

Peptide segment ligation strategy without use of protecting groups

CHUAN-FA LIU AND JAMES P. TAM*

Department of Microbiology and Immunology, Vanderbilt University, A5119 Medical Center North, Nashville, TN 37232-2363

Communicated by R. B. Merrifield, March 22, 1994

ABSTRACT We describe the concept and the verification of a chemical ligation approach to the synthesis of proteins using peptide segments with no protecting groups and no activation of the C-terminal α -carboxyl group. This approach consists of three steps: (i) aldehyde introduction, in which a masked glycolaldehyde ester is linked to the carboxyl terminus of an unprotected peptide by reverse proteolysis; (ii) ring formation, in which the unmasked aldehyde reacts with the N-terminal α -amino group of the second unprotected peptide containing either a cysteine or a threonine residue to form a thiazolidine or oxazolidine ring at an acidic pH; and (iii) rearrangement in which O-acyl ester linkage is transferred to N-acyl amide linkage to form a peptide bond with a pseudoproline structure at higher pH. The feasibility of this scheme was verified by a model study on small compounds and its potential was demonstrated by the synthesis of a 50-residue epidermal growth factor-like peptide containing a preformed disulfide bond.

The chemical synthesis of peptides has been important for the understanding of protein structure–function relationships and the discovery of new therapeutic agents (1). At present, stepwise solid-phase peptide synthesis (2) can efficiently prepare moderate-size peptides of nearly 100 amino acids (3). To extend beyond this range, a practical approach would be a chemical method that ligates unprotected peptide segments to form larger proteins. Such a method would exploit the efficiency of the stepwise synthesis. It would also allow the synthesis of more complex peptides and proteins, such as those with branched structure (4), and the semi-synthesis of proteins consisting of components from synthetic peptides with unusual amino acids and proteins obtained through recombinant technology.

The conventional segment-synthesis approaches (5) usually require a maximal protection strategy which has the limitations of poor solubility, low coupling efficiency, and the possible danger of racemization due to the overactivation of the carboxyl group. Considerable efforts have been made to overcome these limitations through the use of partial or minimal protecting-group strategies and improvements in the activation methods (6–11). However, these approaches do not overcome the problem of the poor coupling efficiency between large peptide segments, because of the intrinsic difficulty of obtaining effective molar concentrations for high molecular weight molecules. Alternative approaches to overcome the difficulty of forming amide bonds use hydrazone (12, 13) or thioether and thioester (14) as surrogate bonds. The advantage of these peptide-bond replacement approaches resides in the high and specific reactivity between the two functional groups not found in amino acids to make the conjugation efficient. The disadvantage is that these linkages are substantially different from the peptide bond, which may lead to conformational distortion, and are usually not stable at either acidic (12) or basic (14) pH.

We have now designed and developed a chemical ligation method by which two unprotected peptide segments can be ligated together through an amide bond. In our design, we make use of a highly regioselective and efficient reaction to link covalently two unprotected peptides. The selectivity of this reaction would also obviate the need for any protecting groups. A peptide bond is then formed in a second step through an intramolecular rearrangement between the two closely neighboring carboxyl and amino groups. The efficiency of the first reaction would solve the slow kinetic problem for reactions between large molecules, and the second reaction, designed as a spontaneous process to form amide bonds without activating the carboxyl group, would overcome the racemization problem of the conventional segment condensation method.

EXPERIMENTAL PROCEDURES

General. ^1H NMR spectra were obtained at 300 MHz on a Bruker AC 300 spectrometer. Analytical HPLC was run on a Shimadzu system with a Vydac column (0.46 cm \times 25 cm, C₁₈ reverse phase, 5 μm) at a flow rate of 1.5 ml/min, with detection at 225 nm; solvent A was 5% CH₃CN in H₂O [0.045% trifluoroacetic acid (TFA)]; solvent B was 60% CH₃CN in H₂O (0.039% TFA). Mass spectra were obtained with the fast-atom bombardment method for the small compounds and with the laser desorption time-of-flight method for large peptides unless otherwise indicated.

Synthesis of Compounds IIb–g. Z-alanine 2,2'-dimethoxyethyl ester (IIb, see Fig. 2; Z, benzyloxycarbonyl) was prepared in 80% yield by reaction of Z-alanine cesium salt with bromoacetaldehyde dimethyl acetal in dimethylformamide (DMF) at 60°C for 48 hr (15). R_f = 0.66 for TLC on silica gel in ethyl acetate/hexanes, 1:1; t_R = 24.8 min for HPLC (20% B–50% B in 30 min); MS [M+H]⁺, m/z 312; ^1H NMR (C₂HCl₃/tetramethylsilane internal standard), δ 7.33 ppm (s, 5H), 5.31 (bd, 1H, J = 6.5 Hz), 5.11 (s, 2H), 4.56 (t, 1H, J = 5.3 Hz), 4.43 (dq, 1H, J_1 = 7.2 Hz, J_2 = 6.5 Hz), 4.17 (2q, 2H), 3.38 (s, 6H), 1.43 (d, 3H, J = 7.2 Hz).

