# REVIEW
# Protein engineering by chemical means?

R.E.Offord

Département de Biochimie Médicale, Université de Genève, 1211 Genève 4, Switzerland

## Introduction

Chemical methods exist that produce modified analogues of natural proteins, precisely of the type that are made by recombinant techniques: are these chemical approaches entitled to be called 'protein engineering'? It seems important to pose this question but, curiously, less important to give a definitive answer to it. It is probably enough to note that the journal *Protein Engineering* has sought to have protein-chemical methods represented on its Editorial Board, and that protein chemists working toward the same goals as those using recombinant techniques are submitting papers to it in reasonable numbers. This seems to indicate that whatever the final decision on terminology, should there ever be one, the community of interest is already recognized.

A more interesting question is: what decides whether the chemical or the recombinant approach is best, and (most interesting of all) what can the two approaches do together that neither can do so well on their own? The present review addresses these points, principally by reference to a few representative papers published in the last three years, the period during which recombinant methods became comparable in their power with chemical ones.

The choice of a paper for inclusion was governed by the wish to be able to illustrate validly and concisely the principal experimental issues involved. Thus, omission of a paper implies no judgement whatever as to either priority or absolute value of the research described.

## The objective

Modern biology and medicine have produced a great demand for proteins having novel structures, not found in nature. Many of the purposes for which this is done have been well discussed, in so far as they affect recombinant preparations, in a previous review in this journal (Leatherbarrow and Fersht, 1986), and the same ideas apply when it is chemical methods that are contemplated. Analogues are invaluable for study of structure–activity relationships, and where questions arise of the importance of non-covalent interactions in protein structure and function. Furthermore, existing knowledge of the underlying principles of protein action makes it possible to design analogues with tailored functional properties. Additionally, the study of a protein in its true biological setting (organ, tissue or cell) often calls for an analogue modified by isotopic substitution, or by a fluorescent chromophore, or by some magnetic-resonance reporter group.

## The choice of methods

One does not make an analogue simply to be able to say one has done so: it is necessary to be able to make the wanted products in acceptable quantity and quality for use in the proposed experiment. The preparation should consume as little time and resources as possible.

When the production of the analogue involves no more than substitutions between coded amino acids, it should now be assumed that this will normally be done by recombinant methods or (an approach that sometimes seems in danger of being forgotten) traditional mutagenesis and selection techniques. On the other hand some very interesting substitutions involve artificial, non-coded amino acids, or their analogues. Although it is possible to envisage ways in which recombinant methods could be made to work even here (e.g. see Offord, 1980), they are not at present a particularly practical proposition. In addition to substitutions that go outside the genetic code, there is a further area in which chemical methods are strong where recombinant methods are sometimes weak; the production of relatively short polypeptides. Also, certain substitutions of coded amino acids are now so simple by chemical methods that they represent a route equally as attractive as recombinant methods.

It should, for example, be remembered that human insulin, the production of which was seen as the first real test of recombinant methodology, was in fact first brought to the market in any quantity by the use of semisynthetic, i.e. chemical, methods. This was not a question of historical inertia, since the development of the semisynthetic methods involved did not begin in any serious way until after the start of the recombinant work. Inspection of the recent literature (see below) suggests that this is by no means an isolated example of the way in which chemical methods will continue to complement recombinant ones.

## Chemical approaches

The choice lies between total synthesis, semisynthesis and the chemical modification of side chains.

### Total synthesis

Total synthesis, as has already been said, is the method of choice for small polypeptides. One of the two main variants, the so-called stepwise or solid-phase method, has been so greatly improved in recent years that it is now within the reach of any molecular biology laboratory that can afford the equipment. On the other hand, only the most expert laboratory could contemplate undertaking the alternative route, known as fragment condensation. Fragment condensation involves the classical organic-chemical synthesis of the peptide by the condensation in free solution, first of the amino acids to form a number of longer fragments of the wanted sequence, and then the successive condensation of the resulting fragments to form ever larger stretches of the target sequence until the complete molecule is obtained.

The solid-phase method, as is widely known, proceeds by anchoring the C-terminal residue to a resin. The amino acids are added in successive cycles of reaction, in reverse order to the normal direction for reading a sequence. The details of how this is done vary from one laboratory to another, but all variants of the technique have in common the need to ensure that the efficiency of incorporation of the new amino acid (the 'coupling

efficiency') is as near 100% as possible. For a target product of $n$ residues, and assuming a constant coupling efficiency of $r$% per cycle, the proportion of molecules of the wanted product in the final mixture is $(r/100)^{n-1} \times 100\%$. If $r$ is 98% (already a relatively impressive figure by the standards of general organic synthesis), then the proportion of correct molecules in a product of 100 residues is 13% and that in a product of 300 residues is 0.24%. It suffices for the average efficiency per cycle to fall to 95% for these two numbers to become 0.62% (100 residues) and 0.00002% (300 residues). Much attention has therefore been devoted to keeping the coupling efficiency as high as possible. Other recent developments have to do with a different choice of protecting groups and solid-phase resins from that made when the method was originally proposed.

