Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Characterization of crystal content by ESI–MS and MALDI–MS

N. Potier,[a]† V. Lamour,[b]† A. Poterszman,[b] J. C. Thierry,[b] D. Moras[b] and A. Van Dorsselaer[a]*

[a]Laboratoire de Spectrométrie de Masse Bio-Organique, 1 Rue Blaise Pascal, 67008 Strasbourg, France, and [b]Institut de Génétique et de Biologie Moléculaire et Cellulaire, 1 Rue Laurent Fries, BP163, 67404 Illkirch CEDEX, C.U. de Strasbourg, France

† N. Potier and V. Lamour contributed equally to this work.

Correspondence e-mail: vandors@chimie.u-strasbg.fr

Received 22 February 2000
Accepted 24 July 2000

A general approach based on mass spectrometry is described for the rapid identification of the content of macromolecular crystals. The experimental procedure was established using lysozyme crystals and then successfully applied to various systems containing specifically bound molecules not easily detectable by other classical techniques. This procedure can be carried out on crystals containing macromolecules of a different nature, such as proteins, nucleic acids and small organic molecules and their non-covalent complexes, grown under various crystallization conditions including PEGs and salts. It can be applied very early on in the crystallization process – as soon as the crystals can be handled. It allows the biologist to control precisely the sequence integrity and homogeneity of the crystallized proteins (in particular at the C-terminus) as well as to verify whether the protein has crystallized with all its expected partners or ligands (nucleic acid molecules, cofactor or small organic molecules).

1. Introduction

Knowledge of the three-dimensional structures of macromolecules is important in understanding their biological functions. It is now widely recognized that a bottleneck in crystallographic structure determination is often the growth of well ordered single crystals containing the molecule of interest. Many methods have been engaged to characterize the species to be crystallized. The homogeneity of the sample can be checked using mass spectrometry. The aggregation state of the solution (monodispersity/polydispersity) can be analysed using dynamic light scattering (Ferré-D'Amaré, 1997). The existence of non-covalent complexes and/or the oligomeric state of the molecules can be investigated using techniques such as gel-filtration chromatography, scattering methods, mass spectrometry, ultracentrifugation *etc* (Ducruix & Giegé, 1999). However, very few techniques are so far available to characterize the crystal content. Indeed, when crystals are observed in the crystallization droplet, the following questions arise.

(i) Did the protein remain intact throughout the course of a prolonged crystallization experiment? Gel electrophoresis allows the presence of a protein in a crystal to be checked and its molecular weight to be estimated. In combination with Edman degradation it allows the user to determine whether the N-terminus of the protein remains intact. However, the mass-measurement accuracy and the resolution obtained by SDS–PAGE are not sufficient to check whether the protein is homogeneous at its C-terminus, especially in the presence of a mixture of degraded proteins. C-terminal sequencing using carboxypeptidase is difficult and would require many crystals.

Table 1
Representative examples of crystals analyzed by mass spectrometry.

The main precipitating agent of the crystallization droplet is indicated for each analysis. ESI mass spectra could be obtained for each macromolecular crystal either in positive (ES^+) or negative (ES^-) ionization mode.

Crystal	Main crystallization agent	Number and size of crystals (μm)	Washings	Results	Comments
RXR–9-*cis* RA (personal communication)	20% propanediol, 10% glycerol	4 crystals, 100 × 10 × 10	2 in mother liquor; 2 in 3 *M* NH_4OAc; 0.05 *M* NH_4OAc	Denaturing conditions (ES^+), monomers; denaturing conditions (ES^-), monomers + ligand	Detection of a small and non-ionic ligand
hAR–NADP (Lamour *et al.*, 1999)	10% PEG 6000	1 crystal, 100 × 150 × 500	2 in mother liquor; 1 in 1 *M* NH_4OAc; 1 in 0.5 *M* NH_4OAc; 0.025 *M* NH_4OAc	Denaturing conditions (ES^+), AR monomer; denaturing conditions (ES^-), NADP cofactor; native conditions (ES^+), AR–NADP complex	Observation of non-covalent protein–ligand complex
TBP–ADN (Nikolov *et al.*, 1996)	10% glycerol	1 crystal, 150 × 150 × 10	2 in mother liquor; 2 in 0.1 *M* NH_4OAc; H_2O	Denaturing conditions (ES^+), protein; denaturing conditions (ES^-), observation ADN	Detection of a small oligonucleotide
Arginyl-tRNA synthetase (personal communication)	2.4 *M* ammonium sulfate	1 crystal, 400 × 200 × 200	2 in mother liquor; 2 in 7.5 *M* NH_4OAc; H_2O	Denaturing conditions (ES^+), intact protein	Crystallization under conditions very unfavourable for MS; sequence-integrity verification
Stromelysin-3 (personal communication)	0.5 *M* ammonium sulfate	7–10 small crystals, 20 × 20 × 50	2 in mother liquor; 2 in 7.5 *M* NH_4OAc; H_2O	Denaturing conditions (ES^+), intact protein	Sequence-integrity verification
ER–E (personal communication)	15% PEG 8000	2 crystals, 50 × 50 × 100	2 in mother liquor; 1 in 1 *M* NH_4OAc; H_2O	Denaturing conditions (ES^+), N-terminal and C-terminal degraded protein	Sequence-integrity verification

(ii) How can crystals containing protein–nucleic acid (DNA or RNA) complexes be distinguished from crystals containing only the protein or the nucleic acid? Gel electrophoresis does not help in the case of small oligonucleotides. The use of several dyes such as methylene blue or methyl violet has been reported to discriminate DNA–protein from protein crystals, but small crystals are often faintly stained and it is difficult to unambiguously determine whether they contain DNA (Wilkosz *et al.*, 1995).

