Protein-Structure Determination

DOI: 10.1002/anie.201007231

(Re)Solution of a Protein Fold Without Solution

*Perdita E. Barran**

gas-phase conformations · mass spectrometry · protein structures · structure elucidation

The technique of mass spectrometry, now nearly 100 years old, was developed by Thomson and Aston[1] at the Cavendish Laboratories, Cambridge University. These early pioneers used the gas-phase environment of their so called "mass spectrograph" as a laboratory for the study of atomic structure, unhindered by bulk interactions. Roll forward to 2010, and the gas phase remains a favorite location for detailed structure analysis, and particularly since the advent of electrospray ionization (ESI)[2] there has been increasing applicability to biological macromolecules.

ESI is uniquely suited to the analysis of high-molecular weight involatile systems, as long as they are soluble in a solvent. As a so-called "soft" method of ionization, it preserves complex unsolvated biological molecules without causing excessive fragmentation. ESI-MS is used extensively in this postgenomic era to determine the primary structure of proteins, but it can also provide information about intramolecular interactions, and hence protein structure. Mass spectrometry has now established itself as a powerful method with which to determine the stoichiometry of very complicated protein complexes, and even to solve the architecture of the subunits[3] or to examine their assembly.[4] There is naturally significant controversy as to whether a solution-phase structure can be retained in its entirety in the solvent-free environment of a mass spectrometer, but for large macromolecular systems bound by many noncovalent interactions, many macroscopic features of the solution structures and even in vivo active structures can be retained. What now emerges as a greater challenge is to see precisely how structure is retained at the atomistic level, and for how long. There is growing evidence to suggest that the gas phase can provide such detail[5] and moreover that dynamic information regarding the intrinsic stability of higher order structure is perhaps best obtained with mass spectrometry.

Breuker et al. have used Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR MS) to examine the three-helix bundle protein KIX.[6] After ESI, proteins are produced in a range of charge states; in the positive ionization mode, the charging species is usually protons. The KIX protein is observed as ions of the general form $[M+n\,\mathrm{H}]^{n+}$, where n lies between 7 and 16. Once trapped in the low vacuum environment of the mass spectrometer, the protein is then subjected to dissociation by using a beam of electrons—so-called electron-capture dissociation (ECD). This technique is non-ergodic, which means that the energy involved in electron capture is consumed locally to cause cleavage more rapidly than the time it takes the bonds to vibrate. As a consequence of this, ECD does not perturb the higher order structure of peptides and proteins. It also tends to cleave selectively the backbone of the protein under analysis, with the bond between the amide nitrogen atom and the α-carbon atom broken. This fast and directed cleavage process can at times give rise to widespread fragmentation of the protein, which can be used to assist in the identification and location of post-translational modifications.

As a result of being exposed to the beam of electrons, each charge state of the protein KIX produces a different signature of fragment ions. In their experiments, Breuker et al. isolate a given charge state of the protein, so that it is trapped in the ion cell of the FT-ICR mass spectrometer. They then expose it to the beam of electrons for a defined period of time, after which, fragments arising from the selected precursor ion are detected. Very little fragmentation is observed for the lowest charge states $n = 7$–9, whereas the amount of fragmentation increases significantly for ions with 10 or more excess protons. Each charge state gives a slightly different pattern of fragment ions, with signatures grouping the charge-state distribution into three distinct regions: $n = 7$–9, $n = 10$–11, and $n =$ 12–16. It is likely that these three regions have similar conformations. As the number of protons located on any protein increases after ESI, so does the likelihood that these protons are close enough to repel one another, thus causing a coulombically driven denaturing of the protein—an effect that is well documented through ion mobility measurements.[7] In light of this, we can appraise the findings of Breuker et al. Low charge states will represent conformation(s) of the protein that are most compact, and most likely to still retain a solution-like fold, whereas those for higher numbers of protons are more likely to be disturbed by electrostatic repulsion, and be more elongated. Since the higher charge-state ions observed here are sprayed from denaturing solutions, it is possible that their gas-phase structures will also resemble those in solution. Therefore, mapping the location of an observed fragment ion in the solution fold is extremely insightful. Detected fragments will come from areas of the protein that are not held by noncovalent bonds, in other words from the parts of the protein that can fray or denature away from any stable fold.

[*] Dr. P. E. Barran
The School of Chemistry, The University of Edinburgh
Kings Buildings, Edinburgh (UK)
Fax: (+44) 131-650-6453
E-mail: perdita.barran@ed.ac.uk

From the solution-state structure of this protein, as determined by NMR spectroscopy, Brueker and co-workers have postulated that up to 10 salt bridges could form to stabilize the structure.[6] As a consequence of the dominance of electrostatic interactions in the absence of solvent, such salt bridges could likely maintain the solution fold. The structure is a tight bundle of α helices, where the N terminus is held in noncovalent contact with the α3 helix. For the $n=7$ ion, very few ECD fragment ions are seen, which suggests that ions of this charge state are still present in the same fold as in the solution structure. Breuker et al. report that the small number of fragment ions observed all arise from the N terminus and not from the bundle region, even when the ECD is preceded by collisional activation. For the $n=8$ ion, ECD produces cleavage products from the N-terminal ends of helices α1 and α2. This is attributed to the start of partial unfolding (in contrast to the $n=7$ ion), with the helix α1 having lost its contact with the bundle and started to unfold at its already less-structured N-terminal end (Figure 1). As one proton is

