Oligopeptides and Copeptides of Homochiral Sequence, via β-Sheets, from Mixtures of Racemic α-Amino Acids, in a One-Pot Reaction in Water; Relevance to Biochirogenesis

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · AUGUST 2008

Impact Factor: 12.11 · DOI: 10.1021/ja709969v · Source: PubMed

6 AUTHORS, INCLUDING:

Gerard Bolbach
Pierre and Marie Curie University - Paris 6
138 PUBLICATIONS **2,071** CITATIONS
SEE PROFILE

Isabelle Weissbuch
Weizmann Institute of Science
113 PUBLICATIONS **3,885** CITATIONS
SEE PROFILE

Meir Lahav
Weizmann Institute of Science
298 PUBLICATIONS **9,685** CITATIONS
SEE PROFILE

J|A|C|S
ARTICLES

Published on Web 06/11/2008

Oligopeptides and Copeptides of Homochiral Sequence, *via* β-Sheets, from Mixtures of Racemic α-Amino Acids, in a One-Pot Reaction in Water; Relevance to Biochirogenesis

Roni A. Illos,[†] Fabricio R. Bisogno,[†] Gilles Clodic,[‡] Gerard Bolbach,[‡] Isabelle Weissbuch,[†] and Meir Lahav*[,†]

Department of Materials and Interfaces, The Weizmann Institute of Science, 76100-Rehovot, Israel and Plate-Forme Spectrométrie de Masse et Protéomique, Université Pierre et Marie Curie, 75252 Paris Cedex 05, Paris France

Received November 11, 2007; E-mail: meir.lahav@weizmann.ac.il

Abstract: As part of our studies on the biochirogenesis of peptides of homochiral sequence during early evolution, the formation of oligopeptides composed of 14−24 residues of the same handedness in the polymerization of DL-leucine (Leu), DL-phenylalanine (Phe), and DL-valine (Val) in aqueous solutions, by activation with *N,N′*-carbonyldiimidazole and then initiation with a primary amine, in a one-pot reaction, was demonstrated by MALDI-TOF MS using deuterium enantio-labeled α-amino acids. The formation of long isotactic peptides is rationalized by the following steps occurring in tandem: (i) creation of a library of short diasteroisomeric oligopeptides containing isotactic peptides in excess in comparison to a binomial kinetics, as a result of an asymmetric induction exerted by the N-terminal residue of a given handedness; (ii) precipitation of the less soluble racemic isotactic penta- and hexapeptides in the form of β-sheets that are delineated by homochiral rims; (iii) regio-enantiospecific chain elongation occurring heterogeneously at the β-sheets/solution interface. Polymerization of L-Leu with L-isoleucine (Ile) or L-Phe with L-^{1}*N*-Me-histidine yielded mixtures of copeptides containing both residues. In contrast, in the polymerization of the corresponding mixtures of L- + D-α-amino acids, the long oligopeptides were composed mainly from oligo-L-Leu and oligo-D-Ile in the first system and oligo-D-Phe in the second. Furthermore, in the polymerization of mixtures of hydrophobic racemic α-amino acids DL-Leu, DL-Val, and DL-Phe and with added racemic DL-alanine and DL-tyrosine, copeptides of homochiral sequences are most dominantly represented. Possible routes for a spontaneous "mirror-symmetry breaking" process of the racemic mixtures of homochiral peptides are presented.

Introduction

The emergence of the homochiral biopolymers, proteins, and nucleic acids, from the racemic prebiotic world is one of the unsolved enigmas in the field of origin of life and early evolution. Homochiral (isotactic) peptides and copeptides composed from residues of the same handedness had presumably played a paramount role as primeval catalysts.[1] Their formation, however, from racemic α-amino acids in water, in the absence of enzymes, still provides a myst,[2,3] since the polymerization of multicomponent systems such as mixtures of different racemic α-amino acids in an ideal solution should yield atactic peptide chains, where the α-amino acid residues of opposite handedness are randomly arranged. Consequently, elaboration of unique nonlinear reaction pathways[4,5] is required for the emergence of isotactic peptides from racemic amino acids.

A viable way to override the process of random polymerization is to perform the reaction at conditions departing from ideality that can be achieved by the formation of supramolecular architectures, either prior or in the course of the reaction.[6–11] Such architectures can serve as stereoselective templates for the formation of long peptides.

In 1957, Wald[12] proposed, based on the experimental results by Lundberg and Doty[13] and Idelson and Blout,[14] that a secondary α-helix structure of a polypeptide chain should exert a very efficient asymmetric induction and bias the selection of

[†] The Weizmann Institute of Science.
[‡] Université Pierre et Marie Curie.

(1) Brack, A. *Chem. Biodiversity* **2007**, *4*, 665–679.
(2) Franck, P.; Bonner, W. A.; Zare, R. N. In *Chemistry for the 21th Century*; Keinan, E., Schecter, I., Eds.; Wiley-VCH: Weinheim, 2000; pp 175−208..
(3) Weissbuch, I.; Leiserowitz, L.; Lahav, M. In *Topics in Current Chemistry*; Walde, P., Ed.; Springer-Verlag: 2005; Vol. 259, pp 123−165..
(4) Girard, C.; Kagan, H. B. *Angew. Chem., Int. Ed.* **1998**, *37*, 2922–2959.
(5) Soai, K.; Kawasaki, T. In *Topics in Current Chemistry*; Springer-Verlag: DOI 10.1007/128_2007_138.
(6) Zepik, H.; Shavit, E.; Tang, M.; Jensen, T. R.; Kjaer, K.; Bolbach, G.; Leiserowitz, L.; Weissbuch, I.; Lahav, M. *Science* **2002**, *295*, 1266–1269.
(7) Nery, J. G.; Bolbach, G.; Weissbuch, I.; Lahav, M. *Angew. Chem., Int. Ed.* **2003**, *42*, 2157–2161.
(8) Nery, J. G.; Bolbach, G.; Weissbuch, I.; Lahav, M. *Chem.—Eur. J.* **2005**, *11*, 3039–3048.
(9) Nery, G. J.; Eliash, R.; Bolbach, G.; Weissbuch, I.; Lahav, M. *Chirality* **2007**, *19*, 612–624.
(10) Eliash, R.; Nery, J. G.; Rubinstein, I.; Clodic, G.; Bolbach, G.; Weissbuch, I.; Lahav, M. *Chem.—Eur. J.* **2007**, *13*, 10140–10151.
(11) Rubinstein, I.; Eliash, R.; Bolbach, G.; Weissbuch, I.; Lahav, M. *Angew. Chem., Int. Ed.* **2007**, *46*, 3710–3713.

homochiral amino acids in peptide chain elongation in the later stages of the propagation reaction. Blair and Bonner[15] have claimed to have supported this hypothesis experimentally in the polymerization of nonracemic Leu-*N*-carboxyanhydride (Leu-NCA) which yields an α-helix in an organic solvent, where they found that the enantiomeric excess (ee) of the residues in the formed peptides (45.4%) was enhanced in comparison with the that of starting monomer (31.1%). However, polymerization of nonracemic valineNCA (ValNCA) in organic solvent yielded peptides with a depletion of the major enantiomer in the chains. Since valine peptides self-assemble as β-sheets, these results led Bonner et al. to assume that the Wald's model is limited to α-helices alone.[16] On the other hand, Brack and Spach[17] proposed that β-sheets composed from alternating hydrophilic and hydrophobic residues could be also implicated in the increase of the ee of the residues composing these peptides. All the above experiments, however, led only to partial enrichment of the peptides but did not provide a scenario for the formation of long isotactic peptides and copeptides from mixtures of racemic α-amino acids.

