A pH-dependent cyanate reactivity model: application to preparative *N*-carbamoylation of amino acids

Jacques Taillades,* Laurent Boiteau, Isabelle Beuzelin, Olivier Lagrille, Jean-Philippe Biron, Willy Vayaboury, Odile Vandenabeele-Trambouze, Olivia Giani and Auguste Commeyras

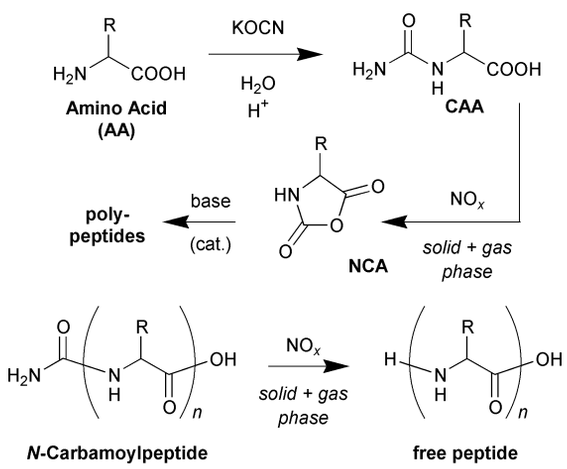
Organisation Moléculaire – Evolution et Matériaux Fluorés, Chemistry Department, University of Montpellier 2, CC 017 – Place E. Bataillon – 34095 Montpellier cedex 5, France. E-mail: lboiteau@univ-montp2.fr

Received (in Cambridge, UK) 20th July 2000, Accepted 30th April 2001
First published as an Advance Article on the web 7th June 2001

Recent developments in peptide synthesis have underlined the importance of optimising, on a preparative scale, the *N*-carbamoylation of amino acids by aqueous cyanate. To this purpose, a theoretical model of aqueous cyanate reactivity was designed. The parameters of the model were evaluated, for various pH and temperatures, from a critical survey of the literature, together with additional experimental data. Computer-simulated kinetics based on this model showed the reaction efficiency to be significantly dependent on pH, and suggested optimum conditions to be moderate temperatures and pH 8.5–9. Discussion of the practical convenience of these theoretical results led us to prefer 40–50 °C and a pH range of 7–8 as reaction conditions, thus maintaining reaction times within a few hours. Various *N*-carbamoyl amino acids (ureido derivatives of glycine, L-valine, L-alanine, L-leucine, DL-methionine, N^{ε}-trifluoroacetyl-L-lysine, β-alanine) were thus successfully synthesised on the gram to kilogram scales.

Introduction

In recent decades the *N*-carbamoylation reaction, especially of amino acids, has been considered to be mostly relevant to biochemistry and metabolic pathways. Interest in this reaction has, however, been renewed since the recent discovery of a new synthetic route to amino acid *N*-carboxyanhydride (NCA), starting from *N*-carbamoyl amino acids (CAAs)[1] (Scheme 1).



Scheme 1

Applied to *N*-carbamoyl peptides or to monoalkylureas, the same reaction leads to the free amino function,[2] thus offering possibilities of using carbamoylation as an N-protective group strategy, whereas the CAAs and alkylureas were considered as very stable compounds, the hydrolysis of which requires drastic basic conditions or the use of enzymes. Both strategies require mastering *N*-carbamoylation on a preparative scale. However in the abundant literature on the subject (mostly biochemical studies), the synthetic procedures described provide little detail concerning the efficiency and selectivity of the reaction, or the purification of the products, hence the need to reconsider the subject.

Though other organic synthetic methods exist, such as trans-carbamoylation from urea degradation,[3] hydantoin hydrolysis by strong bases[4] or enzymatic systems,[5] the simplest and cheapest way to prepare CAAs is by reacting the free amino acid with buffered aqueous mineral cyanate.[6,7] Most authors have worked at pH around 5–6, which appears not to be the optimum conditions, since the *N*-carbamoylation reaction is in competition with cyanate hydrolysis and urea formation. While cyanate hydrolysis only requires the use of a larger excess of cyanate, urea formation (a consequence of cyanate hydrolysis) can be a major disadvantage for further product purification, especially when the CAA is very soluble in water. All reaction rates are highly pH-dependent, which led us to reconsider their kinetics and to draw up a global cyanate reactivity model, which was examined by means of a computer-simulated time evolution.

Cyanate reactivity kinetics

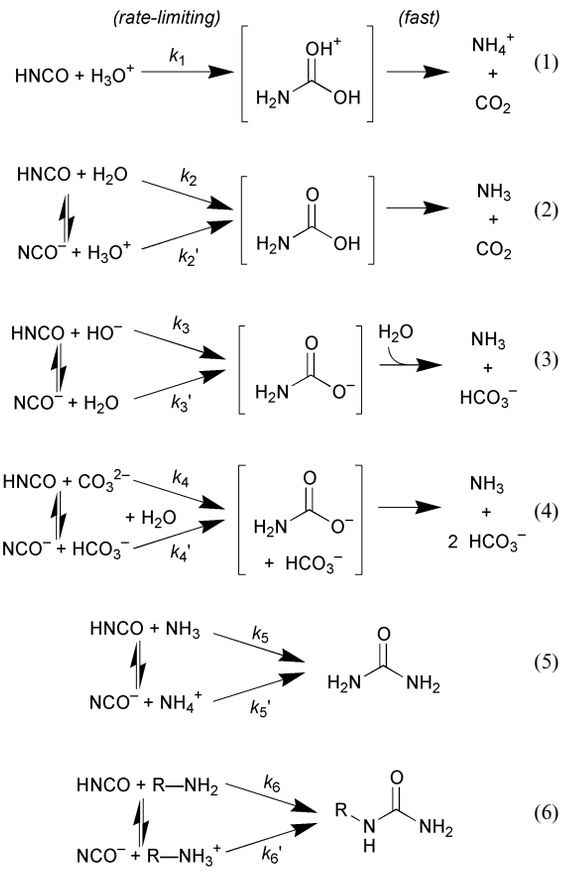
It was not our intention to carry out a complete kinetic study of cyanate reactivity in aqueous solution, the kinetics and mechanism of which have been studied by various authors during the past decades, though under rather different conditions. Our objective was rather to obtain a workable kinetic model over a large pH range for temperatures above 40–50 °C, a temperature range that appeared suitable for amino acid *N*-carbamoylation after a series of preliminary experiments.

