Reactivity and Structure
Concepts in Organic Chemistry 8

Christian Birr

# Aspects of the Merrifield Peptide Synthesis

Springer-Verlag Berlin Heidelberg New York

Christian Birr

# Aspects of the Merrifield Peptide Synthesis

With 62 Figures and 6 Tables



Springer-Verlag
Berlin Heidelberg New York 1978

Priv.-Doz. Dr. Christian Birr

Max-Planck-Institut für medizinische Forschung
Abteilung Naturstoff-Chemie
Jahnstr. 29, D-6900 Heidelberg 1

ISBN-3-540-08872-5 Springer-Verlag Berlin Heidelberg New York
ISBN-0-387-08872-5 Springer-Verlag New York Heidelberg Berlin

Library of Congress Cataloging in Publication Data. Birr, Christian, 1937– Aspects of the Merrifield peptides synthesis. (Reactivity and structure ; v. 8) Bibliography: p. Includes index. 1. Peptide synthesis. I. Title. II. Series. QD431.B57 547'.756 78-16659


Printed in Germany

Typesetting: Elsner & Behrens, Oftersheim
Printing and binding: Konrad Triltsch, Würzburg
2152/3140-5 4 3 2 1 0

# Preface

This book was written in the context of the daily confrontation with problems in the utilization of polymeric supports for the synthesis of peptides. Therefore, views and experiences which usually are not mentioned in scientific journals are collected in these pages. The author has deliberately discussed in detail the possible influence of the polymer phase on the varying reaction conditions in the Merrifield synthesis; this aspect is neglected in most publications dealing with peptide synthesis. However, in view of the growing body of information on the chemistry of polymer-supported peptide syntheses, the international readership should regard the author's arguments as open to discussion.

I am very much indebted to all of my colleagues with whom I have had the opportunity to cooperate in studying the potential of the Merrifield synthesis. Above all I like to express my gratitude to my teacher, Professor Dr. Theodor Wieland, Heidelberg, for his boundless encouragement and support in my efforts in the field of peptide synthesis, particularly in its polymer phase bound version. Last but not at least I wish to thank Miss Hildegard Leyden. With infinite patience and great accuracy she typed the manuscript in addition to her daily duties.

The work which influenced this book was performed under the auspices of the Max-Planck-Gesellschaft, of which the financial support is gratefully acknowledged.

Heidelberg, June 1978 Christian Birr

# Contents


*1 Introduction* . . . . . 1

*2 The Principle* . . . . . 3

2.1 Aspects of the Strategy . . . . . 6
2.2 The Statistical Point of View on Solid Phase Reactions . . . . 9

*3 Chemical Details of the Method* . . . . . 16

3.1 The Polymer Support, Its Choice, Properties, and Preparation . . 16
3.1.1 Introduction and Conversion of Functional Sites on Polymer . . 21
3.1.2 Binding of Amino Acids and Peptides . . . . . 31
3.1.3 Analysis of Support-Bound Amino Acid Derivatives . . . . . 32
3.2 The Deprotection and Deprotonation . . . . . 36
3.2.1 Aspects on Choice of Protecting Groups . . . . . 37
3.2.2 Deprotecting Reagents . . . . . 39
3.2.3 Deprotonation Problems . . . . . 40
3.2.4 Analysis of Deprotection and Deprotonation . . . . . 42
3.3 The Peptide Bond Formation . . . . . 47
3.3.1 Dicyclohexylcarbodiimide Activation . . . . . 47
3.3.2 Symmetric and Mixed Amino Acid Anhydrides . . . . . 50
3.3.3 Active Esters . . . . . 55
3.3.4 Redox Condensation . . . . . 55
3.3.5 Fragment Condensation on Polymer Phase . . . . . 57
3.3.6 Analysis of Peptide Bond Formation . . . . . 58
3.4 Suppression of Deletion Sequences . . . . . 60
3.5 Cleavage of Peptides from Polymer Support . . . . . 62
3.5.1 Acidic Cleavage Conditions . . . . . 62
3.5.1.1 Trifluoroacetic Acid/Hydrogen Bromide . . . . . 62
3.5.1.2 Liquid Hydrogen Fluoride . . . . . 63
3.5.2 Basic Cleavage Conditions . . . . . 65
3.5.2.1 Base Catalyzed Transesterification . . . . . 65
3.5.2.2 Hydrolysis . . . . . 67

