1973

Bruce Merrifield, 1971

The Rockefeller University

THE CHEMICAL SYNTHESIS OF AN ENZYME*

BRUCE MERRIFIELD

The Rockefeller University, New York, New York

I. Introduction

The complexity of the proteins presents an unparalleled challenge to the synthetic chemist, and it is only recently that he has been in a position to respond in a meaningful way. Until now essentially all that has been learned about this class of compounds has come from studies on the naturally occurring proteins themselves. The accumulated body of knowledge is enormous, and very much is known about the composition and structure of the proteins and how they function. There are certain questions, however, which cannot be answered easily in this way and for these questions the synthetic approach should provide a valuable alternative. If we could learn how to synthesize proteins in the laboratory it should become possible to make specific, well defined changes in their structures which could give answers that are hard to derive from studies on the native molecules. In certain special cases, such as the unusually reactive residues at the active sites of enzymes, it has been possible to modify individual amino acid residues selectively, but suppose it were to become important to modify or replace one particular valine residue in the protein; methods simply are not available to carry out such changes in the presence of many other similar groups within the same very sensitive macromolecule.

The three-dimensional structure of a small protein, ribonuclease, is shown in Fig. 1 to illustrate the kind of molecule we are talking about and to illustrate the magnitude of the synthetic task. This crystalline molecule (Kunitz, 1940) contains 124 amino acids in a single polypeptide chain, cross-linked by four disulfide bonds, with 1852 atoms precisely arranged in space in a well defined three-dimensional structure. A chemical synthesis of ribonuclease A means,

* Lecture delivered November 18, 1971.

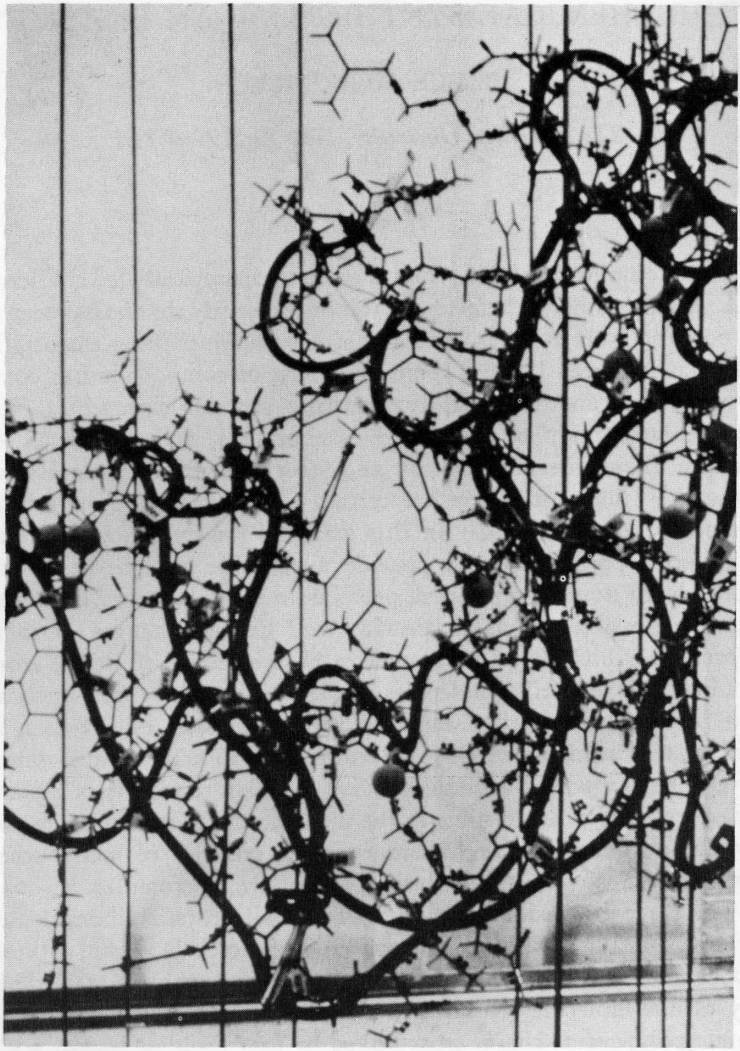


FIG. 1. The three-dimensional structure of bovine pancreatic ribonuclease. Taken from the model of RNase S by Wyckoff *et al.* (1967) which was deduced from X-ray diffraction data at 3.5 Å resolution and chemical sequence data. The heavy tubing outlines the general contours of the polypeptide backbone, and the fine wire structures show the positions of the amino acid side chains. Reproduced by permission of the American Society of Biological Chemists, Inc.

therefore, that each of these atoms must be assembled in the laboratory into exactly this same sequence and structure. An understanding of the problem can be greatly simplified, first, by looking at a two-dimensional picture in which the amino acid residues, rather than the individual atoms, are indicated as the basic units of the structure (Fig. 2) and, second, by recalling the important hypothesis of Anfinsen and his associates (White, 1961; Epstein *et al.*, 1962) that the tertiary structure of this protein is determined by its primary structure. Thus, there was good reason to believe that if the amino acids could be assembled into the correct sequence the resulting molecule would direct its own folding and cross-linking to give the native protein, and a total synthesis of an enzyme would have been achieved.

In principle, a protein can be synthesized by either of two general approaches; by the classical method or by the solid phase method. The classical method (see Schröder and Lübke, 1965) involves the preparation of small peptides (fragments), the combination of these small fragments into larger peptides, and eventually the assembly of the larger fragments into the final protein molecule. All these reactions are carried out in homogeneous solution, and the essential feature of this technique is the opportunity to isolate and purify the product after each reaction. This helps to ensure high purity and careful identification of intermediates.

The most dramatic application of the classical technique to the synthesis of a protein occurred in the period between 1963 and 1965, with the nearly simultaneous synthesis of insulin by groups in Germany, the United States, and China (Meienhofer *et al.*, 1963; Katsoyannis *et al.*, 1964; Kung *et al.*, 1965). At the present time, syntheses of proinsulin, ribonuclease T_1, staphylococcal nuclease, myoglobin, growth hormone and, no doubt, several other proteins are under way. As far as I know none of these has been completed, but two or three can be expected soon. The problems, of course, are enormous. Large numbers of synthetic reactions and purification steps are involved, and much time, expense, and manpower is required.

A variation of the strictly classical method, in which very rapid Leuchs anhydride couplings were employed, was applied successfully to the synthesis of S-protein (residues 21–124 of ribonuclease)

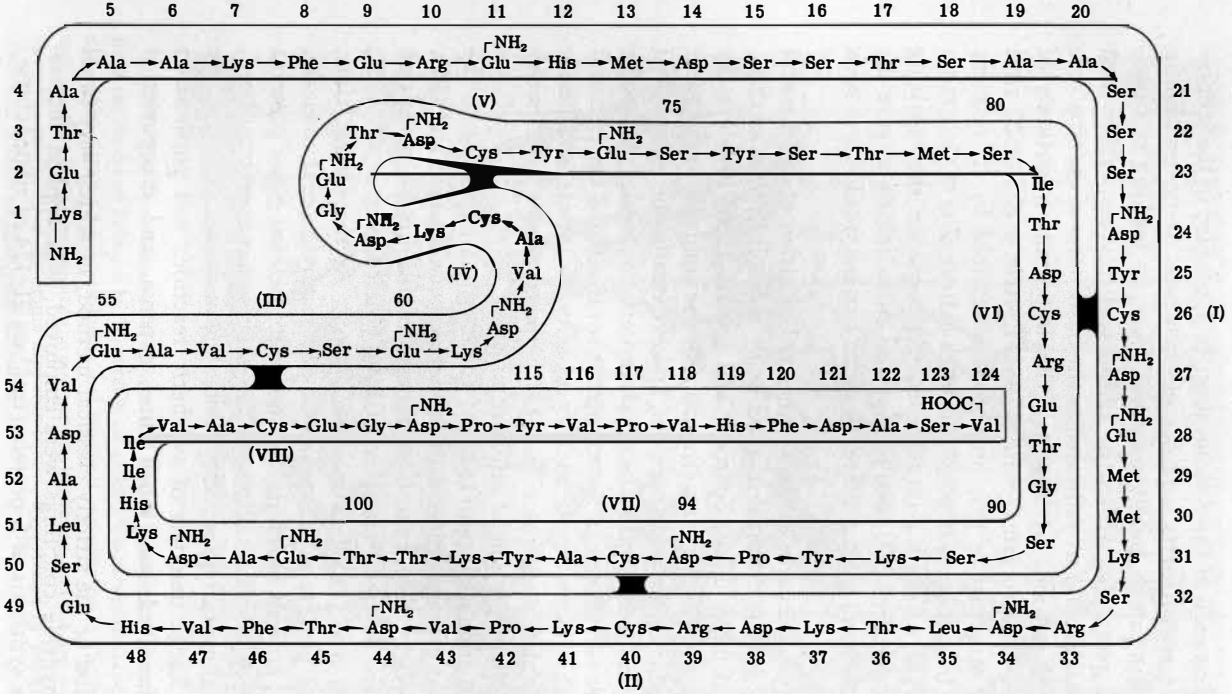


FIG. 2. The two-dimensional structure of bovine pancreatic ribonuclease A, showing the linear, 124 amino acid residue polypeptide chain cross-linked by the four disulfide bonds. The amino acids are indicated by their three-letter abbreviations. From Smyth *et al.* (1963), by permission of The American Society of Biological Chemists, Inc.

by Hirschmann *et al.* (1969). This accelerated technique eliminated the need for isolation of every intermediate, but permitted the isolation and purification of protected fragments. After noncovalent binding with S-peptide (residues 1–20) the resulting ribonuclease S possessed enzymatic activity.

