Homochiral Oligopeptides Generated by Induced "Mirror Symmetry Breaking" Lattice-Controlled Polymerizations in Racemic Crystals of Phenylalanine *N*-Carboxyanhydride

Jose Geraldo Nery,[a] Gerard Bolbach,[b] Isabelle Weissbuch,[a] and Meir Lahav*[a]

Dedicated to Professor Jean-Marie Lehn on the occasion of his 65th birthday

Abstract: The formation of diastereoisomeric libraries of oligopeptides through the heterogeneous polymerization of racemic crystals of phenylalanine *N*-carboxyanhydride (PheNCA) is reported. The diastereoisomeric compositions of the oligopeptides formed on polymerization of (*R*,*S*) crystals incorporating the deuterium-tagged *S* enantiomer were determined by MALDI-TOF mass spectrometry. The racemic mixtures of the oligopeptides longer than pentamers are represented primarily by diastereoisomers of homochiral sequence and with peptides containing only one heterochiral repeating unit. A mechanism comprising the following three sequential steps to account for this unusual observation is proposed: 1) formation of dimers and trimers at a partially damaged liquid/solid interface, 2) chain propagation that takes place within the bulk of the crystal through a lattice-controlled "zipper-like" mechanism between homochiral molecules arranged in a head-to-tail motif to yield crystalline antiparallel β-sheets of alternating oligopeptide chains of homochiral sequence of opposite handedness, and 3) enantiomeric cross-inhibition that results in chain termination. Induced desymmetrization of the racemic mixtures of the formed peptides was achieved by the polymerization of the mixed quasi-racemic crystals of (*R*)-PheNCA, ((*S*)-PheNCA), and (*S*)-ThieNCA (3-(2-thienyl)-alanine *N*-carboxyanhydride) of various compositions. These experiments resulted in the formation of nonracemic libraries of oligopeptides composed of homochiral chains of (*R*)-Phe and copolymers of randomly distributed (*S*)-Phe and (*S*)-Thie sequences. From these findings, we propose a stochastic model for the generation of libraries of nonracemic mixtures of oligopeptides from the polymerization of host (*R*,*S*)-PheNCA with racemic mixtures of other guest NCA amino acids dissolved in limited quantities in the crystal.

Keywords: amino acids · homochirality · mass spectrometry · oligophenylalanine · solid-state reactions · topochemistry

Introduction

Prebiotic scenarios that might have operated for the formation of primordial homochiral biopolymers generated from racemic mixtures of the corresponding activated monomers (or nonracemic mixtures of low enantiomeric imbalance) provide an interesting riddle in the field of the origin of life.[1–9] Two central questions have to be addressed: firstly, how to design feasible prebiotic synthetic routes for the generation of racemic mixtures of polymers composed of repeating units of the same handedness (isotactic polymers of homochiral sequence) and secondly, how to generate enantiopure isotactic chains when starting out from nonracemic mixtures of the corresponding monomers. Polymerization of racemates in solution in most instances yields (with some exceptions[10,11]) complex mixtures of atactic polymers. One way to obtain isotactic polymers comprising homochiral sequences when starting from racemates is to perform the reaction in a crystalline environment. Such reactions in three-dimensional crystals, however, are rarely lattice-controlled beyond the formation of dimers and trimers.[12,13] Apart

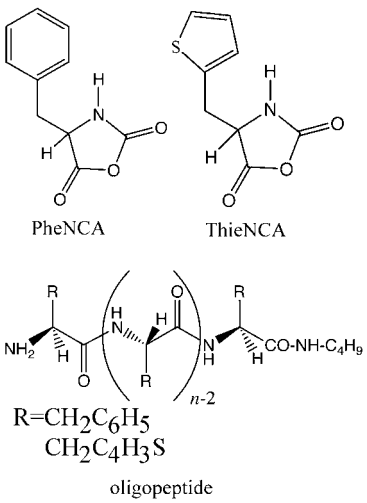
[a] Dr. J. G. Nery, Dr. I. Weissbuch, Prof. M. Lahav
Department of Materials and Interfaces
Weizmann Institute of Science
76100-Rehovot (Israel)
E-mail: meir.lahav@weizmann.ac.il

[b] Dr. G. Bolbach
Laboratoire de Chimie Structurale Organique et Biologique
Université Pierre et Marie Curie
75252 Paris Cedex 05 (France)

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

 DOI: 10.1002/chem.200401097

from the 2π–2π photopolycycloaddition reactions of diethylenes,[14,15] di- and triacetylenes,[16] butadienes,[17–19] and some organometallic monomers[20,21] that have been demonstrated to undergo topochemical reactions, polymerization of most other monomers yields only atactic polymers. Therefore, the generation of homochiral polypeptides through heterogeneous polymerization reactions in three-dimensional crystals as a possible reaction pathway was not regarded as a viable scenario. Recently, though, we reported a possible route for the generation of peptides of homochiral sequence from racemates. It comprised two steps: self-assembly of the monomers into two-dimensional crystallites at the air/solution interface, followed by a lattice-controlled reaction.[22–24] During these studies we noticed that the polymerization of two-dimensional crystallites of γ-stearylglutamic acid *N*-carboxyanhydride took place through an efficient lattice-controlled reaction by a "zipper-like" mechanism between neighboring molecules arranged in a "head-to-tail" pattern along a polar direction.[22] After this result, we anticipated that such a polymerization might also occur through a lattice-controlled process to yield isotactic polypeptides in similar or related three-dimensional crystalline motifs. Kanazawa et al.[25–27] have reported extensive crystallographic and kinetic studies on the solid-state polymerization of *N*-carboxyanhydrides of amino acids. The sequences of the oligopeptides generated in the racemic crystals, however, were not available. More recently, we reported preliminary results



on a stereospecific polymerization of racemic PheNCA in the crystalline state, yielding oligopeptides and with a high fraction of them being racemic mixtures of homochiral sequence.[28]

Here we propose the operation of a cooperative multistep mechanism that might account for the formation of these isotactic oligopeptides. We also report the synthesis of libraries of nonracemic homochiral oligopeptides generated by an efficient induced desymmetrization process.

Results and Discussion

Polymerization of racemic crystals of PheNCA: The crystalline motif of racemic PheNCA[26] ($a = 9.606$, $b = 6.378$, $c = 30.077$ Å, space group $Pna2_1$ $Z = 4$) viewed along the *a* axis is shown in Figure 1a. The crystal contains two independent, almost identical, molecules per unit cell. The molecules form two-dimensional hydrogen-bonded networks of four rows that form a bilayer motif arranged perpendicular to the *c* axis. Within each row, the molecules are related by translation in a head-to-tail motif interlinked by C=O···H–N hydrogen bonds. The two rows composed of homochiral molecules within the bilayer are not symmetry related and assume a polar motif along the *b* axis. The sense of polarity of the *R* molecules is opposite to that of the *S* molecules in each bilayer (Figure 1b). The bilayers are separated from one another by hydrophobic interactions of the phenyl rings and are related by a twofold screw axis along the *c* axis.

