CHIMIA

Volumen 14 Fasc. 12 1960 Pag. 393–434

3. Europäisches Peptid-Symposium

Basel, 5. bis 8. September 1960

Die drei nachfolgend abgedruckten Vorträge wurden am 3. Europäischen Peptid-Symposium in Basel gehalten. Von den übrigen Vorträgen sind Autoreferate im Novemberheft der *Chimia* zu finden*.

The Application of Mass Spectrometry in Amino Acid and Peptide Chemistry

By K. BIEMANN

Department of Chemistry, Massachusetts Institute of Technology, Cambridge (Mass. U.S.A.)

Zusammenfassung

Im Anschluß an eine kurze Einführung über das Wesen der Massenspektrometrie wird deren Anwendung auf folgende Probleme beschrieben: Identifizierung von Aminosäuren, Strukturbestimmung unbekannter Aminosäuren, quantitative Aminosäure-Bestimmung, Sequenzanalyse von Peptiden, Bestimmung N-terminaler Aminosäuren und freier Aminogruppen in Seitenketten.

Am wichtigsten ist die Möglichkeit einer direkten Sequenzbestimmung in Oligopeptiden bis mindestens zum Pentapeptid. In Kombination mit einer gas-chromatographischen Vortrennung lassen sich auch komplexe Peptidgemische bearbeiten.

Massenspektrometrische Analysen erfordern sehr wenig Material. Bei Serienversuchen ist auch der allgemeine Aufwand und der Zeitbedarf gering. Die quantitative Aminosäure-Bestimmung empfiehlt sich, sobald eine Untersuchung, wie etwa eine Stoffwechselstudie oder die Konstitutionsaufklärung eines Proteins, eine große Anzahl von ähnlichen Analysen erfordert. Für gelegentliche Einzelbestimmungen dagegen verwendet man besser andere Analysenverfahren.

At present, the techniques used for the determination of the primary structure of peptides and proteins rely mainly on chromatographic and electrophoretic separations and the identification of the ultimate degradation products, in general amino acids and derivatives thereof, by nonspecific color reactions. Over the last decade these techniques have led to a dramatic advance of the knowledge of protein structure. For the solution of the many more structural problems in the peptide field there seems to be only one obstacle: the amount of time and manpower required due to the slowness of the methods of separation employed, the tedious procedures used for the determination of the amino acid sequence in the degradation-peptides and the sometimes not very reliable identification of the products by their mobility in chromatographic and electrophoretic systems.

It was primarily in consideration of these factors that we began to investigate the possible use of mass spectrometric methods in this field, because it is a technique which combines speed with a high degree of specificity of information, the latter being also strictly related to the structure of the molecule. Furthermore, the minute sample required meets another prerequisite of any method to be of pratical applicability in this field and it was also felt that the introduction of an entirely new approach might be of a general value. In addition, the isolation of peptides from microorganisms and plants, which frequently contain new amino acids, makes it necessary to determine their structure before elucidation of the sequence in the peptide, and also for this purpose mass spectrometry could be used with success.

The elaboration of mass spectrometric methods was, therefore, aimed at the following problems:

qualitative identification of amino aids,
determination of the structure of new amino acids,
quantitative determination of amino acids,
sequence of amino acids in peptides,
determination of N-terminal amino acids and free $-NH_2$ groups,
determination of C-terminal amino acids and free –COOH groups.

The first five of these have thus far been solved in principle[1-6] and it is the purpose of this review to discuss the methods involved, their potentialities and limitations.

* Sonderdrucke, in denen alle Autoreferate und die drei nachstehend abgedruckten Vorträge von K. BIEMANN, H. ZAHN und H. ZUBER zusammengefaßt sind, können bei genügendem Interesse beim Verlag H. R. Sauerländer & Co., Aarau, bezogen werden.

[1] K. BIEMANN, J. SEIBL and F. GAPP, *Biochem. Biophys. Res. Comm. 1* (1959) 307.

[2] K. BIEMANN, C. LIORET, J. ASSELINEAU, E. LEDERER and J. POLONSKY, *Biochim. Biophysica Acta 40* (1960) 369.

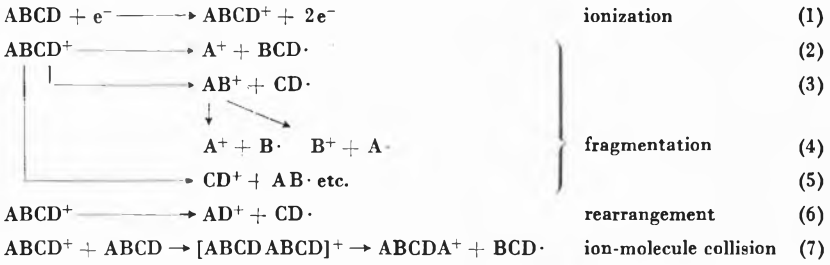
[3] K. BIEMANN and W. VETTER, *Biochem. Biophys. Res. Comm. 2* (1960) 93.

[4] K. BIEMANN, F. GAPP and J. SEIBL, *J. Amer. Chem. Soc. 81* (1959) 2274.

[5] K. BIEMANN and W. VETTER, *Biochem. Biophys. Res. Comm.*, in press.

[6] K. BIEMANN and W. VETTER, unpublished.

