Proceedings of the International School of Physics «Enrico Fermi»
Course CXXVIII, E. Burattini and A. Balerna (Eds.)
IOS Press, Amsterdam 1996

Time-Resolved Protein Crystallography.

H. L. MONACO

Dipartimento di Genetica, Università di Pavia - 27100 Pavia, Italia

1. – Introduction.

Most of the known three-dimensional structures of the macromolecules of biological interest are the result of X-ray diffraction studies on single crystals. Biocrystallography has thus made a very significant contribution to molecular biology: it has provided a more or less static picture of the structure of the macromolecules that play an essential role in biological processes. However, during their functioning, these molecules undergo significant structural changes that in most cases cannot be characterized by conventional X-ray crystallography. It is obvious that it would be desirable to include these changes in the structural description, but this goal has been achieved in only a very limited number of cases [1-4].

The goal of time-resolved X-ray crystallography is to examine those short-lived intermediates which are compatible with the packing present in the crystals that were solved by conventional X-ray crystallography.

Like in every kinetic experiment, in time-resolved X-ray diffraction studies the crucial factor is the data acquisition rate, that is the time required to collect the data has to be negligible compared to the time required by the system which is being studied to evolve. It is for this reason that a synchrotron source is an essential ingredient of time-resolved X-ray diffraction studies. It is only the extremely high brilliance of a synchrotron that can provide enough photons to produce sufficient data to follow these evolving phenomena. But it is not only brilliance that is important, the second factor is the radiation spectrum of the source that offers the possibility to have high intensities at different wavelenghts. For X-ray diffraction studies on single crystals this means that the Laue polychromatic method of data collection can be used [5].

Of the many aspects that have to be considered in an experiment of this kind, I have chosen to focus on essentially two:

a) how the reaction is started in the crystal and

b) how the diffraction data are collected.

We will discuss these two topics and then turn our attention to four case studies chosen to illustrate the possibilities of this rather new methodology.

2. – Reaction initiation.

Perhaps the most difficult experimental problem in time-resolved X-ray crystallographic studies is the finding of an event that will initiate the reaction to be followed in time. Such an event has to satisfy the following requisites:

a) It must not alter the crystal integrity.

b) Initiation must be fast.

c) The process must be spatially uniform.

Of the three requisites listed the first two are more obvious, but the third is also crucial since time-resolved studies require that a large majority if not all the unit cells in the crystal be synchronized with respect to the triggering event. Otherwise the result, which is always spatially averaged (*i.e.* averaged for all the unit cells present in the crystal), becomes uninterpretable.

The events that can be used to initiate a reaction in a time-resolved X-ray diffraction study are the following:

1) reagent diffusion into the crystals,

2) photoactivation of a stable ligand precursor,

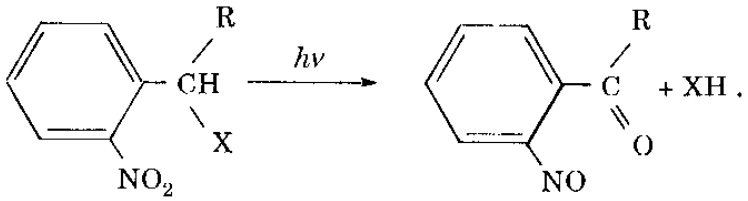
3) temperature jump,

4) photoactivation of an inactive protein.

We recall that protein crystals contain a minimum of 30% by weight of water and that they are on the average roughly 50% water and 50% protein. Thus the large liquid-containing channels that are present in all macromolecular crystals can be used to access the protein active sites. Reagent diffusion is effected by washing the crystal mounted in a flow cell with mother liquor containing the effector that reacts with the protein. This method of starting a reaction works well when the binding of the effector is fast and its modification by the protein relatively slow, but, in spite of these limitations, the method has been used successfully many times as the examples that will be discussed below will show.

