ADVANCES IN PROTEIN CHEMISTRY

EDITED BY

M. L. ANSON
Continental Foods, Inc.,
Hoboken, New Jersey

JOHN T. EDSALL
Harvard Medical School,
Boston, Massachusetts

Associate Editor for the British Isles

KENNETH BAILEY
Trinity College,
Cambridge, England

VOLUME V

1949

ACADEMIC PRESS INC., PUBLISHERS
NEW YORK, N. Y.

PRINTED IN THE UNITED STATES OF AMERICA

CONTRIBUTORS TO VOLUME V

JAMES B. ALLISON, *Rutgers University, New Brunswick, New Jersey*

ALBERT CLAUDE, *The Rockefeller Institute for Medical Research, New York, New York**

JOSEPH S. FRUTON, *Department of Physiological Chemistry, Yale University, New Haven, Connecticut*

K. H. GUSTAVSON, *Swedish Tanning Research Institute, Stockholm, Sweden*

J. W. H. LUGG, *Department of Biochemistry, University of Melbourne, Melbourne, Victoria, Australia*

HAROLD P. LUNDGREN, *Western Regional Research Laboratory, U. S. Department of Agriculture, Albany, California*

THOMAS L. MCMEEKIN, *Eastern Regional Research Laboratory, U. S. Department of Agriculture, Philadelphia, Pennsylvania*

B. DAVID POLIS, *Eastern Regional Research Laboratory, U. S. Department of Agriculture, Philadelphia, Pennsylvania*

G. R. TRISTRAM, *Sir William Dunn Institute of Biochemistry, University of Cambridge, Cambridge, England*

* Present Address: *Institute Jules Bordet, Bruxelles, Belgium.*

CONTENTS

Plant Proteins

By J. W. H. Lugg, *Department of Biochemistry, University of Melbourne, Melbourne, Victoria, Australia*

Synthetic Fibers Made from Proteins

By Harold P. Lundgren, *Western Regional Research Laboratory, U. S. Department of Agriculture, Albany, California*

Some Protein-Chemical Aspects of Tanning Processes

By K. H. Gustavson, *Swedish Tanning Research Institute, Stockholm, Sweden*

Proteins, Lipids, and Nucleic Acids in Cell Structures and Functions

By Albert Claude, *The Rockefeller Institute for Medical Research, New York, New York*

The Synthesis of Peptides

By JOSEPH S. FRUTON

Department of Physiological Chemistry, Yale University, New Haven, Connecticut

Contents

I. Introduction and Nomenclature

The objectives of the search for satisfactory methods of peptide synthesis were clearly stated by Fischer and Fourneau (149) in the

memorable paper which initiated the systematic exploration of this field of study:

"Der Gedanke, die aus den Proteinstoffen durch Hydrolyse entstehenden Aminosäuren durch Anhydridbildung wieder zu grösseren Komplexen zu vereinigen, ist schon seit längerer Zeit von verschiedenen Forschern experimentell behandelt worden. Wir errinern nur an die Anhydride von Schaal (262), ihre Verwandlung einerseits in den kolloidalen Polyasparaginharnstoff von Grimaux (199), anderseits in die Polyaspartsäuren von H. Schiff (263), ferner an die Versuche von Schützenberger (270) über die Vereinigung verschiedener Aminosäuren (Leucine und Leuceine) mit Harnstoff durch Erhitzen mit Phosphorsäureanhydrid, an die ähnlichen Beobachtungen Lilienfelds (238) über die Wirkung von Kaliumbisulfat, Formaldehyd und anderen Kondensationsmitteln auf ein Gemisch von Aminosäureestern und endlich an die Angaben von Balbiano und Trasciatti (39) über die Verwandlung des Glycocolls in ein hornartiges Anhydrid durch Erhitzen mit Glycerin. Aber alle von ihnen beschriebenen Produkte sind amorphe, schwer characteresierbare Substanzen, über deren Struktur man ebensowenig wie über den Grad ihrer Verwandtschaft mit den natürlichen Proteinstoffen etwas sagen kann.

"Will man auf diesem schwierigen Gebiete zu sicheren Resultaten kommen, so wird man zuerst eine Methode finden müssen, welche es gestattet, successive und mit definierbaren Zwischenstufen die Moleküle verschiedener Aminosäuren anhydridartig aneinander zu reihen."

