THE JOURNAL OF Organic Chemistry®

VOLUME 45, NUMBER 4

FEBRUARY 15, 1980

A Quantitative Evaluation of Methods for Coupling Asparagine[1]

Svetlana Mojsov,* Alexander R. Mitchell, and R. B. Merrifield

The Rockefeller University, New York, New York 10021

Received August 6, 1979

A quantitative procedure was developed to evaluate methods for coupling asparagine. Boc-asparagine was coupled to glycyloxymethylphenoxymethyl-copoly(styrene–divinylbenzene) resin under a variety of conditions, and the product, Asn-Gly, and any byproducts were cleaved from the resin with 50% trifluoroacetic acid in methylene chloride. The mixture was then separated on a sulfonated ion exchange column and quantitated by the ninhydrin reaction. Coupling by the dicyclohexylcarbodiimide (DCC) or symmetrical anhydride methods gave large amounts of the dehydration product, β-cyanoalanylglycine, and smaller amounts of β-aspartamidinoacetic acid, α-aspartylglycine, and β-aspartylglycine in addition to the desired asparaginylglycine. Activation of Boc-Asn by DCC plus hydroxybenzotriazole or by the nitrophenyl ester gave 98 to 99% of Asn-Gly, but very low levels of the byproducts were detectable with the sensitive chromatographic method. Protection of the amide function of Boc-Asn with the 4,4′-dimethoxybenzhydryl group avoided completely the formation of the nitrile during DCC coupling. Pure Ala(CN)-Gly was quantitatively reconverted to Asn-Gly by HF. The rehydration of nitrile also occurred in 50% trifluoroacetic acid in dichloromethane, but much more slowly.

Dehydration of amides to nitriles has been observed during coupling of N^{α}-protected asparagine and glutamine, both in solution and in solid-phase peptide synthesis.[2-6] The side reaction was found after coupling with dicyclohexylcarbodiimide, mixed anhydride, pyrophosphite, and other activation methods. It occurs while these amino acids are activated and not during subsequent coupling steps after asparagine or glutamine has been incorporated into the peptide chain. The proposed mechanism[3-5] for the rearrangement postulates that the nitrile compound is formed through a cyclic isoimide intermediate.

Coupling with N^{α}-protected asparagine and glutamine *p*-nitrophenyl esters was reported to circumvent the problem of nitrile formation[7] and has been used extensively for that purpose in solid-phase peptide synthesis.[8,9] However, since the reaction proceeds slowly and sometimes is not complete after prolonged periods of time,[10,11] alternative routes have been explored.

Dehydration was shown to be prevented in solution synthesis when 1-hydroxybenzotriazole was added together with dicyclohexylcarbodiimide for activation and coupling of N^{α}-protected asparagine and glutamine.[12] This method has also been applied with success in solid-phase synthesis.[13,14] Oxidation–reduction coupling of these residues is also reported[15] to proceed without dehydration. An alternative way to avoid the rearrangement has been to protect the amido function with the 4,4′-dimethoxybenzhydryl (Mbh) group.[16] However, no quantitative experimental data by highly sensitive methods have been available to establish the extent of the supression of nitrile formation under any of these conditions.

It has also been proposed[17] that use of the symmetrical anhydride of Boc-Asn or Boc-Gln will avoid formation of the cyclic isoimide intermediate. The feasibility of this coupling approach was deduced by the failure to detect the nitrile derivative of glutamine by amino acid analysis of leucine aminopeptidase digests after the final synthetic peptide was cleaved from the resin support with anhydrous HF. However, treatment of Tos-Ala(CN) with a strong

(1) Nomenclature and symbols generally follow the recommendations of the IUPAC–IUB Commission on Biochemical Nomenclature; see *J. Biol. Chem.* **1972**, *247*, 977. In addition, the following abbreviations are used: Boc-iAsn, *tert*-butyloxycarboxylisoasparagine; Tos-Ala(CN), *p*-toluenesulfonyl-β-cyanoalanine; Mbh, 4,4′-dimethoxybenzhydryl; HOBt, 1-hydroxybenzotriazole; DCU, *N,N′*-dicyclohexylurea.

(2) (a) Gish, D. T.; Katsoyannis, P. G.; Hess, G. P.; Stedman, R. J. *J. Am. Chem. Soc.* **1956**, *78*, 5954. (b) Ressler, C. *J. Am. Chem. Soc.* **1956**, *78*, 5956.

(3) Stammer, C. H. *J. Org. Chem.* **1961**, *26*, 2556.

(4) Kashelikar, D. V.; Ressler, C. *J. Am. Chem. Soc.* **1964**, *86*, 2467.

(5) Paul, R.; Kende, A. S. *J. Am. Chem. Soc.* **1964**, *86*, 4162.

(6) Hagenmaier, H.; Frank, H. *Hoppe-Seyler's Z. Physiol. Chem.* **1972**, *353*, 1973.

(7) Bodanszky, M.; du Vigneaud, V. *J. Am. Chem. Soc.* **1959**, *81*, 5688.

(8) Marshall, G. R.; Merrifield, R. B. *Biochemistry* **1965**, *4*, 2394.

(9) Merrifield, R. B. *Adv. Enzymol.* **1969**, *32*, 221.

(10) Losse, G.; Ulrich, R. *Tetrahedron* **1972**, *28*, 5832.

(11) Karlsson, S.; Lindeberg, G.; Ragnarsson, U. *Acta Chem. Scand.* **1970**, *24*, 337.

