Acta Chemica Scandinavica B 37 (1983) 911–916

Synthesis of Strombine. A New Method for Monocarboxymethylation of Primary Amines

JAN KIHLBERG, ROLF BERGMAN and BÖRJE WICKBERG

Organic Chemistry 2, Chemical Center, The Lund Institute of Technology, Box 740, S-220 07 Lund 7, Sweden

Many common methods for monocarboxymethylation of primary amines give various amounts of dialkylated by-products. In this work the reaction of two equivalents of glyoxylic acid with representative primary aliphatic and aromatic amines, as well as with amino acids and a dipeptide, is shown to give only the *N*-(carboxymethyl)-*N*-formyl derivative of the amine under mild conditions in carboxylic acid solvents. Hydrolysis then produces the monocarboxymethylated primary amine in good to excellent overall yield. Proof that the intermediate product is not obtained *via* the Leuckart reaction is given.

Some monocarboxymethylated amino acids and dipeptides have interesting biological functions. In 1975, strombine *3b* was isolated from the clam *Strombus gigas* by Sangster *et al.* and was found to induce small fish to display so-called "exploratory feeding behaviour" in dilutions down to 10^{-8} g/l.[1,2] Recently, a series of substituted *N*-(carboxymethyl)dipeptides has been prepared and demonstrated to be active as inhibitors of angiotensin converting enzyme, which is involved in the control of blood pressure.[3]

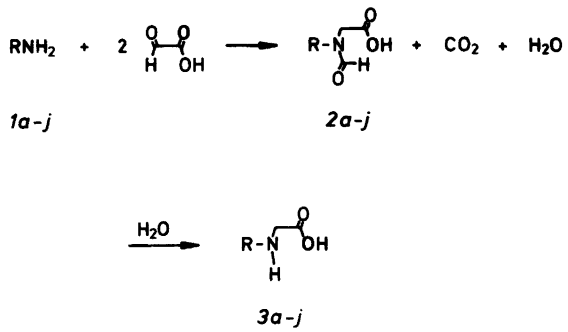
In an attempt to synthesize strombine, we found that alkylation of L-alanine with chloroacetic acid according to Sangster *et al.*[1] produced a very poor yield. This route to *N*-(carboxymethyl)amino acids, as well as their properties and their incorporation into peptides, has recently been extensively studied by Miyazawa.[4]

Seeking an alternative route, we considered a reductive amination of alanine with glyoxylic acid by conventional methods. In an attempted Leuckart reaction with alanine *1b* and glyoxylic acid in 98–100 % formic acid we were surprised to observe evolution of carbon dioxide even at room temperature, since the Leuckart reaction usually requires heating at above 150 °C.[5] The reaction product was *N*-formylstrombine *2b*, which was readily hydrolyzed to give strombine *3b* in a good overall yield. (*cf.* Scheme 1).

In the beginning of this century, Erlenmeyer and co-workers[6] reported a similar reaction between ammonia or $(NH_4)_2CO_3$ and α-oxoacids such as glyoxylic, pyruvic and phenylpyruvic acid, which afforded the corresponding *N*-acylated amino acids, in analogy with the first step of Scheme 1. It thus appeared that our case might represent an extension of the Erlenmeyer procedure, and we therefore set out to study its scope and limitations as a method for monocarboxymethylation of primary amines.

Normal Leuckart conditions usually afford a mixture of mono- and dialkylated products. The monoalkylated products become partly formylated by reaction with excess formic acid. On the other hand, the glyoxylic acid reaction studied here consistently yielded the monoalkylated formamide derivatives *2* as the first isolable products. Furthermore, at the relatively low temperatures employed in the present procedure no formylation of Leuckart monoalkylation products *3* with formic acid was found to occur. It is also significant that the reaction took the same general course in a variety of solvents, such as formic, acetic and trifluoroacetic acid, as well as in *N*,*N*-dimethylformamide. This suggests that the carbon dioxide evolved (and the formyl