3.5.2.3 Hydrazinolysis . . . . . 67
3.5.2.4 Ammonolysis . . . . . 68
3.5.3 Modified Cleavages . . . . . 69

*4 Automatization of the Merrifield Peptide Synthesis* . . . . . 72

4.1 Components of an Automatic System . . . . . 72
4.1.1 Programming . . . . . 73
4.1.2 Transport and Metering of Liquids . . . . . 73
4.1.3 The Reactor . . . . . 74
4.1.4 Control of Functions and of the Synthesis . . . . . 76
4.2 Critical Statement on Peptide Synthesizers . . . . . 79

*5 Critical View on the Applicability of the Merrifield Synthesis* . . 81

5.1 Aspects on Product Isolation and Purification . . . . . 81
5.1.1 Problems on Identification of Impurities . . . . . 83
5.1.2 The Demand for Analytical Investigations . . . . . 83
5.1.3 Synthetic Peptides for Pharmaceuticals . . . . . 84
5.2 Outlook on Disciplines Related to Peptide Chemistry . . . . . 85

*6 Conclusion* . . . . . 87

*7 References* . . . . . 88

*Author Index* . . . . . 94

*Subject Index* . . . . . 97

# 1 Introduction

In most areas of science, knowledge seems to increase not in a continuous process but in a way of certain quanta jumps, from a level of insight to the next better. Surely this is the case in the development of peptide chemistry.

From my point of view, the evolution of the methodology of peptide synthesis may be characterized by four distinct periods. Although syntheses of amino acids and their reactivity were well known before, the finding of E. Fischer [1] in 1901–1903 of how to transform amino acids to render a peptide synthesis possible marked the breakthrough to protein research.

But not until 1932, when M. Bergmann and L. Zervas [2] published the first protecting group for temporary blocking of the N-terminal amino function during peptide bond formation, was a second level in the knowledge of methods in peptide synthesis reached. The use of the carbobenzoxy group considerably simplified the elongation of peptide chains. Consequently the very versatile principle of the urethane type of protection of amino functions was often varied and adapted to very different reaction conditions during the next two decades.

Doubtless, a new field in the development of peptide synthesis was opened when Th. Wieland (1950–51) [3] published the methods of activation of amino acid carboxylic functions both by formation of mixed anhydrides with ethylchlorocarbonate and by active esters with thiophenol [4]. At the same time also, R. Boissonnas [5] and J. R. Vaughan [6] independently described the mixed anhydride method. A most fruitful period of peptide synthesis began, which culminated (for example) in the synthesis of peptide hormones like oxytocin by duVigneaud (1953) [7], ACTH by R. Schwyzer (1963) [8], and insulin by the group of H. Zahn in 1963 [9]. Besides these examples, hundreds of important syntheses of natural occurring peptides were published in that decade, but it would go beyond the scope of this introduction to note even some of them. Although further protecting groups were introduced, and the methods of peptide coupling became more and more subtilized, recognition increased that one has to search for a new principle of peptide synthesis, to reach the goal of building up peptide chains of a hundred or more amino acid residues, by copying nature's principles, to synthesize proteins.

In the beginning sixties, when R. B. Merrifield for the first time described his new idea of solid phase peptide synthesis (1962) [10], the cellular mechanism of ribosomal protein biosynthesis was already known, and was no longer a subject of hypothetic discussions. The ribosomal area of the biological stepwise condensation of amino acids might be described by the term "gel phase", where reactions proceed between solution and solid phase in a

region to be paraphrased by hydrophobic-hydrophilic interplay. It seems plausible to me that Merrifield's idea of solid phase peptide synthesis was influenced by the knowledge of this biological principle of surface reactions on a gelatinous phase. Later on we shall see that from the chemical point of view, the Merrifield peptide synthesis differs drastically from the very rapid reactions of the protein biosynthesis.

