Homochiral Oligopeptides via a Lattice-Controlled Polymerisation in Racemic Crystals of Valine N-Carboxyanhydride Suspended in Aqueous Solutions

7 AUTHORS, INCLUDING:

Gerard Bolbach
Pierre and Marie Curie University - Paris 6
138 PUBLICATIONS   **2,083** CITATIONS

SEE PROFILE

Isabelle Weissbuch
Weizmann Institute of Science
113 PUBLICATIONS   **3,900** CITATIONS

SEE PROFILE

Meir Lahav
Weizmann Institute of Science
298 PUBLICATIONS   **9,725** CITATIONS

SEE PROFILE

Homochiral Oligopeptides via a Lattice-Controlled Polymerisation in Racemic Crystals of Valine *N*-Carboxyanhydride Suspended in Aqueous Solutions

Ran Eliash,[a] Jose G. Nery,[a] Irena Rubinstein,[a] Gilles Clodic,[b] Gérard Bolbach,[b] Isabelle Weissbuch,[a] and Meir Lahav*[a]

Dedicated to Professor David Reinhoudt on the occasion of his 65th birthday

Abstract: As part of our program on the biochirogenesis of homochiral peptides, we report the formation of racemic parallel (*p*) β sheets composed of alternating *R* and *S* chains of up to 14–15 repeat units of the same handedness through the polymerisation of (*R,S*)-valine *N*-carboxyanhydride (NCA) crystals suspended in aqueous solutions of a primary amine as the initiator. The occurrence of such a lattice-controlled reaction accompanied by a reduction in volume implies the operation of a mechanism that differs from that of the common solid-state polymerisation in vinyl systems. The topotacticity of the reaction is explained through the operation of a multistep nonlinear process comprising lattice control coupled with an asymmetric induction in the formation of homochiral short peptides followed by their self-assembly into racemic *p* β sheets, which operate as efficient templates in the ensuing process of enantioselective chain elongation at the polymer/crystal interface. The composition of the diastereoisomeric libraries of oligopeptides was determined by MALDI-TOF and MALDI-TOF-TOF MS analyses of the products obtained from monomers enantioselectively labelled with deuterium. The structure of the *p* β sheets could be determined by initiating the polymerisation reaction with water-soluble esters of enantiopure α-amino acids or short peptides. The same reaction performed with the monomer crystals suspended in hexane yielded a complex mixture of diastereoisomeric oligopeptides, thus highlighting the indispensable role played by water in controlling the stereoselectivity of the reaction. By contrast, polymerisation of (*R,S*)-leucine NCA crystals, with a different packing arrangement that presumably does not endorse the formation of periodic peptide templates, yielded, both in aqueous and hexane suspensions, libraries of peptides dominated by heterochiral diastereoisomers.

Keywords: amino acids · beta-sheet structures · biochirogenesis · lattice-controlled reactions · oligopeptides