The Ehler−Orgel reaction[18–20] on the polymerization of α-amino acids in aqueous solutions, activated *in situ*, by solid *N*,*N*′-carbonyldiimidazole (CDI), is considered a plausible model system for obtaining primeval peptides.[21] Recently, Luisi et al.[22,23] reported the polymerization of racemic tryptophan (Trp), leucine (Leu), and isoleucine (Ile) via the above reaction, in buffered solutions, to yield libraries of short oligopeptides in the range of 6−10 residues, where the isotactic peptides are formed as minor diastereomers, albeit in amounts larger than those predicted by a binomial distribution. These results were rationalized in accordance to a kinetic Markov mechanism in which the homochiral residues at the N-terminus of the peptide exert an asymmetric induction in the chain elongation. The process was recently simulated by a mathematical model based on a first-order Markov kinetics.[24] Such a reaction pathway, however, will lead to a library of diastereoisomeric peptides where the longer isotactic ones are only slightly overexpressed beyond the binomial distribution.

We reported recently a very different distribution of diastereoisomeric peptides in the polymerization of aqueous solutions of racemic LeuNCA (prepared separately in an organic solvent), where long homochiral oligopetides of 25 detectable residues were obtained as the most dominant diastereoisomers, when the reaction was initiated by primary amines.[11] We demonstrated that racemic β-sheets are formed in the course of the reaction and operate as intermediate templates in the stereoselective process of chain elongation.

(12) Wald, G. *Ann. N.Y. Acad. Sci.* **1957**, *69*, 352.
(13) Lundberg, R. D.; Doty, P. *J. Am. Chem. Soc.* **1957**, *79*, 3961–3972.
(14) Idelson, M.; Blout, E. R. *J. Am. Chem. Soc.* **1957**, *79*, 3948–3955.
(15) Blair, N. E.; Bonner, W. A. *Origin Life Evol. Biosphere* **1980**, *10*, 255–263.
(16) However, recent MALDI-TOF MS studies of the oligopeptides obtained from racemic LeuNCA or ValNCA in organic solvent resulted in the formation of atactic polymers; see ref 11.
(17) Brack, A.; Spach, G. *J. Mol. Evol.* **1979**, *13*, 35–46.
(18) Ehler, K. W.; Orgel, L. E. *Biochim. Biophys. Acta* **1976**, *434*, 233–243.
(19) Ehler, K. W. *J. Org. Chem.* **1976**, *41*, 3041–3042.
(20) Hill, A. R.; Orgel, L. E. *Orig. Life Evol. Biosph.* **1996**, *26*, 539–545.
(21) Brack, A. *Origins Life Evol. Biosphere* **1987**, *17*, 367–379.
(22) Blocher, M.; Hitz, T.; Luisi, P. L. *Helv. Chim. Acta* **2001**, *84*, 842–847.
(23) Hitz, T.; Blocher, M.; Walde, P.; Luisi, P. L. *Macromolecules* **2001**, *34*, 2443–2449.
(24) Wattis, J. A.; Coveney, P. V. *J. Phys. Chem. B* **2007**, *111*, 9546–9562.

In order to understand the differences between the two processes it was most challenging to elaborate experimental conditions for the Orgel−Ehler reaction, where long isotactic peptides and copeptides can be formed, via the operation of β-sheet templates. Here we report that, with the use of primary amines as initiators, we can generate peptides and copeptides of a homochiral sequence (isotactic) as the most dominant diastereoisomers.

Experimental Section

Deuterated L-α-amino acids, leucine L(d10)-Leu, valine L(d8)-Val, phenylalanine L(d5)-Phe, alanine L(d4)-Ala, tyrosine L(d4)-Tyr, tryptophan L(d4)-Trp, were purchased from Cambridge Isotope Laboratories USA, and the corresponding D-α-amino acids and the enantiopure L- or D-α-amino acid-methyl ester monohydrochlorides used as initiators, from Sigma-Aldrich. The latter were converted into the corresponding free amines using NH_3 gas in chloroform.

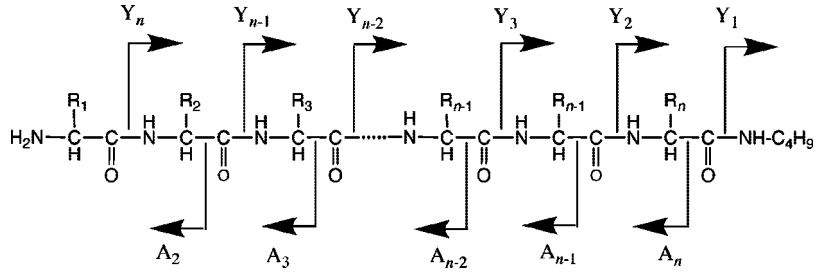
Polymerization Reaction. Aqueous solutions of racemic α-amino acids or their mixtures, DL(d10)-Leu, 40 mM, DL(d8)-Val, 40 mM, DL(d5)-Phe 80 mM, DL(d4)-Tyr 6 mM, DL(d4)-Ala 20 mM, in which the L-enantiomer was enantioselectively labeled with deuterium, were reacted in an ice-cooled bath with solid *N*,*N*′-carbonyldiimidazol (2.5 mol/mol Leu, Val, Trp and 10 mol/mol Phe, Tyr, Ala) for 2 min. Then, 0.25−0.5 mol/mol *n*-butylamine, glycine-ethyl ester, or methyl ester of L-Leu (L-Leu-OMe) was added to initiate the polymerization reaction that proceeded at room temperature with vigorous stirring. X-ray diffraction was measured in situ in a capillary. After 24 h, the reaction mixture was directly freeze-dried under vacuum or after centrifugation, decantation and several washings with water of the precipitate. The diastereoisomeric distribution of the oligopeptides was determined by matrix-assisted laser-desorption-ionization-time-of-flight mass spectrometry (MALDI-TOF MS) analysis. FTIR spectra of the solid products were measured in KBr pellets and the X-ray powder diffraction on the powders deposited on a horizontal holder.

MALDI-TOF MS and MS/MS Analysis, Sample and Matrix Preparation. Representative samples were measured in the analytical laboratories at the Weizmann Institute as well as in Paris VI. The MALDI-TOF MS analysis was performed on a Bruker Reflex II MALDI-TOF mass spectrometer (Bruker, Bremen, Germany), equipped with a 337 nm nitrogen laser and with the SCOUT source (delayed extraction and reflector). Each mass spectrum was generated from the signal average of 300 laser shots. Oligo-Val, oligo-Leu, and oligo-Trp products were dissolved in 20 μL of trifluoroacetic acid (TFA), and the solution was diluted with 80 μL of THF. The isotactic oligo-Phe products beyond the dimers are not soluble in organic solvents; for this reason they were reacted with trifluoroacetic anhydride (TFAA) according to the following protocol: 1 mg of the dried solid products were weighed in a 1.5 mL Eppendorf and then reacted with a mixture of TFAA and THF (3:1 v/v) to yield the *N*-trifluoroacetyl derivatives of the corresponding oligopeptides that are soluble. The obtained completely clear solution was used for the MALDI-TOF MS sample preparation.

The matrix was prepared in the following way: 6 mg of dithranol were dissolved in 125 μL of chloroform, and to the obtained solution 125 μL of a solution prepared from 17 mg of NaI dissolved in 1 mL of THF were added. The final matrix solution was vortexed for 1 min at high speed. The best preparation for the MALDI-TOF MS analysis was achieved by the double deposit procedure: 1 μL of the matrix solution was deposited on the target holder, and then 1 μL of the sample solution was deposited on the matrix layer.