It might be advantageous to first discuss the principle of the mass spectra of organic compounds, without, however, going into the instrumental aspects, reviews of which can be found elsewhere.[7] A mass spectrum is a recording of mass vs. intensity of the positively charged particles formed on electron bombardment of the compound, present in the gaseous state in the ion source of the spectrometer. They are charged molecules and fragments thereof produced on electron impact through the following simplified processes:



The first step leads to the molecular ion, the mass of which discloses the accurate molecular weight. Steps 2 to 5 give rise to fragments representing atoms or groups present as such in the molecule, in contrast to the ion formed in step 6. Reaction 7 produces a particle heavier than the original molecule. The last two processes would seem to limit the usefulness of mass spectrometry for the determination of structure; detailed knowledge of the mechanism of these reactions and the circumstances under which they occur, however, makes it possible to deduce specific structural information also from these ions.

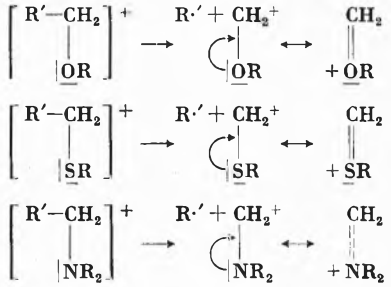
The mass spectrum of even a relatively small organic molecule thus is very complex and contains a large number of peaks, practically one for every possible combination of the atoms present in the molecule. This combined with the high degree of reproducibility of the spectra makes for their usefulness for the proof of identity of two compounds. A consideration of the intensity of the peaks, i. e. the probability of formation of a fragment of a certain mass, is the basis of the interpretation of such a spectrum in terms of the structure of the compound. Since each fragmentation step involves the scission of one or more bonds and the production of two fragments (a positive and a neutral one) the energy required and thus the frequency of the event depends mainly on the strength of the bond to be broken and the stability of the fragments to be formed. The latter factor is the more influential one, and the most intense peaks in the spectrum are due to those fragments in which the positive charge is best stabilized:

[7] L. Jenckel and E. Dörnenburg, in E. Müller, ed., *Methoden der Organischen Chemie* (Houben-Weyl), Vol. III, Part I, pp. 693 to 752, Verlag Georg Thieme, Stuttgart 1955. D. W. Stewart, Mass Spectrometry, in A. Weissberger, ed., *Physical Methods of Organic Chemistry* (Technique of Organic Chemistry, Vol. I) 2nd ed., Part II, pp. 1990–2058, Interscience, New York.

Hydrocarbons or alkyl groups are cleaved at the points of branching due to the increase in carbonium ion stabilization in the order

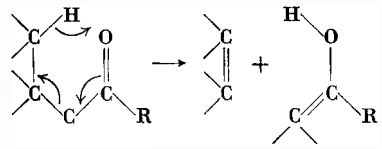
$$CH_3^+,\ RCH_2^+,\ R_2CH^+,\ R_3C^+.$$

Molecules containing atoms with unshared electron pairs are cleaved at the carbon-carbon bond next to this atom due to the resonance stabilization of the positive charge by the available electron pair



These effects are in some respect additive, i. e. the fragmentation of the carbon-carbon bond next to the hydroxyl group is much more pronounced in tertiary alcohols than in primary ones. In other respects they are competitive: in an α-amino alcohol the positive charge is stabilized better on the nitrogen containing fragment which, therefore, appears in the spectrum as a stronger peak than the hydroxylic fragment.

The most common rearrangement reactions of the ions formed in the mass spectrometer are the elimination of H_2O and NH_3 from alcohols and primary amines (NH_2R and NHRR′ from secondary and tertiary amines respectively) and of olefines or similar fragments from carbonyl compounds, esters etc. through



The ion-molecule collisions (step 7 in the above scheme) do not occur in mass spectra obtained under normal conditions with the exception of a peak at one mass unit above the molecular weight which corresponds to a molecular ion which has abstracted a hydrogen atom from another molecule and is thus a protonated molecule which is often more stable than the molecular ion itself

and gives rise to an appreciable "M + 1 peak". Since it arises from a collision of two particles its intensity is not directly proportional to the sample pressure, like all other peaks, but to the square of the pressure. It is also dependent on the conditions in the ion source which in effect alter the relative concentration of the ions. These changes in relative intensity clearly identify the M + 1 peak which thus may be used to establish the molecular weight of a compound which does not exhibit a peak at its molecular weight because of the instability of the molecular ion. A very typical class of such compounds are the polyamino alcohols and polyamines discussed later on.

This rather brief introduction to the fragmentation processes on electron impact will suffice for the following discussion of the mass spectra of derivatives of amino acids and peptides.

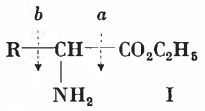
One more point has to be emphasized: In a conventional mass spectrometer for organic materials the sample has to be present in the ion source in the gas phase. For practical purposes, the compound has to be vaporized to produce at least about 10^{-2} mm pressure at the temperature of the inlet system of the instrument which is kept at 140 °C in the author's instrument,* but might be as high as the thermal stability of the sample permits. This necessity of vaporizing the sample without decomposition makes it practically impossible to obtain from free amino acids or peptides mass spectra of the quality required for an unambiguous interpretation. It was, therefore, necessary to convert these materials by specific chemical reactions into more volatile derivatives which, however, still contain all the structural features of the original compounds.

Identification of amino acids

As volatile derivatives of amino acids we have chosen their ethyl esters, because they can be easily prepared, even on a very small scale, by esterification with ethanolic hydrochloric acid and have the advantage over methyl esters of a slower rate of dimerization to diketopiperazines.

