Eur. Phys. J. Special Topics **208**, 15–20 (2012)

DOI: 10.1140/epjst/e2012-01603-6

THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS

Review

The prominent role of resonant elastic scattering for solving the X-ray structure of macromolecules

R. Kahn[1,*], E. Girard[1,a], and R. Fourme[2]

[1] IBS (CEA, CNRS, UJF), 41 rue Jules Horowitz, 38027 Grenoble, France
[2] Synchrotron Soleil, BP. 48 Saint Aubin, 91192 Gif-sur-Yvette, France

Received 25 December 2011 / Received in final form 23 March 2012
Published online 15 June 2012

Abstract. This article is a short overview of basic principles, methods and instrumentation of phasing methods based on anomalous scattering for X-ray macromolecular crystallography (MX).

1 Principle of experimental phasing methods

The inverse problem of X-ray crystallography, i.e. estimating structure factor phases, is particularly acute in the case of macromolecular crystals where scattering units may include several thousand atoms, so that direct methods are generally not applicable [1]. Molecular replacement [2] is a very effective phasing method, albeit applicable only when a model of the whole molecule (or at least a significant fraction) is available. Phasing without prior information (*ab initio* phasing) is generally performed by experimental methods based on measurements of diffracted intensities. Historically, the first *ab initio* method was isomorphous replacement. The Multiple Isomorphous Replacement (MIR) method requires crystals of the « native » structure and of several so-called « heavy atom derivatives » which differ from the native structure only by the presence of a few high-Z atoms [3]. Ideally, except heavy atoms, all these structures should be strictly identical (same unit cell, same light atom structure); if this condition is met, isomorphism is perfect. MIR requires measurements on native and at least two crystals of different derivatives. There are two steps in phasing. First, the heavy atom substructure of each derivative is determined: this is feasible because the scattering by heavy atoms is very strong, in spite of the background scattering from many low-Z atoms. Then the determination of the phase of the structure factor contribution from the main structure can proceed, taking the heavy atom substructures

*Richard Kahn presented an invited oral communication on anomalous scattering and macromolecular crystallography at REXS2011, a field of research where Richard had an eminent contribution. He could not write the associated article, as he died suddenly in october 2011. As close collaborators and friends of Richard, Eric Girard and Roger Fourme wrote this account dedicated to his memory.

[a] e-mail: eric.girard@ibs.fr

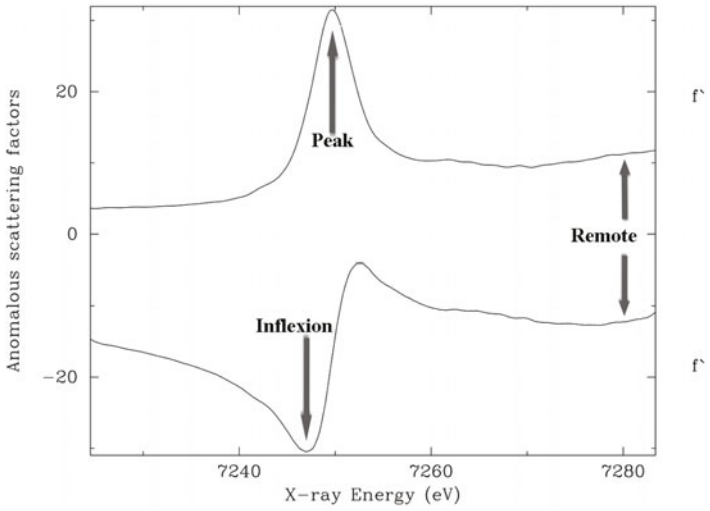


Fig. 1. Variation of the real (f') and imaginary (f'') components of the scattering factor of gadolinium near the L_{III}-absorption edge.