It is now accepted that the reactivity of cyanate with water, as well as ammonia and amines, actually involves a protic or nucleophilic attack of isocyanic acid, forming carbamates or ureas. This slow-rate step is followed by fast decomposition of the carbamates into ammonia and carbon dioxide, whereas ureas are much more stable. All reactions have been found to be first-order with respect to each of their reactants. In addition, the carbonate anion CO_3^{2-} strongly catalyses HNCO acid hydrolysis,[8] the rate increment also being first-order with respect to HNCO and CO_3^{2-}.

In the scope of this study, the possible reversal of reactions (1)–(6) will be neglected: since ammonia and carbon dioxide

DOI: 10.1039/b005856o

concentrations remain low, carbamate formation due to them remains limited; (alkyl)urea hydrolysis is very slow, and is negligible below 50–60 °C for reaction times not exceeding a few hours. The general rate law can thus be written as eqn. (7), with the rate constants k_1–k_6 relating to reactions (1)–(6) respectively.



$$v = [\mathrm{HNCO}] \times (k_1 \times [\mathrm{H_3O^+}] + k_2 + k_3 \times [\mathrm{HO^-}] + k_4 \times [\mathrm{CO_3^{2-}}] + k_5 \times [\mathrm{NH_3}] + k_6 \times [\mathrm{RNH_2}]) \quad (7)$$

Considering proton transfer in acid–base equilibria to be much faster than reactions (1)–(6), the rate law can be equivalently rewritten as eqn. (8), in terms of cyanate anion concentration and related rate constants k_2'–k_5'.

$$v = [\mathrm{HNCO}] \times k_1 + [\mathrm{NCO^-}] \times (k'_2 \times [\mathrm{H_3O^+}] + k'_3 + k'_4 \times [\mathrm{HCO_3^-}] + k'_5 \times [\mathrm{NH_4^+}] + k'_6 \times [\mathrm{RNH_3^+}]) \quad (8)$$

The different rate constants have been determined at various temperatures by different authors, who presented them either as k_i or as k_i'. These data and references are compiled in Table 1. However due to the scarcity of data within the temperature range we are interested in, an extrapolation was necessary, as well as a critical analysis of literature data, since the two series are not equivalent in practice. We performed some complementary measurements at pH 5–6.5 (Table 2), a pH range where reaction (2) is highly preponderant; our results proved to be in agreement with the data from the literature.

Apparent rates

So that it can be more easily used in a pH-dependent kinetic model, the rate law is best expressed as a function of the total species concentrations (acid + base forms), and of the pH-dependent apparent rate constants k_i^a. Thus, with $\mathbf{NCO^t}$, $\mathbf{CO_3^t}$, $\mathbf{NH_3^t}$ and $\mathbf{AA^t}$ standing for the total concentrations of cyanate, carbonate, ammonia and amino acid respectively, the rate law is expressed as follows in eqn. (9),

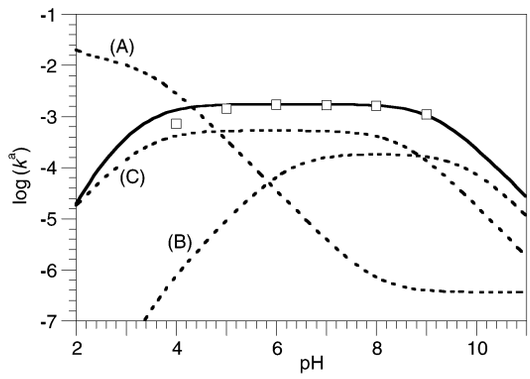


Fig. 1 Apparent rates at 50 °C, as a function of pH, of glycine *N*-carbamoylation k_6^a (solid line; squares: experimental values from Table 5), and of other cyanate reactions (dotted lines, calculated using data from Tables 3 and 4): cyanate hydrolysis k_0^a (A), carbonate catalytic increment k_4^a (B), urea formation k_5^a (C).

$$v = \frac{\mathrm{d}\mathbf{NCO^t}}{\mathrm{d}t} = \mathbf{NCO^t} \times (k_0^a + k_4^a \times \mathbf{CO_3^t} + k_5^a \times \mathbf{NH_3^t} + k_6^a \times \mathbf{AA^t}) \quad (9)$$

k_0^a thus representing the apparent rate of cyanate hydrolysis (including acid- and base-catalysis). The apparent rate constants k_i^a are related to the k_i and k_i' rate constants, and to the respective dissociation constants K_{Cy}, K_w, K_{1Car}, K_{2Car}, K_{Am}, K_{1AA}, K_{2AA} (of cyanic acid, water, 1st and 2nd dissociation of carbonic acid, ammonium, and 1st and 2nd dissociation of amino acid respectively). With h standing for the H_3O^+ concentration, k_0^a–k_6^a are respectively expressed according to eqn. (10)–(13), in which the relationships between k_i and the respective k_i' are also made explicit. The pH-dependent shapes of these apparent rates are exemplified in Fig. 1.

$$k_0^a = \frac{k_1 \times h + k_2 + k_3 \times \frac{K_w}{h}}{1 + \frac{K_{Cy}}{h}} = \frac{k_1}{1 + \frac{K_{Cy}}{h}} + \frac{k_2' \times h + k_3'}{1 + \frac{h}{K_{Cy}}} \qquad \frac{k_2'}{k_2} = \frac{1}{K_{Cy}} \quad \frac{k_3'}{k_3} = \frac{K_w}{K_{Cy}} \quad (10)$$

$$k_4^a = \frac{k_4}{\left(1 + \frac{K_{Cy}}{h}\right) \times \left(1 + \frac{h}{K_{2Car}}\right)} = \frac{k_4'}{\left(1 + \frac{h}{K_{Cy}}\right) \times \left(\frac{h}{K_{1Car}} + 1 + \frac{K_{2Car}}{h}\right)} \qquad \frac{k_4'}{k_4} = \frac{K_{2Car}}{K_{Cy}} \quad (11)$$

$$k_5^a = \frac{k_5}{\left(1 + \frac{K_{Cy}}{h}\right) \times \left(1 + \frac{h}{K_{Am}}\right)} = \frac{k_5'}{\left(1 + \frac{h}{K_{Cy}}\right) \times \left(1 + \frac{K_{Am}}{h}\right)} \qquad \frac{k_5'}{k_5} = \frac{K_{Am}}{K_{Cy}} \quad (12)$$

$$k_6^a = \frac{k_6}{\left(1 + \frac{K_{Cy}}{h}\right) \times \left(1 + \frac{h}{K_{2AA}}\right)} = \frac{k_6'}{\left(1 + \frac{h}{K_{Cy}}\right) \times \left(\frac{h}{K_{1AA}} + 1 + \frac{K_{2AA}}{h}\right)} \qquad \frac{k_6'}{k_6} = \frac{k_{2AA}}{K_{Cy}} \quad (13)$$

