Peptides—Drugs of Tomorrow

Arun Chatterjee*

From time immemorial man is continuously searching for new armaments to fight the diseases. As a result of this search, we have today antibiotics, steroids, prostaglandins, interferon, etc. There is one more field which beckon to us frantically today, one that may solve all the great health problems of mankind, cancer, heart disease, immunological an ethnic weakness, -and that field is the polypeptides and proteins from which a new generation of drugs will come up in place of the drugs of today.

The importance of proteins in the living cell was first recognised by the german chemist G. T. Mulder in 1938. Recent historical research has revealed that this name "Protein" which has come to occupy such an important place in the language of biochemistry, was suggested to Mulder by Berzellius in a letter[1] dated July 10, 1938 and apparently received by Mulder while he was writing the scientific paper in which "Protein" made it's first appearence. During the last two decades, a great deal of the chemical mechanisms underlying physical functions, excepting perhaps the neural, brain and some glandular functions have been discovered. There are in nature a number of peptides and proteins that regulate biological systems and are classified as hormones. Insulin, for example is a regulator of glucose metabolism. The pituitary gland of vertebrates produces a variety of hormones that regulate the growth of the animal, balance of fluid, reproductive cycles and other important biological functions. There are enzymes which are also protein in nature and which selectively catalyze most of the chemical reactions essential for normal cellular function. The RNA's and DNA's, the immunoglobins, the enzyme regulators, and energy carrying substances are complex polymers that depend for their specific function upon moieties which are chains of amino-acids. The RNA and DNA molecules contain polypeptide chains, act as templates for the replication of such chains in the plasma and that the code depends entirely upon the order of amino-acid in the chain.

The highly developed animal world we know exists not only because it's functional systems are valid but also because it contains provisions for corrections of it's malfucnctions. Such provisions also involve activity of specific proteins. Thus, when a doctor prescribes his patients with crude chemicals, or resorts to surgery, he then really is hoping that these measures will mimic the action of some specific proteins to tip the balance in favour of the recovery, in which the body's own protein chemistry is the principal agency. Biochemical research has begun today to identify proteins that can act in that way, and even to strip away parts of the complex molecules

★ Arun Chatterjee, Asst. Quality Control Manager, Dey's Medical Stores Mfg. U. P. Ltd, Naini, Allahabad.

that are found to be redundant. Synthetic vaccines could be the early result of such work. It is believed that genetic disorder of even somatic cells will one day be corrected. These miracles of tomorrow will be possible with the medicines based on peptides and proteins. Already prostaglandins and interferon are on the cards.

The engineer in this field has literally moved ahead of the biomedical researcher, who has the whole business of detailing the polypeptide topography of living systems before him. He has the tools; but all that detail must be done before the golden age of the polypeptide medicines can be entered. It is trully a great challenge. But no effort in history could possibly be as rewarding.

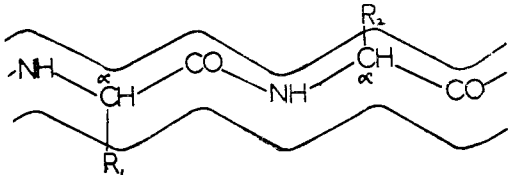
Proteins are macromolecules ranging in molecular weight from 6,000 to several million. Todate 20 amino-acids have been isolated from proteins.

Structure of Protein

All proteins are polymers of amino-acids except two, have an amino group attached to the carbon atom next to the carboxyl group, the α position, i.e., they are α -amino-acids. The amino-acids in a protein molecule are connected by amide linkages (called peptide bonds in this case) made from the carboxyl group of amino-acid and the amino group of the next.

$$H_2\ NCHCONHCHCO$$
$$R_1 \qquad R_2$$

Any one combination of amino-acids is represented by the primary structure of the protein; that is, the primary structure describs the number and order in which the amino acids are arranged in the protein molecule. It deals with the covalent bonding between amino acids in a chain (called a polypeptide chain).



