Determining the Diastereoselectivity of the Formation of Dipeptidonucleotides by NMR Spectroscopy

Olivia Doppleb[+],[a] Jennifer Bremer[+],[a] Maren Bechthold,[a] Carolina Sánchez Rico,[a] Daniela Göhringer,[a] Helmut Griesser,[a] and Clemens Richert*[a]

Abstract: Proteins are composed of L-amino acids, but nucleic acids and most oligosaccharides contain D-sugars as building blocks. It is interesting to ask whether this is a coincidence or a consequence of the functional interplay of these biomolecules. One reaction that provides an opportunity to study this interplay is the formation of phosphoramidate-linked peptido RNA from amino acids and ribonucleotides in aqueous condensation buffer. Here we report how the diastereoselectivity of the first peptide coupling of the peptido RNA pathway can be determined in situ by NMR spectroscopy. When a racemic mixture of an amino acid ester was allowed to react with an 5′-aminoacidyl nucleotide, diastereomeric ratios of up to 72:28 of the resulting dipeptido nucleotides were found by integration of ^{31}P- or ^{1}H-NMR peaks. The highest diastereomeric excess was found for the homochiral coupling product D-Ser-D-Trp, phosphoramidate-linked to adenosine 5′-monophosphate with its D-ribose ring. When control reactions with an *N*-acetyl amino acid and valine methyl ester were run in organic solvent, the diastereoselectivity was found to be lower, with diastereomeric ratios $\leq$62:38. The results from the exploratory study thus indicate that the ribonucleotide residue not only facilitates the coupling of lipophilic amino acids in aqueous medium but also the formation of a homochiral dipeptide. The methodology described here may be used to search for other stereoselective reactions that shed light on the origin of homochirality.

Introduction

Homochirality is a trait of living systems that has fascinated scientists for decades.[1,2] In the cell, all encoded proteins are made up of L-amino acids, all nucleic acids contain D-ribose or D-deoxyribose, and the carbohydrates of primary metabolism are D-sugars, leading to the question of what the origin of homochirality was. There are functional requirements that favor homochirality, both in terms of folding and in terms of substrate specificity of enzymes or ribozymes that produce them, but molecular evolution is likely to have produced (and selected) molecules before they had a specific function.[3] A chemical process must have favored one stereochemical configuration over the other, and functional selection may then have led to homochirality.[4] What that process was, however, remains open to debate.[5,6]

For nucleic acids, replication is an obvious candidate for the process that may have led to homochirality. Enzyme-free copying of homochiral templates with mixtures of enantiomers as monomers has been studied in detail,[7,8,9] and so have the pairing properties of oligonucleotides that may have formed by untemplated oligomerization.[10,11] For peptides, the situation is less clear.[12] Amino acids of one stereochemical configuration may have been enriched by abiotic processes, such as crystallization or sublimation,[13,14] but selection of monomers is just the first step toward the synthesis of homochiral, folded proteins of the >100 residues typical for fully functional enzymes. Chiral amplification can occur during polymerization, but this process is complicated by racemization[15] and polymerization does not readily produce enzymes.

Self-replication of peptides is known,[16,17] including processes that are stereoselective,[18] but nature does not use such self-templated reactions to produce proteins in the cell, employing translation instead. Translation is an RNA-mediated process, and there is no simple path from preselected unencoded proteins to the encoded proteins of extant biology. In translation, an mRNA templates the incorporation of individual amino acid residues into a growing polypeptide chain, catalyzed by the RNA-based ribosomal machinery.[19] It is therefore interesting to ask whether ribonucleotides may have interacted with amino acids in a synergistic process that favored one stereochemical configuration over the other.

Lacey and coworkers had shown that mono- and bis(2′,3′-aminoacyl) esters of AMP can form stereoselectively,[20,21] but this work did not include peptide-forming reactions. We recently found that amino acids and ribonucleotides form

[a] O. Doppleb,[+] J. Bremer,[+] Dr. M. Bechthold, Dr. C. Sánchez Rico, D. Göhringer, H. Griesser, Prof. C. Richert
Institut für Organische Chemie, Universität Stuttgart
Pfaffenwaldring 55, 70569 Stuttgart (Germany)
E-mail: lehrstuhl-2@oc.uni-stuttgart.de

[+] These authors contributed equally to this manuscript.

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202101630

covalently linked species when allowed to react in aqueous condensation buffer.[22] Peptide coupling is accelerated once a phosphoramidate bond between a ribonucleotide and the first amino acid is established,[23] and organocatalysis plays a role in making chain growth efficient, both in the peptide and in the RNA reaction channel.[24] The chemistry producing the covalently linked 'peptido RNAs' in the condensation mixture is compatible with all proteinogenic amino acids,[25] and also induces the formation of cofactors from precursor molecules.[22] The phosphoramidates of amino acids and nucleoside 5′-monophosphates are called 'aminoacidyl nucleotides' or 'amino acidyl nucleotides'[30] to avoid confusing them with 2′/3′-aminoacyl nucleotides that feature an ester link between the carboxy group of the amino acid moiety and the ribose ring of the nucleotide. Alaninyl nucleotides are long-known intermediates of the metabolism of antiviral prodrugs,[26] and are structurally different from the alanyl products of aminoacylation.