Given access to one of the modern automated machines exploiting this chemistry, few workers should hesitate in synthesising a product of, say, 30 residues. For most syntheses, it is hard to imagine any other method, chemical or recombinant, being able to compete. Most non-coded amino acids (including the D-amino acids) can be incorporated at will by this method.

Literally thousands of analogues of polypeptides have been made by one or other variant of these methods; they have provided much of our current understanding of the biology of shorter peptides, and this situation is unlikely to change in the near future. An overview of what is possible can be obtained by reading any of the recent volumes of either the European or the American Peptide Symposia (e.g. Ragnarsson, 1984; Deber et al., 1985).

### Side-chain modification

Space does not permit a discussion of the other major method of protein alteration, side-chain modification without peptide-bond cleavage or synthesis. None the less, this approach, to which we owe much of our present understanding of structure—activity relationships, will continue to give us much information, particularly if one can attempt the specific conversion of just one side-chain, rather than affecting simultaneously many side-chains of the same type. While mentioning the classical experiment of conversion of a serine protease to a cysteine one (Neet and Koshland, 1966; Polgar and Bender, 1966) we should not forget the remarkable advance in specific side-chain conversion reported by Kaiser and Lawrence (1984). As an illustration of a general approach to the production of semisynthetic 'chemical mutations' at the catalytic centres of enzymes, these authors converted papain into a highly efficient oxido-reductase, by means of a series of reactions designed in the light of existing knowledge of structure—activity relationships.

### semisynthesis

This approach combines that of total chemical synthesis with more biologically oriented methods. In it, fragments of naturally occurring proteins are used as ready-made intermediates for the chemical synthesis of proteins. Only the part of the molecule that is to be changed (often just a single residue) need be synthetic In principle, this method has the considerable advantage over total synthesis that most analogues currently required are 99% or more of native structure. The synthesis of these portions can be left to the biosynthetic machinery while one concentrates on the 1% or less into which one wishes to introduce a change. The semisynthetic approach has the advantage over recombinant methods that it is just as easy to incorporate non-coded amino acids as it is to incorporate coded ones. Its limitation as a method derives from the necessity of finding a degradation method (enzyme or chemical reagent) that not only cleaves the protein into a reasonably small number of fragments, but also exposes the site that it is wished to modify. A few of the same problems of side-chain protection arise that are encountered with total synthesis. However, the use of proteolytic enzymes working in reverse as coupling agents (see below) has greatly simplified this problem. The methodology of the several variants of protein semisynthesis has been extensively reviewed (e.g. Offord, 1980; Sheppard, 1979; Chaiken, 1981).

If total synthesis helped us learn most of what we know about smaller peptides, it seems to be the message of the tables and the papers cited below that semisynthesis has contributed a considerable amount to work on larger molecules, where few if any totally synthetic analogues are available. There is no sign of these developments slackening off.

Table I summarizes the principal types of protein semisynthesis. Because of limitation of space, the rest of this review will concentrate principally on semisynthesis via the formation of peptide bonds.

The main experimental requirements that confront us are: (i) a means of efficient cleavage at a few specific sites; (ii) a means for separating the fragments; (iii) a means of ensuring that the fragments, when coupled back together, will combine only through their $\alpha$ carboxyl and amino groups; (iv) a set of reagents and procedures that enable all steps to be carried out under conditions mild enough to avoid permanently compromising the chemical and three-dimensional structure of the product.

The first two of these requirements have been met simply by drawing on the techniques developed by workers in the field of protein sequencing. The third requirement (to ensure coupling through the correct groups) can be met by careful selection from

**Table I.** The principal kinds of protein semisynthesis

| Type | Description | Applied to. |
|---|---|---|
| 1. Non-covalent | The formation of functioning complexes of protein fragments One or more of the fragments can be replaced by a synthetic or modified version | Pancreatic ribonuclease[1], S aureus nuclease[2], cytochrome c[3], thioredoxin[4], somatotropin[5], phospholipase $A_2$[6], myoglobin[7], IgA fragment $F_V$[8] |
| 2 Covalent | | |
| (i) via $-S-S-$ bridges | The reoxidation of reduced chains of a multichain protein, one of the chains having been replaced by a synthetic or modified version | Insulin[9], antibodies[10] |
| (ii) via the peptide bond | The chemical coupling, via alpha amino and carboxyl groups, of fragments of natural sequence with those of altered sequence. The altered sequences can be obtained by total synthesis or by semisynthesis, or by any other method | See Tables II and III |

References (including reviews). 1, Richards (1958), 2, Taniuchi et al (1967), 3, Harris and Offord (1977); 4, Slaby and Holmgren (1975), 5, Li et al (1978); 6, Kihara et al (1981), 7, Hagenmaier et al (1978), 8, Gavish et al (1978), 9, Zahn et al (1981); 10, Rosenberg and Terry (1977)

the available techniques of the classical methodology of protein synthesis to bring about differential protection of side-chains. Alternatively, correct coupling can now be ensured with little or no protection by using the specificity of proteolytic enzymes acting in reverse (Laskowski, 1978; Inouye et al., 1979; Markussen, 1981). This development, which had been impatiently awaited for many years, is not restricted to small substrates: it has even been applied with success to human growth hormone (Graf and Li, 1981) and other protein examples are given below.