IIc was obtained by treatment of IIb with TFA (5% H₂O) in CH₂Cl₂ (1:3) at room temperature for about 20 min. After removal of TFA, the released aldehyde was used right away for the next step. R_f = 0.3, ethyl acetate/hexanes, 1:1; t_R = 9.7 min; ^1H NMR (C₂HCl₃); δ 9.6 (s, 1H), 7.33 (s, 5H), 5.32 (bd, 1H, J = 6.4 Hz), 5.14 (s, 2H), 4.77 (q, 2H, J = 17.5 Hz), 4.54 (dq, 1H, J_1 = 7.2 Hz, J_2 = 6.4 Hz), 1.53 (d, 3H, J = 7.2 Hz).

Ring formation. IIc-Tfa salt was prepared from IIc with 2-aminoethanethiol hydrochloride in CH₃CN/H₂O, 1:3 (pH 5), in 15 min at room temperature and purified by HPLC (82% from IIb). R_f = 0.34, CHCl₃/CH₃OH/HOAc, 96:3:1; t_R = 11.0 min; MS [M+H]⁺, m/z 325. The NMR analysis of IIc-IIg was complicated by the mixture of diastereoisomers and the cis/trans conformations for the rearranged product,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: DMF, dimethylformamide; DMSO, dimethyl sulfoxide; TFA, trifluoroacetic acid; Thz, thiazolidine; HMThz, hydroxymethyl-Thz.

*To whom reprint requests should be addressed.

so an approximate assignment was given for these products. $^1\text{H NMR}$ (C^2HCl_3), δ 7.34 (5H), 6.31 (b, 2H), 5.64 and 5.29 (2 bd, 1H), 5.09 (2H), 4.96 (1H), 4.55–4.30 (3H), 3.67–3.43 (2H), 3.19–3.05 (2H), 1.40 (3H). **Ile** was obtained from **Iib** and cysteine by the same procedure as **Iid** (yield, 85% from **Ib**). $R_f = 0.31$, $\text{CHCl}_3/\text{CH}_3\text{OH}/\text{HOAc}$, 90:8:2; $t_R = 11.1$ and 11.6 min; MS $[\text{M}+\text{H}]^+$, m/z 369; $^1\text{H NMR}$ (C^2HCl_3), δ 7.69 (b, 2H), 7.32 (5H), 5.73 (bd, 1H), 5.10 (2H), 4.96 and 4.79 (2t, 1H), 4.53–4.31 (2H), 4.28–4.10 (2H), 3.31–3.07 (2H), 1.40 (3H).

O-to-N acyl transfer (Iif). Samples of **Iid** (0.1 mmol) were dissolved in 4 ml of 0.1 M phosphate buffer and 2 ml of CH_3CN . The pH of the solution was adjusted to values ranging from 6 to 9. The reaction was monitored by HPLC. $R_f = 0.38$, $\text{CHCl}_3/\text{CH}_3\text{OH}/\text{HOAc}$, 96:3:1; $t_R = 13.2$ and 13.4 min; MS $[\text{M}+\text{H}]^+$, m/z 325; $^1\text{H NMR}$ (C^2HCl_3), δ 7.34 (s, 5H), 5.72 and 5.56 (2bd, 1H), 5.52 and 5.41 (2t, 1H), 5.10 (2H), 4.84 and 4.53 (1H), 4.15–3.40 (4H), 3.20–2.90 (2H), 1.77 (1H), 1.39–1.32 (3H). **Ilg** was obtained from **Iie** in the same procedure as **Iif**. $R_f = 0.34$, $\text{CHCl}_3/\text{CH}_3\text{OH}/\text{HOAc}$, 90:8:2; $t_R = 8.6$ min; MS $[\text{M}+\text{Na}]^+$, m/z 391; $^1\text{H NMR}$ (C^2HCl_3), δ 7.32 (5H), 5.90 and 5.80 (1H), 5.09 (2H), 5.03–4.83 (1H), 4.53–4.31 (2H), 3.95–3.75 (2H), 3.55–3.30 (2H), 3.15 and 3.11 (1H), 1.43–1.39 (3H).