Introduction

The origin of homochirality in the biological world provides an as yet unsolved riddle of prebiotic chemistry and early evolution. Of particular interest is the prebiotic formation of homochiral biopolymers from racemic monomers, such as the primeval polypeptides and polynucleotides composed from repeat units of the same handedness.[1–3] Homogeneous polymerisation of racemic monomers in organic solvents yields libraries of atactic polymers.[4] Under such conditions, the formation of isotactic chains of at least eight homochiral repeat units, commonly considered as necessary for the formation of α-helix or pleated-β-sheet templates to generate long homochiral peptides, is not possible.[4–7] Recently, we proposed a reaction pathway for the formation of isotactic peptides that can override this drawback. This pathway comprises the self-assembly of racemic monomers into crystalline motifs in which the two enantiomers are suitably arranged for a lattice-controlled polymerisation reaction between molecules of the same handedness.[8,9] This requirement, however, encounters intrinsic complications, as the
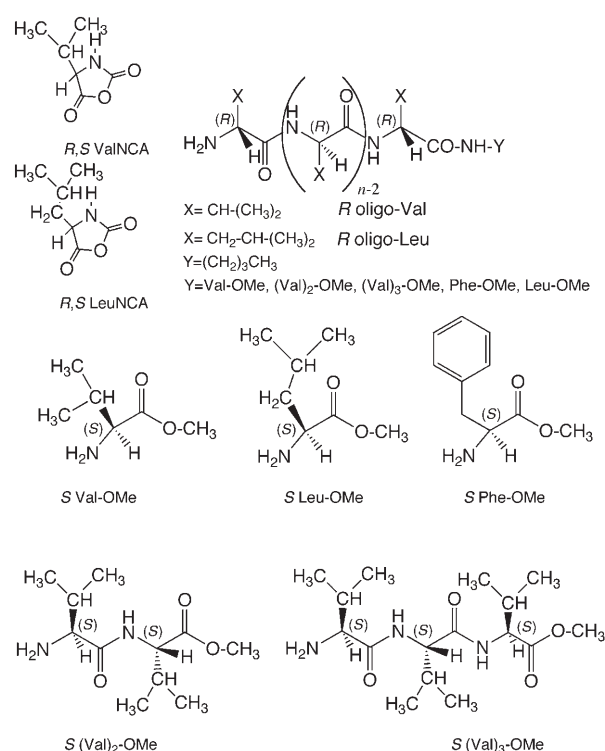
[a] Dr. R. Eliash, Dr. J. G. Nery, I. Rubinstein, Dr. I. Weissbuch,
Prof. Dr. M. Lahav
Department of Materials and Interfaces
The Weizmann Institute of Science
76100 Rehovot (Israel)
Fax: (+972) 8-934-4138
E-mail: meir.lahav@weizmann.ac.il

[b] G. Clodic, Dr. G. Bolbach
Platforme Spectrométrie de Masse et Protéomique
Université Pierre et Marie Curie
75252 Paris, Cedex 05 (France)

emergence of short polymers, which are denser than the crystalline monomer, is accompanied by damage of the periodic order of the latter at the sites where chain elongation proceeds. Note that only a limited number of systems including diethylenes,[10,11] diacetylenes[12] and nickel coordination complexes[13,14] were reported, during years of research, to undergo lattice-controlled polymerisation beyond dimers and trimers. As a result of this destruction, the longer polymers formed in the disordered environment comprise repeat units of both handedness. This disadvantage might be overcome if the short polymers of homochiral sequence (isotactic) would self-assemble into ordered architectures that could exert steric control on the ensuing process of chain elongation occurring at the polymer/crystal interface.[15–17] The operation of such interplay requires an understanding of the role played by the structure of the monomer crystal at the early stages of the reaction, its role in dictating the self-assembly of the formed peptides into supramolecular architectures, such as α helices or β sheets, and the ultimate role played by the latter as stereospecific or enantioselective templates in the ensuing stages of polymer propagation. Here, we provide support for this hypothesis for the solid-state polymerisation of racemic crystals of valine N-carboxyanhydride (ValNCA, Scheme 1) and leucine N-carboxyanhydride (LeuNCA) suspended in hexane or water solutions of a primary-amine initiator. Previous reports on the polymerisation of these crystals suspended in hexane have not described the diastereomeric distribution of the products and the mechanism of the reaction.[18–20]