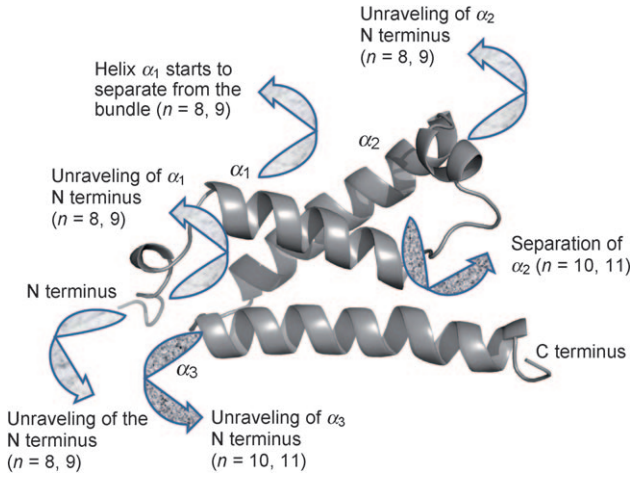

Figure 1. The solution structure of the KIX protein[8] in aqueous solution at pH 5.5. and 27 °C (pdb entry 2AGH, model 1). Figure adapted from Breuker et al.[6] to summarize the loss of the solution fold as revealed by ECD fragmentation. The fragmentation patterns of the $n=8$–11 ions indicate a loss of the tertiary structure, in contrast to the pattern seen for the $n=7$ ion.

added to give the $n=9$ ion, more unraveling of the α1 and α2 helices is observed. A more significant change in the detected fragments occurs with the $n=10$ and 11 ions observed; fragments now arise from regions containing the α2 and α3 helices, which suggests that these were retained intact when $n=7$–9, but are now no longer defined helices in the ions from a more denaturing solution. The fragment yield from the α3 helix is still low, which indicates that this helix is still partially retained even in these high charge states. It is only for ions where $n=12$–16 that significant cleavage products come from the α3 helix. Assessment of the helix stability was made by comparing the fragments from each of the three helical regions, with the order of helix stability found to be α3 > α2 > α1. This is identical to that found in solution.[9]

Breuker et al.[6] go on to show that the low level of fragmentation observed for the $n=7$ ion remains low even after four seconds of trapping prior to ECD, thus showing that the compact solution fold can be retained for significant time periods, even after being subject to collisional activation. So, despite the loss of solvent and the corresponding loss of hydrophobic bonding, the intrinsic noncovalent interactions in this three-helix bundle protein are strong enough to be retained for long periods of time in a solvent-free environment. The postulated salt bridges must be playing a dominant role to stabilize the global fold in the absence of solvent, and the regions where they would be located correspond to the most stable region of the helices where least fragmentation occurs. Additionally, there is likely to be a stabilizing effect arising from the electric dipole moments of each of the helices as a result of charge–helix dipole interactions as well as dipole–dipole interactions between the helices. On the basis of an analysis of the charge density in the solution structure, the order of stability of the helices has been found to be largely dictated by the density of salt bridges, with other electrostatic interactions arising from neutral and ionic hydrogen bonds as well as dipole interactions providing additional stabilization.

This study shows that mass spectrometry can be used to provide information about the macroscopic features that arise from a tertiary fold, as well as exquisite detail on the intrinsic interactions that dictate it. Studies of this sort suggest a change in the way we may elucidate protein structure. Accurate (atomic resolution) determination of protein structure has long been monopolized by X-ray crystallography, but this technique suffers from an inability to assess the many changes in conformation which occur as proteins fold; it is practically impossible to crystallize folding intermediates. Other transient structural changes arising from the activity of a protein are also unobservable by crystallographic methods. NMR spectroscopy has gone some way to address this, but there are limitations to its ability to characterize folding events that occur on very fast time scales; NMR spectroscopy is often carried out synergistically with molecular dynamics calculations. In this approach experimental data is used to provide starting (and finishing) coordinates, and MD methods supply the structural intermediates that occur over short timescales. In silico methods are gaining momentum as viable means to determine tertiary structures; however, effort has been (rightly) directed to model systems, and only a handful of "real" systems have been solved solely by in silico techniques. Breuker et al. have shown that mass spectrometry based methods have a real role to play, and that it is possible to determine a solution-relevant structure, as well as the dynamics between secondary structural elements, in the absence of any solvent at all!

Received: November 17, 2010
Published online: March 2, 2011

[1] J. J. Thomson, *Proc. R. Soc. London Ser. A* **1913**, *89*, 1.
[2] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science* **1989**, *246*, 64.

[3] M. Zhou, A. M. Sandercock, C. S. Fraser, G. Ridlova, E. Stephens, M. R. Schenauer, T. Yokoi-Fong, D. Barsky, J. A. Leary, J. W. Hershey, J. A. Doudna, C. V. Robinson, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 18139.
[4] R. J. Rose, D. Verger, T. Daviter, H. Remaut, E. Paci, G. Waksman, A. E. Ashcroft, S. E. Radford, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12873.
[5] K. Breuker, F. W. McLafferty, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 18145.
[6] K. Breuker, S. Brüschweiler, M. Tollinger, *Angew. Chem.* **2011**, *123*, 903 – 907; *Angew. Chem. Int. Ed.* **2110**, *50*, 873 – 877.
[7] B. J. McCullough, J. Kalapothakis, H. Eastwood, P. Kemper, D. MacMillan, K. Taylor, J. Dorin, P. E. Barran, *Anal. Chem.* **2008**, *80*, 6336.
[8] R. N. De Guzman, N. K. Goto, H. J. Dyson, P. E. Wright, *J. Mol. Biol.* **2006**, *355*, 1005.
[9] P. Schanda, B. Brutscher, R. Konrat, M. Tollinger, *J. Mol. Biol.* **2008**, *380*, 726.

