ARTICLE

# Mass spectrometry as a novel tool in protein structural biology

**Robert H.H. van den Heuvel* and Albert J.R. Heck**
Department of Biomolecular Mass Spectrometry, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, The Netherlands

## Introduction

In 2002, John Fenn and Koichi Tanaka were awarded the Nobel prize in Chemistry for their contribution to the development of the ionisation techniques electrospray ionisation (ESI)[1] and matrix-assisted laser desorption/ionisation (MALDI),[2] respectively, in the late 1980s. These two novel "soft" ionisation techniques have had a great impact on mass spectrometry and have given, in particular, biomolecular mass spectrometry research an enormous impulse. Working independently, Fenn and Tanaka discovered ways to give huge molecules an electric charge without ripping them apart. Since the invention in 1988, ESI and MALDI have become ubiquitous in both academic and industrial laboratories. Nowadays, mass spectrometry is, in addition to its applications in organic and inorganic chemistry, an essential technology in high-throughput proteomics for the identification of primary sequences of proteins, protein profiling, quantification and post-translational modifications. However, these proteomics studies do not provide us with direct information on higher-order protein structures, protein complexes and interactions between proteins and between proteins and ligands.

Protein X-ray crystallography, electron microscopy, nuclear magnetic resonance and small-angle X-ray scattering are excellent resources in providing protein structural information. In the last decades, these techniques have yielded an enormous amount of information, varying from three-dimensional structures and dynamics of smaller proteins, like the solution structures of the integral membrane protein OmpX by nuclear magnetic resonance[3] up to quaternary structures of viruses, like the 7 Å resolution data for hepatitis B virus particles by electron microscopy.[4] Although these methods provide high-quality data, they also have their own intrinsic limitations, such as high sample consumption and low throughput. In addition, nuclear magnetic resonance and X-ray crystallography are mainly restricted to single proteins or protein domains and require homogeneous samples.

In recent years, the newly-developed soft ionisation techniques together with the possibility to measure masses by high-resolution mass analysers with high transmission and with a broad mass-to-charge range have given mass spectrometry the opportunity to add complementary information to the protein structural biology community. Indeed, soon after the introduction of MALDI and ESI, these ionisation methods were already being used to investigate intact proteins and non-covalent interactions between proteins and ligands, such as nucleic acids, lipids, cofactors, substrates and inhibitors.[5,6] Nowadays, we can explore intact heterogeneous and large assemblies by, especially, ESI mass spectrometry.[7,8] For instance, Robinson and coworkers have demonstrated that it is possible to maintain the non-covalent associations in ribosome complexes, consisting of a large number of proteins and RNA (e.g. *Escherichia coli* 30S ribosome consists out of 21 different proteins and 16S RNA).[9] Also, extremely large intact virus particles are nowadays within the scope of mass spectrometry analysis.[10] In the rest of this communication we will focus further on the emerging role of mass spectrometry in structural biology.

By definition, an ESI mass spectrometer consists of an ESI source, a mass analyser and a detector. The success of ESI started when John Fenn and coworkers showed that multiply-charged ions were obtained from proteins, allowing their molecular mass to be determined.[1] The introduction of nanoflow electrospray has given ESI a new impetus, as such a device makes the analysis of minute amounts of sample possible without compromising signal intensity. Moreover, the increased tolerance to buffer salts and the formation of smaller droplets means that lower desolvation energy is required.[11] An electrospray is produced by applying a strong electric field (1–4 kV), under atmospheric pressure, to a liquid passing through a capillary (Figure 1).[12] Due to this electric field, the liquid in the capillary becomes, in positive-ion mode, enriched for positively-charged ions. Evaporation of solvent from the droplets

*Correspondence to: Robert H. H. van den Heuvel, Department of Biomolecular Mass Spectrometry, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands. E-mail: r.h.h.vandenheuvel@chem.uu.nl


# Powering Discovery

## The New Nicolet™ FT-IR

Thermo Electron, the worldwide leader in FT-IR instrumentation, introduces the new line of Nicolet FT-IR spectrometers.