Scheme 1. Chemical formulae of ValNCA, LeuNCA and their oligopeptides of homochiral sequence (isotactic), as initiated by either n-butylamine or methyl esters of various α-amino acids or short peptides.
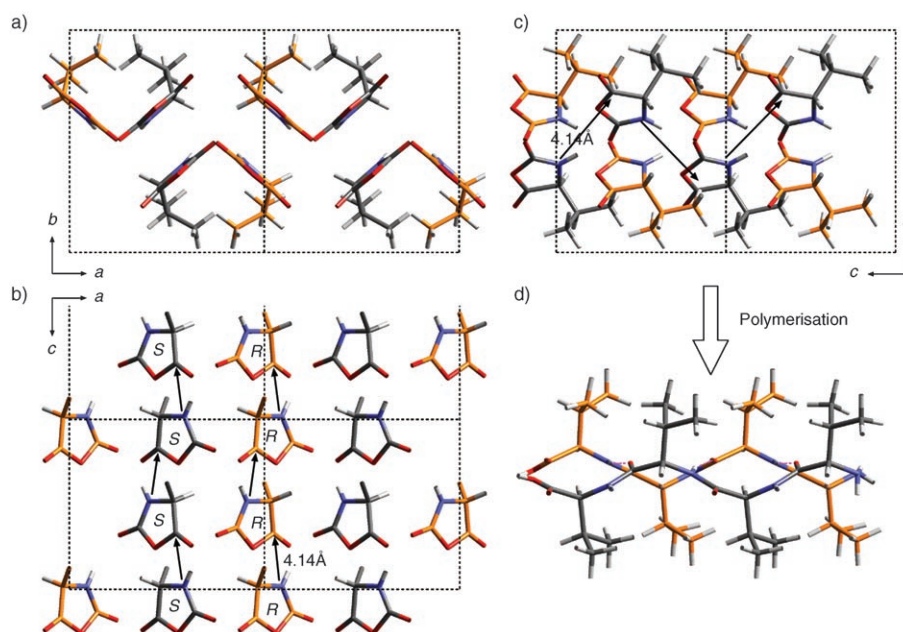
Results and Discussion

Polymerisation of (R,S)-ValNCA crystals suspended in hexane and initiated by n-butylamine: The racemic ValNCA crystal structure reported by Kanazawa and co-workers[21] appears in the orthorhombic polar space group $Pca2_1$ ($a = 8.938$, $b = 10.207$, $c = 7.788$ Å, $Z = 4$) with two R and two S molecules in the unit cell, as shown in Figure 1. The molecules are packed in a bilayer motif perpendicular to the b axis. The bilayers are composed of alternating parallel rows of R and S molecules with a single sense of polarity along the c axis. Within each row, the molecules are related by a twofold screw axis along the c direction, and alternating molecular rows are related to each other by a c glide plane perpendicular to the a axis.

Figure 1. a–c) The packing arrangement of (R,S)-ValNCA viewed along the c, b and a axes. d) Two adjacent isotactic R and S tetrapeptides viewed along the a axis, as resulting from the proposed preferred pathway of the polymerisation reaction along the polar c axis, shown in (c).

On the basis of this packing arrangement, one anticipates that a preferred geometric lattice-controlled reaction should take place along the c axis to produce racemic mixtures of

oligopeptides of homochiral (isotactic) sequence. The spacing that separates the N atom of one molecule from the αC-(carbonyl) atom of a favourably oriented neighbouring molecule of the same handedness along the c axis is 4.14 Å. Another, but less probable, pathway of polymerisation might be along the a axis where the spacing which separates the same N and C atoms of two adjacent unfavourably aligned heterochiral molecules is much longer at 4.88 Å. On the basis of these purely geometric considerations, we anticipated that the packing of the (R,S)-ValNCA crystal is appropriate for the generation of isotactic polymers through a reaction along the polar c axis.

Racemic crystals of (R,S)-ValNCA, in which the S enantiomer was enantioselectively tagged with eight deuterium atoms, were polymerised by suspending them in hexane containing n-butylamine as an initiator. The resulting diastereoisomeric oligopeptides were analysed by MALDI-TOF mass spectrometry and one such spectrum is shown in Figure 2a. Oligopeptides with up to 14–15 repeat units were detected.

In the MALDI-TOF mass spectrum, the oligopeptides of homochiral sequence (isotactic), labelled R_n and S_n, appear at the wings of each group of signals originating from various diastereoisomers of a given length n, in which the signals between the wings represent the oligopeptides that contain repeat units of both handednesses, labelled R_hS_d in which $h+d=n$. Previous studies have demonstrated that the molar fractions of the diastereoisomeric peptides of a given length are proportional to the relative intensity of their signals.[8,9] A histogram of the experimental molar fractions of the oligopeptides of each length and, for comparison, a histogram calculated with the assumption of a theoretical binomial distribution are shown in Figure 3a and b. It is clear that the distribution of the obtained diastereomeric oligopeptides differs substantially from the binomial one. According to the experimental results, racemic mixtures of homochiral oligopeptides were obtained. Moreover, for the tetra- and pentapeptides ($n=4$, 5), the isotactic chains are the dominant diastereoisomers (Figures 2a and 3a), in keeping with the lattice-controlled polymerisation envisaged above. However, the concentration of the heterochiral peptides, as compared to that of the isotactic ones, increases with the increase in length for $n=6$–15, presumably due to the lattice damage. Nonetheless, the departure from the binomial distribution is significant even for the long oligopeptides and it suggests the operation of an enantioselection process that results in the formation of detectable isotactic peptides of 14–15 repeat units.