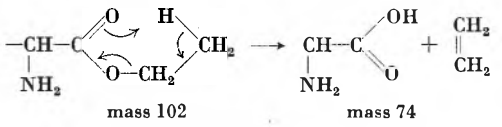
The mass spectra of over thirty of such esters have been investigated[1] and they revealed a number of characteristics which make them suitable for the identification or determination of structure:

If we consider an α-amino ester of the general structure



I

we find, in agreement with the generalizations made above, that either one of the carbon-carbon bonds next to the amino group are easily broken and cleavage at *a* gives rise to the more intense peak (this has also been observed with methyl esters).[8] The latter indicates the size of R and the fragment due to cleavage at *b* will always occur at mass 102, unless it would bear a substituent on the α-carbon atom or the amino group, the size of which could be deduced from the mass number of the corresponding peak after subtraction of 101 mass units. Proline ethyl ester, lacking this group, does not exhibit a peak at m/e 102. Furthermore, this ion of mass 102 always gives rise to a peak at m/e 74 by elimination of ethylene from the alcohol moiety of the ester group



The structure of the group R in I can be deduced from peaks due to preferred fragmentation of bonds in this part of the molecule, e. g. at the points of branching. The esters of leucine (Fig. 1) and isoleucine (Fig. 2) may be distinguished by their peaks at m/e 144, 116, 43 and

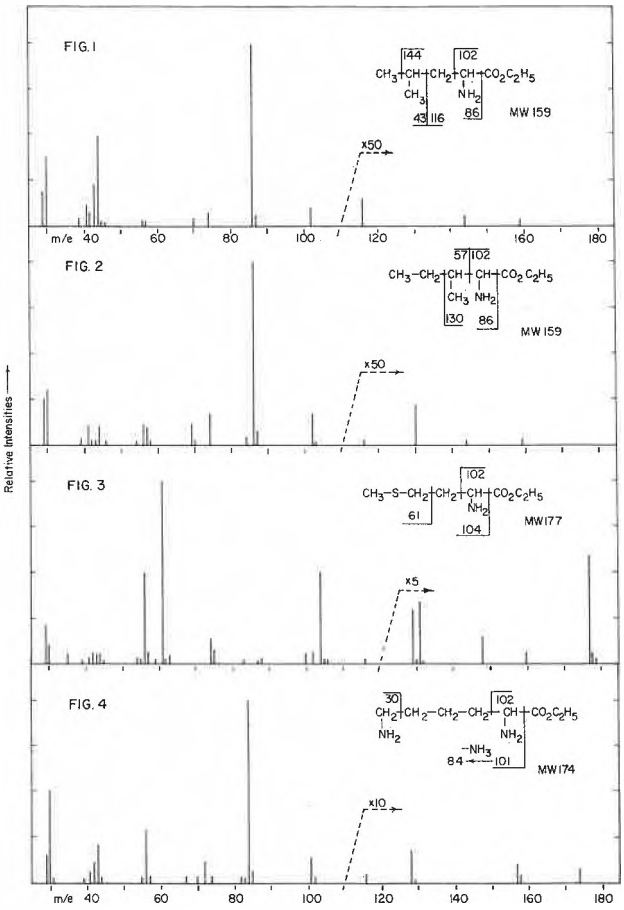


Fig. 1–4. Mass spectra of the ethyl esters of leucine (1), isoleucine (2), methionine (3) and lysine (4)

* A CEC21–103C mass spectrometer has been used for all experiments discussed herein. It has a mass range and useable resolution extending to mass 700. Scanning time m/e 12–700 is 11 to 20 min. A fraction of a milligram of sample is required for a spectrum.

[8] C. O. Andersson, *Acta Chem. Scand. 12* (1958) 1353.

130, 57, respectively.* Furthermore, characteristic differences in these two spectra are due to further decomposition of the fragment (mass 86 in both cases) arising through cleavage at *a* in I. Peaks at mass 44 (C_2H_6N) in leucine and 69 (C_5H_9) in isoleucine ester are due to such processes.

Although in many simple amino acid esters the $R-\underset{\displaystyle NH_2}{\overset{\displaystyle |}{CH^+}}$ peak is the most intense one in the spectrum, the introduction of a functional group into the side chain frequently gives rise to other very strong peaks, e. g. m/e 84 in lysine ester (Fig. 4)

$$\underset{\displaystyle NH_2}{\overset{\displaystyle |}{CH_2}}-CH_2-CH_2-\underset{\displaystyle NH_2}{\overset{\displaystyle |}{CH}}-CH_2-CH^+ \xrightarrow{-NH_3} \underset{\displaystyle NH_2}{\overset{\displaystyle |}{CH_2}}-CH_2-CH_2-CH=CH^+ \quad \text{(or cyclic structure)}$$

or the peak at m/e 61 in methionine ester (Fig. 3), which represents the fragment $CH_3-S-CH_2^+$. Some of the fragments (m/e 56 and 129) in this spectrum are due to elimination of CH_3SH (mass 48) from either the molecular ion or the fragment of mass 104.

The mere identification of an amino acid by a brief interpretation of the mass spectrum of its ester and comparison with the spectrum of an authentic sample is therefore very simple and also highly specific. It is even possible to recognize the individual components in a mixture of amino esters without prior separation by consideration of the intense peaks present in the mass spectrum of the mixture. In this case the use of ionizing electrons of low energy (10–12 eV instead of the otherwise employed 70 eV) is of advantage since the spectrum obtained is much simpler and consists mainly of the peaks due to cleavage at *a* in structure I.

It might be pointed out at this time, that the mass spectra of D and L and racemic forms of optically active molecules are, of course, identical. There are only slight differences in the spectra of aliphatic diastereomeres, in contrast to certain cyclic ones.

Determination of the structure of new amino acids

From the foregoing discussion it is clear that the structure of an amino acid is reflected in detail in the mass spectrum of its ethyl ester, and it should therefore also be possible to deduce the structure of a new amino acid from the corresponding spectrum.