Table 1 Kinetic data from the literature for the cyanate degradation model (the water activity is aggregated into the constants)

T/°C	$\log(k_1$/L mol^{-1} s$^{-1})$	$\log(k_2'$/L mol^{-1} s$^{-1})$	$\log(k_2$/s$^{-1})$	$\log(k_3'$/s$^{-1})$	$\log(k_3$/L mol^{-1} s$^{-1})$	$\log(k_4'$/L mol^{-1} s$^{-1})$	$\log(k_4$/L mol^{-1} s$^{-1})$	$\log(k_5'$/L mol^{-1} s$^{-1})$	$\log(k_5$/L mol^{-1} s$^{-1})$	pK_{Cy}	Ref.
0		(*−0.045*)[bc]	−3.745							3.70[a]	9
0										3.65	10
0										3.92	11
1.5	−1.845										9
4	−1.785	(*−0.761*)[bc]	−4.301							3.54[a]	12
10	−1.545	*−0.042*[b]	−3.582							3.54[a]	12
10										3.68	9
13.1	−1.390										9
18	−1.220	(*0.434*[b]	−3.106			−5.000				3.54	12
25		(*2.188*)[bc]	−1.102	−7.712	2.991			*−4.452*	1.748	3.29	7
25										3.47	13
25.5	−0.907										9
27										3.73	9
30	(−0.364)[c]	*1.319*[b]	−2.478	−7.535		*−4.039*	2.214			3.80[6]	8
30		0.68	−2.78							3.46[a]	Our work
50		1.56	−1.89							3.45[a]	Our work
60		1.826	−1.569	−5.975	3.342	−3.535	2.590	−2.903	1.997	3.70[a]	14
65				−5.745							9
70				−5.548							12
80		2.367	−1.004	−5.062	3.838	−2.925	3.170	−2.035	2.417	3.70[a]	14
80.9				−5.046							9
82				−5.022							12
94				−4.545							12
100				−4.499							8
100				−4.298							9

[a] Extrapolated by author (not measured at this temperature). [b] *Italic* data (k_i') are recalculated from the respective k_i. [c] Data in brackets were excluded from our model fit.

Dissociation constants

Temperature-dependent expressions of the dissociation constants of the species involved were found in the literature[15] and are summarised in Table 3, as second-order polynomials that usually fit experimental data very well (a third-order polynomial was necessary for K_{1Car}).

The dissociation constant of HNCO has never been accurately measured above 40 °C, due to its intrinsic instability in water. Even below this temperature, the data found in the literature (Table 1) show a serious discrepancy: neither a polynomial nor a homographic function of temperature fitting these data yields realistic values when extrapolated above 40 °C. While the expression listed in Table 3 is satisfactory below 30–35 °C, we assumed pK_{Cy} to be roughly 3.45 above 35–40 °C, and found a compromise to minimise the consequences of this approximation.

Rate constants of cyanate degradation, choice of the model

The constants k_i and k_i' cannot be considered as equivalent: as first pointed out by Kemp and Kohnstam,[14] data for k_i' values are probably much more reliable than k_i values.This can be seen in Table 1 in the much better correlation between the k_i' than between the k_i, the latter (usually calculated from measured k_i^a) being highly affected by the uncertainty concerning K_{Cy}. Most measurements of k_2–k_5 were taken above pH 4, where the cyanate anion, rather than cyanic acid, is the major species; and where the $k_i^a : k_i'$ ratio does not involve $K_{Cy'}$, unlike the $k_i^a : k_i$ ratio. Furthermore, when k_i' values not provided by the literature were recalculated from the respective k_i (using the author's values for dissociation constants: data in italics in Table 1), they correlated very well with those directly mentioned in the literature.

This determined our preference for using k_i'-based kinetics in our model, despite the closer physical realism of k_i: inaccuracies in K_{Cy} therefore affect no significant terms above pH 4 (CAA synthesis is inefficient below pH 4, anyway). Extrapolated temperature-dependent expressions of rate constants k_i' in the form $\log(k) = a - b/T$ were thus evaluated using least-square fitting, and are shown in Table 4. Some isolated data were however excluded when notoriously inaccurate, k_2' values found at low temperature by Lister[9] and Jensen,[12] or obviously inconsistent with other data, k_2' value found by Williams and Jencks,[7] k_1' value found by Vogels *et al.* [8] (in brackets in Table 1).

Table 2 Experimental cyanate decomposition rate at various pH and temperatures; $[NCO^-]_0 = 1.6$ mM

	30 °C			50 °C	
pH	5	5.5	6	6	6.5
$\log(k_0^a/s^{-1})$	−4.45	−4.58	−5.42	−4.57	−4.82
$\log(k_2'/L\ mol^{-1}\ s^{-1})$	0.55	0.92	0.58	1.43	1.68
Mean $\log(k_2')$	0.68			1.56	

Rate constant of amino acid *N*-carbamoylation ($k_{6'}$)

The rate of *N*-carbamoylation of various amines and amino acids by cyanate was first studied by Williams and Jencks[7] at 25 °C in the scope of a mechanistic study, with no evaluation of the activation energy. Since we needed to work at higher temperatures we performed a new series of measurements on several amino acids. To minimise the influence of cyanate degradation, initial reaction rates only were evaluated at various pHs. The results are summarised in Table 5 and in Fig. 1 for

Table 3 Temperature (°C) dependent expression of dissociation constants $pK = A - BT + CT^2 - DT^3$ calculated from literature data by least-square fitting

	pK_w	pK_{Am}	pK_{1Car}	pK_{2Car}	pK_{Cy}	pK_{1Gly}	pK_{2Gly}
A	14.9394	10.0820	6.5787	10.3190	3.9208	2.4403	10.487
10^2B	4.2044	3.6123	1.3345	1.4713	2.8700	4.7818	3.071
10^4C	1.6899	1.0536	1.9039	0.99487	4.3014	0.48047	0.95328
10^6D	—	—	0.81746	—	—	—	—

Table 4 Calculated temperature (°C) dependent expressions of rate constants $\log(k_i') = a - (b/T)$, with respective correlation coefficient r^2 and activation energies $E_A/kcal\ mol^{-1} = b \times \ln(10) \times R/kcal\ mol^{-1}\ K^{-1}$. Units for k_i' are the same as in Table 1