Synthesis of Iib and Iif. Peptides **Iib** and **Iif** (see Fig. 3) were synthesized by the solid-phase method (2). **Iib** was prepared with *N*-hydroxypropionyl-4-methylbenzhydrylamine (MBHA) resin (1% divinylbenzene crosslink) and **Iif** was prepared with the conventional Pam resin (16). Both peptides were prepared by standard *t*-butoxycarbonyl (Boc)/benzyl (Bzl) chemistry with benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate (BOP) as coupling reagent (17) and *in situ* neutralization (18). Side-chain protecting groups were Arg(Tos), Asp(OBzl), Cys(4-MeBzl), Glu(OBzl), His(Dnp), Lys(2-ClZ), Ser(Bzl), Thr(Bzl), and Tyr(2-BrZ) (Tos, tosyl; Dnp, 2,4-dinitrophenyl). Cys-10 in **Iie** was protected by an acetamidomethyl group. N^α -acetylation for **Iib** was performed with acetic anhydride in $\text{CH}_2\text{Cl}_2/\text{DMF}$, 1:1, containing 5% *N,N*-diisopropylethylamine (DIEA) for 15 min. The Dnp groups on His residues were removed by treatment (3×8 hr) with 10% thiophenol in DMF containing 5% DIEA. Final deprotection of side chains and cleavage from resin were achieved by treatment of 0.5 g of dried protected peptide resin (N^α deprotected for peptide **Iie**) with 0.75 ml of cresol/0.25 ml of thiocresol and 10 ml of anhydrous HF at 0°C for 75 min. HF was removed by evaporation at 0°C . After washing with cold anhydrous ether, the crude peptide was extracted with 25% CH_3CN in H_2O (1% TFA). Oxidation of **Iib** for disulfide formation was performed directly on the extraction solution from HF cleavage with 10% dimethyl sulfoxide (DMSO) (19) in a total volume of 200 ml diluted with water (pH adjusted to 5.5 by sodium acetate) overnight. The oxidized **Iib** and crude **Iie** were purified by C_{18} reverse-phase HPLC on a preparative Vydac column. The HPLC fractions were lyophilized; t_R (**Iib**) = 25.8 min, t_R (**Iif**) = 16.9 min (gradient 10% B–45% B in 35 min; this gradient was used for all following steps). Amino acid analysis data were consistent with the theoretical composition for both peptides. MS for **Iib** (by electrospray ionization), molecular weight 3859.2 (calc.), 3859.3 (found); for **Iif** ($\text{M}+\text{H}^+$), m/z 1874.1 (calc.), 1874.0 (found).

Preparation of Iic. **Iic** was obtained by hydrogenolysis of **Iib** in tetrahydrofuran with 10% Pd/C for 2 hr at 20°C . The product was dissolved (≈ 1.25 M) in DMF and was kept at -70°C for next use. $R_f = 0.65$, $\text{CHCl}_3/\text{CH}_3\text{OH}/\text{HOAc}$, 6:2:1; MS $[\text{M}+\text{H}]^+$, m/z 178.

Preparation of Iid. **Iib**-3Tfa salt (4 mg, ≈ 1 μmol) was dissolved in 12 μl of 0.25 M Tris-HCl, pH 8.5/20 mM CaCl_2 , and 28 μl of 1.2 M **Iic** in DMF was added, followed by 1 μl of freshly prepared trypsin solution (1 mg/10 μl) in 0.25 M

Tris, pH 8.5/20 mM Ca^{2+} . The mixture was shaken for 75 min at room temperature. HPLC showed a new peak corresponding to coupling product ($\approx 60\%$ estimated by HPLC) eluted at 27.5 min, and the hydrolysis product of the ester handle, which was eluted at about the same time with the starting material, was confirmed by MS analysis. The coupling product was purified by a semipreparative Vydac column and the fractions were lyophilized. Amino acid analysis showed the presence of one more alanine residue. MS for **Iid** ($\text{M}+\text{H}^+$), m/z 3948.2 (calc.), 3949 (found); hydrolysis-product ($\text{M}+\text{H}^+$), m/z 3789.2 (calc.), 3788 (found).

Preparation of Iig. **Iif**-2Tfa salt (4.2 mg, ≈ 2 μmol) was mixed with **Iid** in a 5-ml flask, and 1 ml of cooled TFA (5% H_2O) was added. The reaction mixture was kept at 20°C for about 5 min. TFA was then removed under reduced pressure (water aspirator and then an oil pump). Four hundred microliters of 50 mM acetate buffer (pH 3) was added and the pH was readjusted to about 3 (judged by pH paper) with 0.2 M sodium acetate. HPLC was run immediately, showing the disappearance of all acetal material and the appearance of the released aldehyde **Iie** at 25.7 min. This peak from analytical HPLC was collected and concentrated for MS analysis, which gave a molecular ion of m/z 3902.0, in agreement with the calculated value, 3902.2. The reaction between the free aldehyde **Iie** and **Iif** occurred immediately and a new HPLC peak at 28.8 min was observed, corresponding to the ring-formation product **Iig**. The reaction was $>90\%$ complete after overnight. At pH 3 and under N_2 , no significant oxidation of the free Cys of peptide **Iif** was observed. MS for **Iig**, ($\text{M}+\text{H}^+$), m/z 5757.2 (calc.), 5758 (found). The rearrangement reaction was carried out by increasing the pH to 5. A portion of the ester product was purified and lyophilized, and the rearrangement was performed separately in order to show the cleanness of the reaction (see Fig. 4).