as references. The MIR is a powerful and general method, but has several drawbacks. The insertion of different heavy atoms may be tedious and time-consuming and non-isomorphism resulting from the insertion of heavy atoms is the rule rather than the exception. The solution of each substructure may be difficult, because it is not possible to separate the diffraction by the substructure from the diffraction by the native structure: only approximate values are obtained for substructure structure factors, so that the application of direct methods for solving substructures may be difficult. Finally, the use of several crystals adds complexity to data collection and is a source of experimental errors. MAD (Multiwavelength Anomalous Diffraction) is a variant of MIR which exploits the diffraction from only one crystal containing a substructure of anomalous scattering atoms [4]. Data collection proceeds at several wavelengths chosen close to one of the absorption edges of the substructure atoms. The atomic scattering factor of substructure atoms can be expressed as $f(\lambda) = f^\circ + f'(\lambda) + f''(\lambda)$ where f° is the normal scattering factor (i.e. the Fourier transform of the atomic electron density); $f'(\lambda)$ and $f''(\lambda)$ are the real and imaginary components respectively of resonant scattering. Variations of $f'(\lambda)$ and $f''(\lambda)$ are largest close to an absorption edge (Fig. 1). In such conditions, both amplitude and phase of the wave scattered by the substructure vary substantially with wavelength so that a change of wavelength is equivalent to a change of heavy atom, but the position of centers as well as the crystal content remains invariant. In practice, data are collected at three wavelength chosen at the maximum value of f'', at the inflection point of f' and rather far from the edge respectively.

Small intensity differences are measured. Main defaults of MIR are suppressed: only one crystal is required; isomorphism is perfect, and the structure factor of the substructure can be extracted without approximation [5], so that direct methods of phasing can be used to solve a substructure with more than 200 atoms. The calculated structure factor of the substructure is the pivot for the determination of the phase of the main structure. When anomalous dispersion is significant, reflections, which have normally the same intensity (Friedel pairs with Miller indices hkl and -h-k-l, or equivalent reflections in the crystal space-group), fall into two classes with different intensities: a pair of such reflections with one member in each class is called an anomalous pair. As intensity variations associated to wavelength changes are generally

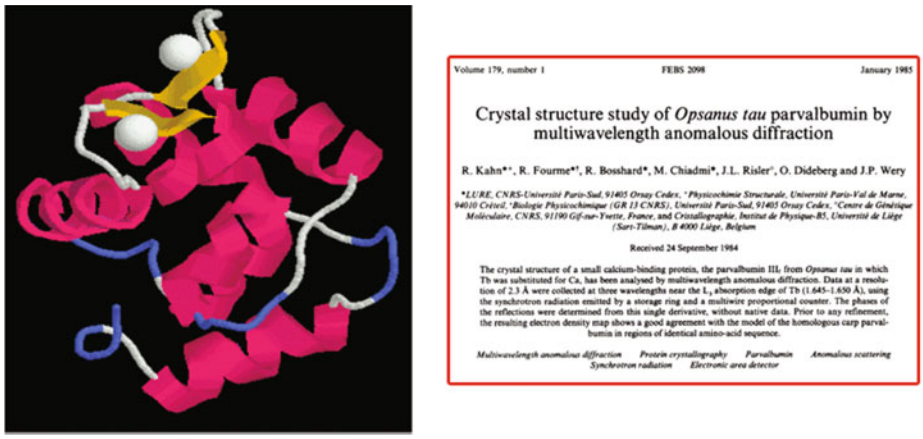
weak, the goal of data collection is the accurate measurement of intensity differences between mates in anomalous pairs. With the help of anomalous pairs, measurements at two wavelengths are sufficient for an unambiguous phase determination by the MAD method. The typical variation of f″ is $5e^-$ at a K edge, $10e^-$ at a L edge and can be in excess of 30 electrons for the L_{III} edges of lanthanides, producing a high phasing power. Many heavy (or not-so-heavy) elements can be used as anomalous centers: either intrinsic centers (such as metals in metalloproteins) or extrinsic centers, purposely introduced in the main structure. In this latter category, the most popular species is by far selenium, which can be substituted to intrinsic sulfur atoms in cysteine or methionine residues by bio-engineering [6,7]; in this case, data are collected at 2 or 3 wavelengths near the Se K edge ($\lambda = 0.98$ Å). Heavy atoms commonly used for MIR can also be used for MAD (for instance Hg at the L_{III} absorption edge, $\lambda = 1.01$ Å). As mentioned previously, lanthanides are excellent extrinsic heavy atoms due to a giant resonance at the L_{III} edge. With measurements of anomalous pairs at a single wavelength only, the phase is not uniquely determined and can take two values. This ambiguity may be solved by injecting additional phasing information from non-crystallographic symmetry or solvent flattening (a method where regions of the electron density identified as solvent are assigned a constant density before Fourier transformation): this is the SAD method. Sulfur is found in many proteins, and can be used for phasing [8]. Extremely accurate and redundant measurements are required to exploit the weak anomalous signal of this species. SAD, and especially sulfur-SAD, is increasingly popular, thanks to progress in instrumentation and pixel detectors.