0302-4369/83 $2.50

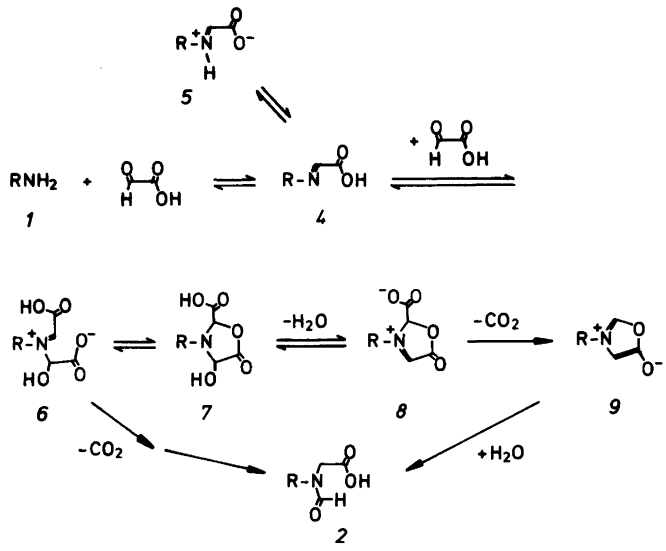


Scheme 1. R= *a*: CH_2CO_2H, *b*: CH_3CHCO_2H, *c*: $PhCHCO_2H$, *d*: $H_2N(CH_2)_4CHCO_2H$, *e*: $CH_3CH_2CH_2$, *f*: $(CH_3)_2CH$, *g*: $(CH_3)_3C$, *h*: *p*-$PhCO_2H$, *i*: *p*-$PhNO_2$, *j*: $-CH(CH_3)CONHCH(CH_3)CO_2H$.

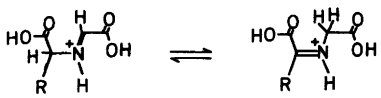
residue on the amine) originates from glyoxylic acid, and not from the formic acid solvent as in the Leuckart reaction. A mechanism that accommodates our observations is proposed in Scheme 2.

Increasing the water content of the carboxylic acid solvent to 20 % strongly retards the reaction, in accord with the first steps of the mechanism leading to the Schiff's base *4*. That a more acidic solvent decreases the rate of reaction by protonating the amine, is demonstrated by the primary aliphatic amines. Reaction in acetic acid at 40 °C is complete within a few hours, whereas use of formic acid requires more drastic conditions (*cf.* Table 1). Unfortunately, decarboxylation of glyoxylic acid giving formaldehyde (proposed to occur through intermediates *4* and/or *5* [7]) occurs to an extent of approximately 50 % in acetic acid, but is suppressed in formic acid. This dependence of decarboxylation on acidity has been described by von Euler *et al.*[7] Analogously, trifluoroacetic acid has to be used instead of formic acid in conjunction with the less basic aromatic amines. Nonacidic solvents seem to enhance by-product formation even more.

Racemization of an optically active amino acid might occur *via* the equilibrium in Scheme 3. However, when products *3* from DL- and L-



Scheme 2.



Scheme 3.

alanine were derivatized with L-proline and analyzed by GLC[8], neither racemization (optical purity >97 %) nor transamination, or aldol condensation with glyoxylic acid was observed.

The reaction of the Schiff's base *4* with a second molecule of glyoxylic acid may conceivably occur *via* alternative routes, *e.g.* by direct attack of the imino nitrogen of *4* at the formyl group of the acid as indicated in Scheme 2, or by a preceding addition of the carboxyl group over the C=N double bond of *4* followed by a ring closure to *7*. The end product *2* could be formed by decarboxylation of the open chain isomer *6* but the latter would be expected to cyclize very readily to the hydroxyoxazolidone *7*, which in turn might be in equilibrium with the dehydration product *8*. This should be prone to decarboxylate, forming the mesoionic oxazolone derivative *9*. Intermediates of this type have been invoked in the Dakin-West reaction[9] and the racemization of *N*-acylamino acids. Spontaneous hydrolysis of the oxazolone would then yield the end product *2*.

The derivatives *2* were obtained in better than 80 % yield (NMR), but did not crystallize,[4b] with the exception of the *p*-aminobenzoic acid derivative. The split formyl proton signals of these products arise from amide rotamers[10] (*cf.* Table 1).