The polypeptide has a backbone[2] and side chains represented by R groups linked to the α-carbons the backbone is not rigid; it is free to rotate around the α-carbons, and by so doing the polypetide can assume any one of many possible shapes, some more stable than the others. A spiral shape (α-helix), for example, is stabilized by forming hydrogen bonds; or if several polypeptides arrange themselves in parallel, held together by hydrogen bonds, they gain stability. The shape resulting from hydrogen bonding of this type is known as the *secondary structure*.

Tertiary structure describes the shape of the folding that the polypeptide assumes as it becomes packed into a three dimensional structure. The foldings are made possibly by mutual interactions of the said chains and interactions of the side chains with the solvent.

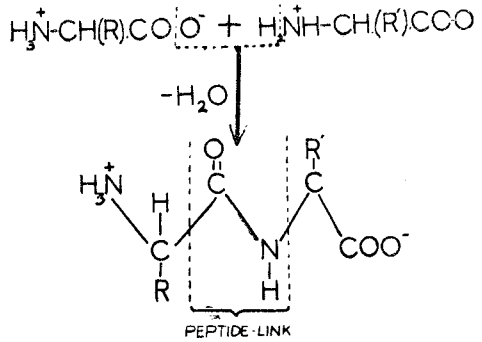
Quarternary structures describe polymeric aggregates resulting when several protein subunits interact and form a larger unit.

Peptide Bond

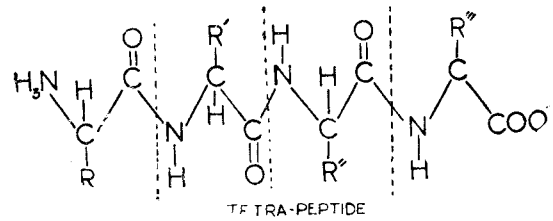
In 1902, Emil Fischer and Franz Hofmeister postulated independently the peptide bond in proteins. In their view a protein consisted of many amino-acids in which α-amino group of one amino-acid and the carboxyl group of another amino-acid is joined, with the elimination of a molecule of water, to form an amide

linkage. The product of such a condensation reaction is a "peptide" and an acid amide type of bond joining two amino-acid is a "peptide bond" or "peptide linkage".

The nature of linkage between the amino acids of protein is as follows[3], —



Moreover, the formation of a peptide bond between the two amino acids results in the loss of two charged groups, ammonium and carboxylate. The peptide structure has one hydrogen atom associated with it; it is very



weakly acidic and ionizes only under special condition such as heavy metal complexing, but it forms hydrogen bond readily, a property which has profound influence on the three dimensional structure of proteins. The existence of the peptide bond in protein is proved by :

(1) it's characteristic U.V. and I.R. absorption spectra which have firmly established that in proteins primary covalent bond is the peptide bond.

(2) the susceptibility of proteins to reagents and enzymes that attack model substances containing peptide bond.

(3) the relative abscence of free amino and carboxylic groups in proteins but their appearence upon hydrolytic cleavages by enzymes and chemical reagents; and finally.

(4) the synthesis of polypeptide using methods that definitely produce the peptide bond.

If one grants the validity of the peptide bond hypothesis, it may next be questioned about the link of amino-acid like lysine, glutamic acid and aspartic acid with the protein. The possibilities are, (a) they may be linked solely through their α-amino or α-carboxyl grops (b) the α-amino groups of the lysine or the carboxyl groups in the side chain of the glutamic acid and aspartic acid involved in the peptide linkage to other amino-acids. The available evidence indicates that, in most proteins, the the α-amino group of lysine, the guanidine group of arginine and the imidazolyl group of histidine do not participate in amide linkage with their amino acids. This conclusion[4] is based in part on the results of titration experiments and in part on the findings that these groups are available for reactions that would not be observed if the groups are substituted.