Since the enunciation of this view, it has become abundantly clear that one of the principal contributions of the organic chemist to the study of the structure and reactions of proteins has been indeed the development of several techniques for the synthesis of compounds in which amino acids are joined to one another "anhydridartig" by means of acid amide linkages. The services of peptide synthesis to protein chemistry have been many and various. The finding of synthetic peptides and peptide derivatives which are hydrolyzek by crystalline enzymes specifically adapted to the hydrolysis of proteins has buttressed the theory, first expressed by Hofmeister (217) and Fischer (126), that in proteins the peptide bond represents the most general type of linkage between the individual amino acid residues.* In addition, modern methods of peptide synthesis have made possible the preparation of special peptides of low molecular weight and known chemical structure

* In stressing the importance of the peptide bond in the structure of proteins, it is not intended to exclude the possibility that other types of covalent linkage may also play a significant role in the architecture of the protein molecule (99,135). Of particular interest in connection with the search for labile bonds which may be involved in the phenomena associated with protein denaturation is the recent suggestion of Linderstrøm-Lang and Jacobsen (239) that cysteine residues of proteins might participate in the formation of thiazoline rings. Similar consideration might be given to the participation of serine (or threonine) residues in oxazoline groupings (46). The recent synthesis, by Ehrensvärd and Davidssohn (119), of labile peptides with thioiminoether linkages is a further experimental approach in this direction.

for use as models in the examination of several physical and chemical properties of proteins. Thus, the use of synthetic peptides has notably facilitated the interpretation of data on the acid–base relationships of proteins (101). Furthermore, in the study of the reactions of proteins with various chemical reagents, parallel experiments with peptides or peptide derivatives have frequently clarified the results observed with proteins. Of the numerous examples in the recent literature one may cite the study of the iodination of tyrosine-containing proteins (212), or the investigation of the action of mustard gas (250) and the nitrogen mustard gases (191) on proteins.

Apart from their importance in providing simple models for study of the enzymatic degradation, physical properties, or chemical reactions of proteins, the newer techniques of peptide synthesis have been invaluable in the final proof of the chemical structure of several physiologically important substances, such as glutathione (202) and carnosine (272). The discovery that the antibiotics gramicidin, tyrocidine, and gramicidin S are peptides (219,282) and the recent reports that there is, in pancreatic hydrolyzates of several proteins, a factor (or factors) of peptide nature which promotes the growth of certain microorganisms and the rat (299,302) offer further opportunity for the fruitful use of the methods of peptide synthesis to establish the chemical structure and to study the physiological action of peptides of biological interest.

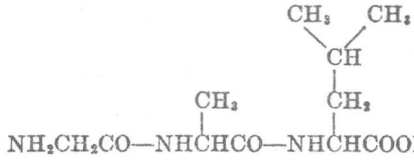
The development, in recent years, of new methods for the separation of amino acids and peptides (103,242) has led to a renewed interest in the products of the partial hydrolysis of proteins (281). It is clear, however, that, whatever methods are used to separate and identify peptides obtained from proteins, the conclusive evidence for the identity of such peptides must come from the comparison of the isolated material with synthetic peptides of known structure. This is the procedure that allowed Fischer and Luniak (159) to establish definitely that the peptide isolated by Osborne and Clapp (253) from a gliadin hydrolyzate was indeed L-prolyl-L-phenylalanine. In a similar manner, Stein, Moore, and Bergmann (278) demonstrated the presence of glycyl-L-alanine and L-alanylglycine in partial hydrolyzates of silk fibroin.

Enough has been said in the foregoing to justify the importance attached to peptide synthesis as a tool in the study of proteins. The purpose of this review is to survey the available methods for the synthesis of peptides. These methods will be evaluated, whenever possible, with regard to their relative difficulty, their adaptability to meet the many problems encountered in amino acid chemistry, and the yield and purity of the products of synthesis. Although the preparation of peptides containing nonprotein amino acids will occasionally be men-

tioned, primary emphasis will be placed on the synthesis of peptides of amino acids definitely known to be formed upon protein hydrolysis.