(iii) Did the protein co-crystallize with its ligand? Gel electrophoresis does not give any information about small molecules bound to the protein. In the course of screening synthetic ligands, verifying their presence in the crystal can be useful before optimization of the crystallization conditions and the start of the structure determination. In other cases, proteins can co-crystallize with a foreign compound which gives an unexpected signal in the electron-density map which needs to be identified.

The development in mass spectrometry of efficient soft ionization methods such as electrospray (ESI–MS; Mann & Wilm, 1995; Fenn *et al.*, 1989) or matrix-assisted laser desorption ionization (MALDI–MS; Karas & Hillenkamp, 1988) has led to a recent breakthrough in the molecular characterization of biological systems. These techniques have been widely shown to be able to characterize and evaluate the purity of protein or oligonucleotide preparations. Detailed protein-sequence information is now easily available by mass peptide mapping of proteolytic digests (Klarskov *et al.*, 1994; Kim *et al.*, 1994). Mass spectrometry is also routinely used for the identification and localization of post-translational modifications as well as for full or partial sequence determination using tandem-MS (Biemann & Scoble, 1987; Mann & Wilm, 1994).

In addition to these now classical applications of molecular mass spectrometry, a new field is emerging where mass spectrometry is used to measure the mass of non-covalent complexes. This approach defined the field of supramolecular mass spectrometry. The feasibility of ESI–MS for probing higher order (tertiary and quaternary) structures has already been well illustrated by a series of examples (Loo, 1997). By the control of several parameters (buffer, pH, voltages *etc.*), specific non-covalent complexes such as protein–protein (Light-Wahl *et al.*, 1994; Green *et al.*, 1999), protein–ligands (Rogniaux *et al.*, 1999; Bruce *et al.*, 1998), protein–DNA (Cheng *et al.*, 1996; Potier *et al.*, 1998), DNA–DNA *etc.* can be studied. Mass spectrometry provides several unique advantages over other methods: (i) each constituent characterized by its molecular weight can be observed in a mixture, (ii) oligonucleotides or ligands can be detected easily in the case of non-covalent complexes either bound to the protein (under near-physiological conditions) or alone (under denaturing conditions) and (iii) a small amount of material (a few picomoles) is required for the analysis and enables the determination of molecular masses with good accuracy (0.01%).

Mass spectrometry is a powerful analytical technique which is however extremely sensitive to the presence of non-volatile salts and other compounds usually present in crystallization buffers. The use of mass spectrometry for analysing crystal-

lized material has already been reported in a study of lysozyme crystal impurities (Hirschler *et al.*, 1998). In this case, dilution of the dissolved crystal and collisionally induced dissociation (CID) were optimized to obtain precise mass measurement despite the presence of NaCl in the crystallization buffer. However, considering the diversity of the compounds used in macromolecules crystallization experiments (salts, PEGs, glycerol *etc.*), we needed a procedure to characterize the crystal content in a more general way.

Here, we report a novel approach for rapid identification of crystal content based on mass spectrometry. Lysozyme was first chosen as a model for establishing a protocol that has been successfully applied to the analysis of crystals containing non-covalent protein–ligand complexes. These include crystals composed of TATA box binding protein–DNA, aldose reductase–NADP and nuclear receptor–ligand complexes.

2. Material and methods

2.1. Crystals

Lyzozyme crystals were obtained by vapour diffusion after equilibration of a drop containing 20 mg ml^{-1} lyzozyme, 1 *M* NaCl and 0.05 *M* sodium acetate pH 4.5 against a 1 ml reservoir containing 2 *M* NaCl and 0.1 *M* sodium acetate pH 4.5 (Ducruix & Giegé, 1999). The other crystals used in this study were generous gifts from members of the laboratory (Table 1).

2.2. Sample preparation

2.2.1. Solvent exchange. The transfer of the crystals from the crystallization droplet to the washing droplet is carried out using a cryoloop (20 µm wire, diameter ranging from 0.05 to 0.7 mm depending on the size of the crystals; Hampton Research). The crystals were systematically washed twice in 10 µl of mother liquor to remove the protein remaining in solution or on the surface of the crystal, followed by *X* washings (*X* varying from 1 to 3) in 5 µl droplets containing *Y M* NH_4OAc pH 7 (*Y* varying from 0 to 12). The crystals were transferred from drop to drop after about 20 s or after a longer period if the crystal was able to withstand these conditions. The concentration of the NH_4OAc washing buffer (*Y*) as well as the number of washings (*X*) were adjusted to prevent crystal disintegration (see Table 1). For analysis under denaturing conditions, crystals were dissolved in 5 µl pure water and diluted in 1:1 H_2O/CH_3CN containing either 1% HCOOH (analyses in positive ionization mode) or 0.5% TEA (analyses in negative ionization mode). For the analyses in native conditions, crystals of hAR–NADP and RXR–9-*cis* RA complexes were dissolved after the last washing in a 5 µl droplet containing 0.025 and 0.05 *M* NH_4OAc pH 7, respectively, in order to preserve the specific non-covalent interactions (Potier *et al.*, 1997).

2.2.2. Evaluation of the sample concentration. For many protein crystals, V_M varies between 1.7 and 3.5 Å^3 Da^{-1}, with most values around 2.15 Å^3 Da^{-1} (Matthews, 1968). The amount of protein present in a crystal is given by the equation $m = v/V_M \times N_A$, where *m* is the quantity of protein present in the crystal in grams, *v* is the volume of the crystal in Å^3 and N_A is the Avogadro constant. As an example, using a mean value of V_M = 2.15 Å^3 Da^{-1}, the quantity of protein in a parallelepipedic crystal of dimensions 600 × 400 × 50 µm is about 9 µg. The molecular weights of the proteins studied in this work range from 20 to 70 kDa. The crystals are dissolved in 5 µl of the appropriate buffer and the sample concentration varies from 0.2 to 1 m*M*. This value is still 10–100 times higher than classical mass spectrometry conditions (1–100 µ*M*).