Preparation of Iih. The lyophilized **Iig** was dissolved in 1 ml of 50 mM acetate buffer (pH 5). The rearranged product showed an HPLC peak that was eluted just before the starting unrearranged ester. The $t_{1/2}$ was about 20 hr. When the product was purified by semipreparative HPLC after 3 days, the yield was about 90% based on HPLC. MS ($\text{M}+\text{H}^+$), m/z 5757.2 (calc.), 5757 (found).

RESULTS AND DISCUSSION

Strategy for Segment Ligation. The basic idea is to bring two unprotected peptides together with unusual regioselectivity through their respective C-terminal α -carboxyl and N-terminal α -amino functional groups. To achieve this, an unusual functional group not found in peptides has to be introduced at the C terminus of one peptide segment that will react selectively with the N-terminal α -amino group of the second peptide component. We have found that simple alkyl aldehydes can fulfill this requirement.

Aldehydes condense with amines in proteins to form imines that are unstable and reversible in aqueous solution unless a ring system is formed. We made use of the reversibility of the reaction of aldehydes with amines and the stability of ring formation with β -functionalized amines in designing our chemical ligation strategy. Thus, an alkyl aldehyde was introduced via an ester linkage to the C terminus of one component to form a thiazolidine (Thz) and oxazolidine ring product, with the second component bearing a β -functionalized N-terminal α -amino group such as those occurring in Cys, Thr, and Ser (Fig. 1). In this way, the C-terminal α -carboxyl function of the first peptide (the N^α component) and the N-terminal α -amino function of the second peptide (the C^α component) are brought very close together to achieve a high effective molarity (20). The ester bond that links the peptide segment and the aldehyde was designed in such a way that an amide bond could then be

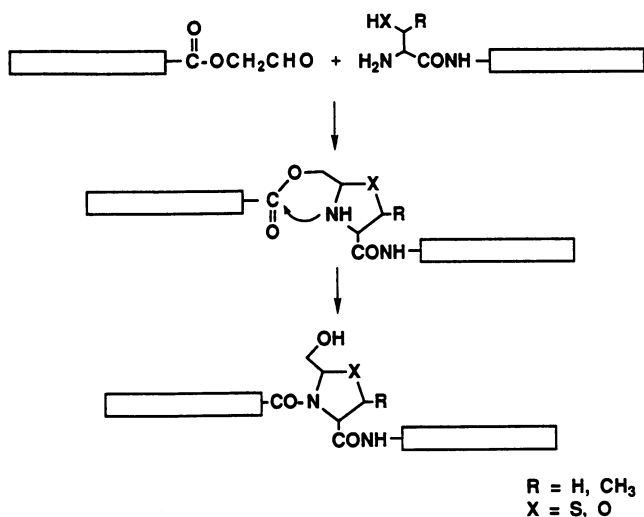


FIG. 1. General scheme of segment ligation strategy.

formed through an intramolecular O-to-N acyl transfer reaction (Fig. 1). We chose the α -acyloxyacetaldehyde system, which is a glycolaldehyde ester. In this ester-aldehyde system, the acyl carbonyl and the ring nitrogen are separated by three atoms to facilitate a five-member ring transition state. The objectives of this work are relatively similar to the elegant chemical scheme proposed by Kemp and coworkers (21), who used the thiol-capture method to mediate a covalent linkage of the N^α and C^α components through an unsymmetric disulfide bridge and subsequent aminolysis of the C^α phenolic active ester to form the peptide bond. Our scheme is different from that of Kemp and coworkers, which requires a large, aromatic template bearing a thiol function on the C^α component and a medium-size ring intermediate for the coupling step.

Model Study on Small Compounds. The aldehyde function was introduced onto Z-Ala as a masked acetal (Fig. 2) through reaction between its Cs^+ salt (IIa) and bromoacetaldehyde dimethyl acetal in DMF (15) to give the acetal Z-Ala-OCH₂CH(OCH₃)₂ (IIb). Acetal is particularly suitable because the ease of its removal and its unreactive side products (CH₃OH) eliminate the need for purification after its deprotection step. Treatment with 30% TFA in CH₂Cl₂ with a trace amount of H₂O at 0°C for 15–20 min smoothly converted the acetal (IIb) to its aldehyde Z-Ala-OCH₂CHO (IIc). TFA in other solvents such as CH₃CN also gave satisfactory results. After TFA and solvents were removed under reduced pressure, the regenerated aldehyde (IIc) was then used without purification with β -mercaptoethylamine and cysteine to form the Thz derivatives II d and II e at pH

4–5, which rearrange at higher pH to the Z-Ala-HMThz (hydroxymethyl-Thz) derivatives II f and II g (Fig. 2).

