Multicomponent Reactions

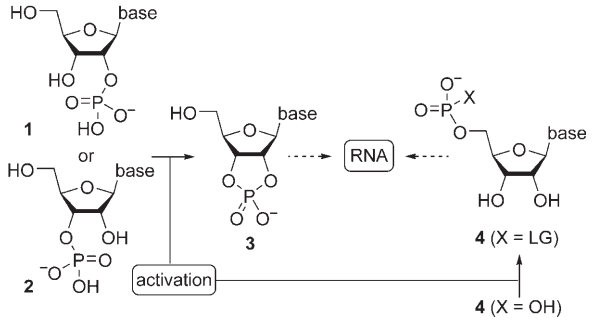
DOI: 10.1002/anie.200702870

# Simultaneous Nucleotide Activation and Synthesis of Amino Acid Amides by a Potentially Prebiotic Multi-Component Reaction**

*Lee B. Mullen and John D. Sutherland**

Activated nucleotides are needed for prebiotic RNA synthesis, and amino acid derivatives are needed for protein synthesis. Efficient nucleotide activation in water with prebiotically plausible reagents has not proved possible until now due to low yields and side-reactions with the nucleobases.[1] Amino acids and their derivatives are formed from hydrogen cyanide, ammonia, and aldehydes in Miller–Urey experiments, but in low yields and along with numerous other products.[2] We wondered if more selective and efficient routes to RNA and protein monomers might be discovered through investigations with additional prebiotic feedstock molecules. Isocyanides can be detected in the interstellar medium[3] and are formed from nitriles under conditions that simulate the chemistry of Titan and comets.[4] We have therefore started to explore early Earth model chemical systems that comprise isocyanides in mixtures with other components in a search for efficient reactions that activate nucleotides or lead to amino acid derivatives.

Upon phosphate activation, 2′-nucleotides **1** and 3′-nucleotides **2** give nucleoside-2′,3′-cyclic phosphates **3** (Scheme 1).[5] These cyclic nucleotides retain a degree of activation due to ring strain and have long been considered as monomers for the synthesis of RNA by polymerization.[6]

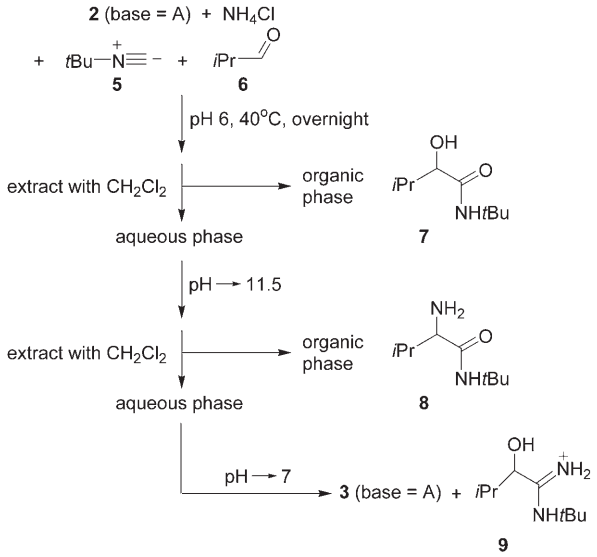



***Scheme 1.*** Activation of nucleoside phosphate monoesters in the prebiotic synthesis of RNA. LG = leaving group.

Activation of 5′-nucleotides **4** (X = OH) does not result in cyclization, and 5′-activated nucleotides **4** (X = leaving group LG) are alternative monomers for RNA synthesis.[7,8]

