Tetrahedron Letters,Vol.23,No.18,pp 1877-1880,1982 0040-4039/82/181877-04$03.00/0
Printed in Great Britain

SEMISYNTHESIS OF A SHORTENED OPEN-CHAIN PROINSULIN

Erika E. Büllesbach

Deutsches Wollforschungsinstitut an der RWTH Aachen
Veltmanplatz 8, D-5100 Aachen FRG

Summary : The openchain proinsulin model : B-chain-Arg-Arg-Ala-Gly-Lys-Arg-A-chain was prepared from partial protected native insulin chains and synthetic peptides.

In the crystal structure of insulin the α-amino group Gly^{A1} and the α-carboxyl group Ala^{B30} are separated by only 8 - 10 Å[1]. The analogue structure of insulin was found in the biological precursor proinsulin[1]. In this molecule the C-terminal amino acid of the insulin B-chain and the N-terminal amino acid of the insulin A-chain are connected with a C-peptide (length: 27 to 35 amino acids). This sequence has a high species-variability. The connecting peptide facilitates the formation of the correct disulphide bridges[2,3]. However this role cannot be explained by its length and variability. Moreover, its structure-inducing function can be reproduced by a linkage between the two amino groups in $Gly^{\alpha A1}$ and $Lys^{\varepsilon B29}$ via one of several dicarbonic acid derivatives[4].

Proinsulin models with shortened connecting peptides should have reoxidation properties similar to $Gly^{\alpha A1}$ and $Lys^{\varepsilon B29}$ linked insulins[4] but, in contrast to these, models with α-peptide bonds can be prepared by gene technology.

The present aim is to synthesize several proinsulin models with connecting peptides differing in length and sequence from the native C-peptide. Kullmann and Gütte[5] prepared a proinsulin analogue in which the B-chain sequence B_{1-29} was connected with a modified A-chain sequence by the solid phase methode.

The present work describes the semisynthesis of the open-chain proinsulin analogue (B-chain)--Arg-Arg-Ala-Gly-Lys-Arg-(A-chain) as hexa-S-sulphonate. The most important steps were: i) the N-terminal elongation of the native insulin A-chain with synthetic peptides, ii) the modification of the B-chain to permit selective activation of the α-carboxyl group, iii) the condensation of the both partially protected, large fragments, iv) the deprotection of the open-chain proinsulin model, and v) the isolation of the single chain peptide as hexa-S-sulphonate.

For N-terminal elongation of native A-chain with synthetic peptides, the azide method[6] (with its low rate of reaction) as well as the faster mixed anhydride method[7] have been used.

The starting material for the proinsulin model was Arg-A-chain*disulphide monomer prepared by a tryptic catalyzed reaction[8]. Further condensation with synthetic peptides were performed usually via mixed anhydride reactions[9]. Carboxyl components were activated by reaction with isobutyloxycarbonylchloride in DMF for 2 min at -10°C. For coupling, the amino components were dissolved in 0.2M N-methylmorpholine/HCl buffer in DMF/water 4:1 at "pH 7.2 and reacted with a

5-fold excess of carboxyl component for 2 min at RT. The reaction was stopped by addition of a 10 to 50 fold excess of $NH_2OH \cdot HCl$ in DMF/water 4:1 at "pH 7.2" for 2 h. Possibly formed O-acyl derivatives were cleaved under these conditions. The protein was isolated after addition of acetic acid by gel-filtration on Sephadex G-25 in 10% acetic acid and lyophilization. The reaction was controlled by paper electrophoresis and end group determination by the Dansyl-technique[10].