The reaction of carbonyl compounds with Cys has been extensively studied (22, 23). The sulfhydryl first reacts with the carbonyl to produce an addition product that condenses with the α -amine of the Cys to form a cyclic Thz derivative in a wide pH range. To avoid hydrolysis of the ester and the unwanted reaction of aldehyde with the side-chain amino groups, this reaction was usually conducted under acidic conditions. At pH 4–5, the Thz product was formed almost immediately, but at lower pH this reaction was slower. At pH 2, the reaction required several hours for completion. The Thz derivatives were stable and were easily purified by normal or reversed-phase HPLC under usual conditions. Application of the aldehyde reaction to N^α Thr-peptide has also been achieved, but further refinement is needed to ensure faster rates and better yields.

O-to-N Acyl Transfer Reaction to Form an Amide Bond. The basic character of the secondary amine in the Thz ring makes it possible for the acyl group to migrate from the ester oxygen to the nitrogen. The O,N acyl transfer reaction has been known since 1923 (24, 25) and is a dominating side reaction (26) in the acidic deprotection step of peptide synthesis during which the acyl moiety of the peptide migrates from the α -amine to the free hydroxyl group on the side chain of a Ser or Thr residue. The transfer reaction is reversible upon base treatment, involving a five-member oxazolidine-like ring intermediate (24). We proposed a similar rearrangement involving a 3,3,0 fused ring intermediate (Fig. 2). We undertook a careful investigation of the kinetics of this rearrangement reaction of small model compounds involving Z-Ala-OCH₂CHO (IIc) with β -mercaptoethylamine and Cys derivatives, which could be easily followed by HPLC. The results (Table 1) showed that O-to-N acyl rearrangement was effective even in the acidic pH range. The weak basicity of the Thz amine ($\text{pK}_a \approx 6.2$) (22) may have contributed to this. Rearrangement was dependent on pH but occurred under all conditions from pH 6 to pH 9 (Table 1). The presence of the α -carboxyl group of Cys in II e participates in the hydrolysis of the ester bond. At higher pH (e.g., pH 9), >60% hydrolysis product was found. In fact, the optimal pH for the rearrangement from the ester Z-Ala-Thz-CO₂H (IIe) to the amide II g was found to be around 7. In the case of peptides where the α -carboxyl group is linked through an amide bond to other amino acid residues, two consequences can be expected that will affect the acyl transfer reaction: (i) the electron-withdrawing effect of the carboxylamide bond decreases the pK_a of the Thz amine, allowing the rearrangement to occur at a lower pH and (ii) the steric constraint imposed by the -CO-Xaa linkage of the Cys carboxyl group will make the rearrangement more difficult. This was found to be true when the Cys was linked to a Leu amide or Val ethyl ester, where

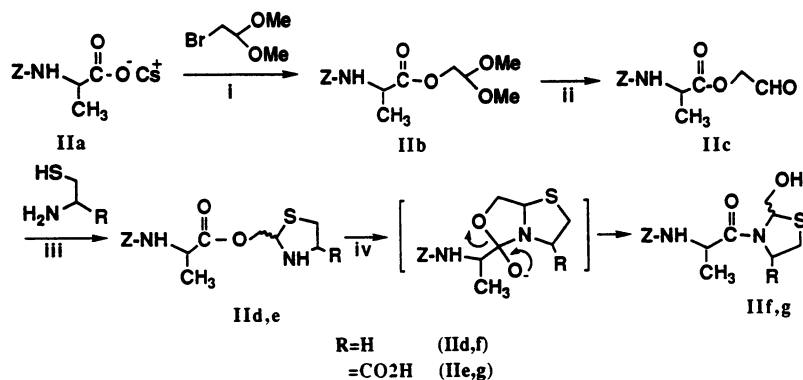


FIG. 2. Model-compound study of segment ligation strategy. Reaction conditions: (i) DMF, 60°C, 48 hr; (ii) 30% TFA in CH₂Cl₂ (2–5% H₂O); (iii) H₂O/CH₃CN, pH 2–5; (iv) pH 6–9.

Table 1. Rates of O-to-N acyl transfer reactions from peptide ester to amide

Compound	$t_{1/2}$, hr				
	pH 6	pH 7	pH 7.4	pH 8	pH 9
Z-Ala-Thz (IIId \rightarrow IIIf)	37.5	22.2	20.2	9.9	3.4
Z-Ala-Thz(CO ₂ H) (IIIe \rightarrow IIIf)	55.0	8.4	9.5	11.0	9.3

the optimal pH found for the rearrangement was 5 and the $t_{1/2}$ was \approx 1 day (unpublished results).

Both the condensation products and the rearranged products gave a mixture of two diastereoisomers due to the creation of a new asymmetric carbon at position 2 of the Thz ring. These diastereoisomers are separable by HPLC. The difference between the ester (before rearrangement) and amide (after rearrangement) forms was distinguished analytically (HPLC), spectrometrically (NMR), and chemically. In ¹H NMR studies, the prominent changes were the disappearance of the proton signals of the secondary amine protons in the Thz ring and an upfield shift of two protons on the methylene carbon linked to the oxycarbonyl which, after rearrangement, became a hydroxyl group. The ester form was susceptible to alkaline hydrolysis under saponification conditions, whereas the amide form was stable. Indeed, treatment of the ester product IIId-1 with 0.1–1 M NaOH gave Z-Ala-OH as the hydrolyzed product along with the rearrangement product within 10 min, whereas the amide form IIIe-1 was stable under the same conditions.

