X-ray diffuse scattering and molecular dynamics in proteins
Ph. Faure, J. Pérez, J. Doucet, J. Benoit

To cite this version:

HAL Id: jpa-00253513
https://hal.archives-ouvertes.fr/jpa-00253513
Submitted on 1 Jan 1994

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
X-ray diffuse scattering and molecular dynamics in proteins
Ph. Faure, J. Pérez, J. Doucet and J.P. Benoit

LURE, Laboratoire CNRS-CEA-MEN, Bât. 209D, Université Paris-Sud, 91405 Orsay cedex, France

Abstract. — In this paper we review the main significant results obtained up to now by X-ray diffuse scattering analysis applied to crystallized proteins. The results are classified in two groups: those assuming intermolecular correlations and those assuming short and long-range intramolecular correlations. We try to perform a synthesis of these results and we specially discuss the diffuse scattering in terms of the low frequency normal modes or the "liquid-like motion". In addition, we also consider the effect of uncoupled rigid body translations of the molecule, sometimes found to be in good agreement with the experimental patterns.

1. INTRODUCTION

It is well established that the biological activity of proteins is related not only to their mean three-dimensional structure, but also to their intramolecular mobility. A well known illustration is the ligation of CO by hemoglobin: the comparison of the quaternary structures of the deoxy and carbonmonoxy forms shows a 15° rotation of the α2β2 dimer relative to the α1β1 dimer [1]. Another good example of high collective intramolecular motion of atoms is the "hinge-bending" motion in the lysozyme molecule [2-3]. The number of atoms involved in correlated motions can vary from a few atoms up to thousands. The frequencies associated with such motions depend on the number of atoms involved and are expected to lie in the range 10⁹ to 10¹⁴ Hz approximately. Various techniques sensitive to dynamics have been used to study these motions, as for example, NMR [4,5], Raman spectroscopy [6,7], Mössbauer spectroscopy [8], neutron spectroscopy [9],...

In the refinement step of the X-ray structure determination, the dynamical information is concentrated in an isotropic B_j temperature factor, introduced for each atom. This factor is related to the mean square amplitude u_j of the small displacements of the atom around its average position [2]. In the harmonic approximation the relation is B_j=8π²u_j². Unfortunately only poor information concerning the correlated motions of atoms can be reached by this technique, since B thermal factors are individual atomic factors summing up all the types of displacements.

The correlation information can be obtained accurately by the X-ray diffuse scattering study, which gives access to the deviations from the ideal mean crystalline structure. This technique was used for the first time in proteins by Phillips et al., who have derived a model of filament motion for tropomyosin[10]. Since then it has been extensively used to study the correlated motions of atoms in proteins. In this article we review and discuss the main results, paying a particular attention to the calculated diffuse scattering deduced from normal mode analysis.
2. WHAT IS DIFFUSE SCATTERING?

2.1 Theory
In a real crystal, the effect of imperfections is to split the scattered intensity into two components \( I_B \) and \( I_D \). The term \( I_B \) corresponds to the set of Bragg reflections which arise from the average periodic structure of the crystal; the intensity of the Bragg peaks can then be derived from the expression valid for a perfect crystal by multiplying each atomic scattering factor by the Debye-Waller factor \( \exp(-B_j S^2/4) \). The second term \( I_D \) represents the diffuse scattering intensity; unlike \( I_B \), this intensity can be spread all over the reciprocal space and its value and location strongly depend on the correlations between atomic displacements. The diffuse scattering intensity can be written as \([11]\):

\[
I_D(S) = N \sum_m \langle (F_n - \langle F \rangle)(F_{n+m} - \langle F \rangle)^* \rangle \exp(-2\pi S \cdot r_m)
\]

where \( F_n \) is the structure factor of the \( n \)-th unit cell, \( S \) is the scattering vector and the sum \( \sum_m \) runs over the relative position vectors \( r_m \) between unit cells (* means complex conjugate). The correlation function \( \langle (F_n - \langle F \rangle)(F_{n+m} - \langle F \rangle)^* \rangle \) is determined by the correlations between the atomic displacements. If the crystal can be divided into \( p \) blocks (or assemblies) of atoms or molecules characterized by correlated displacements, (1) can be written as:

\[
I_D = \sum_p I_D^p \quad \text{with} \quad I_D^p = \langle |F_p|^2 \rangle - 1 \langle |F_p|^2 \rangle.
\]

where \( F_p \) denotes in this case the structure factor of the \( p \)-th block. \( \langle \rangle \) denotes spatial and/or time average running over all the \( p \) blocks that are supposed independent (uncorrelated).