The MS/MS spectra were obtained in MALDI-TOF-TOF (4700 Proteomics Analyzer, Applied Biosystems, US) using an NdYag laser (355 nm, 200 Hz) and 1 keV collision energy (gas N_2 at 5×10^{-7} torr). For the different oligopeptides MS/MS spectra needed to average over a few tens of thousands laser shots. The objective of the MALDI-TOF MS/MS study was to deduce some sequence

Scheme 1. Nomenclature Used for Ion Fragmentation Observed for the Oligopeptides Na^+ Cationized Ions[a]



[a] They are referred to as *Y*-type and *A*-type series. Note *Y*1 fragment starts after the initiator at the C-terminus.

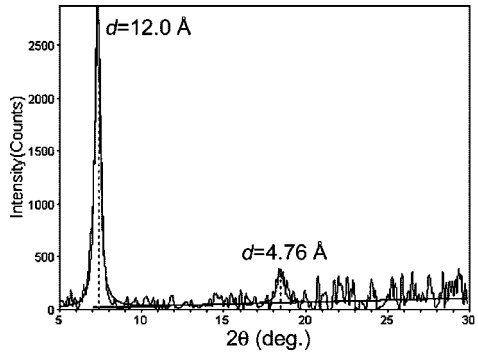


Figure 1. X-ray diffraction pattern, measured in a capillary, of the colloidal phase obtained in the polymerization of DL-Leu (40 mM), activated, in an ice bath, with solid CDI (100 mM), with 50% mol/mol *n*-butylamine initiator, after 3 h of vigorous stirring at room temperature. Note that the pattern is after background subtraction so as to remove the strong contribution of water scattering at $2\theta > 15°$.

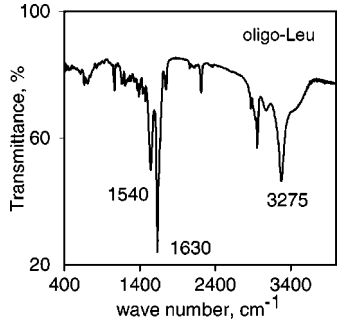


Figure 2. FTIR spectrum of the water-insoluble fraction obtained in the polymerization of DL-Leu with L-Leu-OMe initiator.

information from the spectra taking into account that isotactic and atactic oligopeptides are analyzed. It is well-known that changing only one L- or D-residue by a D- or L- residue results in different fragment ion abundances for protonated peptides. In the present study, working with Na^+ cationized peptides, the systematic analysis of hundreds of different samples containing D- (protonated) residues and L- (deuterated) residues with different initiators lead to the conclusion that, for both fragmentation series (Scheme 1), there is no systematic and significant variation of the ion abundance with the type L (deuterated) or D (nondeuterated) of the residue as described in the Supporting Information available (ESI).

Results

Polymerization of Racemic α-Amino Acids. Aqueous solutions of racemic α-amino acids Leu, Val, or phenylalanine (Phe), enantioselectively tagged with deuterium atoms, cooled in an ice bath, were activated with an excess of solid CDI and polymerized either in the absence or in the presence of amines as initiators added after the activation. The enantioselective

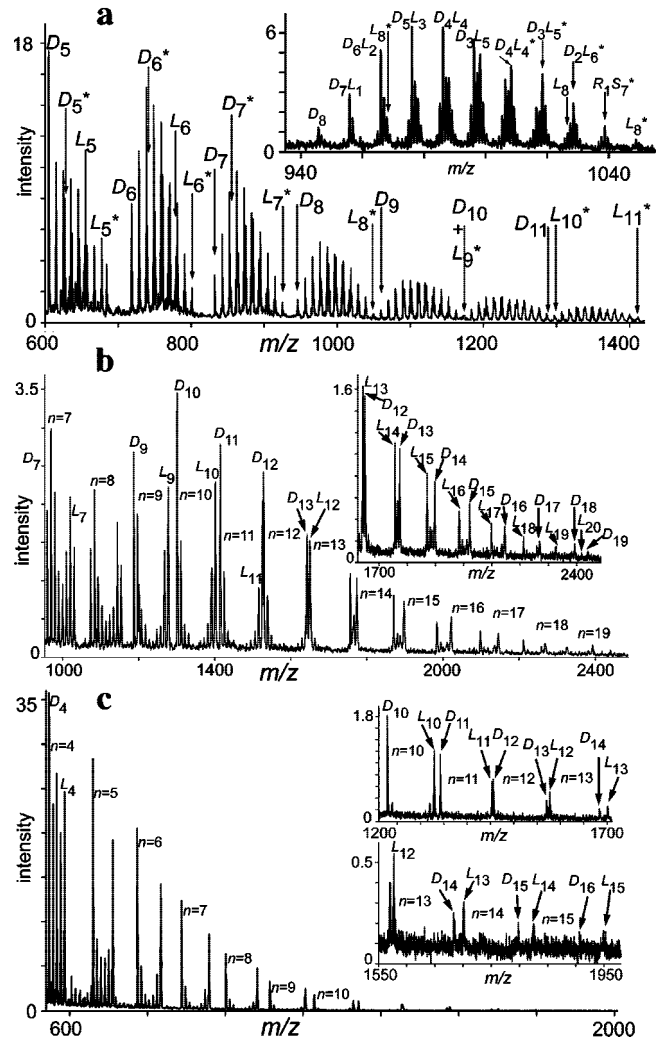


Figure 3. MALDI-TOF spectra of $(M + Na)^+$ cationized oligopeptides obtained by polymerization in water of DL(d10)-Leu (40 mM) activated with solid CDI: (a) without initiator and (b, c) initiated with (50% mol/mol) L-Leu-OMe and *n*-butylamine, respectively. The total number of residues in oligopeptides, *n* (except the L-Leu-OMe initiator linked at C-terminus) is shown. Note in (a), signals corresponding to $(M + 2Na - 1)^+$ cationized oligopeptides labeled with an asterisk. Note that inserts in (b) and (c) show the region where the strongest signal D_n and L_n represent the long isotactic oligopeptides.

labeling permits differentiation between diastereoisomeric oligopeptides in the mass spectrometry analysis.[25] As initiators we used *n*-butylamine or enantiopure α-amino acid-methyl esters, and the polymerization was performed with vigorous stirring at room temperature. The reaction started as a homogeneous solution, a foam appeared after ~10 min, and ~30−60

Table 1. Sequence Analysis of the Oligo-Leu Products Obtained in the Polymerization of DL(d10)-Leu with *n*-Butylamine

		Y-series	*A*-series
$n = 7$	D_6L_1	L-Leu(d10) at N-term	L-Leu(d10) at C-term
	D_1L_6	D-Leu(d0) at N-term	D-Leu(d0) at N-term
$n = 8$	D_7L_1	L-Leu(d10) at N-term	L-Leu(d10) at C-term
	D_1L_7	D-Leu(d0) at N-term	D-Leu(d0) at N-term
$n = 9$	D_8L_1	L-Leu(d10) at N-term	L-Leu(d10) at C-term
	D_1L_8	D-Leu(d0) at N-term	D-Leu(d0) at N-term
$n = 10$	D_9L_1	too weak abundance	too weak abundance
	D_1L_9	D- and L-Leu at N-term	D-Leu(d0) at N-term
$n = 11$	$D_{10}L_1$	too weak abundance	too weak abundance
	D_1L_{10}	D- and L-Leu at N-term	D- and L-Leu at N-term

min later the solution became turbid due to the appearance of a colloidal phase. An X-ray diffraction pattern of this colloidal phase inserted in a capillary displays Bragg peaks at *d*-spacings of 12.0 and 4.76 Å for oligo-Leu, Figure 1.