The rearrangement product now resembles a Z-Ala-Pro structure with a thioether linkage as the isoelectronic replacement of the methylene carbon at position 4 and a hydroxymethyl substitution on position 5 of the Pro ring. Such modification will probably not change the backbone conformation of a Pro-containing peptide chain. Thus, this pseudo Xaa-Pro bond (27) can generally be viewed as a substituent for any of the Xaa-Pro bonds present in protein sequences and further enlarges the scope of application of our strategy. Of course, it would be also desirable if the HMThz derivative could be converted to a natural Cys residue. Although the cleavage of the Thz ring poses little problem and can be accomplished with assistance of an electrophile or heavy metal ion, the removal of the N-substituted glyceryl moiety requires drastic conditions such as base or strong acid.

Aldehyde Ester Introduction onto the Carboxyl Component. To be widely applicable, our scheme must resolve the first step of aldehyde introduction whereby the glycolaldehyde ester has to be introduced to the unprotected peptide segment. With synthetic peptides, this should not pose a problem, because new resins have been developed to give such a linker functional group containing a glycolaldehyde ester on the C^α moiety. For proteins derived from recombinant DNA or natural sources, an initiation step is needed. Because we are trying to develop a chemistry that is generally applicable to protein segment synthesis using unprotected peptide fragments from all sources, we have designed a derivatization scheme as illustrated in the synthesis of a 50-residue epidermal growth factor-like peptide (Fig. 3). The key to our design is the introduction of the masked aldehyde function onto the C terminus of the first component by enzymatic synthesis (28–30). Kinetically controlled aminolysis (31) of a peptide ester by enzymes in the presence of a water-miscible organic solvent can be smoothly and efficiently accomplished if a high concentration of the amino component is used (32). The C^α-component peptide ester was synthesized by the solid-phase method (2) using a newly developed resin from our laboratory (see IIIa in Fig. 3). After the cleavage and removal of all protecting groups, this resin provided a peptide containing a C^α ester (IIIb) for an enzyme-catalyzed coupling. The enzymatic coupling between the peptide and the small

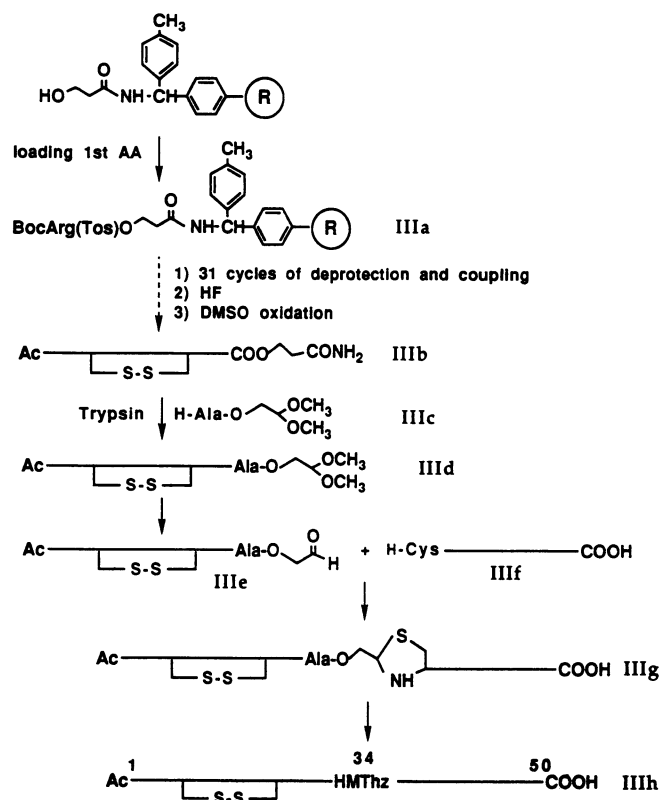


FIG. 3. Synthesis of a model 50-residue peptide. AA, amino acid; R, polystyrene resin. Sequences: IIIb, VVSHFNDCPDSHTQFEF-HGTCRFLVQEDKPAR (32 residues); IIIf, CHSGYVGARC(Ac-m)EHADLLA (17 residues); Ac-m, acetamidomethyl.