### Powerful Performance

Explore new horizons in IR analysis with a system that provides the best signal-to-noise, throughput and resolution performance in its class.

### Intuitive Solutions

Drive productivity in your laboratory with a complete set of tools designed to provide integrated solutions for all of your IR problems.

### Future Proof

The proactive, upgradeable system is designed to anticipate new applications in FT-IR. Couple the power of this instrument with the Nicolet Continuμm™ XL– the new standard in infrared imaging microscopes.

See the future of FT-IR today. Contact Thermo Electron at +1-800-532-4752, email analyze.eu@thermo.com or visit www.thermo.com/spectroscopy

FASTLINK / CIRCLE 005 FOR FURTHER INFORMATION

Analyze • Detect • Measure • Control™

Thermo
ELECTRON CORPORATION

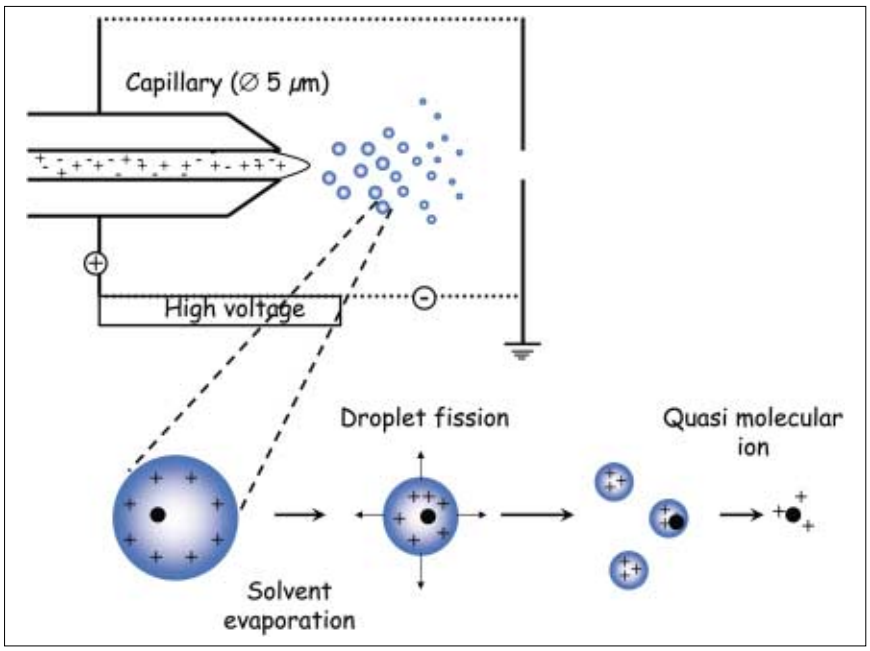



**Figure 1.** Schematic representation of the ESI process.

decreases the radius of the droplets and since the charge is conserved, at a critical radius the repelling Coulombic forces overcome the cohesion forces, thereby causing their divison. Each droplet will then undergo a cascade of ruptures, yielding smaller and smaller droplets. When the electric field on their surface becomes high enough, desorption of ions from the surface occurs. The resulting ions obtained from proteins carry a greater charge when multiple ionisation sites are present. Typically, a protein measured under denaturing conditions will carry one charge per 1000–1500 Da, however, a protein measured under folded, native-like conditions will carry less charges. Therefore, as we will see later, the charge envelope of a protein may give tertiary and quaternary structure information.