When the first X-ray crystal structures of aminoacidyl nucleotides became available, it was found that the carboxy group of the phosphoramidate-linked amino acid can be close to the sugar and base (Figure 1). Since the ribonucleotide is chiral, it could cause a stereochemical induction that results in a more stereoselective coupling of the next amino acid. Whether this is the case was of particular interest, as peptido RNA may have been a transitional species that paved the way for an early form of translation that relied on doubly anchored peptides.[27] The question of stereoselection prompted us to develop an assay for measuring the diastereoselectivity of the reactions forming dipeptido nucleotides. A methodology was called for that was suitable for a range of different amino acids, and that avoided potentially biasing and time-consuming work-ups. Using an approach for in situ monitoring, such as NMR spectroscopy, appeared promising, since the selectivity of other reactions in condensation buffer had recently been determined successfully using nuclear resonance techniques.[28,29] Here we report how monitoring by NMR can be used to determine diastereomeric ratios, together with the results of an exploratory study into the diastereoselectivity of the formation of dipeptido nucleotides in condensation buffer.

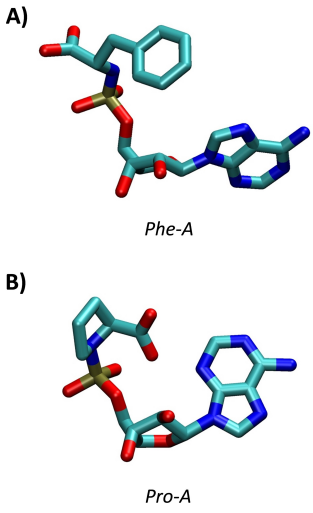


Figure 1. Conformations of aminoacidyl nucleotides found in X-ray crystal structures, with A) phenylalanine and B) proline as amino acid residue and AMP as nucleotide. Graphics were generated in VMD, using coordinates of PDB entries 6XHF and 6XHD.[30]

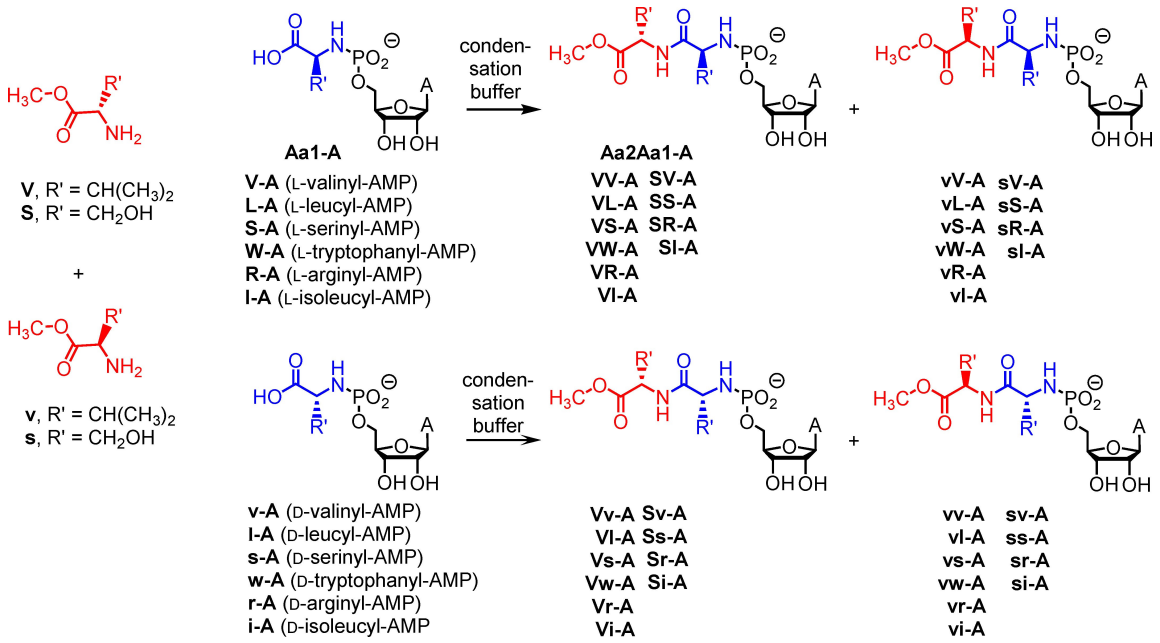
Results

Scheme 1 shows the reactions studied. They involve the first peptide-forming coupling step of the peptido RNA pathway.[22] An amino acid that was captured as a phosphoramidate by a ribonucleotide[23] couples with an amino acid ester. Among the condensation agents that induce the reactions of the peptido RNA pathway, such as carbonyl diimidazole, carbodiimides and cyanamide, the carbodiimides are the most reasonable choice, as they give the fastest conversion.[23] The use of an ester avoids oligomerization and limits the reaction to a single coupling event. Except for glycine, either of the amino acid residues, i.e. that of the aminoacidyl phosphoramidate (Aa-NMP) and the amino acid ester are chiral, so that different diasteromers are formed in the coupling. The ribonucleotide residue is chiral and enhances or decreases the intrinsic diastereoselectivity of the coupling reaction. It also increases the water solubility of the covalently linked amino acid, as well as that of the dipeptide product, facilitating reactions with lipophilic amino acids.

We focused our study on phosphoramidates of adenosine 5′-monophosphate (AMP), which is abbreviated "A" in the shorthand used. Our shorthand further uses upper case single-letters for L-amino acid residues, and lower case single-letters for D-amino acids. The amino acid linked to the ribonucleotide via its amino group is referred to as Aa1, and the amino acid ester reacting with it to form the dipeptido nucleotide is amino acid residue Aa2 in the products. So, the shorthand **Vs-A** stands for the dipeptido AMP with a *C*-terminal L-valine residue linked to a D-serine residue, which is *N*-linked to adenosine 5′-monophosphate.

