How to cite: *Angew. Chem. Int. Ed.* **2024**, *63*, e202319958
doi.org/10.1002/anie.202319958

Activated Nucleotides

Prolinyl Phosphoramidates of Nucleotides with Increased Reactivity

*Adrian Humboldt, Fabian Rami, Franka M. Topp, Dejana Arnold, Daniela Göhringer, Pradeep S. Pallan, Martin Egli, and Clemens Richert**

In memoriam Dr. Clemens Wissing.

Abstract: Nucleoside monophosphates (NMPs) are the subunits of RNA. They are incorporated into growing complementary strands when sequences are copied in enzyme-free reactions using organic leaving groups at the phosphates. Amino acids are rarely considered as leaving groups, but proline can act as a leaving group when *N*-linked to NMPs, so that prolinyl NMPs hydrolyze in aqueous buffer at 37 °C, with half-life times as short as 2.4 h, and they act as monomers in enzyme-free primer extension. Still, their level of reactivity is insufficient for practical purposes, requiring months for some extensions. Herein we report the synthesis of eight substituted prolinyl AMPs together with seven related compounds and the results of a study of their reactivity. A δ-carboxy prolinyl NMP was found to be converted with a half-life time of just 11 min in magnesium-free buffer, and a δ-isopropyl prolinyl NMP was shown to react sevenfold faster than its prolinyl counterpart in enzyme-free genetic copying of RNA. Our results indicate that both anchimeric and steric effects can be employed to increase the reactivity of aminoacidyl nucleotides, i.e. compounds that combine two fundamental classes of biomolecules in one functional entity.

Introduction

Functional molecules that are reactive, but kinetically stable, play important roles in chemistry and biology. Examples for such molecules can be found in nucleic acid chemistry. Best known to biologists are nucleoside triphosphates, such as ATP, which is stable in aqueous solution but readily reacts in the active site of enzymes.[1,2] This feature renders it suitable for its role as energy currency of the cell and as a building block for biosyntheses. Although the mechanism of the reactions of P(III) compounds is different from that of their P(V) counterparts, phosphoramidites of nucleosides are another prominent class of compounds that are unreactive under neutral and basic conditions but are readily converted into highly reactive monomers when an acidic organocatalyst is added during solid-phase synthesis of DNA and RNA.[3,4,5] This makes phosphoramidites the preferred starting materials for oligonucleotide synthesis, with a shelf-life of months to years, coupling times of approximately one minute and yields >99% per chain extension step.[6] Finding other kinetically stable, but readily activated forms of functional molecules is an interesting task.

One area, in which there is a strong motivation to develop new kinetically stable compounds that can undergo efficient transformations is medicinal chemistry. Many inhibitors lack bioavailability in their active form and have to be masked to help them reach their target cells. This applies to some prodrugs of antivirals. Nucleotide prodrugs in which the phosphate is masked as aryl amino acidyl ester phosphoramidate, called ProTides, are successful antivirals in the clinic.[7,8] The ProTide design has been used for anti-HIV drugs[9,10] and anticancer agents.[11,12] The masked form is able to pass membranes, thereby achieving bioavailability, and is then converted to the active triphosphate form in a series of metabolic steps. This is shown for sofosbuvir, a breakthrough drug against hepatitis C, in Figure 1A.[13] The last step of the unmasking of the nucleotide, the conversion of PSI-353707 to PSI-7411, requires enzymatic catalysis and has been described as the rate-limiting step in the overall metabolism to the active form, PSI-7409.[14,15] The hydrolytic cleavage of the amino acidyl phosphoramidate bond is believed to be catalyzed by histidine triad nucleotide binding proteins (Hints).[16] The millimolar substrate affinity of Hint enzymes for alaninyl nucleotides,[17] and the modest expression levels in some cells may help to explain why ProTide inhibitors like remdesivir, whose target organs are something other than the liver, are less successful in the clinic[18] than sofosbuvir.[19]

A recent finding fosters hope that the bottleneck of enzymatic hydrolysis of amino acidyl phosphoramidates may be overcome. Replacing the alanine residue of all ProTides clinically employed[20] with proline accelerates the liberation

[*] M. Sc. A. Humboldt, Dr. F. Rami, M. Sc. F. M. Topp, M. Sc. D. Arnold, D. Göhringer, Prof. C. Richert
Institute of Organic Chemistry, University of Stuttgart
70569 Stuttgart, Germany
E-mail: lehrstuhl-2@oc.uni-stuttgart.de
Dr. P. S. Pallan, Prof. M. Egli
Department of Biochemistry, Vanderbilt University
School of Medicine, Nashville, Tennessee 37232, USA

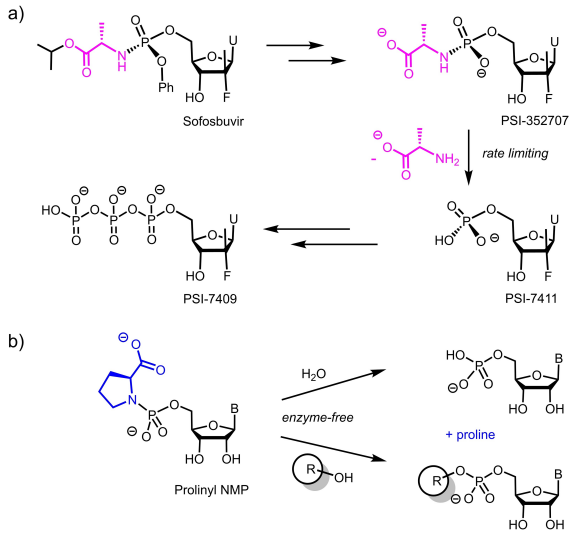


Figure 1. Interconversion of nucleoside phosphates. A) Steps of the intracellular metabolism of sofosbuvir, with the alanine residue shown in purple. B) Release of a nucleoside monophosphate (NMP) from a prolinyl nucleotide, with proline highlighted in blue.

of NMPs[21] up to 45-fold in enzyme-free buffer containing an imidazole. Further, for prolinyl nucleotides, histidine, rather than the histidine triad of Hint enzymes, suffices to induce the reaction to the nucleotide.[21] For other amino acidyl nucleotides, such as phenylalanine, hydrolysis is much slower, indicating that optimization of the amino acid substituent of the phosphoramidates may lead to an efficient, enzyme-free release pathway.