The pH of the solution measured during the reaction ranged between 7.5 and 7.7. After 24 h, the reaction mixture was dried by lyophilization either directly or after centrifugation, supernatant removal, and washing. The X-ray powder diffraction patterns measured for the dried water-insoluble fraction of oligo-Leu (see Supporting Information Figure S2c) was slightly different from that measured *in situ*, Figure 2, presumably due to the aqueous environment. FTIR spectra display strong C═O stretching (amide I) and N—H bending (amide II) vibrations at 1630 cm^{-1} and 1540 cm^{-1}, Figure 2 (for oligo-Phe, see Supporting Information Figure S2a, d). The results obtained with both methods are in keeping with the formation of β-sheets.

The diastereisomeric composition of the oligopeptides was determined by matrix-assisted laser-desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS). For example, mass spectra of samples obtained from the polymerization of DL(d10)-Leu initiated with L-Leu-OMe and in the absence of initiator are shown in Figure 3a,b. The signals corresponding to the oligopeptides of homochiral sequence (isotactic), labeled D_n or L_n where n is the total number of repeating units, appear at the wings of the sets of diastereisomers of each length, whereas those corresponding to heterochiral oligopeptides D_hL_d ($n = h + d$) appear in between.

When the polymerization was performed in the absence of the initiator, a CO_2^- group is present at the C-terminus of the chains. The signals appear at the *m/z* values $(M + Na)^+$ and $(M + 2Na - 1)^+$ where M is the molecular weight of the peptide chain, Figure 3a. In the experiments performed in the presence of initiator, the latter is attached at the C-terminus and *m/z* of $(M + Na)^+$ where M includes also the mass of the initiator, Figure 3b,c. Oligopeptides initiated by free amino acids were not detected.

An inspection of the spectrum of the oligopeptides obtained in the absence of the initiator (Figure 3a) shows a remarkable difference in the distribution of the various diastereoisomers as compared to those obtained with initiator, Figure 3b, c. In the absence of initiator, we obtained mixtures of oligopeptides containing up to 11 Leu residues, and even longer (not shown), where the heterochiral peptides are by far most abundant, although the isotactic ones are overexpressed as compared to a binomial distribution. Similar results were obtained in the polymerization of racemic Phe (see Supporting Information, Figure S3).

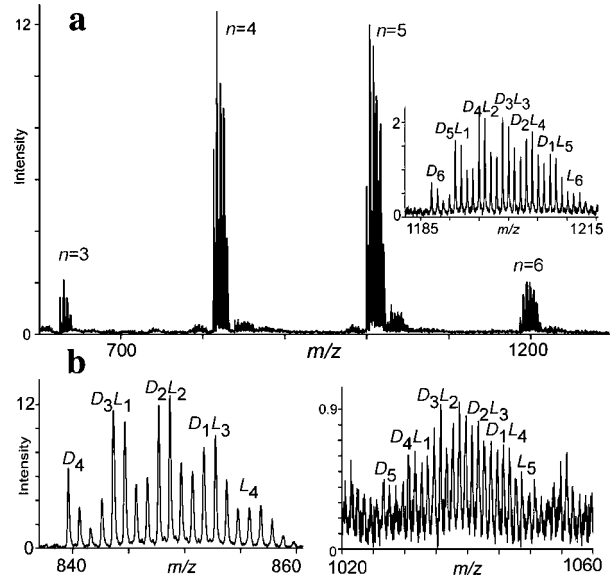


Figure 4. MALDI-TOF spectra of the $(M + Na)^+$ cationized oligopeptides obtained by polymerization in water of DL(d5)-Trp (40 mM) activated with solid CDI: (a) in the absence and (b) in the presence of *n*-butylamine initiator where only the tetra- and pentapeptide regions are shown. Note that in (a) the mixture of the oligopeptides was ethylated at the C-terminus to reduce the number of signals since for CO_2^- terminated peptides; signals for both $(M + Na)^+$ and $(M + 2Na - 1)^+$ cationized molecules are observed.

By contrast, when the polymerization was initiated with either L-Leu-OMe (Figure 3b) or *n*-butylamine (Figure 3c), isotactic oligopeptides of up to 17—21 D- or L-residues could be regularly detected as the dominant diastereoisomers. In a few experiments, isotactic peptides containing 24—30 residues were observed. A significant result is that whereas for the hepta- and octapeptides the heterochiral diastereomers are still formed in relatively high concentrations, beyond decapeptides the homochiral diastereoisomers appear as the dominant fraction. Other examples of the mass spectra of samples obtained from the polymerization of DL(d5)-Phe or DL(d8)-Val with various initiators are shown in the Supporting Information, Figures S4—6.

MALDI-TOF-TOF (MS/MS) Sequence Analysis of the $D_{n-1}L_1$ or D_1L_{n-1} Diastereoisomers. Among the diastereoisomeric long peptides obtained in the presence of initiators we found always that the major fraction of isotactic peptides is accompanied by a minor fraction of those bearing one residue of opposite handedness. According to MALDI-TOF MS analysis of the oligopeptides obtained from DL(d10)-Leu initiated by *n*-butylamine, Figure 3c, the ratio between signal intensities of homochiral D_n or L_n and heterochiral $D_{n-1}L_1$ or D_1L_{n-1} (i.e., containing a single unit of opposite handedness) oligopeptides increases with the increase in chain length, suggesting that the elongation of the homochiral chains is more probable than the termination by incorporation of a repeating residue of opposite handedness. This suggestion was independently confirmed by MALDI-TOF-TOF (MS/MS) analysis.

The analysis of fragmentation pattern in the MS/MS spectrum for all the different oligopeptides showed the two main series of Na^+ cationized fragments independent of the nature of the initiator. A first series named *Y* (Scheme 1) corresponds to the classical *y*-series for protonated peptides[26,27] with ($Y = y +$

(25) Addadi, L.; Gati, E.; Lahav, M.; Leiserowitz, L. *Israel J. Chem.* **1977**, *15*, 116.

(26) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.
(27) Johnson, R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T. *Anal. Chem.* **1987**, *59*, 2621–2625.

