Tetrahedron Vol. 35. pp. 2767 to 2769

0040-4020/79/1201-2767/$02.00/0

PEPTIDES—XXXIII[1]

SYNTHESIS OF A POLYPEPTIDE CHAIN COMPRISING 129 RESIDUES. STRATEGY AND TACTICS

GEORGE W. KENNER†, ROBERT RAMAGE[2] and ROBERT C. SHEPPARD[3]
The Robert Robinson Laboratories, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, England

(*Received in the UK 8 May 1979*)

Abstract—A plan for synthesis of a small protein by fragment condensation with side-chain protection by t-butyl, adamantyloxycarbonyl and acetamidomethyl groups is discussed.

The aim of synthesising proteins has been a goal of organic chemists since the days of Emil Fischer.[4] Many efforts have been described, some achieving considerable success.[5] Nevertheless the problem remains unsolved, despite achievements in the Insulin field, culminating in a brilliant synthesis of crystalline human zinc insulin.[6] Insulin is indeed a protein, and one whose synthesis posed extraordinary difficulties, and yet the conjunction of two chains extending to only 21 and 30 residues respectively limits its relevance to the general problem. At this stage in the development of bio-organic chemistry the aim must be efficient synthesis, to the standards of molecular science rather than of biological science, of peptide chains comprising 100–300 residues. If this could be achieved, we could begin to look forward to the more exciting prospect of synthesising still more valuable proteins, an order of magnitude greater in molecular size.

Hitherto the greatest success in this area has undoubtedly been achieved by R. B. Merrifield and his collaborators by means of the solid-phase method.[7] The excellent work on ribonuclease-A8 shows the power of this method, and yet one may be forgiven for doubting that this approach will lead eventually to the goal of synthesising crystalline small enzymes.

The programme described in subsequent papers of this series arose from discussion with D. C. Phillips of the problem of defining the mechanism of lysozyme action. Much is already known from the pioneer X-ray analysis of the hens egg enzyme and its complex with an inhibitor and from related model building.[9] But the power of synthetic organic chemistry offers the hope of providing more definitive answers by making available tailor-made proteins possessing different binding sites and reactive functions. We therefore set out to synthesise a polypeptide comprising 129 residues, closely modelled on hens egg lysozyme but possessing 28 changes in the sequence.[10] Five of these changes correspond to those in human leukaemic lysozyme.[11] The rationale for these changes was to alleviate the difficulty of synthesis, e.g. by removing residues of arginine, histidine, and methionine, without destroying the intrinsic propensity of the chain to fold in the manner necessary to facilitate formation of the essential disulphide links at 6–127, 30–115, 64–80, and 76–94. These changes, despite being devised after careful examination of the molecular model, carry the risk that folding will be directed in another course and hence the product from dehydrogenation of the octa-sulphhydryl 129 chain will not yield a synthetic enzyme. That would be disappointing, but it is not relevant to the substance of this paper and its immediate congeners.

The real question is whether a polypeptide chain of more than 100 residues can be satisfactorily constructed to acceptable standards. We believe that it can, and we hope to demonstrate this in the subsequent papers in this series. The purpose of this introductory paper is to set the scene for detailed account of the experimental work carried out during the last seven years by our group.

We decided to adopt the strategy of fragment condensation and the tactics of "permanent" side-chain protection by t-Bu groups, as employed with such outstanding success by the CIBA (now CIBA-GEIGY) group in their numerous hormone syntheses. Fragment condensation is, in our view, the only discernible way of achieving the purification of intermediates which is a cardinal requirement in organic synthesis. The target sequence could be divided into 12 fragments all C-terminating in glycine. This offered the opportunity of fragment condensations without risk of racemisation at that chiral centre labilised in activation of the C-terminal carboxyl group. It was our intention to construct all these fragments by the stepwise Bodanszky technique, which is itself essentially free of racemisation risk. In practice this strategy had to be modified, and the next paper in these series shows, for instance, how the 1–16 sequence is more effectively constructed by union of 1–10 and 11–16, instead of 1–4 and 5–16 as originally planned. Nevertheless the original strategy has proved its worth.

As for the t-butyl tactic, we early on decided to

† Deceased, 25.6.1978.