The aminoacidyl nucleotides of Scheme 1 were prepared from the 2-methylimidazolide of AMP and free amino acids in aqueous solution, following a synthetic approach described recently.[30] Details of the synthetic protocol and spectroscopic data are given in the Supporting Information (SI). The group of amino acids linked to the ribonucleotide included aliphatic, cationic and aromatic amino acids to sample a certain breadth of amino acid structure space. The two starting materials were reacted in condensation buffer containing 1,4-piperazine diethanesulfonic acid (PIPES), 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC), and $MgCl_2$ at pH 7.5 and 0 °C, i.e. conditions similar to those used for assays producing peptido RNA in the past.[22,23,24]

The outcome of the reactions was analyzed by NMR spectroscopy. The use of PIPES instead of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) as buffer reduced the number of signals and thus the extent of peak overlap. The racemic mixture of amino acid esters was prepared by mixing equal amounts of the pure enantiomers, as the commercial racemate of valine methyl esters was of lower purity and contained impurities that led to peak broadening and compli-



Scheme 1. Coupling reactions between racemic mixtures of amino acid esters and aminoacidyl nucleotides studied in this work. Single letter code is used for amino acids and the ribonucleotide. The L-amino acids found in proteins are given as upper case letters, D-amino acids are given as lower case letters. The different diastereomeric products are listed as **Aa2-Aa1-A** on the right-hand side. Reactions were performed in aqueous condensation buffer, pH 7.5, containing 500 mM PIPES, 80 mM $MgCl_2$, 800 mM EDC at 0 °C for 3 d.

cated assays. In some cases, peak broadening was additionally reduced by ten-fold dilution of the reaction mixture prior to NMR analysis.

Different NMR experiments were tested for measuring the diastereomeric ratio of the two products of a given coupling reaction. The challenge in this was to identify a pair of signals of sufficiently different chemical shift and little or no overlap with other peaks. Three different types of signals were considered. The first pair considered was that of the phosphorus signals in the one-dimensional ^{31}P NMR spectra. This is the most straightforward type of analysis, as nucleotides contain a single phosphorus atom, and reactions that go to completion without significant side products produce only two peaks. For **V-A**, product peaks were found to be separated from that of the starting material by approx. 1 ppm, so overlap with possible educt peaks was not an issue. In order to assign the peaks of the diastereomers, separate reactions were run with either of the two enantiomers of the amino acid ester alone, followed by overlaying the spectra (Figure 2). Coupling was found to be complete after 24 h, when no signal for the starting material was detectable in ^{31}P NMR spectra of the valine-valinyl AMP case (Figure S8, Supporting Information). In order to be on the safe side for reactions that are even slower than that between those sterically hindered amino acids, we analyzed subsequent assays after 72 h reaction time. There was no significant change in peak intensity between spectra measured after 24 h and after 72 h (Figure S9 of the Supporting Information).

Spectra for quantitative analysis were acquired with a recycle delay of 5 s to allow for near-complete relaxation. For partially overlapping peaks, deconvolution was achieved using the Lorentz/Gauss fit option in the NMR analysis software (Topspin). This is shown in Figure 2D for the assay producing the diastereomers of **(V/v)V-A**. The ratios of the integrals were then used to calculate the diastereomeric ratios.

For some dipeptido nucleotides, the peak overlap was too severe to yield reliable integration data in the ^{31}P channel, even when 150 mM EDTA had been added to chelate magnesium ions. In these instances, signals from other NMR experiments had to be relied upon. So, the second type of signals studied was that of proton resonances in the one-dimensional ^{1}H NMR spectra, with the singlet for the methyl group of the *C*-terminal residue in the chemical shift ranged of 3.4–3.9 ppm as the default choice. Again, the assignment was based on control experiments with the individual enantiomers of the ester (Figure 3). Partially overlapped peaks were deconvoluted by the same fit procedure as for ^{31}P spectra (see above). When the results from the two NMR channels did not agree, or when smaller peaks were suspected to overlap with the signals to be integrated, a third analysis was performed. For this, magnitude mode ^{1}H–^{1}H–COSY spectra were recorded, and cross peaks originating from corresponding protons of the diastereomers were integrated.

In doing so, we assumed that the scalar coupling leading to these cross peaks was similar for the diastereomers, so that cross peak intensities scale with the concentration of the analytes. As shown in Figure 4, the COSY experiment is well suited to identify peak overlap that may not be trivial to spot in the one-dimensional spectrum. The CHα to CHβ cross peaks of the *C*-terminal valine residues were found to be well suited to

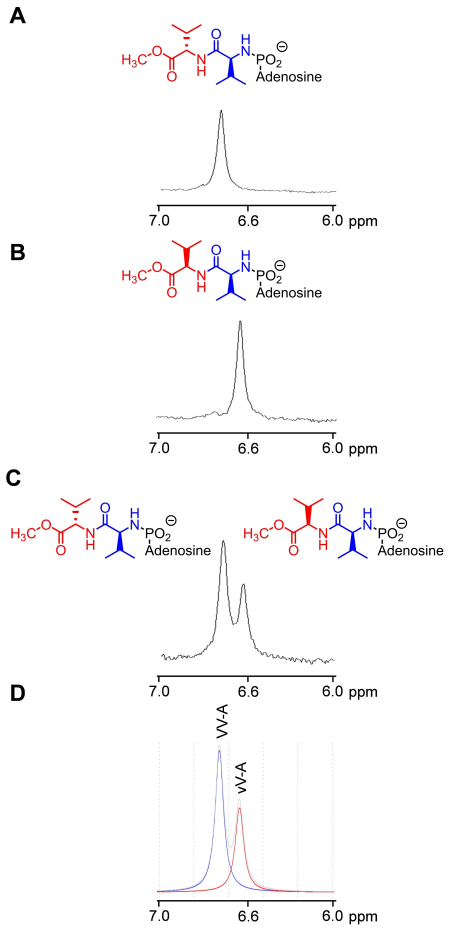


Figure 2. A section of the ^{31}P NMR spectra of the crude of the reaction between L-Val-AMP (**V-A**) and D/L-valine methyl esters is shown. A) reaction of L-Val-AMP (**V-A**) and L-valine methyl ester (**V**). B) reaction between L-Val-AMP (**V-A**) and D-valine methyl ester. C) Reaction between L-Val-AMP (**V-A**) and the racemate of D/L-valine methyl esters. D) Deconvolution of the overlapping peaks of part C with Lorentz-Gauss fit. The blue fit line represents the signal of **VV-A**, and the red fit line represents the signal of **vV-A**. Conditions: 120 mM L-Val-AMP, 240 mM valine methyl esters, 800 mM EDC, 500 mM PIPES, and 80 mM $MgCl_2$, 0 °C, 283 MHz.

