Insulin Peptides. XVI. The Synthesis of a Nonapeptide and a Dodecapeptide Derivative Related to the A Chain of Human Insulin (Positions 1–9 and 10–21)[1,2]

Panayotis G. Katsoyannis, Clyde Zalut, and Andrew M. Tometsko

Contribution from the Division of Biochemistry, Medical Research Center, Brookhaven National Laboratory, Upton, New York. *Received April 3, 1967*

Abstract: Syntheses are described of protected peptides related to the N-terminus and C-terminus of the human insulin A chain. Thus the preparation is given of N-carbobenzoxyglycyl-L-isoleucyl-L-valyl-γ-*t*-butyl-L-glutamyl-L-glutaminyl-S-benzyl-L-cysteinyl-S-benzyl-L-cysteinyl-L-threonyl-L-serine hydrazide and N-carbobenzoxy-L-isoleucyl-S-benzyl-L-cysteinyl-L-seryl-L-leucyl-L-tyrosyl-L-glutaminyl-L-leucyl-L-glutamyl-L-asparaginyl-L-tyrosyl-S-benzyl-L-cysteinyl-L-asparagine *p*-nitrobenzyl ester. The former derivative corresponds to positions 1–9 and the latter to positions 10–21 in the human insulin A chain sequence.

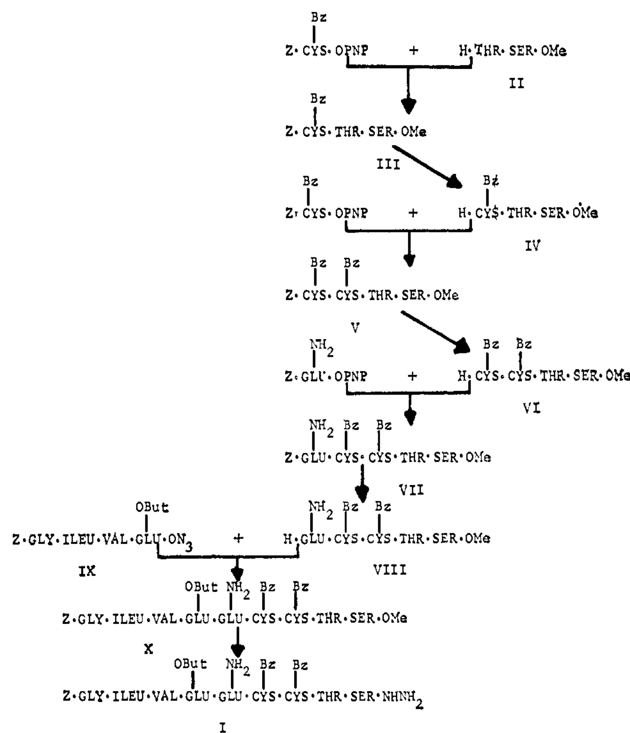
Recent communications[3,4] from this laboratory described the synthesis of peptide derivatives related to the N-terminus of the A chain of sheep insulin and the synthesis of a dodecapeptide derivative containing the C-terminal sequence of that chain. An intermediate in the construction of the latter peptide fragment, namely, the C-terminal nonapeptide portion, embodies the amino acid sequence found at the carboxyl terminus of the A chain of human insulin. The present report describes the synthesis of a protected nonapeptide and a protected dodecapeptide containing the N-terminal and C-terminal amino acid sequence, respectively, of the human insulin A chain.

The amino acid sequence of the reduced A chain of human insulin as determined by Nicol and Smith[5] is glycylisoleucylvalylglutamylglutaminylcysteinylcysteinylthreonylserylisoleucylcysteinylserylleucyltyrosylglutaminylleucylglutamylasparaginyltyrosylcysteinylasparagine. An identical sequence was proposed for the reduced A chain of porcine insulin.[6] The total synthesis of the human insulin A chain in the S-sulfonated form is presented in the following paper[7] and the preparation of two key intermediates used for that synthesis is described in the present report. One of these intermediates is the hydrazide of N-carbobenzoxyglycyl-L-isoleucyl-L-valyl-γ-*t*-butyl-L-glutamyl-L-glutaminyl-S-benzyl-L-cysteinyl-S-benzyl-L-cysteinyl-L-threonyl-L-serine (I), the nonapeptide fragment which contains the amino acid sequence found at the amino terminus of the A chain, and the other intermediate is N-carbobenzoxy-L-isoleucyl-S-benzyl-L-cysteinyl-L-seryl-L-leucyl-L-tyrosyl-L-glutaminyl-L-leucyl-L-glutamyl-L-asparaginyl-L-tyrosyl-S-benzyl-L-cysteinyl-L-asparagine *p*-nitrobenzyl ester (XVI), the dodecapeptide subunit which contains the amino acid sequence found at the carboxyl terminus of that chain.