2.3. Mass spectrometry

2.3.1. Electrospray ionization. All studies were performed on a Bio-Q triple quadrupole mass spectrometer (Micromass, Manchester, England). Samples were dissolved in a 1:1 mixture of water and acetonitrile containing 1% HCOOH at a concentration of 5 µ*M* and were continuously infused into the ion source at a flow rate of 0.5 µl min^{-1} *via* a Harvard Model 11 syringe pump (Havard Apparatus). 0.025 *M* ammonium acetate pH 7 was used for the study of non-covalent complexes. Calibration to m/z = 3500 was performed in the positive ionization mode using the multiply charged ions produced by a separate introduction of egg-white lysozyme dissolved in pure water (MW = 14306 Da). Calibration in the negative ionization mode was performed using a separate introduction of an oligonucleotide (dCp_{12}, MW = 3488 Da) dissolved in a 1:1 mixture of water and acetonitrile and containing 0.5% TEA. CID–MS experiments were performed by increasing the cone voltage (Vc) which determines the kinetic energy of the ions when they collide with residual gaseous molecules in the interface region of the mass spectrometer. These experiments were used to evaluate the strength of non-covalent interactions in the gas phase in the case of the hAR–NADP and RXR–9-*cis* RA complexes.

2.3.2. MALDI. Mass spectra were acquired on a Biflex III mass spectrometer (Bruker, Bremen, Germany) in the linear mode. Ionization was accomplished with a 337 nm beam from a nitrogen laser with a repetition rate of 2 Hz. A mixture of standard peptide was used for external calibration (angiotensin II, MW = 1047.2 Da; ACTH fragment 18–39, MW = 2465.7 Da; bovin insulin, MW = 5734.6 Da). 1 µl of a saturated solution of cyano-4-hydroxycinnamic acid (4HCCA) in acetone was deposited on the probe tip and allowed to spread and evaporate quickly. 0.5 µl of the sample was directly loaded on this thin layer of matrix and covered with 0.2 µl of a second matrix solution consisting of 4HCCA at 7 mg ml^{-1} in a 1:2 mixture of water and acetonitrile (Vorm *et al.*, 1994). After air drying, this preparation was rinsed once with 1 µl of 0.2% TFA which was flushed away after a few seconds.

3. Results

A rough estimation of the amount of protein contained within a crystal (see §2) indicated that in principle it should be possible to analyse its content by mass spectrometry. However, owing to the presence of non-volatile compounds in the crystallization buffer, crystals cannot be directly analysed.

Indeed, the presence of non-volatile salts (even at trace level) has a large impact on the quality of the mass spectra and can even prevent any ion detection. Heterogeneous salt attachment (mainly Na^+ and K^+) dramatically decreases the signal intensity of the analyte as well as the resolution, since the total sample quantity is distributed over a multitude of species. As the mass differences between these species are small and because the resolution of the generally used quadrupole analysers is too low, it is not possible to detect them separately. This feature leads to an inaccurate mass measurement or to no mass measurement at all.

To allow measurement of their content, crystals have to be washed and transferred to a buffer compatible with mass-spectrometry experiments. At this stage, the choice of the washing buffer is critical. It has to be compatible with both the crystal stability (so that the crystal stays intact and does not lose any of its components) and the analytical technique, mass spectrometry. Ammonium acetate (NH_4OAc) appeared to be a good candidate. Its use in mass spectrometry has been widely illustrated, especially for the characterization of non-covalent complexes. Because of its high volatility, no NH_4OAc adducts are observed in the mass spectra, resulting in the obtaining of narrow multiply charged peaks with a good mass accuracy (0.01%). Furthermore, the use of NH_4OAc as a washing buffer for crystals has previously been reported (Petsko, 1975); in this case, crystals were washed in NH_4OAc before transfer in a cryoprotecting buffer and freezing with liquid nitrogen.

4. Description of the procedure: characterization of a lysozyme crystal by ESI–MS and MALDI–MS

Lysozyme has often been chosen as a model both in mass spectrometry and crystallogenesis. In order to establish the experimental conditions allowing accurate mass measurements, crystals obtained in 2 *M* NaCl were used. Such a salt concentration is commonly used for crystallization experiments but constitutes a difficult case for mass spectrometry. We determined the NH_4OAc concentration at which the crystal could be washed without melting by trial and error. Fig. 1 shows the evolution of a lysozyme crystal (300 × 300 × 50 µm) during successive washings (*a*) in the mother liquor, (*b*) in 1 *M* NH_4OAc, (*c*) in 0.5 *M* NH_4OAc, (*d*) in 0.1 *M* NH_4OAc. When 0.1 *M* NH_4OAc was used for washing, the crystal cracked and ultimately disintegrated (Fig. 1*d*). The ionic strength difference between 2 *M* NaCl (the crystallization buffer) and 0.1 *M* NH_4OAc (the washing buffer) was presumably too strong. In contrast, the crystal could be transferred without damage in 1 *M* NH_4OAc (Fig. 1*b*), which allowed us to carry out several efficient washings.

4.1. Analyses by ESI–MS

The ESI mass spectra obtained for a lysozyme crystal (330 × 300 × 50 µm) after three washings in 1 *M* NH_4OAc and a 100-fold dilution in 1:1 H_2O/CH_3CN, 1% HCOOH is shown in Fig. 2(*a*). A unique series of multiply charged ions corresponding to lysozyme is observed (measured mass, 14 306.8 ± 1.6 Da; expected mass, 14 306 Da) with an *m*/*z* ratio distribution identical to that obtained from a non-crystallized sample under the same analytical conditions (data not shown). A peak width of 10 Da for the 9-charged ion and a mass precision of 1.6 Da shows that the washings were very efficient for removing the 2 *M* NaCl used for the crystallization: no accurate mass measurement could in fact be achieved by ESI–MS when the lysozyme was dissolved in a solution containing as little as 10 m*M* NaCl. Moreover, no protein signal could be detected for the last washing droplet (even without any dilution of the solution), showing that very little material was lost during the various steps of washing (Fig. 2*b*) and proving that the signal observed in Fig. 2(*a*) could only arise from the crystallized protein.