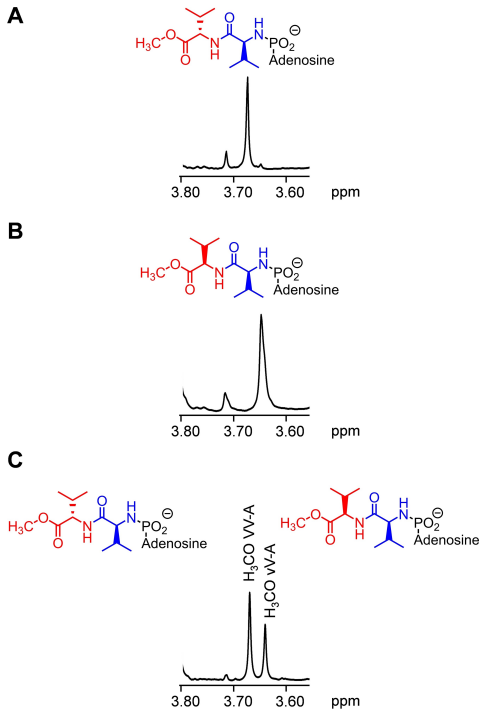


Figure 3. A section of the ^{1}H NMR spectra of the crude of the reaction between L-Val-AMP (**V-A**) and D/L-valine methyl esters is shown. A) reaction between L-Val-AMP (**V-A**) and L-valine methyl ester (**V**) B) reaction between L-Val-AMP (**V-A**) and D-valine methyl ester (**v**). C) reaction between L-Val-AMP (**V-A**) and the racemate of D/L-valine methyl esters (**V:v**). Conditions: 120 mM L-Val-AMP, 240 mM of the valine methyl esters, 800 mM EDC, 500 mM PIPES and 80 mM $MgCl_2$ at 0 °C, 700 MHz.

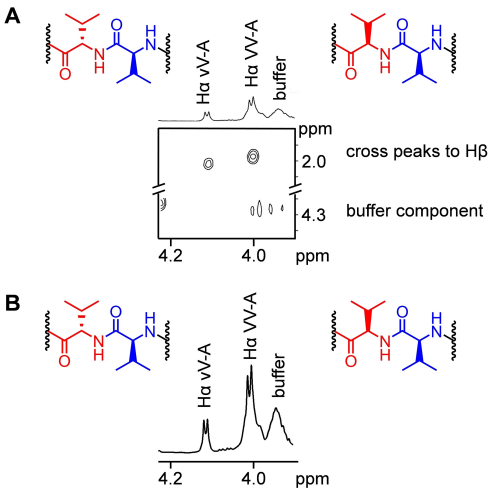


Figure 4. Sections of the 1D and 2D NMR spectra of the reaction mixture involving L-Val-AMP (**V-A**) and D/L-valine methyl esters. A) COSY NMR spectra showing crosspeaks of CHα and CHβ of the valine methyl ester residue of each diastereomer. B) 1D ^{1}H NMR spectrum with partial overlap of signals resolved in the 2D spectrum. Conditions: 120 mM L-Val-AMP, 240 mM valine methyl esters, 0.8 M EDC, 0.5 M PIPES, 80 mM $MgCl_2$, 0 °C, 700 MHz.

obtain signals unencumbered by other components of the reaction mixture, providing the additional set of data, where needed. In the cases where all three NMR experiments were performed, satisfactory agreement was obtained (Table 1).

In Table 1, the results of coupling with the valine methyl esters are compiled, whereas Table 2 lists the diastereoselectivities measured for coupling with serine methyl esters. The assays underlying the first two entries of Table 2 were performed three times to study reproducibility. In either case, the standard deviation was found to be around 5 %. This is similar to the differences found between the results from measurements in different NMR channels used, providing a limit to what we consider a significant difference in diastereoselectivity. Interestingly, the homochiral product was the dominant diastereomer in all assays performed.

Besides the assays involving aminoacidyl nucleotides, we also performed selected control reactions with amino acids bearing an acetyl group at their amino group instead of the ribonucleotide. Attempts to perform these reactions in aqueous condensation buffer were unsuccessful, as the solubility of the amino acid components or the peptide products was too low

Table 1. Diastereoselectivity of peptide coupling between aminoacidyl nucleotides and a racemic mixture of valine methyl esters, as determined by peak integration in ^{1}H, ^{31}P or COSY NMR spectra. See Supporting Information for details.

