Piperazinedione Formation from Esters of Dipeptides containing Glycine, Alanine, and Sarcosine: the Kinetics in Aqueous Solution

By **Jocelyn E. Purdie** *† and **N. Leo Benoiton**, Department of Biochemistry, University of Ottawa, Ottawa, Canada K1N 6N5

Piperazinedione formation from the methyl esters of glycylglycine, glycyl-L-alanine, L-alanylglycine, glycylsarcosine, and sarcosylglycine and from the ethyl ester of glycylglycine in aqueous solution, pH 7·3—8·5, and 25·0° was studied. It was found to be a self-catalysed reaction and other amines also served as catalysts. Some concomitant ester hydrolysis occurred. The kinetics were analysed by regression analysis and the data fitted equation (i). Values of k_{hyd}', $\{k_s + k_{OH}[OH^+]\}$, k_{gb}, and k_{gb}' were obtained. Glycylsarcosine methyl ester

$$d[\text{ester, total}]/dt = [\text{ester, free base}]\{k_s + k_{gb}'[\text{amine, free base}] + k_{OH}[OH^-]\} + [\text{ester, free base}]^2 k_{gb} + [\text{ester, total}] k_{hyd}' \quad (i)$$

cyclised most rapidly ($t_{\frac{1}{2}}$ *ca.* 5 min), by more than an order of magnitude, and this was attributed to the much higher content of *cis*-isomer. The methyl esters of sarcosylglycine and glycylalanine cyclised considerably faster than the glycylglycine ester; sarcosylglycine methyl ester showed evidence of steric hindrance in its smaller self-catalytic (k_{gb}) rate constant. Alanylglycine methyl ester cyclised the slowest and had a markedly low value of k_{gb}. Saponification rates were similar to those of the esters of *N*-acylamino-acids. The mechanism of the cyclisation process, together with the effects of *N*- and *C*-methyl substituents on the reaction at the ester carbonyl carbon and on the ring opening of the piperazinediones are discussed.

It is well known that piperazine-2,5-diones are formed readily from dipeptide derivatives, including esters [1,2] and amides,[3,4] especially when these are not *N*-protected, and particularly those containing *N*-methyl-amino-acids or proline; [5,6] currently this is causing problems in the synthesis of biologically active peptides.[7-9] However, no detailed study has hitherto been made of the kinetics of the cyclisation of dipeptide esters in solution. Piperazinediones are also formed from amino-acid esters in the free base form.[10,11] The kinetics of this slower process have not been investigated either, but if this reaction does proceed through self-condensation to dipeptide esters,[12] kinetic data on the latter would be essential for such a study. We have followed the kinetics of the cyclisation under slightly alkaline conditions of the ethyl and methyl esters of glycylglycine and its *N*- and *C*-methylated homologues. The reaction is complicated by concomitant saponification and as became apparent, by self-catalysis by the uncharged form of the ester. In addition, the piperazinedione products are very susceptible to alkaline hydrolysis.

EXPERIMENTAL

Materials.—Analytical grade reagents and water, triply distilled and stored under nitrogen, were used. Sarcosylglycine, m.p. 205°, was prepared by the method of Levene *et al.*[13] Other dipeptides, piperazinediones, and the methyl and ethyl esters of glycylglycine hydrochloride were commercial products. Their purities were checked by ion-exchange chromatography (Beckman 120B amino-acid analyser) and for the piperazinediones, also by t.l.c. and titration. The methyl ester hydrochlorides of the dipeptides were prepared using methanol and thionyl chloride.[14] Problems were encountered due to the splitting of the dipeptides and consequent contamination of the products by amino-acid esters. This was minimised by very slow addition of fresh samples of thionyl chloride to well-cooled dipeptide suspensions and by careful re-crystallisation of products. The latter was not possible with the deliquescent sarcosine compounds (see below). The m.p. of the methyl ester hydrochlorides of L-alanylglycine and glycyl-L-alanine (both recrystallised from methanol–ether) were 161·5—162·5 and 160—160·8°, respectively.

Kinetic Procedure.—The reaction was carried out under nitrogen, at 25° in a Radiometer pH-stat type PHM 26c/TT11a/SBUI/SBR2/TTA3 standardised as described previously.[15] The solvent was a sodium chloride solution (1M to minimise salt effects) and initial concentration of dipeptide ester hydrochloride was 1—6 × 10^{-2}M. Reaction was started by addition of sodium hydroxide (1N) with rapid stirring to avoid 'spot' saponification and the pH was subsequently kept constant by further addition. Samples were withdrawn at appropriate intervals, acidified to pH 2·2 with hydrochloric acid (which stopped the reaction), and kept in ice until analysed. All analyses were done within 8 h.

Disappearance of ester was measured by a modification [16] of Hestrin's colorimetric method [17] and 20 samples were analysed for each run. Solutions of deliquescent esters were standardised by complete saponification to dipeptides

† *Present address:* Paterson Laboratories, Christie Hospital and Holt Radium Institute, Manchester M20 9BX.

[1] E. Fischer and E. Forneau, *Ber.*, 1901, **34**, 2868.
[2] G. Agren, *Arkiv. Kemi, Mineral. Geol.*, 1940, **14***B*, 1.
[3] H. T. Huang and C. Niemann, *J. Amer. Chem. Soc.*, 1950, **72**, 921.
[4] L. Meriwether and F. H. Westheimer, *J. Amer. Chem. Soc.*, 1956, **78**, 5119.
[5] J. Rudinger in 'Peptides,' ed. G. T. Young, Pergamon, Oxford, 1963, p. 159.
[6] H. N. Rydon and P. W. G. Smith, *J. Chem. Soc.*, 1956, 3642.
[7] B. F. Gisin and R. B. Merrifield, *J. Amer. Chem. Soc.*, 1972, **94**, 3102.
[8] M. C. Khosla, R. R. Smeby, and F. M. Bumpus, *J. Amer. Chem. Soc.*, 1972, **94**, 4721.
[9] M. Rothe and J. Mazánek, *Angew. Chem. Internat. Edn.*, 1972, **11**, 293.
[10] E. Fischer and U. Suzuki, *Ber.*, 1905, **38**, 4173.
[11] E. Abderhalden and S. Suzuki, *Z. physiol. Chem.*, 1928, **176**, 101.
[12] J. P. Greenstein and M. Winitz, 'Chemistry of the Amino Acids,' Wiley, New York, 1961, vol. 2, p. 785.
[13] P. A. Levene, H. J. Sims, and M. H. Pfaltz, *J. Biol. Chem.*, 1924, **61**, 445.
[14] Ref. 12, p. 927.
[15] J. E. Purdie and N. L. Benoiton, *Canad. J. Chem.*, 1971, **49**, 3468.
[16] W. P. Jencks and J. Carriuolo, *J. Amer. Chem. Soc.*, 1961, **83**, 1743.
[17] S. Hestrin, *J. Biol. Chem.*, 1949, **180**, 249.