By these procedures, the peptide Boc-Gly-Lys(Msc)OH was condensed to the Arg-A-chain. No free α-amino group was observed by end group determination[10]. The liberation of the α-amino group by treatment with TFA (99%) was incomplete. Nearly 10% isobutyloxycarbonyl-Arg-A-chain derivative was separated by ion exchange chromatography on SP-Sephadex in 7M urea and 1.5M acetic acid with a linear NaCl-gradient (0-0.15M). After desalting and lyophilization Gly-Lys (Msc)-Arg-A-chain* was isolated in a yield of 73.1% based on Arg-A-chain.* The tripeptidyl-A-chain was further elongated with the protected tetrapeptide Boc-Thr(Bu^t)-Arg-Arg-Ala-OH·2TosOH. Although the amino component reacted incompletly, the isolation of the protected heptapeptidyl-A-chain was not necessary because after treatment with TFA (99%) the charge difference between unreacted amino component, isobutyloxycarbonylated amino component[7], and the deprotected product allowed separation by ion exchange chromatography in the previous system (with exception that the salt gradient was from 0 to 0.4M NaCl). Thr-Arg-Arg-Ala-Gly-Lys(Msc)-Arg-A-chain* was isolated in a yield of 47.5% based on tripeptidyl A-chain. In this case the N-terminal amino acid Thr corresponded to the C-terminal amino acid of the human B-chain.

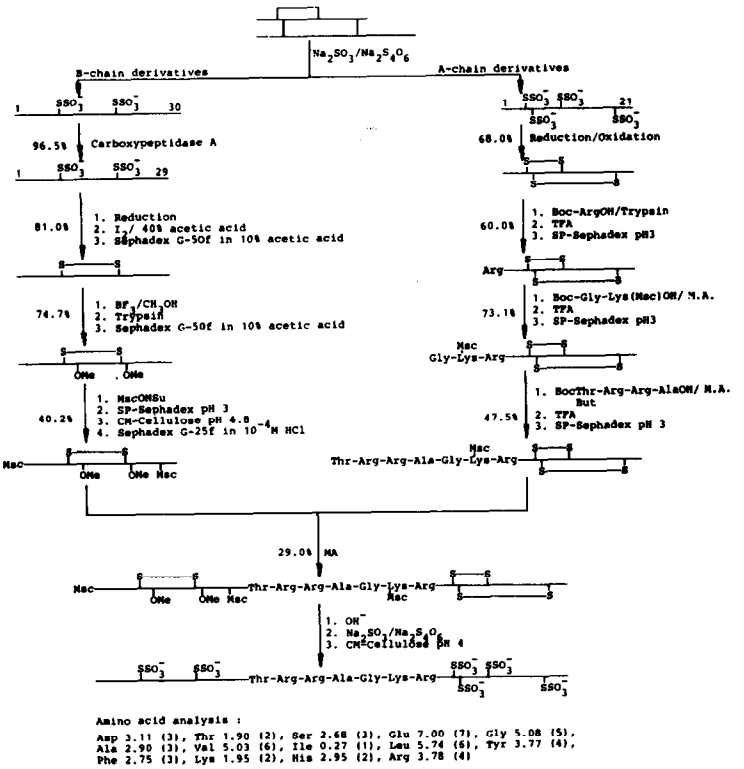
For subsequent N-terminal elongation of this heptapeptidyl-A-chain derivative, des-B30-B-chain had to be activated at the C-terminus. In an earlier work the mixed anhydride activation of an amino protected B-chain derivative has been reported[7]. In this experiment the selective reaction of the α-carboxyl group was not achieved and only low yields resulted[7]. For the selective activation of the C-terminal carboxyl group, protection of side chain carboxyl groups was necessary[11,12]. The preparation of such a protected native B-chain derivative was possible by a route involving several steps[11,12]. Di-S-sulphonate-des-B30-B-chain was prepared by digestion with carboxypeptidase A. The Bunte salt groups were reduced with 2-mercaptoethanol, and the di-thiol-B-chain was separated by gel-filtration on Sephadex G-25 in 30 % acetic acid. Cyclic disulphide B-chain derivative was prepared by oxidation with iodine[12,13] followed by separation on Sephadex G-50f in 10% acetic acid. The esterfication of the three carboxyl groups (Glu B13 and B21, Lys B29) of this derivative was performed in methanol using BF_3[14]. Selective cleavage of the C-terminal methoxy group was achieved by tryptic digestion[15] at pH 5.7 in dioxane/water 1:1, 6h at RT. After reaction with methylsulphonylethyloxycarbonylsuccinimide ester, both amino groups (Phe B1 and Lys B29) were protected. The resulting $N^{\alpha 1}$,$N^{\varepsilon 29}$-bis(methylsulphonylethyloxycarbonyl)-des-B30-B-chain-disulphide-$O^{\gamma 13}$,$O^{\gamma 21}$-di-methyl ester was purified by ion exchange chromatography and converted to the hydrochloride by gel-chromatography on Sephadex G-25f in 10^{-4}M HCl.