Entry No.	Aa1-A	diasteromeric ratio ^{1}H (OMe) (L/L : L/D)	^{31}P (L/L : L/D)	I_{COSY}[e] (L/L : LD)
1	**V-A**	60:40[a]	62:38[b]	62:38
2	**L-A**	(58:42)[c]	63:37	
3	**S-A**	48:52	52:48	60:40
4	**W-A**	57:43	57:43	
5	**R-A**	58:42	(57:43)[c]	58:42
6	**I-A**	59:41	62:38	
		(D/D : D/L)	(D/D : D/L)	(D/D : D/L)
7	**v-A**	54:46	/[d]	
8	**I-A**	58:42	/[d]	
9	**s-A**	56:44	60:40	61:39
10	**w-A**	65:35	70:30	
11	**r-A**	58:42	/[d]	
12	**i-A**	55:45	56:44	

[a] Mean of three independent experiments; standard deviation ±1.8 = 3% for L/L-diastereomer and 5% for D/L-diastereomer. [b] Mean of three independent experiments; standard deviation ±2.7 or 4% for L/L-diastereomer and 7% for D/L-diastereomer. [c] Relative peak height of partially overlapped peaks. [d] Insufficient separation of signals of diastereomers in ^{31}P channel. [e] Ratio of intensities of magnitude-mode COSY cross peaks for Hα-Hβ of *C*-terminal valine residue.

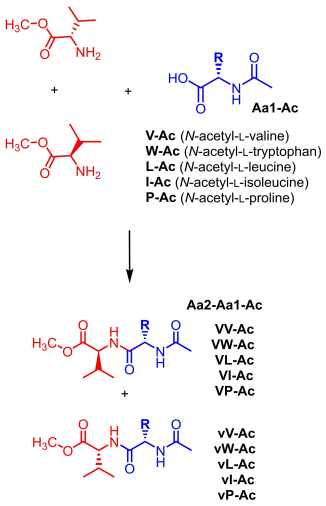
Table 2. Diastereoselectivity of peptide coupling between aminoacidyl nucleotides and a racemic mixture of serine methyl esters, as determined by peak integration in ^{1}H or ^{31}P NMR spectra. Assays were performed following General Protocol A (Supporting Information).

Entry No.	Aa1-A	diasteromeric ratio ^{1}H (OMe) (L/L : L/D)	^{31}P (L/L : L/D)
1	**V-A**	70:30	/[a]
2	**S-A**	53:47	58:42
3	**R-A**	55:45	(54:46)[c]
4	**W-A**	65:35	65:35
5	**I-A**	58:42	(54:46)[c]
		(D/D : D/L)	(D/D : D/L)
6	**v-A**	61:39	64:36
7	**s-A**	57:43	57:43
8	**r-A**	56:44	(56:44)[c]
9	**w-A**	72:28	74:26
10	**i-A**	59:41	58:42

[a] Insufficient separation of signals of diastereomers in ^{31}P channel. [b] Insufficient separation of signals of diastereomers in ^{1}H channel. [c] Estimated, based on peak heights.

under our assay conditions. This forced us to perform the control reactions in $CDCl_3$. Scheme 2 shows the reactions performed, and Table 3 lists the results. We did not detect signs of racemization of the *N*-acetyl amino acids in the NMR spectra of products, where a racemization during activation would manifest itself in peaks for additional diastereomers (compare Figure S12 of the Supporting Information).

Since no phosphorus signal was available in this case, we analyzed one- and two-dimensional ^{1}H NMR spectra only. Again, reproducibility was determined by repeating representative



Scheme 2. Control reactions performed to determine the diastereoselectivity of peptide coupling between acetyl protected L-amino acids and the racemate of D/L-valine methyl esters with 800 mM EDC in $CDCl_3$, 0 °C for 3 d. The general structure of **Aa1-Ac** does not properly represent the cyclic structure of *N*-acetyl proline (**P-Ac**).

Table 3. Diastereoselectivity of peptide coupling between *N*-acetyl-L-amino acids and D/L-valine methyl esters in deuterated chloroform, as expressed in LL *to* LD ratios, determined by peak integration in ^{1}H or ^{1}H,^{1}H–COSY NMR spectra. See the Supporting Information for further details.

Entry No.	Aa1-A	diasteromeric ratio ^{1}H (OMe) (L/L : L/D)	I_{COSY}[a] (L/L : L/D)
1	**V-Ac**	55:45[b]	56:44[c]
2	**W-Ac**	45:55	49:51
3	**L-Ac**	60:40	62:38
4	**I-Ac**	/[d]	47:53
5	**P-Ac**	45:55	47:53

[a] Ratio of intensities Hα-Hβ cross peaks of *C*-terminal valine residue in magnitude mode COSY spectra. [b] Mean of five independent experiments; standard deviation ±1.9 or 4% for L/L-diastereomer or 5% for D/L-diastereomer. [c] Mean of five independent experiments; standard deviation ±2.4 = 4% for L/L-diastereomer or 5% for D/L-diastereomer. [d] Insufficient separation of signals for diastereomers in ^{1}H channel.

assays, providing the standard deviations reported in the footnotes of Table 3.

Unlike the assays involving ribonucleotide-linked amino acids, dipeptides with both the homo- and the heterochiral stereochemistry were found as the more prevalent product. Further, the overall diastereoselectivity was found to be lower than that measured in the nucleotide-based assays.

In a separate set of experiments, it was determined that, under our reaction conditions, the capture of valine methyl esters by free AMP give a diastereomeric ratio of **V-A**:**v-A** of 53:47 (Figure S11 of the Supporting Information). This modest level of diastereoselectivity is not surprising, given that the stereogenic center of the amino acid residue and the C4′ center of the nucleotide are separated by a total of five bonds.

Discussion

Our results show the potential, but also the limitations of the NMR-based detection of diastereoselectivity. Among the strengths of this approach is the ease of performing the assays, which do not require any work-up. A work-up is often the most time-consuming part of a protocol. Further, work-up steps such as extraction and precipitation can bias the result, as the solubility of the products in the different phases can differ significantly, as known from fractionated crystallization as a technique for chiral resolution. Further, the assays are inexpensive, as no isotope-labeled starting materials or reagents are required.