This advance has completely revolutionized the field: many of the more difficult protection strategies used until recently (Offord, 1980) are now needed only under very special circumstances. Ligation by proteases is one of the major factors responsible for semisynthesis having remained competitive in some fields with recombinant methods, even when substitutions by coded amino acids are involved.

The power of the ligation method has recently been improved by using proteases to couple, to the C termini of appropriate fragments, not peptides, but aminoacyl active esters (Rose et al., 1987). Remarkably enough, conditions can be found under which some aminoacyl active esters neither polymerize through their own free $\alpha$-amino groups, nor acylate the side-chain amino groups of the substrate or the ligating protease. Polypeptide active esters resulting from this operation have proved particularly useful in ligating the nicked peptide bonds in non-covalent complexes of fragments of cytochrome $c$ (Wallace and Proudfoot, 1987; K. Rose, private communication; and see below).

The fourth requirement (mild reagents and solvents) was met by trial and error. We now know that many chemical substances that would have been expected to destroy protein activity will not do so, whilst certain others, apparently much milder, are fatal in their effects (see, for example, Offord, 1980).

That the origins of the method lie in primary structure determinations becomes obvious when we look a little closer. Thus, fragment-condensation semisynthesis combines Sanger's original strategy of sequencing (with the additional option of CNBr cleavage) with that of classical solution synthesis. Stepwise semisynthesis combines the strategy of Edman (or that of sequencing by carboxypeptidase digestion) with that of conventional stepwise synthesis.

Tables II and III give a general summary of the fields to which the two methods have been applied, together with selected references to earlier reports and reviews. semisynthetic work published in the last three years is reviewed in the text, according to the principles of selection stated above.

## Recent advances in semisynthesis

*Insulin.* Insulin has for a long time been an excellent example of the fruitful combination of knowledge obtained by study of analogues and derivatives obtained by all possible methods. Extensive studies have been carried out on hundreds of analogues, both synthetic and semisynthetic, on side-chain derivatives, and on evolutionary variants. These results have combined with data from X-ray and other physical methods to tell us a great deal about the structural basis for the action of this hormone, and about its co-operativity. semisynthetic analogues, which are very easy to make by what are practically standardized methods, continue to play their role. Only representative examples can be given here, restricted to the last three years.

Fragment-condensation semisynthesis has been used to probe the role of the region B24-B25-B26. Kobayashi et al. (1984) studied [Ser$^{B24}$]insulin and [Ser$^{B25}$]insulin in terms of receptor binding, negative co-operativity and immunoreactivity. Shoelson et al. (1984) tested the biological activity and pharmacokinetic properties of semisynthetic [Ser$^{B24}$]insulin in dogs. They discuss the abnormal clearance of the analogue *in vivo*, and the causes of hyperinsulinaemia in both diabetic and non-diabetic individuals who secrete this abnormal insulin. Haneda et al. (1985), continuing the studies of Kobayashi et al. (1984), carried out further tests on the potency of the same two serine analogues, which they extended to the investigation of the rates of their degradation. Degradation is slower than for the native sequence, and the

**Table II.** Some earlier examples of fragment-condensation syntheses

| Protein | Principal operations carried out |
|---|---|
| Insulin | Enzymic and chemical recoupling of sequences from the last eight residues of the B-chain (residues important in receptor binding, residues implicated in activity by studies on mutant human insulins)[11,12,13] |
| Proinsulin | Modified open-chain form by combining natural chains with a short synthetic bridge[14] |
| Ferredoxin | Preparation of analogue free of aromatic residues for functional and physical studies of iron–sulphur clusters[11,12,14] |
| Nucleases | Production of analogues of intact protein by chemical or enzymic reforming of nicked peptide bond in non-covalent complexes[14] |
| Cytochrome c | Recombination of CNBr fragments Several analogues made for activity, immunological and physical studies (e g Tyr$^{67}$ changed to p-fluorophenylalanine). Side-chain-specific reactions on separated fragments before reconstructing the protein (gives proteins modified on only some of the side-chains susceptible to the reagent in question)[11,12,15] |
| Myoglobin | Coupling of two synthetic and one natural fragment to give the whole sequence[16] |
| Protease inhibitors | Insertion of new residues at the protease-sensitive site Includes the production by design of new biological specificities[11,14] |
| Phospholipase A₂ | Coupling of short peptides to a large natural fragment Many analogues for activity and physical studies including substitution by nor-leucine[11,12,17] |

References (including reviews) 11, Offord (1980), 12, Sheppard (1979), 13, Riemen et al. (1983), 14, Chaiken (1981), 15, Wallace and Rose (1983); 16, Simmerman et al. (1982), 17, van Scharrenburg et al (1983, 1984)