substrate, the dimethoxy ethyl ester of Ala (IIIc), was catalyzed by trypsin in 70% DMF and completed in 75 min in the presence of high molar concentrations and the large excess of the amino component (IIIc). Under this condition, no hydrolysis of other bonds susceptible to trypsin was observed. We used an N-terminally protected segment, which actually was not necessary because of the largely excessive presence of the small component. After the enzymatic coupling and purification, the ring formation reaction and all subsequent steps were carried out in a manner similar to that described for the model study. The 17-mer IIIIf was premixed with the peptide acetal IIIId (Fig. 4). Deprotection was achieved with 95% TFA containing 5% H₂O for 5 min at 20°C. The regenerated aldehyde IIIe showed a shorter retention time in C₁₈ reverse-phase HPLC. No detectable hydrolysis of the ester bond was found, as shown by HPLC (Fig. 4). After removal of TFA, the reactants were dissolved in an acetate buffer (pH 3) to allow the ring formation reaction, which was also monitored by HPLC. The high reactivity between the aldehyde and the sulfhydryl component, plus the lack of competing reactions, makes it possible to carry out the reaction in a very dilute molar concentration. This is highly desirable because the reaction between macromolecules has to be in a dilute solution due to their large molecular masses. Under our experimental conditions, the concentration of both components is about 1.5 and 5 mM. The expected Thz product formed cleanly without any detectable side reactions. The subsequent O-to-N rearrangement reaction occurred after the solution was adjusted to pH 5 with aqueous acetate buffer, which gave the 50-residue peptide with a HMThz residue at position 34.

We emphasize that all the side-chain functional groups in both peptides were unprotected, including the ϵ -amine of Lys and the β - or γ -carboxyl of Asp or Glu residues, which

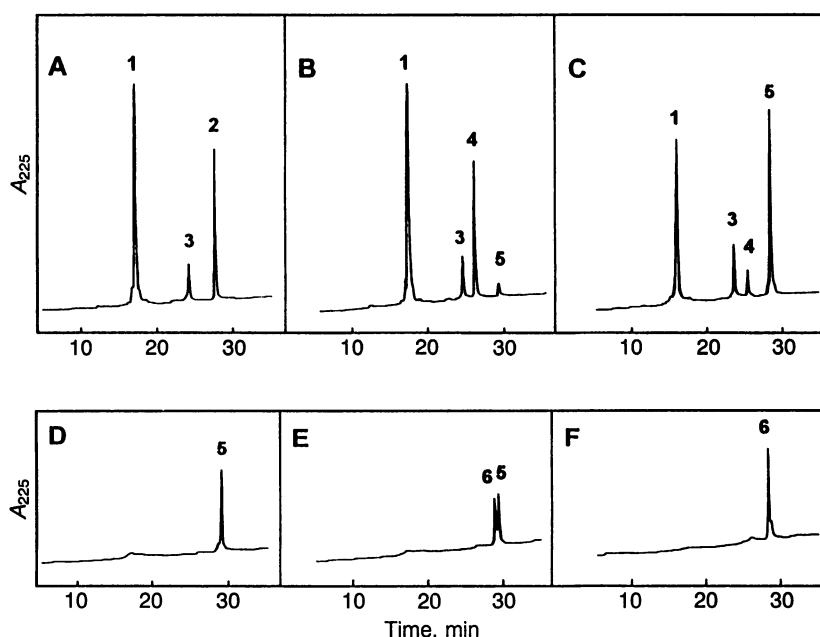


FIG. 4. Progress of the synthesis of the model 50-residue peptide by segment ligation. (A) The 17-mer (**III**f) and 33-mer (as acetal, **III**d). (B) After TFA treatment and immediate redissolution in acetate buffer (pH 3). (C) Reaction of **III**f with aldehyde **III**e after 5 hr at pH 3. (D) Purified **III**g (50-mer) in ester form. (E and F) The O-to-N acyl transfer at pH 5 after 15 hr and 3 days, respectively. Peak 1, amino component of the 17-mer **III**f in excess; peak 2, carboxyl component of the 33-mer in acetal form **III**d; peak 3, dimer formed by oxidation of **III**f during storage; peak 4, the 33-mer in aldehyde form **III**e; peak 5, the 50-mer ring-formation product in the ester form **III**g; peak 6, the O,N acyl-rearranged product **III**h. See *Experimental Procedures* for HPLC conditions.

inevitably have to be protected in the conventional segment coupling approach. The N-terminal α -amine of the first peptide does not necessarily need to be protected either, as we found in other examples. The three steps—acetal deprotection, ring formation, and acyl transfer—could be conducted in the same reaction vessel, requiring only pH changes in aqueous solution and not needing any intermediate purification steps. This has tremendously simplified the experimental procedure.

Our scheme for the segment ligation strategy described above can be generalized for the condensation between large, unprotected peptides or proteins. In general, this scheme can be divided into three steps: (i) aldehyde introduction, to incorporate a masked aldehyde ester function at the C terminus of a peptide through enzymatic or chemical coupling procedures; (ii) deprotection and ring formation, to release the masked aldehyde to react with a β -functionalized α -amino group at the N-terminus of a second peptide; and (iii) rearrangement, to form an amide bond between the two components. Due to the versatility of the enzymatic coupling used for the first step, the building blocks in this strategy are not limited to purified synthetic, disulfide-formed, and folded peptides. They can also be natural proteins. The clean, one-vessel reactions in aqueous solution provide ease of manipulation and simplification. Our method provides a useful example for the application of basic organic reactions in the synthesis of macromolecules.