In a heterogeneous polymerization initiated by *n*-butylamine,[25] we may anticipate that at the initiation step the amine will react with a carbonyl group connected to the chiral carbon atom of either an *R* or an *S* monomer molecule at the solid/liquid interface. Consequently, the reaction within the interior of the crystal should proceed through a "zipper-like" head-to-tail mechanism, preferentially between

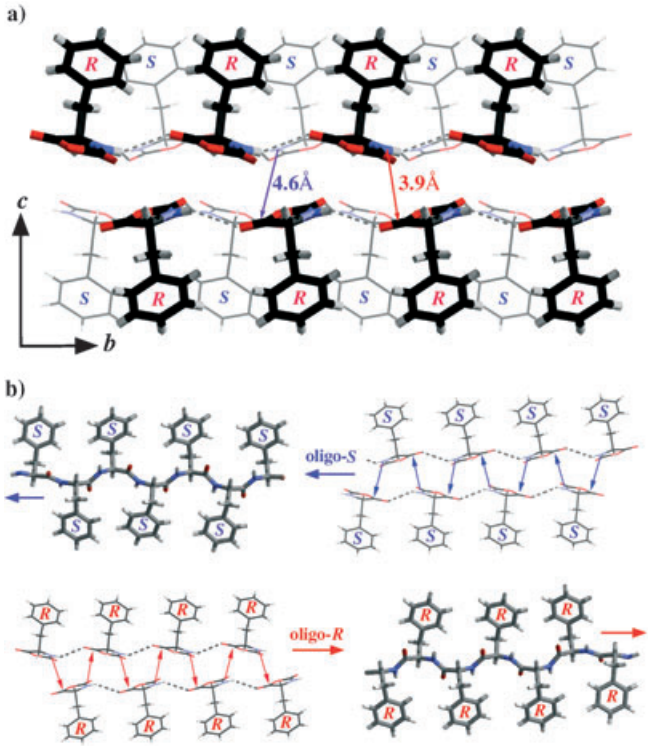


Figure 1. a) Packing arrangement of (*R*,*S*)-PheNCA viewed along the *a* axis, showing four rows of molecules. b) Pathways of the polymerization reaction by a "zipper-like" mechanism for a bilayer of *S* molecules (top), and *R* molecules (bottom) yielding the corresponding oligo-*S* and oligo-*R* peptides with homochiral sequences, as modeled on the basis of the monomer crystal lattice. The arrows signify the direction of polymerization within each bilayer.

the closest homochiral neighboring molecules, since the spacing between their closely situated reacting atoms (H–N to C=O) is shorter (3.90 Å) than the corresponding distance (4.60 Å) for two adjacent heterochiral molecules (Figure 1 a). Polymerization within these layers should yield racemic mixtures of oligopeptides of homochiral sequence, in which adjacent poly-*R* chains in each bilayer point in the opposite direction to those of the *S* chains (Figure 1 b).

Complex mixtures of diastereoisomeric oligopeptides of various lengths, containing up to 27 repeating units, were obtained by *n*-butylamine-initiated polymerization of racemic PheNCA crystals suspended in hexane, and were analyzed by MALDI-TOF mass spectrometry. To probe the diastereoisomeric distribution of the oligopeptides, the polymerization reactions were performed with crystals composed of (*R*)-PheNCA and (*S*)-PheNCA in which the five hydrogen atoms of the phenyl ring of the latter enantiomer were replaced with deuterium.[10,29,30] The MALDI-TOF mass spectrum of an overall sample was taken immediately after the reaction had been completed and the oligopeptides converted into the corresponding soluble *N*-trifluoroacetamides. Oligopeptides bearing up to ten to twelve repeating units could be detected in the mass spectrum with the labeled samples. The *m*/*z* range for penta- to decapeptides is shown in Figure 2, with expanded spectra of the octamer and nonamer ranges given in the two insets, and with the number of *R* and *S* repeating units in each oligopeptide labeled.

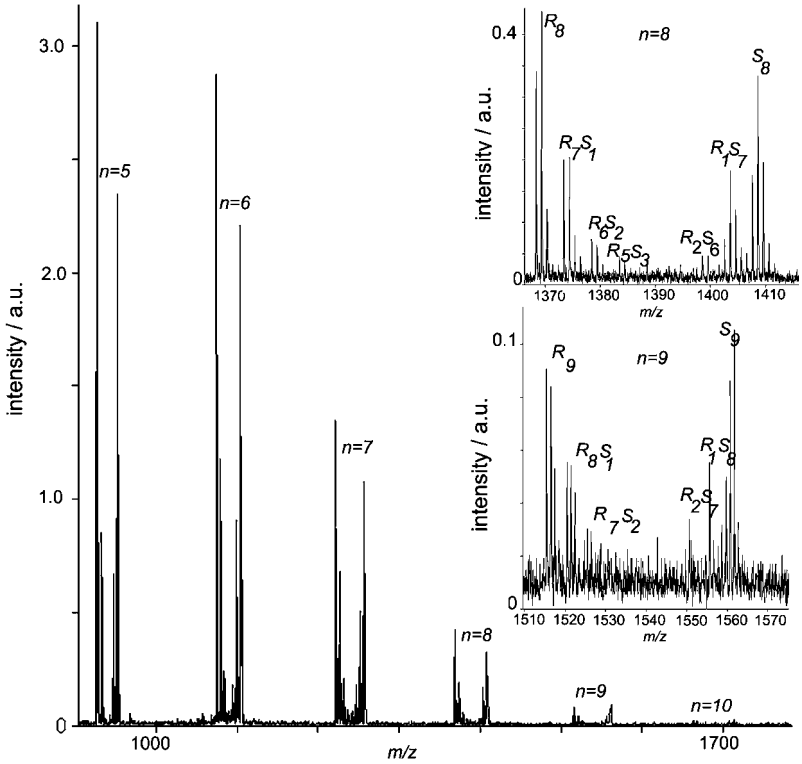


Figure 2. Positive-ion MALDI-TOF mass spectrum of the oligopeptides (Na cationized) obtained in the polymerization of (*R*,*S*)-PheNCA at 22 °C, showing the *m*/*z* range from penta- to decapeptides. The phenyl ring in the *S* enantiomer was deuterated [D_5]. The two insets show expanded spectra of the octamer and the nonamer *m*/*z* ranges.

Since the peaks of the matrix coincided with some of the peaks of the dimers, trimers, and tetramers, their spectra could be better assigned after their extraction with THF. This fraction comprised about 43 % of the entire reacted sample and contained primarily all the dimers, trimers, and tetramers, together with small quantities of long atactic oligomers (Figure 3). The long homochiral oligopeptides are not soluble in THF.

In our analysis of the relative compositions of the different diastereoisomers we assumed that the relative quantities of oligopeptides of a given length should be proportional to the intensities of the *m*/*z* signals. This assumption is reasonable for oligopeptides of the same length because of the same ionization yields (similar chemical properties) and detection efficiencies. The compositions of the dimeric diastereoisomers—1:1:1 *RR*/(*RS*+*SR*)/*SS* (Figure 3 bottom insert)—as well as of the trimers—1:1.6:1.6:1 (Figure 3 middle insert)—and tetramers—1:1.9:2.0:1.9:1 (Figure 3 top insert)—are similar to the corresponding distributions obtained from the polymerization experiments performed with the melted monomer (see Supporting Information). This similarity implies that the crystalline lattice exerts a very weak control, if any, over the formation of the short oligopeptides at the liquid/solid interface. However, the departure of the distribution of short oligopeptides from a binomial distribution implies the operation of some degree of diastereoisomeric induction in the chain-propagation step of the reaction. In contrast, the distribution of the oligopeptides longer than pentapeptides generated in the polymerization of the crystalline monomer (Figure 2) is very different from that observed in the polymerization of the melted monomer (Figure 4). This difference indicates the operation of strict control exerted by the crystalline lattice on the formation of the long oligopeptides.