This has been accomplished[2] in the case of an amino acid recently isolated from the crown gall tissue of a number of plants. By the usual analytical methods it had been shown not to be a primary *a*-amino acid, to contain half its nitrogen in an NH_2-group, one $C-CH_3$, no $N-CH_3$ nor $O-CH_3$ and to be probably a diacid of the composition $C_9H_{18}N_2O_4$, although the elemental analysis was made difficult by the retention of crystal solvent. The small amount of material available made it impossible, however, to determine the structure by chemical degradation.

The mass spectrum of the ethyl ester of this compound (Fig. 5) obtained on esterification of 0.4 mg of the acid and regeneration of the free ester with ammonia in dichloromethane (see below) exhibited intense peaks at m/e 84, 156, 184 and 201.

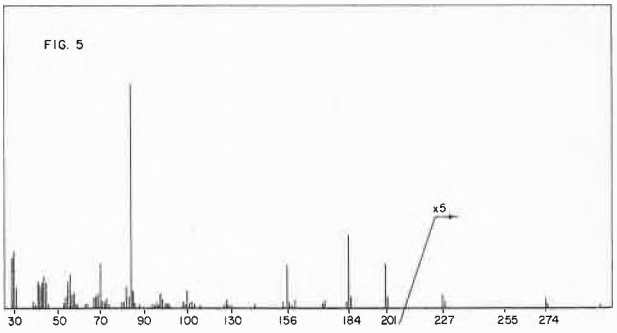


Fig. 5. Mass spectrum of the ethyl ester of lysopine

A small peak at mass 274 was identified as the molecular weight on the basis of the even smaller one at m/e 275 which exhibited the characteristic behaviour of an M+1 peak, mentioned previously. This was corroborated also by the peak at m/e 201, corresponding to a loss of 73 mass units ($-CO_2C_2H_5$). The even molecular weight indicates an even number of nitrogens, most probably two. Mass 201 must still contain a primary amino group because it looses 17 mass units (NH_3) forming the ion of mass 184. The latter eliminates 28 mass units—most probably C_2H_4—to form the particle of mass 156, and it was already pointed out that it is characteristic for fragments containing a carbethoxy group. The very strong peak at m/e 84 is also found in the spectrum of lysine ethyl ester and suggests that also in this new amino acid the two nitrogen atoms are connected by a C_5-chain. The presence of a peak at m/e 30 indicates a $-CH_2-NH_2$ group. These conclusions may be summarized as follows:

Peaks Observed	*Conclusions*
274	molecular weight, even number of nitrogens
↓ $-73\,(COOC_2H_5)$	
201	$R-\underset{\displaystyle N-R'}{\overset{\displaystyle |}{CH}}-COOC_2H_5$
↓ $-17\,(NH_3)$	
184	$-NH_2$ { $R-\underset{\displaystyle N-R'}{\overset{\displaystyle |}{CH}}-COOC_2H_5$
↓ $-28\,(C_2H_4)$	
156	$-NH_2$ { $R-\underset{\displaystyle N-R'}{\overset{\displaystyle |}{CH}}-COOC_2H_5$ } $-COOC_2H_5$
84 $(C_5H_{10}N)$	$H_2N-(CH_2)_4-\underset{\displaystyle N-}{\overset{\displaystyle |}{CH}}-$

and lead to structure

$$H_2N-CH_2CH_2CH_2CH_2\underset{\displaystyle NH}{\overset{\displaystyle |}{CH}}-COOH$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\quad |$$
$$\qquad\qquad\qquad\qquad\qquad\qquad CH_3CH-COOH$$

for this new amino acid which was now named lysopine. Synthesis of this structure proved it to be correct.

* In the mass spectra shown in the figures of this paper all peaks are included which have at least 2% of the intensity of the highest peak. In the magnified regions (on the right hand side of the slanted arrows) the minimum peak shown is decreased for the same factor.

Quantitative analysis of amino acid mixtures

Quantitative analysis by mass spectrometry of complex mixtures is widely used in the petroleum industry and its speed and accuracy, in fact, provided the incentive for the development of the mass spectrometers now commercially available. In the discussion of the mass spectra of amino acid ethyl esters it was already pointed out that they exhibit one or very few peaks which are rather intense. For the different amino acids they fall at different mass numbers with little interference by others in most cases, a fortunate situation for quantitative analysis of multicomponent mixtures. Most often the $(R\text{–}CH\text{–}NH_2)^+$ peak is the most intense and useful one: 30 in glycine ethyl ester, 44 in alanine, 58 in α-aminobutyric acid, 72 in valine etc., whereas fragments involving the R-group are often more pronounced in the esters of amino acids containing further functional groups in the sidechain, e. g. mass 61 in methionine ester, 84 in lysine, 107 in tyrosine etc. The most intense peaks in leucine, isoleucine and hydroxyproline, however, fall at the same mass number (86) and, if more than one of these three are present in the mixture, additional characteristic peaks have to be used (e. g. 44, 69 and 68). The composition of a mixture of amino acid ethyl esters can, therefore, be calculated from a single mass spectrum of the mixture in the simplest way by the subtraction of the contributions of one component after the other once a calibration spectrum of each one of the pure esters is available.

The major problem involved in the application to mixtures of amino acids is, however, not one of mass spectrometry or arithmetic, but the fact that the sample to be analyzed is composed of free amino acids which first have to be converted to the free esters and then introduced into the instrument without any fractionation in spite of the widely different boiling points of the components. Furthermore, the formation of diketopiperazines has to be avoided and, most important, the whole procedure should not require more than 1 to 3 mg of the original mixture.