Since 1962, Merrifield's new idea induced unanticipated activity in peptide chemistry around the world. The synthesis of artificial enzymes and hormones seemed to become possible, and hundreds of organic chemists and biochemists, as well as enzymologists and physicians, devoted themselves to the new method. Unfortunately the potential of the solid phase peptide synthesis was, from the very beginning, erroneously estimated by the international scientific public, both in a positive and negative sense. By emotional arguments more than by objective ones, the peptide chemists until recently were split into two parties, of which one trusted and the other condemned the solid phase peptide synthesis.

The enthusiastic belief in the new method promising rapid success in protein synthesis led to some publications which were in want of the necessary critical care. On the other hand, the methodology of the solid phase synthesis was elaborated and improved by a great number of peptide research groups. Hundreds of excellent investigations, not only concerning chemical details of the new technique, but especially peptide syntheses on solid phase, have been published till today. This progress in the solid phase method of Merrifield is reviewed in several articles and most detailed and comprehensive books [11–35]. In the book submitted here, I therefore shall confine myself to evaluating the chemical aspects of the Merrifield method from the standpoint of present practical use. Several developments of methodical details will be balanced critically against each other, without the intention of reviewing the literature completely or discussing the fundamentals of peptide synthesis from the very beginning.

# 2 The Principle

Contrary to the principles of the ribosomal protein biosynthesis, where the carboxylic functions of amino acids are reactively bound to a polymer – the transfer ribonucleic acid (t-RNA) – which itself is orientated specifically on the ribosomal messenger ribonucleic acid (m-RNA) (Fig. 1; for details see [36]), the basic idea of the Merrifield synthesis depends upon a nonreactive covalent fixation of the C-terminal amino acid of the target peptide on a solid support (Fig. 2). On this insoluble but swollen polymer, the pep-

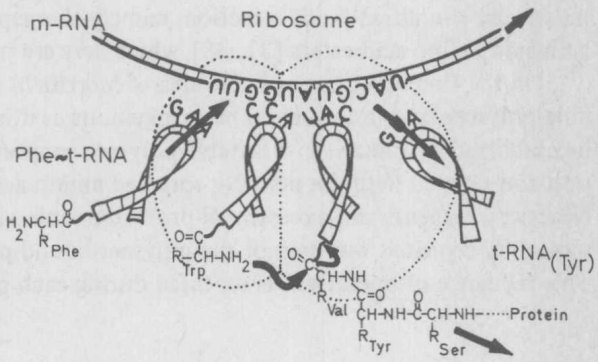


Fig. 1. The ribosomal area of the biological stepwise condensation of amino acids in the protein biosynthesis

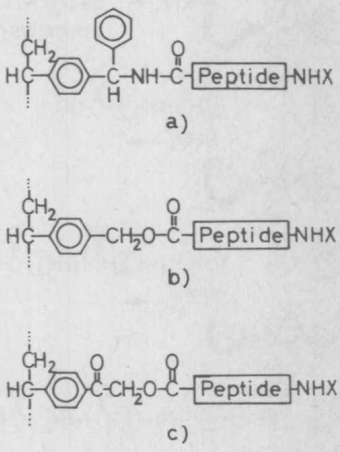


Fig. 2. Three examples of the solid anchoring of amino acids and peptides in polymer phase, nonreactive during chain elongation in the Merrifield peptide synthesis [type a) benzhydryl amine, b) benzyl ester and c) phenacetyl ester]

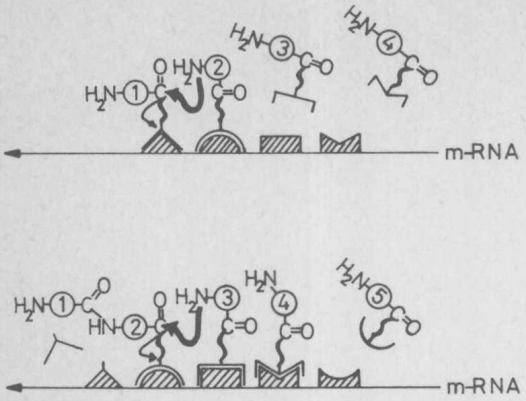


Fig. 3. The aminolytic chain elongation in the protein biosynthesis

tide to be synthesized is assembled in a stepwise manner by N-terminal elongation of the chain with N-protected and C-activated amino acids, which are the dissolved carboxyl components of the solid phase peptide synthesis.