Although two washings in the mother solution were always carried out (to remove the non-crystallized protein), it appeared that a single washing in NH_4OAc could be sufficient to obtain a signal. However, after a single washing the mass spectrometry peaks are broader, revealing a number of salt molecules non-specifically attached to the protein, presumably arising from the non-volatile salts present inside the crystal. This leads to an inaccurate measurement of the molecular mass. Increasing the number of washings in 1 *M* NH_4OAc enhances the quality of the mass spectrum in terms of mass precision, which is essential when the purpose of the analysis is to control the protein integrity.

4.2. Analyses by MALDI–MS

The characterization of biological samples can also be performed by MALDI–MS. MALDI is a good alternative method to ESI if only a small amount of material is available or in the case of very heterogeneous samples (owing to

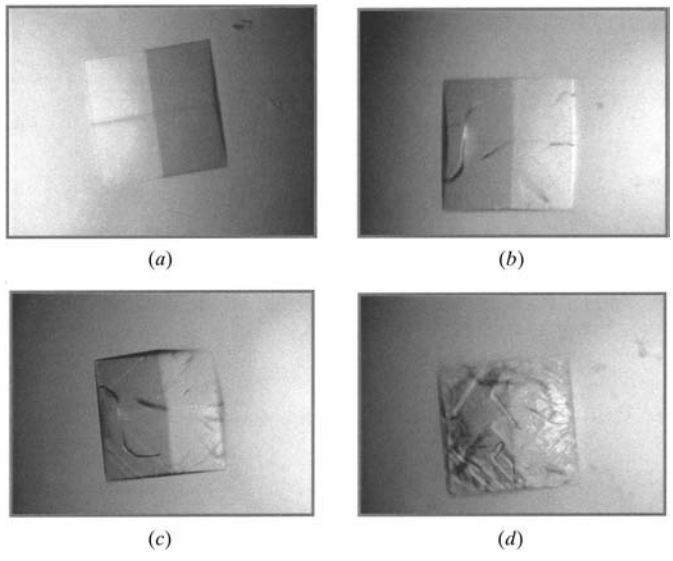

Figure 1
Evolution of a lysozyme crystal (330 × 300 × 50 µm) during washings in 10 µl of (*a*) mother liquor, (*b*) 1 *M* NH_4OAc, (*c*) 0.5 *M* NH_4OAc and (*d*) 0.1 *M* NH_4OAc. Below 0.1 *M* the crystal becomes soft and dissolves.

intensive degradation, for example). The MALDI mass spectrum obtained from the same lysozyme crystal as the one used for the ESI–MS analyses is shown in Fig. 2(*c*). 1 µl of the lysozyme solution prepared for the ESI–MS analyses was used, which corresponds to 1/500 of the crystal. The measured molecular mass of 14 310 Da agreed perfectly with the admitted experimental error of the technique (0.1%).

Some crystals are extremely fragile and cannot be manipulated in NH_4OAc. Considering the fact that MALDI–MS is less sensitive to the presence of salts than ESI–MS, we wondered if washing in NH_4OAc was inevitable prior to mass spectrometry. In order to test the feasibility of a direct measurement, a lysozyme crystal (150 × 50 × 10 µm) was deposited directly on the MALDI target immediately after two washings in the mother liquor. The matrix solution was deposited above the crystal and two washings of the target were performed with 0.2% TFA as described in §2. As shown in Fig. 2(*d*), lysozyme ions were detected. Although the peaks were three times broader than those obtained when the crystal was washed in 1 *M* NH_4OAc (Fig. 2*c*), a mass of 14 323 Da could be measured, reflecting a number of salt molecules still attached to the protein or trapped within the matrix pattern. Additionally, a peak at m/z = 9536 corresponding to the triply charged ions of a lysozyme dimer is detected. The desorption of lysozyme dimer is probably because of a concentration effect (the whole crystal containing at least 10 µg of protein) and results from ion-molecule reactions in the MALDI source. No trace of lysozyme dimer was observed when the solution obtained from the crystal was diluted 100-fold before analysis (Fig. 2*c*). In spite of a lower mass accuracy than with ESI–MS, the possibility of depositing a crystal directly on a MALDI target might be highly useful if the crystal does not withstand washings in NH_4OAc.

The protocol described above allowed us to detect the presence of a protein in a crystal with high reproducibility and accuracy. To determine whether this method could be extended to multi-component systems, we applied it to non-covalent protein–DNA and protein–ligand complex crystals.

5. Characterization of small oligonucleotides in a crystal of a non-covalent protein–DNA complex by ESI–MS and MALDI–MS

While the presence of the protein in a crystal can be quite easily identified by gel electrophoresis, unambiguous proof that the DNA has co-crystallized with the protein is not straightforward, especially in the case of small oligonucleotides. A non-covalent complex between the TATA box binding protein (TBP) and an autocomplementary oligonucleotide containing the specific binding DNA sequence has been crystallized and the resulting crystal analysed by MALDI–MS and ESI–MS. As the crystals were grown at low salt concentration, they could only be washed in 20 m*M* NH_4OAc.

The protein and the oligonucleotide present in the TBP–ADN crystals were detected by MALDI–MS using the matrix 4HCCA and 2,4,6-acetophenone, respectively (data not shown). However, although the mass of the protein could be precisely determined (20 781 Da), it was very difficult to obtain a reliable and reproducible mass measurement for the oligonucleotide.