In what follows, the configuration of the amino acids will be given in accordance with the report of the Editorial Board of the *Journal of Biological Chemistry* (284). The amino acid residues will, in general, be designated by adding the suffix "yl" to the roots of the names of the free amino acids. Thus, for glycine, the term will be glycyl; for proline, prolyl; for tyrosine, tyrosyl; and so forth. Several departures from this rule will be noted, however. The peptides of cysteine will not be designated "cysteyl" but "cysteinyl" peptides, and for the cystine peptides, the designation will be "cystinyl" rather than "cystyl." In the case of peptides of glutamine, the amino acid residue will be termed "glutaminyl" to differentiate it from "glutamyl," which refers to glutamic acid. Similarly, for asparagine peptides, the term will be "asparaginyl" and, for aspartic acid peptides, it will be "aspartyl." In addition, the term "tryptophyl" will be used to designate the amino acid residue of tryptophan.

Since this nomenclature implies that an amino acid has been converted to an acyl group, it follows that the amino acid residues in a polypeptide should be listed in the sequence of substitution at the amino group of the adjacent amino acid. The tripeptide glycyl-L-alanyl-L-leucine has, therefore, the following formula:



A brief discussion of the configurational relationships of peptides appears necessary at this point since frequent mention will be found in the literature of racemic peptides which contain more than one optically active amino acid. It is clear that, in the case of a racemic dipeptide containing two optically active amino acids, *i.e.*, DL-leucyl-DL-alanine, four isomers are possible: (*a*) D-leucyl-D-alanine, (*b*) L-leucyl-L-alanine, (*c*) L-leucyl-D-alanine, and (*d*) D-leucyl-L-alanine. Two racemates may be expected; one composed of forms *a* and *b*, and another composed of forms *c* and *d*. Separation of the two racemates may sometimes be achieved by taking advantage of their differences in solubility. It is then customary to designate the less soluble form by the letter "A" and the more soluble racemate by the letter "B."

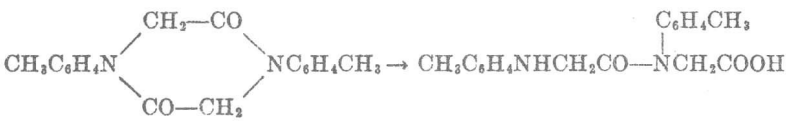
Reports will be encountered in the literature of the synthesis of peptides containing more than one optically active amino acid and in which one amino acid residue is present in a single configuration, while another is given as the DL form, *e.g.*, DL-alanylglycyl-L-glutamic acid. Such preparations are clearly mixtures of diastereoisomers. When attempts to separate the two forms are not successful, three possibilities may be envisaged: (*a*) the two isomers have very similar solubilities, thus making their separation by fractional crystallization difficult; (*b*)

they are isomorphic and form mixed crystals; or (*c*) they form addition compounds in stoichiometric proportions. Since it cannot be predicted beforehand whether, after the synthesis of a mixture of such diastereoisomers, a pure peptide can be isolated, it appears preferable, whenever possible, to perform peptide syntheses with optically active amino acids under conditions where racemization is avoided. Furthermore, in the study of the specificity of proteolytic enzymes or of the properties of model compounds related to proteins, it is usually desirable to have peptides containing optically active, rather than the racemic, forms of the amino acids. Greater interest attaches itself, therefore, to methods which permit the synthesis of such optically active peptides.

II. General Methods of Peptide Synthesis

1. *Synthesis of Dipeptides by Partial Hydrolysis of Diketopiperazines*

In 1888 Abenius and Widmark (38) found that ditolyldiketopiperazine could be partially cleaved by acid hydrolysis:



An analogous reaction of aliphatic compounds was not observed until 1901, when Fischer and Fourneau (149) hydrolyzed glycine anhydride by heating it briefly with concentrated hydrochloric acid and thus obtained glycylglycine, the simplest representative of the group of substances under discussion in this review. The method used by Fischer and Fourneau still is a convenient method for making this dipeptide, especially in view of the recent development of an excellent procedure for the synthesis of glycine anhydride from glycine (261). Fischer noted, however, that brief acid hydrolysis of DL-alanine anhydride and of DL-leucine anhydride did not yield the expected dipeptides (125). Fischer later found that brief hydrolysis with dilute sodium hydroxide at room temperature would also yield dipeptides from diketopiperazines (132). In addition to glycylglycine, one of the two possible racemic forms of DL-alanyl-DL-alanine was prepared in this way. It soon became clear, however, that the method was not generally applicable to the synthesis of dipeptides containing optically active amino acids, for, when Fischer attempted to prepare L-alanyl-L-alanine from the corresponding anhydride, the product which resulted was partially racemic (153). Similarly, treatment of L-tyrosine anhydride with alkali caused appreciable racemization (166). As was shown by the later systematic studies of Levene (237) and Bergmann (82), among others, the racemization of optically

