Pergamon

Tetrahedron Letters 40 (1999) 3415–3418

TETRAHEDRON LETTERS

A Novel Method for Peptide Block Synthesis Using Unprotected Peptides

Yasushi Ishihama, Osamu Ito, Yoshiya Oda

Tsukuba Research Laboratories, Eisai Co., Ltd., 5-1-3 Tokodai, Tsukuba, Ibaraki 300-2635, Japan.

Tatsuyuki Takenawa, Masahiro Iwakura*

National Institute of Bioscience and Human-Technology, 1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan.

Received 28 December 1998; revised 3 March 1999; accepted 5 March 1999

Abstract: An S-cyanocysteine-mediated α-carbon activation reaction was used for the block synthesis of unprotected peptides. The repeated reaction using the S-cyanocysteinyl peptides as building blocks made it possible to ligate several peptide segments in order to synthesize larger peptides with natural amide bonding.

Block synthesis of unprotected peptides is a very interesting approach not only for the total synthesis of proteins,[1-7] but also for developing a combinatorial synthesis of artificial proteins. Although some approaches, such as various chemical ligation methods, have been developed for dovetailing synthetic peptide segments, the produced linkage backbone did not have natural amide bondings[1, 2, 4-7] or a Cys-free peptide could not be provided.[3] Moreover, additional complex steps were often needed for the preparation of the segment functionalities for ligation.

We demonstrate here that an S-cyanocysteine-mediated α-carbon activation is useful for combining two unprotected peptide segments with peptide bonding, and we also propose a strategy for the block synthesis using the S-cyanocysteinyl peptide as a building block.

It is well-known that the conventional cleavage reaction of an S-cyanocysteinyl peptide is caused by nucleophilic attack of a hydroxide ion on the carbonyl carbon atom of the X-cyanocysteine linkage to form the N-terminal peptide (NC-peptide) and the C-terminal peptide with a 2-iminothiazolidine-4-carbonyl amino terminal (CC-peptide), as demonstrated by Catsimpoolas et al.[8] and Jacobson et al.[9] Recently, we reported that the intramolecular attack of an ε-amino group on the activated carbonyl carbon atom was observed in the cyanocysteine-mediated cleavage of dihydrofolate reductase.[10] Based on the same mechanism, the intermolecular attack of an α-amino group would also occur and the binding of an S-cyanocysteinyl peptide with an α-amino group of an N-terminal free peptide would be possible.[10] This intermolecular reaction would provide a novel method for segment ligation of unprotected peptides with the natural peptide bonding. Moreover, this reaction would be extended to a novel synthesis through a repeated segment ligation reaction using the S-cyanocysteinyl peptide as the building block.

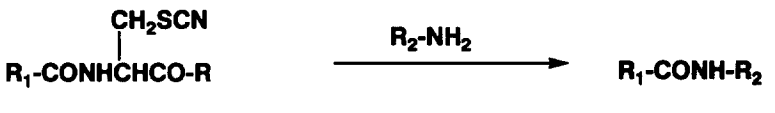
As a feasible test, the intermolecular reactions were evaluated using small synthetic peptides. The synthetic peptides employed in this work were selected by considering the minimization of the side chain effects, and moderate retention for both the preparation and analysis by HPLC (YAAGAAAGGAAYA and its segments). Angiotensin II was also employed as a natural peptide. The standard protocol of the reaction was as follows: 200 nmol of the electrophile peptide was added to 200 μL of 0.1 M borate - 0.1 M NaOH buffer (pH 9.6) containing 400 nmol of the nucleophile. The reaction mixture was kept at room temperature for three hours except as otherwise indicated. The reaction was then monitored using an on-line reversed-phase HPLC/electrospray mass spectrometer (LC/MS).[11] During the conventional cleavage reaction of the S-cyanocysteinyl peptide, it is known that the peptide containing the dehydroalanine residue is produced by the β-elimination of Cys, which often makes the reaction yield low. In this case, the segment ligation peptide (SL-peptide) as well as the β-elimination product

* Corresponding author: e-mail iwakura@nibh.go.jp, FAX (+81-298) 54-6408

0040-4039/99/$ - see front matter
PII: S0040-4039(99)00504-3

(BE-peptide), the NC-peptide and the CC-peptide were observed and the content of other by-products was less than 3 %. In Table 1, the ratio of these products as well as the residual substrate are listed.