Synthesis of Peptides

Mainly there are two types of peptide synthesis,—

(a) conventional method in solutions

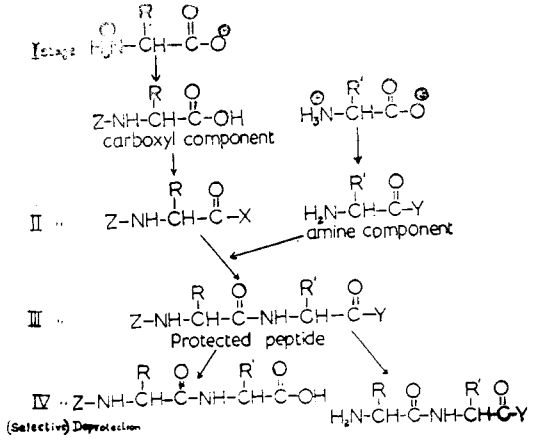
(b) solid phase synthesis

(a) Conventional method in aqueaus solution

General scheme of peptide synthesis :

Fig. 4. Z, Amino protecting group; Y, Carboxyl protecting group; X, Activating substituent.

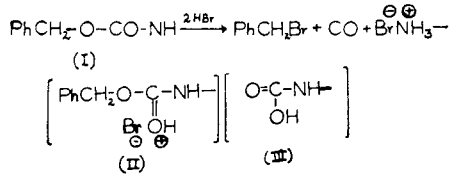
According to Hoffmann and Katsoyannis[5], the synthesis of peptides by conventional method involves the following stages—



(1) Formation of the "carboxyl component" of the peptide by blocking it's reactive amino group with some reagent Z (Fig. 5). Z may be— (a) Toluene-p-sulfonyl (Tosyl) derivatives- one drawback of this method has been the difficulty of seperating the product from sodium ions after removal of the protective group with sodium in liquid ammonia. The "tosyl" group can be removed by treatment with hydrogen bromide and phenol in acetic acid. (b) Benzyloxycarbonyl (Carbobenzoxy) derivatives— it is the most widely used amino protecting group in the conventional method of peptide synthesis. It's value has been significantly increased by the development of a convenient new method for it's removal[6].

This consists in taking the benzyloxycarbonyl derivative (I) with hydrogen bromide in anhydrous acetic acid, whereupon benzyl bromide and carbon dioxide are eliminated leaving the hydrobromide of the amino compound. Presumably the salt (II) and the derived carbamic acid are the intermediates.

One hour at 20°C with N-hydrogen bromide in acetic acid is often sufficient and reaction is correspondingly more rapid in warm solutions;



it's progress can be followed by the evolution of carbon dioxide. Under these mild conditions most other sensitive groups, such as ethyl and benzyl esters, thiol esters, amides and S-benzyl residues are unaffected. Any deesterification can be corrected by treating the crude hydrobromide with alcohol and a further portion of hydrogen bromide solution.

(e) Triphenylmethyl (Trityl) and other aralkyl derivatives.
(d) Phthaloyl Derivative.
(e) Formyl Derivatives.
(f) Trifluroacetyl Derivatives.
(g) Chloroacetyl Derivatives.
(h) Nitrophenoxyacetyl Derivatives.
(i) Phenoxycarbonyl Derivatives.
(j) Alkylthiocarbonyl Derivatives.
(k) Benzylidine Derivatives.

2. In the second stage the carbonyl group of the N-protected acid component is activated by converting it to a mixed anhydride or an active ester. Anhydrides are used at once without isolation, while many active esters can be isolated and stored as well-defined stable solids.

Amino-acids can be esterified conveniently with methyl alcohol and thionyl chloride. The free esters can be obtained from their hydrochlorides by treatment with a slight excess of ammonia in chloroform. The thionyl chlo-

ride method can also be applied to acetyl amino-acids and presumably, peptides and their amino protected derivatives. Alternately ion exchange resins or toluene-p-sulfonic acid can be used as as catalysts of esterifications.

The benzyloxycarbonyl amino-acid can be esterified with benzyl alcohol and toluene-p-sulfonic acid, and then the benzyloxycarbonyl group can be cleaved with hydrogen bromide in acetic acid.

3. Peptide bonds are next formed (3rd stage) by nucleophilic attack by the amine component at the activated carbonyl carbon in an addition-elimination reaction[7]. This may be done through—

(a) Acid hydrolysis
(b) Acid Azides :

The chief merit of the azide method is that apparently no racemization is caused by it. A side reaction, which can be significant or even overwhelming is Curtius Degradation[8] of the acyl azide to an isocyanate, which also reacts with the amino compound.