(12) König, W.; Geiger, R. *Chem. Ber.* **1970**, *103*, 788.

(13) Hruby, V. J.; Muscio, F.; Groginsky, C. M.; Gitu, P. M.; Saba, D. *J. Med. Chem.* **1973**, *16*, 624.

(14) Wolters, E. T. M.; Tesser, G. I.; Nivard, R. J. F. *J. Org. Chem.* **1974**, *39*, 3388.

(15) Mukaiyama, T.; Yeki, M.; Maruyama, H.; Matsueda, R. *J. Am. Chem. Soc.* **1968**, *90*, 4490.

(16) König, W.; Geiger, R. *Chem. Ber.* **1970**, *103*, 2041.

(17) Hemmasi, B.; Bayer, E. *Z. Physiol. Chem.* **1972**, *355*, 481.

0022-3263/80/1945-0555$01.00/0

anhydrous acid such as HBr in acetic acid has been shown to promote rehydration of the nitrile to the corresponding amide,[18] and HF can be expected to behave similarly. Therefore, it is quite possible that the conditions that were applied for coupling of Boc-Gln[17] actually did lead to the formation of the nitrile byproduct but that it was not observed because of the subsequent hydration following HF treatment.

Since it was important to us to select the most convenient and effective way to avoid this side reaction, we have studied the extent of this and other reactions occurring during the coupling of N^{α}-protected asparagine under the various conditions reported to be useful in solid-phase peptide synthesis. This has been done in a more quantitative manner than has been previously reported by using the synthesis of the dipeptide, Asn-Gly, as a model. In order to identify and quantitate the relative amounts of the byproducts, a chromatographic system has been developed in which Asn-Gly can be separated from its dehydration product Ala(CN)-Gly and several related byproducts and quantitated at levels of 0.1% or less. In most previous studies the detection of nitrile has been by fractional crystallization, thin layer chromatography, mass spectrometry, or IR. The IR band at 2260 cm^{-1} is very characteristic for C≡N, but an analytical method based on this weak band is quite insensitive. For example, Hruby et al.[13] found a lower limit of detection of about 5% in a mixture of peptides, and in much earlier work in this laboratory we had found that the lower limit of detection of C≡N in a resin-bound product was 5–10%. In addition, IR could distinguish only one of the several byproducts of the reaction.

The peptide was synthesized on a *p*-alkoxybenzyl alcohol resin support[19] because the resulting ester linkage between the peptide and the resin is more acid labile than the standard benzyl ester prepared from chloromethyl-copoly(styrene–divinylbenzene) resin. Thus, the cleavage of the peptide from the resin could be accomplished with 50% trifluoroacetic acid in dichloromethane without exposure of the peptide to HF or HBr. Therefore, the influence of the synthetic procedure on the formation of nitrile could be separated from the influence of HF on the rehydration back to the amide.

Experimental Section

Commercial starting materials were the following: chloromethyl-copoly(styrene–divinylbenzene) resin (0.75 mequiv of Cl/g, 200–400 mesh (Bio-Rad); Boc-asparagine nitrophenyl ester and isoasparagine (Chemical Dynamics); α-aspartylglycine and β-aspartylglycine (Bachem). Biphenylylisopropyloxycarbonylglycine DCHA salt was prepared according to Feinberg and Merrifield[20] and Boc-(4,4′-dimethoxybenzhydryl)asparagine was synthesized according to Geiger.[16] Other samples were synthesized for the purpose of developing a chromatographic system.

L-Asparaginylglycine. A solution of 0.220 g (0.533 mmol) of carbobenzoxy-L-asparaginylglycine benzyl ester[21a] in 65 mL of 90% aqueous methanol was hydrogenated at 50 psi in the presence of palladium (5%) on carbon for 7 h at room temperature. The catalyst was removed by filtration, and the filtrate was evaporated in vacuo. The resulting residue was washed several times with absolute ethanol and petroleum ether (bp 30–75 °C), yielding 0.076 g (75%) of the title compound, mp 205–206 °C [lit.[21b] 205–206 °C].

Boc-cyano-L-alanylglycine. Boc-cyano-L-alanine[22] was obtained by dehydration of Boc-L-asparagine with dicyclohexylcarbodiimide or isobutyl chlorocarbonate. A solution of Boc-cyano-L-alanine (1.29 g, 6.03 mmol) and triethylamine (0.84 mL, 6.03 mmol) in 30 mL of DMF was cooled to −15 °C. Isobutyl chlorocarbonate (0.78 mL, 6.00 mmol) was added, and the mixture was stirred for 15 min. A solution of glycine (0.454 g, 6.03 mmol) in 12 mL of 0.50 N aqueous sodium hydroxide was added. The reaction mixture was stirred overnight at room temperature. The slightly turbid solution was evaporated in vacuo to a residue, which was dissolved in water (20 mL), chilled in an ice bath, and acidified to pH 2.0 by addition of 3 N hydrochloric acid. The aqueous solution was extracted with three 20-mL portions of ethyl acetate. The combined extracts were washed with five 20-mL portions of saturated aqueous sodium chloride, dried over magnesium sulfate, and freed of solvent to yield an oil (1.03 g). The oil contained the title compound, R_f 0.16 in chloroform–acetic acid (9:1) on silica gel TLC plates, contaminated with Boc-cyano-L-alanine, R_f 0.41. The product was purified by preparative TLC. An oil was isolated that was homogeneous by TLC (0.291 g, 18%). Attempts to crystallize this material were unsuccessful, and the title compound was isolated as the dicyclohexylammonium salt in 44% yield (mp 176–177 °C). The Boc-cyanoalanine and Boc-cyanoalanylglycine gave characteristic weak absorption bands for the C≡N group at 2250 cm^{-1}.