Table 1. Segment Ligation Reaction[a)]



Entry	Electrophile	Nucleophile (2eq.)	Ratio of the products[b)]			Residual electrophile (%)
			SL	NC[a)]	BE[a)]	
1	Ac-AAKC(CN)A	intramolecular-Lys	1[c)]	0.02	0.06	ND[d)]
2	Ac-YAAGC(CN)A	none	-	1	0.31	31.7
3	Ac-YAAGC(CN)A	Gly-NH_2	1	0.02	0.22	67.2
4	Ac-YAAGC(CN)A	angiotensin II	1	1.2	0.74	29.4
5	Ac-YAAGC(CN)A	1-octylamine	1	0.42	0.16	ND
6	Ac-YAAGC(CN)A	Val	1	0.47	0.30	18.4
7	Ac-YAAGC(CN)A	Ala	1	0.52	0.16	15.7
8	Ac-YAAGC(CN)A	Arg	1	0.24	0.11	12.5
9	Ac-YAAGC(CN)A	Asn	1	5.01	2.52	40.6
10	Ac-YAAGC(CN)A	Cys	ND	ND	1[e)]	ND
11	Ac-YAAGC(CN)A	Glu	1	1.88	0.97	42.7
12	Ac-YAAGC(CN)A	Gln	1	0.81	0.31	23.5
13	Ac-YAAGC(CN)A	Gly	1	0.12	0.09	8.0
14	Ac-YAAGC(CN)A	His	1	0.41	0.21	43.3
15	Ac-YAAGC(CN)A	Ile	1	1.04	0.24	14.2
16	Ac-YAAGC(CN)A	Leu	1	1.22	0.29	16.7
17	Ac-YAAGC(CN)A	Lys	1	0.06	0.10	19.2
18	Ac-YAAGC(CN)A	Met	1	1.05	0.28	18.8
19	Ac-YAAGC(CN)A	Pro	1	0.12	0.04	3.4
20	Ac-YAAGC(CN)A	Phe	1	0.42	0.19	16.3
21	Ac-YAAGC(CN)A	Ser	1	0.62	0.18	23.5
22	Ac-YAAGC(CN)A	Thr	1	0.69	0.17	23.3
23	Ac-YAAGC(CN)A	Trp	1	0.31	0.07	8.6
24	Ac-YAAGC(CN)A	Tyr	1	0.80	0.26	16.9
25	Ac-YAAGC(CN)A	AAAGCA	1	0.03	ND	1.1[f)]
26	Ac-YAAGC(CN)	Val	1	1.06	0.10	39.8[g)]

a) The reaction mechanism was described in ref.10. Other products are shown as follows; NC-peptide: R_1-COOH, BE-peptide: R_1-CONHC(=CH_2)CO-R, CC-peptide: NHC(=NH)CHSCH_2CHCO-R

b) The ratios of the products are based on the absorbance at 210 nm.

c) The ratios of the products are calculated from the intensity of the molecular ion signals from MS.

d) Not detected.

e) The mass of the products are the same as that of the CN-free electrophile, whose content is more than 95 %.

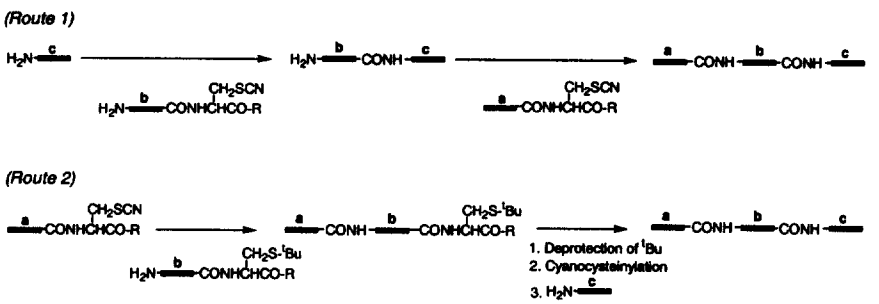
f) The major products are the CN-free electrophile (50.5%) and the dimer (30.6 %).

g) The reaction time is 1 week.

As expected, nucleophilic reactions between the S-cyanocysteinyl peptides and some nucleophiles with the amino group were achieved. The structure of the Ac-AAK lactam, which was the intramolecular reaction product, was confirmed by ^{1}H-NMR, MS and tandem MS.[12] The structures of the SL-peptides such as Ac-YAAG-angiotensin II, Ac-YAAGG-NH_2 and YAAGAAAGGAAYA were determined by the tandem MS sequencer, whereas the other SL-peptides were assigned based on the molecular weight measured by LC/MS. In some cases, the reaction was also achieved in mixed organic solvents such as DMF/diisopropylamine (3:1). As the pH of the reaction solvent increased, the reaction rate became faster and the production