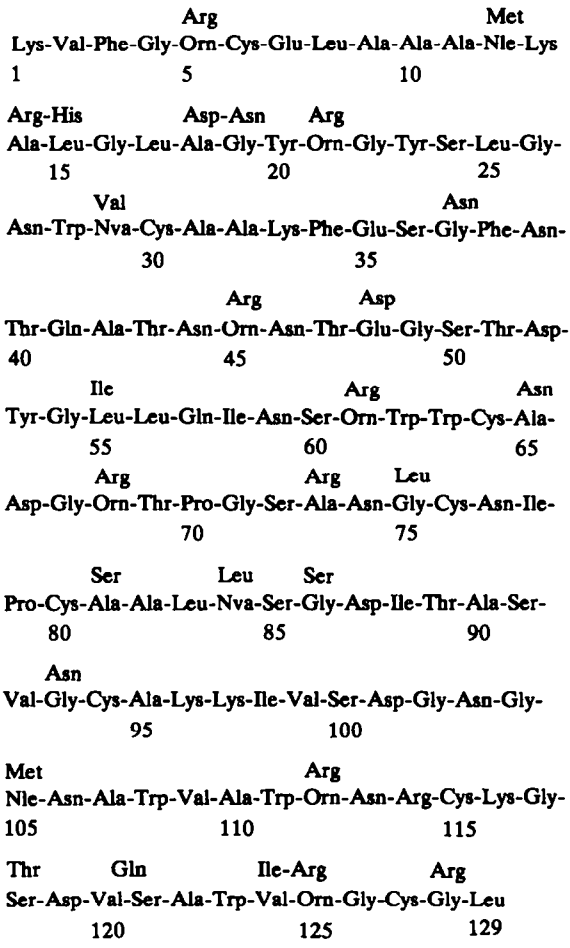

Fig. 1. Sequence of lysozyme analogue being synthesized in Liverpool. Superimposed residues at positions 5, 12, 14, 15, 18, 19, 21, 29, 37, 45, 48, 55, 61, 65, 68, 73, 75, 82, 84, 86, 93, 105, 112, 118, 121, 124, 125, and 128 show sequence of hens egg lysozyme.

modify this by employing adamantyloxycarbonyl protection[12] for the ε-amino-group of lysine and the δ-amino-group of ornithine. Boc groups have, of course, been used many times for such purposes, but the synthesis which we have undertaken imposes stringent requirements of stability in the "permanent" protecting groups, and some preliminary experiments indicated partial loss of side-chain Boc groups. We also firmly adhered to maximal protection of phenolic and alcoholic hydroxyl

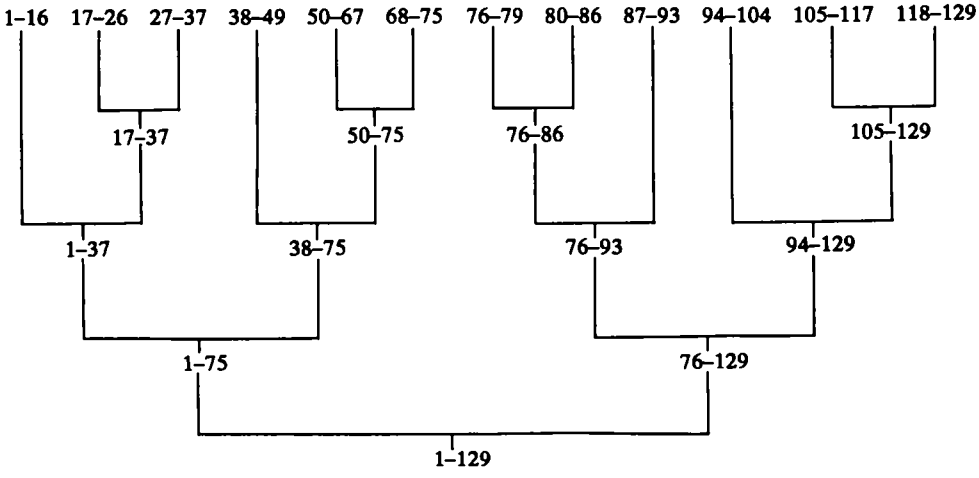

Fig. 2. Scheme for synthesis of target sequence

groups; we regard the attractive possibility of omitting this protection as a snare.[10] On the other hand we have not employed protection of amide side-chains, and perhaps that was an error which should be corrected in subsequent syntheses. Side-chain carboxyl groups were all protected as t-butyl esters in the usual way.