Anal. Calcd for $C_{23}H_{40}N_4O_5$: C, 61.20; H, 8.86; N, 12.32. Found: C, 61.03; H, 8.91; N, 12.38.

Isoasparaginylglycine. Dicyclohexylcarbodiimide (45.9 mg, 0.22 mmol, 3 equiv) in 1 mL of CH_2Cl_2 and 1-hydroxybenzotriazole (34.3 mg, 0.22 mmol, 3 equiv) in 0.5 mL of DMF were mixed for 10 min at 0 °C. Boc-iAsn (51.9 mg, 0.22 mmol, 3 equiv) in 8 mL of DMF was added, and after 10 min at 0 °C the mixture was added to H-Gly-Res (302 mg, 0.075 mmol, 1 equiv). After 120 min at 25 °C the resin was filtered, washed, and dried in vacuo. Peptide-resin was treated with HF for 30 min at 0 °C. After evaporation of HF, the dipeptide was extracted with H_2O and lyophilized, yielding 15.5 mg of iAsn-Gly. It was observed that iAsn-Gly was completely converted by 1% aqueous triethylamine into a mixture of α-Asp-Gly and β-Asp-Gly.

Ion-Exchange Chromatography. The standard amino acid and peptide derivatives just described were used to develop a sensitive chromatographic system that would separate all of the components expected to be produced during the various syntheses of Asn-Gly and after the subsequent acid treatments. The instrument was a Beckman Model 120B amino acid analyzer equipped with an Altex 1.00-mL rotary valve manual injector. The column was 0.9 × 54 cm, packed with Beckman AA15 sulfonated cation exchange resin, thermostated at 57 °C. The buffer was prepared by adjusting 4 L of Beckman (0.2 N, pH 3.25) sodium citrate buffer to pH 2.80 with 6 N HCl (14.5 mL). The buffer flow rate was 61 mL/h, and the ninhydrin flow rate was 32 mL/h. The elution positions of amino acids and peptides were the following: Ala(CN), 55 min; β-Asp-Gly, 70 min; Asn, 103 min; Gly, 170 min; α-Asp-Gly, 185 min; Asn-Gly, 270 min; iAsn-Gly, 283 min; and Ala(CN)-Gly, 300 min.

Peptide Synthesis: Bpoc-glycyloxymethylphenoxymethyl-copoly(styrene–1% divinylbenzene)-resin (Bpoc-Gly-Res). 4-Hydroxymethylphenoxymethyl resin was prepared from chloromethyl-resin and *p*-hydroxymethylphenol according to the procedure of Wang.[19] The resin (2.02 g, 1.2 mmol of $HOCH_2$–) was esterified by treatment with Bpoc-Gly (1.13 g, 3.6 mmol) in 10 mL of CH_2Cl_2, DCC (0.75 g, 3.6 mmol) in 3.5 mL of CH_2Cl_2, and 4-dimethylaminopyridine (3.8 mg, 0.03 mmol) in 1 mL of CH_2Cl_2. The mixture was shaken for 1 h at 25 °C, filtered, and washed three times each with CH_2Cl_2, *i*-PrOH, and CH_2Cl_2. The esterification reaction was repeated one more time. Amino acid analysis after hydrolysis in 12 N HCl–HOAc–phenol (2:1:1, v/v) showed 0.60 mmol of Gly/g of resin.

Samples of the Bpoc-Gly-Res (50–700 mg) were washed three times with 0.5% TFA in CH_2Cl_2 for 1 min and deprotected by a second treatment for 20 min. Neutralization was by three 2-min treatments with 5% diisopropylethylamine in CH_2Cl_2. After thorough washing with CH_2Cl_2, *i*-PrOH, and CH_2Cl_2, the Gly-Res

(18) Zaoral, M.; Rudinger, J. *Collect. Czech. Chem. Commun.* **1959**, *24*, 1993.
(19) Wang, S. S. *J. Am. Chem. Soc.* **1973**, *95*, 1328. Wang, S. S. *J. Org. Chem.* **1975**, *40*, 1235.
(20) Feinberg, R. S.; Merrifield, R. B. *Tetrahedron* **1972**, *28*, 5865.
(21) (a) Ressler, C.; Kashelikar, D. B. *J. Am. Chem. Soc.* **1966**, *88*, 2025. (b) Miller, H. K.; Waelsch, H. *Arch. Biochem. Biophys.* **1952**, *35*, 176.
(22) Grzonka, Z.; Liberek, B. *Bull. Acad. Pol. Sci.* **1969**, *17*, 219.