	$\log(k_1/L\ mol^{-1}\ s^{-1})$	$\log(k_2'/L\ mol^{-1}\ s^{-1})$	$\log(k_3'/s^{-1})$	$\log(k_4'/L\ mol^{-1}\ s^{-1})$	$\log(k_5'/L\ mol^{-1}\ s^{-1})$
a	9.9542	11.7720	9.1574	5.8098	10.9340
b	3249.2	3301.5	5041.1	3082.7	4592.5
r^2	0.99613	0.94579	0.99734	0.93061	0.99859
$E_A/kcal\ mol^{-1}$	14.9	15.1	23.1	14.1	21.0

Table 5 Experimental conditions and results for amino acid *N*-carbamoylation rate measurement. Constants k_6' and k_6^a refer to eqn. (13)

Amino acid	T/°C	pK_{2AA}	pH	$[AA]_0$/mmol L^{-1}	$[KOCN]_0$/mmol L^{-1}	$10^6k_6^a$/mol^{-1} L s^{-1}	$\log(k_6'/mol^{-1}\ L\ s^{-1})$	mean $\log(k_6')$
Gly	30	9.652[15]	5	56	140	340	−3.476	
			5	56	280	319	−3.504	
			5	56	420	326	−3.495	−3.49
Gly	50	9.190[15]	5	56	140	1420	−2.825	
			6	56	100	1700	−2.757	
			7	56	140	1680	−2.760	
			8	56	100	1630	−2.741	
			9	56	100	1080	−2.738	−2.78
Val	50	9.142[15]	6	11.2	15.9	1190	−2.925	
			6.5	11.2	15.9	1300	−2.887	
			6.5	300	373	1300	−2.887	−2.90
β-Ala	50	9.605[15]	8	60	102	1420	−2.85	
Lys(TFA)	50	8.9	8	70	77	1364	−2.87	
Thr	50	8.55[16]	8	50	100	3530	−2.45	

those concerning glycine, measured apparent rates k_6^a being converted into k_6' according to eqn. (13). The first-order kinetics were confirmed for glycine and valine by the linear dependence of rate on initial cyanate concentration at constant pH. Since *N*-carbamoylation of glycine was initially intended to serve as a model for theoretical calculations, this case was more thoroughly investigated. The rate constants obtained for glycine at 30 and 50 °C are consistent with that found by Williams and Jencks at 25 °C, and correspond to an activation energy of 17.2 kcal mol^{-1} for k_6' (7.1 kcal mol^{-1} for k_6, assuming pK_{Cy} to be 3.45 at 50 °C).

This study also aimed to demonstrate a dependence between the amino acid *N*-carbamoylation rate and the amino group dissociation constant, in order to allow more accurate computerised kinetic calculations for various amino acids. In their kinetic study,[7] Williams and Jencks found a good correlation at 25 °C between amine basicity (pK_A) and *N*-carbamoylation rate k_6, with an approximate law $\log(k_6) = 0.15 + 0.3 \times pK_A$, which may also be rewritten, according to eqn. (13), as $\log(k_6') = 3.44 - 0.7 \times pK_A$; this correlation was interpreted as an effect of increasing nucleophilicity. Practically, this correlation means that the maximum apparent *N*-carbamoylation rate (according to eqn. (13), $k_{6\,max}^a = k_6'$) must be expected to decrease with increasing amine basicity.

Unfortunately, although following a similar trend as at 30 °C, the experimental rate constants k_6' we measured for different amino acids at 50 °C exhibited a very bad correlation to amino group pK_A; therefore no realistic linear correlation could be assigned from our measurements. For theoretical calculations we can therefore at least assume that, for common amino acids, the rate constant k_6' will vary by at most one order of magnitude from the value found for glycine.

N-Carbamoylation of amino acids

Competition between different reactions

The pH-to-apparent rate profiles of cyanate reactivity (amino acid is exemplified by glycine) are shown in Fig. 1. The *N*-carbamoylation apparent rate constant k_6^a is maximum and roughly constant within a pH range between pK_{Cy} and pK_{2AA}, where $k_6^a = k_6'$. The cyanate hydrolysis rate (k_0^a) shows continuous decay as pH increases, then remains constant above pH 8. The highest selectivity of the former reaction against the latter thus occurs around pH 8.5.

The urea formation apparent rate constant k_5^a and carbonate-catalysis rate increment k_4^a have similar profiles to k_6^a, with maximum plateaux between pK_{Cy} and pK_{Am} for k_5^a ($= k_5'$), and between pK_{1Car} and pK_{2Car} for k_4^a ($= k_4'$). These maximum apparent rate constants are situated no more than one order of magnitude lower than k_6^a [$\log(k_5') = -3.28$, $\log(k_4') = -3.73$ at 50 °C]; no larger difference in apparent rates can be expected with other amino acids and at any pH. Therefore a good selectivity of reaction (6) *versus* the others requires that the concentrations of ammonia and carbonate should remain very low. Since both species are by-products of cyanate hydrolysis [reactions (1)–(3)], this confirms the necessity of limiting the latter reaction as much as possible, by operating above pH 6 and as close as possible to pH 8.5.

Computer simulation

To refine these conclusions, we designed and performed a computer simulation of the evolution of the system, based on the (temperature-dependent) kinetic and dissociation constants discussed above, using glycine as a model amino acid. A stepwise integration algorithm with variable time increments was implemented to integrate eqn. (7), pH and temperature remaining constant throughout the simulation. The simulation was run until consumption of either glycine or cyanate was (almost) complete, the latter case occurring when molar excess cyanate was introduced. Examples of simulation results are displayed in Fig. 2.

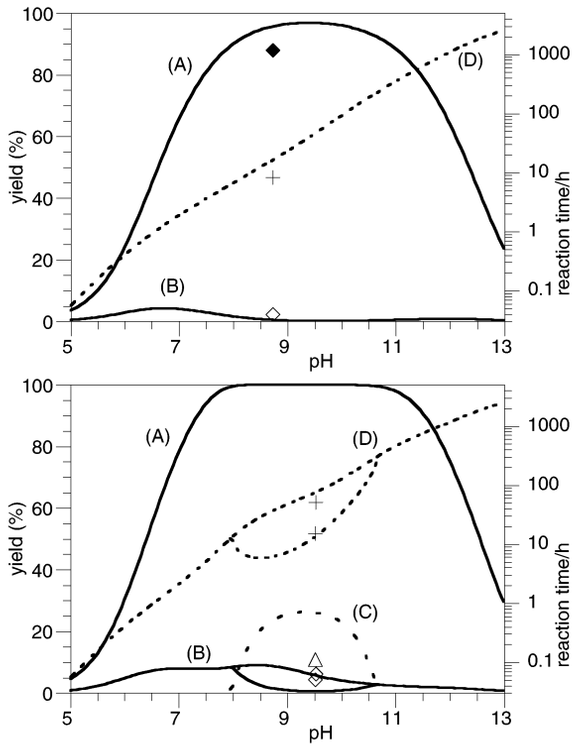


Fig. 2 Computer-simulated *N*-carbamoylation of glycine at 50 °C, with starting conditions $[Gly]_0 = 0.5$ M, $[NCO]_0 = 0.5$ M (top), 0.65 M (bottom); simulations performed at constant pH (curves) or with free pH variations (isolated marks, plotted at final pH), stopped once 99% of initial cyanate is consumed. Product yields (mol% relative to initial glycine) are plotted as a function of pH: *N*-carbamoylglycine (A; ◆), urea (B; ◇), unreacted cyanate (C; △); reaction time (D; +). Split lines (B) and (D), and duplicate marks in the bottom graph correspond to consumption of either 99% initial cyanate (upper) or 99.5% initial glycine (lower) at the end of reaction.