The ESI-MS analysis was carried out on a 150 × 150 × 10 µm crystal. Under these mass-spectrometry conditions, the

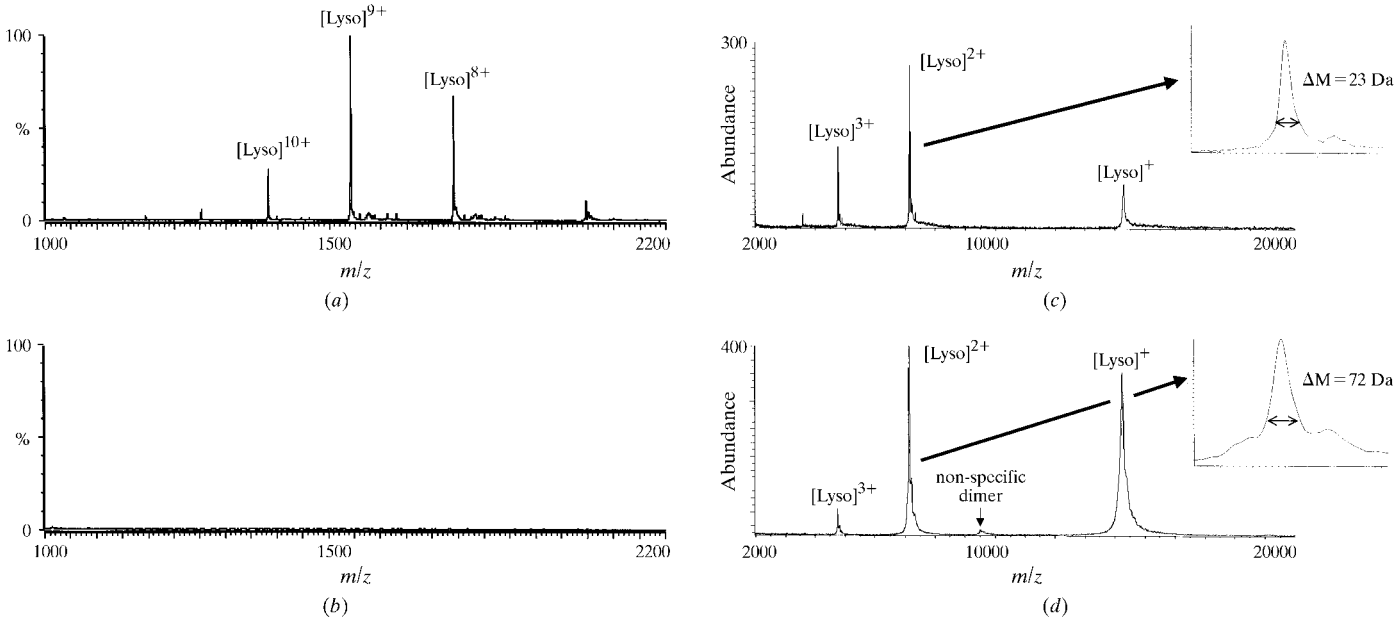

Figure 2
ESI–MS and MALDI–MS analysis of a lysozyme crystal. ESI mass spectra obtained from a single lysozyme crystal (*a*) after three washings in 1 *M* NH_4OAc and (*b*) from the last washing droplet. A mass of 14 306.8 ± 1.6 Da is measured. No signal corresponding to the lysozyme is detected in the last washing droplet, revealing that no material is lost during the washing in NH_4OAc. MALDI mass spectra obtained using a lysozyme crystal after three washings in 1 *M* NH_4OAc and (*c*) dissolution in water or (*d*) direct deposition on the MALDI target. Lysozyme ions are detected with these two preparations, but a lower mass accuracy is achieved if the crystal is not washed with NH_4OAc prior to its deposition on the target.

non-covalent complex is denatured and the two constituents are detected separately. Results obtained in positive ionization mode (protein) and negative ionization mode (oligonucleotide) are shown in Figs. 3(*a*) and 3(*b*), respectively. The measured masses (20 772.3 ± 1.9, 3644.2 ± 1.0 Da) correspond to the expected masses of the protein (20 775 Da) and of the oligonucleotide (3644.2 Da) detected as a single strand because measurements were performed under denaturing conditions (0.5% TEA).

6. Characterization of ligands in crystals containing a non-covalent protein–ligand complex by ESI–MS

Checking the presence of ligands in a crystal before structure determination still remains a difficult problem, especially if the ligand does not display any spectroscopic properties. Here, we present two examples where mass spectrometry appeared to be the only way to verify that the protein had crystallized with its specifically bound ligand.

In the first example, a crystal containing a human aldose reductase hAR–NADP non-covalent complex was characterized. Crystals corresponding to the holoenzyme were obtained in PEG 6000 (Lamour *et al.*, 1999). Results obtained on a 70 × 100 × 550 µm crystal under denaturing or native conditions are shown in Fig. 4. Under denaturing conditions (50% CH_3CN), the two components of the non-covalent complex are dissociated and detected separately in positive (Fig. 4*a*) and in negative (Fig. 4*b*) ionization mode. A mass of 36 137.2 ± 2.3 Da was measured in the positive mode, which corresponds to the expected mass of the protein (36 136 Da). A mass of 743 Da detected in the negative ionization mode confirms the presence of the NADP cofactor. No trace of the cofactor was detected in the last washing droplet, showing that the signal at 743 Da arose from NADP present in the crystal. The peak sharpness as well as the mass precision reveals that two washings in NH_4OAc are sufficient to remove the PEG 6000 precipitating agent as well as the mother-liquor salts (citrate buffer), since no signal of the protein could be detected when the crystal was directly dissolved from its crystallization droplet in H_2O/CH_3CN, 1% HCOOH. As for the preceding protein–DNA complex, analyses in both positive and negative mode are necessary to characterize all constituents of the crystal.