Table I. Distribution of Products Formed in the Synthesis of Boc-Asn-Gly-$OCH_2C_6H_4OCH_2C_6H_4$-Res under Different Coupling Conditions

	coupling method[a]										
	A	B	C	D	E	F	G	H	I	J	K
compound	mol % of the total products										
β-Asp-Gly	0.3	0.1	0.1	0.1	1.7	1.6	tr	tr	tr	0	0
α-Asp-Gly	0.3	0.2	0.1	0.1	1.4	1.3	0.1	tr	0.2	tr	0.2
aspartoyl-Gly	tr	tr	tr	tr	tr	tr	0	0	0	0	tr
Asn-Gly	54.8	69.3	80.7	73.1	14.6	48.2	98.9	>99.8	99.1	97.9	99.8
iAsn-Gly	0	0	0	0	0	0	0	0	0	0	0
Ala(CN)-Gly	39.2	25.9	16.1	22.4	68.2	41.1	0.7	0	0.5	1.9	0
β-aspartamidinoacetic acid	5.5	4.6	3.0	4.3	14.1	7.8	0.5	0	0.2	0.2	0
coupling yield (%)[b]	96.4	72.0	98.3	62.9	76.5	94.5	99.0	78.8	99.0	96.1	93.0

[a] Details of the couplings were the following: (A) Boc-Asn:DCC:Gly-Res 4:4:1; Boc-Asn and Gly-Res equilibrated for 5 min, 25 °C, then DCC added; coupled 120 min, 25 °C. (B) Boc-Asn:DCC:Gly-Res 4:4:1; Boc-Asn and Gly-Res equilibrated for 5 min, 25 °C, then DCC added; coupled 5 min, 25 °C. (C) Boc-Asn:DCC:Gly-Res 8:8:1; Boc-Asn and DCC mixed briefly just prior to addition to Gly-Res; coupled 120 min, 25 °C, two times. (D) Boc-Asn:DCC:Gly-Res 8:8:1; Boc-Asn and DCC mixed briefly just prior to addition to Gly-Res; coupled 3 min, 25 °C, two times. (E) Boc-Asn:DCC:Gly-Res 8:8:1; Boc-Asn and DCC premixed 30 min, 0 °C; DCU filtered and reaction mixture added to Gly-Res; coupled 20 min, 25 °C. (F) Boc-Asn:DCC:Gly-Res 8:4:1; Boc-Asn and DCC premixed 30 min, 0 °C; DCU filtered and reaction mixture added to Gly-Res; coupled 20 min, 25 °C, four times. (G) Boc-Asn:DCC:HOBt:Gly-Res 4:4:4:1; Boc-Asn and HOBt mixed 10 min, 0 °C; then DCC added for additional 10 min, 0 °C, and reaction mixture added to Gly-Res; coupled 120 min, 25 °C. (H) Boc-Asn:DCC:HOBt:Gly-Res 4:4:4:1; Boc-Asn and HOBt mixed 10 min, 0 °C; then DCC added for additional 10 min, 0 °C, and reaction mixture added to Gly-Res; coupled 5 min, 25 °C. (I) Boc-Asn:DCC:HOBt:Gly-Res 4:4:4:1; DCC and HOBt mixed 10 min, 0 °C; then Boc-Asn added for additional 10 min, 0 °C, and reaction mixture added to Gly-Res; coupled 120 min, 25 °C. (J) Boc-Asn-ONp:Gly-Res 8:1; coupled 16 h, 25 °C. (K) Boc-Asn(Mbh):DCC:Gly-Res 4:4:1; Boc-Asn(Mbh) and Gly-Res equilibrated 5 min, then DCC added; coupled 60 min, two times. [b] Based on the amount of uncoupled Gly-Res.

was coupled (1) with Boc-Asn by activation with DCC, symmetrical anhydride, or DCC plus 1-hydroxybenzotriazole, (2) with Boc-Asn(Mbh) plus DCC, or (3) with Boc-asparagine nitrophenyl ester, as described under Table I. The peptide products were simultaneously deprotected and cleaved from the resin in TFA–CH_2Cl_2 (1:1) for 40 min, 25 °C.

Peptide Analysis. Samples were analyzed for product distribution by the chromatographic procedure. They were usually run twice—once heavily overloaded, with the major component at 1–3 μmol/mL to allow detection of 0.1% or less of the minor components, and a second time with the major component at approximately 100 nmol/mL. The ninhydrin peaks were identified by comparison with standards (Figure 1) and were quantitated by correcting for color constants that were determined for weighed samples of pure standards. The distribution of products is shown in Table I.

In some instances the column eluates were diverted to a fraction collector before passing through the ninhydrin reaction coil. The various components were located by the fluorescamine spray procedure.[23] The homogeneity of the isolated components was then verified by rechromatography using the ninhydrin system. A small amount of the material from the peak at 360 min was also isolated in this way and shown to be homogeneous by rechromatography. Its identification as β-aspartamidinoacetic acid is tentatively based on the following information: (1) the peptide eluted at a position expected for a compound more basic than Ala(CN)-Gly or Asn-Gly, (2) acid hydrolysis gave equal amounts of Asp and Gly, (3) ^1H, ^{13}C, and ^{15}N NMR data (to be reported separately) were all consistent with the amidine structure.

The position of aspartoylglycine on the column was determined indirectly.[24] A sample of Boc-Asp(OBzl)-Gly-resin was synthesized and cleaved in HF, and the peptide mixture was extracted into water. Part of the sample was separated on the column just described and showed a major α-Asp-Gly peak, a minor β-Asp-Gly peak, and a third component eluting at 205 min. A second part of the sample was treated for 10 min with 0.1 N NaOH, neutralized with 0.1 N HCl, and separated on the same column. The chromatogram showed the absence of the 205-min peak and a corresponding increase in the β-Asp-Gly peak, with very little change in the α-Asp-Gly peak as expected for an aspartimide after treatment with base.