In contradiction to this, the protein biosynthesis proceeds via aminolysis of activated amino acids reactively bound to t-RNA, where the growing protein is liberated from the polymer RNA support (Fig. 3). This type of biosynthesis was imitated in the way of a drastically simplified model reaction, namely the peptide syntheses with solid phase bound activated amino acid esters [37, 38], which here are not the matter for discussion (Fig. 4).

On the first view, the brilliant idea of Merrifield to synthesize peptides onto an insoluble polymer support seems to be as ingenious as it is simple (Fig. 5). The first amino acid, insolubilized by linkage to a suitable polymer support, is deprotected at its amino function and reacted with the next N-protected amino acid to form the first peptide bond. Necessary reagents and excessive N-protected amino acid derivatives are thoroughly removed by repeated washings of the polymer-bound peptide with appropriate solvents. This sequence of operations is repeated during each peptide elongation stage. At the end

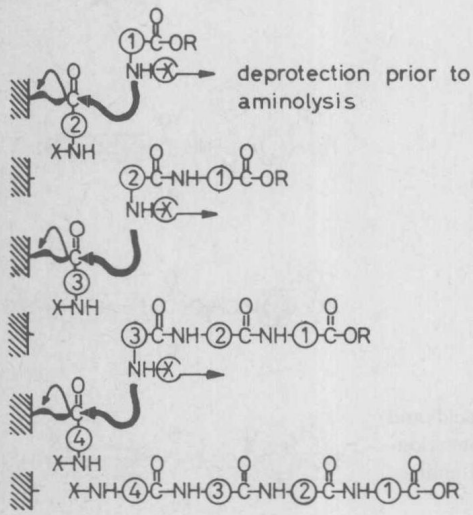


Fig. 4. The stepwise peptide synthesis in gel phase with polymer-supported active esters of amino acids, a model of the protein biosynthesis

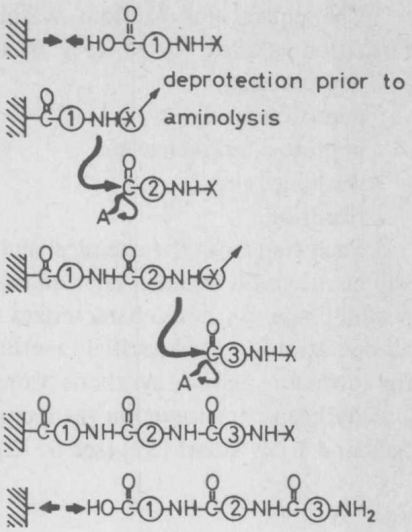


Fig. 5. The principal scheme of the Merrifield solid phase peptide synthesis

of the synthesis on solid phase, the link of the peptide to its polymer support is cleaved, the dissolved peptide is separated from the insoluble resin and has to be purified. It is obvious that this procedure extremely simplifies the manipulations of classical peptide chemistry, because it utilizes the same classical reactions for deprotection, deprotonation, and activation of amino acids, as well as for cleavage of the peptide from the support. All the time-consuming and wasteful operations in conventional peptide synthesis, to derivatize intermediates and to work-up and purify these products by extractions and crystallization, as well as to analyze and characterize each intermediate peptide combined with often repeated transfers of the synthetic materials through several glass-vessels, are eliminated in the Merrifield synthesis, or are substituted for simple intense washings and filtration of the polymer-bound peptide in a single reaction vessel.