For folded proteins it is thought that the charged residue mechanism[13] describes the process in which macro-ions of native proteins are formed. In this mechanism the water molecules in the smallest droplets containing a protein completely evaporate and the protein becomes charged by the residual charges, which have been accumulated during the electrospray process. These excess charges are then transferred to the basic sites at the surface of the protein. Therefore, the number of charges a protein takes up depends on the number of exposed basic sites when the protein is transferred from the solution-phase into the gas-phase. The charged residue mechanism may also predict that the number of charges a protein obtains is dependent on the mass or size of a protein. The calculated charge of a protein can be predicted by the stability limit for Coulombic fission of water droplets[14] with a size equivalent to the protein:

$$Z_R = 8\pi / e \times (\gamma \varepsilon_0 R^3)^{1/2} \quad (1)$$

In which $\gamma$ is the surface tension of the water droplet, $\varepsilon_0$ the electrical permittivity of vacuum, $e$ the elementary charge and $R$ the radius of the droplet. It is reasonable to assume that the radius of a protein is directly correlated to the molecular mass of a protein and that the density of a protein is similar to that of water. From this we can derive a simpler equation in which $M$ is the mass of the protein:

$$Z_R = 0.078 \times M^{1/2} \quad (2)$$

Our group and others have validated this equation by accumulating a large number of experimental data on the ESI of folded globular proteins (Table 1). This table shows, for a number of proteins and protein complexes, the molecular mass and the charge of the most intense ion signal in positive-ion mode. We have also plotted the charges of the base peaks against the molecular masses in Figure 2. It is clear that the predicted number of charges is in close agreement with the experimentally-measured number of charges (≈ 90% of the predicted number). For instance, human tumour suppressor protein Menin type I, having a molecular mass of 68.2 kDa, obtains in positive-ion mode on average 17 charges,

**Table 1.** Number of mean charges a protein obtains in positive ion mode for a number of folded native-like proteins and protein complexes. All mass spectra were obtained in 50 mM ammonium acetate, pH 6.8.

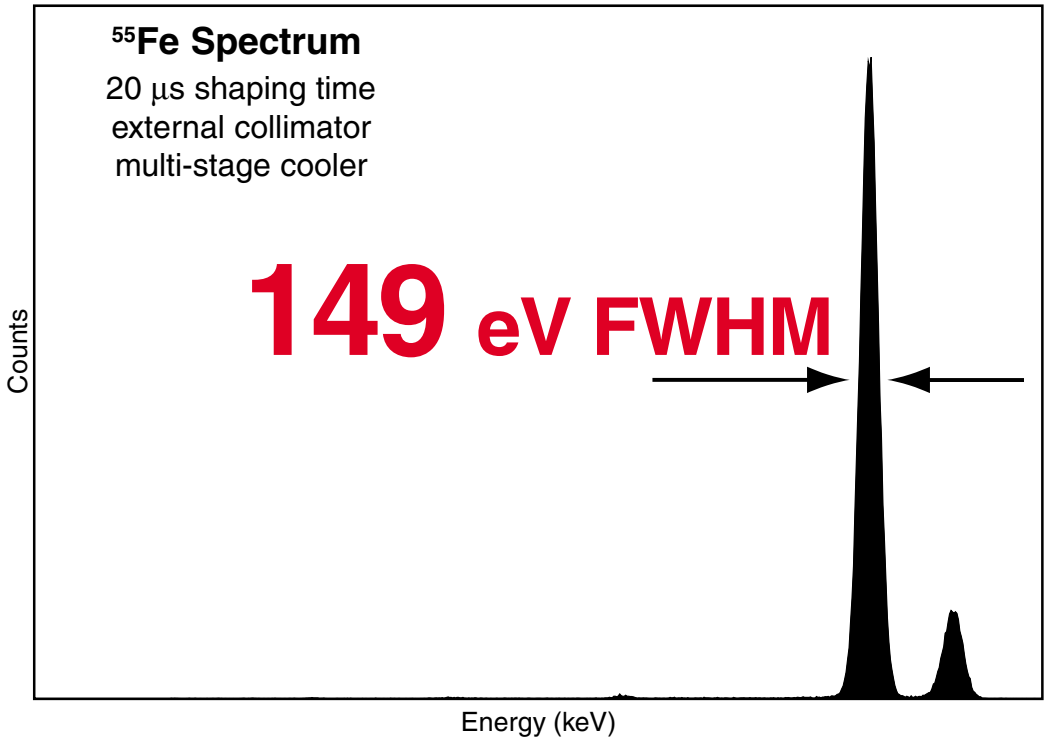
| Protein | Mass (kDa) | Charge (positive ion mode) | Rayleigh[a] |
|---|---|---|---|
| Sumo | 11.0 | 6 | 8.2 |
| PheA2, monomer | 18.4 | 8 | 10.6 |
| PheA2, dimer | 36.9 | 12 | 15.0 |
| EtaA | 61.8 | 16 | 19.3 |
| Menin type I | 68.2 | 17 | 20.3 |
| MutS, monomer | 95.4 | 20 | 24.1 |
| GltS, monomer | 166.2 | 24 | 31.8 |
| MutS, dimer | 190.8 | 28 | 34.0 |
| GltS, dimer | 332.6 | 36 | 45.0 |
| Gp23, hexamer | 335.8 | 41 | 45.2 |
| SR1:GroES | 476.5 | 50 | 53.8 |
| GroEL | 801.3 | 71 | 69.8 |
| Urease | 1064 | 80 | 80.5 |