While no problems were encountered in the acid hydrolysis of the formamides derived from the amino acids *2a–d*, the aliphatic amines *2e–g* or the dipeptide *2j*, hydrolysis of the aromatic formamides *2h,i* required more care. For example, hydrolysis of the *p*-aminobenzoic acid derivative under acidic catalysis led to decarboxylation of the aromatic carboxyl group. Consequently, alkaline hydrolysis was used for these two derivatives.[11]

That the difference in reactivity between primary aliphatic amines and α-amino acids (*cf.* Table 1) allows selectivity is demonstrated by the case of lysine monohydrochloride, which gave N^2-(carboxymethyl)lysine in 62 % overall yield under α-amino acid conditions. The position of reaction was assigned from the ^{1}H NMR spectra of N^2-(carboxymethyl)-N^2-formyllysine *2d*, where the C^2-hydrogen is shifted 0.65 ppm downfield relative to lysine whereas the C^6 hydrogens overlap for the two compounds. The assignment was corroborated by ^{13}C NMR (see Experimental part).

In conclusion, the glyoxylic acid reaction appears to be a generally useful method for monocarboxymethylation of primary amines, amino acids and peptides.

EXPERIMENTAL

^{1}H NMR spectra were run on either a Jeol JNM-PMX60 instrument at 60 MHz or on a Nicolet WB360 instrument at 360 MHz. TMS or 3-trimethylsilylpropanesulfonic acid (TriMS) were used as internal references. Mass spectra were recorded on a single-focusing, non-commercial instrument at 70 eV ionization potential. Melting points were determined with a Mettler FP1 instrument (temperature increase 2 °/min until melting begins), and are uncorrected.

Reactions were followed with TLC and/or HPLC. The following TLC-systems were used;

TLC 1: EtOAc–HOAc–H_2O= 2:2:1 on silica gel

TLC 2: EtOAc–HOAc–H_2O= 2:2:1 on cellulose

TLC 3: EtOAc–HOAc–HCO_2H–H_2O= 18:3:1:4 on silica gel

TLC 4: EtOAc–HOAc= 19:1 on silica gel

HPLC was run on a C-18 column (Merck LiChrosorb 5 μm silanized with octadecyldimethylchlorosilane) with UV-detection at 205 nm. A 20 mM phosphate buffer (pH=5.5) made 20 mM with $(C_4H_9)_4N^+HSO_4^-$ mixed with various amounts of methanol was used as eluent (flow rate: 1 ml/min). The following methanol concentrations were employed: HPLC 1: 5 %, HPLC 2: 25 %, HPLC 3: 32.5 %, HPLC 4: 40 %. The systems used for each amine are specified under the corresponding product in this section.

Amberlite CG-120 (200–400 mesh, H-form) cation exchanger was used for separations.

General synthetic procedure. (Detailed experimental conditions with the specific starting amines *1a–j* are given in Table 1.) To the

Table 1. Experimental conditions, yields and NMR shifts of formyl protons in *2a–j*.

Starting amine	Step 1 Solvent[a]	Temp./ °C	Time/ h	Step 2 Method of hydrolysis[b]	Time/ h	Product (% isolated yield)	NMR shifts of formyl proton[c]
Glycine (*1a*)	A	40	2.5	C	5	*3a*[12] (64)	8.17[d]
L-Alanine (*1b*)	A	40	4	C	5	*3b*[1,4a] (69)	8.27, 8.13[d]
D-Phenylglycine (*1c*)	A	40	4	C	5	*3c*[4a] (77)	8.18, 8.08[d]
L-Lysine HCl (*1d*)	A	40	4	C	5	*3d*[13] (62)	8.05, 8.00[d]
Propylamine (*1e*)	A	70	14	C	5	*3e*[14] (93)	8.18, 8.12[e]
Isopropylamine (1f)	A	70	14	C	5	*3f*[14] (90)	8.22, 8.13[e]
tert-Butylamine (*1g*)	A	70	14	C	5	*3g*[14] (77)	8.53[e]
p-Aminobenzoic acid (*1h*)	B	25	2.5	D	4	*3h*[15] (65)	8.80, 8.43[f]
p-Nitroaniline (1i)	B	25	1.5	D	0.5	*3i*[16] (92)	9.97[f]
L-Alanyl-L-alanine (*1j*)	A	40	4	E	1.5	*3j* (51)	8.30, 8.10[d]