What added to our motivation for the current study was the fact that proline is the very amino acid known to act as organocatalyst for a broad range of coupling reactions forming carbon-carbon bonds.[22,23,24] So, it was interesting to ask what role this fascinating molecule plays in the release of nucleotides. Further, there was the broader question of why five-membered heterocycles and phosphoramidates play such an important role in nucleotide chemistry. For example, the most commonly used form of activated nucleotides in enzyme-free genetic copying have been imidazolides,[25,26,27] and some of the most successful enzyme-free replication assays lead to phosphoramidates.[28] Also it appears that phosphoramidates can play roles not unlike that of triphosphates in enzyme-free oligomerization scenarios,[29] again pointing to a special relationship between nucleic acids and such compounds. By shedding light on the structural basis for these relationships, insights may be gained that have relevance beyond the medicinal chemistry of nucleosides and nucleotides.

Results and Discussion

As stated above, the aim of our project was to identify compounds more reactive than the unmodified prolinyl nucleotides Pro-NMP (Figure 1). To achieve this goal, we considered the effect of structural alterations on the three processes shown in Figure 2, based on our working hypothesis that the opening step of the mechanism is the intramolecular attack of the carboxylate on the phosphorous (I).[21] Several effects were considered that might lower the activation barriers of the individuals steps. Sterically demanding substituents may have a preorganizing effect, favoring conformations, in which the carboxylate is close to the phosphorous.

Further, substituents placed on the pyrrolidine ring can be expected to shift the conformational equilibrium between equatorial and apical positions in the trigonal bipyramid formed upon nucleophilic attack at the phosphorous. This attack may result in a number of different intermediates that are involved in the reaction mechanism, thus lowering or increasing the activation barriers. Finally, through electronic and indirect effects, substituents will affect the basicity of the ring nitrogen, thereby changing the leaving group properties of the proline moiety.

Attempts to crystallize aminoacidyl phosphoramidates alone remained unsuccessful and we turned to a crystallization scaffold approach to trap amidates as complexes with with RNase A.[21] To gain structural insights, we determined crystal structures of RNase A complexes of the following aminoacidyl phosphoramidates: Ala-A, Arg-A, Glu-A, His-A and *N*-butyl-A (Figure 3). A sixth structure was a reference complex between RNase A and AMP. The structures were determined at resolutions between 1.53 and 2.0 Å and contain two RNase A molecules per crystallographic asymmetric unit. Crystal data, data collection and refinement parameters are summarized in Table S1 in the Supporting Information. Interactions of individual phosphoramidates bound to the enzyme are depicted in Figure 3. The two independent RNase A molecules in the asymmetric unit feature different packing interactions. Therefore, access to their active sites for a ligand diffusing into the crystal is

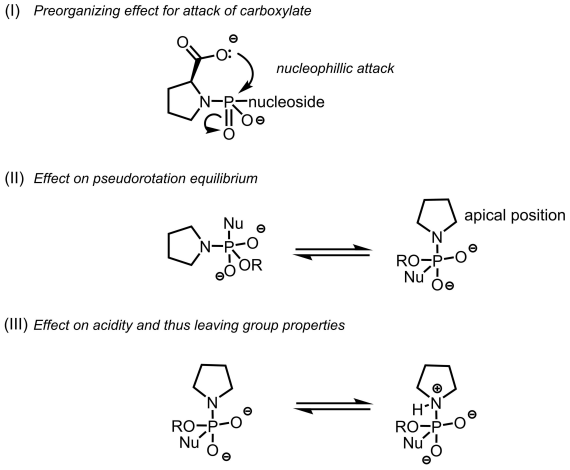


Figure 2. Possible effects of proline substituents on the rate of steps involved in hydrolysis. A) Intramolecular nucleophilic attack of the carboxylate, as a form of neighboring group participation. B) Berry pseudorotation placing the leaving group in an apical position. C) Basicity of the ring nitrogen affecting the leaving group properties of the prolinyl residue. Abbreviations: Nu, nucleophile; R, nucleoside residue.

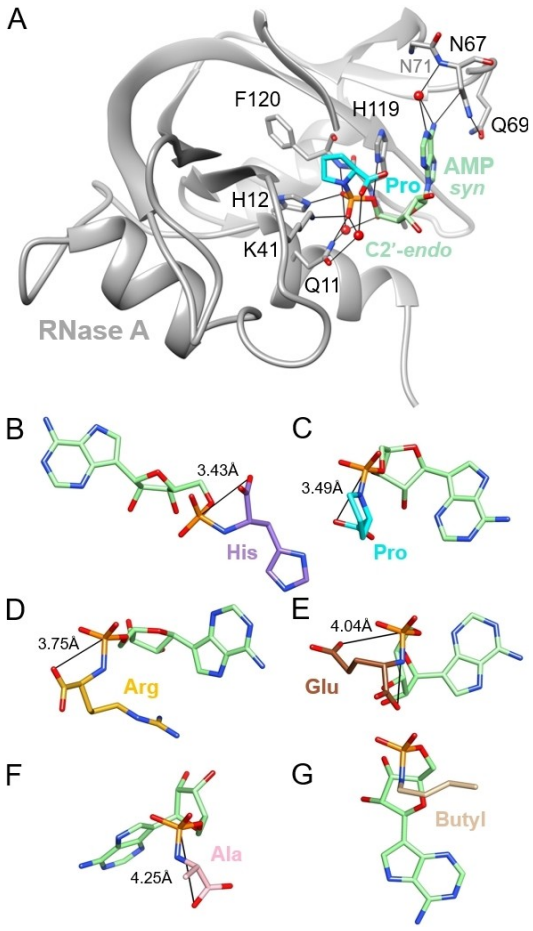


Figure 3. X-ray crystal structures of RNase A-phosphoramidate complexes. A) View of the enzyme with bound Pro-A. B–G) Bound nucleotides found in the active site, presented in the order of increasing distance between carboxylate oxygen and phosphate phosphorous atoms, indicated by gray lines. B) His-A, C) Pro-A, D) Arg-A, E) Glu-A, F) Ala-A, and G) *N*-butyl.