and subsequent determination. Piperazinediones reacted slightly with the hydroxylamine reagent and each observed O.D. reading was corrected according to equation (1) where

$$E_t = \frac{OD_t - C_D[E_o - [\text{dipeptide}]]}{C_E - C_D} \quad (1)$$

E_t = [ester] at time t, E_o = [ester] at time = 0, C_D = colour yield of piperazinedione, C_E = colour yield of ester.

The O.D. readings from the sarcosine dipeptide esters which contained amino-acid esters (<4%) as contaminants, also needed correction for the saponification of the latter. As the dipeptides were stable under experimental conditions, any amino-acid produced during the course of the reaction originated from its ester. Thus a complete analysis of ninhydrin-positive components, at various time intervals, allowed the correction to be made. Ninhydrin-positive components, in 7—8 samples per run, were analysed on a column (17 cm) of Aminex A-5 resin, at 57°, with the eluting buffer at pH 3·25. Under these conditions, the following separations were achieved: glycine from glycylglycine, glycylglycylglycine, and glycylglycylglycylglycine; glycine and sarcosine from glycylsarcosine and sarcosylglycine; and glycine and alanine from glycylalanine and alanylglycine. Glycylglycine and sarcosylglycine were partially resolved. The dipeptide esters were eluted after ammonia, followed by the amino-acid esters. Sarcosine and sarcosylglycine were analysed at half the normal flow rate of buffer under which conditions the colour yield of sarcosine increased to 33 but that of sarcosylglycine was only 2·2. All other components had colour yields close to 20 at the normal buffer flow rate. No detectable hydrolysis of any piperazinedione occurred during passage through the Aminex column.

The piperazinediones were isolated from the acidified reaction mixture by passage through a column of Dowex 50W-X8 in the H^+ form. The product was identified by t.l.c. and by amino-acid analysis before and after ring opening by mild alkali. The unhydrolysed material invariably contained only traces of ninhydrin-positive material.

Determination of pK'.—The apparent dissociation constant (K') is defined by equation (2). The pK' values were

$$K' = a_{H^+}[RNH_2]/[RNH_3^+] \quad (2)$$

determined in sodium chloride solution (1M) at 25°, as described previously.[15] It was not possible to determine the pK' of glycylsarcosine methyl ester as the free base cyclised so rapidly.

RESULTS

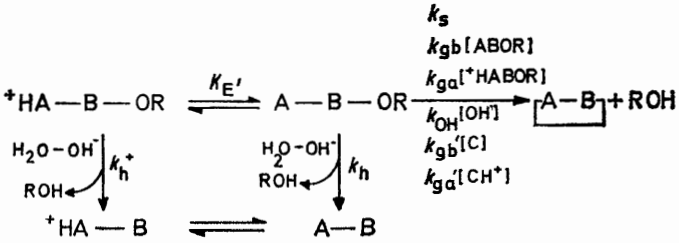
The first experiments showed that with all the dipeptide esters cyclisation competed very effectively with saponification, within the pH range 7·3—8·5. This range was determined at the lower limit by the decreasing rate of cyclisation and at the upper limit by the increasing lability of the product piperazinediones, cyclo-(Gly–Gly–) being the least stable. The cyclisation of glycylsarcosine methyl ester proceeded too rapidly to follow above pH 7·3.

The products from the dipeptide esters of glycine and alanine could be completely accounted for by piperazinedione and dipeptide. Higher peptides could not be detected either in the reaction mixture or in samples of reaction mixture treated with a minimum of alkali so that only saponification and some ring opening occurred. Analytical conditions were such that 0·004 μmol of glycylglycylglycylglycylglycine would have been detected and loads equivalent to 8 μmol starting material and 1 μmol saponified material were analysed. Also, the piperazinedione fraction from the Dowex 50 column yielded only dipeptides after mild alkaline treatment (analytical column load 2 μmol).

cyclo-(Gly–Sar–) was the major product from the methyl esters of glycylsarcosine and sarcosylglycine; the latter also yielded some sarcosylglycine but saponification was unable to compete with the cyclisation of the former. Both esters gave traces of unidentified material.

From the alkaline hydrolysis of the piperazinediones, the concentration ratios [Gly–Sar] : [Sar–Gly] = 3·4 and [Gly–Ala] : [Ala–Gly] = 2·1 were obtained from cyclo-(Gly–Sar–) and cyclo-(Ala–Gly–) respectively. The presence of 'inverted' dipeptide in a cyclisation mixture would indicate ring opening but this was only observed in a 6 h sample from L-alanylglycine methyl ester at pH 8·52 and only as a slight rise in baseline at the glycylalanine position (column load equivalent to 7·3 μmol starting material).

The rate of disappearance of each ester was unequivocably by mixed first- and second-order kinetics. Self-condensation was ruled out by the absence of higher linear or cyclic peptides. Also, the rate of cyclisation of glycylglycine ethyl ester was increased by TRIS, glycine, glycylglycine, and sarcosylglycine, and again the only products detected were the piperazinedione and dipeptide. Hence it was concluded that the second-order component was due to self-catalysis and Scheme 1 was suggested. This is similar to the system describing the amine-catalysed aminolysis of phenyl acetate,[18] but is not a general equation for such a reaction if a kinetically significant intermediate is involved.



SCHEME 1

In Scheme 1, C is an amine catalyst (apparent dissociation constant K_c'), and k_s, k_{gb}, and k_{ga} are the rate constants for the spontaneous or water-catalysed, general base-catalysed, and general acid-catalysed cyclisation respectively.