We first studied 2′/3′-nucleotides **1/2** because cyclization to **3** can be easily detected by $^1$H NMR spectroscopic analysis. Our decision to attempt the activation of a 2′/3′-nucleotide by treatment with an isocyanide, an aldehyde, and $NH_4Cl$ was based upon an analysis of the Ugi reaction.[9] In the first stages of an Ugi reaction, these three components react to give an intermediate which can activate a carboxylate ion, and we wondered if the phosphate group of a 2′/3′-nucleotide **1/2** could take the place of this anion. Furthermore, we postulated that the activated carboxylate and the activated phosphate would undergo different reactions. Before we could test these hypotheses experimentally, we had to decide which isocyanide and aldehyde to use. For ease of handling, and for olfactory reasons, we selected *tert*-butylisocyanide (**5**), and we chose isobutyraldehyde (**6**) as it would give derivatives of the proteinogenic amino acid valine if the transformation we envisaged took place. Addition of four equivalents each of **5** and **6** to a solution of β-D-adenosine-3′-phosphate (**2**, base = A; 100 mM) and $NH_4Cl$ (1M) at pH 6 resulted in a heterogeneous reaction mixture that was stirred at 40 °C overnight.

To analyze the products formed, we then fractionated the mixture (Scheme 2). This showed that the reaction had produced the hydroxy amide **7**, amino amide **8**, and hydroxy amidine **9** in addition to the 2′,3′-cyclic phosphate **3** (base = A) and two minor products. The structures of **7** and **8** were determined by spectroscopic analysis of purified samples, the




***Scheme 2.*** Fractionation of the products of a four-component reaction.

[*] L. B. Mullen, Prof. Dr. J. D. Sutherland
School of Chemistry
The University of Manchester
Oxford Road, Manchester M13 9PL (UK)
E-mail: john.sutherland@manchester.ac.uk

[**] This work was carried out as part of EU COST action D27 "Prebiotic Chemistry and Early Evolution", and was funded by the Engineering and Physical Sciences Research Council through a studentship to L.B.M.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

structure of **3** (base = A) by comparison to an authentic standard, and that of **9** by NMR spectroscopic and mass spectrometric analysis of the mixture containing **9** and **3** (base = A). The yield of **3** (base = A) based on **2** (base = A) was >95% by $^1$H NMR analysis, but, since losses were incurred in the fractionation and purification of **7** and **8**, another method was sought that could reliably give the yield of these species in the reaction. We tried $D_2O$ as a solvent in order to determine the yields by $^1$H NMR analysis directly after the reaction, given that we now had purified standards of the products, but the heterogeneous reaction mixture prevented this. Accordingly, we lyophilized the mixture after reaction, and dissolved the products in a deuterated NMR solvent ($CD_3OD$ or $(CD_3)_2SO$) in which they were all soluble. In this way we were able to determine the yields for all identified products (Table 1). The quantitative cyclization of **2** (base = A) to **3** (base = A), without nucleobase modification, suggests that the phosphate activation is highly selective since the amino group of adenine derivatives can undergo modification by other electrophiles.[10]

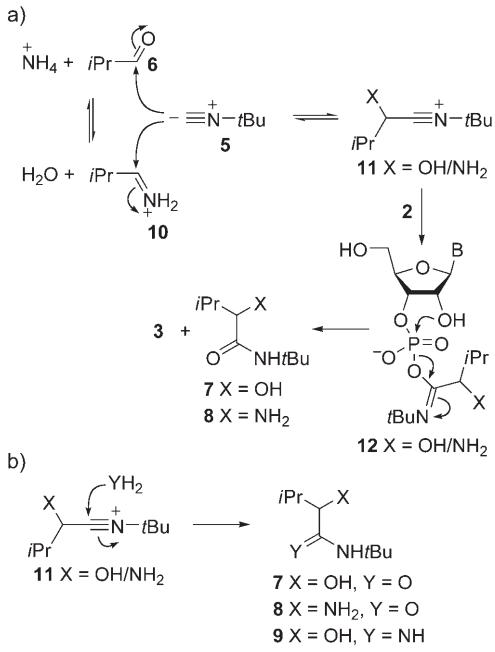
***Table 1:*** The effect of nucleotide structure on the yields of **3** and other products.