Preparation of these intermediates was accomplished following essentially the scheme we employed in the synthesis of the corresponding peptide fragments of the A chain of sheep insulin.[3,4] Thus the synthesis of the nonapeptide derivative I involved the condensation of an activated protected tetrapeptide subunit with a partially protected pentapeptide fragment, and the preparation of the dodecapeptide derivative XVI was brought about by the azide reaction of a tripeptide fragment with a nonapeptide subunit.

The over-all scheme which was used for the construction of the protected nonapeptide hydrazide I is summarized in Chart I. N-Carbobenzoxy-S-benzyl-L-cysteine *p*-nitrophenyl ester[8] was allowed to react with L-threonyl-L-serine methyl ester (II), which was prepared

Chart I



(1) This work was supported by the U. S. Atomic Energy Commission.

(2) A preliminary report describing the results presented in this paper has appeared: P. G. Katsoyannis, A. Tometsko, and C. Zalut, *J. Am. Chem. Soc.*, **88**, 166 (1966).

(3) P. G. Katsoyannis, A. M. Tometsko, and C. Zalut, *ibid.*, **88**, 5622 (1966).

(4) P. G. Katsoyannis, A. M. Tometsko, and C. Zalut, *ibid.*, **88**, 5618 (1966).

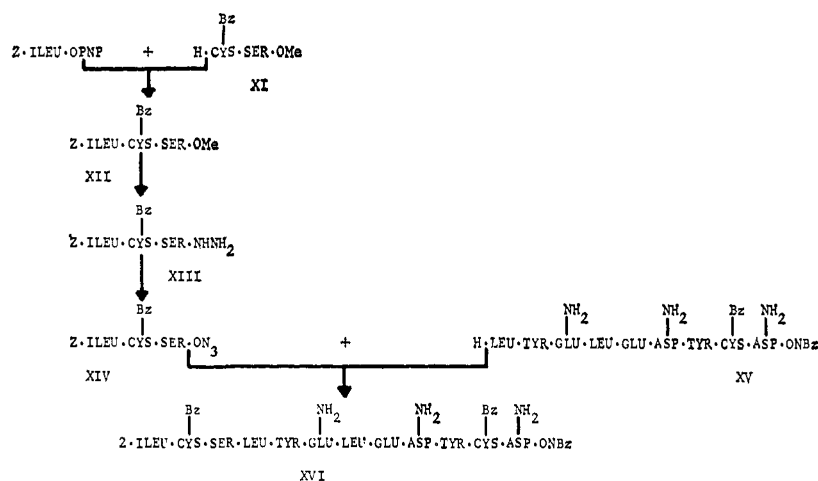
(5) D. S. H. W. Nicol and L. F. Smith, *Nature*, **187**, 483 (1960).

(6) J. I. Harris, F. Sanger, and M. A. Naughton, *Arch. Biochem. Biophys.*, **65**, 427 (1956).

(7) P. G. Katsoyannis, A. M. Tometsko, and C. Zalut, *J. Am. Chem. Soc.*, **89**, 4505 (1967).

(8) M. Bodanszky and V. du Vigneaud, *ibid.*, **81**, 5688 (1959).

Chart II



by hydrogenolysis of its carbobenzoxylated derivative,[9] to give N-carbobenzoxy-S-benzyl-L-cysteinyl-L-threonyl-L-serine methyl ester (III). Decarbobenzoxylation of the protected tripeptide III yielded the derivative IV bearing a free amino group which was condensed with N-carbobenzoxy-S-benzyl-L-cysteine *p*-nitrophenyl ester to yield the protected tetrapeptide V. Removal of the amino-protecting group from V and coupling of the ensuing partially protected derivative VI with the *p*-nitrophenyl ester of N-carbobenzoxy-L-glutamine[8] afforded the pentapeptide derivative VII which, on exposure to HBr in trifluoroacetic acid, yielded the decarbobenzoxylated derivative VIII. The latter compound was completely digested by leucine aminopeptidase as judged by amino acid analysis of the digest. Conversion of N-carbobenzoxyglycyl-L-isoleucyl-L-valyl-γ-*t*-butyl-L-glutamic acid hydrazide, whose synthesis has been described in a previous report,[3] to the corresponding azide IX and coupling of the latter compound with the pentapeptide derivative VIII gave the fully protected nonapeptide methyl ester X in 60% yield. On exposure to hydrazine hydrate X was converted to the desired protected nonapeptide hydrazide I in 78% yield.