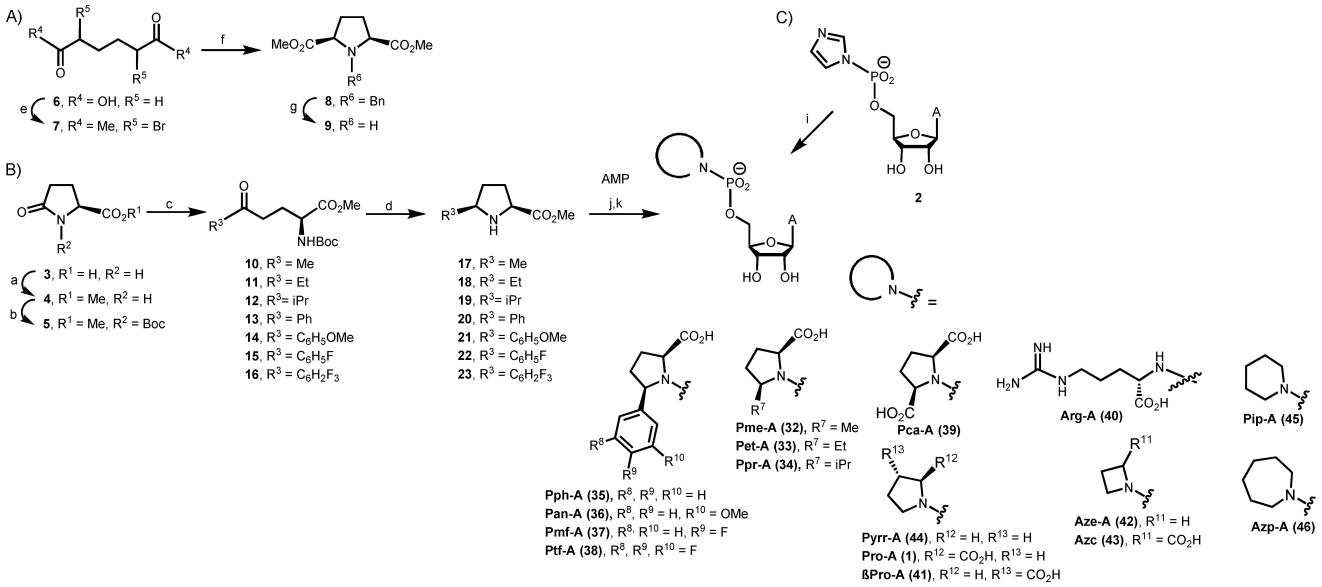
necessarily somewhat different. In four of the five complexes, only one active site harbors a phosphoramidate, whereas the other is either empty or shows electron density that is consistent with the partial occupancy of the ligand. Only the structure of the Glu-A complex shows two bound molecules, whereby the occupancy for the carboxylate moiety of the amino acid side chain is very limited for one of them. Because the complexes were not obtained by co-crystallization but rather by soaking phosphoramidates into enzyme crystals, some atoms/portions of ligands situated at the active site that is normally occupied do not display full occupancy. This is likely not just a result of using the soaking approach but may indicate that the amidate has undergone partial hydrolysis, with AMP diffusing more easily through the crystal compared to the complete aminoacidyl-AMP ligand. Hence, base, sugar and phosphate are well defined in the electron density map drawn at a particular threshold (e.g., 1σ level) compared to the amino acid moiety that is surrounded by weaker (2Fo-Fc) Fourier sum electron density in some cases. Alternatively, the amino acid and the nucleobase are well defined, but the density around some ribose atoms indicates an increased mobility of the sugar moiety.

The RNase A chaperone does not exert strong constraints on the conformations of phosphoramidates as is evident from the panels in Figure 3. The ribose puckers vary between different complexes and adenosine can adopt both the anti and syn conformation. However, the stacking interaction between the H119 side chain and adenine is seen in all complexes. Conversely, the conformations of the ligands can be relatively compact (Arg-A, Figure 3b) or extended (His-A, Figure 3d, and *N*-butyl-A, Figure 3e). The arginine side chain that was fit into the electron density in two alternative conformations folds back toward the base and the closest approach between the guanidino moiety and adenine C8 is just 3.37 Å. For His-A, the imidazole moiety shows no interactions with surrounding RNase A amino acids, unlike the carboxylate that forms a salt bridge with K41, a H-bond with Q11, and is also linked to the phosphate group by a water molecule Figure 3d. The amidate nitrogen accepts in an H-bond to K7, whereas in two other complexes (Ala-A, *N*-butyl-A) the amidate nitrogen acts as an acceptor in an H-bond to the amino group of the H119 side chain.

The versatility of the RNase A scaffold in terms of tolerating a range of orientations by the amino acidyl moiety of the phosphoramidates is also apparent from the environment of the phosphate group that is free of interactions (Ala-A, Figure 3a), or shows non-bridging oxygen atoms engaging in H-bonds (Arg-A, e.g., H12, salt bridge with K41). In addition, in two complexes, the bridging O5' oxygen is H-bonded to the imidazole amino group of H119: Arg-A and Glu-A. In the reference structure with AMP, the phosphate group of the nucleotide is surrounded by water, but does not engage in H-bonds to RNase A amino acid side chains. We conclude that the conformations of phosphoramidates bound to RNase A observed are to some degree intrinsically preferred and are certainly not just sculpted by constraints generated by the nucleotide binding site of the enzyme. This diversity of phosphoramidate conformations is also found for the four previously studied amidates Phe-A, Pro-A, βPro-A and Gly-A.[21]

Looking at the structures of the various phosphoramidates and the closest distance between a carboxy oxygen and the phosphorous atom drawn in Figure 3 and the rate constants for hydrolysis[21] it becomes clear that there is no simple correlation between proximity in the ground state and reactivity.