2.2 Example
Figure 1. represents an experimental pattern showing strong X-ray diffuse scattering. At least three components of diffuse scattering can be observed on this pattern (this is the case for almost all protein crystals):

- a - haloes around each Bragg spot.
- b - diffuse streaks passing through reciprocal lattice nodes.
- c - very low diffuse background patches or diffuse "clouds".

Fig. 1.
Experimental pattern of a tRNA-seryl-synthetase crystal \([12,13]\) recorded by T.Thuene and S.Cusack (EMBL Grenoble Outstation) on the Troika beam line of the ESRF synchrotron at Grenoble. The image corresponds to a still photograph on an imaging plate system at a wavelength of 0.7 Å. The diffuse scattering due to the capillary, the solvent and the air were subtracted according to the method described below.
The three main types of diffuse scattering are indicated by letters.
(Reproduced with authorization of the authors)

2.3 Methods
The diffuse scattering intensity is currently \( 10^2 \) to \( 10^6 \) times weaker than the Bragg intensity. Consequently, its observation requires high intensity beams, sensitive detectors, long exposure times and low background. X-ray synchrotron radiation (used in the typical wavelength range 0.7 to 1.2 Å) associated with an imaging plate system is suitable. The figure 2.a shows a typical experimental pattern. Intense parasitic scattering arises from the solvent (around and inside the crystal), the capillary and the air. These contributions are measured by recording in the same geometry, the scattering by the capillary and air...
alone, then the scattering by the capillary, air and solvent. They are then subtracted of the rough experimental pattern 2.a and a typical "corrected experimental" pattern is obtained on figure 2.b. The subtraction of the Bragg spots intensity was not made mainly because its value is currently $10^5$ times that of diffuse scattering intensity, which implies considerable difficulties in scaling both contributions. Since Bragg intensities cannot be measured with a precision better than 2 or 3%, only a smoothing around the Bragg spots is sometimes made; but in this case the intensity of the haloes surrounding the Bragg peaks is likely to be modified.

3. MAIN RESULTS

We review below the most significant results obtained up to now, classifying them in two groups: the first group includes those corresponding to the long range intermolecular correlations; the second group is relative to the intramolecular correlations, either short range or long range.

3.1. INTERMOLECULAR CORRELATIONS

3.1.1. Study of the diffuse streaks

The first study was made by J. Doucet and J.P. Benoit[14] on the orthorhombic form of lysozyme. The diffuse streaks observed on the patterns do correspond in fact to the intersection between the Ewald's sphere and diffuse reciprocal planes. The authors showed that these diffuse planes are due to the existence in the crystal of independent blocks of molecules aligned along the crystallographic a or c directions moving rigidly in the same directions. The amplitude of the displacements is about 0.25 Å and the correlation extends to about 250 Å.

3.1.2. Study of the haloes

Two studies of the halo intensities surrounding the Bragg peaks have been made: the first for insulin by D. Caspar et al. [15] and the second by J. Clarage et al. for the tetragonal form of lysozyme[16]. These haloes are present in all crystals and are known as TDS (Thermal Diffuse Scattering). In "classical" rigid crystals, the haloes are due to the so-called phonons (harmonic vibrations propagating on very long distances). In the Debye harmonic and isotropic theory, the intensity of these haloes decreases according to $1/q^2$ where $q$ represents the distance to the nearest reciprocal node. For the insulin and the tetragonal lysozyme crystals the authors have found a decrease law $1/q^n$ with $n<2$. They concluded that the correlation length is about 30Å for insulin (with a r.m.s. amplitude of displacement of 0.25Å) and 50Å for lysozyme (with a r.m.s. amplitude of 0.11Å).

These studies suppose isotropic haloes. It is well known that in most cases this is not correct. More recently, A. Kolatkar et al. developed similar calculations using an anisotropic model for a yeast initiator tRNA crystal[17]. They concluded to the existence in the crystal of intermolecular correlations (rigid body motion) extending over distances as long as two unit cells (130Å).
All these studies have shown that many protein crystals display rigid body translations correlated over a few unit cells. Unfortunately, this kind of motions does not give direct information on the interesting "biological" displacements inside the protein.