2 Instrumentation and milestones

To make MAD phasing, the wavelength tunability of the X-ray beam is crucial, and a very bright source is necessary in order to perform multiple data acquisitions quickly. Besides a synchrotron radiation source and a good monochromator, the most important item for MAD data collection is an area detector featuring high detective quantum efficiency (DQE) at the required wavelengths, a low intrinsic noise, a large dynamical range and a fast readout. The development of MAD methods is indeed parallel to the progress of this instrumentation, in addition to progress in software. The first MAD phasing of a protein structure was achieved at the French synchrotron facility LURE at Orsay, and published in 1985 by R. Kahn et al. [9].

The protein was a parvalbumin extracted from the swim bladder of *Opsanus Tau*, a fish living in deep waters. The two Ca^{2+} ions in the so-called EF hand of the protein (Fig. 2) were substituted by lanthanide ions (Tb^{3+}). Data collection was performed at three wavelengths around the Tb L_{III}-edge and phases were derived by a MIR-like statistical method. This experiment was challenging as it combined synchrotron radiation, a two-crystal monochromator and a counting, noise-free, detector. This detector, called Penelope I, was the first product of a collaboration with Georges Charpak at CERN aiming at the fabrication of a parallax-free multiwire proportional chamber (MWPC) for protein crystallography. A few years later, Penelope II [10], mounted on the diffractometer shown in Fig. 3, was used for the determination of an important protein structure using MAD phasing based on anomalous scattering from Ho ions [11]. This work was another milestone as it demonstrated the potential of MAD phasing based on strong anomalous variations and very accurate data.

Wayne Hendrickson introduced the use of selenium for phasing in several steps. The first one was the solution of the streptavidin structure. As this protein has a high affinity for the small molecule biotin, Se-biotin was used as a vehicle to introduce Se



Volume 179, number 1 FEBS 2098 January 1985

Crystal structure study of *Opsanus tau* parvalbumin by multiwavelength anomalous diffraction

R. Kahn*', R. Fourme*', R. Bosshard*, M. Chiadmi*, J.L. Risler', O. Dideberg and J.P. Wery

*LURE, CNRS-Université Paris-Sud, 91405 Orsay Cedex, 'Physicochimie Structurale, Université Paris-Val de Marne, 94010 Créteil, 'Biologie Physicochimique (GR 13 CNRS), Université Paris-Sud, 91405 Orsay Cedex, 'Centre de Génétique Moléculaire, CNRS, 91190 Gif-sur-Yvette, France, and Cristallographie, Institut de Physique-B5, Université de Liège (Sart-Tilman), B 4000 Liège, Belgium

Received 24 September 1984

The crystal structure of a small calcium-binding protein, the parvalbumin IIIb from *Opsanus tau* in which Tb was substituted for Ca, has been analysed by multiwavelength anomalous diffraction. Data at a resolution of 2.3 Å were collected at three wavelengths near the L_3 absorption edge of Tb (1.645–1.650 Å), using the synchrotron radiation emitted by a storage ring and a multiwire proportional counter. The phases of the reflections were determined from this single derivative, without native data. Prior to any refinement, the resulting electron density map shows a good agreement with the model of the homologous carp parvalbumin in regions of identical amino-acid sequence.

Multiwavelength anomalous diffraction *Protein crystallography* *Parvalbumin* *Anomalous scattering* *Synchrotron radiation* *Electronic area detector*

Fig. 2. Structure of the protein parvalbumin. Grey balls are Ca ions which were replaced by Tb ions for MAD phasing [9].