We synthesized the prolinyl nucleotides of Scheme 1 to explore the effect of substituents on reactivity experimentally. Two routes were employed to generate prolines with substituents at the δ-position (general structure **1**), where steric effects are most likely to manifest themselves. They were to be coupled directly or via imidazolide **2**. One route starts from *L*-pyroglutamic acid (**3**), which was esterified to **4**, following literature precedent,[30] and *N*-protection to **5**. On the second route, the amino acid scaffold was elaborated by brominating adipic acid (**6**) to **7**, followed by esterification to **7**, cyclization to **8**, and *N*-deprotection to give **9**. For the first route, the scaffold-forming steps were the ring



Scheme 1. Synthesis of prolinyl phosphoramidates. A) Synthesis of δ-carboxy proline. B) Synthesis δ-substituted alkyl and aryl prolines. C) Elaboration of phosphoramidates from AMP and amino acid building blocks. a) $SOCl_2$, MeOH, −0 °C to r.t., 96%; b) Boc_2O, DMAP, Et_3N, CH_2Cl_2, quant.; c) RMgBr, THF, −40 °C to r.t., 67% (**10**), 77% (**11**), 45% (**12**), 75% (**13**), 71% (**14**), 48% (**15**), 59% (**16**); d) 1. TFA, CH_2Cl_2 2. H_2, Pd/C (10 wt%), MeOH, yields over 2 steps: 67% (**17**), 82% (**18**), 87% (**19**), 73% (**20**), 81% (**21**), 65% (**22**), 87% (**23**); e) $SOCl_2$, MeOH, Br_2, 75 °C to r.t. and then to 75 °C, quant.; f) $BnNH_2$, K_2CO_3, PhMe, H_2O, 100 °C, 36%; g) H_2, Pd/C, MeOH, 96%; i) amino acid, DMAP, 18% (**40**), 18% (**41**), 5% (**42**), 5% (**43**), 19% (**44**), 31% (**45**), 7% (**46**); j) EDC, DMAP, *N*-methylmorpholine, DMF, 20% (**24**), 24% (**25**), 18% (**26**), 22% (**27**), 11% (**28**), 13% (**29**), 11% (**30**), 18% (**31**); k) NaOH, DMF, quant. (**32**), 95% (**33**), 77% (**34**), 68% (**35**), 62% (**36**), 54% (**37**), 64% (**38**), 64% (**39**). Boc_2O = di-*tert*-butyl decarbonate, DMAP = 4-dimethylaminopyridine, TFA = trifluoroacetic acid, $BnNH_2$ = benzylamine, EDC = 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide, DMF = *N*,*N*-dimethylformamide.

opening addition of Grignard reagents to give compounds **10**–**16**, followed by the removal of the Boc group and cyclization to proline esters **17**–**23** as late steps.

The aminoacidyl nucleotides were prepared by coupling the prolines with AMP or CMP, either directly, using EDC, or by transamidation from imidazolide **2** as mentioned above. Esters were converted to the free acids by basic hydrolysis with NaOH in DMF-containing solution. The yields are listed in the legend to Scheme 1, and protocols and analytical data are provided in the Supporting Information. No attempts were made to optimize the synthetic protocols beyond what was required to obtain sufficient material for assays.

With the δ-substituted prolinyl AMPs in hand, we then studied their rate of hydrolysis in aqueous buffer by ^{31}P NMR spectroscopy. The half-life times of the phosphoramidate starting material span a range from 10.9 min for Ppr-A in magnesium-free buffer to 23′000 min for Arg-A in Mg^{2+}-containing buffer (Table 1). The latter value is among the longest measured thus far,[21] whereas the former is more than an order of magnitude shorter than that measured for Pro-A, our lead compound. Among the values measured in the same Mg^{2+}-containing medium, the following trends manifested themselves: i) introducing an alkyl group at the δ-position increased reactivity (**32**–**34**), a phenyl group reduces it slightly (**35**); ii) saturated heterocycles lacking the carboxyl group of proline give less reactive phosphoramidates (**42**–**46**), and so do both electron-donating and electron-withdrawing substituents on a phenyl ring at the δ-position (**35**–**38**); iii) a second carboxy group at the β-position (**39**) significantly enhances reactivity. Furthermore, for the fast-reacting amino acidyl nucleotides, a signal for an organocatalytic imidazolium phosphate intermediate appears in the ^{31}P NMR spectra (species **II** in Figure 4), indicating that the rates of the displacement of the prolinyl leaving group by the organocatalyst and the subsequent hydrolysis of this intermediate are approaching similar levels.

The appearance of detectable levels of the organocatalytic intermediates called for a kinetic treatment with two coupled differential equations, so that the concentration of species **I**–**III** of Figure 4a over time can be modeled, using the approach described in chapter 7 of the Supporting Information. The rate constants and selected half-life times derived from this kinetic treatment are presented in the last three columns of Table 1. Entries 19–21 of the same table again show results for key compounds under modified experimental conditions.