The distribution of the longer oligopeptides was obtained directly from analysis of the whole polymerized sample (see Figure 2). The abundance of these peptides decreases with increasing chain length. The peaks at the left wings of each composition in the spectra represent the untagged isotactic oligopeptides of *R* absolute configuration (*n*,0) (labeled R_8 and R_9 for octamers and nonamers in the insets of Figure 2) and the peaks at the right wings the deuterated oligopeptides of *S* absolute configurations (0,*n*)

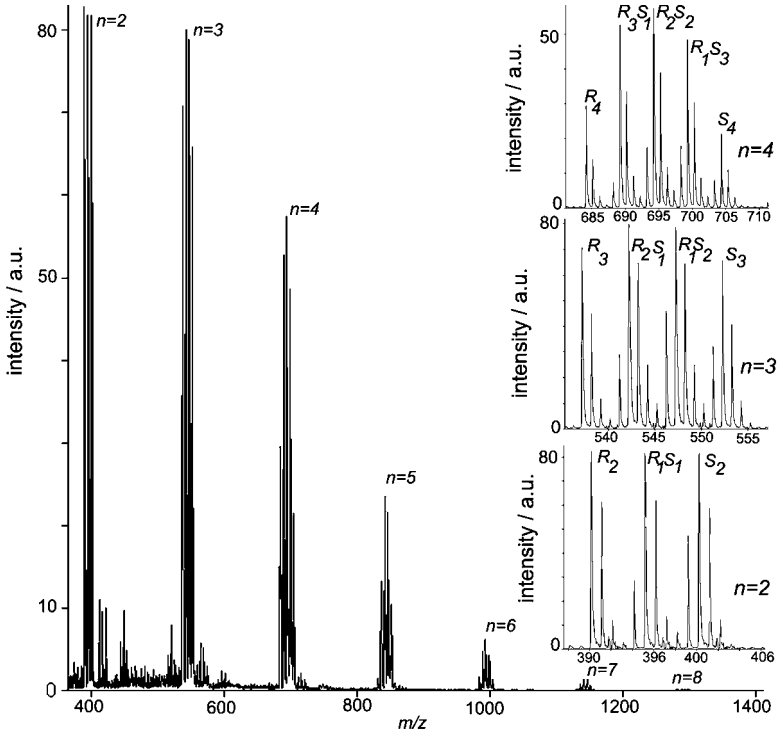


Figure 3. Positive-ion MALDI-TOF mass spectrum of the THF-soluble fraction of the oligopeptides (Na cationized) obtained from the polymerization of (*R*,*S*)-PheNCA at 22 °C, showing the *m*/*z* range from di- to octapeptides. The phenyl ring in the *S* enantiomer was deuterated [D_5]. The three insets show expanded spectra of the dimer (bottom), trimer (middle), and tetramer (top).

(S_8 and S_9), where *n* is the total number of repeating units in each oligopeptide chain. The peaks between the wings represent heterochiral oligopeptides labeled (R_h,S_d) where *h* is the number of *R* repeating units and *d* the number of *S* deuterated ones. Inspection of Figure 2 shows that the peaks of the wings and those of (*n*−1,1) and (1,*n*−1)—R_8, S_8, R_7S_1, and R_1S_7, for example—are the most intense for oligopeptides for all lengths. The high degrees of stereospecificity observed in these reactions, as opposed to their random distribution, would entail the operation of an uncommon lattice-controlled mechanism. The relative abundances of the homochiral oligopeptides as normalized to the corresponding values of a binomial distribution are shown in Figure 5.

To explain this distribution of the products, we allude again to the packing arrangement of the monomer crystal. We propose that once the dimers *Bu-R-R*, *Bu-S-S*, and (*Bu-R-S* and *Bu-S-R*) are formed at the liquid/solid interfaces, they subsequently propagate towards the interior of the crystal by a "zipper-like" head-to-tail mechanism. The *Bu-R-R* and *Bu-S-S* species yield the homochiral oligopeptides *Bu*-$(R)_n$ and *Bu*-$(S)_n$, respectively (Figure 1b). On the other hand, the heterochiral dimers also undergo a similar "zipper-like" polymerization in which the second repeating unit in *Bu-S-R*, formed by the reaction of molecule S^1 (green) with molecule R^2 (pink) (Figure 6a), dictates both the direction of the chain propagation and the absolute configuration of the new monomers to be added to the growing chains. Such a reaction pathway might also account for the formation of oligopeptides *Bu-S*-$(R)_{n-1}$ and *Bu-R*-$(S)_{n-1}$ (Figure 2). The modeled conformation of the tetrapeptide *Bu-S*-$(R)_3$ is shown in Figure 6b.

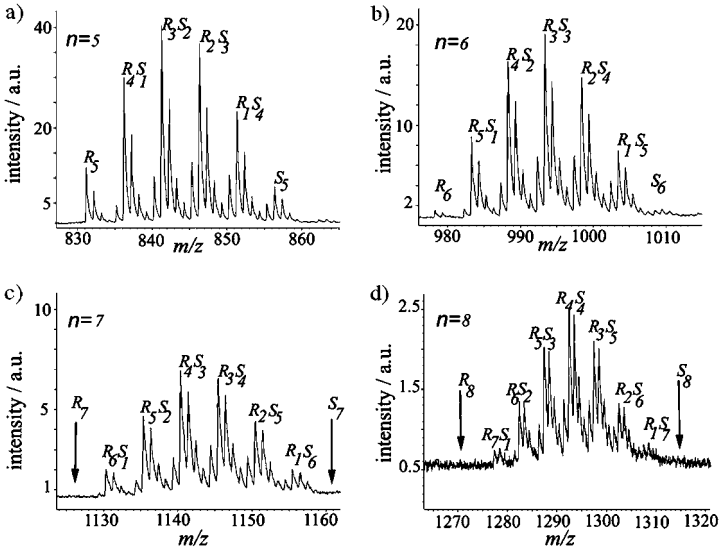


Figure 4. Positive ion MALDI-TOF mass spectrum of the oligopeptides (Na cationized species) obtained from polymerization of melted (*R*,*S*)-PheNCA suspended in heptane at 100 °C, showing the *m*/*z* ranges of: a) penta-, b) hexa-, c) hepta-, and d) octapeptides. The arrows in (c) and (d) indicate the absence of oligopeptides of homochiral sequence.

During the chain propagation process, one might expect that the degree of isotacticity of polymerization should be reduced as a result of partial destruction of the monomer crystal at sites of reaction, owing to the release of CO_2 and decrease in density of the growing polymeric phase relative to that of the monomer. The unusual increase in stereospecificity of the reaction with increasing chain length in this system can be interpreted by the assumption that, once formed, the chiral dimers exert strong diastereoisomeric control over the addition of fresh monomer units to the growing chains within the local environment of chain propagation in the monomer crystal. Were a monomer of *R* absolute configuration—R^7 (yellow), for example—to add to a growing chain composed of all *S* repeating units—S^1-to-S^6 (blue), say (Figure 6a)—such an addition should force the added *R* repeating unit to assume an energetically unfavora-

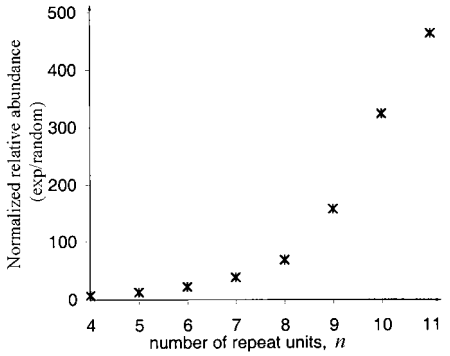


Figure 5. Enhancement of the experimentally observed relative abundances of the homochiral [(*n*,0)+(0,*n*)] oligopeptides normalized to those calculated for a theoretical random process for molecules of any length *n*. Note that deca- and undecapeptides were detected only in some of the polymerization experiments.