Validity limits of the model. Due to uncertainty concerning pK_{Cy}, and therefore HNCO/NCO$^-$ concentrations below pH 5, the results of our model are not very accurate in this pH range, where cyanate hydrolysis is anyway too fast for amino acid *N*-carbamoylation to be efficient. The practical consequences are thus of minor importance. Above 60–70 °C, as well as for long reaction times, reverse reactions of (1)–(6), not taken into account in our model, are no longer negligible. Calculated reaction times and CAA yields are thus probably below those that could be physically observed; however we still consider our simulation results to be qualitatively pertinent.

Simulations results. These showed that no good yield of CAA can be obtained at pH < 6 (except by using a considerable excess of cyanate, which would be very inconvenient on a large scale). At 0.5 M glycine concentration, and using only 1 equivalent cyanate, 95% conversion (at most) is attained after *ca.* 30 hours at 50 °C and pH 9. Using excess cyanate (1.3 equivalent), full glycine conversion can be obtained above pH 8 before complete cyanate hydrolysis (in *ca.* 10 h). Starting with 0.2 M glycine (this corresponds to the solubility limit of the least polar amino acids at 50 °C), and using 1 equivalent cyanate, 95% conversion (at most) is reached at 50 °C and pH 9.4, but after *ca.* 70 hours. Full conversion may then be reached above pH 8.2 within *ca.* 10–20 h, when using 1.5 equivalent cyanate. From these results it obviously appears that pH 5–6, as proposed in the literature,[6] is not the best reaction condition. If the reaction is stopped once 99.5% of initial glycine has been converted, the minimum urea formation occurs at pH 9–10, which is higher than the value (pH 8.5) predicted by the appar-

ent rate ratio k_6^a/k_0^a. This, however, corresponds to the shortest times required to attain 99.5% conversion. Complete excess cyanate consumption leads to an increase in urea production, the maximum of which (*ca.* 30–35 mol% of the initial excess cyanate) occurs around pH 8.5 (Fig. 2). Whatever the pH, CAA/urea selectivity decreases as the temperature increases. However, reactions carried out at temperatures lower than 40–50 °C and/or pH higher than 9 require a considerable time to reach completion (more than 100 h).

Additional calculations were performed to examine the reaction without pH regulation. Considering acid–base equilibria to be much faster than reactions of cyanate, the simulation algorithm was modified so that the pH was corrected after each integration time step, using the first-order differential of the electrical neutrality equation (Σ[anions] = Σ[cations]) against [H^+] concentration. The initial pH of the reaction mixture was calculated using the same equation with initial species concentrations. In accordance with experimental data, calculations show a continuous increase in pH, very fast at the beginning of the reaction, then slower towards an asymptotic value around 9–10, this value mostly depending on the initial stoichiometric excess of cyanate/glycine.

Results in terms of CAA yield, urea formation and reaction time are similar to those obtained using the constant-pH algorithm with the pH fixed at the asymptotic value (isolated marks in Fig. 2). We only observed a slight increase in reaction time and urea formation, the latter mostly occurring at the beginning of the reaction, when pH is still below 8. However, these results must be considered as merely qualitative since this algorithm provides less accurate results than the constant-pH algorithm. For instance, the calculated asymptotic pH (after 10–15 h reaction at 50 °C) from an initial solution of 0.5 M glycine and 0.65 M cyanate is 9.53, while the experimental value under the same conditions is pH 9.98.

Application to preparative synthesis

According to our computer simulations, optimum CAA yield and CAA/urea selectivity (a critical requirement for purification when the CAA is very soluble in water) should be obtained by reacting 1.2–1.5 equivalent cyanate, at moderate temperatures and with pH regulated to 8.5–9, stopping the reaction once *ca.* 99% amino acid is converted. The remaining excess cyanate can be conveniently hydrolysed without producing more urea by rapidly acidifying the medium to pH 2–3. Practically speaking however, these optimum conditions are not very convenient for preparative purposes, since low temperatures and basic pH induce long reaction times (over 100 h). Buffering the reaction at medium to basic pH results in a significant increase in salt content, the separation of which from water-soluble CAA is as difficult as from urea. Compromises had thus to be found.

Concerning low-polarity CAAs, weakly soluble in acidified water (pH < 3), neither urea nor salts are a major problem, provided they are only present in moderate amounts. For practical convenience on a laboratory scale, more cyanate was used (1.5–2 equiv.) than theoretically necessary, and a pH around 7.5 and temperatures of 50 °C were applied, so that the reaction time remained within 10–24 h; when racemization on α-carbon is not a problem, higher reaction temperatures can be used. Under these conditions, the *N*-carbamoyl derivatives **C-Ala**, **C-Val**, **C-Leu**, **C-Phe**, **C-Met**, **C-Lys(TFA)** were prepared and isolated with good yields through a straightforward procedure. In a typical experiment, the warm solution of amino acid was mixed with cyanate, then adjusted if necessary to the desired pH, which is maintained constant by further controlled addition of acid (*e.g.* HCl or HNO_3). The reaction can easily be followed by HPLC. After completion the medium is chilled, rapidly acidified to pH 2–3 in order to hydrolyse the unreacted excess cyanate, as well as to precipitate the CAA, which is recovered by filtration.

Concerning more polar amino acids (which yield very water-soluble CAAs), by-product salts may in theory be removed by the use of cation-exchange resins. However these must be carefully handled since the CAA readily undergoes irreversible cyclisation into the corresponding hydantoin under acidic conditions (pH < 2); such hydantoin formation seems to be enhanced by the presence of hydrochloric acid. However we obtained a good result for **C-Gly** preparation by carrying out the reaction at 50 °C in a 0.5 M glycine–0.65M cyanate solution, without pH regulation, obtaining *ca.* 95% **C-Gly** and less than 10% urea after 10 h reaction. These results are in accordance with our computer simulations for these conditions (Fig. 2). Most of the product was recovered pure, by simple filtration of the crude reaction mixture through a cation-exchange column. Developments concerning this and applications to other polar amino acids will be dealt with in a forthcoming paper.