3.2. INTRAMOLECULAR CORRELATIONS

3.2.1. Intramolecular short range correlations

The less structured c-type diffuse scattering on figure 1 has been interpreted as due to intramolecular displacements correlated over short distances. The authors and the investigated crystals are those mentioned before: Caspar et al. for insulin, Clarage et al. for lysozyme and Kolatkar et al. for t-RNA. Concerning lysozyme, the authors computed the diffuse scattering by means of the following formula:

\[ I(S) = \langle (2\pi S \delta_S)^2 \rangle \left\{ |F_0|^2 (S) \exp \left[ -(2\pi S \delta_S)^2 \right] \right\} \times \text{FT} \left( \exp \left[ -\frac{\pi}{\gamma_S} \right] \right) \]

where * denotes convolution and FT the Fourier Transform. \( F_0 \) is the ideal structure factor for the perfectly ordered crystal. The atomic r.m.s. displacement \( \delta_S \) is characterized by a decay length \( \gamma_S \).

They obtained \( \delta_S = 0.45 \, \text{Å} \) for insulin, \( \delta_S = 0.5 \, \text{Å} \) for the tetragonal form of lysozyme and 0.3 Å for the monoclinic form, with \( \gamma_S = 6 \, \text{Å} \) in all cases. It was concluded to the existence of intramolecular displacements, correlated over a short distance (6 Å), and described as "liquid-like motion". According to the authors, the main part of diffuse scattering should result from this type of motion. In particular, they assert that "any model with rigid body displacement of large domains (corresponding to low frequency normal modes)" would contradict the experimental data. This result will be discussed below.

A. Kolatkar et al. [16] made a more precise analysis in the case of the yeast initiator tRNA and got somewhat different results. The non-structured diffuse scattering observed from this crystal was split into distinct contributions and precise interatomic correlations were evidenced. Unlike the former models, part of the scattering was attributed to "independent motion of a large coherent unit, the anticodon regions of each molecule being displaced independently of all other molecules". Other parts of the scattering were attributed to a correlated movement of the same anticodon loop, polarized along the \( \mathbf{a} \) direction with an amplitude of about 0.8 Å. Finally, part of the remaining scattering was attributed to independent atomic movements (Einstein model), producing diffuse scattering with spherical symmetry, for which the position of the maximum depends exclusively on the amplitudes of the displacements. The agreement between the experimental and the simulated patterns is very good.

3.2.2. Intramolecular long range correlations

a) The analysis of Kolatkar et al. contradicts the conclusions of Clarage et al. about the non-existence of rigid body movements. It also shows how difficult it can be to identify which part of the molecule gives rise to the diffuse scattering process.

Fig. 3. — (a) Experimental diffuse scattering pattern of the orthorhombic form of lysozyme. —(b) Diffuse scattering simulation using the trajectories of the 15 lowest frequency normal modes [18].
This is the reason why we decided to lean on normal mode and/or molecular dynamics calculations so as to find the possible trajectories of atoms or correlation blocks. We worked on the orthorhombic form of lysozyme and showed that the very diffuse (cloudy) scattering could be due to low frequency vibrationnal modes. Figure 3a and 3b present respectively the experimental data and the result of normal mode simulation. Details and conditions of calculations were published elsewhere by Faure Ph et al. [18].

The agreement between model and experiment is rather acceptable and contradicts again the assertion of Clarage et al.. Actually the obviously too structured diffuse scattering due to one mode alone, such as "hinge-bending", is smoothed by the contribution of the fourteen other modes of lowest frequency.

b) We reached an even better agreement with experience including a diffuse scattering component arising from rigid body motion of the isolated molecules. This point is also particularly true for the tetragonal form of lysozyme. Indeed, an important part of the diffuse scattering observed with this form (figure 4a) can be interpreted with only the hypothesis of rigid body motion (figure 4b). A more complete calculation including normal mode technique as well as rigid body motion is now being pursued.

Fig. 4. — (a) Experimental diffuse scattering pattern of the tetragonal form of lysozyme. — (b) Calculated diffuse scattering pattern assuming only independent "rigid body translations" of the molecules in the unit cell.

4. DISCUSSION

We shall essentially examine two points:
1- the coherence between the analysis drawn from normal mode technique and rigid body translations and the previous results about the thermal factors B.
2- the possible coherence between our results and those of Caspar and Clarage.