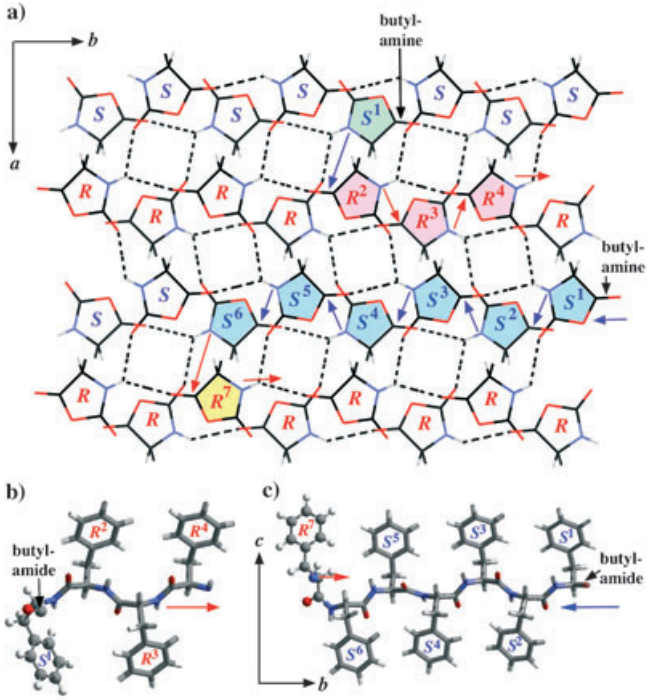


Figure 6. a) Packing arrangement of (*R*,*S*)-PheNCA viewed down the *c* axis, showing: 1) The proposed mechanism of the initiation and chain propagation of *Bu*-*S*-(*R*)$_n$ oligopeptides as follows. Molecule S^1 (green) reacts with molecule R^2 (pink), which further dictates the direction of the chain propagation with molecules R^3, R^4, etc. The conformation of such a *Bu*-*S*-(*R*)$_3$ tetrapeptide modeled on the basis of the monomer crystal is shown in (b). 2) The proposed chain termination of an (*S*)$_6$ hexapeptide formed by the polymerization of molecules S^1 to S^6 (blue) with molecule R^7 (yellow) to yield oligopeptides of sequence *Bu*-(*S*)$_{n-1}$-*R*. The unfavorable conformation after addition of R^7 unit would cause a turn at the propagating end of the heptapeptide that would further enforce a flip in the direction of the chain propagation, thus implementing a chain termination. The modeled conformation of such a *Bu*-(*S*)$_6$-*R* heptapeptide is shown in (c). The arrows indicate the direction of chain propagation.

ble conformation that would alter the structure of the growing end of the oligopeptide. Furthermore, were such a group to be added, it would have to fold the rigid oligopeptide chain and flip the direction of the propagating chain. The modeled conformation of the *R* repeating unit at the growing end of the chain is shown in Figure 6c. Consequently, such addition would presumably result in a chain-termination step that would block additional growth of the polymeric chains. The operation of this termination step should generate oligopeptides of sequence (*n*−1,1) and (1,*n*−1) from the growing chains of the homochiral sequence and oligopeptides (*n*−2,2) and (2,*n*−2) from the growing chains of *Bu*-*R*-(*S*)$_{n-1}$ and *Bu*-*S*-(*R*)$_{n-1}$. From this mechanism, we anticipate that repeating units of opposite absolute configuration should also be found at the *N* end of the peptide in the longer oligopeptides of type *Bu*-(*R*)$_{n-1}$-*S* and *Bu*-(*S*)$_{n-1}$-*R*, and that oligopeptides of type (*n*−2,2) and (2,*n*−2) formed in the crystal should contain the sequences *Bu*-*R*-(*S*)$_{n-2}$-*R* and *Bu*-*S*-(*R*)$_{n-2}$-*S*, respectively.[31]

From the packing arrangement of the monomer crystal and the formation of racemic mixtures of oligopeptides of homochiral sequence as the large fraction, one may predict that the isotactic chains should be able to self-assemble in the form of β-sheets composed of alternating oligo-*R* and oligo-*S* chains aligned in antiparallel directions, induced by the packing arrangement of the monomer crystal, as modeled in Figure 7.

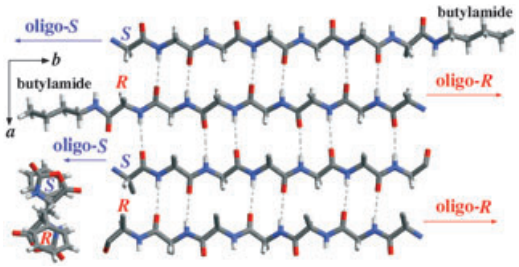


Figure 7. Proposed formation of crystalline antiparallel β-sheets of oligo-*R* and oligo-*S* peptides, as modeled on the basis of the crystal lattice of the (*R*,*S*)-PheNCA monomer, as viewed along the *c* axis. Two unreacted molecules of the monomer, in its crystal lattice, are shown on the left.

The FTIR spectra (reflection mode) of solid samples immediately after reaction are shown in Figure 8, together with those of the THF-soluble fractions and of the THF-insoluble fractions. The THF-soluble fraction shows a very broad peak centered at 1650 cm^{-1}, a value close to the 1655–1668 cm^{-1} region where the amide I (backbone C=O) band of oligopeptides should occur in the cases either of a random coil or of an α-helical conformation.[32] On the other hand, the THF-insoluble fraction displays a very sharp peak at 1632 cm^{-1}, characteristic of the amide I stretching vibration of oligopeptides in the β-sheet conformation. Moreover, the shoulder observed at ≈1685 cm^{-1}, characteristic of antiparallel β-sheet structure, is in agreement with our proposed reaction pathway.

X-ray powder diffraction analysis of the insoluble fraction demonstrates the formation of highly crystalline oligopeptides. Among the various *d* spacings, we observed the 9.4 Å,

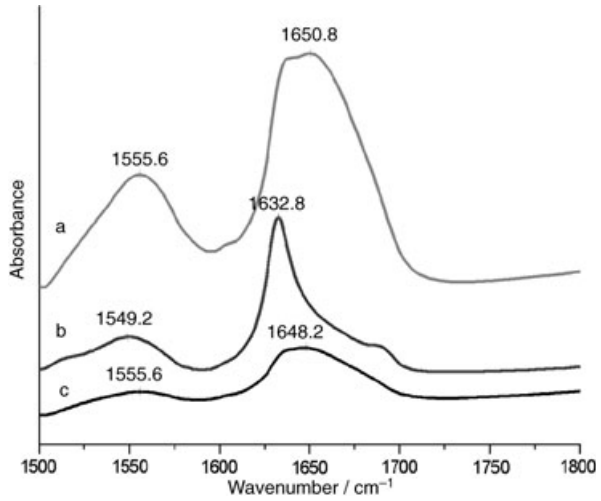


Figure 8. FTIR analysis (in reflection mode) of: a) whole solid sample immediately after polymerization, b) THF-insoluble fraction, and c) THF-soluble fraction.

≈6.9 Å, and very strong 4.7 Å spacings that have been reported in crystalline β-sheet structures.[33] The 4.7 Å spacing corresponds to the inter-strand hydrogen bonding, whereas the 9.4 Å (= 2×4.7 Å) is, in principle, characteristic of an antiparallel arrangement.[34] SEM analysis of the insoluble fraction shows that the oligopeptides organize themselves primarily in the form of long rods (Figure 9).