Conclusion

In this paper we have reported efficient conditions for preparative amino acid *N*-carbamoylation by aqueous cyanate, and have highlighted the important influence of medium to weakly basic pH. While our computer simulations indicate pH 8.5–9 as an optimum, they also exhibit the robustness of the reaction, which may be achieved over a wider pH range (7–10) with moderate excess of cyanate. In many cases, however, pH 7–8 was preferred as the most convenient on a preparative scale. Investigations are in progress for improving polar, high-functional amino acid *N*-carbamoylation.

Amino acid *N*-carbamoylation may have had important consequences in the molecular origins of life, as part of the emergence of prebiotic peptides.[17] These preliminary results are also necessary for a further study of this reaction within this scope, especially in a CO_2 atmosphere, the prebiotic relevance of which is uncontested. This subject will be discussed in more detail in a future paper.

Experimental

Materials and methods

Melting points were measured using a Buchi 520 apparatus. ^{1}H-NMR spectra were recorded on a Varian EM360 (60 MHz) or a Bruker AC250 spectrometer (250 MHz) in DMSO-d_6 solution. pH regulations were performed using a Methrom Titrino 719S autotitrator. HPLC analyses were carried out using a Varian apparatus set up with a Varian 2510 pump, a Varian 2550 UV detector and a Varian 4290 integrator. Nucleosil RP-C18, 5 µm, 250 × 4.1 mm id, analytic columns (obtained from Shandon, France) were used at room temperature under isocratic conditions (detailed conditions are provided below). Chromatographic silica gel (60 mesh) was obtained from Merck.

Pure water (50 MΩ) was obtained using a milli-Q (Millipore) system after column-exchange deionization. Methanol, acetonitrile and hexane were obtained from Baekrout. Glycine, L-alanine, L-valine, L-leucine, L-phenylalanine, fluorenylmethyl chloroformate (FMOC) and trifluoroacetic acid were obtained from Aldrich. β-Alanine was obtained from Fluka. Methionine was obtained from Aventis Animal Nutrition. ε-Trifluoroacetyl-L-lysine and L-threonine were obtained from Degussa. Sodium hydroxide and silver nitrate were obtained from Prolabo. Hydrochloric acid (36%), nitric acid (70%), acetic acid and triethylamine, were obtained from Carlo Erba. All reagents and solvents were used as received except potassium cyanate, which was washed with methanol then dried *in vacuo* prior to use.

Kinetic measurements

Isocyanic acid degradation. A solution of 78 mg (0.96 mmol) potassium cyanate in 60 mL water was stirred in a thermostated reactor; the pH was regulated (5 to 6.5) using 0.2 M HNO_3 by means of the autotitrator. The reaction was followed by reverse silver titration: at regular time intervals, 5 mL aliquots were taken, cooled and added to 1 mL 0.1 M $AgNO_3$. The AgOCN precipitate was filtered off, washed with water, and the remaining Ag^+ in the filtrate was titrated with 0.01 M HCl, using silver electrode potentiometry.

Glycine *N*-carbamoylation. 20 ml of a 0.1–0.42 mM potassium cyanate solution in water was stirred in a thermostated reactor and the pH was adjusted with 4 M nitric acid. A solution of 126 mg (1.68 mmol) glycine in 10 mL water and adjusted to the same pH was immediately added. The pH was regulated throughout the reaction using 4 M nitric acid. At regular time intervals, aliquots were taken and submitted to FMOC derivatisation prior to HPLC analysis.

FMOC-derivatisation. (Adapted from Clapp *et al.*[18]): a 0.5 mL aliquot was taken, cooled and added to 0.5 mL of 0.2 M borate buffer (pH 7.7) containing 1 mM L-alanine (internal reference). To half of this mixture 0.5 mL of a fluorenylmethyl chloroformate solution (15 mM) in acetone was added and the mixture was stirred for 30 s, then the solution was washed 3 times with 2 mL hexane. 0.1 mL of the aqueous layer was then analysed by HPLC (eluent MeOH–MeCN–acetate buffer: 20 : 30 : 50. The acetate buffer was made from 3 mL acetic acid, 1 mL triethylamine in 0.9 L water, adjusted to pH 4.2 with sodium hydroxide then made up to 1 L with water); t_R/min = 8.5 (**FMOC-Gly**), 11.1 (**FMOC-Ala**) at 0.9 mL min^{-1}.

L-Valine *N*-carbamoylation. The same protocol was used as for glycine, except that FMOC derivatisation was not used: 60 mL of a 15.9–37.3 mM potassium cyanate solution in water was stirred in a thermostated reactor. To this a solution of 126 mg (1.08 mmol) valine in 10 mL water and adjusted to the same pH was immediately added. At regular time intervals, aliquots were taken and analysed by HPLC, after dilution to 1 : 10 in HPLC eluent: water–acetonitrile (90 : 10) + 0.05% TFA; t_R/min = 4.8 (**Val**), 10.1 (**C-Val**) at 0.9 mL min^{-1}.

***N*$^{\alpha}$-Carbamoylation of *N*$^{\varepsilon}$-trifluoroacetyl-L-lysine.** Into a stirred solution of 2.55 g (10.5 mmol) *N*$^{\varepsilon}$-trifluoroacetyl-L-lysine in 150 mL water, warmed to 50 °C and adjusted to pH 8 with solid NaOH, was dissolved 946 mg (11.5 mmol) potassium cyanate at time 0. The pH was regulated to 8 throughout the reaction with 6 M HCl using the autotitrator. At regular time intervals, 0.1 mL aliquots were taken, diluted in pure water and analysed by HPLC, eluent: water–acetonitrile (90 : 10) + 0.05% TFA; t_R/min = 5.4 (**Lys(TFA)**, used as internal reference), 9.7 (**C-Lys(TFA)**) at 1.0 mL min^{-1}.

pK_A of *N*$^{\varepsilon}$-trifluoroacetyl-L-lysine. 25 mL of a 0.02 M solution of *N*$^{\varepsilon}$-trifluoroacetyl-L-lysine in pure water and warmed to 50 °C was titrated with 0.1 M sodium hydroxide, the pH being measured by the autotitrator. The pK_A (8.9 at 50 °C) was directly read from the titration curve (pH *vs.* NaOH mol. added). This value is consistent with that found in the literature[19] at 25 °C.