If the ligand is neutral, it cannot be detected separately from its receptor. In order to check if this procedure could be applied even to a non-ionic ligand, an hAR–NADP crystal was washed, dissolved in a non-denaturing solution (0.025 *M* NH_4OAc pH 6.5) and analysed. A mass of 36 880.1 ± 2.3 Da was measured (Fig. 4*c*). This mass corresponds to the expected mass of the holoenzyme (36 879 Da), revealing that the non-covalent hAR–NADP complex has resisted the washing and has remained intact during the electrospray process. Under the same conditions, the spectra obtained from the crystal and from the purified complex in solution (before crystallization)

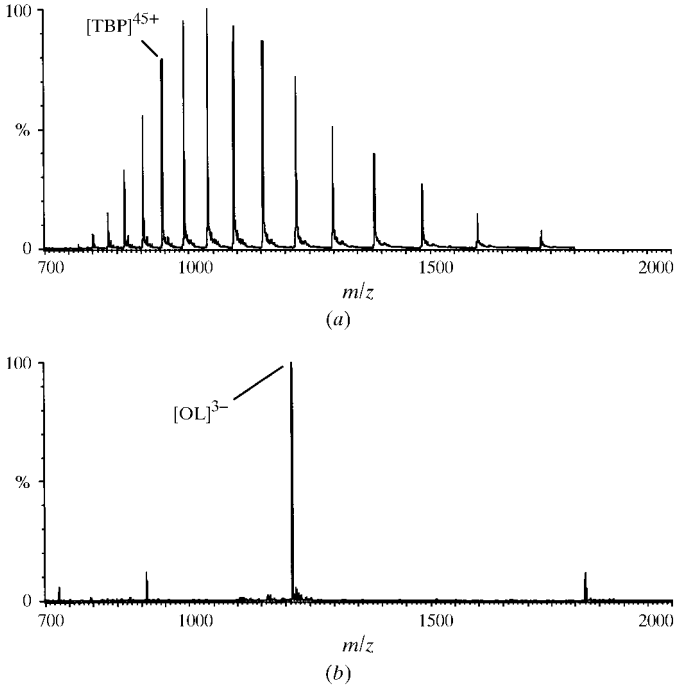

Figure 3
ESI mass spectra obtained from a TBP–DNA complex crystal after three washings in 20 m*M* NH_4OAc and a 100-fold dilution in (*a*) H_2O/CH_3CN, 1% HCOOH (positive ionization mode) and (*b*) H_2O/CH_3CN, 0.5% TEA (negative ionization mode). The analysis in negative ionization mode confirms the presence of the oligonucleotide (OL) in the crystal (MW = 3644.2 Da).

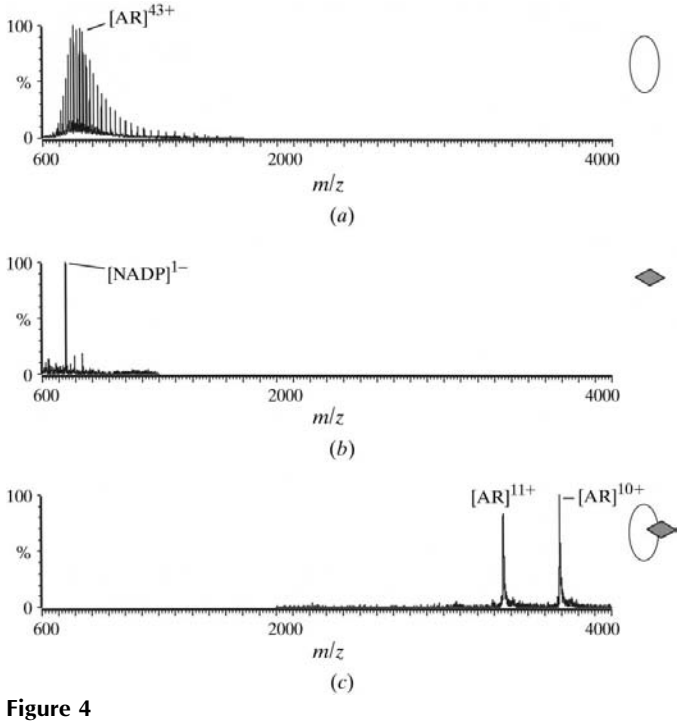

Figure 4
ESI mass spectrum obtained from a crystal containing a non-covalent hAR–NADP complex after two washings in 1 *M* and 0.5 *M* NH_4OAc successively and a 200-fold dilution in (*a*) H_2O/CH_3CN, 1% HCOOH, (*b*) H_2O/CH_3CN, 0.5% TEA, (*c*) 0.025 *M* NH_4OAc. The protein and the NADP cofactor are detected either separately under denaturing conditions (*a* and *b*) or still bound under native conditions (*c*), revealing that the non-covalent complex is stable during the various steps of crystallization, washing, dissolution and analysis by mass spectrometry.

are identical in terms of the charged-state distribution and these two complexes behaved identically with respect to collisionally induced dissociation (CID). In this particular case, increasing the cone voltage (Vc) did not lead to the dissociation of the non-covalent hAR–NADP complex, but to the breakage of a covalent bond of the NADP cofactor still fixed on the protein. 50% of covalent breaking was obtained at Vc = 120 V, as has already been described for the purified holoenzyme (Potier *et al.*, 1997).

In the second example, crystals of a small ligand, the 9-*cis* retinoic acid (9-*cis* RA; MW = 300 Da) bound to a nuclear retinoid X receptor (RXR; MW = 27 234 Da) were characterized. Four crystals (100 × 10 × 10 µm) of the RXR–9-*cis* RA complex were washed in 3 *M* NH_4OAc and dissolved in non-denaturing conditions (0.05 *M* NH_4OAc). A mass of 27 537.6 ± 1.8 Da, which corresponds to the mass of the intact non-covalent RXR–9cis RA complex, was measured (Fig. 5*a*). In this case, mass-spectrometry conditions were carefully controlled to prevent complex dissociation. In particular, the cone voltage Vc had to be reduced to 20 V. About 10% of unbound RXR is also detected even at Vc = 10 V, which is the lowest value that can be applied on our instrument, suggesting that the unbound protein did not result from a CID effect but rather was a consequence of some dissociation occurring during the dissolution of the crystals. In contrast, increasing the cone voltage Vc to 60 V led to a loss of 302 Da, reflecting the dissociation of the ligand (MW = 300 Da) in the gas phase (Fig. 5*b*). The cone voltage applied to obtain 50% ligand dissociation could be determined (Vc_{50} = 36 ± 1 V) and was the same as that measured for a RXR–9-*cis* RA complex before crystallization (data not shown).