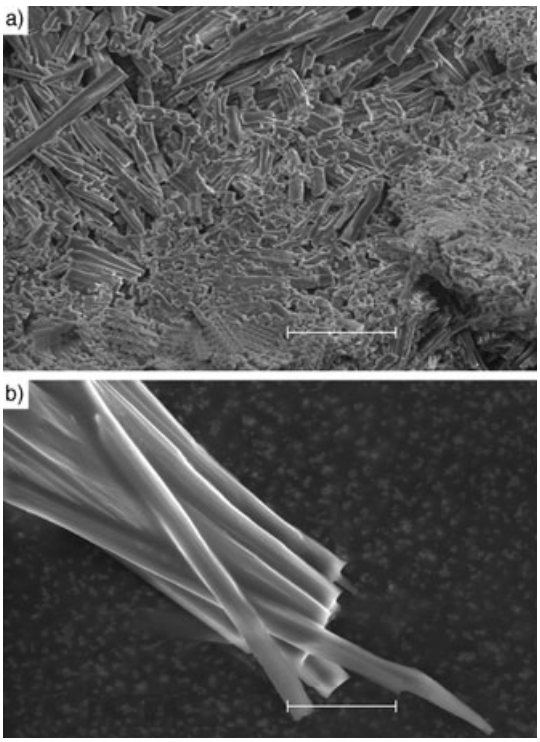


Figure 9. Scanning electron microscopy images of the THF-insoluble fraction at magnifications: a) scale bar 50 μm and b) scale bar 10 μm.

Nonracemic libraries of oligopeptides by induced desymmetrization: The above results suggest that racemic crystals of PheNCA might be most appropriate systems for the generation of nonracemic libraries of isotactic oligopeptides by an induced "mirror symmetry breaking" process. There are several ways to induce departures from the racemic mixtures of the homochiral oligopeptides. Here we demonstrate a scenario comprising the enantioselective insertion of guest homochiral *N*-carboxyanhydrides of amino acids within the chiral sites of racemic crystals of PheNCA. Were an enantiopure guest monomer of *S* absolute configuration, for example, to be occluded in the crystal by occupying sites of (*S*)-PheNCA molecules, the lattice-controlled polymerization in such a mixed crystal should yield nonracemic libraries of oligopeptides composed of monocomponent isotactic peptides of (*R*)-Phe repeating units and bi- or multicomponent isotactic copolymers, as illustrated below.

To make the diastereoisomeric mixtures of oligopeptides amenable to MALDI-TOF mass spectrometric analysis, crystals composed of (*R*)-PheNCA/(*S*)-Phe([D_5])NCA containing varying amounts of (*S*)-ThieNCA were grown after treatment of the corresponding amino acids with bis-(trichloromethyl)carbonate. Mixed crystals of three compositions were prepared and polymerized under the same conditions as reported above for the corresponding (*R*,*S*)-PheNCA crystals. The following compositions were investigated: 1:1 (*R*)-Phe/(*S*)-Thie, 1:0.5:0.5 (*R*)-Phe/(*S*)-[D_5]Phe/(*S*)-Thie, 1:0.9:0.1 (*R*)-Phe/(*S*)-[D_5]Phe/(*S*)-Thie, and 1:0.95:0.05 (*R*)-Phe/(*S*)-[D_5]Phe/(*S*)-Thie. After polymerization, a whole reacted sample was transformed into its corresponding N-trifluoroacetamides and analyzed by MALDI-TOF mass spectrometry, as shown in Figure 10, Figure 11, and Figure 12. The mass spectra obtained from the polymerized quasi-racemic monomer (Figure 10) again indicate the preferential formation of homochiral oligopeptides of *Bu*-((*R*)-Phe)$_n$ (*n*,0) and *Bu*-((*S*)-Thie)$_n$ (0,*n*), together with the

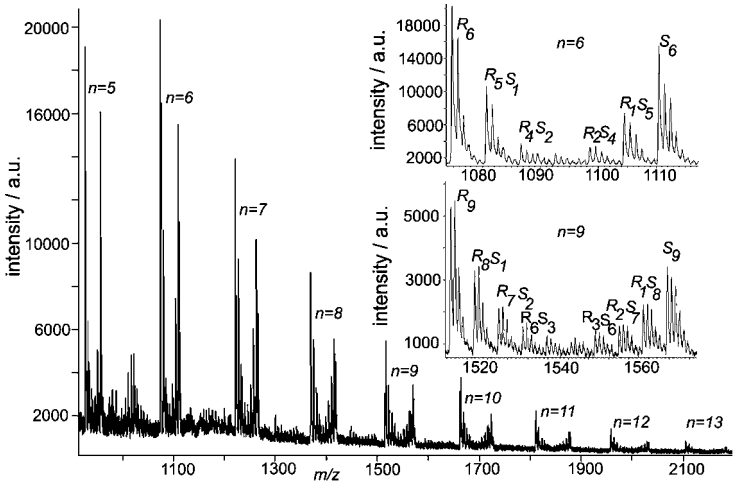


Figure 10. Positive-ion MALDI-TOF mass spectrum of the oligopeptides (Na cationized) obtained from the polymerization of 1:1 (*R*)-PheNCA/(*S*)-ThieNCA, showing the *m*/*z* range from pentamer to 13-mer peptides.

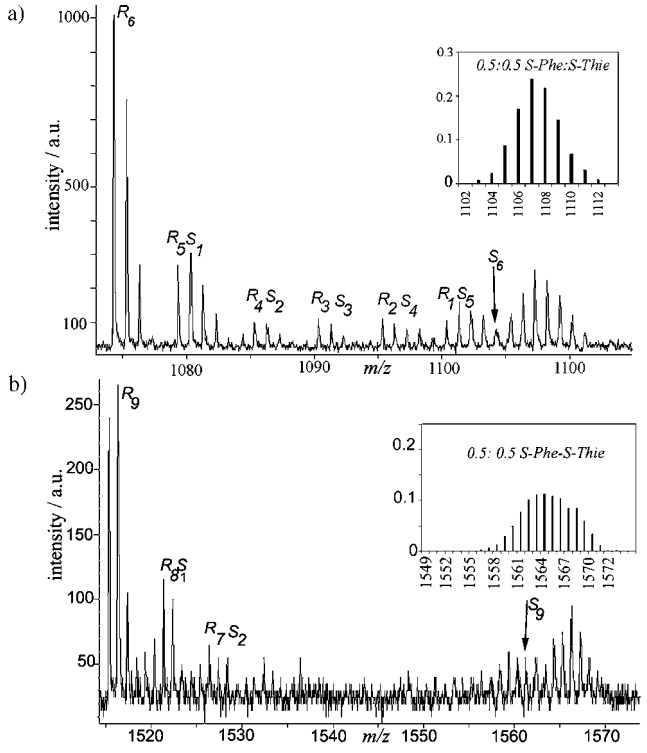


Figure 11. a) Hexamer and b) nonamer m/z regions of the positive ion MALDI-TOF mass spectrum of the oligopeptides (Na cationized) obtained from the polymerization of 1:0.5:0.5 (*R*)-PheNCA/(*S*)-[D_5]PheNCA/(*S*)-ThieNCA. The inserts show the calculated isotopic pattern of a random polymerization for 0.5:0.5 (*S*)-[D_5]Phe/(*S*)-Thie.

