volume 1

CHEMISTRY AND BIOCHEMISTRY OF AMINO ACIDS, PEPTIDES, AND PROTEINS

edited by
Boris Weinstein

CHEMISTRY AND BIOCHEMISTRY OF AMINO ACIDS, PEPTIDES, AND PROTEINS

A Survey of Recent Developments

◄ *Volume 1* ►

Edited by

BORIS WEINSTEIN

DEPARTMENT OF CHEMISTRY
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON

MARCEL DEKKER, INC., New York 1971

MARCEL DEKKER, INC., 95 Madison Avenue, New York, New York 10016

LIBRARY OF CONGRESS CATALOG CARD NUMBER 75-142896

ISBN 0-8247-1768-6

PRINTED IN THE UNITED STATES OF AMERICA

INTRODUCTION TO THE SERIES

The amide bond is one of the less reactive organic functional groups, yet it serves as the cornerstone for the building of the many peptides and proteins found in living systems. The evolving science of molecular biology has served to stress again that the chemistry and biochemistry of amino acids, peptides, and proteins is interwoven into a complex pattern, which on closer examination is found to be dependent on a host of secondary factors associated with individual compounds. There has been a need for a new review series in this area, especially if the interrelationships between the various disciplines are to be discussed in a detailed fashion. In an ideal sense, each volume should contain some chapters on recent developments and applications of established techniques, whereas others might describe the background and problems for topics still under investigation. Too, the subjects encompassed here do permit a variety of treatments without undue duplication or specialization.

One need not remind the reader of the many life processes that are dependent upon specific amino acid, hormone, and enzyme systems. Each functions in a very unique fashion, yet, in the end, they must involve the reactions of fundamental organic chemistry. Sometimes this point is overlooked, and it will be restated in greater detail through the series. To balance the scale, the brief comment is made here that new protecting, labeling, and coupling agents are always desirable, but these must be put to the test by the synthesis or degradation of actual compounds, for which practical use exists in Nature.

It is anticipated that these volumes can be useful both to the specialist and nonspecialist, and may provide a reference point to those who may do research in a broad region, or to the active worker in a small field. Most importantly, these volumes can serve the general purpose of presenting various points of view on the amide bond to interested observers, who, at present, are unknown to one another.

BORIS WEINSTEIN

Seattle, Washington
December, 1970

PREFACE

This volume, the initial one in a new series, ranges over a variety of topics of much interest to active workers in the fields concerned with the biochemistry and chemistry of peptides.

The first chapter by John W. Westley discusses the assignment of configuration to amino acids and some peptides, with special reference to the use of gas-liquid chromatography. The second by Charles H. Stammer surveys the chemistry of the antibiotic cycloserine. The next one by Shumpei Sakakibara summarizes the application of hydrogen fluoride to peptide and protein chemistry. The review by Danute E. Nitecki and Joel W. Goodman gives an account of the intriguing glutamyl peptides produced by certain bacteria. Finally, Mary Païs and F.-X. Jarreau describe a new class of natural products, peptide alkaloids.

The authors are gratefully thanked for their contributions, and any errors, omissions, or delays are the responsibility of the editor.

It is hoped that this first survey, as well as succeeding ones, will be of aid to the many biochemists, biologists, and chemists who are interested in the study of amino acids, peptides, and proteins.

BORIS WEINSTEIN

Seattle, Washington
December, 1970

CONTRIBUTORS TO THIS VOLUME

J. W. Goodman, Department of Microbiology, University of California, San Francisco, California

François-Xavier Jarreau, C.N.R.S.-Institut De Chimie Des Substances Naturelles, 91-Gif-Sur-Yvette, France

D. E. Nitecki, Department of Microbiology, University of California, San Francisco, California

Mary Païs, C.N.R.S.-Institut De Chimie Des Substances Naturelles, 91-Gif-Sur-Yvette, France

S. Sakakibara, Institute for Protein Research, Osaka University, Osaka, Japan

Charles H. Stammer, Department of Chemistry, University of Georgia, Athens, Georgia

John W. Westley, Chemical Research Department, Hoffmann-La Roche Inc., Nutley, New Jersey

CONTENTS

◄ *chapter 1* ►

THE OPTICAL ANALYSIS OF AMINO ACID DERIVATIVES BY GAS CHROMATOGRAPHY

John W. Westley

CHEMICAL RESEARCH DEPARTMENT
HOFFMANN-LA ROCHE INC.
NUTLEY, NEW JERSEY

I. OPTICAL PURITY OF AMINO ACIDS

In order to synthesize a biologically active peptide, the optical purity of the amino acid derivatives involved in the synthesis is just as essential as their chemical purity. For instance, if amino acids containing only 1% D isomer were used to synthesize a polypeptide of 100 residues, only 37% of the product would have the all L configuration(*1*). A more common problem in peptide synthesis than optical impurities in the starting materials is racemization during the coupling steps. For these reasons, it is essential to have very sensitive methods of determining optical impurities in amino acids and their derivatives.