of the by-products was suppressed. Considering the pK_a values of the amino group, the concentration of the nonionic form would be important for the rate and the selectivity. Other side chain functionalities of the natural amino acids, such as the phenolic group and imidazole, did not react with the electrophile (data not shown). As for the alkylamine, only the primary amine reacted while dimethylamine and trimethylamine did not react. However, proline could attack the electrophile because of the bare reactive electron pair on the nitrogen atom of the secondary amino group (entry 19). Among the natural amino acids as nucleophiles, only Cys did not react with the electrophile peptide (entry 10). The Cys-containing peptide also did not react with the electrophile peptide (entry 25). Although the reason was not clear, this might be caused by the reversibility of the addition reaction of the cyano group. However, this limitation could be overcome by using S-alkyl derivatives such as Met and S-^{t}Bu-Cys. Regarding the C-terminal of the electrophile peptide, Cys was not preferable because the rate of the reaction became lower and the selectivity of the reaction was worse (entry 26). Although N-acetyl peptides were initially used to prevent the electrophile peptides from condensing with each other, the self-ligation was suppressed when the electrophile could be surrounded by the nucleophile. Based on the result, a strategy for building block synthesis from the N-terminal to C-terminal was developed as indicated as Route 1 in Scheme 1. Alternative synthetic route with the opposite direction was also possible as indicated as Route 2, although the additional deprotection process of the ^{t}Bu group was necessary and, therefore, overall reaction yield would be lower. The typical example for ligating three peptides by Route 1 and 2 is shown in Table 2. The successive reactions in Route 1 were accomplished with moderate yields, whereas the reaction selectivity was lower in Route 2.

Scheme 1 Building block synthesis



(a, b, and c: peptide segments)

Table 2 Block synthesis of YAAG-AAAG-GAAYA

Route	Reaction[a]	Ratio of the products[b]			Residual electrophile
		SL	NC	BE	(%)
Route 1	c + b → b-c	1	0.28	0.11	30.7
	b-c + a → a-b-c	1	0.21	0.07	11.4
Route 2	a + b → a-b[c]	1	10.2	1.57	15.5
	a-b + c → a-b-c[d]	1	2.90	0.48	31.9

a) Peptide a: YAAG, peptide b: AAAG, peptide c: GAAYA
b) The reaction time is 18 hours.
c) The reaction did not include the deprotection step.
d) The reaction did not include the cyanocysteinylation step.

In this study, we developed the novel reaction for ligating unprotected peptide segments with a natural peptide bond and indicated the strategies for the block synthesis of large peptides and proteins. This reaction seems to have the potential limitation that it is difficult to control the selectivity to the various nucleophiles and that the competitive hydrolysis reaction is included. Fortunately, the hydrolysis of the electrophile could be suppressed using an excess nucleophile and few dimerization products of the electrophile were observed so that the selectivity problem could essentially be avoided by using the repeated strategy. Also, the unreacted nucleophile peptide can be recovered. Thus, approaches such as the solid synthesis would help to recover the unreacted peptides and to improve the yield. For example, a site specific immobilization at the C-terminal of the enzyme on a solid support having a free NH_2 group has been accomplished (Iwakura et al., unpublished observation). Further studies are now in progress in our laboratories.

References and Notes:

(1) Schnolzer, M.; Kent, S. B. H. *Science* **1992**, *256*, 221-225.

(2) Williams, M. J.; Muir, T. W.; Ginsberg, M. H.; Kent, S. B. H. *J. Am. Chem. Soc.* **1994**, *116*, 10797-10798.

(3) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776-779.

(4) Baca, M.; Muir, T. W.; Schnolzer, M.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 1881-1887.

(5) Rose, K. *J. Am. Chem. Soc.* **1994**, *116*, 30-33.

(6) Shao, J.; Tam, J. P. *J. Am. Chem. Soc.* **1995**, *117*, 3893-3899.

(7) Liu, C.; Rao, C.; Tam, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 307-312.

(8) Catsimpoolas, N.; Wood, J. L. *J. Biol. Chem.* **1966**, *241*, 1790-1796.

(9) Jacobson, G. R.; Schaffer, M. H.; Stark, G. R.; Vanaman, T. C. *J. Biol. Chem.* **1973**, *248*, 6583-6591.

(10) Takenawa, T.; Oda, Y.; Ishihama, Y.; Iwakura, M. *J. Biochem.* **1998**, *123*, 1137-1144.

(11) A 10 μL aliquot of the reaction mixture was introduced into the LC/MS system. The LC separation was performed with an acetonitrile gradient containing 0.1 % TFA for an ODS column (150 x 2.1 mm i.d., 5 μm). The electrospray tip was held at a potential of 5000 V. The MS system was operated with an orifice voltage of 120 V and the scanning range was from 200 to 1500 m/z using a step size of 1.0 and a dwell time of 1 ms. Tandem MS spectra for determination of the peptide sequence were obtained as the averaged sum of 20 scans from 50 to 700 using a step size of 1.0 and dwell time of 10 ms with argon as the CID gas.

(12) The ^{1}H-NMR data indicated that the Ac-AAK lactam has four peptide bonds, whereas Ac-AAK has three peptide bonds. Other data also confirmed the structure of the Ac-AAK lactam.