These examples of ESI–MS analysis under denaturing and non-denaturing conditions provide different ways to detect the presence of a bound ligand in a protein crystal.

7. Discussion

We have established a reliable procedure to characterize the content of macromolecular crystals by mass spectrometry using ESI–MS or MALDI–MS. Only conventional mass-spectrometry equipment is required, permitting the mass measurement of all components of the crystal in less than 60 min. The choice of the analytical technique depends on the number and/or size of the available crystals (more material is required in ESI–MS) and on the problem. When several crystals are available ESI–MS is preferable, because a higher accuracy can be achieved which might be important for controlling the protein integrity. ESI–MS is also the most appropriate technique for detecting the presence of low molecular-weight molecules or analysing non-covalent complexes, even if some reports of successful characterization of protein–protein non-covalent complexes by MALDI are now emerging in the literature (Bordini & Hamdan, 1999; Hillenkamp, 1998; Farmer & Caprioli, 1998). Because MALDI–MS is less sensitive to the presence of salts than ESI–MS, it should be used if the crystals are fragile or cannot be manipulated in NH_4OAc solutions. In this case, it is possible to wash the crystal in the mother liquor, directly deposit the crystal on the MALDI target and perform a successful analysis.

Mass spectrometry was successfully applied to a number of crystals whose content could not be unambiguously analysed with biochemical techniques. Representative applications are summarized in Table 1. The sequence integrities of crystallized arginyl–tRNA synthetase, stromelysin-3 and oestradiol receptor (ER) protein were analysed using ESI–MS. Proteins from the two first crystals were shown to be intact, while in the case of oestradiol receptor protein N- and C-terminal heterogeneity was demonstrated. When crystals diffract poorly or are difficult to reproduce, such analyses are important as the sample homogeneity is known to affect crystal growth and/or quality.

The presence of a specifically bound ligand in a protein crystal was investigated in the case of TBP–DNA, hAR–NADP and RXR–9-*cis* RA crystals. The easiest way to reveal the presence of a specific non-covalently bound ligand by mass spectrometry is to dissociate the complex at extreme pHs and perform an analysis under these denaturing conditions. To detect the ligand, either the positive or the negative ionization mode can be used, depending on its charge. For example, the presence of a negatively charged DNA molecule is classically controlled at high pH (pH = 12), while positively charged molecules are analysed at low pH (pH = 3). Moreover, we have demonstrated for the hAR–NADP and the RXR–9-*cis* RA crystals that the masses of non-covalent complexes could also be obtained when the spectra were recorded at neutral pHs. The possibility of directly characterizing non-covalent complexes might be extremely useful in the case of non-ionic ligands undetectable by mass spectrometry when dissociated from their receptors.

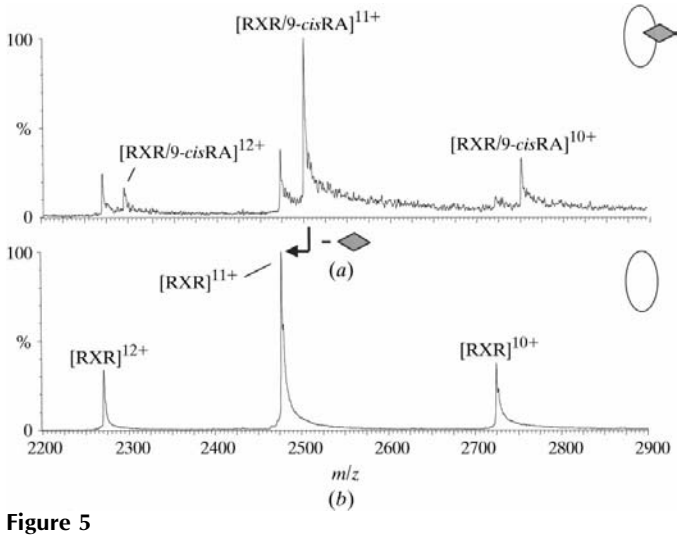

Figure 5
ESI mass spectra obtained from four crystals containing a RXR–9-*cis* RA non-covalent complex after three washings in 3 *M* NH_4OAc and dissolution of the crystal in 0.05 *M* NH_4OAc. When the cone voltage is set to Vc = 50 V (*a*) or lower, the 9-*cis* RA ligand is detected bound to its receptor. An increase of this value to Vc = 60 V (*b*) leads to ligand dissociation (loss of 302 Da).

Mass spectra were obtained from crystals grown in the presence of commonly used crystallization agents. These conditions are considered to be extremely unfavourable for mass spectrometry: the crystals of arginyl-tRNA synthetase were grown in 2.4 *M* $(NH_4)_2SO_4$, the crystals of hAR–NADP in 10% PEG 6000 and the crystals of RXR–9-*cis* RA in 2.5 *M* NaOOCH in the presence of 20% propanediol and 10% glycerol. Determination of adequate washing conditions in the NH_4OAc solution prior to mass measurement was a critical step. Although an exhaustive survey has not been made, it has been observed that many protein crystals can be transferred from their normal mother liquor to NH_4OAc solutions without causing drastic damage. The number of washings as well as the NH_4OAc concentration were adjusted to prevent cracking and ultimate dissolution of the crystal. Although it is not certain that osmotic shock is responsible for the crystal cracking, we observe that the transfer proceeds much more smoothly when the ionic strength of the washing solution is at least as high as the ionic strength of the mother liquor. NH_4OAc being extremely soluble (>10 *M*), it is possible to choose the most appropriate concentration to transfer the crystals without damage. For example, in the case of arginyl-tRNA synthetase crystals, 7.5 *M* NH_4OAc was necessary to wash the crystal grown in 2.4 *M* $(NH_4)_2SO_4$. The use of a high concentration of NH_4OAc for the crystal washing does not spoil the mass-spectrum quality, as shown by the fact that the measured mass of the tRNA synthetase (69 397.7 $\pm$ 7 Da) is in agreement with that expected (69 394 Da).