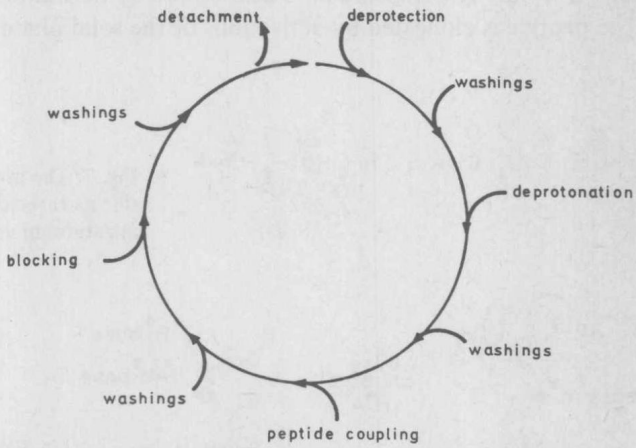


Fig. 6. The cyclic sequence of operations in the Merrifield peptide synthesis

The sequence of reactions, washings, and filtrations necessary for one peptide bond formation is called "cycle" (Fig. 6) and consists of several "steps", which are, for example,

deprotection,
deprotonation,
peptide bond formation,
washings, and
filtration.

Each step from the chemical and practical view has its own problems and therefore will be discussed in detail later on in separate sections. The succession of "cycles" during a solid phase synthesis characterizes the various "stages" of the peptide elongation. Since all operations in the Merrifield method are often repeated, the procedure is predestined for automatic peptide synthesis. Consequently the first apparatus for this purpose had already been constructed in the incipience of the methodical development by R. B. Merrifield and J. M. Stuart [39] (see p. 72).

## 2.1 Aspects of the Strategy

Merrifield's original idea was based on the general scheme of stepwise condensation of N-protected amino acids to the first one, which is linked with its carboxyl function by an ester bond to the insoluble polymer support. This way of solid phase peptide synthesis resulted from the well-known risk of racemization during activation of peptidic carboxyl components, which is minimized in activated amino acid derivatives, N-acylated by urethane-type protecting groups [40] (Fig. 7). Depending on the chosen method, the C-terminal activation of N-protected peptides tends to racemize a certain amount of the material because of the possible formation of an oxazolinone intermediate [41] (Fig. 8).

For that reason, the applicability of the other strategy introduced by R. L. Letsinger and M. J. Kornet in 1963 [42] is in contrast to the Merrifield procedure limited. In the second method the stepwise peptide synthesis starts from the opposite direction, beginning from the N-terminal amino acid bound by its amino function to an insoluble support. The peptide is elongated by activations of the solid phase carboxyl component on each

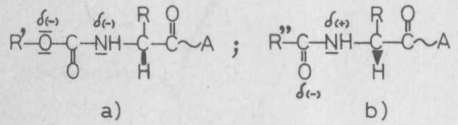


Fig. 7. The urethane type of N-protecting groups (a) shields carboxyl activated amino acids against racemization, in contrast to common N-acyl residues (b)

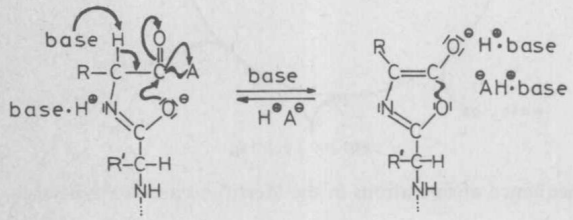


Fig. 8. The oxazolin-5-one mechanism of the C-terminal peptide racemization

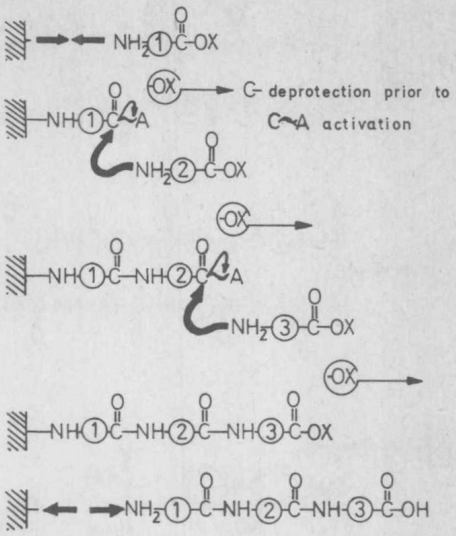


Fig. 9. The schematic peptide synthesis on polymer by C-terminal chain elongation according to Letsinger and Kornet