| Nucleotide (100 mM) | ±$NH_4Cl$ (1 M) | Yield [%][a] **3** | **7** | **8** | **9** |
|---|---|---|---|---|---|
| **2** (base = A) | + | 100 | 100 | 50 | 90 |
| **2** (base = C) | + | 90 | 162 | 41 | 81 |
| **1** (base = U) | + | 84 | 117 | 36 | 79 |
| **1** + **2** (base = G)[b] | + | 65 | 135 | n.d.[c] | n.d.[c] |
| **1** + **2** (base = G)[d] | +[e] | 60 | 200 | n.d.[c] | 82 |
| **2** (base = A) | − | 100 | 160 | – | – |
| **2** (base = C) | − | 100 | 201 | – | – |
| **1** (base = U) | − | 100 | 167 | – | – |
| **1** + **2** (base = G) | − | 70 | 188 | – | – |

[a] Based on starting nucleotide; for **7**–**9** this leads to yields higher than 100% in some cases (because four equivalents of both **5** and **6** were used), but allows direct comparison of the relative amounts of the nucleotide product and the other products. [b] Ratio of **1** (base = G) and **2** (base = G) ca. 1:2. [c] Yield could not be determined due to signal overlap. [d] 10 mM. [e] 100 mM $NH_4Cl$.

Encouraged by this result, we next studied the reaction of **2** (base = C) since the amino group of cytosine derivatives is also prone to electrophilic modification. Again we observed cyclization in high yield with no nucleobase modification, and again we saw the same range of non-nucleotide products including the amino acid derivative **8**. To ascertain whether the reaction would also work with a 2′-nucleotide, we then subjected **1** (base = U) to the same reaction conditions. Once more the cyclic nucleotide was formed in excellent yield along with the other products. To demonstrate that mixtures of **1** and **2**, that would form by slow hydrolysis of **3**,[6,7] could be cyclized back to **3**, we treated a mixture of **1** (base = G) and **2** (base = G) with the phosphate activation reagents. In this case, conversion to **3** was less efficient, though the cyclic nucleotide was still formed in 65% yield, and the yields of **8** and **9** could not be determined because of signal overlap in the $^1$H NMR spectrum. Furthermore the reaction mixture had been viscous (presumably due to aggregation of the guanine nucleotides), so we repeated the reaction 10-fold diluted. In this way, we were not able to quantify the amount of **8** formed, but key $^1$H NMR signals for **9** could be clearly discerned and integrated. Compared to the reaction with the normal concentration of substrate and reagents, the more dilute reaction gave increased amounts of **7**, but the cyclization yield was only marginally decreased showing that the transformation is still high-yielding at lower concentrations. The constitution of **7** implies that its formation does not require the presence of $NH_4Cl$, so we carried out additional reactions in the absence of this salt (Table 1). These experiments gave no **8** or **9**, but showed increased production of **7**, and extremely efficient cyclization of **1**/**2** to **3**. Isocyanides alone have been shown to activate phosphoric acid derivatives in pyridine,[11] so, in other control experiments, we treated **1**/**2** with **5** in water. At pH 6 we only observed very slow cyclization to **3**, the rate being insufficient to account for the formation of **3** in reactions of **1**/**2** with **5** and **6** in the presence or absence of $NH_4Cl$.

We have not investigated the mechanism of the reactions in detail, but the constitution of the various products provides clues as to how they might be formed (Scheme 3). In the Ugi reaction, and the closely related Passerini reaction, nitrilium ions formed from isocyanides are the activating agents,[9] and it seems likely that this is also the case in the process we have uncovered. Reversible reaction of **6** and an ammonium ion would give the iminium ion **10**, and reaction of **10** with **5** would give the amino nitrilium ion **11** (X = $NH_2$). Alternatively (or exclusively in the absence of $NH_4Cl$), direct reaction of **5** with **6** would give the hydroxy nitrilium ion **11** (X = OH). Subsequent reaction of either nitrilium ion with **2** would lead to the imidoyl phosphates **12**, and the simultaneous formation of **3**, and **7** or **8**, can be explained by intramolecular attack of the 2′-OH group of **12** at the phosphorus atom. The hydroxy