[a] Charge predicted by the Raleigh model, $Z_R = 0.078 \times M^{1/2}$

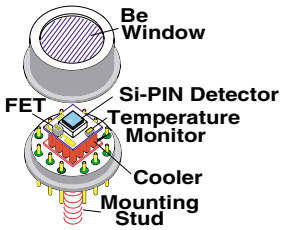
# X-Ray Detector

**No Liquid Nitrogen**
**Solid State Design**
**Easy to Use**

## SUPER XR-100CR




**Area: 6 mm²**
**Thickness: 500 µm**
**P/B Ratio: 3000/1**
**Counting Rates >100000**




**APPLICATIONS**
- OEM
- X-Ray Fluorescence
- Process Control
- Research & Teaching
- Art & Archaeology
- Environmental Monitoring
- Lead Detectors
- Heavy Metals in Plastic
- Semiconductor Processing
- Smoke Stack Analysis
- Plating Thickness
- Coal & Mining Operations
- Sulfur in Oil & Coal Detection
- Jewelry Analysis
- Forensic Investigations

**High Performance at Low Cost**

**OEM's #1 Choice**

**Complete XRF System**

**XR-100CR** X-Ray Detector System
**ECLIPSE-II** X-Ray Tube System
**MCA8000A** Multichannel Analyzer
**XRF-FP** Quantitative Analysis Software

AMPTEK


**AMPTEK Inc.** 6 De Angelo Drive, Bedford, MA 01730-2204 USA
Tel: +1 (781) 275-2242 Fax: +1 (781) 275-3470 E-mail: sales@amptek.com **www.amptek.com**


FASTLINK / CIRCLE 006 FOR FURTHER INFORMATION

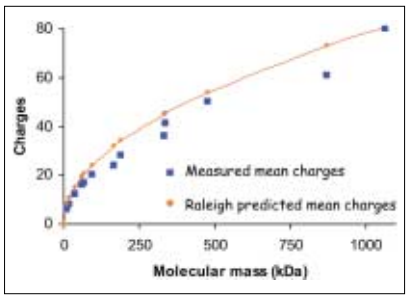



**Figure 2.** Number of mean charges of a number of globular proteins and protein complexes in positive-ion mode compared with the Rayleigh limit model predicted charge. The number of observed charges is close (≈ 90%) to the Raleigh limit model predicted charge. All mass spectra were obtained in 50 mM ammonium acetate, pH 6.8 in positive ion mode.

whereas 20 charges are predicted by Equation 2. Thus, the number of mean charges a protein obtains is a measure of the tertiary and quaternary conformation of a protein and may be used to study protein conformational changes, as reviewed recently.[15] In the second part of this communication, we will focus on the quaternary structure of one protein complex to show the potential of ESI mass spectrometry in structural biology.