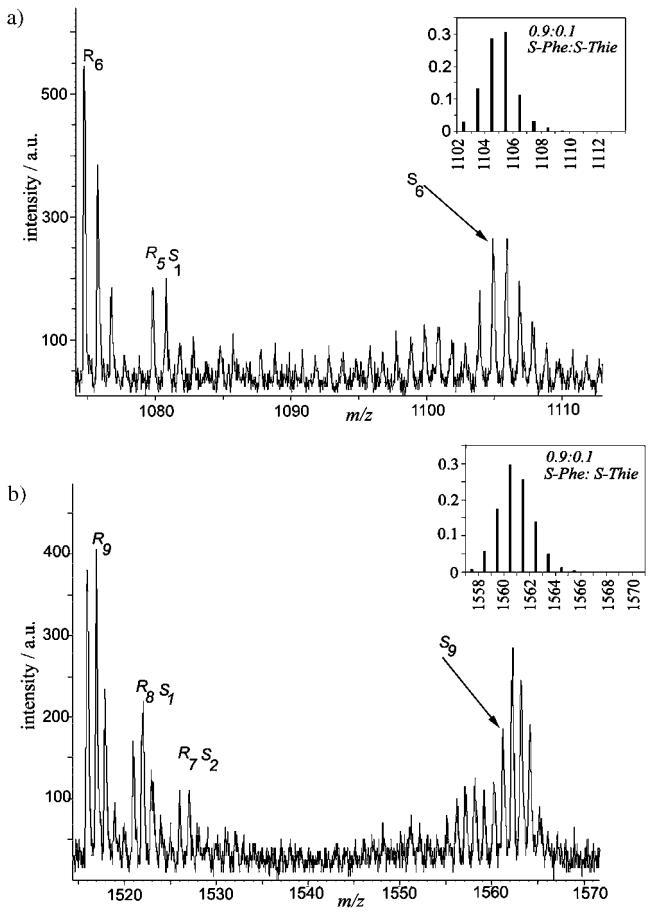


Figure 12. a) Hexamer and b) nonamer m/z regions of the positive ion MALDI-TOF mass spectrum of the oligopeptides (Na cationized) obtained from the polymerization of 1:0.90:0.10 (*R*)-PheNCA/(*S*)-[D_5]PheNCA/(*S*)-ThieNCA. The inserts show the corresponding calculated isotopic pattern of random distribution of 0.90:0.10 (*S*)-[D_5]Phe/(*S*)-Thie.

corresponding peaks of *Bu*-((*R*)-Phe)$_{n-1}$-((*S*)-Thie)$_1$ (n−1,1) and *Bu*-((*R*)-Phe)$_1$((*S*)-Thie)$_{n-1}$ (1,n−1).

Figure 11 and Figure 12 show a comparison between the experimentally obtained MALDI-TOF MS spectra of the hexamer (Figure 11a, Figure 12a) and of the nonamer (Figure 11b, Figure 12b) m/z ranges obtained by polymerization of monomer compositions of (*R*)-Phe/(*S*)-[D_5]Phe/(*S*)-Thie of 1:0.5:0.5 and 1:0.9:0.1, respectively. For comparison, the inserts in each of these figures also show simulated spectra of oligo-(*S*)-[D_5]Phe/(*S*)-Thie copolymers calculated on the assumption of a random distribution of the (*S*)-[D_5]Phe and (*S*)-Thie repeating units and with allowance made for the natural isotopes of sulfur and the 98% deuteration of (*S*)-[D_5]PheNCA (see Experimental Section). The simulated spectra coincide with the experimentally observed ones, in agreement with the proposed random replacement of the (*S*)-[D_5]PheNCA sites of the host monomer crystal with (*S*)-ThieNCA molecules.

As a result of the random distribution of the two repeating units within the homochiral *S* copolymers, the degree of departure from the nonracemic composition varies with chain length and the starting composition of the monomer mixture. All oligopeptide chains containing one or more repeating unit of the (*S*)-Thie are enantiopure. On the other hand the homochiral oligopeptides of (*R*)-Phe are enantiomerically enriched. The proportion of the homochiral oligopeptides of *S* configuration for each length and composition, and consequently the enantiomeric excess for each chain, can be calculated by applying the binomial theorem (1), where $\frac{n!}{m!(n-m)!}$ are the binomial coefficients; that is, the number of ways of picking up unordered outcomes from n possibilities, and a and b are the fractions of (*S*)-PheNCA and (*S*)-ThieNCA molecules in the mixed crystals of the monomer.

$$(a+b)^n = \sum_{m=0}^{n} \frac{n!}{m!(n-m)!} a^m b^{(n-m)} \qquad (1)$$

From the fractions of the homochiral cooligopeptides of *S* configuration, we calculated the enantiomeric excess (*ee*; *ee* = 100(*R*−*S*)/(*R*+*S*)) of the homochiral oligo-*R* peptides of each chain length. Figure 13 shows the correlation between the computed *ee* for homochiral *R* dimers to *R* 100-mers for the three studied compositions. From this analysis we can note that the *ee* of the homochiral oligopeptides increase

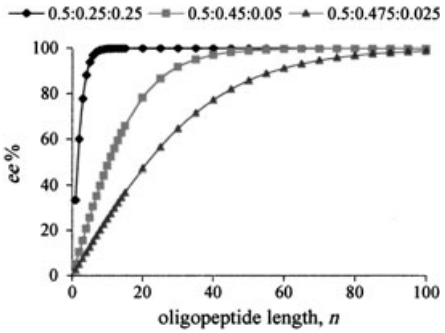


Figure 13. Calculated *ee* values of the *R*-(Phe)$_n$ peptides of each chain length *n* for each of the three starting (*R*)-PheNCA/(*S*)-[D$_5$]PheNCA/(*S*)-ThieNCA compositions of the monomer crystals.

with increasing oligopeptide chain length. For example, in the systems of 1:0.5:0.5 (*R*)-PheNCA/(*S*)-[D$_5$]PheNCA/(*S*)-ThieNCA composition, the dimers have *ee* = 60%, whereas oligomers such as the decamers or longer are almost enantiopure. On the other hand, in mixtures low in ThieNCA, such as 0.95:0.05 (*S*)-[D$_5$]Phe/(*S*)-Thie, only the oligopeptides beyond twenty units will contain at least one repeating unit of this amino acid.

To provide independent structural information on the position of the (*S*)-ThieNCA groups in the mixed crystals, the crystal structure of the quasi-racemate of (*S*)-ThieNCA with (*R*)-PheNCA was determined: $a = 6.308$, $b = 28.96$, $c = 9.574$ Å, $\beta = 90.06°$, space group $P2_1$ $Z = 2$, four independent molecules (see Experimental Section). The packing arrangement is isostructural with that of (*R*,*S*)-PheNCA, as shown in Figure 14. There are very small differences in cell dimensions between the two crystals. The crystal structure of racemic (*R*,*S*)-ThieNCA is also isostructural with that of (*R*,*S*)-PheNCA ($a = 9.5530$, $b = 6.2730$, $c = 28.721$ Å, space group $Pca2_1$ $Z = 4$, two independent molecules (see Experimental Section).

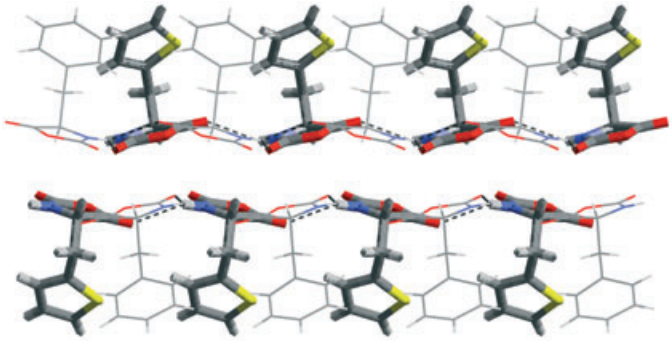


Figure 14. Packing arrangement of the quasi-racemate of (*S*)-ThieNCA with (*R*)-PheNCA, viewed along the *a* axis and showing four rows of molecules.