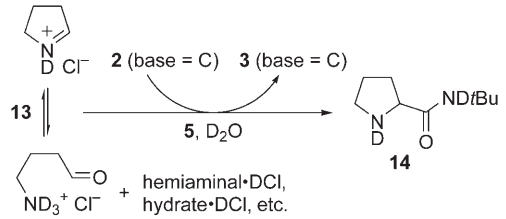



***Scheme 3.*** Proposed mechanisms for a) the three- and four-component reactions, and b) the competing reactions of the nitrilium ions **11**.

amidine **9** is presumed to result from competing attack of ammonia on **11** (X = OH), and the fact that the combined yields of **7** and **8** are 1.5–2.7 times higher than the yield of **3** suggests that the nitrilium ions **11** also undergo competing attack by water. Consistent with this, the products **7**–**9** were again formed in similar yield in the absence of nucleotides, but the reaction proceeded more slowly.

In the four-component reaction, the formation of **7** competes with the formation of **8** presumably because the equilibrium between **10** and **6** is not completely displaced in favor of **10**. Increasing the concentration of ammonia would be expected to displace the equilibrium in favor of **10**, but also to increase the rate of formation of the by-product **9**. However, if an amine group is tethered to the aldehyde such that intramolecular iminium ion formation can occur, amino amide formation ought to be selectively favored. To test this hypothesis, we investigated the reaction of a four-fold excess of 4-aminobutyraldehyde deuterochloride/1-pyrroline deuterochloride (**13**) with **2** (base = C) and **5** in $D_2O$ solution (Scheme 4). $^1$H NMR analysis showed that the reaction was




***Scheme 4.*** Selective formation of an amino amide from a cyclic iminium ion.

complete within 30 min and the products remained unchanged after 16 h. The yield of **3** (base = C) was quantitative, and only one major co-product was formed in 149% yield based on **2** (base = C). This co-product was shown to be deuterated proline *tert*-butylamide **14** by spiking the NMR sample with a synthetic standard of L-**14**. Although we have not determined the solution-structure profile of **13**, the rapid reaction with **13** coupled with the observation of **14** as the only major co-product nevertheless suggests that iminium ions are more reactive than aldehydes towards isocyanides.

In the activation of 5′-nucleotides **4** (X = OH) by **5**, **6**, and $NH_4Cl$, cyclization to nucleoside-3′,5′-cyclic phosphates is not to be expected, so successful phosphate activation would not be evidenced by stable nucleotide products. However, formation of the non-nucleotide products **7**–**9** more rapidly than in the minus-nucleotide control would indicate that transient activation to imidoyl phosphates, followed by hydrolysis, had taken place. We investigated this by adding four equivalents of **5** and **6** to solutions of **4** (base = C, X = OH) at pH 6 in the presence or absence of $NH_4Cl$. In the presence of 1M $NH_4Cl$, all three products were formed (yields based on **4**: **7**, 233%; **8**, 30%; **9**, 31%), but in its absence only **7** (370%) was observed. These reactions were significantly faster than the minus-nucleotide control, and strongly suggest that transient activation to **4** (X = OC(*t*BuN)CH(OH/$NH_2$)*i*Pr) had taken place. The subsequent hydrolysis of these imidoyl phosphates implies that such species could only be intermediates in a templated oligomerization of RNA in aqueous solution, in which case the 3′-OH group of a growing chain could have a high effective molarity relative to water and therefore would be able to compete as a nucleophile.