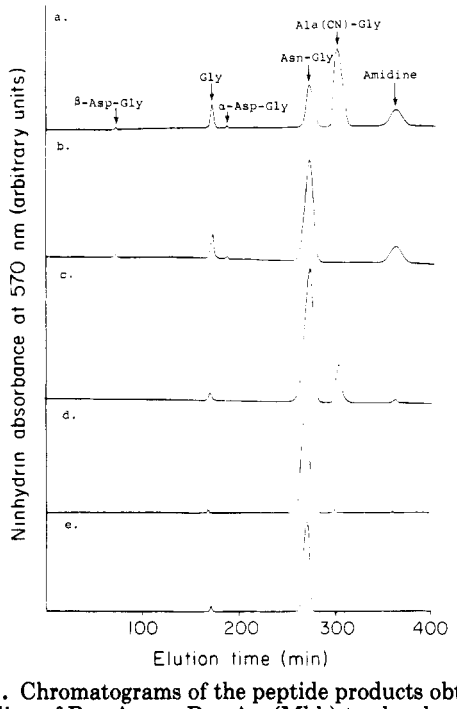


Figure 1. Chromatograms of the peptide products obtained in the coupling of Boc-Asn or Boc-Asn(Mbh) to glycyloxymethylphenoxymethyl-resin under a variety of conditions. (Details of the synthesis are given in Table I.) (a and b) Preformed symmetrical anhydride (run E); (a) peptide-resin treated with 50% TFA/CH_2Cl_2 for 40 min at 25 °C; (b) peptide-resin treated with HF for 30 min at 0 °C; (c) DCC mediated coupling (run A); (d) DCC plus 1-hydroxybenzotriazole (run G); (e) Boc-Asn (Mbh) activated with DCC (run K). The relative areas of the peaks do not represent the relative amounts of the peptide components present in the sample because they have not been corrected for the differences in their color constants.

Rehydration of Ala(CN)-Gly to Asn-Gly. (a) Ala(CN)-Gly was obtained by ion-exchange chromatography and examined in

(23) Lai, C. Y. *Methods Enzymol.* **1977**, *52*, 240–241.
(24) Yang, C. C.; Merrifield, R. B. *J. Org. Chem.* **1976**, *41*, 1032.

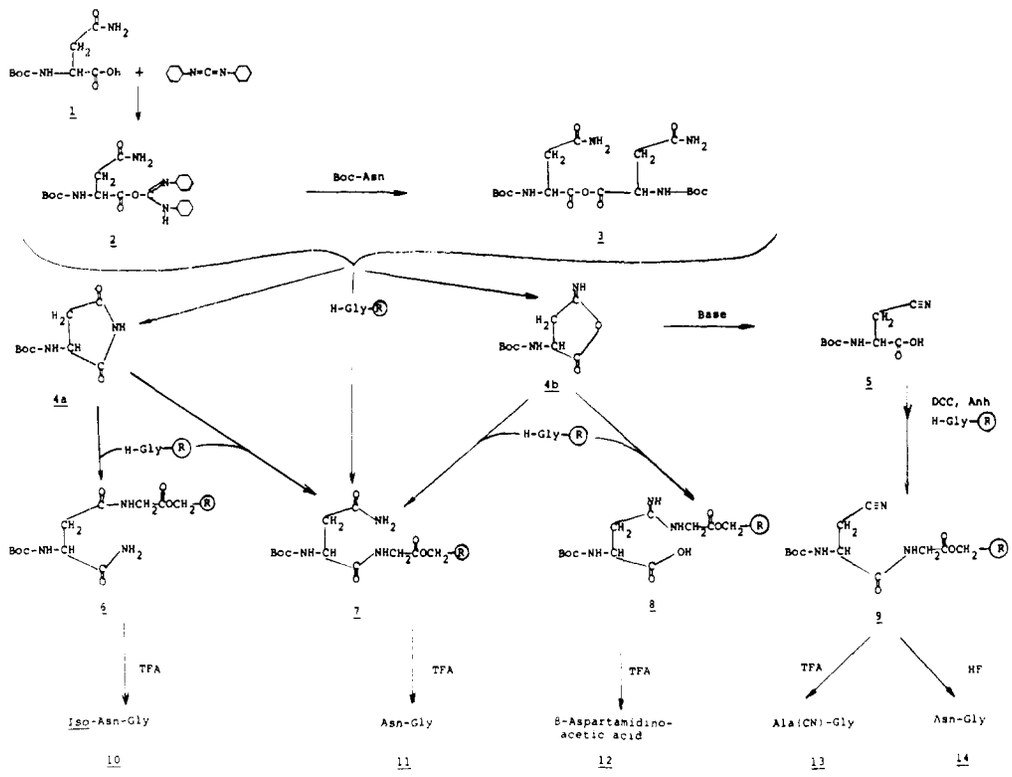


Figure 2. Routes to byproducts formed during DCC or symmetrical anhydride activation and coupling of Boc-Asn to glycyloxymethylphenoxymethyl-resin.

the following series of experiments. A 0.56 mmol sample was dissolved in 3 mL of water, and five 0.5-mL aliquots were taken to dryness and treated as follows: 1, control, no treatment; 2, treated with 1 mL of anhydrous HF, 30 min, 0 °C; 3, treated with 1 mL of anhydrous TFA, 30 min, 25 °C; 4, treated with 1 mL of anhydrous TFA, 29 h, 25 °C; 5, treated with 1 mL of TFA–CH_2Cl_2 (1:1, v/v), 40 min, 25 °C. The samples were evaporated, dissolved in pH 2.80 citrate buffer, and run on the ion-exchange column described above.