Working near O°C is likely to reduce this and other side reactions.

(c) Mixed Anhydrides with carboxylic acid.
(d) Mixed anhydrides with sulfuric acid.
(e) Mixed anhydrides with phosphoric acid.
(f) Activated esters—This mehod is similar to (c)
(g) Thiolesters
(h) Isocyanates
(i) Derivatives of phosphoric acid
(j) Derivatives of arsenious acid
(k) Carbodiimides
(l) Alkoxyacetylenes
(m) Ketenimines

4. In the fourth stage, protecting groups are removed as required. Partial selective clevage affords intermediates for further condensation. At the end of a synthesis, all remaining protective groups are removed. Most peptide synthesis are carried out in this way. Large peptides (i.e., proteins) are prepared by alternating repetations of stages 2→3 and 4.

Methods of synthesis differing from the usual pattern are—

(a) Through carboxyanhydrides
(b) Through other cyclic anhydrides
(c) Through oxazolidones
(d) Through chloroacetyl Peptides
(e) Through diketopiperazines
(f) Through derivatives of salicylic Acid.

Cyclization of peptides

Occassionally a cyclic rather than a linear peptide may be the object of synthesis. Cyclic peptides are formed from N-carboxyamino acid anhydrides under special conditions but obviously the general synthetic path must be through derivatives of one of the corresponding linear peptides. Among these are oxazolones, 2-thio-5-thiazolidones and N-carboxyanhydrides. The N-carboxyanhydrides have been useful for the preparation of polymeric peptides containing so many as 100 to 200 amino-acid residues per chain[9].

(b) Solid phase synthesis

One of the most vexing atteributes of peptide synthesis is the large number of man-hours required, mostly for workup and purification. Moreover during condensation step, there are more chances of racemization in the aqueous media, and, of course hydrolysis of the active reagent may be serious. The only method

which offers good yield without serious risk of racemization seems to be that of mixed sulfuric anhydrides. Even here both hydrolysis and, to a less extent, racemization are encountered but the process is reasonably satisfactory.

All the above drawbacks of the aqueous phase synthesis of peptides, prompted Merrifield to design a procedure roundabout 1964. which could be automated.

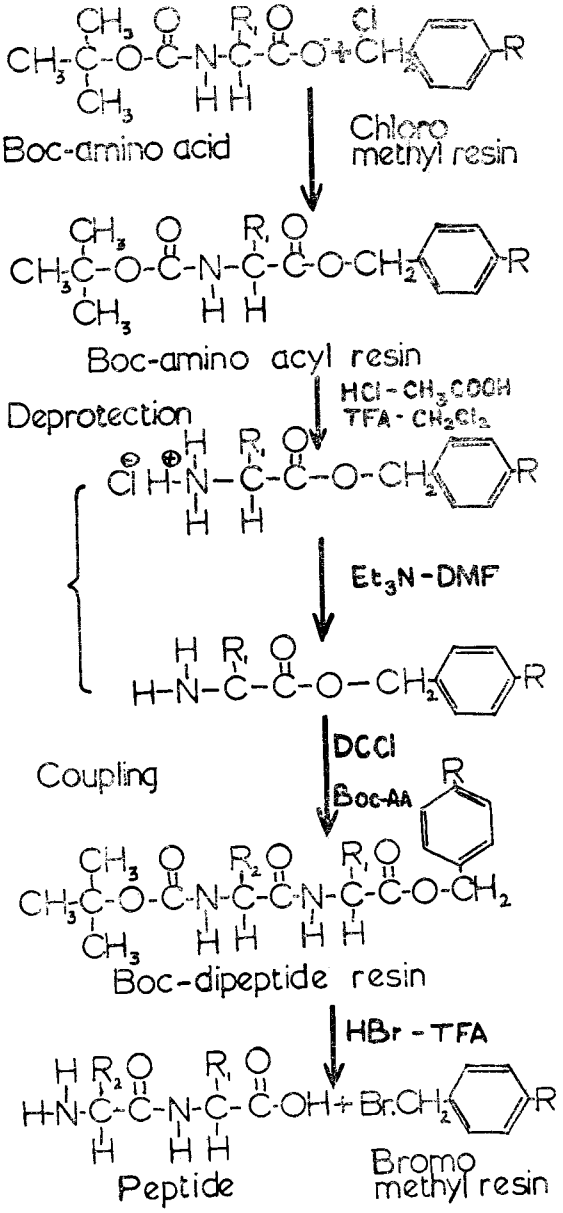
The standard solid phase procedure is schematically depicted in Fig. 6.