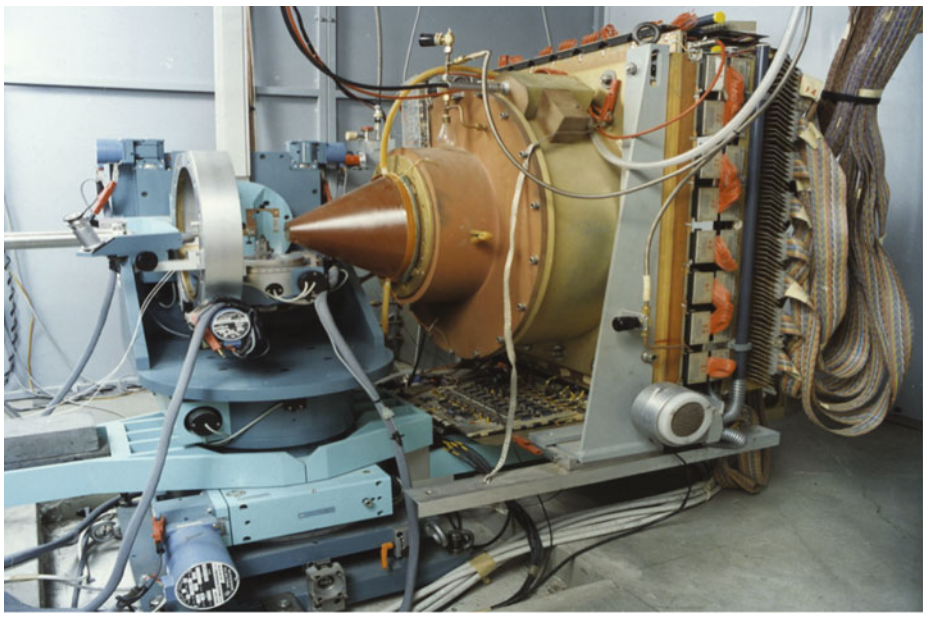


Fig. 3. Diffractometer on the D23 synchrotron beamline at LURE, showing the goniometer and the MWPC detector PENELOPE II.

in the protein, and the structure of the complex was solved by MAD at the Se K-edge [12]. Later, a general method to introduce Se in selected protein residues was described, as already mentioned [6]. The use of other labels has also made many progresses. In particular, lanthanides are of particular interest for anomalous-based experiment. Compared to one selenium atom, and assuming that the diffraction data are collected at the respective absorption edge, a lanthanide atom will allow a protein that is nine times larger to be phased [13]. As mentioned, lanthanide ions were used in early MAD studies [9,11] as they can substitute for Ca^{2+}. However, incorporation of lanthanides by using lanthanide salts often damages protein crystals, owing to the preferred nine-based coordination of lanthanide ions. One solution to overcome this problem is to use lanthanide complexes made of a ligand that surrounds the

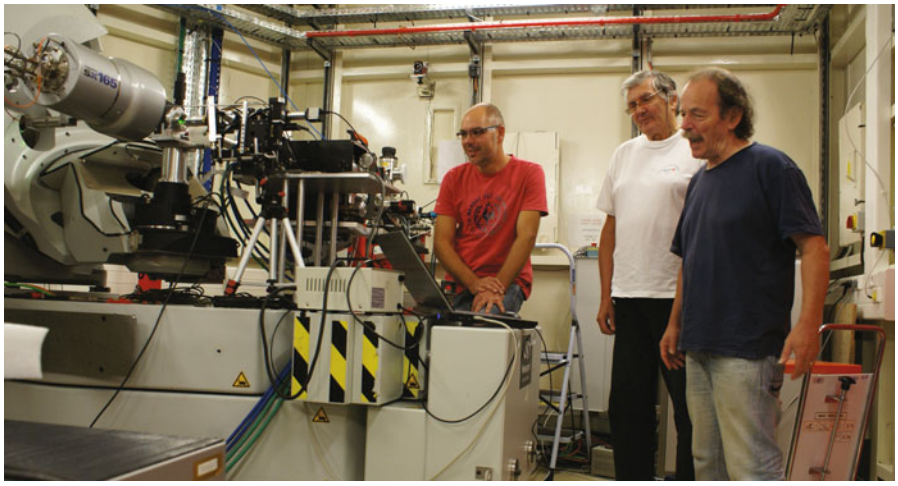


Fig. 4. Richard was also one of the pioneers of high pressure macromolecular crystallography since 2000. The photograph shows our team during a high pressure data collection on the CRISTAL beamline at SOLEIL (25 September 2011).