The classical methods of determining the optical purity of amino acid derivatives have been polarimetry and various enzymatic procedures. The disadvantages of the polarimetric method are its lack of sensitivity(*1*) (2–3% error) and the inability of the technique to discriminate between optical and chemical impurities. The optical purity of a compound can be determined from polarimetry only if the rotation

of the pure enantiomer is known. Although semiempirical calculations can be used to determine the absolute rotation from structural considerations, this approach at present has limited applicability and the rotation of a given compound is usually just the value on which most authors agree. Enzymatic methods have been used extensively. The amino acid oxidases and decarboxylases and the proteolytic enzymes are all highly stereospecific but lack chemical specificity. There was therefore a need for a technique that would analyze amino acid derivatives for both optical and chemical impurities.

Raban and Mislow(*2*) have reviewed some of the modern approaches to the problem of optical purity determination. These include kinetic resolution, nuclear magnetic resonance, differential microcalorimetry, and the isotope dilution method. Of the techniques they describe, the one that has stimulated most attention has been the separation of enantiomers by chromatography. The advantage of this method, when successful, is that it provides a direct measurement of the amount of each enantiomer. In addition, there is a very high probability that chemical impurities are separated on the column from the amino acid derivative under investigation. The most efficient and sensitive of the chromatographic techniques is gas–liquid chromatography (GLC), which can be routinely used quantitatively at the 0.1 to 1 μg level. Thin layer chromatography has also been usefully applied in a number of cases for qualitative analysis.

II. OPTICAL ANALYSIS BY GAS CHROMATOGRAPHY

In spite of the nonvolatile nature of amino acids, many derivatives have now been found that are amenable to gas chromatographic analysis. In a recent review(*3*) by Blau on the analysis of amino acids by GLC, 196 references were cited up to the middle of 1965. This review and an earlier one by Weinstein(*4*) included discussions of the GLC separation of optical isomers. This type of resolution has been achieved by two approaches. The one investigated to the greatest extent involves the separation of enantiomers as diastereoisomeric derivatives. This is analogous to the classical resolution of asymmetric compounds, where separation was achieved as a result of the different solubility of diastereoisomeric salts. In the case of GLC, separation is attributable to both the solubility difference of covalently bonded diastereoisomers in the stationary phase and the difference in volatility of the two isomers.

A more direct method that has been developed exclusively by Gil-Av and his associates, relies on the difference in solubility of the enantiomers in an optically active stationary phase in order to achieve resolution. This second approach appears to be simpler and therefore more attractive but as it relies solely on the solubility effect, the selection of a suitable stationary phase is extremely critical.

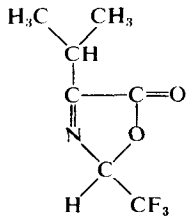
A. Resolution of Diastereoisomers

1. N-Trifluoracetyl peptide methyl esters. The first report on the application of gas chromatography to the separation of diastereoisomeric dipeptides came from Weygand's group(*5*) in 1960, when they succeeded in resolving N-trifluoracetyl (TFA)-L-alanyl-L-phenylalanine and N-TFA-L-alanyl-D-phenylalanine as their methyl esters on a 2-m long column. Using Golay-type capillary columns (50 m long), they had much greater success, resolving 34 peptides(*6*) of the general type, N-TFA-dipeptide methyl ester. The most successful column was coated with polyphenyl ether and operated at temperatures between 184° and 221°C. Resolution of some peptides was achieved on a column coated with polypropylene glycol and operated at 180°C. Only resolutions greater than 90% were useful in determining amounts of the minor diastereoisomer at levels below 1%. Using this criterion, about half the resolutions reported could be used to determine the degree of racemization in peptide synthesis. In those cases where the configuration of the two peaks in the GLC analysis was assigned, the LL diastereoisomer had a shorter retention time than its DL isomer. The only exception was N-TFA-alanyl-alanine methyl ester where the order was reversed.