A caged effector is simply a stable molecule that will release the ligand of interest when activated with light of the appropriate wavelength. This type of molecule is usually the molecule of interest bound to a photolabile protecting

group that renders it biologically inactive. In this method of activation, the inert precursor is diffused into the crystals and is let to equilibrate for a time sufficiently long so that all the crystal channels may contain the same precursor concentration. The reaction is then triggered by a flash of light that liberates the active ligand in the entire crystal volume simultaneously. The photochemical reaction that has been used most often to release biologically important metabolites from their inactive precursors is the following:



Using this photochemical reaction the following biologically important effectors can be released [6]:

a) protons,

b) nucleotides and analogs,

c) divalent cations,

d) phosphate and phosphate esters,

e) amine neurotransmitters, etc.

In an ideal temperature jump experiment one would

a) lower the crystal temperature to a value at which the reaction of interest becomes negligible,

b) diffuse the effector at that temperature and give it enough time to equilibrate with the entire crystal volume and

c) initiate the reaction by raising the temperature to an appropriate value.

Although the method is, in principle, no more complicated than the others, it has so far met with limited success. Part of the reason is that the crystal mother liquor has to be substituted by a cryosolvent that can modify the functional behaviour of the protein. In addition, gross temperature changes result often in partial or total loss of the crystalline order [7].

Photoactivation of a caged protein requires the preparation and crystallization of an inactive form of the protein which has one or more modified essential amino acids at the active site. The modification can be done by chemically reacting the native protein with a compound that specifically blocks the active site or by introducing caged residues during the synthesis of the protein. In both cases the blocking moiety has to be susceptible to release by a light pulse. This ex-

perimental approach has not yet been used in time-resolved X-ray crystallographic studies although some promising preliminary work has been done.

3. – Data collection methods.

The first time-resolved X-ray crystallographic studies used monochromatic radiation and the standard rotation (oscillation) geometry as applied in conventional macromolecular crystallography. In this method, the crystal is rotated about an axis perpendicular to the X-ray beam through an angular range of no more than a few degrees, while the diffraction data are recorded in a plane perpendicular to the X-ray beam [8]. The advantage of this method is that data processing and interpretation are relatively simple. Its disadvantage is that many exposures on different planes (film or area detector) are required to obtain a complete data set. With the existing synchrotron sources the rate of data acquisition can be of up to 1000 reflections per second. Depending on the crystal unit cell and the resolution of the diffraction pattern, several tens of thousands of reflections may have to be recorded and data collection will thus require several tens of seconds.

Another monochromatic method that has been used to collect data very fast on a synchrotron source uses a very large Weissenberg camera in which again the crystal is rotated about an axis perpendicular to the beam, but in which the detector is a cylinder whose axis is coincident with the rotation axis. In addition, in this method the detector is translated along its axis and this allows the use of larger rotation ranges.

Although monochromatic methods are presently suitable only for relatively slow reactions, the advent of more powerful synchrotron sources may render them competitive again.

The first X-ray diffraction experiment, communicated by FRIEDRICH, KNIPPING and VON LAUE in 1912, was performed on a stationary copper sulphate crystal using polychromatic radiation. This method of recording a diffraction pattern, with a stationary crystal and polychromatic radiation in a wavelength interval, later became known as the Laue method [9]. At one point in the development of X-ray crystallography it fell largely into disuse due mainly to two reasons:

a) the difficulty to obtain sufficiently intense radiation in a reasonable wavelength interval on a conventional X-ray source and

b) the difficulties encountered in the interpretation of the diffraction pattern.

With the advent of synchrotron radiation sources the first difficulty was completely removed and the road was open to the developments that led

to a revitalization of the method which is currently used for most time-resolved X-ray diffraction studies.

Figure 1 is a representation of the geometry of data collection using the Laue method. The two Ewald spheres corresponding to the two extreme wavelenghts in the radiation interval limit the volume of reciprocal space that contains the nodes that will produce diffraction when the specimen is hit with radiation in the wavelength interval between λ_{max} and λ_{min}.