The synthesis of the dodecapeptide intermediate is illustrated in Chart II. S-Benzyl-L-cysteinyl-L-serine methyl ester (XI), which was prepared from its carbobenzoxylated derivative[10] by treatment with HBr in trifluoroacetic acid, was condensed with N-carbobenzoxy-L-isoleucine *p*-nitrophenyl ester[8] to give N-carbobenzoxy-L-isoleucyl-S-benzyl-L-cysteinyl-L-serine methyl ester (XII). On exposure to hydrazine hydrate the protected tripeptide XII was converted to the hydrazide XIII which, in turn, in the usual manner, gave the corresponding azide XIV. The protected nonapeptide N-carbobenzoxy-L-leucyl-L-tyrosyl-L-glutaminyl-L-leucyl-L-glutamyl-L-asparaginyl-L-tyrosyl-S-benzyl-L-cysteinyl-L-asparagine *p*-nitrobenzyl ester, whose synthesis was described previously,[4] was decarbobenzoxylated by treatment with HBr in acetic acid, and the resulting derivative XV bearing a free amino group was coupled with the tripeptide azide XIV to give the desired protected dodecapeptide XVI in 89% yield. Elemental analysis and amino acid analysis of the protected peptide and paper chromatography of the decarbobenzoxylated derivative established the chemical purity of compound XVI. Complete digestibility of the decarbobenzoxylated derivative by leucine aminopeptidase established its stereochemical homogeneity.

(9) E. Schroder and H. Gibian, *Ann.*, **656**, 190 (1962).
(10) S. Guttmann and R. A. Boissonnas, *Helv. Chim. Acta*, **43**, 201 (1960).

Experimental Section

Capillary melting points were determined for all compounds and are corrected.

For paper chromatography the protected peptides were deblocked on exposure to 2 *N* HBr in acetic acid. Serine-containing peptides were deblocked by treatment with HBr in trifluoroacetic acid.[11] The resulting hydrobromides were chromatographed on Whatman No. 1 filter paper. $R_f{}^1$ values refer to the Partridge system;[12] $R_f{}^2$ values refer to the system[13] 1-butanol–pyridine–acetic acid–water (30:20:6:24) and are expressed as a multiple of the distance traveled by a histidine marker. The amino acid analyses of acid hydrolysates were carried out by the method of Moore, Spackman, and Stein[14] with a Beckman–Spinco amino acid analyzer, Model 120B, to which a digital readout system, Model CRS-10AB (Infotronics Corp., Houston, Texas), has been attached. Enzymatic analyses (LAP) were carried out according to the method of Hill and Smith[15] using a crystalline enzyme (Worthington). Optical rotations were taken with a Zeiss photoelectric precision polarimeter.

The following abbreviations are used: DMF, dimethylformamide; DMSO, dimethyl sulfoxide; HMPA, hexamethylphosphoramide; TFA, trifluoroacetic acid; Z, carbobenzoxy; Bz, benzyl; PNP, *p*-nitrophenyl; Me, methyl; NBz, *p*-nitrobenzyl; N_3, azide; But, *t*-butyl.

N-Carbobenzoxy-S-benzyl-L-cysteinyl-L-threonyl-L-serine Methyl Ester (III). N-Carbobenzoxy-L-threonyl-L-serine methyl ester (21 g) was dissolved in methanol (300 ml) containing glacial acetic acid (7 ml) and hydrogenated for 2 hr over 10% palladium–charcoal catalyst (4 g). The catalyst was filtered off and the filtrate was evaporated to dryness *in vacuo*. The remaining product was dried by the addition of anhydrous methanol followed by evaporation under reduced pressure. To a solution of the residue in DMF (300 ml) made slightly alkaline with triethylamine was added N-carbobenzoxy-S-benzyl-L-cysteine *p*-nitrophenyl ester (27 g). After 24 hr at room temperature the reaction mixture was poured into ethyl acetate (1200 ml) and water (300 ml). The organic layer was separated, washed successively with 1 *N* NH_4OH, water, 1 *N* HCl, and water, and dried with $MgSO_4$. The crystalline product formed on removal of the ethyl acetate *in vacuo* was recrystallized from methanol–water; 24 g (76%); mp 185–187°; $[\alpha]^{27}D$ −16.4° (*c* 1, DMF); after HBr in TFA treatment: $R_f{}^1$ 0.74, $R_f{}^2$ 4.1 × His; single ninhydrin-positive spot.