(b) Samples of Boc-Ala(CN)-Gly-resin (10–12 mg) were treated with TFA–CH_2Cl_2 (1:1, v/v) at 25 °C for periods of 30 min to 16 days and analyzed chromatographically.

Results and Discussion

The model used here was designed to allow us to distinguish between the products actually formed on the resin during the coupling reaction and those obtained at the end of the synthesis after HF cleavage. The results of this study, summarized in Table I, show in a quantitative manner that the introduction of α-amino-protected asparagine into a peptide can proceed without formation of appreciable amounts of byproducts when the activation is carried out under appropriate conditions.

Coupling of Boc-Asn by DCC activation or as the preformed symmetrical anhydride generated by the DCC reaction gave large amounts of the resin-bound nitrile and amidine derivatives (runs A through F). For identification and quantitation of these peptide products, they were cleaved from the resin with TFA–CH_2Cl_2 (1:1), under conditions where they had been shown to be stable, and then separated chromatographically. The close similarity in the nature and quantity of the byproducts identified in the synthesis of the dipeptide, Asn-Gly, by these two methods suggests that the same reaction mechanism operates in each. The data can be explained[3-5] by the intermediate formation of aspartylisoimide (**4b**) (Figure 2), which could then react in one of several ways: (1) opening of the ring by a base to give Boc-Ala(CN)-OH (**5**), which could then become reactivated and couple to the Gly-Res to give Boc-Ala(CN)-Gly-Res (**9**); (2) the free amino group from the Gly-Res could attack at the α-carbonyl, giving the desired Boc-Asn-Gly-Res (**7**); (3) the free amino group from the Gly-Res could attack at the β-carbonyl, giving the amidine derivative **8**. Similar attack of Gly-Res on Boc-aspartimide (**4a**) would yield Boc-Asn-Gly-Res (**7**) and Boc-iAsn-Gly-Res (**6**), although the latter was not detected. The traces of aspartoylglycine observed may have been formed through the acid-catalyzed cyclization of **6** or **7**,[25] and subsequent ring opening would give the small amounts of α-Asp-Gly and β-Asp-Gly that were found. Boc-Asn-Gly-Res could also be formed in part through the direct reaction of **2** or **3** with Gly-Res.

From this mechanism one would not expect that a shorter coupling time would significantly reduce or alter the proportion of the side reactions,[26] and our experiments showed that the distribution of the nitrile and other byproducts was little changed by reducing the normal DCC coupling time from 120 min to 5 min or reducing the symmetrical anhydride coupling time from 120 min to 3 min. It should be pointed out, however, that the short reaction time with Boc-Asn activated in either way markedly reduced the coupling rate and final yield.

When Boc-Asn anhydride was formed by premixing Boc-Asn and DCC for 30 min in the absence of the amine component, the amounts of byproducts formed after 20-min coupling were increased, and the yield was reduced. For example, it was necessary to perform four coupling reactions in order to achieve a 95% yield (run F, Table I). The low coupling rate is presumably due to the formation of additional amounts of the less reactive isoimide intermediate **4b** (Figure 2) before the addition of the amine. The proportion of byproducts was also enhanced when DCC was in excess (equimolar with Boc-Asn). The DCC

(25) Bornstein, P.; Balian, G. *Methods Enzymol.* **1977**, *47*, 132–145.
(26) Corley, L.; Sachs, D. H.; Anfinsen, C. B. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 1353.

Table II. Effect of Acids on the Rehydration of Ala(CN)-Gly[a] to Asn-Gly

acid	temp, °C	time, min	products, %[b] Ala(CN)-Gly	products, %[b] Asn-Gly
none			100	0
HF	0	30	0	99
TFA	25	30	87	13
TFA	25	1740	26	68
TFA/CH_2Cl_2 (1:1)	25	40	100	0

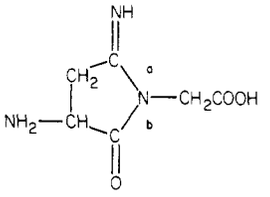
[a] Pure Ala(CN)-Gly was isolated from an ion-exchange column. After treatment with acid, the samples were evaporated to dryness in vacuo and dissolved in aqueous buffer for chromatographic analysis. [b] (Moles of product/mole of starting material) × 100. The ninhydrin values were corrected for color constants.

can serve as a base to promote nitrile **5** formation, and in addition it can activate the product to give **9** in a competing reaction with activated intermediates **2**, **3**, and **4b**. Even if DCC is not in excess, it is clear that **5** can be activated by the various active species in the coupling mixture and lead to the incorporation of nitrile into the peptide.

The product distribution was dramatically changed when activation was with the nitrophenyl ester (run J) or with DCC and 1-hydroxybenzotriazole (run G). Under these conditions the yield of the desired Asn-Gly product rose to 98–99%. However, the sensitive chromatographic method used here was able to detect the presence of 0.5–2% of nitrile and 0.1–0.3% of amidine and Asp-Gly. These low levels of byproducts have not been reported before and could not have been detected by the methods previously applied to the problem. The order of addition of reactants in the DCC/HOBt coupling made no significant difference (run I). Blocking of the amide function of Boc-Asn with the dimethoxybenzhydryl (Mbh) group (run K) provided essentially complete protection from the side reactions under study and gave a product of 99.8% homogeneity in high yield.