**Table III.** Some earlier examples of stepwise syntheses

| Protein | Principal operations carried out |
|---|---|
| Insulin | Very many analogues with normal amino acids In addition, introduction of $^{13}$C, of D-amino acids in sites critical for activity, selective tritiation, selective iodination. Conversion of porcine insulin into human insulin[11,12] |
| Ferredoxin | Numerous analogues for activity and physical studies of the iron–sulphur clusters[11,12] |
| Phospholipase A₂ | Many substitutions for physical and activity studies. Introduction of $\beta$-alanine and $^{13}$C amino acids (both D and L)[17,18] |
| Glucagon | Substitutions for physical and activity studies[19] |

References (including reviews): 18, Offord and Di Bello (1978); 19, Flanders et al (1984) Other refs, see footnote to Table II.

authors discuss the pharmacokinetic implications of this finding. These analogues could also have been made by recombinant methods: in the event it proved more convenient to do so by semisynthesis.

Nakagawa and Tager (1986) have described a magnificent series of analogues in this same region of the molecule, many of which could not have been made by existing recombinant techniques. These consist of: (i) forms of the hormone in which 5, 6 or 7 residues have been deleted from the C terminus of the B-chain; (ii) analogues in which $Phe^{B25}$ has been replaced by natural or unnatural amino acid residues; and (iii) analogues lacking the C-terminal five residues of the B-chain and in which $Phe^{B25}$ has been replaced by $\alpha$-carboxaminated amino acids. Among the unnatural amino acids used were naphyl-(1)-alanine, naphyl-(2)-alanine, and homophenylalanine. These were of particular interest because of the different bulk of their side-chain relative to that of phenylalanine itself. Among the many conclusions drawn from the experiments carried out on these analogues there is only room to mention here the proposal that steric hindrance involving the C-terminal domain of the B-chain plays a major role in directing the interaction of the hormone with its receptor.

Degradation of insulin at the target cell is an important adjunct to the expression of the biological activity of the hormone. In the author's laboratory we have produced a number of mono-isotypically labelled insulins in order to be able to characterize better the products of insulin degradation under various conditions: [$^3$H-Phe$^{B1}$]insulin (Halban and Offord, 1975); [$^3$H-Gly$^{A1}$]insulin (Davies and Offord, 1985); [$^{18}$O-B$^{29}$]insulin (Rose et al., 1984); [$^2$H-Gly$^{A1}$]insulin (L.A.Savoy and K.Rose, unpublished results); and [$^3$H-Ala$^{B30}$]insulin (C.Bradshaw, J.G.Davies, R.E.Offord and K.Rose, unpublished results). These analogues have been shown to behave authentically in a variety of systems. The degradation products liberated from them under various conditions have been identified either by radiochemical sequencing (e.g. Muir et al., 1986; Davies et al., 1986) or mass spectrometry (Rose et al., 1984; K.Rose and L.A.Savoy, unpublished data). It has, for example, been possible to identify with some confidence the sites of cleavage of the hormone by the enzyme insulin proteinase. This enzyme may be responsible for insulin degradation in intact tissues (Duckworth and Kitabchi, 1981). The results obtained using the semisynthetic analogues, when taken into account with those from studies in cells, organs and the whole animal, do indeed offer some support for the physiological significance of this enzyme. The possible secondary biological activities of the fragments liberated by the enzyme (which as a result of the above work have been well characterized as regards their structure and purification) therefore merit further study.

The 50 or so papers devoted to the clinical behaviour of semisynthetic human insulin are outside the scope of this review.

*Cytochrome c.* This protein lends itself particularly well to the fragment-condensation version of the semisynthetic method. Two of the most reliable cleavage agents, trypsin and CNBr, have their points of attack at convenient positions. Additionally, several instances are known of the non-covalent association of large fragments to form active complexes.

The earlier work having already been reviewed (Offord, 1983, and the references cited therein), only more recent advances will be mentioned.

Ten Kortenaar et al. (1985) have used the fragment-condensation approach to prepare a number of analogues, including

[fluoroPhe$^{67}$]cytochrome c, [Val$^{78}$]cytochrome c, [Leu$^{82}$]cytochrome c and [Tyr$^{97}$]cytochrome c. Following the earlier work of this group on n.m.r. of such analogues, the paper in question concentrates principally on the measurements of the redox potentials of the analogues produced, and the implications of the results obtained for the mechanism of electron transfer by this protein.

In the author's own laboratory, the most recent of the many publications of C.J.A.Wallace and co-workers on the subject (Wallace and Proudfoot, 1987) reports the preparation of a cytochrome c analogue from which the so-called 'bottom loop' of the native structure had been excised. As pointed out by Leszczynski and Rose (1986), this loop (residues 40−55) is one of the principal differences between mitochondrial cytochrome c and the bacterial cytochrome $c_{555}$, and may account for the differences in functional properties between the two proteins. The loop is a relatively self-contained structure, with its N-terminal and C-terminal portions in close proximity. Wallace and Proudfood (1987) were therefore easily able to produce a semisynthetic cytochrome c analogue that lacked it, chemically coupling the appropriate fragments of the native protein by means of the enzyme-directed, active-ester method mentioned previously. They have characterized the biological properties of their product and have commented on the original hypothesis of Leszczynski and Rose.