Among the limitations of the approach is the reliance on separated peaks for the pair of diastereomeric products. While one such pair could be identified in most spectra of reaction mixtures, peak overlap with resonances from the buffer components or with educt peaks can limit the ability to determine the relative concentration of the two diastereomers. Also, unambiguously assigning peaks requires an analytical effort. The resonances of the diastereomers listed in Table S1 (Supporting Information) may be used to avoid this effort in future studies. Likewise, the signals found for the aminoacidyl nucleotides with D- or L-amino acids prepared by us may be used in studies on the diastereoselectivity of the capture reaction linking the first amino acid (**Aa1**) to the ribonucleotide, an aspect of stereoselection[31] not studied in sufficient detail by us.

Peak overlap prevented us from determining the diastereoselectivity of coupling to prolinyl AMP, a reaction that deserves to be studied in detail, as the release of peptides with an *N*-terminal proline residue can be faster than that with other amino acids.[30] It is reasonable to assume that the lack of dispersion for the peaks observed in this case may be reduced by resorting to spectrometers with higher magnetic field, optimized assay conditions, and/or the addition of shift reagents. Additives, such as a relaxation enhancer or a chelating agent have been employed in similar cases.[24]

Despite the limitations of the current study, effects worth discussing were observed. Figure 5 shows a comparison of the diastereomeric excess (% *de*) measured for *N*-acetyl amino acids and the corresponding ribonucleotide-linked amino acid residues. Except for leucine, for which comparable levels of diastereoselectivity were found in the two systems, significantly lower *de* values were found in the former case, including two cases with a modest preference for the heterochiral product (Figure 5a).

This indicates that the ribonucleotide has a significant effect on the diastereoselectivity of peptide formation, at least if one assumes that the change in the reaction medium (aqueous vs. organic solvent) does not change the intrinsic diastereoselectivity. Both solvents are achiral, and the temperature at which the assays were performed was the same, but the solvation of the transition states will be different. The effect found in the peptidonucleotide case may be due to the steric bulk of the nucleotide residue and/or a stereochemical induction caused by the D-riboside. If the latter was the case, it should manifest itself

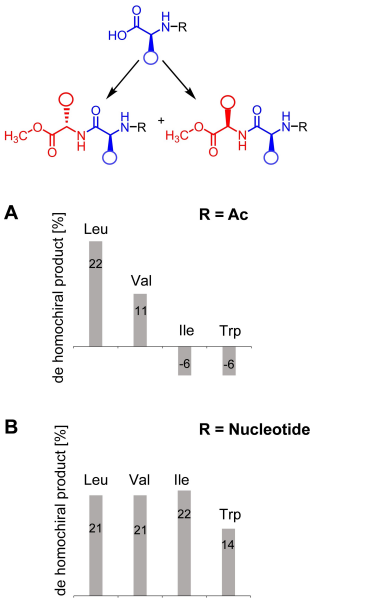


Figure 5. Bar graphs of the effect of the nucleotide residue on the diastereoselectivity of peptide coupling. A) Percent diastereomeric excess of the homochiral product in control reactions in chloroform. B) Percent diasteromeric excess of homochiral dipeptide in reactions with nucleotide residues. The numeric values are given on each bar. See Tables 1 and 3 for details.

in the couplings with opposite chirality of **Aa1**, but the same chirality of AMP. Figure 6 shows a plot of the percent diastereomeric excess of the homochiral product for each stereochemical configuration of **Aa1**. The largest *de* values were

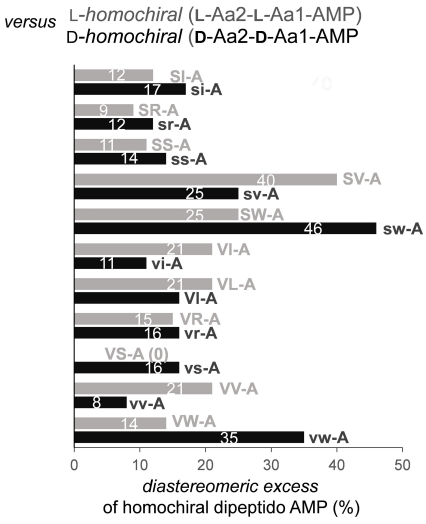


Figure 6. Bar graph of the percent diastereomeric excesses of the homochiral products of the reactions producing dipeptido nucleotides, with pairs of homochiral dipeptides of the same sequence shown next to each other. The ^{1}H- and ^{31}P-NMR results of Tables 1 and 3 were averaged, where more than one value was available, and were plotted in alphabetical order of the Aa1 residues, with the couplings to serine esters shown above those with valine esters. The numerical value of each *de* value is given on the respective bar.

observed with D-tryptophan as **Aa1**, but the L/L-homochiral peptido nucleotide **SV-A** also gave a *de* of 40%, and there are no simple trends among the amino acids studied. Still, the overall number of cases where the D/D dipeptides are formed with higher *de* than the L/L dipeptides is larger (7 versus 4 cases). The level of diastereoselectivity is similar to that found by Pascal and colleagues in their study on peptide couplings via 5(4H)-oxazolones.[32] Our work is different from this, as our phosphoramidates do not form oxazolones, and there is no sign of racemization in our products in the control reactions (Figure 2 A/B and 3 A/B), but the predominance of homochiral peptide chains is the same. Further, the increase in diastereoselectivity observed by Pascal et al. for longer peptides bodes well for homochirality of longer peptides formed in the peptido RNA pathway. Exploratory experiments involving the hydrolysis of the *C*-terminal methyl esters of **(V/v)V-A** to the corresponding free carboxylates, followed by the acquisition of NMR spectra, gave resolved CHα to CHβ cross peaks for the diastereomers in the COSY spectrum (Figure S10, Supporting Information), indicating that our methodology is not limited to esters. The maximum length of the peptido chain for which diastereomers may be resolved in future studies will depend on the field strength of the NMR spectrometer and the sequence of amino acids involved.