stage of the stepwise synthesis (Fig. 9). To what extent the utilization of this method might be improved by racemization suppressing additives [43] like 1-hydroxybenzotriazole and others, which recently have been introduced by R. Geiger and W. König [44], is not the matter for discussion here. Rather, I want to relate to another approach of the stepwise peptide synthesis on solid phase, which combines the potential of Merrifield's and Letsinger's strategies. Since almost fifty percent of the naturally occuring amino acids are trifunctional ones like glutamic acid, asparagine, tyrosine, serine, cysteine, histidine, lysine and others, a starting amino acid for the purpose of solid phase peptide synthesis can be linked covalently to the polymer support by its third function [45–48]. This aspect not only includes the possibility of peptide synthesis in two directions (Fig. 10), but also is advantageous in peptide cyclization on solid phase (Fig. 11). To consider the polymer support a special reagent to protect, for example, the sulfhydryl function of cysteine or cysteine containing peptides [49], this view opens peculiar chances to synthesize peptides for subsequent disulfide cyclization. J. D. Glass, R. Walter, and J. L. Schwartz [50] demonstrated a further variation of binding of trifunctional amino acids like histidine by using a dinitrofluorophenyl-modified polymer support (Fig. 12) for bidirectional solid phase synthesis. The method promises mild conditions of peptide cleavage from the polymer and therefore will be discussed in a separate chapter. From the strategic point of view, it is hard to evaluate the advantage of this bridging method because of the early stage of

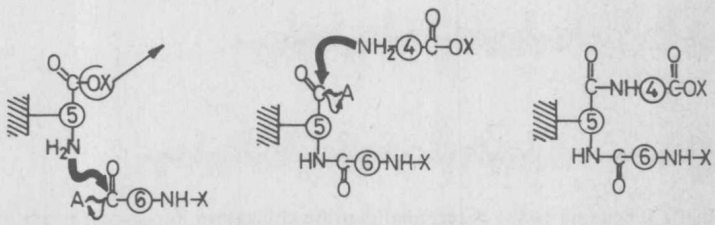

Fig. 10. The bidirectional peptide synthesis in polymer phase

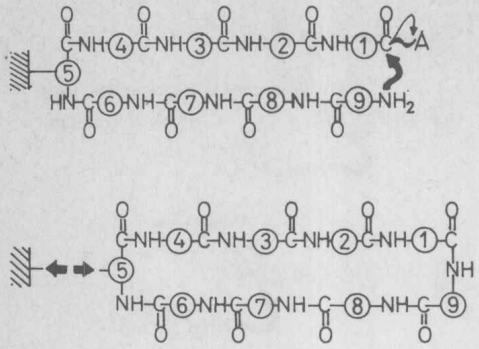

Fig. 11. The peptide cyclization in polymer phase after bidirectional synthesis of the linear precursor

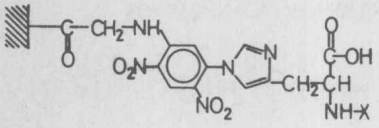

Fig. 12. The 2,4-dinitrophenylene bridge links histidine to the modified support for bidirectional synthesis

experiences. Obviously it is still problematic to transform in high yield the carboxyl functions of peptides which are bound to a polymer support into activated esters or other intermediates of sufficient reactivity, without aggravating the racemization tendency.

Although this argument is concerned even more with the activation of dissolved peptides, the original strategy of Merrifield involving stepwise amino acid condensation is being modified occasionally in terms of stepwise peptide fragment condensation on solid phase, which differs from the conventional technique in dissolved systems, because of the exclusive N-terminal peptide elongation on polymer support (Fig. 13). As discussed in the

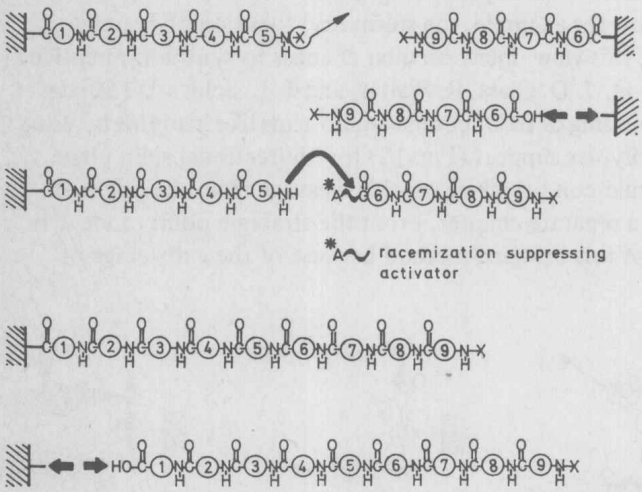


Fig. 13. Scheme of the N-terminal peptide elongation in polymer phase by stepwise fragment condensation