Writing [E′] = [E] + [EH⁺] for total dipeptide ester, and [C′] = [C] + [CH⁺] for total amine, equation (3) is

$$-d[E']/dt = a[E'] + b[E']^2 \quad (3)$$

derived where equations (4) and (5) hold. Integration of equation (3) gives (6).

$$a = \frac{K_{E'}}{(K_{E'} + [H^+])}\left\{k_s + k_h + k_{gb}'\frac{K_c'[C']}{(K_c' + [H^+])} + k_{ga}'\frac{[H^+][C']}{(K_c' + [H^+])} + k_{OH}[OH^-]\right\} + \frac{k_h^+[H^+]}{(K_{E'} + [H^+])} \quad (4)$$

$$b = k_{gb}\left\{\frac{K_{E'}}{K_{E'} + [H^+]}\right\}^2 + k_{ga}\frac{[H^+]K_{E'}}{(K_{E'} + [H^+])^2} \quad (5)$$

$$\frac{[E_0']}{[E']} = e^{at} + \frac{b[E_0']}{a}(e^{at} - 1) \quad (6)$$

[18] W. P. Jencks and J. Carriuolo, *J. Amer. Chem. Soc.*, 1960, **82**, 675.

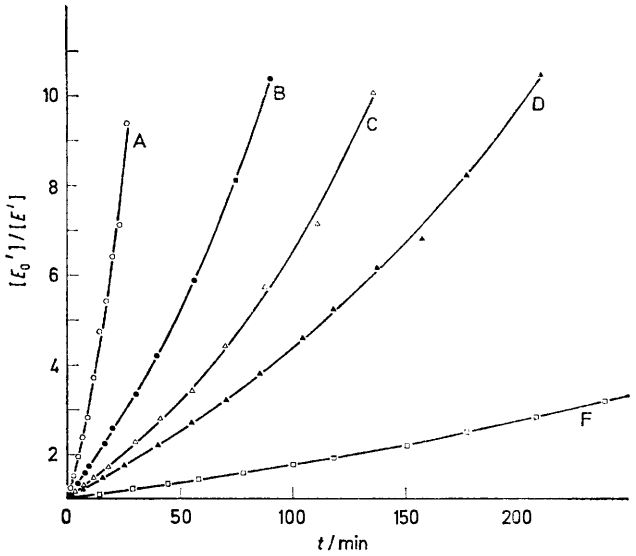
As the hydrolysis of the piperazinediones was negligible under experimental conditions, d[dipeptide]/dt = k_h'[E'] where equation (7) applies. Integration gives equation (8).

$$k_h' = \frac{k_h^+[H^+]}{(K_E' + [H^+])} + \frac{k_h K_E'}{(K_E' + [H^+])} \quad (7)$$

$$[\text{dipeptide}] = \frac{k_h'}{b} \log_e \left[1 + \frac{bE_0'}{a}(1 - e^{-at})\right] \quad (8)$$

Readings of $[E_0']/[E']$ *vs.* time were subjected to regression analysis using the BMD [19] programs X85 (nonlinear least squares) and 06R (asymptotic regression) as described in the Appendix. Only values of $[E'_0/E'] \leqslant 10$ (*i.e.* between 3 and 4 half-lives) were used; this corresponded approximately to a decrease in O.D. from *ca.* 1·4 to 0·14, the error in the reading thus increasing with time. This is a problem inherent in the measurement of disappearance of starting material. In our case it was not practical to analyse a large number of samples for the piperazinedione products. These were not sufficiently lipophilic to be readily extracted from aqueous solution for analysis by g.l.c. Determination of methanol posed a similar problem. However by carefully standardising experimental procedures, the colorimetric assay was precise enough to give $[E_0']/[E']$ *vs.* time plots which were superimposable in repeat runs.

The experimental data fitted equation (6) very closely for all the esters as can be seen in the Figure. The worst fit occurred at pH 7·92 with glycylglycine ethyl ester (10^{-2}M) and L-alanylglycine methyl ester (4×10^{-2}M); both these reactions were slow, with half-lives *ca.* 450 and 400 min respectively and could not be followed for much longer together with pK' values and Table 2 shows the effect of amine catalysts on the cyclisation of glycylglycine ethyl ester.



Plots of $[E_0']/[E']$ against t. The points are experimental values, the lines are the regression predicted curves: pH = 7·30 A, Gly–Sar–OMe; pH = 8·52 B, Gly–L–Ala–OMe; C, Sar–Gly–OMe; D, Gly–Gly–OMe; F, L–Ala–Gly–OMe

Parameter b.—As the fraction of free base ($\alpha = K_E'/K_E' + [H^+]$) will increase with pH, it follows from equation (5), for any value of k_{gb} and k_{ga} and $pH_2 > pH_1$, that $\alpha_2^2/\alpha_1^2 - b_2/b_1 \geqslant 0$. For $pH_1 = 7{\cdot}92$, $pH_2 = 8{\cdot}52$, the values of α_2^2/α_1^2 and b_2/b_1, respectively, are for glycylglycine ethyl ester, 2·63 and 2·64 ± 0·18; glycylglycine methyl

TABLE 1

Titrimetric data for dipeptide esters containing glycine, sarcosine, and alanine and kinetic parameters * for their concurrent saponification and cyclisation to piperazinediones in 1M-sodium chloride solution at 25·0°

Ester	pK' Ester	pK' Dipeptide	pH	$10^5 a$*/s^{-1}	$10^3 b$*/l mol^{-1} s^{-1}	$10^6 k_h'$/s^{-1}	$10^2 k_{gb}$/l mol^{-1} s^{-1}
Gly–Gly–OMe	8·00	8·20	7·92	3·66 ± 0·18	2·97 ± 0·23	3·67 ± 0·05	1·44
			8·52	7·71 ± 0·50	9·23 ± 0·83	13·0 ± 0·2	1·56
Gly–Gly–OEt	7·94	8·20	7·92	1·43 ± 0·13	2·00 ± 0·12	1·85 ± 0·08	0·84
			8·52	4·13 ± 0·33	5·28 ± 0·13	6·67 ± 0·37	0·84
Gly–Ala–OMe	8·11	8·27	7·92	5·43 ± 0·23	7·11 ± 0·52	1·66 ± 0·05	4·62
			8·52	15·1 ± 0·31	23·7 ± 0·8	6·28 ± 0·33	4·58
Ala–Gly–OMe	8·03	8·17	7·92	2·00 ± 0·25	0·41 ± 0·05	2·98 ± 0·12	0·22
			8·52	4·49 ± 0·12	1·52 ± 0·08	10·1 ± 0·3	0·265
Gly–Sar–OMe	†	8·65	7·30	59·1 ± 8·1	67·5 ± 7·3	†	
Sar–Gly–OMe	8·35	8·68	7·92	4·30 ± 0·17	2·46 ± 0·17	3·67 ± 0·05	3·35
			8·52	14·1 ± 1·3	12·5 ± 1·9	12·2 ± 0·5	3·52

* From 2—10 runs. † See text.