The second general approach to the chemical synthesis of proteins is called solid phase peptide synthesis (Merrifield, 1963, 1969). The method was devised in our laboratory in an effort to overcome certain of the difficulties encountered with the classical methods. It is much faster and simpler to carry out and introduces sizable savings in manpower and materials. It has made possible our synthesis of ribonuclease A (Gutte and Merrifield, 1969, 1971). I would like to describe this technique in some detail and to discuss some of the difficulties and problems associated with it as well as some of the achievements.

II. Solid Phase Peptide Synthesis

The basic idea of the solid phase method is outlined in Fig. 3. It depends on the use of an insoluble solid support. The plan was to begin with an insoluble particle, indicated by the large circles, to functionalize it with a group, x, and to attach the first amino acid of the proposed peptide chain with a stable covalent bond by reaction at x. In practice we attach at the carboxyl of the amino acid (with the amino group protected), although in principle the attachment can also be at the amino group. The amino group is then deprotected, and the second amino acid is added to the first by a suitable peptide-forming reaction. In a similar way the subsequent amino acids are combined in a stepwise manner until the entire sequence has been assembled. Finally, the bond holding the peptide chain to the solid support is selectively cleaved, together with most side chain protecting groups, and the peptide is liberated into solution.

The advantage that I envisioned for this approach was the opportunity to purify intermediates after each reaction by simple, rapid washing procedures. Since the growing peptide chain would be completely insoluble as a consequence of its attachment to the large insoluble support, it could be filtered and washed with large volumes of solvents to remove reagents and by-products with no danger of losing the desired molecule. A related advantage, which was the real driving force behind the project, was the possibility that the

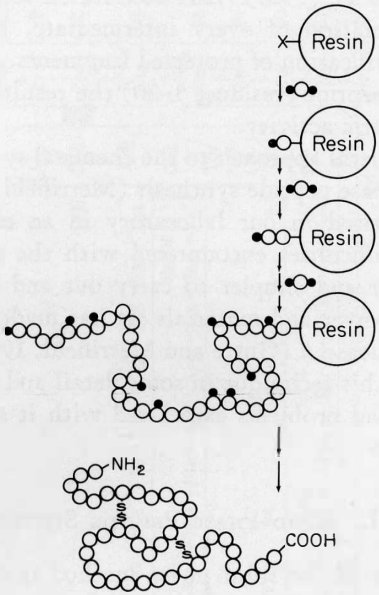


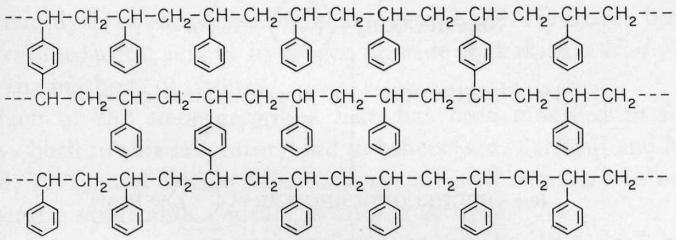
FIG. 3. The basic idea of solid phase peptide synthesis. The large circle represents the solid support, the small open circles represent amino acid residues, and the dark circles represent protecting groups and activating groups. From Merrifield, 1971, by permission of Gordon and Breach, New York.

process could be automated and thereby introduce further efficiency into the synthesis. It only remained to translate the general idea into a workable set of reactions.

The support which we have found to be most satisfactory is a copolymer of styrene and divinylbenzene, containing 1 or 2% of the cross-linking agent. The resulting beads are illustrated in Fig. 4. They are about 50 μm in diameter when dry and swell to nearly 8 times their volume in organic solvents such as the halogenated hydrocarbons. This means that the polystyrene chains are highly solvated during the chemical reactions and are readily accessible to diffusing reagents. The reactions occur not only at the surface of the bead, but, in major part, within the interior matrix of the bead. This could be readily demonstrated by autoradiography (Merrifield and Littau, 1968) of cross sections of beads containing a peptide in

which tritiated proline had been incorporated. The gross distribution of peptide chains was quite uniform throughout the sections. The technique did not, however, have the resolving power to show the distribution at the molecular level and heterogeneity at that resolution is probable.

Many detailed chemical schemes have been devised to carry out solid phase synthesis. An early one which proved to be very useful is shown in Fig. 5. The aromatic rings of the styrene can be derivatized in several ways in order to provide an attachment site for the first amino acid. The first successful derivative was obtained by



Styrene–divinylbenzene Copolymer

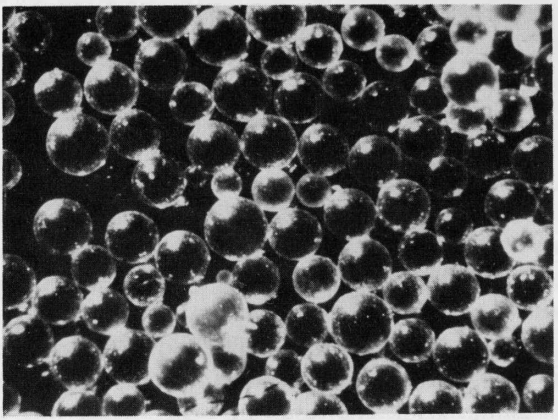


FIG. 4. The styrene-divinylbenzene copolymer support. The spheres average about 50 μm in diameter.

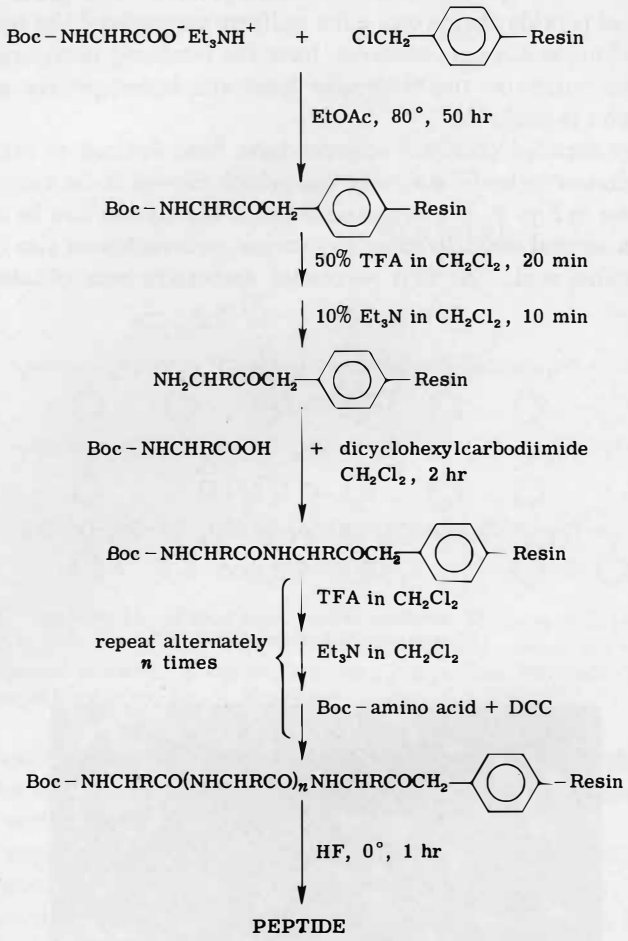


FIG. 5. A scheme for solid phase peptide synthesis.

chloromethylation to give a substituted benzyl chloride which could be esterified by reaction with the triethylammonium salt of an N^{α}-protected amino acid. The resulting benzyl ester was carefully chosen because it is quite stable for the remainder of the synthesis but can be selectively cleaved at the end with strong anhydrous

acid. In order to be compatible with the anchoring bond the amino protection should be relatively labile to acid, and the *t*-butyloxycarbonyl (Boc) group is commonly employed for that reason. It can be removed by 1 *N* HCl or by 20% trifluoroacetic acid in methylene chloride, for example. The resulting amine salt is neutralized by a tertiary amine, such as triethylamine, and is then ready to couple to the next Boc-amino acid to form the first peptide bond. Any one of a number of peptide-forming reagents can be applied. We have found dicyclohexylcarbodiimide (Sheehan and Hess, 1955) to be especially useful. Subsequent amino acids are incorporated in exactly the same way by alternately repeating the deprotection and coupling reactions. The last step is the cleavage of the peptide from the resin by acidolysis of the anchoring benzyl ester bond. Hydrogen bromide in trifluoroacetic acid or hydrogen fluoride (Sakakibara *et al.*, 1967) are the methods of choice.