Anal. Calcd for $C_{26}H_{33}N_3O_8S$: C, 57.1; H, 6.03; N, 7.7. Found: C, 57.3; H, 6.19; N, 8.0.

(11) S. Guttmann and R. A. Boissonnas, *ibid.*, **41**, 1852 (1958).
(12) S. M. Partridge, *Biochem. J.*, **42**, 238 (1948).
(13) S. G. Waley and G. Watson, *ibid.*, **55**, 328 (1953).
(14) S. Moore, D. H. Spackman, and W. A. Stein, *Anal. Chem.*, **30**, 1185 (1958).
(15) R. L. Hill and E. L. Smith, *J. Biol. Chem.*, **228**, 577 (1957).

N-Carbobenzoxy-S-benzyl-L-cysteinyl-S-benzyl-L-cysteinyl-L-threonyl-L-serine Methyl Ester (V). N-Carbobenzoxy-S-benzyl-L-cysteinyl-L-threonyl-L-serine methyl ester (19.8 g) was dissolved in TFA (35 ml) and HBr was passed through the solution for 1 hr at 0°. Addition of anhydrous ether to the reaction mixture caused the decarbobenzoxylated tripeptide ester to precipitate as the hydrobromide which was filtered off, washed with ether, and dried over KOH *in vacuo*. To a solution of this product in DMF (120 ml), triethylamine (8.0 ml) was added followed by N-carbobenzoxy-S-benzyl-L-cysteine *p*-nitrophenyl ester (16.3 g). After 24 hr at room temperature the reaction mixture was poured into ethyl acetate (1200 ml) and water (200 ml). The organic layer was separated, washed successively with 1 *N* NH_4OH, water, 1 *N* HCl, and water, and dried with $MgSO_4$. The crystalline product formed on removal of the ethyl acetate *in vacuo* was recrystallized from acetic acid–water; 19.0 g (74%); mp 174–178°; $[\alpha]^{27}D$ −26.0° (*c* 1, DMF); after HBr in TFA treatment: R_f^1 0.87, R_f^2 4.45 × His; single ninhydrin-positive spot.

Anal. Calcd for $C_{36}H_{44}N_4O_9S_2$: C, 58.4; H, 6.00; N, 7.6. Found: C, 57.9; H, 6.04; N, 7.3.

N-Carbobenzoxy-L-glutaminyl-S-benzyl-L-cysteinyl-S-benzyl-L-cysteinyl-L-threonyl-L-serine Methyl Ester (VII). N-Carbobenzoxy-S-benzyl-L-cysteinyl-S-benzyl-L-cysteinyl-L-threonyl-L-serine methyl ester (9 g) was dissolved in TFA (30 ml) and water (0.5 ml), and HBr was passed through this solution for 1 hr at 0°. The tetrapeptide ester hydrobromide was precipitated with anhydrous ether, filtered off, washed with ether, and dried over KOH *in vacuo*. To a solution of this solid in DMF (50 ml), triethylamine (3.0 ml) was added, followed by N-carbobenzoxy-L-glutamine *p*-nitrophenyl ester (5.0 g). After 48 hr at room temperaure the reaction mixture was poured into cold 1 *N* NH_4OH (400 ml). The precipitate was filtered off, washed successively with 1 *N* NH_4OH, water, 1 *N* HCl, and water, and dried; 8.0 g (74%); mp 220–221°. For analysis a sample was reprecipitated from acetic acid–water; mp 222–223°; $[\alpha]^{27}D$ −26.9° (*c* 1, DMF); after HBr in TFA treatment: R_f^1 0.75, R_f^2 4.36 × His; single ninhydrin-positive spot; amino acid ratios in acid hydrolysate: $Thr_{1.1}Ser_{0.9}Glu_{1.0}$S-benzylcysteine$_{2.0}$; amino acid ratios of an LAP digest: $Thr_{1.0}$-$Ser_{1.0}Glu(NH_2)_{0.9}$S-benzylcysteine$_{2.1}$. Serine was separated from glutamine in a 30° chromatographic run.

Anal. Calcd for $C_{41}H_{52}N_6O_{11}S_2$: C, 56.7; H, 6.00; N, 9.7. Found: C, 56.4; H, 6.27; N, 9.2.