All these requirements were met by the following method:[3] The mixture of amino acids is refluxed with ethanolic hydrochloric acid for three hours, the solvent evaporated in a desiccator and the remaining ester hydrochlorides dissolved or suspended in dichloromethane. Dry ammonia is then passed briefly through this mixture liberating the free esters while ammonium chloride precipitates which is then filtered off. The dichloromethane is now distilled off in a special tube which is then connected with the port on the inlet system of the mass spectrometer. The tube with the ester mixture and some remaining solvent is now cooled to $-60°$ and the latter is pumped off through the pumping system of the instrument. After closing the pump valves the tube is heated quickly to 160° which quantitatively introduces the esters into the mass spectrometer. Since the amino esters, when free, are either in a dilute solution or at very low temperature or in the vapor phase under low pressure the dimerization is prevented. Dichloromethane was chosen as a solvent because it dissolves most of the ester hydrochlorides, has a low boiling point and its mass spectrum interferes least with the spectra of the esters which is important because the last traces of solvent are difficult to remove.

Results obtained on synthetic mixtures are summarized in Table I. The values found are in good agreement with the theoretical ones; tyrosine and ornithine deviate most, the former probably because of the size and polarity of the molecule, the latter because ornithine ester cyclyzes to 3-amino-piperidon-2, the spectrum of which is not very favorable for such an analysis. Difficulties are still encountered with sulfur containing amino acids (except methionine) although it should be stated that we have analyzed lanthionine in mixtures after desulfurization with Raney-nickel to alanine, which in principle is also possible with cystine and cysteine.

Table I: Results of quantitative analyses on four synthetic mixtures of amino acids

Amino Acid	Mass Number[a]	Mixture A		Mixture B		Mixture C		Mixture D	
		Calculated	Found *h*	Calculated	Found *i*	Calculated	Found *j*	Calculated	Found *j*
gly	30[b]	2.64[g]	2.63	2.16	2.09	–	–	–	–
ala	44[b]	1.73	1.75	1.31	1.28	–	–	–	–
abu	58[b]	1.85	1.84	–	–	–	–	–	–
val	72[b]	1.87	1.83	–	–	–	–	1.60	1.65
leu	44[c]	–	–	–	–	1.41	1.38	–	–
ileu	69[c]	–	–	–	–	1.66	1.68	–	–
pro	70[b]	1.30	1.32	–	–	1.18	1.17	–	–
ser	60[b]	–	–	1.44	1.42	–	–	–	–
thr	74[b]	–	–	0.98	0.98	–	–	–	–
opro	86[b]	1.40	1.41	1.01	1.00	–	–	–	–
met	61[d]	–	–	–	–	1.43	1.44	–	–
phe	120[b]	1.00	1.00	–	–	–	–	1.00	1.03
tyr	107[e]	–	–	–	–	–	–	0.80	0.76
asp	116[b]	1.03	1.05	1.00	1.00	1.00	1.02	1.19	1.21
glu	84[c]	1.40	1.41	–	–	–	–	–	–
orn	69[f]	3.10	3.27	–	–	–	–	–	–

(*a*) m/e of peak used for calculation. (*b*) $(R-CH-NH_2)^+$ fragment of $R-CH(NH_2)CO_2C_2H_5$. (*c*) Originates through further fragmentation of (*b*). (*d*) $(CH_3-S-CH_2)^+$. (*e*) $(HO\text{-}C_7H_6)^+$. (*f*) Fragment of orn-lactam. (*g*) Values represent molar ratios. (*h*) Average of 2, (*i*) average of 4, (*j*) average of 3 determinations.

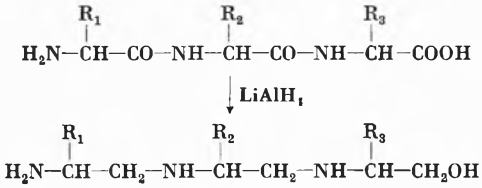
The major advantage of the mass spectrographic analysis over existing methods is the relatively short instrument time (30 to 40 min.) required, while the preparation of the samples itself can be done in larger series. It is thus possible to analyze accurately a large number of samples in a relatively short time, which might be necessary in connection with the structure determination of a protein, studies of metabolism, multiple analyses to increase the accuracy by statistical treatment of the results, etc. The method has little advantages, however, for only occasional, single analyses which might just as well be done with simpler instruments.

Determination of the amino acid sequence in small peptides

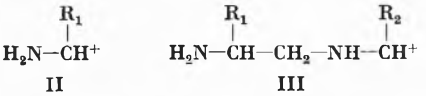
This problem is an important one for the determination of the structure of peptides and proteins which is

commonly attacked by enzymatic and/or hydrolytic degradation to, in most cases very complex, mixtures of small peptides, the amino acid sequence of which has to be determined in order to fit them back to the original structure. This approach obviously involves the separation of as many as possible of these peptides and sequence analysis on them.