[a] A:98–100 % HCO_2H; B:98–100 % CF_3CO_2H. [b] C:20 ml 1M HCl; D:20 ml 2M NaOH; E:30 ml 0.5 M HCl. [c] When two signals appear the largest is given first. [d] D_2O, TriMS as reference. [e] Acetone-d_6, TMS as reference. [f] DMSO-d_6, TMS as reference.

appropriate solvent (100 ml) at the specified temperature the amine (10 mmol) was added first, and then glyoxylic acid monohydrate (21 mmol, 1.93 g). The reaction was allowed to proceed with magnetic stirring and a reflux condenser was used when necessary. The reactions of the aromatic amines *1h* and *1i* were carried out under an inert nitrogen atmosphere. Evaporation of the solvent after the time specified gave the formyl derivative *2a–j*. Formyl derivatives *2d* and *2h* were purified before the hydrolysis step. The formyl derivative was then hydrolyzed at 100 °C, with the exception of *2i*, which underwent hydrolysis at 25 °C. The hydrolysate was then evaporated to dryness, with the exception of *3h* and *3i* which were precipitated with 6 M aqueous HCl. Portions of water were added and evaporated to remove excess hydrochloric acid before the final purification of the product, as specified for each compound below.

N-*(Carboxymethyl)glycine* 3a. The free amino acid was obtained by elution from a cation exchanger with 0.5 M aqueous NH_3 and purified by recrystallization from water–acetone. *3a* was analyzed by systems HPLC 1 and TLC 1; m.p. 229–230 °C(dec.). Anal. $C_4H_7NO_4$: C, H, N, O. ^{1}H NMR [60 MHz, TriMS, D_2O]: δ 3.88 (4 H, s).

N-*(Carboxymethyl)*-L-*alanine hydrochloride (strombine) 3b* was obtained by recrystallization from ethanol–ether and was analyzed by systems HPLC 1 and TLC 1; m.p. 168–169 °C(dec.), $[\alpha]_D^{20}$ +6.81° (*c* 5, 6 M HCl). Found: C 32.25; H 5.37; N 7.75; Cl 19.25; O 34.7. Calc. for $C_5H_{10}NClO_4$: C 32.71; H 5.49; N 7.63; Cl 19.31; O 34.86. ^{1}H NMR (60 MHz, TriMS, D_2O): δ 1.63 (3 H, d, *J* 7.0 Hz), 4.07 (2 H, s), 4.23 (1 H, q, *J* 7.0 Hz). The optical purity of synthetic *3b* was determined to be better than 97 % by convertion of the amino acid to its dimethyl ester followed by derivatization with *N*-(trifluoroacetyl)-L-prolyl chloride.[8] The diastereomeric derivatives were then analysed by GLC on a 47 m SE-52 glass capillary column, temperature programmed as follows: initial 105 °C (3 min), linear increase from 105 to 300 °C by 10 °C/min.

N-*(Carboxymethyl)*-D-*phenylglycine* 3c. The free amino acid was obtained by elution from a cation exchanger with 0.5 M aqueous NH_3 and was purified by recrystallization from water–ethanol–ether. *3c* was analyzed by system HPLC 2; m.p. 209–210 °C(dec.), $[\alpha]_D^{20}$ −139.4° (*c* 5, 6 M HCl). Anal. $C_{10}H_{11}NO_4$: C, H, N. ^{1}H NMR (60 MHz, TriMS, DMSO-d_6): δ 3.30 (2 H, s), 4.60 (1 H, s), 7.43 (5 H, s).

N^2-*(Carboxymethyl)*-L-*lysine monohydrate* 3d. *2d* was separated from unreacted lysine on a cation exchanger. After hydrolysis *3d* was converted to the free amino acid by elution from a cation exchanger with 0.5 M aqueous NH_3 and purified by recrystallization from water–ethanol. *3d* was analyzed by systems HPLC 2 and TLC 2; m.p. 248–250 °C(dec.), $[\alpha]_D^{20}$ +21.0° (*c* 5, 6 M HCl). Anal. $C_8H_{18}N_2O_5$: C, H, N. ^{1}H NMR (60 MHz, TriMS, D_2O): δ 1.63 (6 H, m), 3.05 (2 H,

t, *J* 6.5 Hz), 3.65 (2 H, s), 3.72 (1 H, t, *J* 5.0 Hz). ^{13}C NMR (100 MHz, TriMS, D_2O): δ 41.2 (C6), 64.1 (C2), [values for lysine: δ 40.9 (C6), 56.6 (C2)].