Each of the steps described here has been modified in several ways both in this laboratory and in others (see Marshall and Merrifield, 1971), and certain other changes are still necessary before the generally applicable, "ideal" system is evolved.

The reaction sequences just described can be illustrated by the synthesis of bradykinin, which is outlined in Fig. 6 (Merrifield, 1964). Bradykinin is a nonapeptide, derived from a plasma protein, with strong hypotensive and smooth muscle contracting properties. The solid phase synthesis began by attaching the protected C-terminal amino acid, Boc-nitroarginine, to the polystyrene support and proceeded by carrying out the three basic steps of deprotection, neutralization and coupling, alternately, eight times each until the final peptide sequence was assembled. The peptide was released from the resin by anhydrous HBr in trifluoroacetic acid, and at the same time some of the protecting groups were also removed. The remainder of the side chain protecting groups, in this case from the two nitroarginines, were removed by catalytic hydrogenation to give the crude bradykinin product. Purification by ion exchange chromatography finally gave analytically pure bradykinin, which was shown to have full biological activity when compared with the natural hormone.

Following this early synthesis large numbers of analogs of bradykinin (Stewart and Woolley, 1966) and several other small peptide

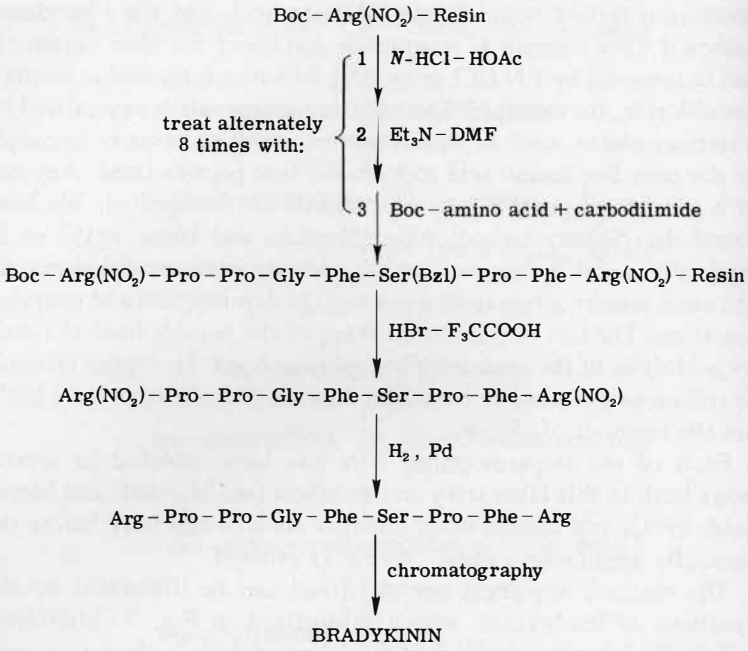


FIG. 6. A solid phase synthesis of bradykinin.

hormones, such as angiotensin (Marshall and Merrifield, 1965) and oxytocin (Manning, 1968), have been synthesized by this same procedure and by various modifications of it. The method has been extended to insulin (Marglin and Merrifield, 1966) and to molecules such as ACTH, parathyroid hormone, secretin, valinomycin, trypsin inhibitor, several hormone-releasing factors, and a rather sizable number of model peptides for use in a variety of special problems.

The requirements for a successful synthesis by the solid phase method are: (1) rapid reactions, (2) quantitative reactions, (3) no side reactions. We can approach these requirements, but obviously we do not quite realize them even in the best case. In spite of the fact that these are heterogeneous reactions, they are very fast (Fig. 7), and there is no real problem in that regard. Some second-order rate constants for the coupling reaction have been measured which

indicate that 99% of the reaction between the hindered amino acids isoleucine and valine required 140 seconds, while the coupling was 99% complete within 10 seconds for most unhindered amino acids.

At the moment the most pressing problem is to devise precise, sensitive, and rapid methods to monitor the coupling reaction, because an indication of quantitative bond formation is crucial to the success of the method. A number of techniques are at hand, but no one of them meets all the requirements satisfactorily. It is obviously not sufficient to measure the disappearance of the amino acid reagent because it is present in excess at the outset of the reaction and precision will only be of the order of a few percent, whereas we

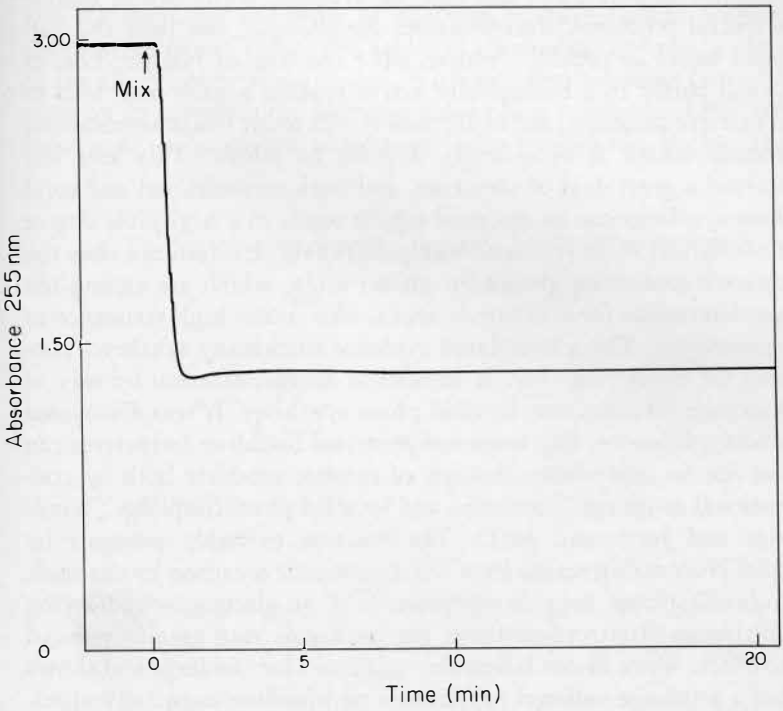


FIG. 7. A rate curve for the coupling reaction. Coupling reaction: Bpoc-Gly-anhydride (1.6×10^{-2} M) + Val-resin (0.8×10^{-2} M). $k_2 = 4.0$ liters/mole·sec. Bpoc = biphenylisopropyloxycarbonyl.

are looking for an accurate measure of the last tenths or even hundredths of a percent of the reaction. A more sensitive test is the determination of the amount of unreacted peptide chain, and several ways to measure the terminal amino group have been devised. It can be estimated by the ninhydrin reaction (Kaiser *et al.*, 1970) or, more quantitatively, by electrometric titration with perchloric acid (Brunfeldt *et al.*, 1969) or by titration of its hydrochloride (Dorman, 1969) or picrate salt (Gisin, 1972) or by conversion to a Schiff base and colorimetric assay of the aldehyde component (Esko *et al.*, 1968).