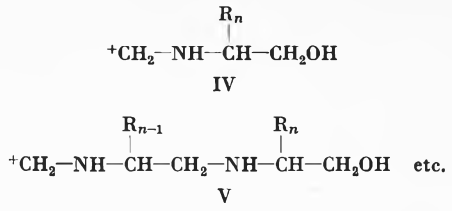
For the mass spectrometric determination of the sequence it is necessary to convert the free peptide into a more volatile derivative, which is of even greater importance for peptides containing polyfunctional amino acids. In consideration of the much lower boiling points of amines compared with the corresponding amides we have chosen to remove completely all the amide and carboxyl groups of the peptide by reduction to the corresponding amino alcohol with lithium aluminum hydride in tetrahydrofuran.[4] The polyamino alcohol thus obtained from a peptide still has the same structure as the parent peptide as far as type and location of the side chains are concerned except that the —CO— groups are replaced by $-CH_2-$:



In these molecules we now have a succession of —NH—C—C—NH— groups, the carbon-carbon bonds of which are easily cleaved on electron impact as discussed previously in the cases where there was only one nitrogen atom attached. These polyamino alcohols, therefore, give rise to fragments due to cleavage of any one of these bonds and since the positive charge is better stabilized on the more substituted carbon atom the fragment pertaining to the N-terminal part of the molecule will be somewhat more pronounced. In general, the spectrum of such an amino alcohol will exhibit a very strong peak due to the N-terminal fragment (II)



the mass number of which identifies the N-terminal amino acid (except differentiation of leucine and isoleucine). A somewhat less intense peak will be found $42 + R_2$ mass units higher (fragment III) which establishes now the amino acid next to the N-terminal and so forth. The reason for the decrease in intensity towards higher mass numbers is, aside from some instrumental factors, due the increased number of possibilities for further fragmentation of the larger particles. The sequence deduced in this way is further borne out by the corresponding peaks due to the fragments originating from the C-terminal part of the molecule:



The peak due to fragment IV found at mass $73 + R_n$ reveals the C-terminal amino acid, the one found $42 + R_{n-1}$ mass units higher (fragment V) the next amino acid and so on. The molecular weight of the polyaminoalcohol is deduced from the "M + 1 peak" which is small but very characteristic in these compounds and further corroborates the interpretation as do some peaks due to loss of any one of the side chains from the otherwise intact molecule. Such a spectrum is shown in Figure 9. It should be noted that the spectrum is of the amino alcohol obtained from acetyl-L-leucyl-L-alanyl-L-proline by reduction with lithium aluminium deuteride and therefore contains a $-CD_2-$ group for each —CO— group of the peptide which thus shifts the

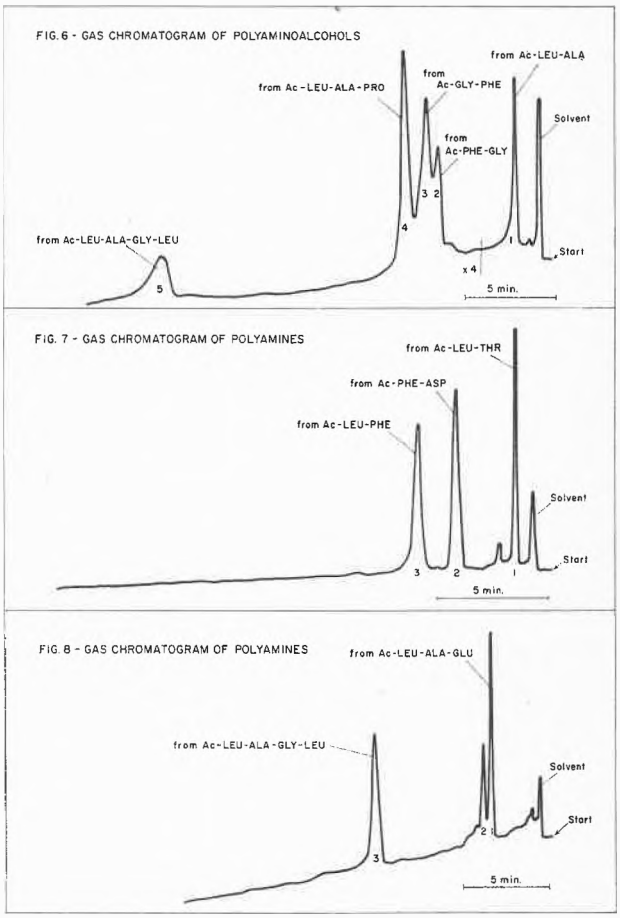


Fig. 6. Gas chromatogram of the polyamino alcohols obtained from a mixture of five acetyl-di- to tetrapeptides

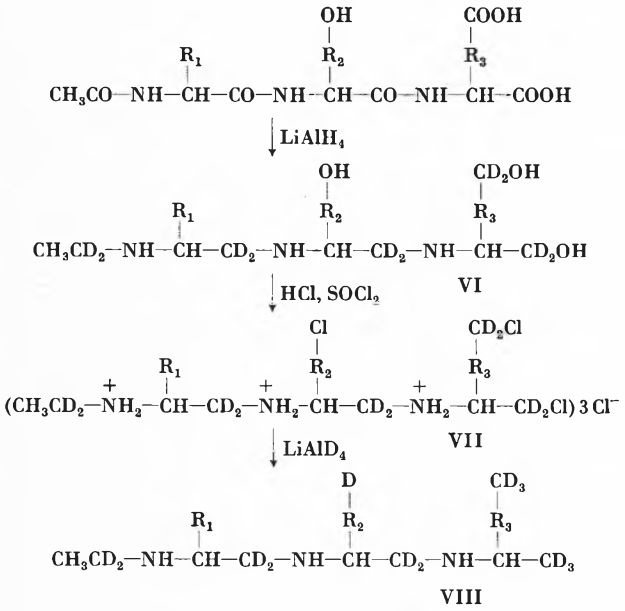
Fig. 7. Gas chromatogram of the polyamines obtained from a mixture of acetyl-peptides containing polyfunctional amino acids

Fig. 8. Gas chromatogram of the polyamines obtained from a mixture of an acetyl-tripeptide containing glutamic acid and an acetyl-tetrapeptide

mass numbers of the fragments up for two mass units per such group contained in it. For reasons outlined below we prefer the use of the deuteride over the hydride. The acetyl peptides were employed in our studies because these were easier synthesized as model compounds and also seem to give better results in general. The largest peptide we have as yet tested successfully with this method was a pentapeptide, Ac-leu-ala-gly-leu-pro.