N-Carbobenzoxyglycyl-L-isoleucyl-L-valyl-γ-*t*-butyl-L-glutamyl-L-glutaminyl-S-benzyl-L-cysteinyl-S-benzyl-L-cysteinyl-L-threonyl-L-serine Methyl Ester (X). Compound VII (2 g) was dissolved in TFA (25 ml) and water (0.5 ml), and HBr was passed through this solution for 1 hr at 0°. The hydrobromide was precipitated with anhydrous ether (300 ml), filtered, washed with ether, and dried over KOH *in vacuo*. To a solution of this product in DMF (30 ml) made slightly alkaline with triethylamine (0.8 ml) was added the tetrapeptide azide prepared as follows. A suspension of N-carbobenzoxyglycyl-L-isoleucyl-L-valyl-γ-*t*-butyl-L-glutamic acid hydrazide (2 g) in DMF (40 ml) was cooled to −15° (Dry Ice–acetone) and brought into solution by the addition of 2 *N* HCl (4.5 ml.) To this solution $NaNO_2$ (220 mg) in water (1 ml) was added. After 5 min at −15° the reaction mixture was poured into cold (0°) half-saturated NaCl (250 ml) and the precipitated tetrapeptide azide was isolated by filtration, washed with cold water, and dried at 0° over P_2O_5 *in vacuo*. This azide was then added to the solution of the amino component which was prepared as described previously. The reaction mixture was stirred at 0° for 48 hr and then poured into methanol (300 ml). The precipitated product was isolated by filtration, washed with methanol and water, and dried; 1.8 g (60%); mp 257° dec. For analysis a sample was reprecipitated from dimethylformamide–water; mp 263–264° dec; $[\alpha]^{27}D$ −30.85° (*c* 1, DMF); after HBr in TFA treatment: R_f^1 0.71, R_f^2 4.39 × His; single ninhydrin-positive spot; amino acid ratios in acid hydrolysate: $Thr_{1.0}Ser_{0.9}Glu_{2.0}Gly_{1.1}Val_{0.8}Ile_{0.8}$S-benzylcysteine$_{1.9}$.

Anal. Calcd for $C_{63}H_{90}N_{10}O_{17}S_2$: C, 57.2; H, 6.80; N, 10.6. Found: C, 57.2; H, 6.62; N, 10.7.

N-Carbobenzoxyglycyl-L-isoleucyl-L-valyl-γ-*t*-butyl-L-glutamyl-L-glutaminyl-S-benzyl-L-cysteinyl-S-benzyl-L-cysteinyl-L-threonyl-L-serine Hydrazide (I). A solution of the nonapeptide ester derivative X (0.9 g) in DMF (35 ml) containing hydrazine hydrate (4 ml) was stirred for 24 hr at 42° and for 48 hr at 25°. The reaction mixture was then poured into H_2O (100 ml), and the pH of the mixture was adjusted to 6 with acetic acid. The precipitated hydrazide was filtered, washed with water, dried, and reprecipitated from hexamethylphosphoramide–water; 0.7 g (78%); mp 258° dec; $[\alpha]^{28}D$ −20.0° (*c* 1, HMPA).

Anal. Calcd for $C_{62}H_{90}N_{12}O_{16}S_2$: C, 56.3; H, 6.80; N, 12.70; S, 4.84. Found: C, 56.4; H, 7.19; N, 12.73; S, 5.40.

N-Carbobenzoxy-L-isoleucyl-S-benzyl-L-cysteinyl-L-serine Methyl Ester (XII). N-Carbobenzoxy-S-benzyl-L-cysteinyl-L-serine methyl ester (13.5 g) was dissolved in TFA (40 ml) and water (0.6 ml) and HBr was passed through the solution for 1 hr at 0°. Addition of dry ether to the reaction mixture caused the precipitation of the peptide ester hydrobromide as a heavy oil. The ether was decanted and the residue dried over KOH *in vacuo*. To a solution of this material in DMF (100 ml) containing triethylamine (7 ml) was added N-carbobenzoxy-L-isoleucine *p*-nitrophenyl ester (11.0 g). After 24 hr the reaction mixture was diluted with 1 *N* NH_4OH (2 ml), stirred for 30 min, and poured into cold 1 *N* NH_4OH (300 ml). The precipitated product was collected by filtration, washed successively with 1 *N* NH_4OH, 1 *N* HCl, and water, and crystallized from methanol–water; 8.0 g (48%); mp 196–198°; $[\alpha]^{27}D$ −31.0° (*c* 1, DMF); after treatment with HBr in TFA: R_f^1 0.84, R_f^2 5.24 × His; single ninhydrin-positive spot.