## Glutamate synthase in action

The enzyme glutamate synthase is a crucial protein in the early stages of ammonia assimilation in plants in which it is catalysing the conversion of the amino acid L-glutamine into the amino acid L-glutamate.[16,17] Glutamate synthase is a large (165,000 Da per monomer) and complex enzyme, containing two cofactors and two active site cavities that are separated from each other by 30 Å. Furthermore, the protein requires two electrons originating from two iron-sulfur containing ferredoxin molecules to produce the final L-glutamate product. We studied the quaternary structures of glutamate synthase and its interaction with the binding partner ferredoxin. Initially, we measured the exact mass of glutamate synthase under denaturing conditions using a spray solution of 50/50/0.1% acetonitrile/water/formic acid by nanoflow ESI mass spectrometry [Figure 3(a)]. The mass determined from this spectrum is

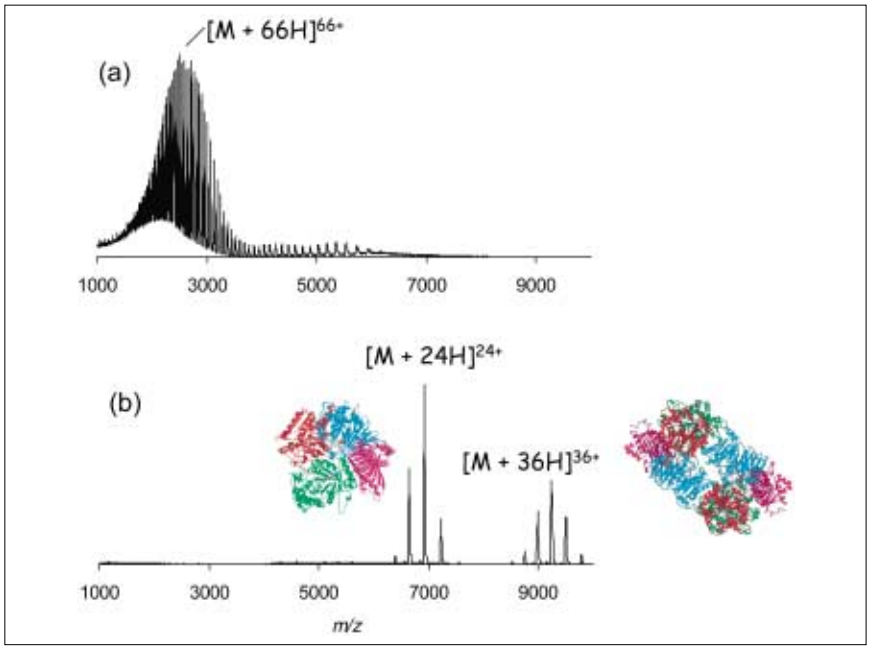



**Figure 3.** Mass measurement in positive-ion mode and quaternary structure of glutamate synthase. (a) Glutamate synthase (2 µM) as measured in the unfolded state in 50/50/0.1% acetonitrile/water/formic acid and (b) glutamate synthase (2 µM) as measured in the folded native-like state in 50 mM ammonium acetate, pH 6.8.

165,526 Da, which is very close to the mass calculated from the primary sequence (165,477 Da). Due to the denaturing conditions, the protein accommodates many positive charges with a distribution centred around $[M + 66H]^{66+}$. This experiment was followed by the measurement of folded glutamate synthase under native-like conditions in aqueous ammonium acetate buffer at pH 6.8. The mass spectrum showed two glutamate synthase species [Figure 3(b)]. The first species has a narrow charge distribution around $[M + 24H]^{24+}$ with a measured mass of 166,222 Da. Indeed this mass corresponds very well with the mass of glutamate synthase in complex with its two cofactors (calculated mass 166,230 Da). The fact that the mean charge state is reduced from 66+ to 24+ reveals that the protein is, under the latter conditions, more folded. The second species with a charge distribution around $[M + 36H]^{36+}$ has a mass very close to the calculated mass of the folded dimeric enzyme.