The relatively good volatility of these derivatives made it attractive to attempt the separation of mixtures of these by gas chromatography and this was accomplished by the use of Apiezon L as stationary phase on columns operated at 260°. Figure 6 shows such a separation of a five component mixture of polyamino alcohols derived from di-, tri- and tetrapeptides. The fractions emerging from the column can be collected and are ready for introduction into the mass spectrometer without any further treatment. Figure 9, in fact, is the spectrum obtained from fraction 4 of this chromatogram. Very satisfactory spectra were obtained even from not completely resolved fractions (2 and 3) by collecting only the less overlapping portions.

The gas chromatographic separation of these products makes it possible to accomplish the reduction of many peptides in one experiment followed by separation into the individual components in a very fast and efficient way if compared with the rather tedious one of separating the peptides themselves by column or paper-chromatography or electrophoresis which always yield the fractions in the form of very dilute, aqueous solutions frequently containing large amounts of salts.



If the peptide contains one or more hydroxy or dicarboxylic amino acids the resulting reduction product will contain two or more hydroxyl groups. Such compounds emerge very late, if at all, from the gas chromatograph because of their high boiling point and polarity. The polyamines resulting by reductive elimination of all hydroxyl groups were expected to be very well suited for this type of separation, and to give good mass spectra which still allow the unambiguous elucidation of the sequence in the original peptide if the position of all the former hydroxyl groups were marked by deuterium. This was accomplished by a second reductive step involving the conversion of the polyamino alcohol (VI) in form of its hydrochloride into the alkyl chloride (VII) and its reduction with lithium aluminum deuteride to the amine (VIII).[5]

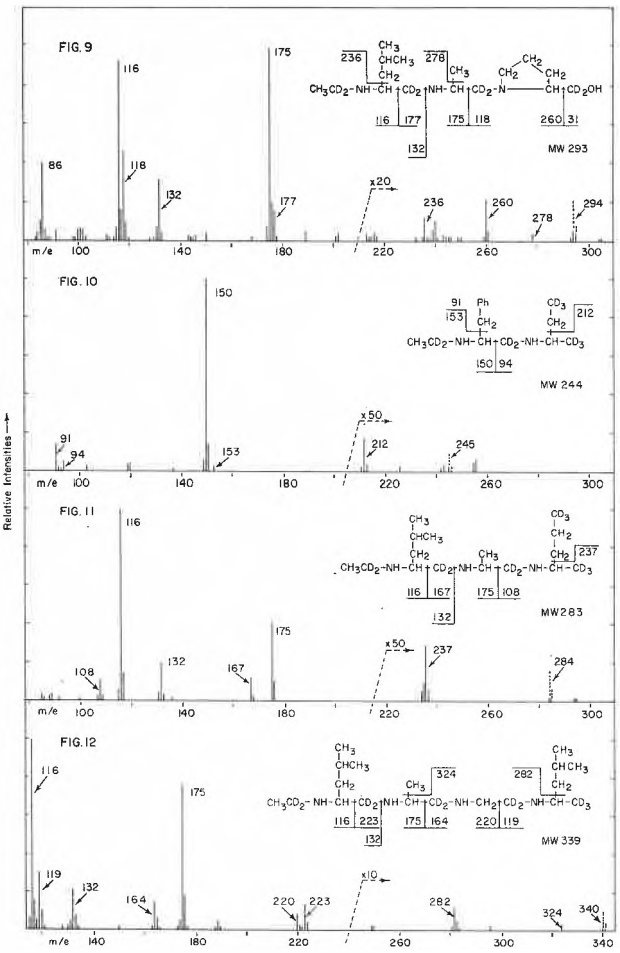


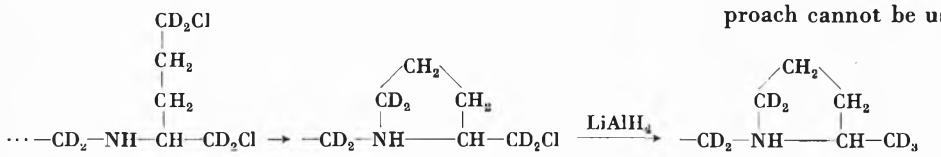
Fig. 9. Mass spectrum of fraction 4 of Figure 6*
Fig. 10. Mass spectrum of fraction 2 of Figure 7*
Fig. 11. Mass spectrum of fraction 1 of Figure 8*
Fig. 12. Mass spectrum of fraction 3 of Figure 8*

Figures 7 and 8 exhibit gas chromatographic separations of such polyamines obtained from mixtures containing peptides of hydroxy- or dicarboxylic amino acids. The reduced trailing of the peaks and the much shorter retention time of the tetramine in Figure 8 compared with the corresponding tetramino alcohol in Figure 7 should be noted.

* The dotted peaks indicate the increase in relative intensity of the M + 1 peak and its isotope peak on switching the ion source from "focused" to "non-focused".

The mass spectra of fraction 2 in Figure 7 and fractions 1 and 3 in Figure 8 are shown in Figures 10, 11 and 12 respectively. The fragmentation pattern closely follows the one outlined above for the polyamino alcohols and the peaks due to cleavage of the —NH—C—C—NH— are even more pronounced. In all these cases both reduction steps were done with lithium aluminum deuteride which makes it possible to differentiate between the side chains due to α-aminobutyric acid, threonine and aspartic acid which are converted to CH_3CH_2—, CH_3CHD— and CD_3CH_2— respectively, or between valine and glutamic acid, proline and hydroxyproline, and alanine and serine in a similar manner.