Proudfoot et al. (1986) have prepared a semisynthetic analogue of fully acetimidylated horse cytochrome c in which the peptide chain is nicked between residues Gly$^{37}$ and Arg$^{38}$. In contrast with a complex previously described (Harris and Offord, 1977), in which the break in continuity was between residues Arg$^{38}$ and Lys$^{39}$, the new analogue had a nearly normal redox potential and could more fully restore succinate oxidation to mitochondria depleted of cytochrome c. Studies of this and other semisynthetic analogues led the authors to propose a structural reason for the lower biological activity of the earlier complex, and they put forward an explanation for the evolutionary invariance of Arg$^{38}$.

Wallace et al. (1986) have made and characterized some semisynthetic chimeras of cytochrome c (horse−yeast, cow−yeast). Their results permit a number of deductions about the nature of the haem crevice.

Wallace and Corthésy (1986) have replaced residue Glu$^{66}$ of horse cytochrome c with nor-valine, glutamine and lysine. As a necessary control they put back by the same means the authentic residue, glutamic acid. Residue Glu$^{66}$, which is quite strongly conserved in cytochrome c, forms part of a cluster of acidic residues that occurs in all cytochromes c but whose function is obscure. Comparative studies of the analogues mentioned above suggest that Glu$^{66}$ contributes significantly to the stabilization of the active conformation of the protein, probably by salt-bridge formation, and it appears to participate in the redox-state-dependent binding site for ATP. The results support the idea that the charged surface residues in cytochrome c influence redox potential through the field intensity experienced at the redox centre, while they affect the kinetics of electron transfer through the dipole moment that they generate.

Veloso et al. (1984) wished to investigate the part of the apocytochrome c molecule that is recognized by cytochrome c synthetase. They prepared, by solid-phase synthesis, tritium-labelled fragment (1−25), as well as the same fragment in which the two cysteine residues had been replaced by glycine. On incubation of these fragments with the enzyme, and with hemin, the authentic sequence, but not the analogue, combined with the hemin and cytochrome fragment (23−104) to give a semisynthetic non-

covalent complex of the cytochrome. This product had properties practically indistinguishable from that of the complex of entirely natural origin. The authors draw from their observations certain conclusions concerning the recognition of the apocytochrome by the enzyme.

*Myoglobin.* Continuing their series of myoglobin analogues modified at or near the amino terminus, Busch *et al.* (1985) have used stepwise methods to prepare a number of semisynthetic analogues enriched with $^{13}$C at position 1. The replacements include $^{13}$C Gly, Ala, Val, Leu and Ile. After extensive characterization, the analogues were used in a number of n.m.r. experiments. The results elucidated the nature of the interaction between the side-chain of residue 1 with the invariant hydrophobic cluster Leu$^2$, Leu$^{137}$, and the effect that this cluster exerts on the hydration volume available to the alpha amino group.

*Haemoglobin S.* Seetharam and Acharya (1986) have found a fragment-condensation method for resealing the nick in a non-covlent complex of fragments of the α-subunit of haemoglobin S. The way is now open for the production of analogues of the α-chain.

*Precursor of oxytocin-bovine neurophysin I.* A semisynthetic precursor, of 105 residues in length, was obtained (Kanmera and Chaiken, 1985) by coupling a synthetic oxytocinyl peptide to native neurophysin. The precursor contained the full polypeptide sequence deduced from the corresponding DNA except for omission of the C-terminal histidine residue. Kanmera and Chaiken studied the binding of their product to various immobilized ligands, as well as the resistance of the disulphide bridge to reduction. The results led them to deduce that the precursor self-associates, and that the ligand-binding site of the neurophysin domain is occupied intramolecularly by the hormone domain. They conclude that both the self-association surface and hormone-binding site are established during the initial folding-up of the precursor, and that the general structure of the precursor is established before enzymic processing liberates mature hormone and neurophysin.

*Glucagon.* semisynthetic work continues on this peptide which, in spite of its small size, tends to be intractable by virtue of its unfavourable solubility properties. Flanders *et al.* (1984) have reported the production of stepwise semisynthesis of [Phe$^1$] and [Ala$^1$] analogues. Studies on these analogues, taken together with those on the *des*-His$^1$ glucagon produced *en route*, indicate that the α-amino group is involved more in signal transduction than in binding, but that the aromatic nature of the side-chain of His$^1$, and its hydrogen-bonding capacity, are important for both binding and transduction. In addition, circular dichroism studies gave evidence for unexpected conformational changes throughout the structure of the analogues, possibly having relevance for function−activity relationships.

*Nucleases and pancreatic phospholipase A$_2$.* Like some of the other proteins mentioned, these enzymes have a long history of intensive study by all available means; only recent, representative work is mentioned here.