TABLE 2

Kinetic parameters for the base-catalysed cyclisation of glycylglycine ethyl ester at pH 8·52 in 1M-sodium chloride solution at 25·0°

Amine	10^3[Amine]/M	pK' Amine	$10^5 a$/s^{-1}	$10^3 b$/l mol^{-1} s^{-1}	$10^6 k_h'$/s^{-1}	$10^2 k_{gb}'$/l mol^{-1} s^{-1}
Gly–Gly–OEt		7·94	4·13 ± 0·33	5·28 ± 0·13	6·7 ± 0·4	0·84
Gly	20·0	9·65	7·90 ± 0·6	4·90 ± 0·47	6·3 ± 0·4	3·45
Gly–Gly	19·4	8·20	14·4 ± 0·65	4·91 ± 0·45		0·983
Sar–Gly	19·0	8·68	11·1 ± 1·1	4·48 ± 0·97	*	1·13
TRIS	18·8	8·22	10·6 ± 0·5	4·95 ± 0·40	6·3 ± 0·41	0·65

* See text.

periods without introducing errors due to drifting in the pH-stat.

Values of a and b obtained from the computer programs were substituted into equation (8) together with values of [dipeptide] and time (taken from smoothed progress curves) to give k_h'. These composite constants are shown in Table 1

[19] B. M. D. Biomedical Computer Programs, University of California Press, Berkeley, 1970.

ester, 2·86 and 3·15 ± 0·38; glycylalanine methyl ester, 3·37 and 3·34 ± 0·17; alanylglycine methyl ester, 3·02 and 3·68 ± 0·49; and for sarcosylglycine methyl ester, 4·85 and 5·09 ± 0·41. Thus the term in k_{ga} cannot be important. To test further the sensitivity of the ratio to the term in acid catalysis it was assumed that $k_{gb} = mk_{ga}$ where $10 > m > 0{\cdot}2$. For each ester it was ascertained that the resultant b_2/b_1 value would be significantly different from the experimental value. Thus the term in acid catalysis could be ignored in this case and k_{gb} values were calculated (see Table 1).

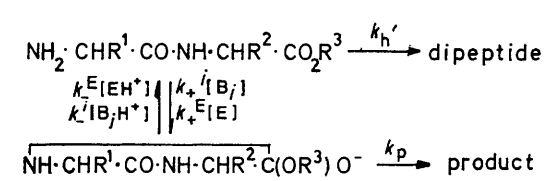
Parameter a.—In the absence of added catalysts, from equation (4) $a - k_h' = \alpha\{k_s + k_{OH}[OH^-]\}$ and, if $k_{OH} > 0$, $[(a_2 - k_{h2}')/(a_1 - k_{h1}')] > \alpha_2/\alpha_1$. The ratios of $[(a_2 - k_{h2}')/(\alpha_1 - k_{h1}')]$ and α_2/α_1 respectively are: for glycylglycine ethyl ester, 2·78 and 1·62; glycylglycine methyl ester, 1·94 and 1·78; glycylalanine methyl ester, 2·74 and 1·84; alanylglycine methyl ester, 2·01 and 1·74; and for sarcosylglycine methyl ester, 3·27 and 2·20. Thus catalysis by hydroxide ion is not negligible in spite of its very low concentration and k_{OH} must be between 2 and 3 orders of magnitude greater than k_{gb}. According to equation (4) k_{gb}' can be determined from $(a - k_h')$ in the presence and absence of added catalyst, assuming k_{ga}' is negligible. From Table 2 it can be seen that b did not change significantly in accordance with the suggested reaction scheme, and k_{gb}' values were calculated. The rate of saponification was unaffected by added base. A Brønsted plot showed a linear relation for glycylglycine ethyl ester, glycylglycine, and glycine (pK' range 7·9—9·65) with slope 0·35. Sarcosylglycine and TRIS showed marked negative deviation.

DISCUSSION

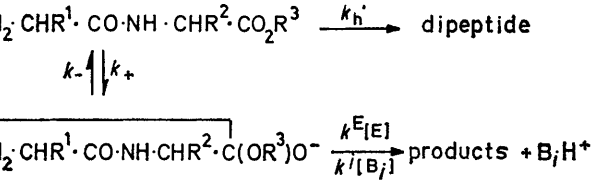
Some aspects of the kinetics of the aminolysis of esters are best explained by assuming the existence of a tetrahedral intermediate.[20,21] Specific examples include the aminolysis of methyl formate,[22] the lactamisation of the ethyl esters of some amino-acids,[23] and the aminolysis of *S*-ethyl thiobenzoate.[24] As these reactions are general base-catalysed, the question arises whether the catalysis occurs in the formation (and back reaction) of the intermediate, or in its breakdown to products, or in both steps. For methyl formate at high pH, and the ethyl ester of glutamic acid (which have poor leaving groups) the formation of the intermediate is believed to be catalysed and its breakdown to products uncatalysed. The reverse situation is postulated for *S*-ethyl thiobenzoate (which has a good leaving group). The mechanisms for the cyclisation and saponification of dipeptide esters are portrayed in Schemes 2 and 3. The latter reaction probably also proceeds *via* an intermediate but this is unimportant here. For Scheme 2 the assumption that steady state conditions prevail leads to equation (9).

$$\frac{-d[E']}{dt} = k_h'[E'] + \frac{k_p[E]\{k_+^E[E] + \sum_i k_+^i[B_i]\}}{k_p + k_-^E[EH^+] + \sum_i k_-^i[B_iH^+]} \quad (9)$$

If $k_p \gg k_-^E[EH^+] + \sum_i k^i[B_iH^+]$ equation (9) reduces to the form of equation (3). For Scheme 3, again assuming



SCHEME 2



SCHEME 3

steady state conditions, equation (10) holds. If $k_- \gg k^E[E] + \sum_i k^i[B_i]$, equation (10) reduces to the form of

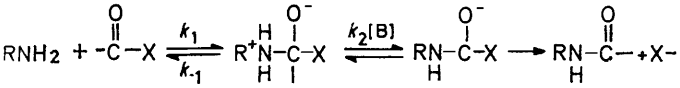
$$\frac{-d[E']}{dt} = k_h'[E'] + \frac{k_+[E]\{k^E[E] + \sum_i k^i[B_i]\}}{k_- + k^E[E] + \sum_i k^i[B_i]} \quad (10)$$

equation (3) except that an equilibrium constant (k_+/k_-) is included. If $k^E[E] + \sum_i k^i[B_i] \gg k_-$ equation (10) reduces to equation (11).

$$-d[E']/dt = k_h'[E'] + k_+[E] \quad (11)$$

Our data fit equation (3) so closely that the reaction could be of the concerted type, or either Scheme 2 or 3 could apply but our experimental conditions were such that the kinetics were of the form in equation (3). Martin *et al.*[23] did not observe self-catalysis (by ester) of lactam formation but they used low ester concentrations ($\sim 10^{-4}$M).