It was important for us to know whether or not Ala(CN)-Gly could be rehydrated to Asn-Gly by treatment with various acids followed by aqueous workup and whether such treatment gives quantitative conversion to a single product or to a mixture. The treatment of pure Ala(CN)-Gly with the strong acid HF led to its quantitative rehydration to the desired Asn-Gly dipeptide (Table II). The nitrile was much more stable to trifluoroacetic acid. In TFA–CH_2Cl_2 (1:1) no Asn-Gly was formed in 40 min. Even in anhydrous TFA only 13% of Asn-Gly was formed in 30 min and 68% in 1740 min, with 87% and 26% of the starting material being recovered.

When samples of Boc-Ala(CN)-Gly-resin were treated with TFA in methylene chloride (1:1 v/v) at 25 °C for 40 min, Ala(CN)-Gly was quantitatively released, and again no Asn-Gly could be detected. After 24 h it was found that the nitrile was converted into a roughly equimolar mixture of Asn-Gly (3.7%) and β-aspartamidinoacetic acid (3.0%). After 16 days these values rose to 28% and 13%, respectively. The proportion of rehydration products therefore is quite different from that found after HF treatment. The appearance of amidine after extended treatment of Boc-β-cyano-Ala-Gly-Res with TFA suggests that the rehydration reaction proceeds by way of a ring closure to an intermediate such as **15**. During workup hydrolysis at bond a would give Asn-Gly and at bond b could give amidine. The reaction is probably sequence dependent, and the extent of the rehydration in HF may be exaggerated



15

in the case of Asn-Gly, which is known to form cyclic compounds readily under acidic conditions.

Since treatment with 50% TFA–methylene chloride, under the conditions used for cleavage of the ester bond to *p*-alkoxybenzyl-resin, did not modify Ala(CN)-Gly, we conclude that the products cleaved and worked up in this way give an accurate picture of the resin-bound peptides that were actually formed in the coupling reaction. The data reported here show that in the case of Boc-Asn the nitrile does form during DCC or anhydride coupling but will not be observed if the peptide is cleaved in HF. The previous work[17] in which the unexpected conclusion was drawn that coupling of Boc-Gln symmetrical anhydride does not produce the nitrile might be explained in a similar way.

It would be better to avoid the nitrile formation entirely rather than to depend on acid-catalyzed regeneration of the amide, because side reactions during the coupling that lead to formation of nitrile also give rise to the amidine byproduct which was shown to be stable under the conditions of the acid-catalyzed regeneration of the nitrile to amide. The data from the Boc-Asn-ONp coupling reaction confirm the original findings[7] that activation with *p*-nitrophenyl esters largely avoids nitrile formation and justifies the early use of this procedure in solid-phase synthesis.[8,9] However, the couplings are relatively slow, and the ester must be freed of contaminating nitrile before it is used for a synthesis. The low level of nitrile found in this work may well have been produced during the preparation of the Boc-Asn-ONp rather than during the coupling reaction. Blocking of the amide nitrogen with the Mbh group[16] was also very effective in preventing nitrile formation and other side reactions. This alternative approach was fully satisfactory in the experiments reported; however, in our hands it has not always given high coupling yields, and it also requires the preparation of an additional protected amino acid. The third synthetic method that effectively avoided the nitrile side reaction was the DCC/HOBt approach of König and Geiger.[12] Our data agree entirely with their findings and conclusions, but extend them in a quantitative way. The chromatographic system used here is considerably more sensitive than any of the methods applied previously to detect and quantitate the products of the reactions.[12-14] We conclude that the DCC/HOBt coupling reaction is very fast and virtually quantitative, with only small amounts of byproducts being produced. In addition, this method of activation can be applied when the acid-labile Bpoc group is used for N^{α} protection. Although 1-hydroxybenzotriazole is acidic (pK_a = 4.2), at concentrations of 0.05 M, which are normally employed for the coupling reaction, it caused less than 0.2% deprotection of the Bpoc group. For these reasons we prefer the DCC/HOBt approach for the solid-phase synthesis of asparaginyl and glutaminyl peptides.

Acknowledgment. We are especially indebted to Dr. George Barany for his help and advice during this work and to Drs. David Live and David Cowburn and Mr. William Wittbold for the NMR analyses. This work was supported in part by Grants AM 01260 and AM 24039 from the U.S. Public Health Service and by a grant from

the Hoffmann-La Roche Foundation.

Registry No. L-Asparaginylglycine, 67576-72-1; carbobenzoxy-L-asparaginylglycine benzyl ester, 20902-64-1; Boc-cyano-L-alanylglycine, 72378-46-2; Boc-cyano-L-alanine, 45159-34-0; glycine, 56-40-6; isoasparaginylglycine, 72378-47-3; Boc-*i*-Asn, 72390-11-5; Boc-Gly, 23650-19-3; Boc-Asn, 7536-55-2; Boc-Asn(Mbh), 72378-48-4; β-aspartamidinoacetic acid, 72378-49-5; β-Asp-Gly, 3790-52-1; α-Asp-Gly, 3790-51-0; aspartoyl-Gly, 72378-50-8; Ala(en)-Gly, 72378-51-9; β-aspartimidinoacetic acid, 72378-52-0; HF, 7664-39-3; TFA, 76-05-1.