Sasakin *et al.* (1985) and Stern and Doscher (1986) have studied structure−function relationships in bovine pancreatic ribonuclease, the latter publication demonstrating conclusively a role for Asp$^{121}$ in the catalytic site. Taylor *et al.* (1985) designed an analogue of ribonuclease S on the basis of existing knowledge of structure−activity relationships, and then prepared it by the non-covalent variant of semisynthesis. The analogue had the predicted properties.

The group of Slotboom and de Haas have for a long time used semisynthetic analogues of pancreatic phospholipase A$_2$ to study various aspects of structure−activity relationships, including the differential response of the enzyme to the various types of lipid−water interface that it encounters, an essential feature of the control of its activity (see references in Tables II and III). To this end, they have introduced, as well as coded amino acids, many non-coded and isotopically substituted residues. Their more recent work is described in van Scharrenburg *et al.* (1984) and Jain *et al.* (1986).

*Antibodies.* Chang *et al.* (1985) have made some interesting observations on the selectivity of proteolytic cleavage of antibody light chains at the regions joining variable with constant regions, which they propose to use for chemical semisynthesis of human−mouse light-chain chimeras. If this proves easy to do, it might constitute a useful adjunct to recombinant techniques.

*Protease inhibitors.* In a recent study, Tschesche *et al.* (1987) prepared and characterized 10 analogues of aprotinin, the Kunitz trypsin inhibitor from bovine mast cells. As is well known, this inhibitor functions by engaging the active site of the protease in an abortive (or extremely slow) cleavage of a particular peptide bond (Tschesche and Kupfer, 1976; cf. Laskowski and Kato, 1980, and references cited in these publications). The bond in question is that between residues 15 and 16 of the inhibitor. Much of the specificity of this interaction depends on the side-chain of residue 15, which is a lysine in the natural inhibitor. Previous methods of producing analogues of inhibitors of this type have involved enzymically catalysed insertion of the new residue (e.g. Sealock and Laskowski,1969), or a variety of chemical procedures (e.g. Jering and Tschesche, 1976; cf. Kowalski, 1978). Recombinant methods have also been used to prepare four analogues of α$_1$-proteinase inhibitor at the corresponding key site (Rosenberg *et al.*, 1984; Courtney *et al.*, 1985; Matheson *et al.*, 1986) in that inhibitor. These alterations, naturally enough, were limited to exchanges between natural, coded amino acids.

In the study mentioned above (Tschesche *et al.*, 1987), there is no limitation either to coded amino acids, or to those able to participate in enzymically catalysed insertion. The key bond, Lys$^{15}$−Ala$^{16}$, was first split by an enzyme less susceptible than trypsin to the inhibiting action of aprotinin, or by rendering this peptide bond susceptible to fast tryptic cleavage by selective cleavage of the neighbouring Cys$^{14}$−Cys$^{38}$ disulphide bond (Jering and Tschesche, 1976), and then all the side-chain carboxyl groups of the protein were protected by esterification so that they could not participate in any of the subsequent steps.

Although it was unavoidable that the newly liberated carboxyl group of residue Lys$^{15}$ was esterified as well, it could be selectively deprotected by using a protease as a specific esterase. Carboxypeptidase B then removed the lysine, leaving the carboxyl group of the residue Cys$^{14}$ as the only free one in the whole molecule. Appropriate derivatives of new amino acids were attached to this uniquely reactive group by a variety of the conventional techniques of protein semisynthesis, whilst the closure of the peptide bond between residues 15 and 16 was accomplished either by chemical or enzymic coupling.

The substitutions that were carried out involve the replacement of the −(CH$_2$)$_4$−NH$_2$ of the lysine side-chain by a range of neutral groups of different size and degree of branching. In some instances this involved replacement of the lysine by coded amino acids (Gly, Ala, Val, Leu, Ile), whilst in others the replacement was by non-coded ones [side-chains −CH$_2$CH$_3$, −(CH$_2$)$_2$CH$_3$, −(CH$_2$)$_3$CH$_3$, −C(CH$_3$)$_3$, −CH$_2$C(CH$_3$)$_3$]. The spectrum of

proteases inhibited by these analogues differs radically from one analogue to the other. Quite new biological specificities are created, and the theoretical and practical implications are discussed by the authors (Tschesche *et al.*, 1987). The processes involved in producing these analogues seem no more complex than the various steps of the corresponding recombinant methods (in those cases in which they can be applied, i.e. to coded amino acids). There is also no problem of expression, refolding, or the formation of the correct disulphide bridges, as there would be with recombinant products. Yields are reasonable and the methods lend themselves to scaling up.

Bode *et al.* (1984) have reported the refined 2.2 Å X-ray crystal structure of the ternary complex formed by bovine trypsinogen, Val−Val, and semisynthetic [Arg¹⁵] bovine pancreatic trypsin inhibitor. The inhibitor analogue was obtained by enzyme-catalysed fragment-condensation. Replacement of the reactive site Lys¹⁵ by an arginine residue is accompanied in the complex by small movements within the specificity pocket of polar side-chains belonging to trypsinogen. Solvent molecules, also, are moved. It was possible to detect the expulsion from the pocket of a specific water molecule, which allowed the bulkier guanadinyl group to approach the carboxylate group of Asp¹⁸⁹ of the trypsinogen.