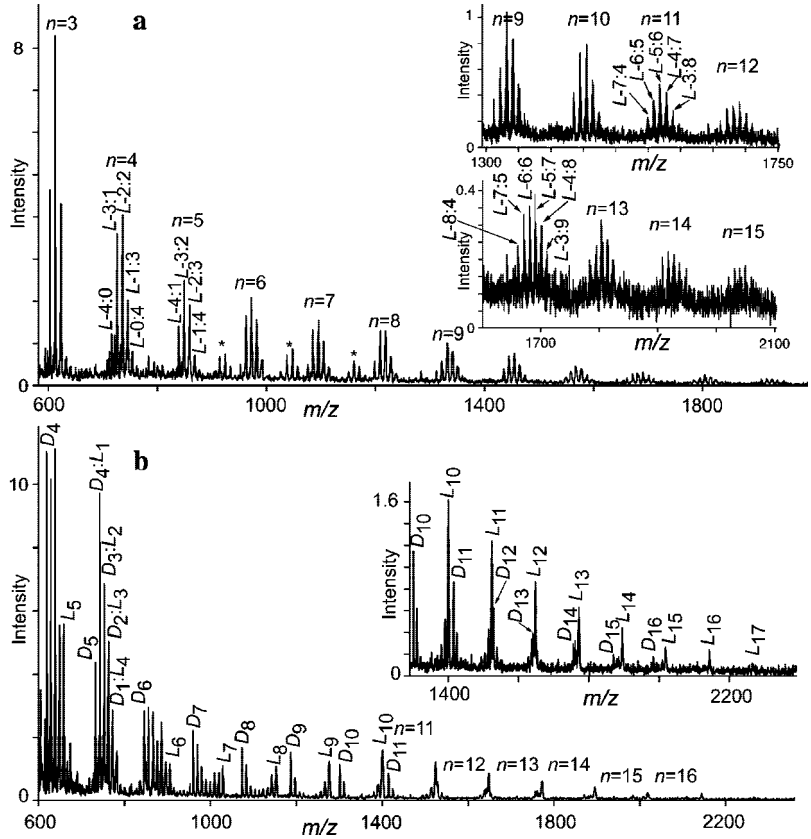


Figure 5. MALDI-TOF MS of the $(M + Na)^+$ cationized oligopeptides obtained from the polymerization in water of mixtures of α-amino-acids: (a) L-Ile + L-Leu(d10), (b) D-Ile + L-Leu(d10), with 25%mol/mol L-Leu-OMe initiator. Note that in (a) the *m/z* values correspond to the *N*-trifluoroacetyl (CF_3-CO-) derivatized N-terminus of the Ile/Leu copeptides. The signals corresponding to diastereomers of each length *n* are labeled. In (a) individual copeptides, of total length $n = a + b$, are labeled L-*a*:*b*, where *a*:*b* are the number of Ile and Leu residues, respectively. In (b), single component isotactic peptides are labeled L_n or D_n and heterochiral copeptides are labeled D_a:L_b, where $n = a + b$. Some partial fragmentation observed in (a) was labeled with asterisks.

Na−H) associated with the cleavage of the peptide bond. The second series named *A*, corresponds to the *a*-series for protonated peptides[26,27] ($A = a + Na-H$) previously observed for Na^+ cationized peptides[28,29] associated with the cleavage of the $C^{\alpha}-CO$ bond. Systematically, the second series was found 2-fold higher in abundance than the first one.

The analysis of the fragment *An*-1 abundances with D(d0)-Leu and L(d10)-Leu at the C-terminus, respectively, allows us to determine if some sequences are overexpressed or not and similarly for the abundance of *Yn* for D(d0)-Leu and L(d10)-Leu at the N-terminus, respectively. The other fragments *A*i and *Y*i are also used to obtain more details on the overexpressed/underexpressed sequences. Table 1 contains a summary of this sequence analysis of the oligo-Leu products obtained in the polymerization of DL(d10)-Leu with *n*-butylamine and shows that, for oligopeptides of $n \geq 7$, the overexpressed sequences found by both series of fragmentations lead to the same conclusion: the single residue of opposite handedness is located essentially at either N- or C- terminus.

Moreover, when the polymerization was initiated by L-Leu-OMe, the MALDI-TOF-MS/MS sequence analysis showed that in the D_1L_{n-1} chains the single D-residue is located at the N-terminus, whereas in the $D_{n-1}L_1$ chains the single L-residue is located either at the N-terminus, according to the *Y*-type of fragmentations, or adjacent to the L-initiator at the C-terminus, according to the *A*-type of fragmentations. Therefore, in addition to an asymmetric induction exerted by the enantiopure initiator in the early stages, since the D_1L_{n-1} and $D_{n-1}L_1$ chains were formed in smaller concentrations than the corresponding isotactic oligopeptides, we conclude that the elongation of the latter was favored and chain termination occurred by incorporation of a residue of opposite handedness.

Proposed Polymerization Route. On the basis of these results we propose the following reaction pathway, Scheme 2:

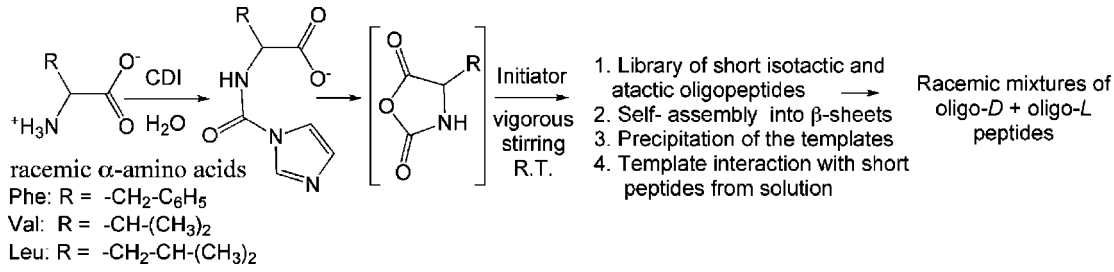
(i) The reaction of racemic Phe, Leu, or Val with solid CDI in an ice bath produces *N*-(imidazolyl-1-carbonyl)amino acid (ICA) in ∼95%, 82%, and 72% yields, respectively, as determined by NMR.

(ii) After the addition of the initiator, the progress of the polymerization reaction, presumably via the *N*-carboxyanhydride (NCA) derivative[21] as the transient reactive intermediate, that could not be detected, was monitored by NMR (see Supporting Information Figure S7).[30] The rate of the disappearance of ICA was much faster when the reaction was performed in the presence of the initiator than in its absence. Complete disappearance of ICA occurred after ∼5 h with *n*-butylamine as initiator and ∼24 h without initiator (Supporting Information Figures S8,9).

(iii) Once short peptides are formed, the isotactic diastereisomers, perhaps together with those containing a single heterochiral residue, self-assemble into β-sheet aggregates that

(28) Barr, J. M.; Stipdonk, J. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 566–578.
(29) Sabareesh, V.; Balaram, P. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 618–628.
(30) Note that the chemical shifts of the H linked to the asymmetric C-carbon atom for ICA and NCA derivatives of Leu and Val were found different and PheNCA is not water soluble.

Scheme 2. Polymerization Pathway



precipitate and serve as templates for the chain elongation process proceeding heterogeneously at the peptide template/liquid interface.[11] X-ray diffraction and FTIR measurements (Figures 1, 2 and Supporting Information Figure S2) support the formation of the β-sheets. However, these methods do not allow us to differentiate between the antiparallel or parallel β-sheets, be they racemic (rippled) (composed from alternating isotactic L- and D-chains)[31] or pleated (composed from chains of the same handedness). Nevertheless, homochiral rims, where the growing NH_2 groups are located, delineate all of these types of β-sheets. Such rims engender asymmetric induction in the ensuing steps of chain elongation.

In order to provide additional and independent support for the formation of the β-sheet templates, we performed the polymerization of DL(d5)-Trp under the same conditions. We expected that the bulky indole group of Trp should prevent the self-assembly of the short oligo-Trp into β-sheets. Consequently, long isotactic oligo-Trp was not anticipated to be formed. Indeed, polymerization of racemic Trp, in either the presence or absence of initiator (Figure 4), yielded up to penta- and hexapeptides only, with a diastereomeric distribution very similar to that reported by Luisi et al.,[22] where the fraction of the isotactic ones departs slightly from a binomial distribution following a previously suggested Markov kinetics. The FTIR spectra of the oligo-Trp show strong peaks at 1647 cm^{-1} and 1515 cm^{-1}, and the X-ray powder pattern shows a *d*-spacing of 4.9 Å characteristic of hydrogen bonding and a very broad peak at *d*-spacing of 8.9 Å, which are different from those obtained from Leu and Phe (see Supporting Information Figure S2b,d).

Enantio-Selection by β-Sheet Templates for Different α-Amino Acids. In order to test the feasibility of generating isotactic copeptides via the β-sheet templates composed from hydrophobic residues, we investigated the enantioselectivity in the polymerization of mixtures of various α-amino acids. Thus, for example, polymerization of 1:1 mixtures of L(d10)-Leu with either L- or D-isoleucine (Ile) was studied. The MALDI-TOF mass spectrum of oligo-copeptides obtained from L-Ile with L(d10)-Leu, Figure 5a, indicates a random distribution of Ile and Leu residues, and oligopeptides of a single component were not detected beyond pentamers.