following section, this modification has its extraordinary importance in questions of possible syntheses of pure proteins – peptides consisting of more than 100 amino acid residues – on solid phase, because of the enlarged molecular weight of the peptidic building blocks in stepwise fragment condensation, compared to amino acid units in the original strategy of Merrifield.

## 2.2 The Statistical Point of View on Solid Phase Reactions

Behind the elegance and the ingenious simplicity of Merrifield's idea of peptide synthesis a serious problem is hiding, which is inherent in the principle itself, namely the demand for completion of all chemical reactions on solid phase [51].

In the early stage of application of the Merrifield method to polypeptide syntheses, either this supreme requirement was not perceived or its consequences were disregarded. This sin of omission, on the one hand, fundamentally discredited the method temporarily with groups of international peptide chemists, as indicated in the Introduction. But on the other hand, from its very first enunciation, this demand signified a strong provocation to numerous organic chemists engaged in peptide syntheses to overcome several shortcomings of the Merrifield method. These shortcomings are derived from the difficulty of realizing complete reactions on a polymer support and therefore will be the matter for discussion in diverse sections of this book.

The interpretation of the problem depends upon some statistical aspects of chemical transformations on insoluble polymers. Contrary to the homogeneous situation in real dis-

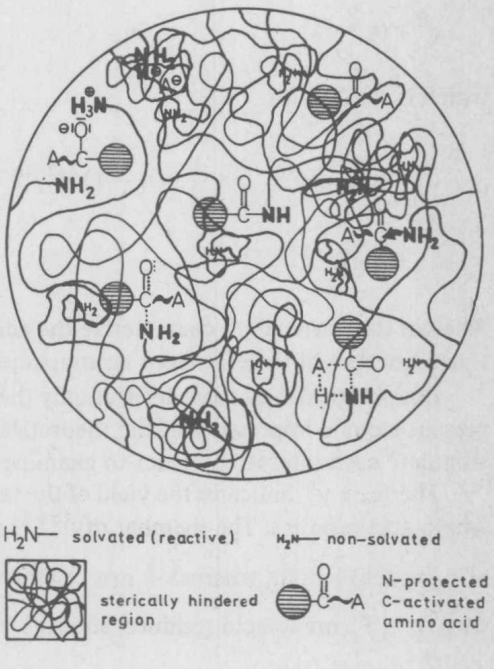


Fig. 14. The statistical distribution of functional transformations during peptide synthesis in heterogeneous gel phase

solved phases, the bimolecular reaction conditions in a Merrifield support vary by different microenvironments of the reactive functions in the inhomogeneous polymer matrix (for further discussion see Sect. 3.1, page 16). This fact results not only in slower reaction rates by phase separation in general, but in a statistical distribution of the degree of all transformations under distinct conditions for each step of a solid phase peptide synthesis (Fig. 14). For this reason it is very difficult to reach the goal of complete transformations on an insoluble support, and even more difficult to analyze the reaction since the same uncertainty of reaction conditions seriously concerns the direct analytical control of 100% conversions on solid phase by additional chemical reactions.

During the ninth European Peptide Symposium 1968, in Orsay, France, the reasonable view was published [52] that even very small deviations – less than 1% – of reaction rates from completion (a very optimistic assumption as we already know from reactions in real dissolved systems) have to generate a statistical distribution of by-products accompanying the target peptide. To get an impression of these inhomogeneities which depend on factors such as incomplete deprotection, deprotonations, and peptide bond formations on solid phase, several groups of peptide chemists tried to calculate these effects roughly as a function of constant yield of peptide elongations in each cycle of a Merrifield synthesis. We compared our results [53] with those of E. Bayer's group – which were published later [54] – and found equivalent expressions on the basis of the binominial proposition.