A second important and continuing effort is directed toward the identification and elimination of side reactions. Several have been examined in detail and can now be avoided, while others remain as special problems. Racemization, for example, has been the universal worry of peptide chemists since the time of Fischer. Loss of optical purity in a biologically active peptide usually will lead to an inactive molecule, and in all cases it will result in a diastereomeric mixture which is exceedingly difficult to purify. This area has received a great deal of attention, and both conventional and solid phase syntheses can be designed which result in a negligible degree of racemization. It is a remarkably fortunate circumstance that the urethane protecting groups for amino acids, which are among the best derivatives for synthetic reasons, also confer high resistance to racemization. The accumulated evidence from many syntheses confirms the belief that there is essentially no racemization by way of oxazolone intermediates in solid phase syntheses. It was discovered recently, however, that urethane-protected histidine derivatives can give rise to appreciable amounts of racemic products both by conventional couplings in solution and by solid phase couplings (Windridge and Jorgensen, 1971). The reaction probably proceeds by direct proton abstraction from the asymmetric α-carbon by the basic imidazole group, because the presence of an electron withdrawing substituent (dinitrophenyl) on the imidazole ring greatly reduced the effect. Work in our laboratory confirms their findings and shows that a *p*-toluene sulfonyl substituent on histidine essentially eliminates the racemization problem (Lin *et al.*, 1972).

Dr. Gisin has recently discovered another side reaction during a solid phase synthesis and, by studying its mechanism, has devised a

way to eliminate it (Gisin and Merrifield, 1972). During the preparation of Boc-D-Pro-D-Val-L-Pro-resin it was observed that over 70% of the peptide had been lost from the resin. By examining each of the steps in the synthesis, it was found that the loss occurred during the customary 10-minute equilibration treatment of D-Val-L-Pro-resin with Boc-D-Val prior to addition of the coupling reagent. It was then shown that the product was D-Val-L-Pro-diketopiperazine and that its formation was strongly catalyzed by carboxylic acids—in this case by Boc-D-Pro (Fig. 8). Imino acid-containing dipeptides, such as those containing proline or *N*-methylalanine, are especially susceptible to the diketopiperazine formation because of the small energy difference between the cis and trans amide conformations. Once the problem was understood, a simple solution was forthcoming; the order of addition of reagents was reversed, the presence of free carboxylic acid was eliminated, and the catalysis of the ring closure was avoided. With this improvement Gisin has gone on to complete the stepwise solid phase synthesis of the very interesting valinomycin analog -[L-Val-D-Pro-D-Val-L-Pro]$_3$-. This

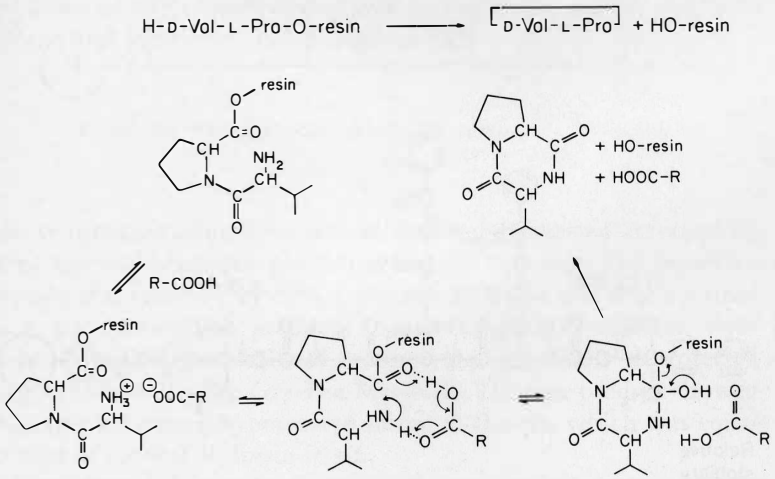


FIG. 8. The proposed mechanism for the acid-catalyzed formation of a diketopiperazine. From Gisen and Merrifield (1972), by permission of the American Chemical Society, Washington, D.C.

is the first cyclic peptide composed entirely of amino acids that can selectively bind potassium and carry it from an aqueous solution into a hydrophobic organic phase.

Because the solid phase method in its usual form does not allow isolation of intermediates, some people feel that it can never be applied to the synthesis of high molecular weight products. I believe that it can be utilized for such problems at the present time and that, through the improvements which are constantly being made, the continuous, stepwise approach will eventually be the method of choice. For those who prefer to isolate protected intermediates, we have devised a new solid phase-fragment approach (Wang and Merrifield, 1969) in which small protected peptides are prepared on a resin support and, following cleavage and purification, are coupled in solution to give larger peptides. This technique, therefore, combines some of the best features of the solid phase and classical approaches into a unified procedure. In its initial and usual form, as already described here, the solid phase method produces partially

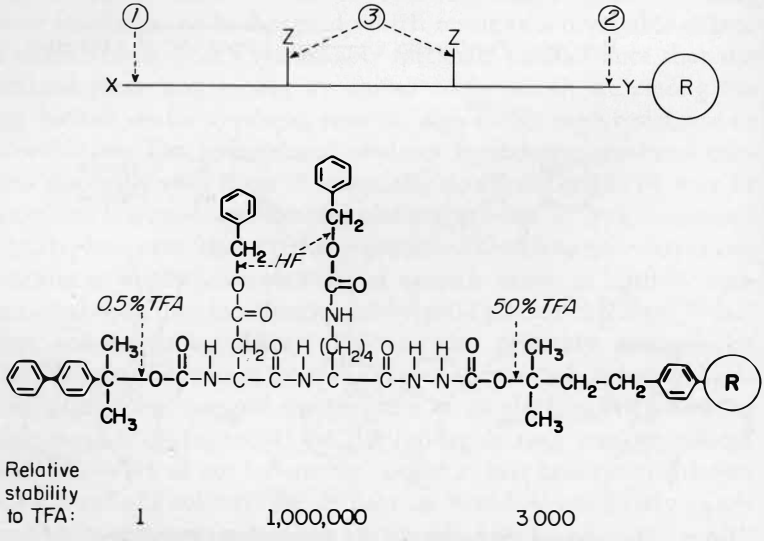
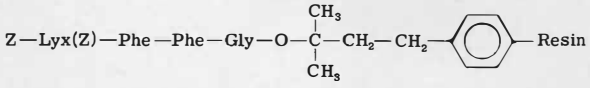


FIG. 9. A solid phase-fragment scheme. From Wang and Merrifield (1972), by permission of Munksgaard, Copenhagen.

or completely deprotected peptides following the cleavage step in strong acid. These products, of course, are not suited for further couplings because they contain functional side chains. To overcome the problem, three classes of protecting groups are needed which can be selectively removed, as illustrated in Fig. 9. The group on the α-amine must be readily removed at each step of the synthesis without loss of side chain protection or of the anchoring bond to the resin, and at the end of the synthesis it must be possible to cleave the anchoring bond without loss of side chain protection. Finally, after the intermediate protected peptide has been purified and coupled to other fragments it must be possible to completely deprotect the finished polypeptide. The scheme developed by Dr. Wang makes use of three greatly different degrees of acid stability. He uses the very labile biphenylisopropyloxycarbonyl (Bpoc) group of Sieber and Iselin (1968) for α-amine protection, relatively stable benzyl or tosyl groups for side chain protection, and a new linkage of intermediate stability for the carboxyl attachment. These groups differ by 3 and 6 orders of magnitude in their susceptibility to acid and, therefore, allow the satisfactory synthesis of protected fragments. A recent synthesis of an eledoisin analog will illustrate this general approach (Wang and Merrifield, 1972). The analog



was synthesized using Bpoc derivatives for the intermediate coupling steps and carbobenzoxy protection for the last step. The protected peptide was removed by 50% trifluoroacetic acid and was obtained as a pure crystalline product. It was coupled in solution with H-Leu-Met-OMe, crystallized, aminolyzed and finally deprotected to give H-Lys-Phe-Phe-Gly-Leu-Met-NH_2. The pure product showed the expected strong hypotensive action in the rat, which was equal to that of natural eledoisin itself.

Until now I have not discussed the mechanical aspects of this technique except to say that it was because of this potential feature that the solid phase method was developed in the first place. The ability to purify by simple filtration and washing and the fact that

all reactions could be conducted within a single reaction vessel appeared to lend themselves ideally to a mechanized and automated process. Initially, a simple manually operated apparatus was constructed (Fig. 10), and later, together with Dr. Stewart, we designed and built the automated instrument shown in Fig. 11 (Merrifield *et al.*, 1966). The essential features were the reaction vessel containing the resin with its growing peptide chain and the necessary plumbing to enable the appropriate solvents and reagents to be pumped in, mixed, and removed in the proper sequence. The mechanical events were under the control of a simple stepping drum programmer and a set of timers. This instrument has been used for the synthesis of many small peptides and for the synthesis of ribonuclease A. In the past few years at least nine commercial instruments have been constructed in several countries. They differ considerably in detail, but are designed to carry out the same chemistry.