Cambridge Data File in Organic Chemistry. Applications to Transition-State Structure, Conformational Analysis, and Structure/Activity Studies

Stephen R. Wilson* and John C. Huffman

Department of Chemistry and the Molecular Structure Center, Indiana University, Bloomington, Indiana 47405

Received July 23, 1979

A description of the Cambridge Crystallographic Data File (CCDF) is illustrated with applications to organic chemistry. The use of chemical substructure searching techniques for the retrieval of three-dimensional coordinates for the more than 24 000 published organic crystal structures allows for inexpensive computer graphics and molecular modeling. Simple editing procedures and readily available programs from the X-ray crystallography literature allow the generation of transition-state models, the examination of bond lengths, etc., and direct molecular comparisons by least-squares fitting. The CCDF is a data base with a vast store of molecular information readily available to the organic chemist.

X-ray crystallography is widely recognized by organic chemists as the single most powerful structural tool currently available. A wealth of structural information can be obtained from a single-crystal X-ray diffraction experiment. However, once the stereochemical or other structural detail of interest to the organic chemist is established, the crystal structure itself is usually relegated to the archives. Little use of crystallographic information is made by most organic chemists despite the long and successful experience of biochemists, who use X-ray crystal structure data as a starting point for a multitude of structural studies.[1]

Part of the problem has been that structural data of the type useful to organic chemists is not readily accessible. Manual searches for suitable model compounds are laborious to impossible. A convenient method for *substructure searching*[2] is necessary in order to find molecular fragments of interest which may be buried in more complex molecules. Such techniques are now available. The Cambridge Crystallographic Data File (CCDF)[3-5] is a computer-readable data base containing the more than 24 124 published organic crystal structures. Substructure search techniques can be used to access both bibliographic information and molecular coordinates. A vast amount of data *as yet unrevealed* is readily accessible. We now discuss some uses of the CCDF in organic chemistry.

Molecular Models. Organic chemists have always been model builders.[6] The importance of molecular models to the organic chemist is particularly evident since Barton's contributions to conformation analysis[7] only 30 years ago.[8] How often has one seen in the literature rationalization of results "by inspection of models"?[9] The CCDF is a unique source of accurate geometric data for the easy construction by computer of molecular models. Although many computer techniques for building structures exist,[10] most programs are too complex for the average organic chemist. The CCDF allows rapid access to "real" molecules and rapid generation of stereoscopic pairs with essentially no prior knowledge of computers.[3] Moreover, many programs are readily available to obtain hard copies

(1) Gurd, F. R. N.; Rothgeb, T. M. *Adv. Protein Chem.*, in press.

(2) A method of substructure searching using Wisswesser line notation (WLN) is available from ISI. Granito, C. E.; Rosenberg, M. D. *J. Chem. Doc.* **1971**, *11*, 251. See also: Graf, W.; Kaindl, H. K.; Kniess, H.; Schmidt, B.; Warszawski, R. *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 51.

(3) Available on-line through the NIH-EPA Chemical Information System, Interactive Sciences, Corp., 918 16th St. NW, Suite 500, Washington, DC 20006. The most recent update (July 1979) contains 24 124 compounds. The Cambridge Crystallographic Data Centre (Cambridge, England) maintains the file and makes every effort to obtain the data for any X-ray crystal structure referred to in the literature. Authors should expect to be contacted regarding X-ray data and should make every effort to cooperate by providing such data to the Centre. Every year, the NIH leases the crystal data base from the Crystal Data Centre in Cambridge, England. This lease is on behalf of the entire U.S. and permits the NIH to distribute the data base within the U.S. If you would like to obtain a copy of the Cambridge Crystal Data Base on tape, please contact Fein-Marquart Associates, 7215 York Rd, Towson, MD 21212.

(4) (a) Kennard, O.; Allen, F. H.; Brice, M. D.; Hummelink, T. W. A.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Pure Appl. Chem.* **1977**, *49*, 1807. (b) Murray-Rust, P.; Motherwell, S. *Acta Crystallogr., Sect. B* **1978**, *34*, 2518. (c) Murray-Rust, P.; Bland, R. *Ibid.* **1978**, *34*, 2527. (d) Murray-Rust, P.; Motherwell, S. *Ibid.* **1978**, *34*, 2534.

(5) "Molecular Structure and Dimensions, Guide to the Literature 1935-1976. Organic and Organometallic Crystal Structures"; Kennard, O., Allen, F. H., Watson, D. G., Eds.; Cambridge Crystallographic Data Center: Cambridge, England, 1977.

(6) Perhaps the earliest molecular models were constructed of sticks and croquet balls by A. W. Hofmann (*Proc. R. Inst. G.B.* **1865**, *4*, 421) and used in a lecture delivered before the Royal Institution in 1865. This reference (found in: Ihde, A. J. "The Development of Modern Chemistry"; Harper and Row: New York, 1964) was kindly provided by E. E. Campaigne.

(7) (a) Barton, D. H. R. *J. Chem. Soc.* **1953**, 1027. (b) Construction of "Barton Models" was described [Barton, D. H. R. *Chem. Ind. (London)* **1956**, 1136], and photographs of such models appeared in the literature: Nare, H. R.; Turner, R. B. *J. Am. Chem. Soc.* **1953**, *75*, 4063.

(8) One point which should be emphasized is that not only are most "common" molecular fragments present in the CCDF but also many unusual and unique fragments are present. This is, of course, true because molecules which are unusual or unique are the ones which are more likely candidates for single-crystal X-ray studies.

(9) Fieser, L. F.; Fieser, M. "Steroids"; Reinhold: New York, 1959; p 7 and a multitude of papers.

(10) Wipke, W. T.; Heller, S. R.; Feldmann, R. J.; Hyde, E. "Computer Representation and Manipulation of Chemical Information"; Wiley: New York, 1974.

0022-3263/80/1945-0560$01.00/0