lanthanide ions as a cage [14,15]. These complexes can be easily introduced in protein crystals and turn out to be efficient vehicles for phasing using anomalous dispersion, in particular for large macromolecular assemblies (for a short overview of protein structures determined using these lanthanide complexes, see [13]).

There are other applications of anomalous scattering in macromolecular crystallography. As anomalous scattering is element specific, it can be used to identify unambiguously the nature of particular atoms (e.g. metal atoms in a cluster). We also mention a parent method of MAD, called MASC, where anomalous centers are in the liquid phase of the crystal [16]. Changes in contrast resulting from wavelength-dependent scattering can be exploited to get the molecular envelope, which may be an important step for the resolution of the structure of large macromolecular systems.

3 Conclusion

Today, developments in instrumentation and software on modern synchrotron radiation Macromolecular Crystallography (MX) beamlines make it possible to collect diffraction data on cryocooled crystals containing anomalous scattering centers and derive the molecular skeleton in a matter of a few hours, and in some cases, of a few minutes. Modern detectors, in particular counting Si-pixel detectors, provide highly accurate data, which support the development of SAD phasing instead of MAD, including Sulfur-SAD. Overall, the progress since the pioneering works is extraordinary, making systematic and high-throughput structure determinations required for structural genomics possible. Resonant elastic scattering has been a major player in this revolution.

References

1. H. Hauptman, Curr. Opin. Struct. Biol. **7**, 672 (1997)
2. J. Drenth, *Principles of protein X-ray crystallography* (Springer, 1999)
3. M. Perutz, Acta Cryst. **9**, 867 (1956)

4. J. Karle, Int. J. Quant. Chem.: Quantum Biol. Symp. **7**, 357 (1980)
5. W.A. Hendrickson, Transactions ACA **21**, 11 (1985)
6. W.A. Hendrickson, J.R. Horton, D.M. LeMaster, EMBO J. **9**, 1665 (1990)
7. S. Doublie, C.W. Carter Jr, in *Crystallization of nucleic acids and proteins. A practical approach*, edited by A. Ducruix and A. Giege (Oxford University Press, Oxford, 1992)
8. W.A. Hendrickson, M.M. Teeter, Nature **290**, 107 (1981)
9. R. Kahn, R. Fourme, R. Bosshard, M. Chiadmi, J.L. Risler, O. Dideberg, J.P. Wery, FEBS Letters **179**, 133 (1985)
10. R. Kahn, R. Fourme, R. Bosshard, V. Saintagne, Nucl. Instrum. Meth. **246**, 596 (1986)
11. W.I. Weis, R. Kahn, R. Fourme, K. Drickamer, W.A. Hendrickson, Science **254**, 1608 (1991)
12. W.A Hendrickson, A. Pahler, J.L. Smith, Y. Satow, E.A. Merritt, P.M Phizackerley, Proc. Natl. Acad. Sci. USA **86**, 2190 (1989)
13. R. Talon, R. Kahn, M.A. Durá, O. Maury, F.M.D. Vellieux, B. Franzetti, E. Girard, J. Synchrotron Rad. **18**, 74 (2011)
14. E. Girard, M. Stelter, J. Vicat, R. Kahn, Acta Cryst. D **59**, 1914 (2003)
15. G. Pompidor, A. D'Aléo, J. Vicat, L. Toupet, N. Giraud, R. Kahn, O. Maury, Angew. Chem. Int. Ed. **47**, 3388 (2008)
16. R. Fourme, W. Shepard, R. Kahn, G. L'Hermite, I. Li de La Sierra, J. Synchrotron Rad. **2**, 36 (1995)