In the Laue method, different cross-sections of a node of reciprocal space are excited by radiation of slightly different wavelength and the result is an intensity integrated over the wavelength rather than over the volume sweeping through a single Ewald sphere, characteristic of all of the monochromatic methods.

Since many nodes of reciprocal space are excited simultaneously, the Laue method is extremely fast and efficient. In addition, the percentage of reflections that are multiple, *i.e.* that contain contributions from more than one node of reciprocal space, is surprisingly low, less than 17% under reasonable experimental conditions. Another advantage is that in most cases the crystal must not be rotated, that is sufficient data can be recorded from only one still photograph if the crystal orientation has been properly chosen. These important advantages are counterbalanced by two severe limitations:

a) The method is extemely sensitive to disorder in the crystal.

b) The low-resolution data are lost.

The first problem is encountered in almost all time-resolved studies since disorder or increase in mosaicity is to be expected whenever a successful reac-

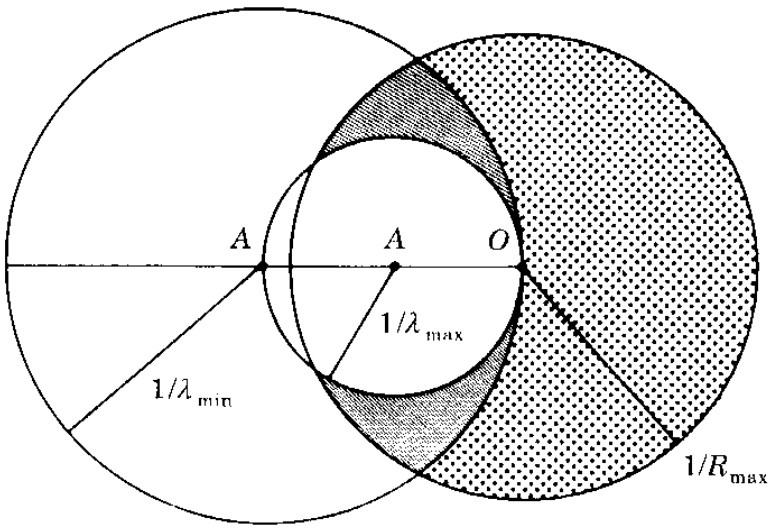


Fig. 1. – The Laue method. A stationary specimen is irradiated with X-rays covering the wavelength interval λ_{max}–λ_{min}. The shaded area represents the volume of reciprocal space that produces diffraction.

tion is initiated in a crystal. The second limitation is also quite serious. Examination of the shape of the volume limited by the two Ewald spheres shows that, as we approach the origin of reciprocal space, this volume becomes smaller and smaller. The result is that the data sets collected with this method normally do not contain the reflections with resolution of less than about (4–5) Å. In spite of these limitations, the method has been and is being used for virtually all time-resolved studies because it is very fast and works well at high resolution with perfect crystals. In all of the examples that we discuss below the Laue method was used to record the time-resolved X-ray diffraction data.

4. – Four case studies.

1) *Glycogen phosphorylase.* – The enzyme glycogen phosphorylase exists in two interconvertible forms: the active form *a* and the inactive form *b*. The inactive form *b* is converted into the active form *a* by phosphorylation of a single serine, Ser 14. X-ray crystallographic studies of the two forms by two different groups have revealed the molecular architecture [10, 11]. The enzyme is a dimer of two identical polypeptide chains of 841 amino acids each. Each monomer is folded into three domains: an N terminal domain of 310 residues, a glycogen-binding domain of 160 residues and a C terminal domain of 371 residues. The catalytic site is located in a crevice found at the interface between the three domains. The enzyme uses the cofactor pyridoxal phosphate (vitamin B_6) which is bound through a Schiff base to a Lys in the C terminal domain, in an area close to the active site. The protein binds to the glycogen particle at a site which is fairly distant from the catalytic site. In addition to these sites, there is an allosteric effector binding site which binds adenosine monophosphate (AMP), an allosteric activator of phosphorylase *b*. All these sites were identified by conventional diffraction studies before any time-resolved studies were undertaken.