By contrast, polymerization of D-Ile with L(d10)-Leu yielded a significantly different product distribution, Figure 5b. As short peptides, we obtained mixtures of oligo-D-Ile (labeled D_n) and oligo-L(d10)-Leu (labeled L_n) as well as copeptides (labeled D_a:L_b, where a and b are the number of Ile and Leu residues, respectively and $n = a + b$), whereas beyond the octamers, the single component chains are almost the only detectable products. Thus, we obtained oligo-D-Ile and oligo-L(d10)-Leu containing up to 16 detectable residues.

Similar results were obtained in the polymerization of hydrophobic L-(or D)-Phe with the hydrophilic L-N^1-Methyl-Histidine[32] (M^1His), where oligopeptides up to 10 residues were detected.[33] In the polymerization of the L-Phe with L-M^1His, Figure 6a, we obtained oligo-L-Phe, but not oligo-L-M^1His, as well as copeptides of the two L-components comprising up to six residue of L-M^1His (labeled L-*a*:*b*, where *a*, *b* are the number of Phe and M^1His repeating units, respectively). On the other hand, when D-Phe was polymerized with L-M^1His, the MALDI-TOF spectra (Figure 6b) demonstrate that only oligo-D-Phe and sometimes those containing a single L-M^1His residue were formed. Similar results were obtained with Phe/M^3His and Phe/Thienyl-alanine mixtures of α-amino acids (not shown).

These results show the high enantioselectivity of the polymerization reaction which are in keeping with the formation of the short peptides that self-assemble in the form of β-sheet templates composed from strands of a homochiral sequence.

Copeptides from Mixtures of Two or Three Racemic α-Amino Acids. Polymerizations of 1:1 mixtures of hydrophobic DL(d10)-Leu+DL(d8)-Val, DL(d5)-Phe+DL(d10)-Leu, and DL(d5)-Phe+ DL(d8)-Val in aqueous solutions, where the L-amino acids were tagged with deuterium atoms and initiated with *n*-butylamine, were performed. The MALDI-TOF MS of the oligopeptides are shown in Figure 7a−c. The dominant signals are those that correspond to oligopeptides and oligo-copeptides composed from residues of the same handedness. Polymerization of ternary 1:1:1 mixtures of DL(d10)-Leu + DL(d8)-Val + DL(d5)-Phe yielded isotactic peptides accompanied by two and three component isotactic copeptides composed from residues of the same handedness as the strongest signals, listed in Table 2.

Similar results were obtained also in the polymerization of DLPhe with the more hydrophilic amino acids DL(d4)-Ala and DL(d4)-Tyr. Tyrosine is scarcely water-soluble; therefore, we used in the reaction a ratio of 3.3:1 Phe/Tyr. The MALDI-TOF mass spectra (Figure 8a,b) show, in addition to oligo-Phe of a homochiral sequence, the formation of copeptides containing a single homochiral Tyr or Ala residue in any chain up to a total of seven or ten residues, respectively. The oligo-Tyr and oligo-Ala were not detected. The enantioselectivity of the copeptides was not affected, and those of homochiral sequence (labeled D-*n*:1 and L-*n*:1) are the dominant products. Figure 8c shows the MALDI-TOF MS of the oligopeptides obtained from the polymerization of racemic Phe/Ala/Tyr where again the isotactic peptides and copeptides are the dominant diastereoisomers.

(31) Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1953**, *39*, 253–256.

(32) Lamy, C.; Lemoine, J.; Bouchu, D.; Goekjian, P.; Strazewski, P. *ChemBioChem* **2008**, *9*, 710–713.

(33) We used *N*-methyl-histidines since histidine does not yield NCAs and does not polymerize at these conditions; ref 32.

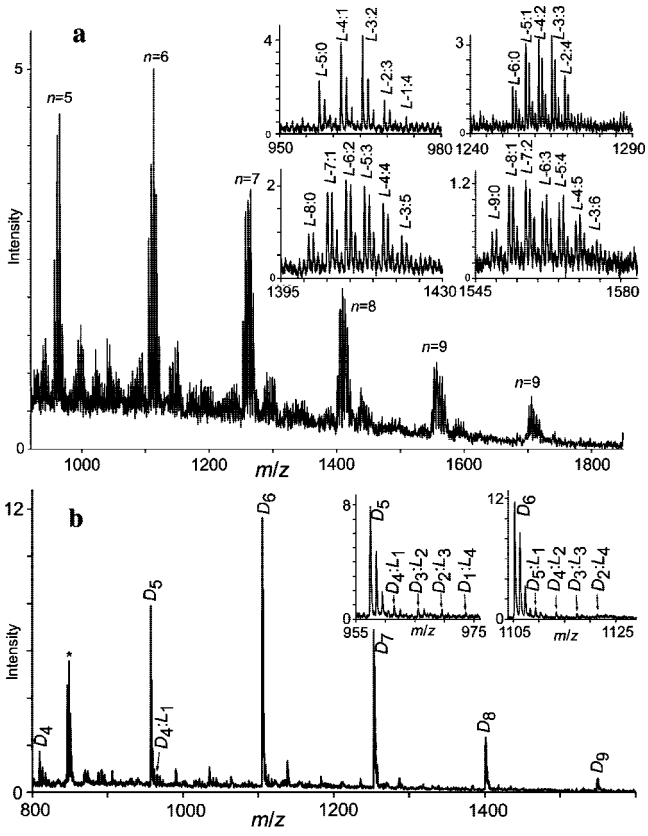


Figure 6. MALDI-TOF spectra of the $(M + Na)^+$ cationized oligopeptides obtained MALDI-TOF Ms spectra of the $(M + Na)^+$ cationized oligopeptides obtained from the polymerization in water, with 25% mol/mol glycine-methyl ester initiator, of mixtures of α-amino acids: (a) L-Phe + L-N^1-Methyl-His (M^1His) and (b) D-Phe + L-M^1His. In (a) individual copeptides, of total length $n = a + b$, are labeled L-a:b, where a:b are the number of Phe and M^1His residues, respectively. In (b), single component isotactic peptides are labeled L_n or D_n and heterochiral copeptides are labeled D_a:L_b, where $n = a + b$. Note that the signals correspond to m/z values of the trifluoroacetyl derivatized N-terminus of the peptides.

Finally, the polymerization of quaternary mixtures of racemic Phe/Ala/Tyr/Val yielded homochiral copeptides comprising three and four types of amino acid residues shown in Supporting Information Figure S10.

Discussion

The formation of isotactic peptides and copeptides as the dominant diastereoisomers, in a one-pot reaction, from mixtures of amphiphilic racemic amino acids in aqueous solutions was demonstrated. The addition of primary amines as initiators was indispensable for the formation of the β-sheet templates that exert regio-enantiospecific control in the formation of the long isotactic oligopeptides.

In the absence of initiator the polymerization is initiated by a free α-amino acid[34] that was not converted to ICA or as a result of partial hydrolysis of the NCA. At pH 7.5 of the reaction, the amino acid appears in its zwitterionic form, with only a very minor fraction 0.8% as free amine, as calculated with the Henderson−Hasselbach equation. Similarly, at this pH, the concentration of the free NH_2 groups of the short zwitterionic peptides needed for chain elongation is also very low. Consequently, in the absence of the added initiators, the concentration of the isotactic penta- and hexapeptides, needed for the formation of the β-sheet templates, is very low. As a result, the polymerization at these conditions proceeds under a regime of Markov kinetics to yield a complex library of diastereoisomeric peptides, where the isotactic peptides are formed in concentrations slightly higher than those anticipated for a binomial distribution.[35]

By contrast, in the presence of primary amines as strong nucleophiles, the initiation is much faster resulting in the formation of short isotactic peptides in sufficient concentrations for their self-assembly into β-sheet templates, already in the early stages of the reaction. Since the initiator is linked covalently at the C-terminus as an amide group, the free NH_2 end group of the oligopeptides operates as an efficient catalyst for chain elongation. Once the β-sheets are formed, they separate from the solution in a colloidal-like form, since the hydrophobic isotactic penta- and hexapeptides are less soluble than the corresponding heterochiral diastereisomers, as demonstrated in the polymerization of pure LeuNCA and ValNCA.[11] Consequently, the chain elongation process takes place under heterogeneous conditions at the interface between the colloidal particles and the monomer in solution.