Scheme of solid phase peptide synthesis

The solid support of choice is beaded chloromethylated styrene-1% divinylbenzyl copolymer[8], which can form a benzyl ester bond with the first (C-terminal) amino-acid. The tert-butyloxycarbonyl (t-BOC) group[8,10] is mainly used for the protection of the α-amino function. t-BOC-amino acids are coupled to to the resin-bound amino acids or peptides by dicyclohexylcarbodiimide[11] within 2 to 6 hrs. period, 3 to 5 molar excesses to drive reactions towards completion. Clevage of the BOC group by 30 to 50% trifluoroacetic acid (TFA) in methylene chloride or by HCl in acetic acid or dioxane and treatment with triethylamine prepares the peptide-resin for condensation with the next BOC amino-acid. All operations are carried out in the same reaction vessel, where the peptide-resin remains throughout the synthesis. Solvents and reagent solutions are introduced in timed sequence and removed through a filter device by suction or pressure. Peptides are taken out from the resins by breaking the bond.

Comparison of the two methods

The aqueous phase synthesis employs amino protecting residues, the most useful of which are benzyloxycarbonyl, tosyl and trityl. Tosyl derivatives often crystalize more easily than do the benzyloxycarbonyl anaogs and they are more stable. On the other hand, removal



of the tosyl group is rather more troublesome. Probably the chief advantage of the trityl group is it's removal by exceptionally mild

hydrolysis. The phthaloyl group is convenient for dipeptide synthesis, but it has been adversely criticized[12]. It seems easier to manipulate intermediates containing a group like trityl.

For condensation the most tested route is through acyl azides. It is noteworthy that apparently it does not cause racemization during the condensation of acetyl-L-leucine with glycine ester. Another synthesis which seems to avoid racemization completely is that through mixed sulfuric anhydrides with coupling in anhydrous dimethylformamide. The carbodiimide method offers the most direct and simplest condensation.

The synthetic steps would be saved if the amino components employed in the condensation were the free peptides or amino-acids instead of their esters. As to the esters themselves, it will be noted that, for good reasons, benzyl esters have become more popular. In this connection the aryl esters, which are more easily hydrolysed may well prove valuable.

The classic combination of protection by benzyloxycarbonyl and condensation through azides, which was favoured by Bergmann in 1932, has withstood the tests of time and competition.

Some trouble may arise because of the branched side chain of the amino acids which hinders the approach of the reagents toward both the carbonyl and amino groups. The hindrance may be overcome in an extended period of reaction. More serious problems are caused by reactive groups in the side chain.

The most attractive features of the solid phase procedure are, (i) high speed of reaction,-permitting upto six amino acid residues to be coupled per day (ii) insolubility problems that hamper conventional synthesis of larger peptides are eliminated (iii) the procedure is easy and convenient to perform (iv) mechanization and automation allow round the clock operation.

SYNTHETIC ACHIEVEMENTS

Interest in the development of new methods of peptide synthesis received considerable impetus during the period 1940-1955, when a large number of peptides were found in the nature.

In 1921, Hopkins isolated from yeast glutathione, later found to be widely distributed in the animal and plant cells.

One very important factor in the rise of interest in peptides was the discovery that various strains of microorganisms elaborate peptides that have antibacterial activity towards other micro-organisms. Among the first of these was a material named "Gramicidin" obtained from *Bacillus brevis* by Dubos and Hotchkiss.