Entry 18 is for Ppr-A (**34**) in the presence of 1-(2-hydroxyethyl)imidazole instead of 1-ethylimidazole. A detectable difference in the rate is observed, despite the small change in the chemical structure of the organocatalyst, confirming the involvement of the organocatalytic step in rate-limiting steps of the mechanism. Entries 19–21 are for **34**, **39** and **44** in magnesium-free buffer, again at 37 °C. For both prolinyl derivatives, there is an approximately fourfold acceleration over 80 mM $MgCl_2$, whereas for pyrrolidine amidate **44** there is none. This finding is consistent with the

Table 1: Kinetic data for the conversion of phosphoramidate (**I**) and the release of free nucleotide (**III**) in buffer at pH 7.5 and 37 °C, as determined by ^{31}P NMR spectroscopy.[a]

Entry	Phosphoramidate	$MgCl_2$	Conversion of **I** [h^{-1}]	$t_{1/2}$ Conversion of **I** [min]	Build-up of **II** k'[c] [mM^{-1} h^{-1}]	Appearance of **III**[c] k'' [h^{-1}]	$t_{1/2}$ Appearance of **III** [min]
1	Pro-A (**1**)[b]	+	0.28	144	1.8×10^{-3}	1.8	24.1
2	Pme-A (**32**)	+	0.57	72.8	–	0.57	72.8
3	Pet-A (**33**)	+	0.65	64.2	3.6×10^{-3}	3.90	10.7
4	Ppr-A (**34**)	+	0.96	43.2	7.8×10^{-3}	3.90	10.7
5	Pph-A, (**35**)	+	0.28	150	1.7×10^{-3}	3.90	10.7
6	Pan-A (**36**)	+	0.25	167	–	0.25	167
7	Pmf-A (**37**)	+	0.23	184	–	0.23	184
8	Ptf-A (**38**)	+	3.1×10^{-2}	1.33×10^{3}	–	3.1×10^{-2}	1.33×10^{3}
9	Pca-A (**39**)	+	0.83	50.2	–	3.80	50.2
10	Arg-A (**40**)	+	2.4×10^{-3}	2.3×10^{4}	–	2.4×10^{-3}	2.3×10^{4}
11	βPro-A (**41**)	+	0.09	480	–	0.09	480
12	Aze-A (**42**)	+	0.10	414	–	0.10	414
13	Pip-A (**45**)	+	0.13	330	–	0.13	330
14	Pyrr-A (**44**)[a]	+	0.09	480	–	0.09	480
15	Azp-A (**46**)	+	3.5×10^{-2}	1.2×10^{3}	–	3.5×10^{-2}	1.2×10^{3}
16	Azc-A (**43**)	+	0.21	198	1.4×10^{-3}	2.43	0.28
17	Ppr-C (**49**)	+	0.97	42.8	7.8×10^{-3}	4.06	10.2
18	Ppr-A (**34**)[d]	+	0.62	72.6	–	0.62	72.6
19	Ppr-A (**34**)	–	3.83	10.9	5.7×10^{-2}	3.90	10.7
20	Pca-A (**39**)	–	3.80	11.0	–	3.90	10.7
21	Pyrr-A (**44**)	–	0.09	480	–	0.09	480

[a] Conditions: HEPES (0.5 M), 1-ethylimidazole (0.15 M), with or without $MgCl_2$ (0.08 M). [b] From Ref. [21]. [c] Second-order rate constant to ethylimidazolium intermediate, if detectable by ^{31}P NMR spectroscopy. [d] Assay with 1-(2-hydroxyethyl)imidazole instead of 1-ethylimidazole.

assumption that the divalent cation partially blocks the intramolecular attack of the carboxylate(s) that is/are absent in **44**. This blocking effect is apparently stronger than the accelerating effect expected for transphosphorylation reactions, where magnesium ions are known to play an important role.[31,32]

The ability to achieve rapid conversion in the absence of double-digit millimolar magnesium ion concentrations was very welcome, as such concentrations are not found in the cell. Furthermore, the need to apply such concentrations in enzyme-free genetic copying with ribozymes limits the lifetime of the biocatalyst, as RNA is susceptible to ion-catalyzed hydrolysis.[33]

The hydrolytic reactivity of the newly prepared prolinyl AMP phosphoramidates made it likely that such amino acidyl nucleotides would also be reactive toward the 3′-alcohols of the primer terminus in template-directed reactions. Alcohols and water are similar in reactivity, and the proximity effect induced by the template should compensate at least partially for the much lower concentration of the primer hydroxy group compared to the competing solvent molecules.

To test this experimentally, we used an adaptation of the primer extension system employed to demonstrate the reactivity of unmodified prolinyl NMPs.[34] As before, assays were performed at 4 °C to achieve a strong template effect. Figure 5a shows the reaction system employed in our tests. The sequences employed call for a CMP monomer, as G is the templating base at the extension site. Accordingly, three CMP phosphoramidates were prepared, namely known Pro-C (**48**) as reference compound, Ppr-C (**49**) as alkyl-substituted prolinyl building block. and Pca-C (**50**) as its δ-carboxyproline counterpart. As for all other newly prepared amino acidyl nucleotides, the analytical data as well as representative spectra are shown in the Supporting Information.

Figure 5b shows a MALDI-TOF mass spectrum of the assay solution, measured after 22 days reaction time with **49** as monomer. The result of an exploratory assay with a more poorly binding primer is shown in Figure S86 of the SI. Figure 5 indicates clean and near-quantitative conversion of the primer to the extended strand. The high level of conversion was achieved using a two-phase protocol involving the addition of a second quantity of monomer when the initial extension phase had ended, and little, if any active monomer was expected in the solution.

The kinetics of the reactions are shown in Figure 5c. It can be discerned that for either of the activated nucleotides a plateau had been reached after approximately 10 days reaction time. Addition of the second helping of phosphoramidate then reactivated the primer extension reaction, with the accompanying jump of conversion from 59 % to 92 % for Ppr-C (**48**). Numerical values for reactions with the three building blocks are compiled in Table 2. No two-stage assays have been performed with Pca-C (**50**) yet, but exploratory assays indicate that primer extension can be achieved with this monomer in the absence of magnesium ions. After 9 days at 4 °C, 20 mM Pca-C in 10 mM MOPS buffer, pH 7.5 showed 15 % conversion to the extended primer, whereas Pro-C and Ppr-C gave <5% conversion under the magnesium-free conditions after 25 days. This is an important finding, as primer extension in the absence of a significant concentration of magnesium ions is the exception, not the rule.[35,36]