These experiments thus suggest that glutamate synthase is in equilibrium between the monomeric and dimeric conformation. In order to validate these results we also studied the protein by X-ray crystallography and small-angle X-ray scattering. Crystallography data collected at the European Synchrotron Radiation Facility in Grenoble (France) revealed that glutamate synthase is described in terms of four distinct domains with different function and topology with a total size of about 110 × 100 × 100 $Å^3$ per monomer[18] [Figure 3(a)]. The crystal packing also indicates that the enzyme forms a dimer within the crystalline state, which agrees well with our mass spectrometry data. On the other hand, the small-angle scattering experiments performed in the solution-phase at the Deutsches Elektronen Synchrotron in Hamburg (Germany) have suggested that glutamate synthase is mainly monomeric at high protein concentrations.[18] Taken together, the data strongly indicate that glutamate synthase is a dimeric protein, which is in equilibrium with its monomeric form. It should be noted here that results obtained by using different techniques might be strongly dependent on the experimental conditions such as pH, protein concentration and ionic strength. In the experiments presented here the pH was always neutral, but protein concentrations varied from 2 µM in mass spectrometry up to 90 µM in

# SHIMADZU


ME-DUS.com

# Automatic Accessory Identification.

See. Identify. Respond. Adapt. That's nature's survival strategy. And it's the strategy Shimadzu employs to guarantee the reliability and accuracy of your quality analyses. For example, with our new FTIR system, the IRPrestige-21.

The IRPrestige-21 from Shimadzu lets you switch easily from near to far infrared with outstanding results: Using the ADA system, the IRPrestige-21 can be converted swiftly and simply. This user-friendly system identifies automatically aligns replaced components, allowing you to harness the full range of available accessories with maximum efficiency and reliability.

Designed for a broad range of uses, the IRPrestige-21 is also ideal for highly sensitive applications. Extremely short warm-up and adjustment times make for fast, automatic analyses at low cost.

What's more, the IRPrestige-21 is an efficient and adaptable system that also delivers the goods when things get a little more complex. So the strategy for survival in the animal world can also make your life easier in the world of industry and R&D – and keep you on the cutting edge.

*Precise identification from near to far infrared: The IRPrestige-21 FTIR system.*

**FASTLINK / CIRCLE 007 FOR FURTHER INFORMATION**

SHIMADZU
*Solutions for Science*
since 1875

**Internet:** www.shimadzu.com, **E-mail:** shimadzu@shimadzu.de

Marketing & Customer Support Center, Albert-Hahn-Str. 6-10, D-47269 Duisburg, Germany ☎ +49 - 203 - 7687-0, Fax +49 - 203 - 766625

**Austria · Benelux · Croatia · Czech Republic · Denmark · Finland · France · Germany · Greece · Hungary · Ireland · Israel · Italy**
**Norway · Poland · Portugal · Russia · Slovakia · Spain · Sweden · Switzerland · United Kingdom**