In order to explain the non-linear Brønsted plots observed in some aminolysis reactions,[25] Jencks has proposed [26] that the zwitterionic intermediate breaks down in two steps, the first, general base-catalysed, producing an anionic entity which then decomposes unassisted to



SCHEME 4

products as shown in Scheme 4. The base functions to prevent the formation of an unstable *N*-protonated

[20] W. P. Jencks, 'Catalysis in Chemistry and Enzymology,' McGraw-Hill, New York, 1969.
[21] T. C. Bruice and S. J. Benkovic, 'Bioorganic Mechanisms,' Benjamin, New York, 1966, vol. 1.
[22] G. M. Blackburn and W. P. Jencks, *J. Amer. Chem. Soc.*, 1968, **90**, 2638.
[23] R. B. Martin, A. Parcell, and R. L. Hedrick, *J. Amer. Chem. Soc.*, 1964, **86**, 2406.
[24] B. Boopsingh and D. P. N. Satchell, *J.C.S. Perkin II*, 1972, 1702.
[25] J. P. Fox, M. I. Page, A. Satterthwait, and W. P. Jencks, *J. Amer. Chem. Soc.*, 1972, **94**, 4729.
[26] W. P. Jencks, *Chem. Rev.*, 1972, **72**, 705.

amide product, and this rate-limiting proton transfer step gives rise to the typically curved Brønsted plots [27] with limiting slopes approaching 0 and 1·0. This mechanism will be described by Scheme 2 if the amino-group addition is relatively fast so that the equilibrium constant can be included in a rate constant. Jencks further suggested that a concerted mechanism would occur where either intermediate is particularly unstable with respect to breakdown to *N*-protonated product.

Values of β of 0·9—1·0 have been reported for the general base-catalysed aminolysis of methyl formate [22] and phenyl acetate.[28] In these cases, the catalysis was detected experimentally by the appearance of a second-order term with respect to amine and it was not possible to vary the attacking amine and amine catalyst separately. In our case, added amine functioned only catalytically and so the derived β value is lower and should give a more direct measure of proton removal in the rate-limiting step. Our value of *ca.* 0·35 is comparable with that of *ca.* 0·4 derived from the lactam formation from amino-acid esters.[23] As our data were obtained over a narrow range of pK (7·9—9·6) it could be part of the curved plot indicative of a simple proton transfer rate-limiting step, or it may remain linear over a wide range of pK in which case bond making and breaking at a carbon atom could be involved as in a concerted mechanism. The latter possibility is preferred for reasons discussed below.

The cyclisation of dipeptide esters to piperazinediones is interesting as the peptide bond must be in the *cis*-form so that the terminal amino-group and ester carbonyl can interact to form the six-membered ring which then contains two *cis*-bonds. The planar unprotonated peptide bond in open-chain compounds normally exists almost exclusively in the *trans*-conformation with a barrier to rotation of *ca.* 20 kcal mol^{-1}.[29] With *N*-methylated peptide bonds, the *cis*-conformation is relatively less unfavourable and n.m.r. studies have shown that an equilibrium occurs, in aqueous solution, between *cis*- and *trans*-forms in peptides of sarcosine [30] and *N*-methylalanine.[31,32] For *N*-acetyl-D-alanyl-L-*N*-methylalanine methyl ester, the population of *cis*-isomer was 20%.[32] Of the esters we studied, only glycylsarcosine methyl ester should have a measurable equilibrium content of *cis*-isomer and its very fast cyclisation must be a reflection of this. Rydon and Smith [6] observed that glycylproline ethyl ester cyclises much more readily than the prolylglycine ester and attributed this to geometrical factors and ease of assumption of *cis*-conformation.

In spite of the very unfavourable equilibrium of *cis*-isomer in the dipeptide esters (other than glycylsarcosine ester) the intramolecular aminolysis proceeded so fast that any intermolecular interaction resulted in catalytic action only under our experimental conditions. Perhaps this is not surprising; Bruice [33] has pointed out that the facility of intramolecular reactions generally makes it difficult to compare their rates with those of the analogous bimolecular reactions and few such cases have been studied. The reasons why intramolecular reactions proceed so fast have recently been the subject of much debate, especially the concept of 'orbital steering' and its role in enzymic reaction.[34] Page and Jencks [35] have argued that a large part of the enhancement in rate of the intramolecular over the intermolecular process can be accounted for by entropy factors arising from translational and overall rotational motion, other factors such as orientation, solvation, and steric effects being usually of secondary importance. Space filling models (CPK) of dipeptide esters in the *cis*-conformation show that certain rotations would produce a situation in which the attacking amino-nitrogen and target carbonyl carbon atoms are in contact within their van der Waals radii, the C–N line perpendicular to the carbonyl plane. Thus, from purely geometrical considerations the *cis*-isomer can adopt a conformation very favourable for ring formation, although obviously, in a flexible molecule the preferred rotational conformation, determined by non-bonded interactions and specific solvation effects will give the statistical distance between the reactive groups. This, together with the favourable entropy factor mentioned above must largely determine the predominance of the intramolecular reaction over its intermolecular counterpart even though the latter would take place with the much more probable *trans*-isomer. Also, the self-catalytic nature of the cyclisation process would ensure its precedence over self-condensation, even in concentrated solution.

