of human erythrocyte carbonic anhydrase B. Three dimensional structure at a nominal 2.2 Å resolution. *Proc Natl Acad Sci USA* 72: 51-55
- KANNAN KK, PETEF M, FRIDBORG K, CID-DRESDNER H, LÖVGREN S (1977) Structure and function of carbonic anhydrases. Imidazole binding to human carbonic anhydrase B and the mechanism of action of carbonic anhydrases. *FEBS Lett* 73: 115-119
- KELLY JA, SIELECKI AR, SYKES BD, JAMES MNG, PHILLIPS DC (1979) X-ray crystallography of the binding of the bacterial cell wall trisaccharide NAM-NAG-NAM to lysozyme. *Nature*, London 282: 875-878
- KNOWLES JR, LEADLEY PF, MEISTER SG (1972) Triosephosphate isomerase. Isotope studies on the mechanistic pathway. *Cold Spring Harbor Symp Quantit Biol* 36: 157-164
- KREBS JF, RANA F, DLUHY RA, FIERKE CA (1993) Kinetic and spectroscopic studies of hydrophilic amino acid substitutions in the hydrophobic pocket of human carbonic anhydrase II. *Biochemistry* 32: 4496-4505
- LIANG J-Y, LIPSCOMB WN (1990) Binding of substrate CO₂ to the active site of human carbonic anhydrase II: a molecular dynamics study. *Proc Natl Acad Sci USA* 87: 3675-3679
- LIANG Z, XUE Y, BEHAVAN G, JONSSON B-H, LINDSKOG S (1993) Importance of the conserved active-site residues Tyr7, Glu106 and Thr199 for the catalytic function of human carbonic anhydrase II. *Eur J Biochem* 211: 821-827
- LILJAS A, HÅKANSSON K, JONSSON BH, XUE Y (1994) Inhibition and catalysis of carbonic anhydrase. Recent crystallographic analyses. *Eur J Biochem* 219: 1-10
- LINDHAL M, SVENSSON LA, LILJAS A (1993) Metal poison inhibition of carbonic anhydrase. *Proteins* 15: 177-182
- LINDSKOG S, HENDERSON LE, KANNAN KK, LILJAS A, NYMAN PO, STRANDBERG B (1971) Carbonic anhydrase. *The Enzymes* (3rd ed) 5: 587-665
- LOLIS E, PETSKO GA (1990) Crystallographic analysis of the complex between triosephosphate isomerase and 2-phosphoglycolate at 2.5 Å resolution: Implication for catalysis. *Biochemistry* 29: 6619-6625
- LOLIS E, ALBER T, DAVENPORT RC, ROSE D, HARTMAN FC, PETSKO GA (1990) Structure of