The formation of the β-sheets as stereoselective templates is in keeping with the FTIR and X-ray data as well as with observed increase in the concentration of the isotactic peptides with the increase of their chain length by MALDI-TOF MS. These templates can also insert a limited number of hydrophilic α-amino acid residues in the peptide chains. The formation of long isotactic peptides implies that the templates reduce the energy of activation in the propagation step, with very low probability for chain termination by heterochiral residues. Furthermore, when β-sheets could not be formed as, in the case of Trp, due to steric hindrance of the indole groups, long isotactic peptides were not detected.

The type of β-sheets could not be determined from the FTIR and X-ray powder diffraction experiments. However, in previous studies on the enantiomeric excess of the isotactic oligopeptides obtained in the polymerization of crystals of DLPheNCA[8,9] and in aqueous solutions of DLLeuNCA or DLValNCA[11] as initiated by enantiopure esters of α-amino acids, we could assign the structures of the templates as antiparallel racemic β-sheets, composed from alternating oligo-D and oligo-L chains. The asymmetric induction is exerted the by homochiral rims that delineate these sheets. For antiparallel β-sheets to grow in a stereospecific fashion, and to preserve the structure of the growing site, the once formed sheets must interact with short peptides simultaneously with their process of crystal growing. In the one-pot polymerization of the α-amino acids, as a result of a different kinetics in comparison to the polymerization of the pure α-amino acid NCAs, it was not possible to differentiate experimentally between the various motifs. Nevertheless, since the pleated and rippled sheets are energetically similar, as reported by Pauling and Corey,[31] the probability to form the latter on kinetics grounds is more favorable, since the rate of self-assembly depends on the concentration of the short isotactic oligopeptides of both D- and L- handedness, whereas those of the pleated depend only on oligopeptides of single handedness.

The heterochiral residues that are attached at the N-terminus of the peptide chains, as determined by MALDI-TOF MS/MS, induce an enantiomeric cross-inhibition in the step of chain elongation. The stereoselection of amino acids of the same handedness by the growing chains, on one hand, and the cross-

(34) Bartlett, P. D.; Jones, R. H. *J. Am. Chem. Soc.* **1957**, *79*, 2153–2159.

(35) One cannot exclude the possibility that, at these conditions, a small fraction of the isotactic peptides is formed via β-sheets.

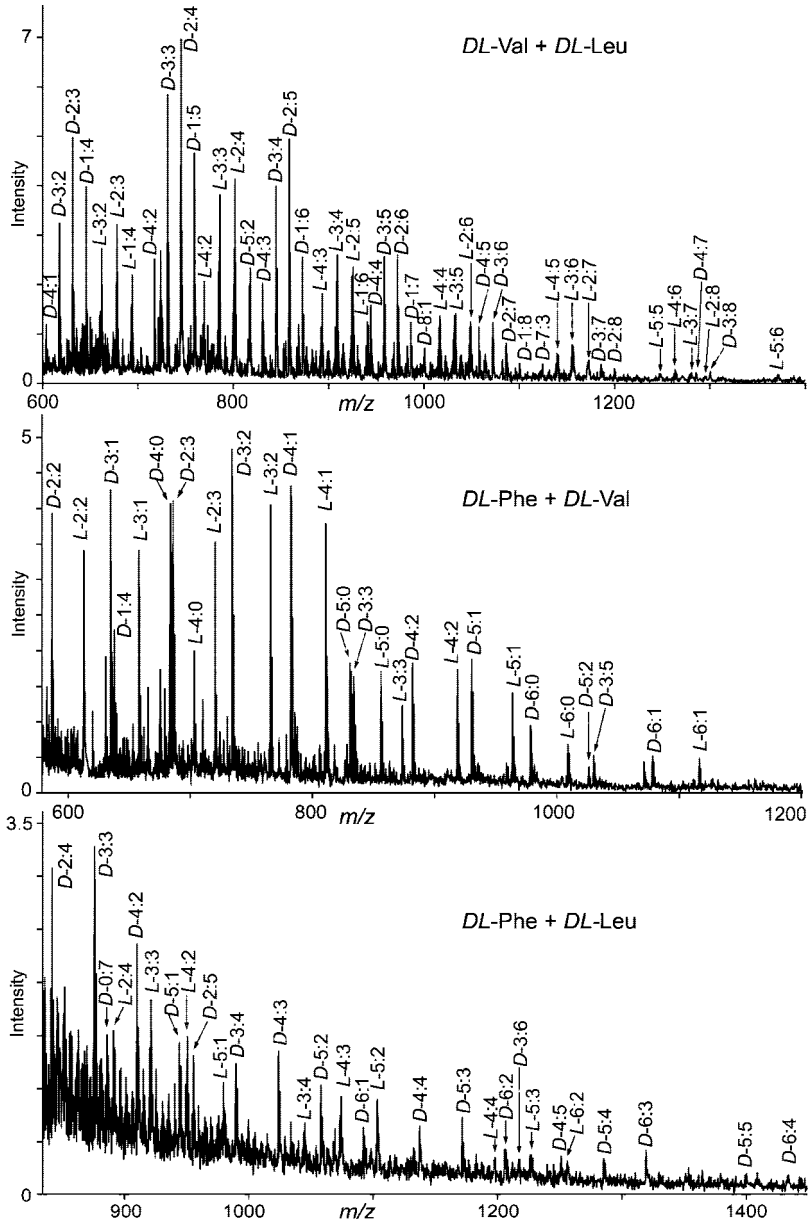


Figure 7. MALDI-TOF spectra of the $(M + Na)^+$ cationized oligopeptides obtained obtained from the polymerization in water, with 25% mol/mol *n*-butylamine initiator, of mixtures of racemic α-amino-acids: (a) DL(d8)-Val + DL(d10)-Leu, (b) DL(d5)-Phe + DL(d8)-Val, (c) DL(d5)-Phe + DL(d10)-Leu. Isotactic copeptides are the dominant signals.

enantiomeric inhibition induced on these chains by amino acids residues of opposite handedness on the other are in keeping with the two central elements present in the Frank model[36] for stochastic "mirror-symmetry" breaking. The operation of such mechanism suggests that slight fluctuations from the racemic state in the racemic amino acids should result in the formation of nonracemic mixtures of the isotactic peptides. "Mirror-symmetry" breaking is also anticipated in the formation of the isotactic copeptides. As a result of the two different C- and N-ends of the peptides, each residue in the chain differs from the others. Consequently, at conditions of low concentrations of some of the amino acids in the polymerization mixtures, one may anticipate that these cannot populate equally all sites of the enantiomeric peptide chains, thus yielding diastereoisomeric libraries rather than racemic mixtures of the isotactic copeptides. Such a scenario has common features with the model proposed by Eschenmoser[37] and Siegel.[38]

Conclusions

The demonstration that isotactic oligopeptides and copeptides of up to 24−30 residues of the same handedness starting from racemic α-amino acids, via the formation of β-sheets as intermediate templates in a one-pot reaction in water, might provide a plausible route for the formation of the prebiotic homochiral peptides. Such studies might support a scenario where homochirality could had been evolved prior to the emergence of the primitive living system, as proposed for nucleic acids by Joyce et al.[39] and Goldanskii et al.[40] The high degree of enantioselectivity in chain elongation of the β-sheets

(36) Frank, F. C. *Biochem. Biophys. Acta* **1953**, *11*, 459.