Carbobenzoxy-dipeptide methyl esters were found to be not amenable to direct GLC analysis. Therefore to monitor racemization during the synthesis of carbobenzoxy-L-valyl-L-valine methyl ester, it was necessary to replace the carbobenzoxy group by TFA prior to GLC analysis. Using this technique, no racemization was observed when carbobenzoxyvalylvaline methyl ester was prepared under most conditions. This in turn excluded racemization during the replacement of the carbobenzoxy group. However, when N-TFA-valylvaline methyl ester was prepared directly by peptide coupling there was considerable racemization using the activated ester method (cyanomethyl and thiophenyl esters), the dicyclohexylcarbodiimide method, the carbonyl diimidazole method, or Woodward's reagent. Optically pure (>98%)

peptide was obtained only when the azide or vinyl ester method was employed. As vinyl esters are difficult to prepare from peptides, Weygand's group concluded that to minimize racemization in peptide synthesis, the amino terminal group should be protected by urethane type (e.g., carbobenzoxy, *tert*-butyloxycarbonyl)acyl groups and coupling should be carried out by the Curtius azide procedure.

It had been assumed that racemization during dicyclohexylcarbodiimide coupling was attributable to the formation of oxazolone intermediates, and this was confirmed by Weygand. 2-Trifluoromethyl-4-isopropylpseudooxazol-5-one (**1**) was prepared from D, L-valine and reacted with L-valine methyl ester to give 74% LL and 26% DL N-TFA-dipeptide methyl ester on GLC analysis.



(1)

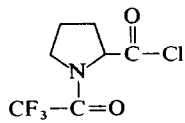
The presence of 2-trifluoromethyl-4-isopropylpseudooxazol-5-one was in fact demonstrated by GLC during the synthesis of N-TFA-valyl-valine methyl ester from N-TFA-valine and valine methyl ester using the dicyclohexylcarbodiimide procedure.

The coupling of racemic oxazolones and pseudooxazolones with L-amino acid methyl esters has been further studied (*7*) to determine the extent of sterically controlled asymmetric induction in this reaction. For instance, N-benzoyl-dipeptide methyl esters were prepared from racemic 4-isopropyl-2-phenyloxazolone and various L-amino acid methyl esters (*8*). After conversion to the corresponding TFA derivatives, they were analyzed by GLC. In all the tested examples of this reaction, the DL-dipeptide was the major product (cf. Ref. *6*).

To analyze for the optical purity of peptides containing more than two amino acid residues, preliminary degradation with methanolic HCl was necessary. The resulting peptide methyl esters were trifluoracetylated and the TFA-dipeptide methyl esters resolved by GLC. Using this technique, the performance of 16 coupling methods in the formation of carbobenzoxy-L-leucyl-L-phenylalanyl-L-valine *tert*-butyl

ester(9) has been appraised. Racemization during the partial hydrolysis step (24 hr at 70° with 8.5 *N* methanolic HCl) was found(*10*) to be less than 5%.

Following the success of Weygand in resolving TFA-dipeptide methyl esters, the Stanford group, because of their interest in the stereospecificity of biochemical processes(*11*), selected N-trifluoracetyl prolyl chloride (TPC)(*12*), (**2**) for study. This reagent was found to be an excellent resolving agent for asymmetric amines including amino acid esters(*13*, *14*) by GLC.



(**2**)

The choice of TPC as a resolving agent was based on the observation that proline does not racemize during acylation or peptide synthesis (oxazolone formation is not possible). In addition, the coupling reaction was rapid and quantitative, and the rigid conformation of the prolyl peptide bonds was expected to enhance differences in the physical properties of its diastereoisomers. The diastereoisomeric derivatives were prepared by first esterifying the amino acid with methanol and then coupling the ester with TPC in the presence of triethylamine. A control experiment was carried out first in which racemic reagent was coupled with the amino acid ester. In the case of an L-amino acid ester (AA), the reaction could be summarized:

$$\text{DL-TPC} + \text{L-AA} \rightarrow \text{D-TPC-L-AA} + \text{L-TPC-L-AA}$$

Resolution of the mixture of diastereoisomers was then attempted on various columns over a range of temperatures to determine the most suitable GLC conditions for optical analysis. When this had been established, optically pure reagent was reacted with the ester:

$$\text{L-TPC} + \text{L-AA} \rightarrow \text{L-TPC-L-AA}$$

Any D-amino acid present in the sample was converted to the L-TPC-D-AA derivative, which on GLC analysis has the same retention time as D-TPC-L-AA. This had been established in the control experiment,

and hence the relative amounts of L and D amino acid could be measured.