We next returned to the reactions of 3′-nucleotides **2** and focused on stereochemical issues, in particular the possibility of a link between the chirality of **2** and that of the amino acid amide **8**. To determine whether **8** was produced enantioselectively, we analyzed a sample we had originally purified from the reaction of **2** (base = A) with **5** and **6** in the presence of $NH_4Cl$ by gas chromatography (GC) with a chiral column (see the Supporting Information). Through comparison of the GC data for this sample with those of commercially available L-**8**, and a synthetic sample of D-**8**, we found the **8** produced in the four-component reaction to have an *ee* value of 0.8% in favor of the L-isomer. This is a low value, and, within experimental error, it indicates that the sample was racemic, or close to racemic, but we note that prebiotically plausible mechanisms exist for the amplification of small enantiomeric excesses in amino acids.[12]

It is particularly noteworthy that the activating agents in this transformation—presumed to be the nitrilium ions **11** and **12**—do not modify any of the four nucleobases, whereas other prebiotically plausible electrophiles react with certain nucleobases. Thus, for example, cyanoacetylene is a reasonable phosphate activating agent and reacts with both cytosine and adenine derivatives;[13,14] reversibly in the former case, but irreversibly in the latter. Furthermore, the fact that the amino acid derivatives **8** are co-products of the four-component reaction makes this process particularly appealing in a prebiotic context.

## Experimental Section

Reagents were of the highest quality commercially available and were used without purification. Operations involving the isocyanide **5** were carried out in well-ventilated fumehoods to avoid the stench of this compound. Small-scale reactions were carried out in plastic tubes with shaking to ensure vigorous agitation of the heterogeneous mixtures, large-scale reactions were carried out in glassware with rapid stirring. Less efficient mixing resulted in lower rates and lower yields. Hydroxy amide **7** purified from a large-scale reaction, commercial 2′,3′-cyclic nucleotides **3**, and amino amide **8**, were used to spike NMR samples to confirm the presence or absence of these compounds among the reaction products. The presence or absence of the hydroxy amidine **9** was determined by comparison of $^1$H NMR spectra of crude reaction products to those of fractionated samples containing **3** and **9**. For full experimental details and characterization of the products see the Supporting Information.

Received: June 28, 2007
Published online: September 13, 2007

**Keywords:** amino acids · multicomponent reactions · nucleotides · phosphate activation · prebiotic chemistry

[1] J. P. Ferris, W. J. Hagan, Jr., *Tetrahedron* **1984**, *40*, 1093.
[2] S. L. Miller in *The Molecular Origins of Life: Assembling Pieces of the Puzzle* (Ed.: A. Brack), Cambridge University Press, Cambridge, UK, **1998**, pp. 59–85.

[3] A. J. Remijan, J. M. Hollis, F. J. Lovas, D. F. Plusquellic, P. R. Jewell, *Astrophys. J.* **2005**, *632*, 333.
[4] R. L. Hudson, M. H. Moore, *Icarus* **2004**, *172*, 466.
[5] D. M. Brown, D. I. Magrath, A. R. Todd, *J. Chem. Soc.* **1952**, 2708.
[6] M. Renz, R. Lohrmann, L. E. Orgel, *Biochim. Biophys. Acta Nucleic Acids Protein Synth.* **1971**, *240*, 463.
[7] L. E. Orgel, R. Lohrmann, *Acc. Chem. Res.* **1974**, *7*, 368.
[8] J. P. Ferris, A. R. Hill, Jr., R. Liu, L. E. Orgel, *Nature* **1996**, *381*, 59.
[9] M. C. Pirrung, K. Das Sarma, *Tetrahedron* **2005**, *61*, 11456.
[10] G. M. Blackburn in *Nucleic Acids in Chemistry and Biology* (Eds.: G. M. Blackburn, M. J. Gait), Oxford University Press, Oxford, UK, **1996**, pp. 283–327.
[11] Y. Mizung, J. Kobayashi, *J. Chem. Soc. Chem. Commun.* **1974**, 997.
[12] M. Klussmann, H. Iwamura, S. P. Mathew, D. H. Wells, Jr., U. Pandya, A. Armstrong, D. G. Blackmond, *Nature* **2006**, *441*, 621.
[13] M. A. Crowe, J. D. Sutherland, *ChemBioChem* **2006**, *7*, 951.
[14] Y. Furukawa, O. Miyashita, M. Honjo, *Chem. Pharm. Bull.* **1974**, *22*, 2552.