N-*Propylglycine hydrochloride 3e* was purified by recrystallization from ethanol–ether and analyzed by systems HPLC 2 and TLC 3; m.p. 172–173 °C(dec.). Anal. $C_5H_{12}NClO_2$: C, H, N. ^{1}H NMR (60 MHz, TriMS, D_2O): δ 0.98 (3 H, t, *J* 7.0 Hz), 1.67 (2 H. m), 3.12 (2 H, t, *J* 7.0 Hz), 3.98 (2 H, s).

N-*Isopropylglycine hydrochloride 3f* was purified and analyzed in the same way as *3e*; m.p. 182–183 °C(dec.). Anal. $C_5H_{12}NClO_2$: C, H, N. ^{1}H NMR (60 MHz, TriMS, D_2O): δ 1.33 (6 H, d, *J* 6.0 Hz), 3.52 (1 H, m), 3.95 (2 H, s).

N-tert-*Butylglycine hydrochloride 3g* was purified and analyzed in the same way as *3e*; m.p. 223–224 °C(dec.). Anal. $C_6H_{14}NClO_2$: C, H, N. ^{1}H NMR (60 MHz, TriMS, D_2O): δ 1.40 (9 H, s), 3.93 (2 H, s).

N-*(p-Carboxyphenyl)-N-formylglycine* 2h. After evaporation of the CF_3CO_2H solvent, unreacted *p*-aminobenzoic acid was removed by washing with 1 M aqueous HCl. *2h* was analyzed with system HPLC 3; m.p. 226–228 °C(dec.). Anal. $C_{10}H_9O_5N$: C, H, N. ^{1}H NMR (60 MHz, TriMS, DMSO-d_6) δ 4.53 (2 H, s), 7.47 (2 H, d, *J* 8.4 Hz), 8.03 (2 H, d, *J* 8.4 Hz), 8.80 (1 H, s). The dimethyl ester of *2h* was obtained by treatment of the amino acid with diazomethane. MS [*m/e* (% rel. int.)]: 251 (10, M), 223 (31), 220 (6.6), 192 (12), 164 (100), 132 (11), 104 (7.0), 77 (6.9). ^{1}H NMR (360 MHz, $CDCl_3$): δ 3.78 (3 H, s), 3.94 (3 H, s), 4.55 (2 H, s), 7.26 (2 H, d, *J* 8.6 Hz), 8.10 (2 H, d, *J* 8.6 Hz), 8.62 (1 H, s).

N-*(p-Carboxyphenyl)glycine 3h* was recrystallized by dissolving in aqueous Na_2CO_3 followed by precipitation with 2 M aqueous HCl. *3h* was analyzed by system HPLC 3; m.p. 233–234 °C(dec.). ^{1}H NMR (60 MHz, TriMS, DMSO-d_6): δ 3.97 (2 H, s), 6.67 (2 H, d, *J* 8.4 Hz), 7.80 (2 H, d, *J* 8.4 Hz). The dimethyl ester of *3h* was obtained by treatment of the amino acid with diazomethane. MS [*m/e* (% rel. int.)]: 223 (33, M), 192 (12), 178 (5.1), 164 (100), 132 (10), 120 (6.0), 105 (10), 104 (6.7). ^{1}H NMR (360 MHz, $CDCl_3$): δ 3.81 (3 H, s), 3.85 (3 H, s), 3.97 (2 H, s), 6.57 (2 H, d, *J* 8.6 Hz), 7.89 (2 H, d, *J* 8.6 Hz).