A few groups of peptides have attracted particular interest because of their biological activity or because they represent interesting directions of peptide chemistry or bio-chemistry. For example, there are recently discovered biological triggers, such as the hypothalamic hormone releasing factors, i.e., β-corticotropin, [D Ser^1 N Lu^4] β-Corticotropin-(1-25)-pentacosapetide amide, [D Ser^4, Lys^{19},]-β-Corticotropin-(1-18) octadecapeptide amide, [β-Ala^4, Lys^{17}]-β-Corticotropin (1-17)-(4-amino-n-butylamide). These exert their effects in nanomolar quantities, and therefore occur in the gland in such minute amounts that synthesis had to be employed to determine the exact structures. Hopes for fertility control by inhibiting the lutenizing hormone follicle stimulating hormone-

releasing hormone have in many laboratories prompted a dramatic surge of analog synthesis in search for an inhibitor that would block the fertility chain.

Calcitonin are hormones which control serum calcium. These molecules have been synthesized by conventional solution methods. The tissue hormone, angiotensin II, which increases blood pressure and causes hypertension was isolated, analyzed and synthesized almost 25 years ago. An incidental discovery during the synthesis of tetragastrin, showed that H-Asp-Phe-OCH_3 is 100-200 fold sweeter than sucrose, posseses a good flavour without unpleasant aftertaste. The potential importance of this sweet peptide for weight watchers and diabetics is obvious.

Oxytocin was synthesized by du Vigneand *et. al.* Oxypressin is a cyclic peptide which is a combination of the cyclic penta peptide from vasopressin and the tripeptide side chain from oxytocin. It posseses oxytocic and pressor biological activities.

Table I shows the biologically active peptides synthesized by the conventional method in solution and Table II by the Merrifield solid phase method.

Table I

Peptide	Size	Year
Ribonuclease S-protein	104	1969
Proinsulin-(17, 18), Porcine	67	1972
Ribonuclease T_1 (1-47) and	47	1972
-(48-80)	33	1972
ACTH	39	
Porcine		1963
Human		1967,'72
Calcitonin	32	
Porcine		1968
Human		1968,'70
Salmon		1969
Insulin	30,21	1963-'65
Glucagon	29	1967
Secretin	27	1966
Gastric inhibitory polypeptide	27	1972
Gastrin	17	1964

Table II

Human growth hormone	188	1970
Ribonuclease A	124	1969
Cytochrome C	104	1969
Ribonuclease T_1	104	1972
Acyl carrier protein-(1-74)	74	1971
Cobrotoxin	62	1972
bovine basic trypsin inhibitor	58	1971
Ferredoxin	55	1968
Staphylococcal nuclease-(6-47)	42	1969
Parathyroid hormon-(1-34)	34	1971
Insulin	21,30	1966
Adrenocorticotropin-(1-19)	19	1972
Gramicidin A,B,C	15	1972
α-Melanotropin	13	1971
Antamanide	10	1969
Gramicidin-S	10	1971

SOME NATURAL PRODUCTS AND DERIVATIVES

Oxytocin Cys-Tyr-I leu-Glu-Asp-Cys-
NH_2 NH_2
Pro-Leu-Gly

Arginine Cys-Tyr-Phe-Glu-Asp-Cys-
Vasopressin NH_2 NH_2
Pro-Arg-Gly

Lysine Cys-Tyr-Phe-Glu-Asp-Cus
Vasopressin NH_2 NH_2
Pro-Lys-Gly

Gami- Val-Orn-Leu-D-Phe-Pro-Val-
cidin-S HCl Orn-Leu-D-Phe-Pro
Tos HCl
Tos

IMPORTANCE OF PEPTIDE SYNTHESIS

Present age demands more of peptide synthesis due to the following reasons :-

(1) **Economic Necessities** : For many years production of oxytocin has been more economic by synthesis than by isolation. Synthetic ACTH has been economically a profitable product. There is every possibility, that in future there will be shortage of natural insulin. So it is high time to make regular production of synthetic insulin.

(2) **Modification of Pharmacological Response** : Analogs possessing enhanced potency, protracted action or inhibitory effects, highly active ACTH fragment analogs, and angiotensin II inhibitors, are being sought for their therapeutic potential.