Conclusion

Polymerization of racemic NCA amino acids suspended in hexane proceeds through two different routes: a reaction in disordered regions, presumably at the solid/liquid interface, to yield short oligopeptides with distributions resembling that of the polymerization in the melt, and a major enantioselective polymerization resulting in the formation of longer oligopeptides of homochiral sequence. Consequently, the distribution of the products depends upon the quality of the crystals.

A cooperative mechanism comprising the formation of dimers, trimers, and possibly tetramers in noncrystalline regions of the crystal, followed by a "zipper-like" enantiospecific propagation and a cross-enantiomeric inhibition step, is proposed to explain the unexpected observation of the lattice-controlled polymerization that gives rise to an augmentation of the degree of diastereospecificity with increasing molecular weight of the oligopeptides.

We have also shown that the packing arrangement of the (*R*,*S*)-PheNCA crystal is most appropriate for the generation of libraries of nonracemic mixtures of homochiral oligopeptides by desymmetrization of the racemic crystals with optically pure ThieNCA. The polymerization within such mixed crystals gives rise to the formation of diastereoisomeric libraries of nonracemic oligopeptides of homochiral sequence composed of enantiopure (Phe)$_n$ and homochiral co-oligopeptides of opposite handedness. The *ee*s of the homochiral oligopeptides are always higher than that of the starting monomer. The *ee* values increase with increasing length of the oligopeptides.

Since each oligopeptidic chain has a C and an N terminus, each repeating unit within each oligopeptide chain of a homochiral sequence differs from the others. On the basis of the mechanism of enantioselective insertion of guest NCA monomers within the host racemic crystals of PheNCA and their random distribution within the homochiral sites to form co-peptides of random distribution of the guests within the host, one may envisage a plausible scenario for the generation of libraries of diastereoisomeric mixtures of peptides by use of racemic mixtures of other NCA amino acids as guests. Under such circumstances, the *S* guests will occupy the sites of the *S* host crystal whereas the *R* guests will occupy the *R* sites of the host. Under regimes in which the numbers of the racemic guests were insufficient to populate all different sites of the chains of the oligopeptides, one would generate a library of diastereoisomeric mixtures of oligopeptides rather than the racemic ones. This mechanism has some features in common with related scenarios proposed recently[35–37] in which a deficient supply of material that does not suffice for complete population of all the sites of libraries of diastereoisomeric polymers results in a stochastic symmetry-breaking process. A detailed analysis of such a process will be presented elsewhere.

In these studies we have used alkanes as suspending solvents; preliminary results indicate that homochiral oligopeptides can also be obtained by performing the polymerization reactions in water, as well as by using other NCA amino acids.

Finally, these reactions provide only a model system, so their prebiotic relevance remains as a proof of principle.

Experimental Section

Synthesis of racemic *N*-carboxyanhydrides: NCA monomers were prepared by a reported procedure.[38] (*R*)-Phe (Aldrich, 1 mmol) and (*S*)-Phe (ring D5 98%, Cambridge Isotope Laboratory, 1 mmol) were suspended in dry THF and heated to 40°C under argon. Solid bis-(trichloromethyl)-carbonate (3.5 mmol) was added gradually for about 1 h until a clear solution was obtained; the reaction was then allowed to proceed for an additional 3 h. The reaction mixture was cooled down and about 80% of the THF was evaporated in vacuum. Hexane was then added to the solution and the NCA crystals were formed overnight at −5°C. The thin, platelike (*R*,*S*)-PheNCA crystals obtained after filtration were analyzed by FTIR spectroscopy and X-ray powder diffraction.

Synthesis of quasi-racemic *N*-carboxyanhydrides: (*R*)-Phe (1 mmol) and 3-(2-thienyl)-(*S*)-alanine (Fluka, 1 mmol) were suspended in dry THF and heated to 40°C under argon. Solid bis-(trichloromethyl)carbonate (3.5 mmol) was added gradually for about 1 h until a clear solution was obtained; the reaction was then allowed to proceed for an additional 3 h. The reaction mixture was cooled down, about 80% of the THF was evaporated in vacuum, hexane was then added to the reaction mixture, and the quasi-racemate (*R*)-PheNCA/(*S*)-ThieNCA crystals were formed overnight at −5°C. The material was recrystallized from a mixture of dichloromethane with hexane (1:1 v/v) avoiding contamination by moisture and single crystals were analyzed by X-ray diffraction

Racemic (*R*,*S*)-ThieNCA crystals were prepared in a similar way and their structure determined.

CCDC-252855 and CCDC-252856 contain the supplementary crystallographic data for these structure. These data can be obtained free of charge via www.ccdc.cam.ac.uk/conts/retrieving.html.

Solid-state polymerization: Solid-state polymerization was performed in the following way: NCA monomer crystals (20 mg, 0.12 mmol) were suspended in hexane (50 mL), and *n*-butylamine (60 μL, 0.605 mmol) was then used as the initiator. The reactions were performed at room temperature with stirring for 72 h. After the completion of the reaction, the hexane was evaporated, and the solution was dried under high vacuum at room temperature. The gel-like or sometimes solid products were then prepared for MALDI-TOF mass spectrometry analysis. Polymerization of the melted monomer was performed by suspending the crystalline monomers in heptane previously heated to 100°C, followed by addition of *n*-butylamine. Attempted polymerization of the melted PheNCA in the absence of the initiator did not yield the oligopeptides.

Dissolution of the oligopeptides: Samples for MALDI-TOF MS analysis were prepared in the following way: the dried solid products (4 mg) were weighed into a 1.5 mL Eppendorf and a total of 10 extractions with dry THF were performed. The resulting extracted part (1.6 mg, 43.5%) was analyzed by MALDI-TOF MS as the THF-soluble fraction. The final insoluble part (2.4 mg, 56.5%) was dissolved with a mixture of trifluoroacetic anhydride and THF (3:1 v/v) until a clear solution was obtained and was analyzed by MALDI-TOF MS as the THF-insoluble fraction. The same mixture of trifluoroacetic anhydride (TFAA) and THF (3:1 v/v) was used for the analysis of the whole sample.

MALDI-TOF MS analysis—sample and matrix preparation: The matrix preparation was done in the following way: NaI (17 mg) was dissolved in THF (1 mL; solution I). Dithranol (6 mg) was dissolved in chloroform (0.99 mL) and trifluoroacetic acid (TFA, 10 μL; solution II). Solution I (125 μL) was mixed with solution II (125 μL). The final matrix solution was vortexed for 10 min at high speed. The best preparation for the MALDI-TOF MS analysis was achieved by the double deposit procedure: the matrix solution (0.5 μL) was deposited on the target holder, and the sample solution (0.5 μl) was then deposited on the matrix layer. 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 500 laser shots.