4.1. - Coherence between B-factor analysis and normal mode technique coupled with molecular rigid body motions

The derivation of the thermal factors from normal mode analysis has been examined by several authors as, for example, Levitt et al.[19], Diamond [20], or Kidera and Go [21]. It was shown that agreement could be achieved between the B-factors evaluated from crystallographic data and the B-factors calculated by that method when taking into account only the 8 or 10 lowest frequency modes. In other respects, Sternberg et al.[22], Schomaker et al.[22], or J. Kuriyan et al. [23], have shown that, in certain cases, most of the B-factors values could be explained by the very simple hypothesis of molecular rigid body motion. This hypothesis was poorly exploited. Diffuse scattering seems to give it a new impulse.
4.2. - Is there any coherence between normal mode and "liquid like" analyses?

We have shown that diffuse scattering observed with orthorhombic lysozyme could be satisfactorily accounted for by low frequency vibration modes, involving a large number of atoms, thus corresponding to a correlation length quite higher than the 6Å proposed by Caspar and Clarage. Although the investigated crystals were not the same, it seems likely to us that this result is rather general and will probably be corroborated in many other cases.

Moreover, a quite simple estimation points out that, if considering a volume V containing N atoms, the diffuse scattering \(I_V\) due to a correlated displacement of all the atoms is roughly \(p\) times more intense than the diffuse scattering \(I_p\) due to \(p\) correlated groups (of volume \(V/p\)). Taking for \(V\) the approximate volume of one lysozyme molecule (about \([30\text{Å}]^3\)), and for \(V/p\) the typical volume of correlation for liquid-like motion (about \([6\text{Å}]^3\)), we obtain the intensity ratio:

\[
\frac{I_V}{I_p} = \frac{N^2}{p(N/p)^2} = p = 125
\]

The diffuse scattering produced by collective low frequency modes, involving a high number of atoms, is thus quite more intense than that produced by "more local" modes, involving only a limited number of atoms. Obviously, we do not assert that liquid-like movements should not exist, but we consider that the diffuse scattering they produce is likely to be hidden by that due to low frequency normal modes.

5. CONCLUSION

In addition to its potentialities for the analysis of the intermolecular displacements, X-ray diffuse scattering can be considered as a good probe for the study of the intramolecular correlated displacements in a crystallized protein. This method is sensitive enough to determine whether a given low frequency mode, as calculated by normal modes analysis, is active or not in the protein. Inversely the diffuse scattering technique can be used to test the validity of the dynamical simulations, either normal modes or molecular dynamics calculations, and their dependence to the initial conditions and hypotheses. The coupling between the diffuse scattering analysis and the molecular dynamics simulation will be in the near future a powerful tool for the study of intramolecular motion in crystallized proteins.

Bibliography

[3] Faber H.R. and Wiirthrich K., N2 4.2. - Is there any coherence between normal mode and "liquid like" analyses ?

We have shown that diffuse scattering observed with orthorhombic lysozyme could be satisfactorily accounted for by low frequency vibration modes, involving a large number of atoms, thus corresponding to a correlation length quite higher than the 6Å proposed by Caspar and Clarage. Although the investigated crystals were not the same, it seems likely to us that this result is rather general and will probably be corroborated in many other cases.

Moreover, a quite simple estimation points out that, if considering a volume V containing N atoms, the diffuse scattering \(I_V\) due to a correlated displacement of all the atoms is roughly \(p\) times more intense than the diffuse scattering \(I_p\) due to \(p\) correlated groups (of volume \(V/p\)). Taking for \(V\) the approximate volume of one lysozyme molecule (about \([30Å]^3\)), and for \(V/p\) the typical volume of correlation for liquid-like motion (about \([6Å]^3\)), we obtain the intensity ratio:

\[
\frac{I_V}{I_p} = \frac{N^2}{p(N/p)^2} = p = 125
\]

The diffuse scattering produced by collective low frequency modes, involving a high number of atoms, is thus quite more intense than that produced by "more local" modes, involving only a limited number of atoms. Obviously, we do not assert that liquid-like movements should not exist, but we consider that the diffuse scattering they produce is likely to be hidden by that due to low frequency normal modes.

5. CONCLUSION

In addition to its potentialities for the analysis of the intermolecular displacements, X-ray diffuse scattering can be considered as a good probe for the study of the intramolecular correlated displacements in a crystallized protein. This method is sensitive enough to determine whether a given low frequency mode, as calculated by normal modes analysis, is active or not in the protein. Inversely the diffuse scattering technique can be used to test the validity of the dynamical simulations, either normal modes or molecular dynamics calculations, and their dependence to the initial conditions and hypotheses. The coupling between the diffuse scattering analysis and the molecular dynamics simulation will be in the near future a powerful tool for the study of intramolecular motion in crystallized proteins.

Bibliography