β-Alanine and L-threonine *N*-carbamoylation. Into a stirred solution of amino acid and α,ω-dicarbamoyllysine (8 mM, internal reference) in 150 mL water, warmed to 50 °C and adjusted to pH 8 with solid sodium hydroxide (for Thr) or 6 M hydrochloric acid (for β-Ala), was dissolved potassium cyanate at time 0. The pH was regulated to 8 with 4 M HCl using the autotitrator. At regular time intervals, 0.1 mL aliquots were taken, diluted if necessary in pure water, and analysed by HPLC, eluent water + 0.05% TFA; t_R/min = 4.5 (**C-Thr**), 5.0 (**C-β-Ala**), 13.5 (α,ω-dicarbamoyllysine) at 1.0 mL min^{-1}.

Preparation of CAAs

***N*-Carbamoyl-L-valine C-Val.** To a warm (50 °C) stirred solution of 175.5 g (1.5 mol) L-valine in 1 L water was added 184.5 g (2.25 mol) potassium cyanate. The pH was then regulated to 7.5 using the autotitrator to control continuous 6 M HCl addition. Reaction completion was checked by HPLC, eluent water–acetonitrile (95 : 5) + 0.05% TFA; t_R/min = 4.15 at flow rate 1 mL min^{-1}. After completion (about 15 h), the reaction mixture was cooled by means of an ice bath and acidified by adding 6 M HCl until reaching pH 2. After filtration, the crude product was recrystallised from boiling water and dried in a heating desiccator, yield 84%. Mp 207–209 °C. δ_H 0.78 (3H, d, J = 7 Hz, CH_3), 0.82 (3H, d, J = 7 Hz, C′H_3), 1.93 (1H, m, H_β), 3.95 (1H, dd, H_α, $J_{\alpha\beta}$ = 5 Hz, $J_{\alpha NH}$ = 9 Hz), 5.55 (2H, s, N′H_2), 6.11 (NH, d, J = 9 Hz), 12.45 (1H, s, COOH). δ_C 19.89 (C_γ), 21.05 ($C_{\gamma'}$), 29.90 (C_β), 63.83 (C_α), 162.94 (NCON′), 179.69 (COOH).

***N*-Carbamoyl-L-alanine C-Ala.** Same procedure as **C-Val**, using 135 g (1.5 mol) L-alanine in 1 L water and 184.5 g (2.25 mol) potassium cyanate, reaction at 50 °C, pH 7.5 for 15 h. Eluent for HPLC analysis: water–acetonitrile (95 : 5) + 0.05 % TFA; t_R/min = 8.4 at 0.8 mL min^{-1}. The crude product was recrystallised from boiling water to yield 240 g (82%) pure material. Mp 165–167 °C. δ_H 1.22 (3H, d, H_β, $J_{\alpha\beta}$ = 7.24 Hz), 4.04 (1H, dq, H_α, $J_{H\alpha\text{-}NH} = J_{\alpha\beta}$ = 7.35 Hz), 5.60 (2H, s, N′H_2), 6.23 (1H, d, NH, $J_{NH\text{-}H\alpha}$ = 7.7 Hz), 12.47 (1H, s, COOH). δ_C 19.05 (C_β), 48.79 (C_α), 159.06 (NCON′), 176.28 (COOH).

3-Ureidopropionic acid C-β-Ala. Same procedure as **C-Val**, using 13.35 g (150 mmol) β-alanine in 0.5 L water and 13.36 g (165 mmol) potassium cyanate. pH regulated to 7.5, reaction time 18 h. Eluent for HPLC analysis: water–acetonitrile (80 : 20) + 0.05% TFA; t_R/min = 3.6 at 1.0 mL min^{-1}. The crude product was recrystallised from boiling water to yield 11.9 g (60%) pure material. δ_H 2.37 (2H, t, H_α, J = 6.0 Hz), 3.18 (2H, dt, H_β, $J_d = J_t$ = 5.9 Hz), 5.40 (2H, s, N′H_2), 6.00 (1H, t, NH, J_t = 5.4 Hz), 12.20 (1H, s, COOH). δ_C 35.0 (C_α), 35.7 (C_β), 159.0 (NCON′), 174.0 (COOH).

***N*-Carbamoyl-L-leucine C-Leu.** Same procedure as **C-Val**, using 30 g (0.23 mol) L-leucine in 160 mL water and 30 g (0.37 mol) potassium cyanate, at 50 °C, pH 7.5 for 18 h. The crude product was recrystallised from boiling water to afford 30.45 g (76%) pure material. Mp 216–218 °C. δ_H 0.89 (6H, dd, H_δ, $J_1 = J_2$ = 6.0), 1.45 (2H, m, H_β), 1.65 (1H, m, H_γ), 4.08 (1Hd, td, Hα, $J_{H\alpha\text{-}NH}$ = 8.4 Hz, $J_{\alpha\beta}$ = 6.3 Hz), 5.55 (2H, s, N′H_2), 6.18 (1H, d, NH, J = 8.4 Hz), 12.44 (1H, s, COOH). δ_C 22.42 (C′$_\delta$), 23.71 (C_δ), 25.16 (C_γ), 41.89 (C_β), 51.56 (C_α), 159.19 (NCON′), 176.12 (COOH).

***N*-Carbamoyl-L-phenylalanine C-Phe.** Same procedure as for **C-Val**, using 11 g (67 mmol) phenylalanine and 20 g (246 mmol) potassium cyanate in 500 mL water; at 50 °C, pH 7 for 15 h. 10.8 g (78%) pure material was obtained after recrystallisation from water. Mp 200 °C. δ_H 2.8 (2H, dd, H_β), 4.35 (1H, m, H_α), 5.10 (2H, s, N′H_2), 6.25 (1H, d, NH), 7.25 (5H, m, H_{Ar}), 11.80 (1H, s, COOH).