Characterization of the oligopeptides: The mass spectra of oligopeptides of a given length *n* allow study of the relative abundance of each diastereoisomeric oligopeptide, labeled *(m,p)* for ((*R*)-Phe, (*S*)-Phe) and *(m,n,p)* for ((*R*)-Phe, (*S*)-Phe, (*S*)-Thie). Indeed, the similar chemical properties and very close masses of these oligopeptides give similar ionization yields and detection efficiencies. However, a complex isotope pattern is to be expected due to the possible interference of many species. In particular, the repeating units (*u*) of (*R*)-Phe (147.068 *u* monoisotopic), (*S*)-Phe(d_5) (152.099 *u* monoisotopic), and (*S*)-Thie (153.0245 *u* monoisotopic), as well as the deuteration yield of (*S*)-Phe (≈98% estimated from the dimer mass spectrum) mean that a given peak *m*/*z* is composed of different components, except for the five first isotopes of the lighter oligopeptide (*n*,0,0). Thus, a home-made program (Visual Basic 6.0) was developed to take account of the isotopic pattern of each (*m*,*n*,*p*) oligopeptide and its interference with the other oligopeptides for the compositions of the monomers in binary or ternary mixtures. The total mass spectrum for a given length can be calculated and compared to the experimentally obtained mass spectrometry data by adjusting the abundance of each *(m,n,p)* component. The isotopic pattern of each oligopeptide can be determined with the Isopro2 (download at http://members.aol.com/msmssoft) isotopic pattern calculator by direct entry of molecular formula. The home-made program takes into account the deuteration yield, derivatization, and cationization in order to compare directly with experimental data. Typically, for the hexamer, 28 different oligopeptide isotope patterns, each composed of at least seven isotopes, have to be taken into account. This program is especially useful for verifying that a random association (binomial law) exists between two compounds ((*S*)-Phe and (*S*)-Thie).

Acknowledgements

We thank the Israel Science Foundation for financial support and Dr. Alla Shainskaya and her team from the Mass Spectrometry Laboratory and Dr. Linda Shimon from the Crystallography Laboratory at the Weizmann Institute. J.G.N. thanks the Hugh David Stone Foundation and the GMJ Schmidt Minerva Center for a postdoctoral Fellowship. This work is part of the COST D-27 program of the European Community.

[1] F. C. Frank, *Biochim. Biophys. Acta* **1953**, *11*, 459.
[2] G. Wald, *Ann. N. Y. Acad. Sci.* **1957**, *69*, 352.
[3] G. F. Joyce, G. M. Wisser, C. A. A. van Boeckel, J. H. van Boom, L. E. Orgel, *Nature* **1984**, *310*, 602.
[4] P. Franck, W. A. Bonner, R. N. Zare, in *Chemistry for the 21st Century*, (Eds.: E. Keinan, I. Schecter), Wiley-VCH Weinheim **2000**, p. 175.
[5] L. E. Orgel, *Nature* **1992**, *358*, 203.
[6] L. E. Orgel, *Origins Life Evol. Biosphere* **1998**, *28*, 91.
[7] A. Brack, *NATO ASI Ser., Ser. C*, Chirality and the Origins of Life, **1993**, pp. 345.
[8] J. R. Cronin, S. Pizzarello, *Science* **1997**, *275*, 951.
[9] S. Pizzarello, A. L. Weber, *Science* **2004**, *303*, 1151.
[10] M. Blocher, T. Hitz, P. L. Luisi, *Helv. Chim. Acta* **2001**, *84*, 842.
[11] T. Hitz, P. L. Luisi, *Helv. Chim. Acta* **2003**, *86*, 1423.
[12] M. Lahav, G. M. J. Schmidt, *J. Chem. Soc. B* **1967**, 312.
[13] B. S. Green, M. Lahav, G. M. J. Schmidt, *J. Chem. Soc. B* **1971**, 1552.
[14] L. Addadi, M. Lahav, *Pure Appl. Chem.* **1979**, *51*, 1269.
[15] M. Hasegawa, *Chem. Rev.* **1983**, *83*, 507.
[16] G. Wegner, *Makromol. Chem.* **1971**, *145*, 85.
[17] A. Matsumoto, T. Tanaka, T. Tsubouchi, K. Tashiro, S. Saragai, S. Nakamoto, *J. Am. Chem. Soc.* **2002**, *124*, 8891.
[18] T. Itoh, S. Nomura, T. Uno, M. Kubo, K. Sada, M. Miyata, *Angew. Chem.* **2002**, *114*, 4482; *Angew. Chem. Int. Ed.* **2002**, *41*, 4306.
[19] S. Nagahama, T. Tanaka, A. Matsumoto, *Angew. Chem.* **2004**, *116*, 3899; *Angew. Chem. Int. Ed.* **2004**, *43*, 3811.
[20] K. Cheng, B. M. Foxman, *J. Am. Chem. Soc.* **1977**, *99*, 8102.
[21] M. J. Vela, V. Buchholz, V. Enkelmann, B. B. Snider, B. M. Foxman, *Chem. Commun.* **2000**, 2225.

[22] H. Zepik, E. Shavit, M. Tang, T. R. Jensen, K. Kjaer, G. Bolbach, L. Leiserowitz, I. Weissbuch, M. Lahav, *Science* **2002**, *295*, 1266.
[23] I. Weissbuch, G. Bolbach, H. Zepik, E. Shavit, M. Tang, J. Frey, T. R. Jensen, K. Kjaer, L. Leiserowitz, M. Lahav, *J. Am. Chem. Soc.* **2002**, *124*, 9093.
[24] I. Weissbuch, G. Bolbach, H. Zepik, E. Shavit, M. Tang, T. R. Jensen, K. Kjaer, L. Leiserowitz, M. Lahav, *Chem. Eur. J.* **2003**, *9*, 1782.
[25] H. Kanazawa, Y. Ohashi, *Mol. Cryst. Liq. Cryst.* **1996**, *277*, 45.
[26] H. Kanazawa, H. Uekusa, Y. Ohashi, *Acta Crystallogr. Sect. C* **1997**, *53*, 1154.
[27] H. Kanazawa, *Acta Crystallogr. Sect. C* **2000**, *C56*, 469.
[28] J. G. Nery, G. Bolbach, I. Weissbuch, M. Lahav, *Angew. Chem.* **2003**, *115*, 2207; *Angew. Chem. Int. Ed.* **2003**, *42*, 2157.
[29] L. Addadi, E. Gati, M. Lahav, L. Leiserowitz, *Isr. J. Chem.* **1977**, *17*, 116.
[30] M. Smrcina, S. Vyskocil, V. Hanus, M. Polasek, V. Langer, B. G. M. Chew, D. B. Zax, H. Verrier, K. Harper, T. A. Claxton, P. Kocovsky, *J. Am. Chem. Soc.* **1996**, *118*, 487.
[31] Attempts to obtain independent information on the sequence of these oligopeptides by MALDI-TOF-TOF tandem mass spectrometry were performed. In spite of the low resolution for ion selection, preliminary results on ($n-1$,1) species gave indication that the *S* residue is not randomly distributed but lies at one end of the oligopeptide, in good agreement with the proposed model. New measurements are currently under investigation.
[32] N. H. Lee, C. F. Frank, *Langmuir* **2003**, *19*, 1295.
[33] M. T. Krejchi, E. D. T. Atkins, A. J. Waddon, M. J. Fournier, T. L. Mason, D. A. Tirrell, *Science* **1994**, *265*, 1427.
[34] H. Rapaport, K. Kjaer, T. R. Jensen, L. Leiserowitz, D. A. Tirrell, *J. Am. Chem. Soc.* **2000**, *122*, 12523.
[35] M. Bolli, R. Micura, A. Eschenmoser, *Chem. Biol.* **1997**, *4*, 309.
[36] V. Avetisov, V. Goldanski, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11435.
[37] J. S. Siegel, *Chirality* **1998**, *10*, 24.
[38] W. H. Daly, D. Poche, *Tetrahedron Lett.* **1988**, *29*, 5859.

Received: October 29, 2004
Published online: March 15, 2005