The yield y(%) of each cycle c of a synthesis of n stages may deviate from complete conversions (100%) in a constant failure of f(%),

$$y(\%) = 100 - f(\%), \quad c = y + f = 1$$

The product composition after n cycles c of the synthesis is then described by the binominial expression,

$$c^n = (y + f)^n$$

which is resolved to

$$c^n = y^n + nfy^{(n-1)} + \frac{1}{2}\,n(n-1)^2 f^2 y^{(n-2)} + \frac{1}{6}\,n(n-1)(n-2)^2 f^3 y^{(n-3)} + \ldots$$

$$+ nyf^{(n-1)} + f^n$$

wherein the terms of $c^n$ characterize the additive statistical distribution of defective peptide sequences in molefractions accompanied by the yield of the target peptide.

For practical reasons we inspect only the first three members of the equation, since they present a quick impression of the theoretical course of a stepwise synthesis, and should stimulate each interested reader to examine synthetic problems of his own choice.

The term $y^n$ indicates the yield of the target peptide, which has a chain length of (n + 1) amino acid residues. The member $nfy^{(n-1)}$ represents the yield of all defective sequences of n residues length, whereas $\frac{1}{2}\,n(n-1)^2 f^2 y^{(n-2)}$ equals the yield of all false sequences with (n – 1) amino acid residues, and so forth.

Table 1. Theoretical end product composition of an incremental peptide synthesis of 100 stages with a constant failure of 1% per peptide elongation

| Chain lengths | Sequence Variations V | Yield (%) |
|---|---|---|
| 101[a] | 0[a] | 36 |
| 100[b] | 100 | 36[c] |
| 99[b] | 5,000[c] | 18[c] |
| 98[b] | 160,000[c] | 6[c] |
| 97[b] | 4,000,000[c] | 1.5[c] |
| 96[b] | 80,000,000[c] | 0.3[c] |

[a] Target peptide. [b] By-products. [c] Rounded off.

Expressed in real numbers a Merrifield synthesis involving, for example, n = 100 stages with a constant failure f = 1% in each cycle would yield the target peptide containing 101 residues in 36% (since $y^n = 0.99^{100} = 0.36$), mixed up with by-product sequences of 100 residues length in a molefraction of likewise about 36% (namely $nfy^{(n-1)} = 100 \cdot 0.01 \cdot 0.99^{99} \sim 0.36$), and with by-products combining all possible sequences with a chain length of 99 amino acid residues in 18%.

Since the quantity of all possible sequence variations V in each additive member of the binominial equation, which symbolizes the molefraction of the by-product group X (all sequences X residues shorter than the target peptide. X = exp. of f = 0, 1, 2 ...) in a synthesis of n stages, is indicated by $V = n!/X!\,(n - X)!$, the theoretical composition of the above-mentioned stepwise synthesis is roughly determined by the figures of Table 1.

The fundamentality of the demand for completion of all reactions is saliently elucidated within the background of these simple estimations on the theoretical result of a solid phase synthesis. It is obviously impossible to separate about 84,165 million extremely similar peptidic by-products – whose molecular weights do not differ by more than five amino acid residues of weight – from the main sequence by any presently available method of purification.

From more accurate calculations of J. M. A. Baas et al. [55] of theoretical end product compositions of natural occurring peptides (Table 2) – which in the meantime nearly all were taken as goals for solid phase syntheses – one has to appreciate that Merrifield syntheses of sequences larger than 15–25 amino acid residues, without guaranteed yield of at least 99.5% in each stage of the procedure must result in peptide end products of indefinite chemical composition and of uncertain scientific value, even though one confines the judgement on biological activity of the material in question: The target peptide, e.g., a synthetic enzyme, a hormone or even parts of these, may contain millions of closely related defective sequences of similar biological activities. Without knowledge of any distinct sequence, the additive activities of these by-products create the impression of the success of the synthesis. The target peptide, however, may scarcely range in the material.

In spite of these murky aspects, which concern the adventures in the field of syntheses of polypeptides and proteins on solid phase by incremental chain elongation, the method of Merrifield is invaluable in the quick and convenient preparation of smaller peptides