We thank Dr. I. Blair and B. Nobes for the MS analysis of some of the compounds. This work was in part supported by Public Health Service Grants CA36544 and AI28701.

- Offord, R. E. (1990) in *Protein Design and the Development of New Therapeutics and Vaccines*, eds. Hook, J. B. & Poste, G. (Plenum, New York), pp. 253–282.
- Merrifield, R. B. (1963) *J. Am. Chem. Soc.* **85**, 2149–2154.
- Gutte, B. & Merrifield, R. B. (1969) *J. Am. Chem. Soc.* **91**, 501–502.
- Tam, J. P. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5409–5413.

- Du Vigneaud, V., Ressler, C., Swan, J. M., Roberts, C. W., Katsouyanis, P. G. & Gordon, S. (1953) *J. Am. Chem. Soc.* **75**, 4879–4880.
- Hirschmann, R., Nutt, R. F., Veber, D. F., Vitali, R. A., Varga, S. L., Jacob, T. A., Holly, F. W. & Denkwalter, R. C. (1969) *J. Am. Chem. Soc.* **91**, 507–508.
- Kiyama, S., Fujii, N., Yajima, H., Moriga, M. & Takagi, A. (1984) *Int. J. Pept. Protein Res.* **23**, 174–186.
- Blake, J. & Li, C. H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4055–4058.
- Blake, J. (1986) *Int. J. Pept. Protein Res.* **27**, 191–200.
- Yamashiro, D. & Li, C. H. (1988) *Int. J. Pept. Protein Res.* **31**, 322–334.
- Aimoto, S., Mizoguchi, N., Hojo, H. & Yoshimura, S. (1989) *Bull. Chem. Soc. Jpn.* **62**, 524–531.
- King, T. E., Zhao, S. W. & Lam, T. (1986) *Biochemistry* **25**, 5774–5779.
- Gaertner, H. F., Rose, K., Cotton, R., Timms, D., Camble, R. & Offord, R. E. (1992) *Bioconjugate Chem.* **3**, 262–268.
- Schnölzer, M. & Kent, S. B. H. (1992) *Science* **256**, 221–225.
- Wang, S. S., Gisin, B. F., Winter, D. P., Makofske, R., Kulesha, I. D., Tzougraki, C. & Meienhofer, J. (1977) *J. Org. Chem.* **42**, 1286–1290.
- Mitchell, A. R., Kent, S. B. H., Engelhard, M. & Merrifield, R. B. (1978) *J. Org. Chem.* **43**, 2845–2852.
- Castro, B., Dormoy, J. R., Evin, G. & Selve, C. (1975) *Tetrahedron Lett.* 1219–1222.
- Le Nguyen, D., Seyer, R., Heitz, A. & Castro, B. (1985) *J. Chem. Soc. Perkin Trans. 1*, 1025–1031.
- Tam, J. P., Wu, C.-R., Liu, W. & Zhang, J.-W. (1991) *J. Am. Chem. Soc.* **113**, 6657–6662.
- Kirby, A. J. (1980) in *Advances of Physical Organic Chemistry*, eds. Gold, V. & Bethell, D. (Academic, New York), Vol. 17, pp. 183–278.
- Fotouhi, N., Galakatos, N. G. & Kemp, D. S. (1989) *J. Org. Chem.* **54**, 2803–2817.
- Ratner, S. & Clarke, H. T. (1937) *J. Am. Chem. Soc.* **59**, 200–206.
- Tome, D. & Naulet, N. (1981) *Int. J. Pept. Protein Res.* **17**, 501–507.
- Bergmann, M., Brand, E. & Weinmann, F. (1923) *Z. Physiol. Chem.* **131**, 1–17.
- Pierce, J. S. & Lunsford, C. D. (1951) *J. Am. Chem. Soc.* **73**, 2596–2598.
- Sakakibara, S., Shin, K. H., Schneider, W. & Hess, G. P. (1962) *J. Am. Chem. Soc.* **84**, 4921–4928.
- Haack, T. & Mutter, M. (1992) *Tetrahedron Lett.* **33**, 1589–1592.
- Bergmann, M. & Fraenkel-Conrat, H. (1937) *J. Biol. Chem.* **119**, 707–720.
- Bergmann, M. & Fruton, J. S. (1938) *J. Biol. Chem.* **124**, 321–329.
- Homandberg, G. A., Mattis, J. A. & Laskowski, M., Jr. (1978) *Biochemistry* **17**, 5220–5227.
- Barbas, C. F., III, Matos, J. R., West, J. B. & Wong, C. H. (1988) *J. Am. Chem. Soc.* **110**, 5162–5166.
- Morihara, K., Oka, T. & Tsuzuki, H. (1979) *Nature (London)* **280**, 412–413.