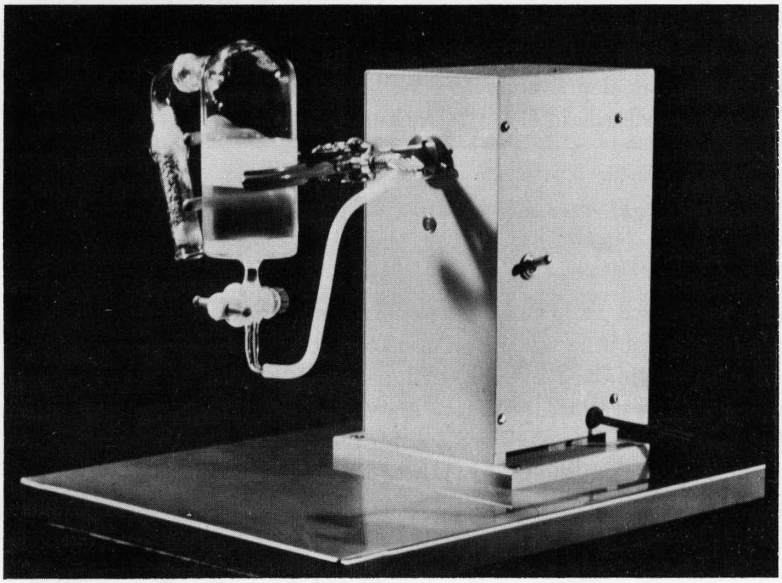


FIG. 10. A manually operated reaction apparatus, showing the glass reaction vessel with a fritted-glass filter disk and side arm. The resin is suspended by a rocking motion about the horizontal axis.

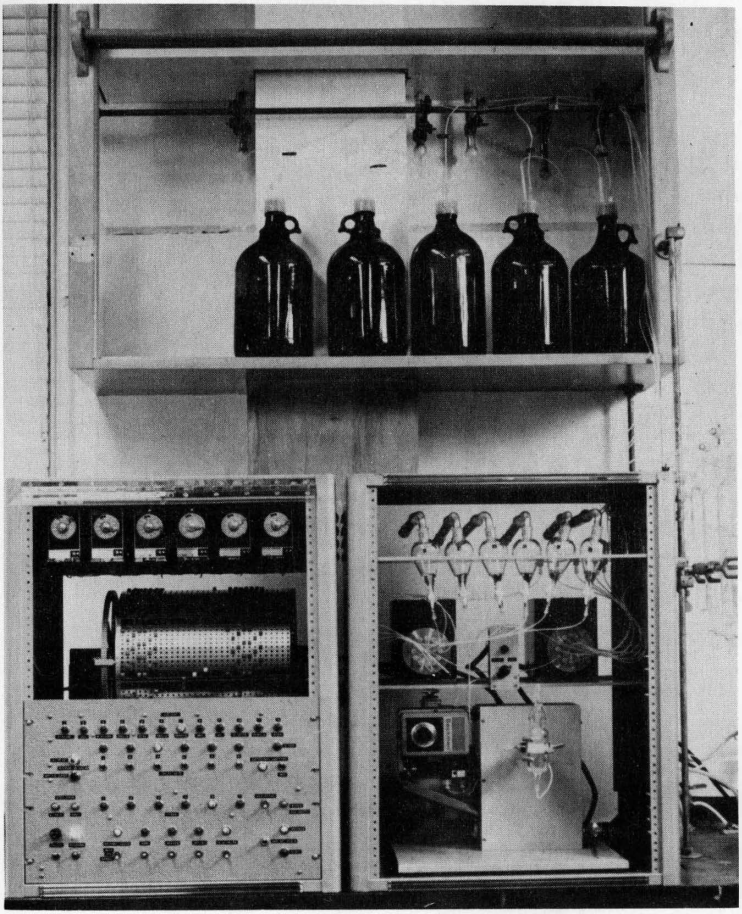


FIG. 11. The automated peptide synthesizer.

III. The Synthesis of Ribonuclease A

As a first step toward bringing to bear the synthetic approach to the study of enzymes we undertook the total synthesis of bovine pancreatic ribonuclease A (Gutte and Merrifield, 1969, 1971). This molecule, which has been so closely associated with this institution, was well suited for the synthetic work because it is small and

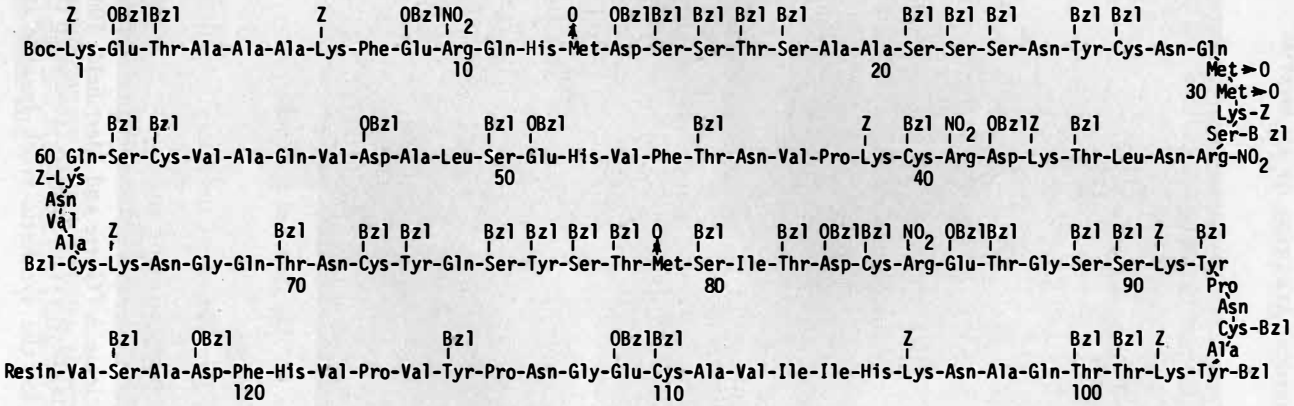


FIG. 12. Protected ribonuclease A.

stable and its amino acid sequence (Hirs *et al.*, 1960; Smyth *et al.*, 1963) and X-ray crystallographic structure were known (Kartha *et al.*, 1967; Wyckoff *et al.*, 1967, 1970).

The formula in Fig. 12 shows the protected derivative of ribonuclease which we wished to synthesize. It contains a total of 67 side chain protecting groups, largely based on benzyl derivatives for stability, and has a molecular weight of 19,791. The synthesis carried out by Dr. Gutte followed essentially the same procedure as that outlined in Fig. 5, while the workup, identification, and analysis of the product utilized a variety of techniques. The synthesis, which has already been described in complete detail (Gutte and Merrifield, 1971), is summarized in Table I. It shows that the overall yield was about 3%, i.e., of the first amino acid, valine, originally attached to the resin 3% was eventually recovered in the

TABLE I

RIBONUCLEASE A SYNTHESIS

Stage of synthesis	Overall yield Mg	%
Boc-Val-Resin	2000	100
Deprotect, Neutralize, Couple — Repeat 123 times		
17% ↓		
Protected RNase-Resin	3430	17
Cleave and deprotect HF		
71% ↓		
Crude RNase $(SH)_8$	697	12
Sephadex G-75		
53% ↓		
RNase A (monomer fraction)	373	6.4
69% Trypsin digestion, Sephadex G-50		
RNase A (Trypsin resistant fraction)	256	4.4
66% $(NH_4)_2SO_4$ fractionation		
RNase A	169	2.9

final purified protein. A large accumulated loss occurred during the synthesis itself. This means that the benzyl ester bond anchoring the peptide to the resin was not entirely stable to the repeated acidic conditions of the synthesis and was lost to the extent of about 1.4% at each cycle. I must emphasize that this has nothing to do with the individual coupling yields, which we believe were high. The cleavage step and the numerous workup steps taken together accounted for another 80% loss of material. The quantities of ribonuclease indicated in Table I are the calculated amounts of product that would have been obtained if all the material had been carried through all these steps. During this work the material was divided into many fractions in order to work out the best procedures and to obtain analytical samples, so that the most ribonuclease available at any time was of the order of 25 mg. When the crude HF-cleaved protein was oxidized and fractionated on an IRC-50 column in phosphate buffer it showed a large peak containing the enzymatic activity and two small inactive peaks. The specific activity could not be raised above 13% by recycling through this column. We then decided to take advantage of the known stability of ribonuclease A to trypsin digestion (Dubos and Thompson, 1938). The reasoning was that material with the correct structure would come through the treatment intact, while molecules with incorrect structures due to incomplete sequences or improper folding would give rise to a mixture of small molecules that could readily be separated from the native molecule (Fig. 13). To our great satisfaction the activity of a crude sample treated in this way increased from 8% to 61%, while only 25% of the protein was actually removed. A further purification step involved an ammonium sulfate fractionation in which some amorphous material could be separated, and a lyophilized preparation possessing nearly 80% specific activity was finally obtained. Table II summarizes the activity data obtained at various stages of purification of the synthetic enzyme. It also shows how the total number of units of RNase activity increased as the purification proceeded.