(37) Bolli, M.; Micura, R.; Eschenmoser, A. *Chem. Biol.* **1997**, *4*, 309.
(38) Siegel, J. S. *Chirality* **1998**, *10*, 24–27.

Table 2. Racemic Mixtures of Homochiral Tricomponent Copeptides Obtained in the Polymerization of 1:1:1 Mixtures of DL(d5)-Phe, DL(d10)-Leu, and DL(d8)-Val, (60mM), with *n*-Butylamine Initiator

D-Phe/Leu/Val[a]		L-Phe/Leu/Val[a]	
1:1:1	1:1:2	1:1:1	1:1:2
1:2:1	0:3:2	1:2:1	1:3:0
3:0:1	1:1:3	2:1:1	0:1:4
1:2:2	1:3:1	1:1:3	1:2:2
2:1:2	2:2:1	1:3:1	2:1:2
2:3:0	3:1:1	2:2:1	3:1:1
3:2:0	1:2:3	1:0:5	2:0:4
4:0:1	1:4:1	0:2:5	2:3:1
2:2:2	3:0:3	3:2:1	0:1:7
1:1:5	3:2:1	5:1:0	1:5:1
1:2:4	1:3:3	1:6:0	2:5:0
1:4:2	4:1:1	1:2:5	2:3:3
2:2:3	2:3:2		
5:1:0	0:4:4		
1:2:5	2:0:6		
2:2:4	3:2:3		
3:3:2	4:3:1		

[a] Number of Phe, Leu, and Val repeating units in copeptides of each length.

provides a simple template that would have enjoyed a considerable selective advantage in a prebiotic environment.[39] In spite of the tendency of hydrophobic peptides such as those of leucine to generally form α-helices, the formation of β-sheets from racemic mixtures of short peptides, as demonstrated in the present study, might support the hypothesis of Brack and Spach[17] that β-sheets could have preceded the formation of the α-helices.[2]

Current studies attempt to use templates of the hydrophobic α-amino acids in order to obtain copeptides containing residues of hydrophilic acids, such as glutamic or aspartic acid. Would such copeptides be formed, one might anticipate that among the libraries of such mixtures one might find diastereoisomers that display catalytic properties[41,42] or operate as templates for the organization of short isotactic peptides that can be ligated enantioselectively.[43] Studies along these lines are in progress.

Acknowledgment. This work was supported by the Israel Science Foundation and the Clore Center for Biological Physics. G.B. thanks Le Conseil Regional Ile-de-France and UPMC. We thank Dr. Alla Shainskaya and her team from the MS laboratory of the Weizmann Institute of Science.

Supporting Information Available: MALDI-TOF MS/MS section and Figure S1; FTIR and PXRD spectra of oligopeptides, Figure S2; MALDI-TOF MS of oligopeptides obtained by polymerization in water of DL(d5)-Phe in the absence and in the presence of L-Leu-OMe initiator and of DL(d8)-Val in the presence of *n*-butylamine and L-Val-OMe initiator, respectively, Figures S3−S6; ^{1}H NMR spectra of L-LeuNCA and L-LeuICA, Figure S7, and of the DLPhe polymerization system in the absence/presence of *n*-butylamine initiator, in D_2O, Figures S8 and 9; MALDI-TOF MS of oligopeptides obtained from a quaternary mixture of amino acids, Figure S10.This material is available free of charge via the Internet at http://pubs.acs.org.

JA709969V

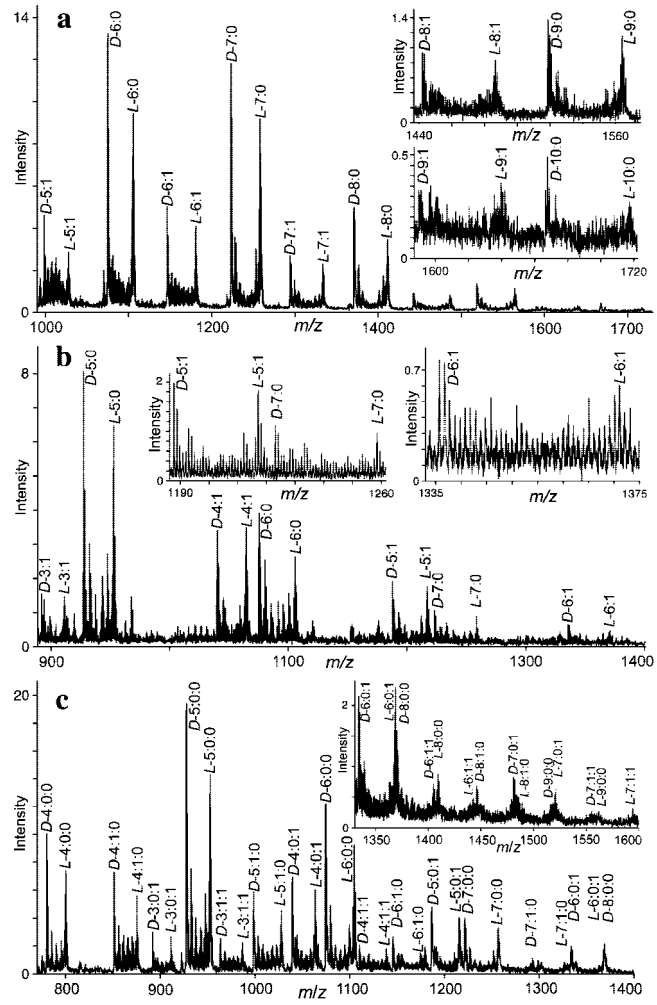


Figure 8. MALDI-TOF spectrum of the $(M + Na)^+$ cationized copeptides obtained in the polymerization in water, with 25% mol/mol *n*-butylamine initiator, of various binary and ternary mixtures of racemic α-amino acids: (a) DL(d5)-Phe (20 mM) + DL(d4)-Ala (20 mM); (b) DL(d5)-Phe (20 mM) + DL(d4)-Tyr (6 mM); (c) DL(d5)-Phe (20 mM) + DL(d4)-Ala (20 mM) + DL(d4)-Tyr (6 mM). The copeptides of homochiral sequence, with total length $n = a + b + c$, are labeled D-*a*:*b*:*c* and L-*a*:*b*:*c*, where *a*, *b*, and *c* represent the number of Phe, Ala, and Tyr repeating units, respectively. Note that the signals correspond to *m*/*z* values of the trifluoroacetyl *N*-termimus derivatized peptides and *O*-trifluoroacetyl-Tyr residues.

(39) Joyce, G. F.; Visser, G. M.; van Boeckel, C. A.; van Boom, J. H.; Orgel, L. E.; van Westrenen, J. *Nature* **1984**, *310*, 602–604.
(40) Goldanskii, V. I.; Avetisov, V. A.; Kuzmin, V. V. *FEBS* **1986**, *207*, 181.
(41) Carrea, G.; Colonna, S.; Kelly, D. R.; Lazcano, A.; Ottolina, G.; Roberts, S. M. *Trends Biotechnol.* **2005**, *23*, 507–513.
(42) Cordova, A.; Zou, W.; Dziedzic, R.; Ibrahem, I.; Reyes, E.; Xu, Y. *Chem.—Eur. J.* **2006**, 5383–5397.
(43) Sagathelian, A.; Yokobyashi, Y.; Soltani, K.; Ghadiri, M. R. *Nature* **2001**, *409*, 797.