A full theoretical treatment of the causes of stereoselectivity observed is outside the scope of our experimental study, but there is little doubt that the attack of the amino group of the incoming amino acid on the activated carboxylate is the selectivity-determining step of the reaction. Further, it is reasonable to assume that the energetically most favorable trajectory for the attack will occur at an angle close to what is called the "Bürgi-Dunitz angle".[33] Hints of what the stereochemical situation is for such an attack may be gleaned from the exploratory modeling results shown in Figure 7. The L-valine-L-valine dipeptide *N*-terminally linked to an acetyl group (Figure 7a) or AMP was built using commercial software. Figure 7a shows the lowest energy conformation found for **VV-Ac** after two molecular dynamics runs in Chem3D, using the MM2 option, followed by force field minimization. This conformation has the *N*-terminal acetyl group located above the the peptide bond, a location where the much more bulky ribonucleotide may cause a steric conflict. The structure obtained for **VV-A** (Figure 7b) is a different conformer, with the *C*-terminus of the peptide on the same side of the plane defined by the peptide amide, as the AMP residue. Here, the isopropyl side chains of the sterically demanding valine residues found at what would be called a 'gauche conformation', if the two CH atoms were connected by a single bond.

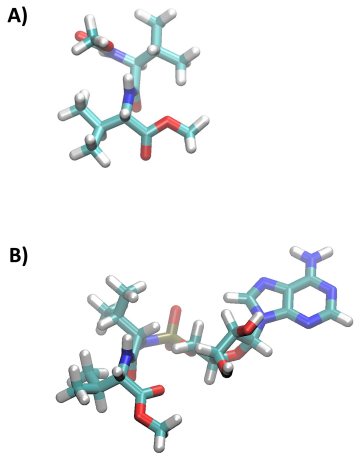


Figure 7. Conformation of the methyl ester of the homochiral dipeptide L-valine-L-valine in acetylated or phosphoramidate-linked form, as obtained by molecular modeling, using the molecular builder of Chem3D, version 15.1. Shown is the view along the peptide bond from the *C*-terminus to the *N*-terminus. A) *N*-Acetyl dipeptide, in the preferred conformation obtained by molecular dynamics and energy minimization in the MM2 force field. B) The same dipeptide ester with AMP as *N*-terminal substituent.

It is too early to conclude what the relevance of these ground state models are for the diastereomorphic transition states that lead to the different products, but it is tempting to conclude that the ribonucleotide is reasonably well positioned to cause a specific inducing effect. It will be interesting to see how the stereoselectivity develops as the peptide chain grows further, reaching a length for which significant folding or intermolecular assembly can be expected. As the peptide chain grows, the induction provided by the nucleotide as chiral substituent on the *N*-terminal amino acid residue can be expected to wane, but folding may bring the *C*-terminus into closer proximity again, so that the riboside may act as a chiral nucleus of what is eventually a stereochemically enriched product, with a stereochemical bias that may tip the balance toward one stereochemistry. The numerical values observed for the first peptide coupling here are significant, but not sufficient to predict what the preferred combination of D/L-residues in the RNA and peptide structure space are that could have manifest themselves as the "winners" of an evolutionary selection.

Independent of such considerations, it is important to note that the results from our exploratory experiments with the NMR-based method provide a very limited view of the structure space of diastereomeric dipeptides. The full matrix of the four canonical ribonucleotides and 2×20 proteinogenic amino acids of either stereochemical configuration almost certainly offer an opportunity to discover more stereoselective reactions. Further, it is likely that the selectivity can be affected by organocatalysis.[24] Both chiral or achiral organocatalysts may be employed, and pre-activation rather than in situ activation can be used to induce peptide formation. Other experimental parameters, such as temperature, pH and salt content may also be optimized to identify reactions that give even higher levels of diastereoselectivity than observed here. These reaction conditions may be chosen to best reflect our understanding of the primitive Earth. We also note that processes other than peptido RNA formation can induce a strong selection of enantiomers.[34,35] Finally, a covalent anchoring of *both* amino acids producing the dipeptide to (oligo)ribonucleotides and coupling under the constraints of a template-driven, early form of translation may have enhanced selectivity.[27] Still, well before this, a strong preference for homochirality may have manifested itself in coupling reactions of the peptido RNA pathway.

Conclusions

Here we show that the diastereoselectivity of reactions producing dipeptide chains from an amino acid ester and a phosphoramidate-linked amino acid can be determined in situ by NMR spectroscopy. The covalent link of the amino acid emerging as the *N*-terminus of the peptide chain to the ribonucleotide ensures sufficient reactivity and solubility in aqueous media. Further, the spectra of the coupling products between phosphoramidates and amino acid esters of a single stereochemical configuration indicate that little racemization occurs in the carbodiimide-driven couplings. A preference for homochiral dipeptides was observed in the more than twenty coupling reactions studied, and the diastereoselectivity was strongest when the amino acid linked to the nucleotide was either tryptophan or valine. It is hoped that our findings will spur a more detailed exploration of stereoselection processes that may have led to homochirality. The field is wide open for searching the vast structure space of peptido nucleotides, including searches with L-nucleotides and peptide lengths beyond the dipeptides studied here.