We interpret these data to mean that inhibitory peptides were present which were giving rise to low enzyme activity values and the crude preparation actually contained more enzyme than we had thought. The purified RNase A was then placed on a CM-cellulose

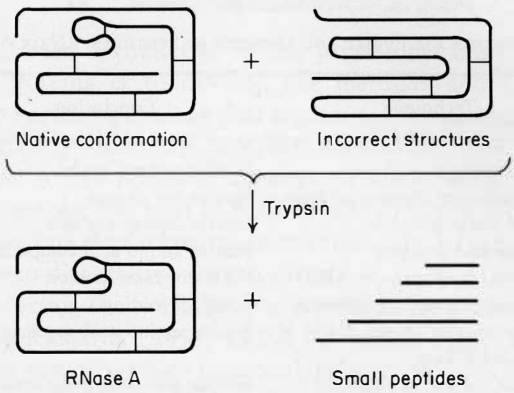


FIG. 13. Schematic view of the purification of synthetic ribonuclease A by trypsin digestion.

column and compared with natural RNase A and reduced-reoxidized natural RNase A. They were identical by this criterion, which was the one first used by White (1961) to show that RNase A after reduction and reoxidation was indistinguishable from native RNase. His was the demonstration that led to the hypothesis that the primary structure of the protein determined its tertiary structure. Our synthesis provides a new kind of evidence for this hypothesis. The fact that the only information put into the synthesis was the linear sequence means that the primary structure must be sufficient

TABLE II

RIBONUCLEASE A ACTIVITY

Purification stage	Specific activity (%)	Total activity (mg RNase/2 gm resin)
HF cleavage	2	14
Sephadex G-75	9	33
IRC-50	13	53
Trypsin	61	156
$(NH_4)_2SO_4$	78	132

TABLE III

EVIDENCE FOR PURITY AND IDENTITY OF SYNTHETIC RNASE A

Technique	Conclusion
1. Gel filtration	Similar molecular size
2. Ion-exchange chromatography	Similar net charge
3. Paper electrophoresis	Similar charge and size
4. Amino acid analysis	Similar amino acid composition
5. Digested by papain + APM	No major racemization
6. Resistant to trypsin digestion	Compact folding
7. Fingerprint of tryptic digest of oxidized RNase	Similar amino acid sequence
8. Neutralization of antibody	Similar antigenic determinants
9. Enzymatic assay	High specific activity and similar substrate specificity

to direct the final folding of the molecule into its active tertiary structure. By starting with free amino acids any objections to the original experiments based on retention of some residual conformation after the reduction step which might serve to initiate the refolding are eliminated. The synthesis of an active enzyme containing no substituents except amino acids also provides a new proof for the now well-established belief that enzymatic activity can be a consequence of a simple protein and that no other components either known or unknown need be present. Not too many years ago such a view was vehemently opposed.

What we cannot claim is that this synthesis constitutes a proof of structure for ribonuclease A. Such a proof in the classical organic chemistry sense will be very difficult indeed because it is virtually impossible with present methods to establish absolute identity between two macromolecules of such complexity.

The various kinds of evidence that have been obtained to establish the identity and purity of the synthetic RNase A are listed in Table III. Clearly, each of these tests has limitations and we can only say that within these limits the synthetic preparation closely resembles native ribonuclease A in its chemical and physical properties, that it has a high specific activity and that it shows the expected substrate specificity.

IV. Structure Function Studies

As I indicated earlier, one of the prime objectives of synthetic work on proteins is to develop the methods and capability to investigate various questions that are not easy to examine by studies of the naturally occurring proteins themselves. The synthesis of ribonuclease A has provided answers to some rather fundamental questions, and we feel that it has laid a foundation for new studies on structure–function relations in enzymes. As a bare beginning, I would like to discuss two such studies.

In considering the well known S-peptide/S-protein system discovered by Richards (1955, 1958), in which these two inactive components can recombine noncovalently to produce a fully active enzyme, we wondered whether the recombination would occur if the protein component were shortened at its amino terminus by the removal of residues 21 through 25, and whether the complex would be enzymatically active. To test this question a sample of peptide-resin containing residues 25–124 was removed during the RNase synthesis, and another containing residues 21–124 was removed after the addition of 5 more residues (Fig. 14). These two synthetic proteins, corresponding to des-21–25 S-protein and to S-protein,

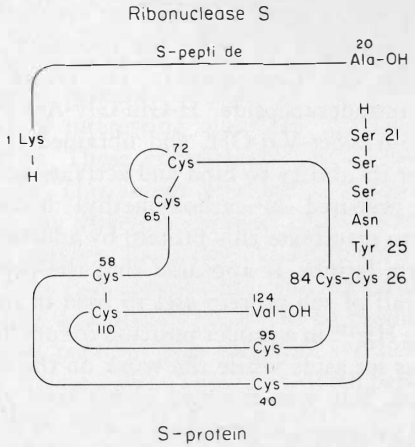


FIG. 14. The S-peptide, S-protein, des 21–25 S-protein system. From Gutte and Merrifield (1970), by permission of The American Society of Biological Chemists, Inc.

were partially purified on Sephadex and then reduced at their 4 disulfide bonds. The resulting random coils were mixed with S-peptide (natural or synthetic), reoxidized and assayed for enzyme activity. The crude mixtures were both found to have approximately 25% as much activity as the product derived from native S-protein by the same treatment. From these data we can say, qualitatively, that the first five residues, 21–25, are definitely not necessary for the binding and reactivation to occur and, quantitatively, that there is probably no appreciable effect whether they are present or absent. Wyckoff and Richards had already predicted from their X-ray data that the three serines at positions 21, 22, and 23 would probably not be necessary since there was no indication of noncovalent interactions with other residues in the crystal. In contrast, asparagine 24 and tyrosine 25 appeared to be involved in a total of 5 hydrogen bonds, and it was reasonable to expect that they would be important in stabilizing the complex between S-peptide and S-protein and might be necessary for the formation of an active complex. Our data indicate that those bonds are not required either for binding or activity.

About 5 years ago I became interested in the question of whether or not a peptide component from the carboxyl end of ribonuclease might function in a manner analogous to that of S-peptide at the amino end. This was a good problem for synthesis because there was no enzyme available to make the necessary cleavage in the vicinity of residue 111. Consequently, Dr. Gutte synthesized the RNase 111–124 tetradecapeptide, H-Glu-Gly-Asn-Pro-Tyr-Val-Pro-Val-His-Phe-Asp-Ala-Ser-Val-OH, and obtained it in a high state of purity. To test its ability to bind and activate an inactive protein component he prepared N^1-carboxymethyl histidine-119-RNase. Many attempts to reactivate this protein by addition of the peptide met with uniform failure. It appeared that the peptide was unable to displace the tail of the protein and to bind in its place in a way that would place His^{119} in a proper position to function catalytically. The problem was set aside while the work on the total synthesis of RNase proceeded, but was resumed about two years later when Lin *et al.* (1968) succeeded in preparing a series of shortened RNases. They made RNase 1–120 by peptic digestion and RNase 1–119 and RNase 1–118 by further digestions with carboxypeptidase A. When

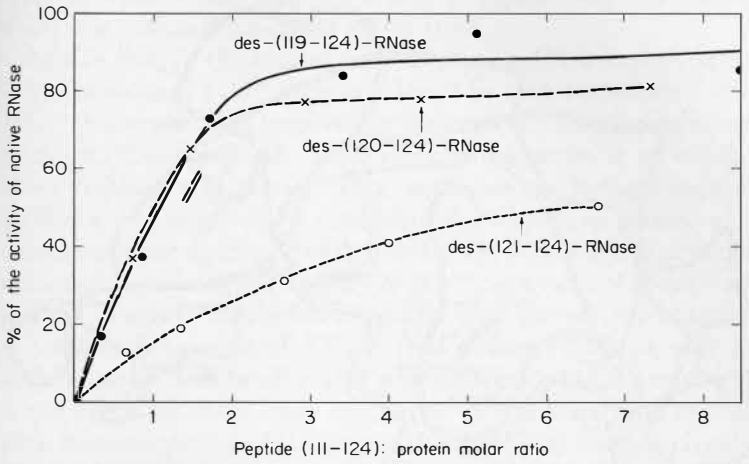


FIG. 15. Enzymatic activity of the synthetic RNase 111–124 tetradecapeptide with shortened ribonucleases. From Lin *et al.* (1970), by permission of The American Society of Biological Chemists, Inc.