## Conclusion

As far as the relative merits of different approaches are concerned, the facts should be left to speak for themselves. In science it is generally true that no one method is universally applicable. Many laboratories (including the author's own) already use both chemical and recombinant methods, as appropriate. This trend is likely to continue, and it is very much to be hoped that the present prospect, of the two approaches complementing each other in an ever more powerful way, will in fact be realized.

## Acknowledgements

## References

Bode,W , Walter,J , Huber,R., Wenzel,H.R and Tschesche,H. (1984) *Eur J. Biochem.*, 144, 185−190.

Busch,M.R , Maskalick,O G , Neireiter,G W , Harris,D E and Gurd,F R. (1985) *Biochemistry*, 24, 6707−6716

Chaiken,I.M (1981) *CRC Crit. Rev Biochem.*, 11, 255−301.

Chang,J.Y., Alkan,S.S , Hilschmann,N. and Braun,D G (1985) *Eur J Biochem.*, 151, 225−230

Corthésy,B E and Wallace,C J.A (1986) *Biochem. J*, 236, 359−364

Courtney,M., Jallet,S , Tessier,L.H , Benaverte,A , Crystal,R G and Lecocq,J.P (1985) *Nature*, 313, 149−151

Davies,J G. and Offord,R E (1985) *Biochem. J*, 231, 389−392

Davies,J G , Muir,A V. and Offord,R.E. (1986) *Biochem. J*, 240, 609−612

Deber,C M., Hruby,V J and Kopple,K D. (1985) *Peptides, Structure and Function* Proc. *9th American Peptide Symposium.* Pierce Chemical Co., Rockford, IL

Duckworth,W C and Kitabchi,A E. (1981) *Endocrinol. Rev*, 2, 210−233

Flanders,K C , Horowitz,E M and Gurd,R S (1984) *J Biol Chem.*, 259, 7031−7037.

Gavish,M , Zakut,R , Wilchek,M. and Givol,D. (1978) *Biochemistry*, 17, 1345−1351.

Graf,L and Li,C.H (1981) *Proc Natl. Acad. Sci USA*, 78, 6135−6141

Hagenmaier,H., Ohms,J.P , Jahns,J and Anfinsen,C B. (1978) In Offord,R.E and Di Bello,C (eds), *Semisynthetic Peptides and Proteins* Academic Press, London, pp 23−25.

Halban,P A and Offord,R E (1975) *Biochem. J*, 151, 219−225.

Haneda,M., Kobayashi,M , Maegawa,H , Watanabe,B , Takata,Y , Ishibashi,O , Shigeta,Y and Inouye,K (1985) *Diabetes*, 34, 568−573.

Harris,D E and Offord,R E (1977) *Biochem. J*, 161, 21−25

Inouye,K., Watanabe,K , Morihara,K , Tochino,T , Kanaya,T , Emura,J and Sakakibara,S. (1979) *J. Am. Chem. Soc*, 101, 751−752

Jain,M K , Mahval,B P , De Haas,G H and Slotboom,A J (1986) *Biochim. Biophys. Acta*, 860, 448−461

Jering,H and Tschesche,H (1976) *Eur J Biochem.*, 61, 443−452.

Kaiser,E T and Lawrence,D S (1984) *Science*, 226, 505−511.

Kanmera,T and Chaiken,I M (1985) *J Biol Chem.*, 260, 8474−8482

Kihara,H., Ishimaru,K and Ohno,M. (1981) *J Biochem*, 90, 363−370

Kobayashi,M , Haneda,J , Maegawa,H , Watanabe,N , Takada,Y , Shigeta,Y and Inouye,K (1984) *Biochem. Biophys Res. Commun.*, 119, 49−57

Kobayashi,M , Takata,Y , Ishibashi,O., Sasaoka,T , Twasaki,T M , Shigeta,Y and Inouye,K (1986) *Biochem. Biophys Res Commun.*, 137, 250−257

Kowalski,D (1978) In Offord,R E and Di Bello,C (eds), *Semisynthetic Peptides and Proteins* Academic Press, London, pp. 262−282

Kubick,T and Cowburn,O. (1986) *Int J. Pept Protein Res*, 27, 514−521

Laskowski,M , Jr (1978) In Offord,R E and Di Bello,C (eds), *Semisynthetic Peptides and Proteins* Academic Press, London, pp 255−262

Laskowski,M., Jr and Kato,I (1980) *Annu Rev Biochem.*, 49, 593−626

Leatherbarrow,R.J and Fersht,A (1986) *Protein Eng.*, 1, 7−16

Leszczynski,J F and Rose,S D (1986) *Science*, 234, 849−855

Li,C H , Blaker,J and Hayashida,T. (1978) *Biochem. Biophys Res Commun*, 82, 217−222

Markussen,J (1981) European Patent Application

Matheson,N R., Gibson,H L., Hallewell,R A , Barr,P J and Travis,J (1986) *J Biol. Chem.*, 261, 10404−10409.