The glycogen phosphorylase *b* crystals were the system used for the first time-resolved studies using both the monochromatic [12] and the Laue polychromatic method of data collection [13, 14]. In the study that used monochromatic radiation, the enzymatic reaction was slowed down to follow the conversion of heptenitol to heptulose-2-phosphate, the phosphorolysis of maltoheptaose to yield glucose-1-phosphate and the oligosaccharide synthesis reaction that involves maltotriose and glucose-1-phosphate. In order to start these reactions, the substrates were diffused into the crystals using a flow cell and allowing for about 15 min equilibration time before the data collection was started. The data collection was done with the standard oscillation method as used in conventional macromolecular crystallography and required times ranging from 30 min to 2 h in a synchrotron radiation source. It was only because the kinetics of the processes studied is very slow that it was possible to follow the conformational changes associated with catalysis. Although these studies were initiated

with a conventional source, the use of synchrotron radiation proved to be essential since it reduced the data collection times from about one week to less than one hour.

In the first crystallographic study with the Laue method of data collection the wavelength range used was (0.2–2.1) Å and the exposure time was 3 s. Again, the substrate was diffused into the crystals using a flow cell and allowing for an equilibration time of 60 min. The difference Fourier map, showing the substrate maltoheptaose bound to phosphorylase *b*, clearly demonstrated the feasibility of this type of experiment.

2) *The catalytic domain of the* H-ras p21 *protein.* – The ras oncogenes are a common group of genes that can be isolated from human cancer cells and that encode for a family of proteins which are called the ras p21 proteins. Although these proteins are believed to be involved in cell growth, normal and malignant, their exact physiological role has not been proven. They are all single polypeptide chains, 188 amino acids long, that have been shown to be active when bound to guanosine triphosphate (GTP) and inactive if bound to guanosine diphosphate (GDP). The catalytic domain of the H-ras p21 protein is a truncated form of the entire protein that includes amino acids 1-166 and maintains the functional ability to bind the nucleotide. The three-dimensional structure of this protein was determined independently by an American group complexed to GDP and a European group complexed to GTP [15, 16]. The protein has a typical nucleotide-binding fold, *i.e.* there is a central parallel six-stranded beta-sheet and five alpha-helices positioned on both sides of it. The loops that connect these elements of secondary structure are involved in the nucleotide binding. The conventional crystallographic work has also identified the specific amino acids involved in the ligand binding.

The time-resolved X-ray crystallographic study on this protein [17] used caged GTP to study the conformational changes that follow the

a) liberation of the GTP molecule by photolytic removal of the protecting group,

b) hydrolysis of the GTP ligand with the consequent formation of the p21-GDP complex.

After removal of the protecting group, the newly formed GTP is hydrolyzed with a half-life of about 40 min at room temperature. The data required to determine the structure of this short-lived p21-GTP complex were collected using the Laue method within 5 min after the photolytic formation of the complex. The electron density map showed clearly the position of the ligand and the way it interacts with an essential Mg ion and the relevant protein side chains.

Another data set collected 14 min after the triggering event showed that there was a clear loss of electron density in the position of the gamma

phosphate and that substantial structural changes were beginning to take place in the protein.

3) *Trypsin.* – The three-dimensional structure of the protease trypsin was determined about 20 years ago by conventional X-ray diffraction studies [18, 19]. The molecule is 223 amino acid residues long and contains six disulfide bridges. The folding of this enzyme is complex and can be described by saying that the protein is made up of two closely related domains, each of which is a Greek key beta barrel. Thus the costituent chains tend to be fully extended and to run parallel to one another forming hydrogen bonds between them.