The choice of *S*-protection for the eight cysteine residues is central to design of the synthesis. The benzyl group has been popular, ever since the pioneering studies of du Vigneaud on glutathione[13] and oxytocin.[14] Nevertheless it is widely recognised that removal with sodium in liquid ammonia can cause side-reactions. These may be acceptable in a peptide with 10 residues, but be intolerable with 100 residues. We therefore chose the acetamidomethyl group, introduced by the Merck group.[15] In practice this has proved to be unexpectedly advantageous in conferring increased solubility in dipolar aprotic solvents, such as dimethylformamide. An important potential advantage of the acetamidomethyl group is that it survives the acidolytic cleavage of t-Bu and adamantyloxycarbonyl groups. There is thus the attractive possibility of purifying thoroughly the octa-*S*-acetamidomethyl polypeptide of 129 residues, before setting free the sulphhydryl groups.

The choice of "temporary" protecting groups, i.e. those for the amino and carboxyl functions which are to be incorporated in the chain, is also vital to the strategy. Wherever possible we have used the classical, still unsurpassed benzyloxycarbonyl ("carbobenzoxy") group of Bergmann and Zervas.[16] When this was precluded by presence of cysteine residues, we resorted to the biphenylisopropyloxycarbonyl group of the CIBA laboratories[17] which can now be removed so smoothly at constant "pH" in trifluoroethanol.[18] For carboxyl protection we adopted the phenyl ester group, for reasons which are fully discussed in the preceding paper of this series.[1]

We believe that the subsequent papers will demonstrate the essential soundness of this strategy and these tactics. Nevertheless it is easy to see the desirability of improvements,[10] and it is our intention to develop them.

Without anticipating the detailed discussion of our experimental work, it is appropriate to list below the sequence under attack and the outline scheme of fragment condensation.

Acknowledgements—This work has been supported by grants from the Science Research Council, Imperial Chemical Industries Ltd., Roche Products Ltd., and exceptionally generous provision of laboratory space, research studentships, post-doctoral appointments and technical appointments at the University of Liverpool.

REFERENCES

[1]Part XXXII: I. J. Galpin, P. M. Hardy, G. W. Kenner, J. R. McDermott, R. Ramage, J. H. Seely and R. G. Tyson, *Tetrahedron* submitted for publication.
[2]Present address: Chemistry Department, UMIST, Sackville Street, Manchester.
[3]Present address: M.R.C. Laboratory of Molecular Biology, Cambridge.
[4]E. Fischer, *Aus meinem Leben*. Julius Springer, Berlin (1922).
[5]e.g. R. G. Denkewalter, D. F. Veber, F. W. Holly and R. Hirschmann, *J. Am. Chem. Soc.* **91,** 502 (1969) and consecutive papers.
[6]P. Sieber, B. Kamber, A. Hartmann, A. Jöhl, B. Riniker and W. Rittel, *Helv. Chim. Acta* **57,** 2617 (1974).
[7]B. W. Erickson and R. B. Merrifield, *The Proteins* (Edited by H. Neurath and R. L. Hill) Vol. 2, p. 257. Academic Press, New York (1976).
[8]B. Gutte and R. B. Merrifield, *J. Biol. Chem.* **246,** 1922 (1971).
[9]C. C. F. Blake, L. N. Johnson, G. A. Mair, A. C. T. North, D. C. Phillips and V. R. Sarma, *Proc. Roy. Soc.*, **B167,** 378 (1967).
[10]G. W. Kenner, *Proc. Roy. Soc.* **A353,** 441 (1977).
[11]R. E. Canfield, S. Kammerman, J. A. Sobel and F. J. Morgan, *Nature New Biol.* **232,** 16 (1971).
[12]W. L. Haas, E. V. Krumkalns and K. Gerzon, *J. Am. Chem. Soc.* **88,** 1988 (1966).
[13]V. du Vigneaud and G. L. Miller, *J. Biol. Chem.* **116,** 469 (1936).
[14]V. du Vigneaud, C. Ressler, J. M. Swan, C. W. Roberts and P. G. Katsoyannis, *J. Am. Chem. Soc.* **76,** 3115 (1954).
[15]D. F. Veber, J. D. Milkowski, S. L. Varga, R. G. Denkewalter and R. Hirschmann, *Ibid.* **94,** 5456 (1972).
[16]M. Bergmann and L. Zervas, *Ber. Dtsch. Chem. Ges* **65,** 1192 (1932).
[17]P. Sieber and B. Iselin, *Helv. Chim. Acta* **51,** 622 (1968).
[18]B. Riniker, B. Kamber and P. Sieber, *Ibid.* **58,** 1086 (1975).