The rate constants for the cyclisation of glycylsarcosine methyl ester cannot be directly compared with those of the other esters as its pK' value could not be determined. The pK' of the dipeptide is the same as that of sarcosylglycine although the methylamino-group is a stronger base than the amino-group. The pK' of the carboxy-group is also lower than that of sarcosylglycine and glycylglycine.[36] This is probably due to the stabilisation of the zwitterion by electrostatic interaction between the charged end-groups which is only possible in the *cis*-isomer. This effect cannot operate in the ester and its pK' value may be close to that of glycylalanine methyl ester. Assuming a value of 8·2 for pK' would give $k_{gb} = 5{\cdot}5$ l mol^{-1} s^{-1} and $(a - k_h') = 1{\cdot}8 \times 10^{-3}$ s^{-1} at pH 7·92, ignoring the increase in contribution from $k_{OH}[OH^-]$.

[27] M. Eigen, *Angew. Chem. Internat. Edn.*, 1964, **3**, 1.
[28] T. C. Bruice, A. Donzel, R. W. Huffman, and A. R. Butler, *J. Amer. Chem. Soc.*, 1967, **89**, 2106.
[29] G. N. Ramachandran and V. Sasisekharan, in *Adv. Protein Chem.*, 1968, **23**, 287.
[30] B. Liberek, K. Steporowska, and E. Jereczek, *Chem. and Ind.*, 1970, 1263.
[31] M. Goodman and N. S. Choi, 'Peptides: Proc. Ninth European Peptide Symp.', North Holland Publishing Co., 1968, p. 1.
[32] V. F. Bystrov, S. L. Portnova, V. I. Tsetlin, V. T. Ivanov, and Yu. A. Ovchinnikov, *Tetrahedron*, 1969, **25**, 493.
[33] T. C. Bruice in 'The Enzymes,' ed. P. D. Boyer, Academic Press, New York, 1970, vol. II, ch. 4.
[34] *Chem. Eng. News*, 1971, no. 12, 49.
[35] M. I. Page and W. P. Jencks, *Proc. Nat. Acad. Sci. U.S.A.*, 1971, **68**, 1678.
[36] W. L. Koltun, M. Fried, and F. R. N. Gurd, *J. Amer. Chem. Soc.*, 1960, **82**, 233.

'Spontaneous' or Water-catalysed Reaction.—The constant $(a - k_h')$ contains a term $k_{OH}[OH^-]$ as well as k_s. At pH 7·92 the contribution from $k_{OH}[OH^-]$ would be about four times less than that at pH 8·52. So at the lower pH, $(a - k_h')$ would be a closer estimate of k_s. At pH 7·92, $[(a - k_h')/\alpha]$ values are (in s^{-1}): glycylglycine methyl ester, $0{\cdot}073 \times 10^{-3}$; glycylalanine methyl ester, $0{\cdot}134 \times 10^{-3}$; glycylsarcosine methyl ester, *ca.* $5{\cdot}3 \times 10^{-3}$; alanylglycine methyl ester, $0{\cdot}039 \times 10^{-3}$; sarcosylglycine methyl ester, $0{\cdot}145 \times 10^{-3}$; glycylglycine ethyl ester, $0{\cdot}026 \times 10^{-3}$. If K_I is the equilibrium constant for the conversion *trans* ⇌ *cis*, then the rate constants for the reacting species are $(1 + 1/K_I)k_{OH}$ and $(1 + 1/K_I)k_s$. As K_I is unknown it is difficult to evaluate the effect of structure on the 'unassisted' nucleophilic attack, unless this has a parallel effect on K_I. Obviously *N*-methylation of the peptide bond has the largest effect on K_I but amongst the esters not containing such a bond, *C*-substituents could also exert small effects with consequent effects on the observed rate constants.

The 'spontaneous' cyclisation of glycylsarcosine methyl ester would proceed at pH 7·92 *ca.* 30—40 times faster than that of the most labile *trans*-esters. The *cis*-*N*-methyl group of the peptide bond is held away from the glycyl function and the molecule would appear to be almost as free from steric constraints in ring formation as the *cis*-isomer of glycylglycine ester. The methyl substituent in alanylglycine methyl ester would offer some steric hindrance in the interaction of the amino-group with the ester carbonyl; this may be the main reason why this ester cyclises much slower than the glycylglycine ester although the two have almost identical p*K* values. The *N*-methyl group of sarcosylglycine methyl ester offers more serious steric hindrance but its greater basicity and nucleophilicity over-rides this; perhaps the formation of the *cis*-*N*-methyl peptide bond is also relatively less unfavourable. The behaviour of glycylalanine methyl ester appeared anomalous (identical results were obtained with two preparations) as it seemed unlikely that the high reactivity was due solely to the slightly higher p*K*.

For cyclisation reactions in general, methyl substitution should raise the initial state free energy relative to that of the ring compound due to entropy effects arising from the restriction in numbers of initial state conformations.[37] This was quantitatively formulated [37] for the equilibrium hexane ⇌ cyclohexane, in which system cyclisation results in a lesser increase in *gauche*-interactions for any of the methylated hexanes than for the parent hydrocarbon. This led to decreases in ΔH and considerable increases in ΔS (both calculated and observed) for the forward reaction, the magnitude depending on the degree and position of substitution. If cyclisation occurs in the rate-determining step, then corresponding changes in $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ would be expected. This effect would tend to promote the cyclisation of the sarcosine and alanine dipeptide esters relative to that of glycylglycine ester. However, other factors must counteract this to cause the observed order of reactivity for the glycine and alanine peptide esters, especially as it was later found that the more substituted compound, L-alanyl-L-alanine methyl ester cyclised slightly more slowly than the L-alanylglycine ester.[38] Perhaps the *C*-methyl group in glycyl-L-alanine methyl ester results in a particularly profitable population of rotamers of the *cis*-form, or in a higher K_I, or in both. Solvation effects, such as those observed in the hydrolysis of peptides [39] and amino-acid esters [15] could also be especially favourable, although the net effect of any of these factors appears to be nullified by *C*-methylation of the *N*-terminal residue.

Self-catalytic Terms.—Again, assuming a p*K'* of 8·2 for glycylsarcosine methyl ester, the k_{gb} would be *ca.* 5·5 l mol^{-1} s^{-1}. Self-catalysis with this ester would be by both *cis*- and *trans*-forms whereas with the other esters only the contribution from the *trans*-form need be considered. In the extended conformation, steric effects at the amino-group appear similar for the *cis*- and *trans*-form of the glycylsarcosine ester, but their basicities could differ.