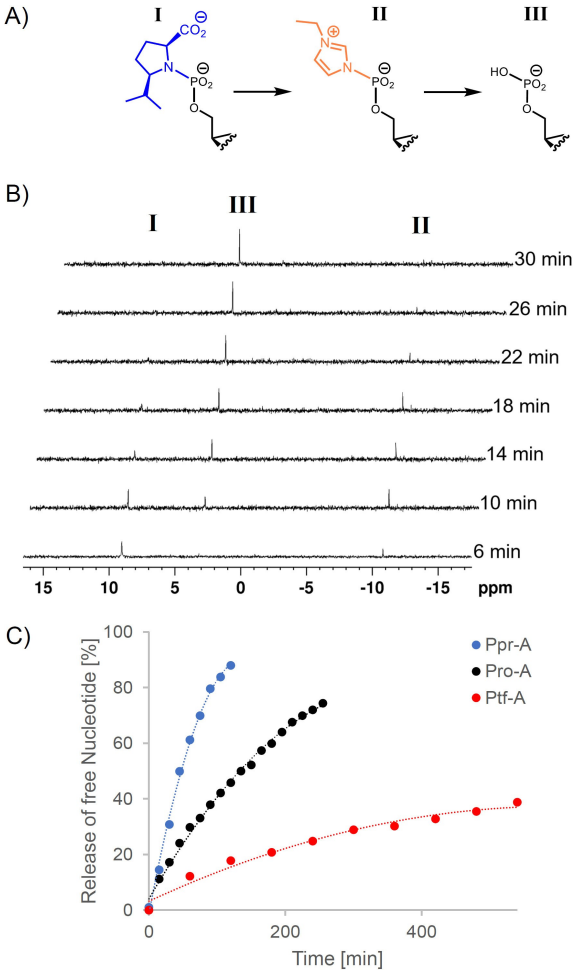


Figure 4. Representative results from hydrolysis assays. A) Two-step pathway proposed for hydrolysis of Ppr-A (**I**) via imidazolium phosphate (**II**) from which the release AMP (**III**) occurs. B) Stack plot of ^{31}P NMR spectra for Ppr-A (**34**), with signals of **I–III** labelled for a reaction carried out in assay buffer without $MgCl_2$. C) Kinetics of hydrolysis for three representative phosphoramidates, Ppr-A (**34**), Pro-A (**1**) and Ptf-A (**38**) in the presence of $MgCl_2$, lines are monoexponential fits. Assay conditions: HEPES (0.5 M), EtIm (0.15 M), with or without $MgCl_2$ (0.08 M), H_2O/D_2O 9:1, pH 7.5, 37 °C.

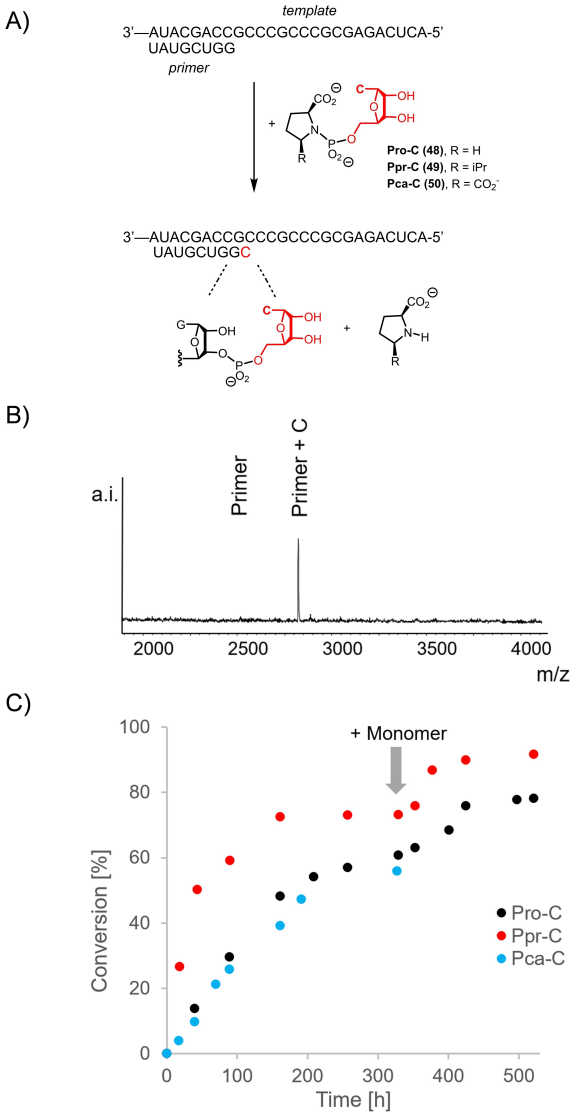


Figure 5. Primer extensions with aminoacidyl phosphoramidates of CMP. A) Components of the genetic copying reaction studied. B) MALDI-TOF spectrum of a full extended primer after 22 days, obtained for Ppr as leaving group, with addition of a second portion of monomer Ppr-C (**48**) after 14 days reaction time. C) Kinetics of primer extensions Ppr-C, Pro-C and Pca-C. Conditions: 40 μM primer, 60 μM template, 150 mM EtIm, 20 mM monomer, 80 mM $MgCl_2$ and 10 mM MOPS, pH 7.5, 4 °C. For Pro-C and Ppr-C a second dose of monomer (20 mM) was added after 14 days, when the conversion achieved with the initial amount of monomer had reached a plateau.

Taken together, these results show how adding a substituent at the δ-position of proline's pyrrolidine ring can enhance reactivity. For the alkyl derivatives **32**–**34**, the increase in reactivity is most likely the result of a Thorpe–Ingold effect.[37,38] The additional steric hindrance favors conformations that are productive in terms of a cyclization, that is, the intramolecular attack of the carboxylate on the phosphorous (step **I**, in Figure 2). Attempts to prepare a δ-*tert*-butylproline nucleotide have proceeded to the reductive cyclization of the amino acid, as described in the Supporting Information, but attempts to couple this more sterically hindered ring to AMP have been unsuccessful thus far. Because the nucleophilicity of the carboxylate is dampened upon complexation to Mg^{2+}, the acceleration of hydrolysis is observed that manifests itself in the values of entries 4 and 19 of Table 1.