To prepare the shortened proinsulin model, 17.5 µmol B-chain derivative were dissolved in 0.75ml DMF and 4 equivalents N-methylmorpholine were added. Under these conditions the α-carboxyl group and both histidine side chains were deprotonated. The mixed anhydride was formed by reaction with 2 equivalents isobutyl chloroformate by -10°C for 30 sec. 5 µmol heptapeptidyl-A-chain derivative were dissolved in 0.37ml 0.2 M-methylmorpholine in DMF/water 4:1 at "pH

7.2", and added to the mixed andydride. The coupling reaction was stopped after 2 min by addition of 0.5ml 1 M $NH_2OH \cdot HCl$ in DMF/water 4:1 at "pH 7.2". The mixture was allowed to stand 12h, then separated by gel-filtration on Sephadex G-50f in 1M acetic acid and lyophilized (yield of protected open-chain proinsulin model was 29% based on heptapeptidyl-A-chain).

In numerous experiments the reaction conditions were varied but the described yields could not be improved. Two equivalents of isobutyl chloroformate were necessary, because two reactions of the carboxyl component were possible: first the formation of the mixed anhydride and second the incomplete formation of N^{im}-isobutyloxycarbonyl derivative of the histidine residues. With higher excess of isobutyl chloroformate the formation of N^{α}-isobutyloxycarbonyl derivatives of the amino component dominated. By electrophoretic studies of several coupling reactions both side reactions were always observed, but to different extents. The irreversible blocking of the amine component was low under the described conditions. The formation of N^{im}-isobutyloxycarbonyl side product was reversible by the treatment with NH_2OH[12] and in the alkaline medium[16] during the cleavage of the Msc- and methyl protecting groups.

The deprotection of the open-chain proinsulin model was performed in 0.1M NaOH for 30min at 0°C under N_2 at a protein concentration of 10mg/ml. The reaction was stopped by addition of glacial acetic acid, and the deprotected derivative was purified by gel-filtration on Sephadex G-50f in 1M acetic acid, lyophilized and converted to the hexa-S-sulphonate derivative by oxidative sulphitolysis[17]. This was purified by ion exchange chromatography on CM-cellulose at pH 4 in 7M urea, 0.1M acetic acid and eluted with a NaCl-gradient (0-0.1M). The shortened open-



Scheme: Pathway of the semisynthesis of the open-chain proinsulin model

chain proinsulin derivative was characterized by electrophoresis at pH 2.2 and 4.8, by amino acid analysis, end group determination[10], by tryptic digestion followed by electrophoresis (Pauli positive fragments A_{1-21}, B_{1-22}, B_{1-29} and B_{23-29}), and digestion with trypsin and carboxypeptidase B[18] followed by amino acid analysis (1Lys and 3 Arg were determinated from the "connecting peptide"). All fragments were also formed by the analogue digestion of the hexa-S-sulphonate bovine proinsulin.

In summary with the N-terminal elongation of the A-chain[7,8] and the C-terminal elongation of the protected B-chain derivative[12] a semisynthetic method was developed which can be used to prepare several shortened proinsulin derivatives for systematic investigations of the re-oxidation properties. These experiments should allow further elucidation of the role of the C-peptide in the biological precursor, proinsulin.

Acknowledgements

This work was financially supported by the Deutsche Forschungsgemeinschaft.