In comparing the k_{gb} values for the esters it is noteworthy that the glycylsarcosine ester has a particularly high value and the alanylglycine ester a particularly low one, relative to $(a - k_h')$. The values of $k_{gb}\alpha/(a - k_h')$ at pH 7·92 are (in l mol^{-1}): glycylsarcosine methyl ester, *ca.* 1000; glycylalanine methyl ester, 333; alanylglycine methyl ester, 47·5; sarcosylglycine methyl ester, 231; glycylglycine methyl ester, 178; glycylglycine ethyl ester, 285. These are much larger than those of $k_{gb}/k_n < 10$ l mol^{-1} reported for the aminolysis of phenyl acetate,[26] as expected from the disparity in leaving group tendencies. They are also greater than those of *ca.* 40 l mol^{-1} for methyl formate.[22] It has been pointed out [40] that in the aminolysis of phenyl acetate, steric effects are amplified in k_{gb} terms over those in k_n (k_{H_2O}) terms, and that generally secondary amines do not exhibit k_{gb} terms. Also Bruice and McMahon [41] found evidence that steric requirements were especially stringent in nucleophilic attack at the acetyl carbonyl atoms of *N*-acetyldehydrophenylalanyl-L-proline piperazinedione, and even more so at the ring carbonyl atom. Thus, although the spontaneous ring closures of glycylalanine and sarcosylglycine esters proceed at similar rates, the self-catalysed pathway of the glycylalanine ester is much more ready. Steric factors should not be so marked with alanylglycine ester but these are not offset by a high p*K'* value. So although the 'unassisted' cyclisation proceeds slower than with the glycylglycine ester, the self-catalysed reaction is very much slower. L-Alanyl-L-alanine methyl ester behaves similarly to L-alanylglycine ester.[38]

[37] N. L. Allinger and V. Zalkow, *J. Org. Chem.*, 1960, **25**, 701.
[38] J. E. Purdie and N. L. Benoiton, unpublished data.
[39] D. A. Long and J. E. Lillycrop, *Trans. Faraday Soc.*, 1963, **59**, 907.
[40] T. C. Bruice, A. F. Hegarty, S. M. Felton, A. Donzel and N. G. Kundu, *J. Amer. Chem. Soc.*, 1970, **92**, 1370.
[41] T. C. Bruice and D. M. McMahon, *Biochemistry*, 1972, **11**, 1273.

Catalysis of Added Amine.—The markedly reduced catalytic efficiency of TRIS, and to a lesser extent, of sarcosylglycine (as shown by negative deviations in the Brønsted plot) can be attributed largely to steric factors. TRIS and t-butylamine both exhibited lower activities as nucleophiles in the aminolysis of *p*-nitrophenyl acetate [42] and this was likewise explained. The apparent sensitivity of the cyclisation process to steric hindrance together with the intermediate value of β (*ca.* 0·35) suggests that the rate-limiting step does not consist of a simple proton transfer such as, for example, from a preformed zwitterionic tetrahedral intermediate.

Saponification of the Esters.—The saponification of alkyl esters of carboxylic acids is considered to be a nucleophilic reaction [43] and alkyl esters of amino-acids, *N*-substituted and *N*-unsubstituted, are readily saponified by dilute alkali at 25°.[15] The rates for the cationic forms of the esters were *ca.* 100 times greater than for the unprotonated forms [44] and this factor was even larger for *N*-methylamino-acid esters.[15] For the dipeptide esters, it appears that the cationic form is still saponified faster than the uncharged form as the ratio of hydroxide ion concentration at pH 8·52 to that at 7·92 is 3·98 whereas the corresponding ratio of k_h' varies from 3·32 to 3·79. Rough estimates gave $k_h^+/k_h \leqslant 2$. This factor was expected to be less than for the amino-acid esters as there can be no corresponding inductive effect at the ester carbonyl atoms (which was suggested as a major factor with the amino-acid esters [15]) and, in the *trans*-forms, the terminal $\overset{+}{N}H_3$ cannot approach the reaction site. Assuming that the saponification at pH 7·92 proceeds exclusively by nucleophilic attack of hydroxide ion, and obtaining the concentration of the latter from titration curves of the solvent, the second-order saponification constants for the methyl esters of glycylglycine and glycylalanine are, respectively, 2·7 and 1·2 l mol^{-1} s^{-1}, compared with 1·72 and 0·94 l mol^{-1} s^{-1} for the methyl esters of *N*-acetylglycine and *N*-acetylalanine respectively.[15] That is, the ester carbonyl carbon atom in glycylalanine ester is less reactive than that in glycylglycine ester as regards nucleophilic attack by hydroxide ion although the intramolecular aminolysis proceeds with greater facility. This may be partly due to the more rigid geometry of the system imposed by ring closure; Dreiding models show that the methyl group is forced back as the terminal NH_2 approaches.

N-Methylation of esters of *N*-acetylamino-acids reduced their saponification rates [15] by a factor of two, so glycylsarcosine methyl ester should be more stable towards saponification than glycylglycine methyl ester. This could not be verified.

The saponification rates of the esters of alanylglycine and sarcosylglycine were similar to those of glycylglycine methyl ester, presumably because the side-chain methyl groups are remote from the reaction site.

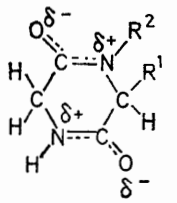
[42] W. P. Jencks and J. Carriuolo, *J. Amer. Chem. Soc.*, 1960, **82**, 1778.
[43] S. L. Johnson, *Adv. Phys. Org. Chem.*, 1967, **5**, 237.
[44] R. W. Hay and P. J. Morris, *J. Chem. Soc.* (*B*), 1970, 1577.

The rate constant for glycylglycine methyl ester is about twice that of unprotonated glycine methyl ester ($k = 1{\cdot}11$ l mol^{-1} s^{-1}), the most labile of the amino-acid esters.[45]

Although esters activated in their acyl group are subject to general base-catalysed hydrolysis,[16] there was no evidence of this with glycylglycine ethyl ester under our experimental conditions.

Ring-opening of the Piperazinediones.—The direction of this reaction was interesting. In 0·1—1·0N-alkali at room temperature, complete ring-opening was possible with negligible dipeptide fission. The greater lability (towards hydrolysis) of piperazinediones than *endo*-peptide bonds in polypeptides has been taken as evidence of strain in the ring.[46] In aqueous solution, the ring is probably not a rigid structure [47] but several nonplanar conformations are possible, at room temperature for both cyclo-(Gly–Gly–) and its *C*-methylated derivatives.

