PROTEIN SEQUENCING BY MASS SPECTROMETRY


Department of Biochemistry, Imperial College of Science and Technology, London SW7. U.K.

INTRODUCTION
Mass spectrometry is playing an increasingly important part in the expanding field of biomolecular structure determination. Indeed, over the past ten years the technique has graduated from the organic chemistry service laboratories to reach a position where today total protein sequence information may be derived by MS techniques alone (1). Mass spectrometry is probably best known for its contribution to the structure elucidation of biologically active substances where some unusual structural feature is either known or suspected to be present. Some examples of this type are shown in Fig. 1, which shows the structures, determined in this laboratory, of the enkephalins, Leukotriene D (Slow-Reacting Substance of Anaphylaxis), γ-carboxy glutamic acid, antifreeze glycopeptides 'AF8' and locust Adipokinetic Hormone, (1,2). This work was carried out on materials of unknown structure usually at the low nanomole level, and clearly shows the power of a technique which, unlike "classical" methodologies, is completely independent of compound class.

Protein/Peptide Analysis
Proteins and peptides probably represent the most difficult examples of structural analysis for mass spectrometry. They are in general polar, highly complex molecules, (often multiply charged at pH7) whose complexity may be further compounded by post-translational modification (e.g. glycosylation, phosphorylation etc.). Further their molecular
Fig. 1. Important biological substances whose structures were determined by MS.

weights are well in excess of the mass range of conventional mass spectrometers (1000 mu). The reasons for studying protein structures in general are many and varied - proteins playing both structural and control (enzymatic/hormonal) roles throughout living systems - but why use mass spectrometry when both classical and recombinant DNA sequencing methods are available? The answer lies in the unique characteristics of mass spectrometry in terms of 1) sensitivity, 2) specificity and 3) mixture handling capability. Because MS is independent of sample class, it is particularly well suited to the analysis of peptides or proteins with unusual structural features.

The method most widely used to generate amino acid sequence is electron impact (EI) MS analysis of volatile N-acetyl N,O permethyl derivatives (1). These derivatives are volatile in the ion source - an essential prerequisite for EI analysis - and most importantly, fragment in a defined and specific way (Fig. 2). Fragmentation occurs at each amide bond resulting in N-terminal sequence ions; the mass difference between the sequence ions corresponds to the