the synthetic peptide 111–124 was assayed in the presence of these inactive proteins, high enzymatic activity was regenerated (Fig. 15). Through a joint effort with Lin, Moore, Gutte, and Caldi, we have studied the combination of a number of carboxyl-terminal peptides with the shortened protein components and have been able to answer a number of interesting questions (Lin *et al.*, 1970, 1972; Gutte *et al.*, 1972). These experiments can be discussed best by referring to Fig. 16. This shows how we visualize the peptide inserting itself into the three-dimensional structure of the protein to produce an active enzyme. As can be seen there are 8–10 overlapping residues between the two components, and we naturally wondered how long a peptide was actually required to produce the effect. A series of C-terminal peptides was therefore synthesized and tested in the presence of RNase 1–118 (Table IV). With peptides containing 7 or 8 residues there was very little activity, but the addition of Val^{116} made a dramatic increase to 60%. Further lengthening of the chain increased the binding gradually to a maximum of 98%. The role of Val^{116} is now being examined in more detail to try to de-

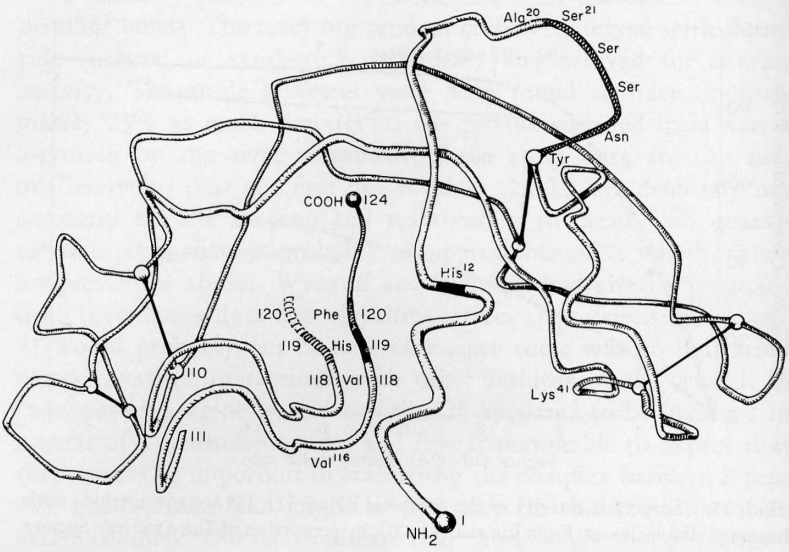


FIG. 16. Schematic representation of peptide-protein complexes in reconstituted ribonuclease. The drawing is based on the X-ray structure of RNase A by Kartha *et al.* (1967).

TABLE IV

EFFECT OF CHAIN LENGTH ON ACTIVATION OF RNASE 1–118 WITH COOH-TERMINAL PEPTIDES

Chain length	Peptide sequence	Activity[a] (%)
7	Val-His-Phe-Asp-Ala-Ser-Val	1
8	Pro-Val-His-Phe-Asp-Ala-Ser-Val	1.5
9	Val-Pro-Val-His-Phe-Asp-Ala-Ser-Val	60
10	Tyr-Val-Pro-Val-His-Phe-Asp-Ala-Ser-Val	70
11	Pro-Tyr-Val-Pro-Val-His-Phe-Asp-Ala-Ser-Val	80
12	Asn-Pro-Tyr-Val-Pro-Val-His-Phe-Asp-Ala-Ser-Val	90
14	Glu-Gly-Asn-Pro-Tyr-Val-Pro-Val-His-Phe-Asp-Ala-Ser-Val	98

[a] Maximum activity, in presence of excess peptide, relative to an equimolar amount of native RNase A. From Gutte *et al.* (1972), by permission of the Amerian Society of Biological Chemists, Inc.

termine whether it is simply a matter of peptide length or whether there is a unique function for valine itself.

Notice that in the combination of peptide 111–124 with protein 1–119 or protein 1–120 there is a His^{119} in both components. Since His^{119} is known to be involved in the catalytic mechanism of ribonuclease (Crestfield *et al.*, 1963) the question arose as to which of these residues was participating in the reconstituted complexes. Evidence was obtained by examining the alkylation pattern of the complexes after treatment with iodoacetate (Table V). Native ribonuclease alkylates either at His^{12} or His^{119}, in a ratio of about 1 to 7, and the extent of alkylation correlates with the enzymatic activity of various derivatives of RNase; thus RNase 1–120 has very low activity and is poorly alkylated while RNase 1–119 is inactive and is not alkylated under these conditions. It was found that the complex between peptide 111–124 and RNase 1–120 was alkylated at His^{119} of the protein component and also at His^{119} of the peptide in a ratio of about 1 to 4, and the complex between peptide 111–124

TABLE V

ALKYLATION OF HISTIDINE RESIDUES IN THE RECONSTITUTED COMPLEXES

Components		Site of alkylation		
		His^{12}	His^{119}	
Protein	Peptide	Protein	Protein	Peptide
RNase 1–124 (native RNase A)		0.12	0.77	—
RNase 1–120		0.20	0.06	—
RNase 1–119		0	0	—
RNase 1–120 + RNase 111–114		0.14	0.24	0.06
RNase 1–119 + RNase 111–124		0.10	0.50	0.40
RNase 1–118 + RNase 111–124		0.10	—	0.70
RNase 1–118 + [Ile^{120}]RNase 111–124		0.10	—	0.07

[a] Determined by measurement of the number of moles of N^3-His(Cm) or N^1-His(Cm) per mole of protein. The N^3-His(Cm) is assumed to occur only at His^{12} of the protein and N^1-His(Cm) only at His^{119} of the protein or peptide. From Gutte *et al.* (1972), by permission of The American Society of Biological Chemists, Inc.

and RNase 1–119 alkylated to nearly the same extent in both components. We interpret this to mean that these complexes exist in at least two conformations in which the peptide His^{119} is functioning catalytically part of the time and the protein His^{119} is functioning part of the time. Since dissociation constants could be measured, the two conformational forms must be in equilibrium with one another. In the case of peptide 111–124 and protein 1–118 there is only one His^{119}, and all of the alkylation occurred in the protein.

A third question which we have studied by synthesis concerns the role of Phe^{120}. This residue is known to stabilize ribonuclease and was expected to be involved in binding the synthetic peptides to the shortened RNase molecules. There is also evidence from X-ray and nuclear magnetic resonance studies that Phe^{120} is located very close to the pyrimidine ring of the substrate and may be involved in binding substrate to RNase. We have obtained evidence on these points by synthesizing replacement analogs and combining them with RNase 1–118, (Table VI). Phe^{120} in the peptide 111–124 was replaced by the hydrophobic but nonaromatic residues leucine and isoleucine and by the large aromatic residue tryptophan. It was found that the dissociation constant, K_d, increased by a factor of 12 when Leu^{120} was present, by 20 when Ile^{120} was present, and by

TABLE VI

Effect of Phe^{120} Replacements on the Dissociation Constant of the Peptide–Protein Complexes and on the Michaelis Constant and Inhibitor Dissociation Constant[a]

Components of the complex	K_d (M)	K_m for C > p (M)	K_i for 2′ Cp (M)
RNase A	—	0.7×10^{-3}	0.9×10^{-5}
[Phe^{120}]RNase 111–124 + RNase 1–118	2.0×10^{-7}	0.7×10^{-3}	—
[Leu^{120}]RNase 111–124 + RNase 1–118	2.5×10^{-6}	1.0×10^{-3}	—
[Ile^{120}]RNase 111–124 + RNase 1–118	4.0×10^{-6}	1.2×10^{-3}	1.1×10^{-5}
[Trp^{120}]RNase 111–124 + RNase 1–118	3.5×10^{-5}	—	—

[a] From Lin *et al.* (1972), by permission of The American Society of Biological Chemists, Inc.

170 when Trp^{120} was present, suggesting a failure to fit properly into the hydrophobic pocket normally occupied by Phe^{120}. Furthermore, even when the peptide was present in saturating concentrations, full enzymatic activity could not be regenerated by these peptide analogs. It therefore appears that phenylalanine plays a very important role in binding the carboxyl end of ribonuclease to the remainder of the protein and in properly aligning the catalytic site.