Chymotrypsin, trypsin and elastase are the three best-known members of the serine protease family. Although the three enzymes have different specificity, they share extended sequence homology, have quite similar three-dimensional structures and exhibit a catalytic mechanism which is nearly identical. The specificity of trypsin, which is more pronounced than that of the other members of the family, requires the presence of a basic amino acid residue, an arginine or a lysine, in the polypeptide chain that will be cleaved. All the serine proteases break a polypeptide chain using a mechanism that requires the presence of three key residues at the active site and in the case of trypsin are Ser 195, His 75 and Asp 102. Like all Ser proteases, trypsin hydrolyzes its substrate via a two-step mechanism. In the first step, the peptide bond cleavage begins with an attack of the oxygen of the hydroxyl group of the active serine on the carbonyl carbon atom of the peptide bond that will be hydrolyzed. The peptide bond is cleaved and an acyl-enzyme intermediate is formed. The attacking Ser is activated by the presence of the nearby His and Asp which together with it form the catalytic triad. In the second step of this mechanism, a water molecule reacts with the acyl-enzyme intermediate releasing the bound reaction product and regenerating the active site Ser.

In the time-resolved X-ray study [20] the species present in the crystals was a stable intermediate with the reactive Ser acylated. The triggering event was a *p*H jump from 5.5 to 8.5 which was done in a flow cell and resulted in the hydrolysis of the acylated protein. Three data sets were collected using the Laue method: the first before the *p*H jump, the second inmediately thereafter and the third after 90 min. The main result of this study is the observation of the water molecule which hydrolyzes the stable intermediate and of very subtle displacements of the side chains of the residues involved in catalysis.

4) *The cytochrome-c peroxidase compound* I. – Cytochrome-c peroxidase (CCP) is an enzyme found in yeast mitochondria that catalyzes the oxidation of ferrocytochrome-c by peroxides

$$2H^+ + ROOH + 2\ \text{cyt c}(Fe^{2+}) \overset{CCP}{\Longrightarrow} 2\ \text{cyt c}(Fe^{3+}) + ROH + H_2O\,.$$

The enzyme is a single 294 residue polypeptide chain with one heme group that

is folded into two domains that together present 10 long and 3 very short alpha-helices. The beta-structure present in the molecule is only about 12% of the residues: a small 3-stranded beta-sheet and two short antiparallel pairs [21].

The structure of the complex formed by the enzyme and two different cytochromes, yeast and horse heart, not ony has revealed the mode of interaction of these two proteins, but it has also suggested a possible electron transfer pathway between CCP and cytochrome-c [22].

If the oxidant participating in the reaction catalyzed by CCP is hydrogen peroxide, the overall reaction

$$2H^+ + H_2O_2 + 2 \text{ cyt c}(Fe^{2+}) \overset{CCP}{\Longrightarrow} 2 \text{ cyt c}(Fe^{3+}) + 2H_2O$$

can be broken down into three major steps:

1) $CCP(Fe^{+3}) + H_2O_2 \Rightarrow$ compound I + H_2O,

2) compound I + cyt c (Fe^{2+}) + $H^+ \Rightarrow$ compound II + cyt c (Fe^{3+}),

3) compound II + cyt c (Fe^{2+}) + $H^+ \Rightarrow CCP(Fe^{+3})$ + cyt c (Fe^{3+}) + H_2O.

In the first step, the enzyme reacts with hydrogen peroxide to generate the oxidized intermediate called compound I which contains the iron atom in the oxidized state 4 and in which a second electron present on the oxygen has been abstracted from a protein side chain, Trp 191, thus generating a semi-stable free radical (Fe^{4+}, O^{2-}, Trp 191).

In the time-resolved study [23], compound I was produced by reacting the enzyme crystals with H_2O_2 in a flow cell. The data were collected using the Laue method before 15 min had elapsed after the triggering event. Previously, it had been shown using spectroscopic techniques that the intermediate was stable for about 30 min. The structure of the intermediate has shown very clear changes in the enzyme at the peroxide binding site but no perceptible modifications of the Trp 191 electron density.