Acknowledgements

The authors thank D. Jovanovic, S. Motsch, and M. Räuchle for discussions. Support by Deutsche Forschungsgemeinschaft, project-ID 364653263 – TRR 235 and Volkswagen Foundation (Life? initiative, grant Az 92 768) is gratefully acknowledged. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: amino acids · peptides · prebiotic chemistry · stereoselectivity · RNA

[1] J. Podlech, *Angew. Chem. Int. Ed.* **1999**, *38*, 477–478; *Angew. Chem.* **1999**, *111*, 501–502.
[2] J. Podlech, T. Gehring, *Angew. Chem. Int. Ed.* **2005**, *44*, 5776–5777; *Angew. Chem.* **2005**, *117*, 5922–5924.
[3] C. De Duve, *Nature* **2005**, *433*, 581–582.
[4] Y. Chen, W. Ma, *PLoS Comput. Biol.* **2020**, *16*, e1007592.
[5] M. Frenkel-Pinter, M. Samanta, G. Ashkenasy, L. J. Leman, *Chem. Rev.* **2020**, *120*, 4707–4765.
[6] A. Brandenburg, *Origins Life Evol. Biospheres* **2019**, *49*, 49–60.
[7] G. Joyce, G. Visser, C. van Boeckel, J. van Boom, L. Orgel, J. van Westrenen, *Nature* **1984**, *310*, 602–604.
[8] I. A. Kozlov, S. Pitsch, L. E. Orgel, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 13448–13452.
[9] I. A. Kozlov, P. K. Politis, S. Pitsch, P. Herdewijn, L. E. Orgel, *J. Am. Chem. Soc.* **1999**, *121*, 1108–1109.
[10] M. Beier, F. Reck, T. Wagner, R. Krishnamurthy, A. Eschenmoser, *Science* **1999**, *283*, 699–703.
[11] V. K. Rajwanshi, A. E. Hakansson, M. D. Sorensen, S. Pitsch, S. K. Singh, R. Kumar, P. Nielsen, J. Wengel, *Angew. Chem. Int. Ed. Engl.* **2000**, *39*, 1656–1659.
[12] S. Toxvaerd, *J. Theor. Biol.* **2017**, *429*, 164–169.
[13] A. V. Tarasevych, A. E. Sorochinsky, V. P. Kukhar, J.-C. Guillemin, *Chem. Commun.* **2015**, *51*, 7054–7057.
[14] A. V. Tarasevych, A. E. Sorochinsky, V. P. Kukhar, A. Chollet, R. Daniellou, J.-C. Guillemin, *J. Org. Chem.* **2013**, *78*, 10530–10533.
[15] C. Blanco, D. Hochberg, *Phys. Chem. Chem. Phys.* **2012**, *14*, 2301–2311.
[16] D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, M. R. Ghadiri, *Nature* **1996**, *382*, 525–528.
[17] S. Yao, I. Ghosh, R. Zutshi, J. Chmielewski, *J. Am. Chem. Soc.* **1997**, *119*, 10559–10560.
[18] A. Saghatelian, Y. Yokobayashi, K. Soltani, M. R. Ghadiri, *Nature* **2001**, *409*, 797–801.
[19] H. F. Noller, *Cold Spring Harbor Perspect. Biol.* **2012**, *4*, a003681.
[20] N. S. M. D. Wickramasinghe, M. P. Staves, J. C. Lacey, *Biochemistry* **1991**, *30*, 2768–2772.
[21] J. C. Lacey, R. D. Thomas, M. P. Staves, C. L. Watkins, *Biochim. Biophys. Acta* **1991**, *1076*, 395–400.
[22] M. Jauker, H. Griesser, C. Richert, *Angew. Chem. Int. Ed.* **2015**, *54*, 14564–14569; *Angew. Chem.* **2015**, *127*, 14772–14777.
[23] H. Griesser, P. Tremmel, E. Kervio, C. Pfeffer, U. E. Steiner, C. Richert, *Angew. Chem. Int. Ed.* **2017**, *56*, 1219–1223; *Angew. Chem.* **2017**, *129*, 1239–1243.
[24] P. Tremmel, H. Griesser, U. E. Steiner, C. Richert, *Angew. Chem. Int. Ed.* **2019**, *58*, 13087–13092; *Angew. Chem.* **2019**, *131*, 13221–13226.
[25] H. Griesser, M. Bechthold, P. Tremmel, E. Kervio, C. Richert, *Angew. Chem. Int. Ed.* **2017**, *56*, 1224–1228; *Angew. Chem.* **2017**, *129*, 1224–1248.
[26] J. Balzarini, A. Karlsson, S. Aquaro, C.-F. Perno, D. Cahard, L. Naesens, E. De Clerq, C. McGuigan, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 7295–7299.
[27] B. Jash, C. Richert, *Chem. Sci.* **2020**, *11*, 3487–3494.
[28] S. Motsch, P. Tremmel, C. Richert, *Nucleic Acids Res.* **2020**, *48*, 1097–1107.
[29] S. Motsch, D. Pfeffer, C. Richert, *ChemBioChem* **2020**, *21*, 2013–2018.
[30] D. Jovanovic, P. Tremmel, P. S. Pallan, M. Egli, C. Richert, *Angew. Chem. Int. Ed.* **2020**, *59*, 20154–20160; *Angew. Chem.* **2020**, *132*, 20329–20335.
[31] D. X. Han, H. Y. Wang, Z. L. Ji, A. F. Hu, Y. F. Zhao, *J. Mol. Evol.* **2010**, *70*, 572–582.
[32] D. Beaufils, G. Danger, L. Boiteau, J.-C. Rossi, R. Pascal, *Chem. Commun.* **2014**, *50*, 3100–3102.
[33] H.-B. Bürgi, J. D. Dunitz, J.-M. Lehn, G. Wipff, *Tetrahedron* **1974**, *30*, 1563–1572.
[34] R. Breslow, Z.-L. Cheng, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9144–9146.
[35] K. Tamura, *Int. J. Mol. Sci.* **2011**, *12*, 4745–4757.

Manuscript received: May 7, 2021
Accepted manuscript online: July 22, 2021
Version of record online: August 21, 2021