The other way to increase reactivity is apparently to decorate proline with an additional carboxy group, as in Pca-A and Pca-C, so that an intramolecular attack of a carboxy oxygen can occur from either half of the pyrrolidine ring. The two carboxy groups give a double anchimeric effect. Here, the susceptibility to Mg^{2+} levels is even greater, as seen when comparing entries 9 and 20 of Table 1, as two Mg^{2+} ions can be complexed by the prolinyl residue, resulting in a double blocking effect.

Steric effects, which can be expected to affect pseudorotation equlibria of pentavalent intermediates (**II** in

Table 2: Results from enzyme-free primer extension reactions with substituted prolinyl phosphoramidates of CMP in aqueous buffer at pH 7.5 and 4 °C.[a]

Monomer	Initial Conversion to extended primer [%][b]	1st order Rate Constant k 10^{-3} [h^{-1}]	Final Conversion to extended primer [%][c]
Pro-C (**48**)[b]	44	5.49	78
Ppr-C (**49**)	59	38.9	92
Pca-C (**50**)	43	5.22	n.a.

[a] Conditions: Primer (40 μM), template (60 μM), monomer (20 mM), 1-ethylimidazole (150 mM), $MgCl_2$ (80 mM) and MOPS (10 mM). [b] Data measured after 10 days. [c] Observed at 22 days, after adding of 20 m M additional monomer after 14 days, as indicated by the gray arrow in Figure 5c.

Figure 2), appear to play a minor role, resulting in fairly similar rates of hydrolysis for amidates with ring sizes of four to six (**42**, **44**, **45**). Only when the heteroalicyclic ring reached a size of seven members (Azp-A, **46**) did reactivity go down significantly.

The acceleration of hydrolysis in the absence of magnesium ions observed for reactive prolinyl nucleotides confirms the hypothesis that the main reaction pathway for those phosphoramidates involves a nucleophilic attack of a carboxylate. It also promises productive release pathways for prodrugs of nucleotides, as discussed in the Introduction. Further, a reactivity level sufficient to induce genetic copying without the need for high Mg^{2+} levels removes a constraint from geochemical prebiotic scenarios for RNA replication in the absence of enzymes, and allays concerns that such reactions were unfeasible in more diverse geochemical settings. While we note that the more reactive prolinyl nucleotides discussed here are not found in extant biology, we suspect that they are not unreasonable as compounds in earlier stages of the molecular evolution of life.[39]

Conclusions

Amino acids play pivotal roles in biology, and five-membered heterocycles play prominent roles in synthesis, e.g. by acting as auxiliaries or organocatalysts.[24] Aromatic heterocycles were known to act as leaving groups in activated nucleotides.[40] We demonstrate here that substituted prolines, the one saturated five-membered heterocycle among the proteinogenic amino acids, have leaving group qualities that produce nucleotide monomers with more than marginal reactivity. The proline scaffold is promising for developing further improved leaving groups in reactions that occur in aqueous buffer, templated by RNA. Independent of more theoretical considerations, the observation that fast releasing and well reacting amino acidyl nucleotides exist, adds a tool to the toolbox of medicinal chemistry and related fields.[41,42]

Acknowledgements

The authors thank Dr. B. Miehlich und Dr. M. Ringenberg for computational work performed during early phases of the project and Dr. E. Kervio for discussions. This work was supported by the German Research Foundation (DFG), project-ID 364653263-CRC 235, grant RI 1063/18-1, project-ID 513030456-INST 41/1175-1 FUGG, and Volkswagen Foundation (grant Az 92 768). Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: RNA · amino acids · template effects · genetic copying · prebiotic chemistry

[1] F. H. Westheimer, *Science.* **1987**, *235*, 1173–1178.
[2] J. R. Knowles, *Annu. Rev. Biochem.* **1980**, *49*, 877–919.
[3] R. L. Letsinger, W. B. Lunsford, *J. Am. Chem. Soc.* **1976**, *98*, 3655–3661.
[4] L. J. McBride, M. H. Caruthers, *Tetrahedron Lett.* **1983**, *24*, 245–248.
[5] N. Usman, K. K. Ogilvie, M. Y. Jiang, R. J. Cedergren, *J. Am. Chem. Soc.* **1987**, *109*, 7845–7854.
[6] C. B. Reese, *Org. Biomol. Chem.* **2005**, *3*, 3851–3868.
[7] C. McGuigan, R. N. Pathirana, N. Mahmood, K. G. Devine, A. J. Hay, *Antiviral Res.* **1992**, *17*, 311–321.
[8] C. McGuigan, R. N. Pathirana, N. Mahmood, A. J. Hay, *Bioorg. Med. Chem. Lett.* **1992**, *2*, 701–704.
[9] C. McGuigan, H. W. Tsang, P. W. Sutton, *Antiviral Chem. Chemother.* **1998**, *9*, 109–115.
[10] D. Saboulard, L. Naesens, D. Cahard, A. Salgado, R. Pathirana, S. Velazquez, C. McGuigan, E. De Clercq, J. Balzarini, *Mol. Pharmacol.* **1999**, *56*, 693–704.
[11] Y. Mehellou, R. Valente, H. Mottra, E. Walsby, K. I. Mills, J. Balzarini, C.r McGuigan, *Bioorg. Med. Chem.* **2010**, *18*, 2439–2446.
[12] C. McGuigan, P. Murziani, M. Slusarczyk, B. Gonczy, J. V. Voorde, S. Liekens, J. Balzarini, *J. Med. Chem.* **2011**, *54*, 7247–7258.
[13] M. J. Sofia, D. Bao, W. Chang, J. Du, D. Nagarathnam, S. Rachakonda, P. G. Reddy, B. S. Ross, P. Wang, H.-R. Zhang, S. Bansal, C. Espiritu, M. Keilman, A. M. Lam, H. M. M. Steuer, C. Niu, M. J. Otto, P. A. Furman, *J. Med. Chem.* **2010**, *53*, 7202–7218.
[14] E. Murakami, T. Tolstykh, H. Bao, C. Niu, H. M. M. Steuer, D. Bao, W. Chang, C. Espiritu, S. Bansal, A. M. Lam, M. J. Otto, M. J. Sofia, P. A. Furman, *J. Biol. Chem.* **2010**, *285*, 34337–34347.
[15] E. Murakami, T. Wang, D. Babusis, E.-I. Lepist, D. Sauer, Y. Park, J. E. Vela, R. Shih, G. Birkus, D. Stefanidis, C. U. Kim, A. Cho, A. S. Ray, *Antimicrob. Agents Chemother.* **2014**, *58*, 1943–1951.