REFERENCES

[1] CRUICKSHANK D. W. J., HELLIWELL J. R. and JOHNSON L. N., Editors, *Time Resolved Macromolecular Crystallography* (Oxford University Press for the Royal Society, Oxford) 1992.

[2] MOFFAT K., *Annu. Rev. Biophys. Biophys. Chem.*, **18** (1989) 309.

[3] HAJDU J. and JOHNSON L. N., *Biochemistry*, **29** (1990) 1669.

[4] HAJDU J. and ANDERSSON I. A., *Annu. Rev. Biophys. Biomol. Struct.*, **22** (1993) 467.

[5] HELLIWELL J. R., *Macromolecular Crystallography with Synchrotron Radiation* (Cambridge University Press, Cambridge) 1991.

[6] MCCRAY, J. A. and TRENTHAM, D. R., *Annu. Rev. Biophys. Biophys. Chem.*, **18** (1989) 239.

[7] MAKINEN M. and FINK A. L., *Annu. Rev. Biophys. Bioeng.*, **6** (1977) 301.
[8] GIACOVAZZO C., Editor, *Fundamentals of Crystallography* (Oxford University Press, Oxford) 1992.
[9] AMORÓS J. L., BURGER M. J. and CANUT DE AMORÓS M., *The Laue Method* (Academic Press, New York, N.Y.) 1975.
[10] WEBER I. T., JOHNSON L. N., WILSON K. S., YEATES D. G. R., WILD D. L. and JENKINS J. A., *Nature (London)*, **274** (1978) 433.
[11] FLETTERICK R. and MADSEN N. B., *Annu. Rev. Biochem.*, **49** (1980) 31.
[12] HAJDU J., ACHARYA K. R., STUART D. I., MC LAUGHLIN P. J., BADFORD D., OIKONOMAKOS N. G., KLEIN H. and JOHNSON L. N., *EMBO J.*, **6** (1987) 539.
[13] HAJDU J., MACHIN P. A., CAMPBELL J. W., GREENHOUGH T. J., CLIFTON I. J., ZUREK S., GOVER S., JOHNSON L. N. and ELDER M., *Nature (London)*, **329** (1987) 178.
[14] HAJDU J., ACHARYA K. R., STUART D. I., BARFORD D. and JOHNSON L. N., *TIBS*, **13** (1988) 104.
[15] DE VOS A. M., TONG L., MILBURN M. V., MATIAS P. M., JANCARIK J., NOGUCHI S., NISHIMURA S., MIURA K., OHTSUKA E. and KIM S.-H., *Science*, **239** (1988) 888.
[16] PAI E. F., KABSCH W., KRENGEL U., HOLMES K. C., JONH J. and WITTINGHOFER A., *Nature (London)*, **341** (1989) 209.
[17] SCHLICHTING I., ALMO S. C., RAPP G., WILSON K., PETRATOS K., LENTFER A., WITTINGHOFER A., KABSCH W., PAI E. F., PETSKO G. A. and GOODY R. S., *Nature (London)*, **345** (1990) 309.
[18] STROUD R. M., KAY L. M. and DICKERSON R. E., *J. Mol. Biol.*, **83** (1974) 185.
[19] KRIEGER M., KAY L. M. and STROUD R. M., *J. Mol. Biol.*, **83** (1974) 209.
[20] SINGER P. T., SMALÅS A., CARTY R. P., MANGEL W. F. and SWEET R. M., *Science*, **259** (1993) 669.
[21] FINZEL B. C., POULOS T. L. and KRAUT J., *J. Biol. Chem.*, **259** (1984) 13027.
[22] PELLETIER H. and KRAUT J., *Science*, **258** (1992) 1748.
[23] FÜLÖP V., PHIZACKERLEY R. P., SOLTIS M. S., CLIFTON I. J., WAKATSUKI S., ERMAN J., HAJDU J. and EDWARDS S. L., *Structure*, **2** (1994) 201.