active diketopiperazines is favored by alkali. It is obvious, therefore, that this procedure is essentially limited to work with glycine anhydride. In fact, it is frequently convenient to treat glycine anhydride with sodium hydroxide and to use the resulting solution of the sodium salt of glycylglycine for condensation reactions involving the amino group of the dipeptide.

The racemization of diketopiperazines, so pronounced in alkaline media, is notably less in acid. In some cases, it is possible, therefore, to achieve a satisfactory synthesis of optically active dipeptides by cautious treatment with hydrochloric acid. Thus, Greenstein (195) was able to make L-cysteinyl-L-cysteine from L-cysteine anhydride by hydrolysis of the latter with concentrated hydrochloric acid at room temperature. It must be added, however, that some diketopiperazines, such as histidine anhydride and tyrosine anhydride, are fairly stable when they are heated with strong acid, while alanine anhydride and leucine anhydride require prolonged treatment with hot concentrated acid to effect cleavage of the ring.

To the difficulties encountered in the partial cleavage of diketopiperazines containing like amino acid residues, must be added those met in the splitting of diketopiperazines derived from two different amino acids. In the latter case, two different dipeptides may result, as in the hydrolysis of glycyl-DL-leucine diketopiperazine, which yielded a mixture of glycyl-DL-leucine and DL-leucylglycine (166). The separation of such mixtures is usually difficult. If, however, the formation of only one of the two possible dipeptides is favored, this difficulty may be eluded. One of several instances in which a mixed dipeptide could be prepared in pure form by hydrolysis of a diketopiperazine was described by Bergmann and Tietzman (68), who obtained L-prolyl-L-phenylalanine from L-prolyl-L-phenylalanine diketopiperazine. The dipeptide had the same rotation as that of the product obtained by Fischer and Luniak (159), who coupled L-prolyl chloride with L-phenylalanine ethyl ester, and then saponified the dipeptide ester.

As the first available method for the synthesis of free peptides, the procedure discussed in this section has considerable historical interest. It is clear from the foregoing, however, that it has many limitations. Despite the occasional successes of this method, and its value as an adjunct to other synthetic procedures, it may be expected that, even for the synthesis of dipeptides, preference will be given to other methods.

2. *Condensation of Peptide Esters*

Curtius noted in 1883 (105) that glycine ethyl ester undergoes spontaneous transformation to yield glycine anhydride and a substance

which gives a positive biuret test. For this reason, the latter product was termed "biuret base." Later studies by Curtius (107) showed that, if moisture is excluded, the formation of the anhydride is suppressed and the "biuret base," which he formulated as (triglycyl)glycine ethyl ester, is the chief product. In a similar manner, Fischer (134) converted (diglycyl)glycine methyl ester to (pentaglycyl)glycine methyl ester by heating the former at 100°.

$$2\ NH_2CH_2CO{-}NHCH_2CO{-}NHCH_2COOCH_3 \rightarrow NH_2CH_2CO{-}(NHCH_2CO)_4{-}NHCH_2COOCH_3 + CH_3OH$$

More recently Pacsu and Wilson (254) and Frankel and Katchalski (176) have shown that, under suitable conditions, long-chain polycondensation products may be obtained from amino acid and peptide esters. The materials obtained represented mixtures of homologous peptide esters containing 20 to 100 amino acid units.