Efficient resolution of TPC amino acid methyl esters has been achieved (*15*) on short (5 ft × ⅛ in.), packed columns with 0.5% ethylene glycol adipate as the stationary phase. The technique has been used to demonstrate the stereospecific action of microorganisms in the soil (*16*). In all cases tested, the LD dipeptide was found to have a lower retention time than its LL isomer. It has been proposed (*17*) that this is attributable to the smaller molecular volume of LD dipeptides (*18*). However, the reverse order of retention had been previously observed by Weygand (*6*) using capillary columns to resolve N-TFA-dipeptide methyl esters. From another investigation, it now appears that TPC derivatives may be exceptional in their order of retention on packed columns. It was found (*19*) that GLC analysis of leucylleucine, leucylphenylalanine, and phenylalanylphenylalanine as their N-TFA-methyl esters resulted in the LL isomer always having the shorter retention time on a column (5% QF-1 on 60/80 DCMS-treated Chromosorb W) that gave the reverse order for all TPC derivatives tested. This anomalous behavior of prolyl peptides on GLC analysis has also been observed recently (*20*, *21*) in the case of diketopiperazines.

The sensitivity of the TPC method in determining the optical purity of amino acids depends on the degree of resolution of the particular TPC derivative under investigation. Thus aliphatic amino acids such as alanine and valine are easily resolved, whereas L-TPC-DL-phenylalanine methyl esters are considerably more difficult to separate. Another factor which determines the sensitivity of the method is the optical purity of the reagent. Commercially available TPC reagent (*22*) was found to be only 90% optically pure in the hands of the reviewer although it is possible to prepare reagent of at least 98.8% optical purity (*23*).

In a recent report (*24*), Weygand's group described the separation of TFA-L-prolyl-L-valyl-L-valine and TFA-L-prolyl-D-valyl-L-valine as their methyl esters on a 2-m packed column containing 0.5% FFAP (*25*) on Chromosorb G. The LDL-tripeptide had a shorter retention time than its LLL-diastereoisomer, which was consistent with the behavior of TPC-amino acid methyl esters (*13*, *14*).

2. 2-Chloralkanoyl-amino acid methyl esters. In 1965 the Stanford group reported the resolution of a number of N-2-chloralkanoyl-valine methyl esters and N-2-chloroisovaleryl amino acid methyl esters (*26*). Optically pure 2-chloralkanoic acids of known configuration were

prepared by Renard's procedure(*27*) from amino acids, followed by conversion to the acid chloride with thionyl chloride and coupling with amino acid methyl esters. These derivatives had the advantage of being more volatile than the analogous TPC diastereoisomers. A correlation was revealed between the steric bulk of the alkyl group and the degree of resolution of the diastereoisomeric N-chloralkanoyl valine methyl esters(*28*). Thus 2-chloro-3,3-dimethylbutanoic acid was found to be the best resolving agent of this type.

Lande and Landowne(*29*) have used the GLC resolution of L-2-chloropropionyl-DL-amino acid methyl esters to study the effect of the penultimate amino acid on the racemization of L-phenylalanine when protected aminoacyl-L-phenylalanine dipeptides are coupled with glycine. After peptide synthesis, the tripeptide was hydrolyzed (48 hr at 110° with constant boiling HCl in sealed, evacuated tubes). The hydrolysate was acylated using a mixed anhydride from L-2-chloropropionic acid and pivaloyl chloride, and esterified with diazomethane. GLC analyses were carried out on capillary columns. One of the columns used was coated with a polycyano silicone grease, XE-60, and a second with a polyalkylene glycol, 50-HB-2000. The results of this investigation showed that whether the N-terminal amino acid derivative of the peptide was carbobenzoxyglycine, bis-carbobenzoxylysine, or carbobenzoxy-γ-benzylglutamic acid, the amount of racemization on peptide coupling with dicyclohexylcarbodiimide was approximately the same (about 25%).

3. Tertiary-butyloxycarbonyl-amino acylamides. Although Weygand (*6*) demonstrated that urethane-type protecting groups were the derivatives of choice in peptide synthesis, his procedure was not capable of directly analyzing carbobenzoxy-dipeptide methyl esters. In 1967, the resolution of *tert*-butyloxycarbonyl (*t*-Boc)-amino acyl-4-methyl-2-pentylamides by GLC(*30*) was reported. This made it possible to investigate directly the racemization of the widely used *t*-Boc-amino acids under various coupling conditions. *t*-Boc derivatives of alanine, α-aminobutyric acid, valine, leucine, isoleucine, and proline were coupled to (−)-4-methyl-2-pentylamine using the carbonyldiimidazole, dicyclohexylcarbodiimide, water-soluble diimide, Woodward's reagent, mixed anhydride, and the *p*-nitrophenyl and N-hydroxysuccinimido active ester methods. The resulting *t*-Boc-amino acylamides were analyzed by GLC, and racemization ($> 2\%$) was observed only using the water soluble diimide procedure.