The mass spectrum of fraction 2 in Figure 8 revealed it to be the product of a side reaction in the reduction of the alkyl chloride intermediate derived from the peptide Ac-L-leu-L-ala-L-glu:



The formation of the five-membered ring appears to be so fast that it competes with the reductive elimination of the chlorine atom although the normal product is still the major one. The presence of the two deuterium atoms makes it possible to distinguish such compounds from the ones derived from proline.

As was already implied earlier the spectra of polyamino alcohols or polyamines derived from peptides differing only in the replacement of a leucyl residue with isoleucyl will have the same major peaks because their side chains have the same mass. They will not, however, be identical because the loss of the side chain is more favored in the case of the isoleucyl derivative due to the difference in the position of the methyl branch. Such peptides can be differentiated only if an authentic sample of one of the two compounds is available.

In the polyamine spectra, as with the polyamino alcohols, no appreciable peak is found for the molecular weight but the characteristic one at one mass unit higher. It might be mentioned at this point that there is consistently a peak at 11 mass units above the molecular weight (in the deuterated compounds a doublet at M + 11 and 12) which, though as yet unexplained, is a further indication of the molecular weight 11 and 12 mass units lower.

It is of importance for the practical implications of this work that only about 1 to 1.5 mg of each original peptide in the mixture is required to obtain a good mass spectrum after the two-step reduction and separation and it will probably be possible to use even smaller amounts if necessary. Because of the small amounts of materials and solvents (0.5 to 2.0 ml) involved, all reactions are simply carried out in sealed ampoules and worked up in centrifuge tubes. The price of the lithium aluminum deuteride is no obstacle because of the minute amounts required.

ANDERSSON indicated in a brief note[8] lacking any data that he obtained a useful mass spectrum from trifluoroacetyl-ala-phe-OMe, and this approach has been extended to a simple tripeptide (Val-gly-ala).[9] It will be of interest to see whether this method can be used for larger peptides and those containing polyfunctional amino acids which present a more severe volatility problem.

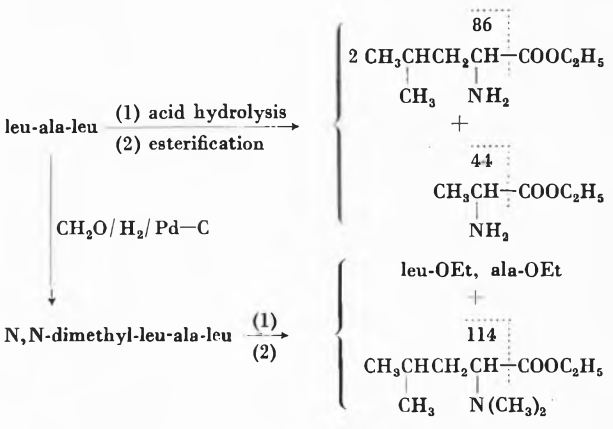
Determination of N-terminal amino acids and free amino groups

Although the N-terminal amino acid in a small peptide can be deduced from the mass spectrum of the corresponding polyamino alcohol or polyamine, this approach cannot be used for peptides which are too large to be subjected to this method without prior degradation to smaller peptides in which case the N-terminal amino acid has to be determined first.

The well-known dinitrophenylation of SANGER[10] seems to have only two minor drawbacks, the cleavage of the DNP-amino acids which occurs to some extent under the conditions of acid hydrolysis giving rise to difficulties in the quantitative estimation of the end groups and, secondly, the identification of the products which, on a small scale, has to be done by R_f-values.

We have combined the dimethylation of the free amino groups[11] with mass spectrometric analysis of the mixture of the amino acid ethyl esters obtained after methylation of the peptide, total hydrolysis and esterification.[6] The dimethyl amino groups have the definite advantage of higher stability towards acid hydrolysis. In the mass spectra of the esters, both the disappearance of the peaks due to the N-terminal amino acid and the appearance of certain peaks 28 mass units higher are used in the interpretation.

We have used a number of small peptides as models for this approach as illustrated on leucyl-alanyl-leucine:



Comparison of the spectra of the ester mixtures obtained before and after dimethylation showed, in the one derived from the methylated peptide, the intensity at mass 44 unchanged whereas mass 86 was decreased to one-half and a new peak at 114 appeared, correspondingly.

Since some of the dimethylated amino esters give rise to peaks isomeric with the ones from unmethylated esters (e. g. dimethylalanine and valine) quantitative evaluation of the spectra is necessary in the more complex cases. The use of isotopically labeled formaldehyde will be of added advantage.

The free amino groups in lysine, which would lead to N_ε-dimethyllysine, could be determined in the same manner.

Conclusions

Mass spectrometric methods have been devised for the solution of a number of basic problems involved in the determination of the structure of peptides and proteins. We feel that the most important contribution appears to be the one for the direct determination of the amino acid sequence in small peptides—up to pentapeptides or, possibly, even higher—since it is much faster and at least as reliable as conventional methods. Gas chromatographic separation of the derivatives used in the sequence determination further facilitates the work on the complex peptide mixtures obtained on partial hydrolysis of proteins. The mass spectrometric quantitative analysis of amino acid mixtures might find its place in those laboratories where a large number of such determinations is required in a relatively short period of time. The value of mass spectrometry for the elucidation of structure as illustrated in the case of lysopine needs no additional comment.

[9] Private communication (by Professor F. WEYGAND) of results obtained by the Swedish group.

[10] F. SANGER, *Biochem. J. 39* (1945) 507.

[11] V. M. INGRAM, *J. Biol. Chem. 202* (1953) 193.