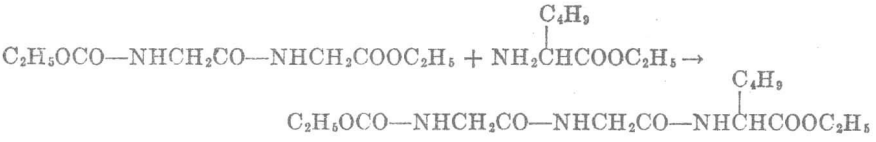
It may be noted at this point that, in general, amino acid esters and dipeptide esters readily yield diketopiperazines rather than polycondensation products. In fact, when a diketopiperazine composed of like amino acid residues is desired, it is most convenient to prepare it from the corresponding amino acid ester with ammonia in alcohol. In this manner, there have been synthesized a variety of diketopiperazines, such as histidine anhydride, lysine anhydride, and serine anhydride (172). Proline methyl ester, in particular, cyclizes with great ease (223). For the preparation of diketopiperazines derived from two different amino acids, treatment of the dipeptide ester with ammonia in alcohol usually leads to the desired product. Examples of this procedure are, among others, the synthesis of L-leucyl-L-alanine diketopiperazine from L-leucyl-L-alanine methyl ester (136) and glycyl-L-valine diketopiperazine from glycyl-L-valine methyl ester (164). A modification of this method, employing a lithium hydroxide solution saturated with carbon dioxide, was useful for the preparation of L-glutamylglycine diketopiperazine from α-L-glutamylglycine ethyl ester (78).

Fischer's first attempts to develop a general method of peptide synthesis led him to prepare the carbethoxy derivative of glycylglycine ethyl ester by treatment of the dipeptide ester with ethyl chlorocarbonate (149).

$$C_2H_5OCOCl + NH_2CH_2CO{-}NHCH_2COOC_2H_5 \rightarrow C_2H_5OCO{-}NHCH_2CO{-}NHCH_2COOC_2H_5$$

It was his hope, in this way, to introduce a substituent which would protect the reactive amino group from further attack in the course of condensation reactions and which could also be removed without hydrolysis

of the peptide bonds of the synthetic product. When the carbethoxy peptide noted above was heated for 36 hours with DL-leucine ethyl ester, ethyl alcohol was eliminated, and the resulting product was carbethoxyglycylglycylleucine ethyl ester (124).

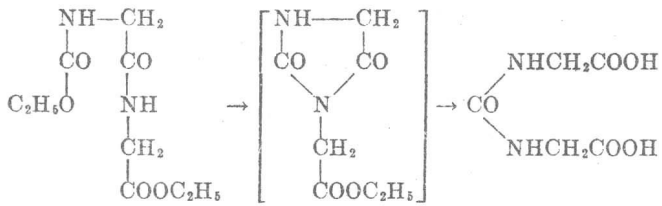


Fischer soon realized, however, that such condensation reactions are increasingly difficult as the peptide chain is lengthened and that even in the preparation of smaller peptides the yields are low. For this reason he turned his efforts to the development of other methods of peptide synthesis.

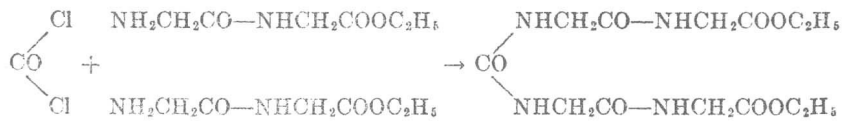
Of some interest in this connection is the behavior of carbethoxyglycylglycine ethyl ester on prolonged hydrolysis with alkali. Fischer (124) obtained a product which he formulated as glycylglycine carbamino acid:

$$HOOC{-}NHCH_2CO{-}NHCH_2COOH$$

It appeared, therefore, that the carbethoxy group could be removed from the acylated peptide ester without cleavage of the peptide bond. However, the complex nature of the reaction led Fischer to abandon further study of its mechanism. It remained for Wessely (296) to show that the product obtained by Fischer was actually carbonyl-bis-glycine, and the following sequence of reactions was suggested to explain its formation:



It may be added that, in the same paper in which he described the hydrolysis of the carbethoxyglycylglycine ethyl ester, Fischer also reported the synthesis of carbonyl-bis-glycylglycine by the treatment of the dipeptide ester with phosgene (in toluene), followed by saponification:



Such compounds may also be prepared by the reaction of the sodium salts of amino acids or peptides with phosgene.

3. *Synthesis of Peptide Derivatives by Means of Acylamino Acid Chlorides and Azides*

While working on the synthesis of hippuric acid, Curtius found, in 1881 (104), that one of the products of the interaction of benzoyl chloride and glycine silver was the substance benzoylglycylglycine. As pointed out by Fischer (129) in the course of a polemic with Curtius, this reaction is rather complex in nature, and, although it may be considered to represent the first recorded synthesis of a well-defined peptide derivative, the method is not suitable for general application.