***N*$^{\alpha}$-Carbamoyl-*N*$^{\varepsilon}$-trifluoroacetyl-L-lysine C-Lys(TFA).** Same procedure as for **C-Val**, using 20 g (82.6 mmol) *N*$^{\varepsilon}$-trifluoroacetyl-L-lysine and 10 g (123.5 mmol) potassium cyanate in 150 mL water at 50 °C; the pH was regulated to 7.5 during the reaction. After stirring at 50 °C for 1 h 30 min, another 3.35 g (41.4 mmol) potassium cyanate was added to the mixture, which was again stirred at regulated pH for 1 h 30 min. The

chilled mixture was then acidified to pH 2 using concentrated HCl, and the precipitated product was recovered by filtration and dried *in vacuo* to yield 17.37 g (60.9 mmol, 73.8%) **C-Lys(TFA)** (11.78 g (50%) after recrystallisation from water). Mp 169–170 °C. δ_H 1.30 (2H, m, H_γ), 1.58 (4H, m, $H_\beta + H_\delta$), 3.17 (2H, dt, H_ε, $J_d = J_t = 6.5$ Hz), 4.04 (1H, m, H_α), 5.59 (2H, s, $N'H_2$), 6.22 (1H, d, $N^\alpha H$, $J_d = 8.2$ Hz), 9.43 (1H, t, $N^\varepsilon H$, $J_t = 6.5$ Hz), 12.51 (1H, s, COOH). δ_C : 23.3 (s, C_γ), 28.7 (s, C_β), 32.5 (s, C_δ), 40.0 (s, C_α), 52.9 (s, C_α), 116.8 (q, CF_3, $J_q = 288.3$ Hz), 157.0 (q, $COCF_3$, $J_q = 35.8$ Hz), 159.2 (s, CON_2), 175.5 (s, COOH). *m*/*z* (FAB^+/GT) : 571 ($2M + H^+$), 324 ($M + K^+$), 286 ($M + H^+$). ν_{max}(cm^{-1}) (film) 1160, 1185, 1213 (C–F), 1555 (N–H), 1642 ($C{=}O_{urea}$), 1699 ($C{=}O_{acid}$), 3273, 3329, 3444 (N–H). $[\alpha]_D = +10°$ (EtOH at 0.5 g/100 mL, 20 °C, wavelength = 589 nm).

***N*-Carbamoylglycine C-Gly.** To a warm (50 °C), stirred solution of 3.75 g (50 mmol) glycine in 50 mL water was added 5.27 g (65 mmol) potassium cyanate. The reaction was followed by HPLC analysis (eluent: aqueous KH_2PO_4 (10 mM) + sodium hexanesulfonate (5 mM), adjusted to pH 2.5 with H_3PO_4, UV detection at 195 nm). After 10 h stirring at 50 °C (the pH meanwhile rose from 7.12 at the beginning to 10.0 at the end of the reaction), the mixture was chilled, and filtered through a column containing AG-50-X2 cation-exchange resin (H^+ form; further elution with water). After water evaporation *in vacuo*, then drying in a heating desiccator, the major fraction afforded 4.88 g (41.3 mmol, 83%) pure **C-Gly** (hydantoic acid: mp 165 °C. δ_H 3.66 (2H, d, $J = 5.6$ Hz, H_α), 5.64 (2H, s, NH_2), 6.14 (1H, t, $J = 5.6$ Hz, NH), identical physical data to an authentic sample from Sigma). A further fraction (824 mg) contained *ca.* 90 mol% **C-Gly** (6.6 mmol) and 10 mol% urea (NMR). Overall yield **C-Gly** 95%.

***N*-Carbamoyl-DL-methionine C-Met.** 447 g (3.0 mol) DL-methionine and 323 g (4.0 mol) potassium cyanate were suspended in 3.5 L water, then stirred and heated to reflux for 30 min (dissolution was completed during heating). The solution was then cooled to 5–10 °C and 400 mL 12 M HCl was slowly added, while keeping the temperature at 10 °C. The mixture was then stirred for another 1 h, then filtered. The solid was washed 4 times with 100 mL water, then dried *in vacuo*, to yield 447 g (2.44 mol, 81%) pure material, which presented identical physical characteristics to an authentic sample supplied by Aventis Animal Nutrition.

Acknowledgements

We are grateful to the CNRS, the French Ministries of Education and Research, and Degussa-Hüls AG (Germany) for their financial support, and to Adam Lindsey-Clark for his help in revising the English expression of this article.

References

1 H. Collet, C. Bied, J. Taillades, L. Mion and A. Commeyras, *Tetrahedron Lett.*, 1996, **37**, 9043.
2 H. Collet, L. Boiteau, J. Taillades and A. Commeyras, *Tetrahedron Lett.*, 1999, **40**, 3355.
3 V. F. Gonko, N. D. Shustova, G. M. Anoshina, T. E. Zubova and L. B. Radina, *Pharm. Chem. J.*, 1978, **12**, 601.
4 A. Rousset, M. Lasperas, J. Taillades and A. Commeyras, *Tetrahedron*, 1980, **36**, 2649.
5 A. Yamashiro, K. Kobuta and K. Yokozeki, *Agric. Biol. Chem.*, 1988, **52**, 2857; O. Keil, M. P. Schneider and J. P. Rasor, *Tetrahedron: Asymmetry*, 1995, **6**, 1257.
6 D. G. Smyth, *J. Biol. Chem.*, 1967, **242**, 1579.
7 A. Williams and W. P. Jencks, *J. Chem. Soc., Perkin Trans. 2*, 1974, 1753.
8 G. D. Vogels, L. Uffink and C. Van der Drift, *Recl. Trav. Chim. Pays-Bas*, 1970, **89**, 500.
9 M. W. Lister, *Can. J. Chem.*, 1955, **33**, 426.
10 K. Taufel, C. Wagner and H. Dunwald, *Z. Elektrochem.*, 1928, **34**, 115.
11 R. Naumann, *Z. Elektrochem.*, 1910, **16**, 773.
12 M. B. Jensen, *Acta Chem. Scand.*, 1958, **12**, 1657.
13 R. Caramazza, *Gazz. Chim. Ital.*, 1958, **88**, 308.
14 I. A. Kemp and G. Kohnstam, *J. Chem. Soc.*, 1956, 900.
15 G. Kortüm, W. Vogel and K. Andrussow, *Dissociation Constants of Organic acids in Aqueous Solution*, Butterworths, London, 1961.
16 P. K. Smith, A. T. Gorham and E. R. B. Smith, *J. Biol. Chem*, 1942, **144**, 737.
17 J. Taillades, I. Beuzelin, L. Garrel, V. Tabacik, C. Bied and A. Commeyras, *Orig. Life Evol. Biosphere*, 1998, **28**, 61; J. Taillades, H. Collet, L. Garrel, I. Beuzelin, L. Boiteau, H. Choukroun and A. Commeyras, *J. Mol. Evol.*, 1999, **48**, 638.
18 C. H. Clapp, J. S. Swan and J. L. Poetchmann, *J. Chem. Educ.*, 1992, **69**, 122.
19 E. E. Schallenberg and M. Calvin, *J. Am. Chem. Soc.*, 1955, **77**, 2779.