N-*(p-Nitrophenyl)glycine 3i* was recrystallized in the same way as *3h* and analyzed by systems HPLC 4 and TLC 4; m.p. 218–219 °C(dec.). ^{1}H NMR (60 MHz, TriMS, DMSO-d_6): δ 4.04 (2 H, s), 6.72 (2 H, d, *J* 9.0 Hz), 8.08 (2 H, d, *J* 9.0 Hz). The methyl ester of *3i* was obtained by esterification with diazomethane. MS [*m/e*, (% rel. int.)]: 210 (26, M), 151 (100), 135 (3.4), 121 (3.4), 105 (44), 76 (6.5). ^{1}H NMR (360 MHz, $CDCl_3$): δ 3.84 (3 H, s), 4.00 (2 H, s), 6.56 (2 H, d, *J* 8.6 Hz), 8.12 (2 H, d, *J* 8.6 Hz).

(N-(Carboxymethyl)-L-alanyl)-L-alanine hydrochloride 3j was purified by recrystallization from ethanol–ether and analyzed by system TLC 1; m.p. 166–168 °C(dec.). ^{1}H NMR (60 MHz, TriMS, D_2O): δ 1.47 (3 H, d, *J* 11.2 Hz), 1.60 (3 H, d, *J* 8.0 Hz), 3.97 (2 H, s), 4.18 (1 H, q, *J* 11.2 Hz), 4.43 (1 H, q, *J* 8.0 Hz). The dimethyl ester of *3j* was obtained by esterification with diazomethane. MS [*M/e* (% rel. int.)]: 246 (2.3, M), 187 (22), 130 (60), 117 (10), 116 (100), 70 (9.6), 59 (10), 57 (7.9), 56 (79). ^{1}H NMR (60 MHz, acetone-d_6): δ 3.73 (6 H, s, methyl esters).

Acknowledgements. We thank Dr. Robert E. Carter for helpful linguistic criticism.

REFERENCES

1. Sangster, A. W., Thomas, S. E. and Tingling, N. L. *Tetrahedron 31* (1975) 1135.
2. a. Haynes, L. J., Sangster, A. W., Steven, D. M. and Thomas, S. *Comp. Biochem. Physiol. 20* (1967) 755; b. Steven, D. M. *J. Exp. Biol. 36* (1959) 261.
3. a. Greenlee, W. J. *Tetrahedron Lett. 23* (1982) 3995; b. Patchett, A. A. *et al. Nature 288* (1980) 280.
4. a. Miyazawa, T. *Bull. Chem. Soc. Jpn. 53* (1980) 2555; b. Ibid. (1980) 3661.
5. Moore, M. L. *Organic Reactions,* Wiley, New York 1949, Vol. 5, p. 317.
6. a. Erlenmeyer, E., Jr. and Kunlin, J. *Justus Liebigs Ann. Chem. 307* (1899) 146; b. Erlenmeyer, E., Jr. and Kunlin, J. *Ber. Dtsch. Chem. Ges. 35* (1902) 2438; c. Erlenmeyer, E., Jr. *Ber. Dtsch. Chem. Ges. 36 (2)* (1903) 2525.
7. von Euler, H., Hasselquist, H. and Eriksson, E. *Justus Liebigs Ann. Chem. 578* (1952) 188.
8. Halpern, B. In Blau, K. and King, G., Eds., *Handbook of Derivatives for Chromatography,* Heyden, London 1977, p. 464.
9. Jones, J. H. In Barton, D. and Ollis, W. D., Eds., *Comprehensive Organic Chemistry,* Pergamon, Oxford 1979, Vol. 2, p. 828 and references therein.
10. Jackman, L. M. and Sternhell, S. *Applications of NMR Spectroscopy in Organic Chemistry,* 2nd Ed., Pergamon, Oxford 1969, p. 361.
11. Greene, T. H. *Protective Groups in Organic Synthesis,* Wiley, New York 1981, p. 250.

12. Chase, B. H. and Downes, A. M. *J. Chem. Soc.* (1953) 3874.
13. Matsutani, H., Kusumoto, S., Koizumi, R. and Shiba, T. *Phytochemistry 18* (1979) 661.
14. Greco, C. V., Nyberg, W. H. and Cheng, C. C. *J. Med. Pharm. Chem. 5* (1962) 861.
15. Musante, C. and Fabbrini, L. *Gazz. Chim. Ital. 84* (1954) 595; see *Chem. Abstr. 50* (1956) 968b.
16. Passeron, S. and Brieux, G. A. *Bull. Soc. Chim. Fr.* (1963) 35.

Received February 7, 1983.