[16] T.-F. Chou, J. Baraniak, R. Kaczmarek, X. Zhou, J. Cheng, B. Ghosh, C. R. Wagner, *Mol. Pharmaceutics* **2007**, *4*, 208–217.
[17] P. A. Furman, E. Murakami, C. Niu, A. M. Lam, C. Espiritu, S. Bansal, H. Bao, T. Tolstykh, H. M. Steuer, M. Keilman, V. Zennou, N. Bourne, R. L. Veselenak, W. Chang, B. S. Ross, J. Du, M. J. Otto, M. J. Sofia, *Antiviral Res.* **2011**, *91*, 120–132.
[18] J. H. Beigel, K. M. Tomashek, L. E. Dodd, A. K. Mehta, B. S. Zingman, A. C. Kalil, E. Hohmann, H. Y. Chu, A. Luetkemeyer, S. Kline, D. Lopez de Castilla, R. W. Finberg, K. Dierberg, V. Tapson, L. Hsieh, T. F. Patterson, R. Paredes, D. A. Sweeney, W. R. Short, G. Touloumi, D. C. Lye, N. Ohmagari, M. Oh, G. M. Ruiz-Palacios, T. Benfield, G. Fätkenheuer, M. G. Kortepeter, R. L. Atmar, C. B. Creech, J. Lundgren, A. G. Babiker, S. Pett, J. D. Neaton, T. H. Burgess, T. Bonnett, M. Green, M. Makowski, A. Osinusi, S. Nayak, H. C. Lane, *N. Engl. J. Med.* **2020**, *383*, 1813–1826.
[19] D. Babusis, M. P. Curry, B. Kirby, Y. Park, E. Murakami, T. Wang, A. Mathias, N. Afdhal, J. G. McHutchison, A. S. Ray, *Antimicrob. Agents Chemother.* **2018**, *62*, e02587–17.
[20] M. Slusarczyk, M. Serpi, F. Pertusati, *Antiviral Chem. Chemother.* **2018**, *26*, doi: 10.1177/2040206618775243.
[21] D. Jovanovic, P. Tremmel, P. S. Pallan, M. Egli, C. Richert, *Angew. Chem. Int. Ed.* **2020**, *59*, 20154–20160.
[22] B. List, *Synlett* **2001**, *26*, 1675–1686.
[23] H. Kotsuki, H. Ikishima, A. Okuyama, *Heterocycles* **2008**, *75*, 493–529.
[24] J. Seayad, B. List, *Org. Biomol. Chem.* **2005**, *3*, 719–724.
[25] R. Lohrmann, L. E. Orgel, *Nature* **1976**, *261*, 342–344.
[26] J. A. Rojas Stütz, E. Kervio, C. Deck, C. Richert, *Chem. Biodiversity* **2007**, *4*, 784–802.
[27] L. Li, No. Prywes, C. P. Tam, D. K. O'Flaherty, V. S. Lelyveld, E. C. Izgu, A. Pal, J. W. Szostak, *J. Am. Chem. Soc.* **2017**, *139*, 1810–1813.
[28] D. Sievers, G. von Kiedrowski, *Nature* **1994**, *369*, 221–224.
[29] P. Tremmel, H. Griesser, U. E. Steiner, C. Richert, *Angew. Chem. Int. Ed.* **2019**, *58*, 13087–13092.
[30] S. Gharpure, L. Nanda, D. Kumari, *Eur. J. Org. Chem.* **2017**, 3917–3920.
[31] A. S. Mildvan, L. A. Loeb, *CRC Crit. Rev. Biochem.* **1979**, *6*, 219–244.
[32] T. A. Steitz, J. A. Steitz, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 6498–6502.
[33] A. Wochner, J. Attwater, A. Coulson, P. Holliger, *Science* **2011**, *332*, 209–212.
[34] F. Welsch, E. Kervio, P. Tremmel, C. Richert, *Angew. Chem. Int. Ed.* **2023**, *62*, e202307591.
[35] S. Motsch, P. Tremmel, C. Richert, *Nucleic Acids Res.* **2020**, *48*, 1097–1107.
[36] S. Motsch, D. Pfeffer, C. Richert, *ChemBioChem* **2020**, *21*, 2013–2018.
[37] R. M. Beesley, C. K. Ingold, J. F. Thorpe, *J. Chem. Soc. Trans.* **1915**, *107*, 1080–1106.
[38] B. L. Shaw, *J. Am. Chem. Soc.* **1975**, *97*, 3856–3857.
[39] C. De Duve, *Nature* **2005**, *433*, 581–582.
[40] J. A. Rojas Stütz, E. Kervio, C. Deck, C. Richert, *Chem. Biodiversity* **2007**, *4*, 784–802.
[41] O. Adelfinskaya, P. Herdewijn, *Angew. Chem. Int. Ed.* **2007**, *46*, 4356–4358.
[42] A. Giraut, P. Herdewijn, *ChemBioChem* **2010**, *11*, 1399–1403.

Manuscript received: December 23, 2023
Accepted manuscript online: February 1, 2024
Version of record online: February 20, 2024