Muir,A V , Offord,R E and Davies,J G (1986) *Biochem. J.*, 237, 631−637

Nakagawa,S.H. and Tager,H S (1986) *J Biol Chem*, 261, 7332−7341

Neet,K.E. and Koshland,D E (1966) *Proc Natl. Acad. Sci. USA*, 56, 1606−1611

Offord,R E (1980) *Semisynthetic Proteins* John Wiley, Chichester

Offord,R E (1982) In Blaha,K and Malon,P (eds), *Peptides 1982* De Gruyter, Berlin, pp 31−42

Offord,R.E (1985) *Proc. 9th American Peptide Symposium.* Pierce Chemical Co , Rockford, IL, pp 325−334

Offord,R E and Di Bello,C (eds) (1978) *Semisynthetic Peptides and Proteins* Academic Press, London

Polgar,L and Bender,M L (1966) *J Am. Chem. Soc*, 88, 3153−3154

Proudfoot,A E I., Wallace,C.J A , Harris,D.E and Offord,R E (1986) *Biochem. J.*, 239, 333−337

Ragnarsson,U (ed ) (1984) *Peptides 1984 Proc 18th European Peptide Symposium.* Almqvist and Wiksell International, Stockholm, pp 235−238

Richards,F M (1958) *Proc Natl Acad Sci USA*, 44, 162−166

Rose,K , Pochon,S and Offord,R E. (1984) In Ragnarson,U (ed.), *Peptides 1984.* Almqvist and Wiksell International, Stockholm

Rose,K , Herrero,C., Proudfoot,A E.I , Wallace,C J A and Offord,R E (1987) In Theodoropoulos,D (ed ), *Proceedings of the 19th European Peptide Symposium.* W de Gruyter, Berlin

Rosenberg,S and Terry,W D (1977) *Adv Cancer Res.*, 25, 323−388

Rosenberg,S , Barr,P J , Najarian,R.C and Hallewell,R.A (1984) *Nature*, 312, 77−80

Sasakin,D.H , Kelly,C E , Martin,P D , Edwards,B F and Doscher,M S (1985) *Arch. Biochem. Biophys*, 241, 132−140

Sealock,R.W and Laskowski,M , Jr (1969) *Biochemistry*, 8, 3703−3710

Seetharam,R and Acharya,A S (1986) *J Cell Biochem*, 30, 87−99

Seino,A., Funakoshi,A , Fu,Z Z and Vinik,A (1985) *Diabetes*, 34, 1−7

Sheppard,R C (1979) In Gross,E and Maienhofer,J (eds), *The Peptides* Academic Press, London, Vol 2, pp 441−484

Shoelson,S E , Polonsky,K S , Zeidler,A , Rubenstein,A H and Tager,H S (1984) *J. Clin. Invest*, 73, 1351−1358.

Simmerman,H.K B , Wang,C.C , Horowitz,E M , Berzofsky,J A and Gurd,F.R N. (1982) *Proc Natl Acad. Sci USA*, 79, 7739−7743

Slaby,I. and Holmgren,A (1975) *J Biol Chem.*, 250, 1340−1347.

Stern,M J and Doscher,M S (1986) *FEBS Lett*, 171, 253−256

Taniuchi,H , Anfinsen,C B and Sodja,A (1967) *Proc Natl Acad Sci USA*, 58, 1235−1242.

Taylor,W C , Komonya,A and Chaiken,I.M (1985) *Proc Natl. Acad. Sci USA*, 82, 6423−6426.

Ten Kortenaar,P.B , Adams,P J and Tesser,G.I (1985) *Proc Natl. Acad. Sci USA*, 82, 8279−8283

Tschesche,H and Kupfer,S (1976) *Hoppe-Seyler's Z Physiol. Chem.*, 357, 769−776

Tschesche,H , Beckman,J , Mehlid,A , Schnabel,E and Wenzel,H R (1987) *Biochim. Biophys Acta*, in press

van Scharrenburg,G J M , Jansen,E H , Egmon,M.R , de Haas,G.H and Slot-
  boom,A J. (1984) *Biochemistry*, **23**, 6235−6294.
van Scharrenburg,G J.M , Puijk,W C., de Haas,G H. and Slotboom,A.J (1983)
  *Eur J. Biochem.*, **133**, 83−89
Veloso,D , Juillerat,M. and Taniuchi,H. (1984) *J. Biol. Chem.*, **259**, 6067−6073.
Wallace,C.J.A. and Rose,K. (1983) *Biochem. J.*, **215**, 651−658.
Wallace,C J.A. and Corthésy,B.E (1986) *Protein Eng*, **1**, 23−27
Wallace,C J.A. and Proudfoot,A E I. (1987) *Protein Eng*, **1**, 246
Wallace,C J A , Corradin,G , Marchiori,F and Bonn,G (1986) *Biopolymers*,
  **25**, 2121−2132
Zahn,H , Naithani,V K., Gattner,H G , Büllesbach,E E. and Thamm,P M. (1981)
  *Naturwissenschaften*, **68**, 56−62