In summary, we have established a general approach for rapid identification of the crystal content based on mass spectrometry. This procedure was applied to crystals composed of species of a different nature (proteins, nucleic acids, small organic molecules) and grown under different crystallization conditions (high salt concentration, PEG, glycerol *etc.*). This approach is now used routinely in the laboratory and can be carried out successfully even on small crystals (>10 µm), as demonstrated in the case of stromelysin-3. So far, this approach has been applied to determine the integrity and homogeneity of the crystallized protein as well as to verify the presence of the expected bound ligands. In the future, useful applications of this technique should concern the analysis of crystals from large macromolecular assemblies: ribosome, proteasome, RNA polymerase II *etc*. These crystals are grown from endogenous material and are likely to be post-translationally modified. The method can also be extended to the detection and characterization of new unexpected bound ligands present in protein crystals by mass spectrometry.

The Conseil Regional d'Alsace is acknowledged for the acquisition of the MALDI BIFLEX III BRUKER mass spectrometer. We thank Sylvie Duclaud, Anne-Laure Gall, Pascal Egea, Bénédicte Delagoutte and Eduardo Howard for generous gift of crystals of ER, stromelysin-3, RXR, arginyl-tRNA synthetase and hAR, respectively. We thank Julie Thompson for careful reading of the manuscript. This work was supported by grants from the CNRS, the Association pour la recherche sur le Cancer, la Ligue contre le Cancer and la Société de Secours des Amis de la Science.

References

Biemann, K. & Scoble, H. A. (1987). *Science*, **237**, 992–998.
Bordini, E. & Hamdan, M. (1999). *Rapid Commun. Mass Spectrom.* **13**, 1143–1151.
Bruce, J. E., Smith, V. F., Liu, C., Randall, L. L. & Smith, R. D. (1998). *Protein Sci.* **7**, 1180–1185.
Cheng, X. H., Harms, A. C., Goudreau, P. N., Terwilliger, T. C. & Smith, R. D. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 7022–7027.
Ducruix, A. & Giegé, R. (1999). Editors. *Crystallization of Nucleic Acids and Proteins*, pp. 21–36. Oxford University Press.
Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. (1989). *Science*, **246**, 64–71.
Ferré-D'Amaré, A. R. & Burley, S. K. (1997). *Methods Enzymol.* **276**, 157–171.
Green, B. N., Bordoli, R. S., Hanin, L. G., Lallier, F. H., Toulmond, A. & Vinogradov, S. N. (1999). *J. Biol. Chem.* **274**, 28206–28212.
Farmer, T. B. & Caprioli, R. M. (1998). *J. Mass Spectrom.* **33**, 697–704.
Hillenkamp, F. (1998). *New Methods for the Study of Biomolecular Complexes. NATO-ASI Series C*, Vol. 510, edited by W. Ens, K. G. Standing & I. V. Chernushevich, pp. 217–224. Dordrecht: Kluwer.
Hirschler, J., Halgand, F., Forest, E. & Fontecilla-Camps, J. C. (1998). *Protein Sci.* **7**, 185–192.
Karas, M. & Hillenkamp, F. (1988). *Anal. Chem.* **60**, 2299–2301.
Klarskov, K., Roecklin, D., Bouchon, B., Sabatie, J., Van Dorsselaer, A. & Bischoff, R. (1994). *Anal. Biochem.* **216**, 127–134.
Kim, H. Y., Wang, T. C. & Ma, Y. C. (1994). *Anal. Chem.* **66**, 3977–3982.
Lamour, V., Barth, P., Rogniaux, H., Poterszman, A., Howard, E., Mitschler, A., Van Dorsselaer, A., Podjarny, A. & Moras, D. (1999). *Acta Cryst.* D**55**, 721–723.
Light-Wahl, K. J., Schwartz, B. L. & Smith, R. D. (1994). *J. Am. Chem. Soc.* **116**, 5271–5278.
Loo, J. A. (1997). *Mass Spectrom. Rev.* **16**, 1–23.
Mann, M. & Wilm, M. (1994). *Anal. Chem.* **66**, 4390–4399.
Mann, M. & Wilm, M. (1995). *Trends Biochem. Sci.* **20**, 219–224.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Nikolov, D. B., Chen, H., Halay, E. D., Hoffman, A., Roeder, R. G. & Burley, S. K. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 4862–4867.
Petsko, G. A. (1975). *J. Mol. Biol.* **96**, 381–392.
Potier, N., Donald, L. J., Chernushevitch, I., Ayed, A., Werner, E., Arrowsmith, C. H., Standing, K. G. & Duckworth, H. (1998). *Protein Sci.* **7**, 1–8.
Rogniaux, H., Barth, P, Barbanton, J., van Zandt, M., Chevrier, B. Howard. E., Mitschler, A., Potier, N., Urzhumtseva, L., Biellmann, J. F., Van Dorsselaer, A., Moras, D. & Podjarny, A. (1999). *J. Am. Soc. Mass Spectrom.* **10**, 635–647.
Vorm, O., Roepstorff, P. & Mann, M. (1994). *Anal. Chem.* **66**, 3281–3287.
Wilkosz, P. A., Chandrasekhar, K. & Rosenberg, J. M. (1995). *Acta Cryst.* D**51**, 938–945.