(3) **Structure-Activity Correlation** : Hundreds of analogs of biologically active peptides have been synthesized to determine which structural parameters might cause the biological effect elicited by natural molecules. Spectral methods have allowed structure-activity conformation correlation.

(4) **Structure Determination synthesis** : Thyrotropin releasing factor[13,14] and bradykinin are examples, where synthesis came to the aid of the structure analyst.

(5) **Proof of structrue** : Hormones (pituitary, gastrointestinal and others such as oxytocin, ACTH, gastrin and calcitonin etc.) and kinins, (angiotensin) have been synthesized and shown to be identical with the natural products.

The Centennial of the first protein synthesis is nearly upon us[15], while modern peptide synthesis can be dated from du Vigneand's (1953) preparation of oxytocin[16]. Now the solid phase procedure of Merrifield disclosed in 1963, appeals by virtue of it's speed and convenience. It has been already used to synthesize enzymes. Nevertheless, there are still significant experimental problems in the routine synthesis of proteins.

There are three hundred antibiotics and microbial products which are peptide in nature. Their importance in drug research is quite obvious. It is now essential to lay emphasis on the commercial production of more and more peptides of biological and clinical interest as they are likely to be the only drugs of choice for the future. Since a number of drugs, presently in use, are being gradually discarded due to the various side effects and toxic manifestations associated with them, so it is quite necessary that the synthetic competence and expertize for the production of peptides is developed in our country to meet the demand of the future. A concerted effort in this direction has been made already by our peptide chemists in the Central Drug Research Institute, Lucknow; the All India Institute of Medical Sciences, Delhi; University of Bangalore; Wellcome Research Institute, Vellore; and there are many other institutions. CDRI during the past few years has designed and synthesized a number of biologically active peptides. Some have been obtained, which indicate that it may also be possible to find simple peptide substitutes for the intricate and complex molecules of nature such as enzymes and hormones.

The time is now ripe for the Indian Pharmaceutical Industry to exploit this new line of biologically active substances and to find out the commercial feasibility of this venture. If it is successful, then the undesirable dependence on the import of these drugs of future will be completely avoided.

References

1. Vickery, H.B., Yale, J., Biol and Med. **22**, 387 (1950)
2. Awapara, J., Introduction to Biological Chemistry, Prentice Hall of India Pvt. Ltd. (1974)
3. West, E.S., Todd, W.R., *et al.*, Text Book of Biochemistry, The Macmillan Company, 1963.
4. Putman, F.W., Neurath, H., and Bailey, K., The Proteins Vol. I B, Academic Press, N. Y., 1953
5. Hofamann, K., and Katsoyannis, P.G., The Proteins, 2nd Ed., Neurath, H. (Ed.), Academic Press, N. Y., 1963.
6. Bailey, K., *et al,* Advances in Protein Chemistry, Vol. XII, Anfinsen, C.B. Jr. (Ed.) Academic Press. N. Y., 1957.
7. Boclanszky, M. and Ondetti, M.A., Peptide Synthesis, Interscience Publ., N.Y., 1966.
8. Gutte, B., and Merrifield, R.B., J. Amer. Chem. Soc., **91**, 501, (1969).
9. Anderson, G.N., and Mc Gregor, A.C., *ibid*, **97**, 6180 (1957).
10. Mckey, F.C., and Albertson, N.F., *idib*, **97**, 4686 (1957).
11. Sheeham, J.C., and Hess, G.P., ibid, **77**, 1057 (1955).
12. Leach, S.J., and Lindley, H., Australian J. Chem., **7**, 173 (1954).
13. Boler, J., Enzemann, F., Folkner, K., *et al,* Biochem. Biophys., Res. Communs, **37**, 705, (1959).
14. Burgus, R., Dunn, T.F., *et al.*, Endocronology, **86**, 573, (1970).
15. Curtius, T.J., Prakt. Chem., **26**, 175 (1982).
16. du Vigneand, V., Ressler, C.S., Roberts, J.M., *et al.*, J. Am. Chem. Soc., **75**, 879 (1953).