- yeast triosephosphate isomerase at 1.9 Å resolution. *Biochemistry* 29: 6609-6618
- LÖVGREN S, BERGSTEN PC, LILJAS A, PETEF M, BENGTTSSON U (1971) Crystal structure of human erythrocyte carbonic anhydrase C. VIII Metal bonding to the apo enzyme. Cited by Bergstén *et al* (1972)
- MATHEWS BW, SIGLER PB, HENDERSON R, BLOW DM (1967) Three-dimensional structure of Tosyl- α -chymotrypsin. *Nature*, London 214: 652-656
- MORENO-HAGELSIEB G, SOBERON X (1996) Protein engineering as a powerful tool for the chemical modification of enzymes. *Biol Res* 29: 127-140
- NOBLE MEM, WIERENGA RK, LAMBEIR A-M, OPPERDOES FR, THUNNISSEN A-MWH, KALK KH, GROENDIJK H, HOL WGJ (1991) The adaptability of the active site of trypanosomal triosephosphate isomerase as observed in the crystal structures of three different complexes. *Proteins* 10: 50-69
- NOSTRAND B (1974) The high resolution structure of human carbonic anhydrase, form B. PhD Thesis, University of Uppsala, Sweden
- NOSTRAND B, WAARA I, KANNAN KK (1975) Structural relationship of human erythrocyte carbonic anhydrase isozymes B and C. In: *Isozymes I. Molecular Structure*. New York: Academic Press. pp 575-599
- PHILLIPS DC (1967) The hen egg-white lysozyme molecule. *Proc Natl Acad Sci USA* 57: 585-495
- PINCUS MR, ZIMMERMAN SS, SCHERAGA HA (1976) Prediction of three-dimensional structures of enzyme-substrate and enzyme-inhibitor complexes of lysozyme. *Proc Natl Acad Sci USA* 73: 4261-4265
- PINCUS MR, ZIMMERMAN SS, SCHERAGA HA (1977) Structures of enzyme-substrate complexes of lysozyme. *Proc Natl Acad Sci USA* 74: 2629-2633
- RICHARDSON J (1985) Describing patterns of protein tertiary structures. *Meth Enzymol* 115: 341-358
- ROSE IA (1975) Mechanism of the aldose-ketose isomerase reactions. *Adv Enzymol* 43: 491-517
- ROSSMAN MG (1972) *Molecular replacement method*. New York: Gordon & Breach
- RUPLEY JA, BUTLER L, GERRING M, HARTDEGEN FJ, PECORARO R (1967) Studies on the enzyme activity of lysozyme. III. The binding of saccharides. *Proc Natl Acad Sci USA* 57: 1088-1095
- RÜHLMANN A, KUKLA D, SCHWAGER P, BARTELS K, HUBER R (1973) Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. Crystal structure determination and stereochemistry of the contact region. *J Mol Biol* 77: 417-436
- SHOTTON DM, WATSON HC (1970) Three-dimensional structure of Tosyl-elastase. *Nature*, London 225: 811-816
- SPRANG S, STANDING T, FLETTERICK RJ, STROUD RM, FINER-MOORE J, XUONG N-H, HAMLIN R, RUTTER WJ, CRAIK CS (1987) The three-dimensional structure of Asn102 mutant of trypsin: Role of Asp102 in serine protease catalysis. *Science* 237: 905-909
- STEITZ TA, HENDERSON R, BLOW DM (1969) Structure of crystallin α -chymotrypsin. III. Crystallographic studies of substrates and inhibitors bound to the active site of α -chymotrypsin. *J Mol Biol* 46: 337-348
- STROUD RM, KAY LM, DICKERSON RE (1974) The structure of bovine trypsin: electron density maps of the inhibited enzyme at 5 Å and at 2.7 Å resolution. *J Mol Biol* 83: 185-208
- STUART DI, JINBI D, MEIZHEN L, JUMMING Y, TODD R, JIAYAO L, JIAHUI W, DONGCAI L (1984) X-ray studies on the structure of des-pentapeptide insulin at 2.4 Å resolution. Solution of the phase problem using the anomalous scattering of cadmium and the derivation of a model structure. *Scientia Sinica (B series)* 27: 572-582
- TU C, SILVERMAN DN, FORSMAN C, JONSSON B-H, LINDSKOG S (1989) Role of His64 in the catalytic mechanism of human carbonic anhydrase II studied with a site-specific mutant. *Biochemistry* 28: 7913-7918
- UNGE T, LILJAS L, STRANDBERG B, WAARA I, KANNAN KK, FRIDBORG K, NORDMAN CE, LENTZ PJ Jr (1980) Satellite tobacco necrosis virus structure at 4.0 Å resolution. *Nature* 285: 373-377
- WIERENGA RK, NOBLE MEM, VRIEND G, NAUCHE S, HOL WGJ (1991a) Refined 1.83 Å structure of trypanosomal triosephosphate isomerase crystallized in the presence of 2.4 M-ammonium sulphate. A comparison with the structure of the trypanosomal triosephosphate isomerase-glycerol-3-phosphate complex. *J Mol Biol* 220: 995-1015
- WIERENGA RK, NOBLE MEM, POSTMA JPM, GROENDIJK H, KALK KH, HOL WGJ, OPPERDOES FR (1991b) The crystal structure of the "open" and the "closed" conformation of the flexible loop of the trypanosomal triosephosphate isomerase. *Proteins* 10: 33-49
- WIERENGA RK, NOBLE (1992) Comparison of the refined crystal structures of liganded and unliganded chicken, yeast and trypanosomal triosephosphate isomerase. *J Mol Biol* 224: 1115-1126
- XUE Y, VIDGREN J, SVENSSON, LA, LILJAS A, JONSSON B-H, LINDSKOG S (1993a) Crystallographic analysis of Thr200-His human carbonic anhydrase II and of its complex with the substrate, HCO₃⁻. *Proteins* 15: 80-87
- XUE Y, LILJAS A, JONSSON B-H, LINDSKOG S (1993b) Structural analysis of the Zinc hydroxide-Thr199-Glu106 hydrogen-bond network in human carbonic anhydrase II. *Proteins* 17: 93-106

BOOKS OR CHAPTERS RECOMMENDED FOR FURTHER READING

- BLUNDELL TL, JOHNSON LN (1976) *Protein Crystallography*. New York: Academic Press
- BUERGER MJ (1960) *Crystal Structure Analysis*. New York: John Wiley & Sons
- DICKERSON RE, GEIS I (1969) *The Structure and Action of Proteins*. New York: Harper & Row
- DUKE PJ, MICHETTE AG (eds) (1990) *Modern Microscopies: Techniques and Applications*. New York: Plenum Press
- FERSHT A (1985) *Enzyme Structure and Mechanism*. 2nd ed. San Francisco, CA: WH Freeman
- GLUSKER JP, TRUEBLOOD KN (1985) *Crystal Structure Analysis. A Primer*. 2nd ed. Oxford: Oxford University Press
- HARBURN C, TAYLOR CA, WELBERY TR (1975) *Atlas of Optical Transforms*. London: G Bell & Sons
- LIPSON H (1972) *Optical Transforms*. London: Academic Press
- RAMACHANDRAN GN, SRINIVASAN R (1970) *Fourier Methods in Crystallography*. New York: Wiley-Interscience
- RHODES G (1993) *Crystallography Made Crystal Clear*. San Diego, CA: Academic Press