To obtain information about the role of Phe^{120} in binding substrate, the Michaelis constant, K_m, for 2′,3′-cyclic cytidylic acid and the substrate inhibition constant, K_i, for 2′-cytidylic acid were examined. These values are related to the strength of binding of substrate to the enzyme complex. Within the experimental error, K_m, was essentially the same for native RNase and for the complexes containing Phe, Leu, or Ile at position 120, indicating that phenylalanine had no unique role in binding substrate, and that if it played any part in the binding its role could be equally well assumed by leucine or isoleucine. A similar conclusion could be reached from the inhibition constant for 2′-cytidylic acid which was nearly the same whether Phe^{120} or Ile^{120} was present.

The final experiment I would like to mention concerns the assembly of a three-component enzyme (Fig. 16). We discovered that S-peptide (1–20), containing His^{12}, and the C-terminal peptide (111–124), containing His^{119}, could be mixed in solution with RNase (21–118), containing Lys^{41}, and that they would bind noncovalently to produce a complex possessing about 30% specific activity. Thus, three separate peptides, each containing one of the amino acids known to function in the catalytic mechanism of ribonuclease, could reassociate in a very specific way to reconstitute the catalytic site and the substrate binding site and produce the active enzyme.

V. Conclusions

I have described a new approach to peptide synthesis, called solid phase peptide synthesis, and have indicated some of its virtues and some of the disadvantages and problems associated with it. The application of this technique to the synthesis of small peptides was described, and its use in the first total synthesis of an enzyme was

discussed. Finally, some examples were presented to show how the synthetic approach can be used to answer questions about the structure and function of proteins.

It is now possible to synthesize true proteins with real enzymatic activity beginning with simple amino acids. I believe that in the years ahead we will see a great deal of progress in this area which will contribute in a significant way to our understanding of this very important class of compounds.

ACKNOWLEDGMENTS

I want to thank my friends who have worked with me during the past twelve years on the subject of this Harvey Lecture. The contributions of Bernd Gutte, although discussed in some detail, must be reemphasized here, and, especially, I wish to recall the help and guidance of Dr. D. W. Woolley. From its outset this research has been supported by Grant AM 1260 from the National Institutes of Health.

REFERENCES

Brunfeldt, K., Roepstorff, O., and Thomsen, J. (1969). *Acta Chem. Scand.* **23,** 2906.

Crestfield, A. M., Stein, W. H., and Moore, S. (1963). *J. Biol. Chem.* **238,** 2413.

Dorman, L. C. (1969). *Tetrahedron Lett.*, 2319.

Dubos, R. J., and Thompson, R. H. S. (1938). *J. Biol. Chem.* **124,** 501.

Epstein, C. J., Goldberger, R. F., Young, D. M., and Anfinsen, C. B. (1962). *Arch. Biochem. Biophys.*, Suppl. 1, 223.

Esko, K., Karlson, S., and Porath, J. (1968). *Acta Chem. Scand.* **22,** 3342.

Gisin, B. F. (1972). *Anal. Chim. Acta.* **58,** 248.

Gisin, B., and Merrifield, R. B. (1972). *J. Amer. Chem. Soc.* **94,** 3102.

Gutte, B., and Merrifield, R. B. (1969). *J. Amer. Chem. Soc.* **91,** 501.

Gutte, B., and Merrifield, R. B. (1971). *J. Biol. Chem.* **246,** 1922.

Gutte, B., Lin, M. C., Caldi, D. E., and Merrifield, R. B. (1972). *J. Biol. Chem.* **247,** 4763.

Hirs, C. H. W., Moore, S., and Stein, W. H. (1960). *J. Biol. Chem.* **235,** 633.

Hirschmann, R., Nutt, R. F., Veber, D. F., Vitali, R. A., Varga, S. L., Jacob, T. A., Holly, F. W., and Denkewalter, R. G. (1969). *J. Amer. Chem. Soc.* **91,** 507.

Kaiser, E., Colescott, R. L., Bossinger, C. D., and Cook, P. I. (1970). *Anal. Biochem.* **34,** 595.

Kartha, G., Bello, J., and Harker, D. (1967). *Nature (London)* **213,** 862.

Katsoyannis, P. G., Fukuda, K., Tometsko, A., Suzuki, K., and Tilak, M. (1964). *J. Amer. Chem. Soc.* **86,** 930.

Kung, Y. T., Du, Y. C., Huang, W. T., Chen, C. C., Ke, L. T., Hu, S. C., Jiang, R. Q., Chu, S. Q., Niu, C. I., Hsu, J. Z., Chang, W. C., Chen, L. L., Li, H. S., Wang, Y., Loh, T. P., Chi, A. H., Li, C. H., Shi, P. T., Yieh, Y. H., Tang, K. L., and Hsing, C. Y. (1965). *Sci. Sinica* **14,** 1710.

Kunitz, M. (1940). *J. Gen. Physiol.* **24,** 15.
Lin, M. C., Stein, W. H., and Moore, S. (1968). *J. Biol. Chem.* **243,** 6167.
Lin, M. C., Gutte, B., Moore, S., and Merrifield, R. B. (1970). *J. Biol. Chem.* **245,** 5169.
Lin, M. C., Gutte, B., Caldi, D. G., Moore, S., and Merrifield, R. B. (1972). *J. Biol. Chem.* **247,** 4768.
Manning, M. (1968). *J. Amer. Chem. Soc.* **90,** 1348.
Marglin, A., and Merrifield, R. B. (1966). *J. Amer. Chem. Soc.* **88,** 5051.
Marshall, G. R., and Merrifield, R. B. (1965). *Biochemistry* **4,** 2394.
Marshall, G. R., and Merrifield, R. B. (1971). *In* "Biochemical Aspects of Reactions on Solid Supports" (G. Stark, ed.), p. 111. Academic Press, New York.
Meienhofer, J., Schnabel, E., Bremer, H., Brinkhoff, O., Zabel, R., Sroka, W., Klostermeyer, H., Brandenburg, D., Okuda, T., and Zahn, H. (1963). *Z. Naturforsch. B* **18,** 1120.
Merrifield, R. B. (1963). *J. Amer. Chem. Soc.* **85,** 2149.
Merrifield, R. B. (1964). *Biochemistry* **3,** 1385.
Merrifield, R. B. (1969). *Advan. Enzymol.* **32,** 221.
Merrifield, R. B., and Littau, V. (1968). *In* "Peptides" (E. Bricas, ed.), p. 179. North-Holland Publ., Amsterdam.
Merrifield, R. B., Stewart, J. M., and Jernberg, N. (1966). *Anal. Chem.* **38,** 1905.
Richards, F. M. (1955). *Trav. Lab. Carlsberg* **29,** 322.
Richards, F. M. (1958). *Proc. Nat. Acad Sci. U.S.* **44,** 162.
Sakakibara, S., Shimonishi, Y., Kishida, T., Okada, M., and Sugihara, H. (1967). *Bull. Chem. Soc. Jap.* **40,** 2164.
Schröder, E., and Lübke, K. (1965). "The Peptides." Academic Press, New York.
Sheehan, J. C., and Hess, G. P. (1955). *J. Amer. Chem. Soc.* **77,** 1067.
Sieber, P., and Iselin, B. (1968). *Helv. Chim. Acta* **51,** 622.
Smyth, D., Stein, W. H., and Moore, S. (1963). *J. Biol. Chem.* **238,** 227.
Stewart, J. M., and Woolley, D. W. (1966). *In* "Hypotensive Peptides" (E. G. Erdos, N. Back, and F. Sicuteri, eds.), p. 23. Springer-Verlag, Berlin and New York.
Wang, S. S., and Merrifield, R. B. (1969). *J. Amer. Chem. Soc.* **91,** 6488.
Wang, S. S., and Merrifield, R. B. (1972). *Int. J. Peptide Protein Res.* **4,** 309.
White, F. H., Jr. (1961). *J. Biol. Chem.* **236,** 1353.
Windridge, G. C., and Jorgensen, E. C. (1971). *Intra-Sciences Chem. Rep.* **5,** 375.
Wyckoff, H. W., Hardman, K. D., Allewell, N. M., Inagami, T., Johnson, L. N., and Richards, F. M. (1967). *J. Biol. Chem.* **242,** 3984.
Wyckoff, H. W., Tsernoglou, D., Hanson, A. W., Knox, J. R., Lee, B., and Richards, F. M. (1970). *J. Biol. Chem.* **245,** 305.