Attempts to resolve *t*-Boc-dipeptide methyl esters have been success-

ful only when one of the amino acids is proline(*31*). Resolution of the proline peptides (Table 1) was carried out on 5 ft × $\frac{1}{8}$ in. columns containing 3% EGS on Aeropak 30 (operated at 210° and with a nitrogen flow rate of 30 ml/min) and 5% QF-1 on Gas Chrom Q (at 240° and nitrogen flow of 30 ml/min).

TABLE 1
GLC RESOLUTION OF *t*-BOC-DIPEPTIDE METHYL ESTERS

Peptide	Column	Retention time, of diastereoisomers, min		Ratio of retention times $\frac{LL}{LD}$
		LD	LL	
Prolylalanine	QF-1	3.7	3.9	1.05
	EGS	2.4	2.8	1.07
Alanylproline	GF-1	3.5	3.7	1.06
	EGS	4.6	5.1	1.11
Prolylvaline	QF-1	3.3	3.7	1.12
	EGS	2.3	2.6	1.13
Prolylleucine	QF-1	3.7	4.0	1.08
	EGS	2.7	3.0	1.11
Prolylproline	QF-1	6.2	6.5	1.05
	EGS	9.0	10.1	1.12
Prolylphenylalanine	QF-1	11.3	12.4	1.10

The fact that only proline peptides were resolved and the order of retention (LD before LL) again demonstrates the exceptional behavior of proline derivatives on GLC analysis. This must in some way be associated with the rigid nature of the proline ring structure. The GLC analysis of these derivatives has been used(*32*) to confirm the steric homogeneity of *t*-Boc-dipeptide methyl esters after removal from Merrifield solid phase resin by transesterification with an anion exchange resin.

4. N-Trifluoracetyl-amino acid secondary alkyl esters. In the first three classes of diastereoisomers discussed, the two asymmetric centers were separated by an amide bond. Diastereoisomeric esters of amino acids have also been extensively investigated by GLC. These compounds are generally less well resolved than the amides, but because of their greater volatility they are more suitable derivatives for

the simultaneous optical analysis of mixtures of amino acids of widely differing structures. There is now a method available for the complete GLC analysis of protein amino acids(*33*) as their N-TFA-1-butyl esters. It should therefore be possible in the near future to carry out this analysis, with 2-butyl esters for instance, and measure steric purity of the amino acids in addition to their relative amounts.

The first report on the GLC resolution of diastereoisomeric amino acid esters appeared in 1963. Charles, Fischer, and Gil-Av(*34*) of the Weizmann Institute, in a paper describing the resolution of 2-*n*-alkanols as their α-hydroxypropionates, mentioned the separation of alanine and valine as their N-TFA-2-butyl esters on a capillary column coated with polypropylene glycol. They extended this work in 1965 to the resolution of alanine, valine, leucine, proline, phenylalanine, and glutamic acid(*35*) as their N-TFA-2-alkyl esters. Alanine, valine, and glutamic acid were resolved as their 2-butyl esters on a capillary column (150 ft × 0.01 in.) coated with trifluoropropylmethyl polysiloxane (FS-1265) that was operated at 140° and with a nitrogen flow rate of 1.5 ml/min. The column was also successful in resolving the 2-*n*-octyl esters of all the amino acids tested except glutamic acid, which would have too high a molecular weight as its di-2-*n*-octyl ester for rapid GLC analysis. Less efficient resolutions were carried out on a capillary column coated with polypropylene glycol. The authors concluded that 2-*n*-octanol is a convenient reagent for the optical resolution of amino acids since its enantiomorphs are commercially available (*36*), but 2-butanol is preferred for the acidic amino acids.

Pollock, Oyama, and Johnson(*37*) tested a number of capillary and packed columns for their efficiency in resolving N-TFA-amino acid 2-butyl esters. Ucon and the Carbowaxes were found to be the best stationary phases for this purpose. No resolution was obtained on capillary columns coated with Apiezon or Quadrol or packed columns with β-CDXA or FFAP as the stationary phase. Using a capillary column (300 ft × 0.01 in.) coated with Ucon and operated at 140° and with a helium flow of 2 ml/min, they were able to resolve the N-TFA-2-butyl esters of alanine, α-aminobutyric acid, valine, norvaline, isoleucine, leucine, alloisoleucine, norleucine, methionine, and phenylalanine. The threonine, serine, and aspartic acid derivatives were also partially resolved. A mixture containing six of the racemic amino acids was separated simultaneously on the Ucon capillary column.

In a later publication, Gil-Av and his co-workers(*38*) added hydroxyproline and lysine to the growing list of amino acids resolvable as their