Anal. Calcd for $C_{28}H_{37}N_3O_7S$: C, 60.1; H, 6.62; N, 7.5. Found: C, 60.7; H, 6.91; N, 8.1.

N-Carbobenzoxy-L-isoleucyl-S-benzyl-L-cysteinyl-L-serine Hydrazide (XIII). To a solution of N-carbobenzoxy-L-isoleucyl-S-benzyl-L-cysteinyl-L-serine methyl ester (3 g) in warm methanol (350 ml) was added hydrazine hydrate (8 ml). After 48 hr at room temperature the precipitated hydrazide was isolated by filtration, washed with methanol, and reprecipitated from dimethylformamide–water; 2.4 g (80%); mp 224–225°; $[\alpha]^{27}D$ −22.0 (*c* 1, DMF).

Anal. Calcd for $C_{27}H_{37}N_5O_6S$: C, 58.0; H, 6.62; N, 12.5. Found: C, 58.6; H, 6.79; N, 12.0.

N-Carbobenzoxy-L-isoleucyl-S-benzyl-L-cysteinyl-L-seryl-L-leucyl-L-tyrosyl-L-glutaminyl-L-leucyl-L-glutamyl-L-asparaginyl-L-tyrosyl-S-benzyl-L-cysteinyl-L-asparagine *p*-Nitrobenzyl Ester (XVI). A suspension of N-carbobenzoxy-L-leucyl-L-tyrosyl-L-glutaminyl-L-leucyl-L-glutamyl-L-asparaginyl-L-tyrosyl-S-benzyl-L-cysteinyl-L-asparagine *p*-nitrobenzyl ester (4 g) in acetic acid (30 ml) was mixed with 4 *N* HBr in acetic acid (30 ml). After 1 hr the resulting solution was poured into dry ether (600 ml), and the precipitated hydrobromide of the partially protected nonapeptide was collected by filtration, washed with ether, and dried over KOH *in vacuo*. To a cooled (0°) solution of this product in DMF (60 ml), triethylamine (1.7 ml) was added, followed by the tripeptide azide prepared as follows. N-Carbobenzoxy-L-isoleucyl-S-benzyl-L-cysteinyl-L-serine hydrazide (2.2 g) was dissolved in a mixture of DMF (70 ml) and 2 *N* HCl (5.0 ml). After cooling this solution to −15° $NaNO_2$ (300 mg) dissolved in cold water (2 ml) was added. The reaction mixture was stirred at −15° for 10 min, and then poured into an ice-cold, saturated solution of NaCl (350 ml). The precipitated azide was filtered off, washed successively with cold (0°) water, 1 *N* $NaHCO_3$, and water, and dried for 1 hr at 0° over P_2O_5 *in vacuo*. The tripeptide azide was then added to the solution of the nonapeptide derivative which was prepared as described previously. The reaction mixture was stirred for 48 hr at 0° and then poured into 1 *N* HCl (400 ml). The precipitated product was filtered off, washed with water, and triturated with hot methanol; 4.5 g (89%); mp 258–260° dec. A sample for analysis was precipitated from dimethylformamide–water; mp 263–265° dec; $[\alpha]^{27}D$ −28.1° (*c* 1, DMSO); after treatment with HBr in TFA: R_f^1 0.83, R_f^2 4.76 × His. Amino acid analysis of an acid hydrolysate of the protected dodecapeptide by the automatic analyzer showed the following composition expressed in molar ratios: $Asp_{2.0}Ser_{0.9}$-$Glu_{2.0}Ile_{1.0}Leu_{2.0}Tyr_{1.5}$S-benzylcysteine$_{1.8}$.

Anal. Calcd for $C_{92}H_{118}N_{16}O_{25}S_2$: C, 57.8; H, 6.18; N, 11.7. Found: C, 57.5; H, 6.61; N, 11.3.

For evaluation of stereochemical homogeneity a sample of the decarbobenzoxylated dodecapeptide was digested with LAP. Amino acid analysis of the digest gave the following composition expressed in molar ratios: $Ser_{1.0}Glu_{1.0}Ile_{1.2}Leu_{1.8}Tyr_{1.7}$S-benzylcysteine$_{1.9}$. Glutamine and asparagine emerge at the same position, and were not determined. Serine was determined in a 30° chromatographic run.

Acknowledgment. The authors wish to express their appreciation to Miss Karin Scheibe for the amino acid analyses and enzymatic analyses.