I am grateful to Prof. Dr. H. Zahn and Dr. H.-G. Gattner for their constant encouragment during the course of this work. I thank Dr. M. Casaretto for her help in the peptide synthesis, Dr. J. Föhles for the amino acid analysis, and Dr. D. Saunders for the critical reading of this manuscript. I thank Mrs.G. Ortmann for her competent assistance.

Abbreviations : Boc = tert.butyloxycarbonyl-, But = tert.butyl-, CM = carboxymethyl-, DMF = dimethylformamide, M.A. = mixed anhydride, Msc = methylsulphonylethyloxycarbonyl-, SP = sulphopropyl-, TFA = trifluoroacetic acid

References

1. Blundell,T., Dodson,G., Hodgkin,D. & Mercola,D. (1972) Advances Prot. Chem. 26, 279-402.
2. Steiner,D.F. & Clark,J.L (1968) Proc.Natl.Acad.Sci.USA 60, 622-629 .
3. Frank,B.H., Pettee,J.M., Zimmerman,R.E. & Burck,P.J. (1981) Proc. 7thAmer.Peptide Sympos. (Rich,D.E. ed.) Pierce Chemical Co. Rockford Ill.,in press .
4. Geiger,R. (1979) Chemiker Zeitung 100, 111-129.
5. Kullmann,W. & Gutte,B. (1975) Biochem.Soc.Trans. 3, 899-902 .
6. Aiba,H. & Shimonishi,Y. (1980) Bull.Chem.Soc.Jpn 53, 192-196, 197-200, 201-204 .
7. Naithani,V.K., Gattner,H.G., Büllesbach,E.E., Föhles,J. & Zahn,H. (1979) in "Peptides" Proc. 6th Amer.Peptide Sympos. (Gross,E. & Meienhofer,J. eds.) pp 571-576, Pierce Chemical Co. Rockford Ill. .
8. Gattner,H.G., Danho,W. & Naithani,V.K. (1980) "Insulin" Proc. 2nd Int. Insulin Sympos. (Brandenburg,D & Wollmer,A.,eds.) pp 117-123 Walter de Gruyter Berlin-New York .
9. Wieland,T. & Bernhard,H. (1951) Justus Liebigs Ann.Chem. 572, 190-194; Boissonnas,R.A. (1951) Helv.Chim.Acta 34, 874-879; Vaughan,J.R. & Osato,R.L. (1951) J.Amer.Chem.Soc. 73 3547 .
10. Gray,W.R. (1967) Methods Enzymol. 11, 139-151 .
11. Schmitt, E.W. (1977) Doctoral thesis RWTH Aachen .
12. Büllesbach,E.E., Schmitt,E.W., Gattner,H.G., Naithani,V.K. & Föhles,J. (1982) Proc. 3rd USSR-FRG Sympos. on Chemistry of Peptides and Proteins (Voelter,W., Wünsch,E., Ivanov,V.T., Ovchinnikov,Yu.A.,eds.) Walter de Gruyter Berlin-New York,in press.
13. Büllesbach,E.E., Schmitt,E.W. & Gattner,H.G. (1982) Int.J.Peptide Protein Res.,in press .
14. Gattner,H.G., Schmitt,E.W. & Naithani,V.K. (1975) Hoppe Seyler's Z.Physiol.Chem. 356, 1465-1467.
15. Coletti-Previero,M.A., Previero,E.& Zuckerkandl,J. (1969) J.Mol.Biol. 39 493-501.
16. Gronvald,F.C., Johansen,N.L. & Lundt,B.F. (1979) in "Peptides" Proc. 6th Amer.Peptide Sympos. (Gross,E. & Meienhofer,J., eds.) pp 309-312, Pierce Chemical Co. Rockford Ill..
17. Bailey,J.L. & Cole,R.D. (1959) J.Biol.Chem. 234, 1733-1739.
18. Kemmler ,W., Peterson,J.D. & Steiner,D.F. (1971) J.Biol.Chem. 246, 6786-6791.

* The compensating ion of the arginine side chain was not determinated.

(Received in Germany 4 December 1981)