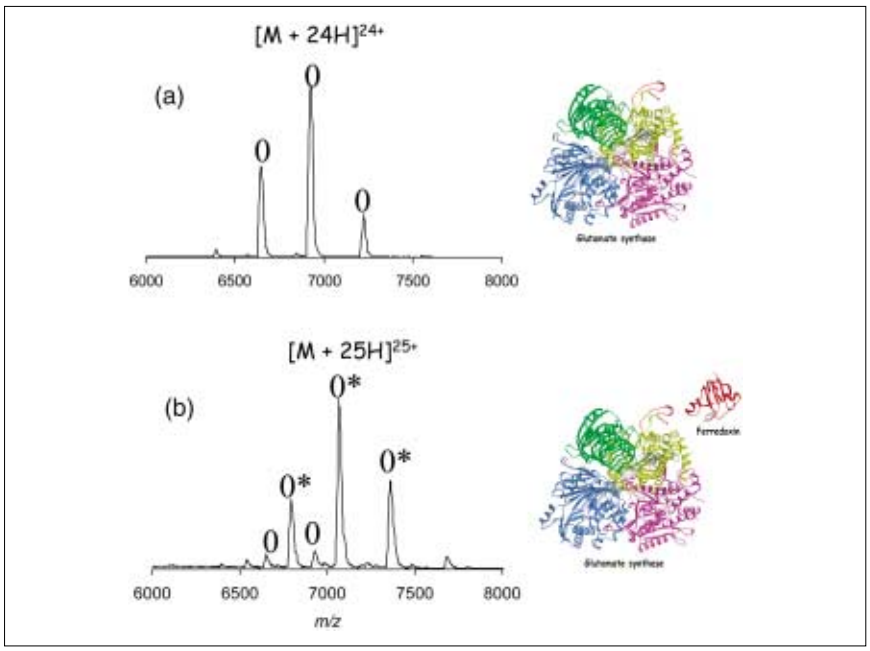



**Figure 4.** Mass measurement and quaternary structure of glutamate synthase in complex with ferredoxin. The mass spectrum was obtained in 50 mM ammonium acetate, pH 6.8 in positive-ion mode. (a) Glutamate synthase (2 µM) and (b) glutamate synthase (2 µM) mixed with ferredoxin (2 µM). (0) Indicates folded glutamate synthase monomer and (0*) indicates the non-covalent complex between glutamate synthase and ferredoxin.

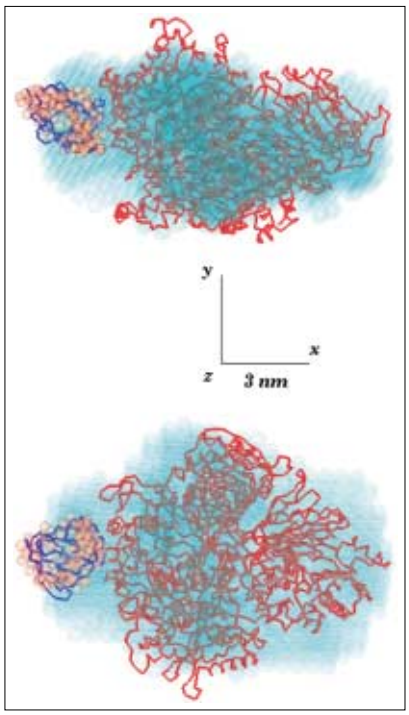



**Figure 5.** Quaternary structure of glutamate synthase in complex with ferredoxin as determined by small-angle X-ray scattering. The low-resolution models were calculated by DAMMIN and GASBOR and superimposed on the high-resolution structures of glutamate synthase and ferredoxin.[19] Bottom view is rotated counter-clockwise by 90° around the X-axis.

small-angle X-ray scattering. Moreover, the ionic strength of the solution used in X-ray crystallography was far higher then in mass spectrometry and small-angle X-ray scattering.

As introduced above, ESI mass spectrometry is a promising technique to study not only oligomerisation behaviour and interactions between proteins and smaller ligands such as cofactors, but also between two or more different proteins. Glutamate synthase is known to receive two electrons, used for the formation of the final product L-glutamate, from its binding partner ferredoxin. Studying the interaction between glutamate synthase and ferredoxin by X-ray crystallography appeared to be inconceivable due to the low stability of the complex at high ionic strength. Therefore, we undertook the assessment to probe the interaction between the two proteins by ESI mass spectrometry. When we mixed glutamate synthase with ferredoxin in a molar ratio of 1 : 1, an equimolar complex of glutamate synthase and ferredoxin was observed (Figure 4). The measured molecular mass of 176,698 Da is in close agreement with the calculated mass of an equimolar complex, including the cofactors (166,230 Da for glutamate synthase + 10,468 Da for ferredoxin). Even upon the addition of a three-fold molar excess of ferredoxin, we only observed an equimolar complex between the two proteins. This is the first direct analysis of a complex between glutamate synthase and ferredoxin, revealing that the two proteins interact very tightly with each other and only form a 1 : 1 complex. These results were recently validated by synchrotron radiation scattering experiments (Figure 5).