For cyclo-(Gly–Sar–) (I; $R^1 = H$, $R^2 = Me$) the ratio of rate constants for the splitting of the bonds –Sar–Gly– and –Gly–Sar– was 3·4 : 1. This is in the expected direction in view of the inductive and steric effects of the *N*-methyl group. In addition, if the rate limiting step is the breakdown of the tetrahedral intermediate as in



(1)

amide hydrolysis,[48] fission of the –Sar–Gly– bond would be favoured as the glycyl amino would be the intrinsically better leaving group. Furthermore, as the piperazinedione contains *cis*-peptide bonds, the product containing the higher proportion of *cis*-isomer will be energetically favoured. In *cyclo*-(Ala–Gly–) (I; $R^1 = Me$, $R^2 = H$) inductive and steric effects would favour fission of the –Gly–Ala– bond but in fact glycylalanine is produced 2·1 times as fast as alanylglycine. Perhaps the unexpectedly high rate of cyclisation of glycylalanine methyl ester and the preferential ring opening to glycylalanine are not unrelated.

APPENDIX

Computation of the Kinetic Parameters a and b (with S. Walter, *Department of Epidemiology, University of Ottawa*).—Equation (6) can be written as (12) where $A =$

$$Y = A + BR^T \qquad (12)$$

$-bE_0'/a$, $B = 1 + bE_0'/a$, and $R = e^a$.

By making observations on Y at various times we may

[45] R. W. Hay and L. J. Porter, *J. Chem. Soc.* (*B*), 1967, 1261.
[46] B. D. Sykes, E. B. Robertson, H. B. Dunford, and D. Konasewich, *Biochemistry*, 1966, **5**, 697.
[47] F. S. Richardson, R. Strickland, and D. D. Shillady, *J. Phys. Chem.*, 1973, **77**, 248.
[48] R. L. Schowen, H. Jayaraman, L. Kerschner, and G. W. Zuorick, *J. Amer. Chem. Soc.*, 1966, **88**, 4008.

estimate the parameters of interest, a and b, through some regression procedure. Because of the constraint (13) we

$$B = 1 - A \quad (13)$$

would expect the estimators of a and b to have an asymptotic correlation, -1. This, combined with the fact that we have a non-linear (in fact exponential) model, implies that the estimates of a and b are also statistically dependent, even though they may have no physical relationship. Of the BMD computer programs, the two which are most suited to handle this model are the X85 and 06R programs.

The former is a general regression program for any non-linear model, and computes a weighted least-square fit to any user-supplied function by using the numerical Gauss–Newton procedure on the partial derivatives with respect to the model parameters, also user-supplied. The 06R, as an alternative, deals specifically with asymptotic models (as we have here), but has the disadvantage that it is not possible to specify any constraints of dependency between the parameters A and B. However, the method of computation is superior, using standard maximum likelihood principles to derive explicitly the information matrix and its inverse, the variance–covariance matrix. Again an iteration technique on the estimates is used, with these matrices being calculated at each stage. In practice, because of the higher specificity of the 06R program to this situation, we would expect a faster convergence to stable estimates of the model parameters, and we might also expect a better ' fit ' to the data than in the X85.

Both programs provide output on the original observed data values together with the regression predicted values, the estimates of the parameters, their standard errors, and either a variance–covariance matrix (06R) or a correlation matrix (X85). However, the 06R also provides a visual plot of the observed and predicted values, and a useful analysis of the variance table which could be used to test the overall ability of the model to describe the data adequately. The X85 does provide a matrix of the asymptotic correlations of the parameter estimates, but in this case there seems to be no simple way of testing the adequacy of the model by, for example, inspecting the departure of the asymptotic correlation of A and B from the ideal value of -1 which we would expect in view of the statistical dependence mentioned earlier.

Estimates of b were obtained from second-order plots for the initial part of the reaction and thence a was estimated using equation (6). The X85 program requires initial estimates of both these values and the 06R program needs the approximate value of R (*i.e.* e^a). The X85 provides standard errors on a and b directly but with the 06R, these have to be evaluated from the standard errors on R, A, and B. This was done by use of equations (14) and (15). It is less straightforward to estimate the standard error of b. If it is accepted that a and b have a perfect negative correlation of -1 (this is borne out by the X85 output and also from considerations of the statistical dependence) then the

$$\mathrm{Var}(a) = \mathrm{Var}(R)\,\frac{1}{R} + \mathrm{O}(R^2) \quad (14)$$

$$\therefore \text{standard error } (a) \simeq \frac{\text{standard error } (R)}{R} \quad (15)$$

estimates of b and a should have a constant ratio; this may be estimated by A/E_0', thus giving equation (16).

$$\text{standard error } (b) \simeq \frac{A}{E_0'} \,.\, \text{standard error } (a) \quad (16)$$

For any set of data, the X85 gave (worked out manually) a sum of squares about the fitted curve significantly larger than the corresponding sum from 06R indicating that the latter achieved the better fit. Also, the X85 tended to consistently underestimate data values at early times and overestimate at subsequent times rather than giving a random residual pattern which in general was obtained with 06R. This is because X85 minimises the sum of squares about the regression curve, usually with equal weight being given to each point; thus the goodness of fit to the data in the early part of the curve depends on the number of observations included from the later asymptotic region. The inclusion of too many points, each conveying approximately the same information about the parameters, will be to the detriment of the regression fit at earlier points on the curve. Even so, both programs easily provide a statistically adequate fit although 06R would probably give the more reliable parameter estimates. The superior computation method of 06R led to faster convergence to the final parameter estimates (but this does not necessarily result in less computing time as the procedure also involves the inversion of the several information matrices).

The ideal program for our situation is one which would estimate the parameter A in the model (17) which is

$$Y = A + (1 - A)e^{aT} \quad (17)$$

equivalent to (12) with constraint (13) built in. It is possible in principle to follow a similar argument to that used in the 06R program using the method of maximum likelihood, but the non-linearity of the model presents great difficulties as compared with the linear situation, where some constraints can be built into the exact estimation procedure.

This work was supported by a grant from the Medical Research Council of Canada of which N. L. B. is an associate. We thank Dr. R. E. Williams, National Research Council, Ottawa, for advice on syntheses and Mrs R. Demayo for preparing the sarcosine compounds.

[3/923 *Received*, 7*th May*, 1973]