Curtius' studies, during the period 1890–1900, on the reactions of hydrazides and azides, led him to use these in the synthesis of peptide derivatives. After the report by Fischer and Fourneau (149) that dipeptides could be made by the partial hydrolysis of diketopiperazines, Curtius (106) described the use of azides of benzoylamino acids or peptides according to the following reaction, illustrated for the case of the synthesis of benzoylglycylglycylglycine:

$$C_6H_5CO{-}NHCH_2CON_3 + NH_2CH_2CO{-}NHCH_2COOH \rightarrow$$
$$C_6H_5CO{-}NHCH_2CO{-}NHCH_2CO{-}NHCH_2COOH + HN_3$$

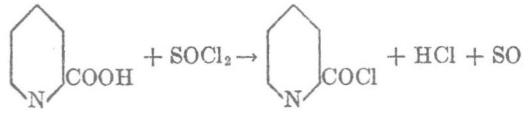
In order to obtain the azide, the corresponding ester was treated with hydrazine hydrate, thus yielding a hydrazide, which was in turn converted to the azide by means of nitrous acid:

$$C_6H_5CO{-}NH\underset{\displaystyle R}{\underset{|}{C}}HCOOC_2H_5 \rightarrow C_6H_5CO{-}NH\underset{\displaystyle R}{\underset{|}{C}}HCONHNH_2 \rightarrow$$
$$C_6H_5CO{-}NH\underset{\displaystyle R}{\underset{|}{C}}HCON_3$$

These reactions proceeded smoothly and, frequently, with excellent yield. Curtius and Levy (111) were able to make benzoyl(tetraglycyl)-glycine ethyl ester by the condensation of benzoyl(diglycyl)glycine azide with glycylglycine ethyl ester, and with his collaborators, Curtius extended this method to the synthesis of benzoylated peptides containing alanine (110), aspartic acid (108), and aminobutyric acid (109).

As was noted in the previous section of this review, Fischer's first attempts to develop a general method of peptide synthesis led him to prepare the carbethoxy derivative of glycylglycine ethyl ester (144). When he concluded that condensation of such esters with esters of amino acids was not feasible as a general procedure, he decided to convert the carbethoxyamino acids to the corresponding acid chlorides by means of

thionyl chloride (127), a reaction which had been found by Meyer (246) to be suitable for the preparation of the acid chloride of pyridine carboxylic acid:



By warming carbethoxyglycine or carbethoxyglycylglycine with thionyl chloride at 35–40°, noncrystalline products were obtained which were used directly for coupling in ethereal or in chloroform solution with amino acid or peptide esters. From the resulting carbethoxy peptide esters, the corresponding acids could be prepared by saponification. In several cases, these acids could be converted to acid chlorides with thionyl chloride, and the peptide chain lengthened by coupling with amino acid or peptide esters.

In principle, the above methods of Curtius and Fischer provide the basis for the further development of the techniques of peptide synthesis. All the subsequent procedures for lengthening the peptide chain have involved the conversion of the carboxyl group of an amino acid into forms which permit reaction with the amino group of another amino acid. Of the various derivatives of carboxylic acids which have proved useful for this purpose, the azides and chlorides have been of the most general value. Indeed, in many cases, the older azide method of Curtius is preferable to the chloride method, particularly in coupling reactions involving acyl peptides (*cf.* page 26).

Although Fischer showed in 1905 (132) that it was possible to convert free amino acids to acid chlorides, it was realized that, in order to permit smooth coupling reactions, the amino group had to be blocked by acylation, or otherwise modified to avoid complicated side reactions in the course of the conversion of the carboxyl group to an acid chloride. In addition to the carbethoxy and benzoyl groups mentioned above, a variety of acyl substituents were introduced into peptide chemistry. In addition to others to be discussed later may be mentioned the naphthalenesulfonyl (144), phenylureido (170), benzenesulfonyl (145), and methanesulfonyl (207) groups. However, as long as it was necessary to remove an acyl substituent by hydrolysis, it could not be used in the synthesis of free peptides, since attempts to eliminate the acyl group in this manner invariably led to either partial or complete cleavage of the linkages between the amino acids. The essential problem of peptide synthesis thus became the development of methods which would obviate the necessity for the hydrolytic removal of an acyl substituent at the end of a series of coupling reactions.