## Discussion and future perspectives

The development of ESI in the late 1980s has led to the breakthrough of biomolecular mass spectrometry in the field of structural biology, next to X-ray crystallography, nuclear magnetic resonance, small-angle X-ray scattering and electron microscopy. It is now perfectly possible to transfer very large heterogeneous protein complexes and viruses from the solution-phase into the gas-phase with retention of their higher-order structures and non-covalent interactions. Moreover, we can speculate on the conformational changes of proteins on the basis of the number of charges a protein obtains in the ESI process.

To date, the mass spectrometry community has mainly focused on relatively homogeneous protein samples, such as single proteins, protein–ligand interactions and protein oligomers consisting of one or two different subunits. Recent technological progression has provided a unique niche for mass spectrometry to focus on large and heterogeneous protein complexes and their dynamics. Mass spectrometry cannot give atomic details as protein crystallography and nuclear magnetic resonance can, but provides the researcher with complementary structural information, such as higher-order structures, stoichiometry and dynamics of association. Major advantages of mass spectrometry over all other structural biology techniques is the relatively high throughput, the sensitivity and the possibility to measure each constituent

present in a mixture as long as there is a difference in mass. The major challenge for mass spectrometry is, therefore, to detect, validate and analyse novel proteins and protein complexes and to evaluate dynamics of assembly and disassembly.

## Acknowledgements

We thank Dmitri I. Svergun and Maxim V. Petoukhov (European Molecular Biology Laboratory, Germany) and Andrea Mattevi (University of Pavia, Italy) for their help with the experiments and in preparing some of the figures.

## References

1. J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong and C.M. Whitehouse, *Science* **246,** 64–71 (1989).
2. K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida and T. Yoshida, *Rapid Commun. Mass Spectrom.* **2,** 151–153 (1988).
3. C. Fernandez, C. Hilty, S. Bonjour, K. Adeishvili, K. Pervushin and K. Wuthrich, *FEBS Lett.* **504,** 173–178 (2001).
4. B. Bottcher, S.A. Wynne and R.A. Crowther. *Nature* **386,** 88–91 (1997).
5. T.D. Veenstra, *Biophys. Chem.* **79,** 63–79 (1999).
6. J.A. Loo, *Mass Spectrom. Rev.* **16,** 1–23 (1997).
7. A.J.R. Heck and R.H.H. van den Heuvel, *Mass Spectrom. Rev.*, in press (2004).
8. M.G. McCammon and C.V. Robinson, *Curr. Opin. Chem. Biol.* **8,** 60–65 (2004).
9. A.A. Rostom, P. Fucini, D.R. Benjamin, R. Juenemann, K.H. Nierhaus, F.U. Hartl, C.M. Dobson and C.V. Robinson, *Proc. Natl. Acad. Sci. USA* **97,** 5185–5190 (2000).
10. S.D. Fuerstenau, W.H. Benner, J.J. Thomas, C. Brugidou, B. Bothner and G. Siuzdak, *Angew. Chem. Int. Ed. Eng.l* **40,** 542–544 (2001).
11. M. Wilm and M. Mann, *Anal. Chem.* **68,** 1–8 (1996).
12. P. Kebarle, *J. Mass Spectrom.* **35,** 804–817 (2000).
13. J. de la Mora, *Anal. Chim. Acta* **406,** 93–104 (2000).
14. L. Rayleigh, *Philos. Mag.* **14,** 184 (1882).
15. X. Yan, J. Watson, P.S. Ho and M.L. Deinzer, *Mol. Cell Proteomics* **3,** 10–23 (2004).
16. R.H.H. van den Heuvel, B. Curti, M.A. Vanoni and A. Mattevi, *Cell. Mol. Life Sci.* **61,** 669–681 (2004).
17. M.A. Vanoni and B. Curti. *Cell. Mol. Life Sci.* **55,** 617–638 (1999).
18. R.H.H. van den Heuvel, D.I. Svergun, M.V. Petoukhov, A. Coda, B. Curti, S. Ravasio, M.A. Vanoni and A. Mattevi, *J. Mol. Biol.* **330,** 113–128 (2003).
19. M.H. Koch, P. Vachette and D.I. Svergun, *Q. Rev. Biophys.* **36,** 147–227 (2003).