Polymerisation of (R,S)-ValNCA crystals suspended in aqueous solutions of n-butylamine: A salient difference was found in the distribution of the diastereoisomeric oligopeptides obtained when the polymerisation reaction was performed on (R,S)-ValNCA crystals suspended in an aqueous solution of the initiator, under otherwise identical conditions. Optical photographs of crystals under such reaction conditions (Figure 4a and b) also show CO_2 bubbles. The distributions of the diastereoisomeric oligopeptides, as determined from the MALDI-TOF mass spec-
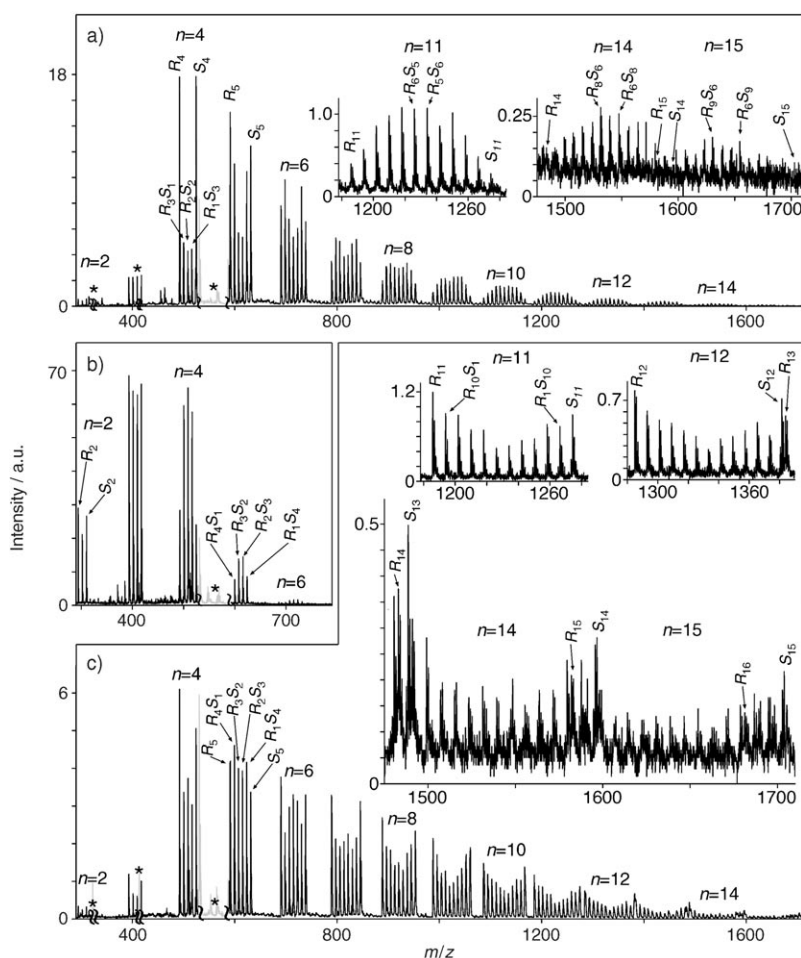

Figure 2. MALDI-TOF mass spectra of the products of polymerisation, initiated with n-butylamine, of (R,S)-ValNCA crystals suspended a) in hexane (insoluble product) and b) and c) in an aqueous solution of the initiator (water-insoluble product and water-soluble product, respectively). The signals corresponding to various diastereoisomers of each length n are labelled: Isotactic oligopeptides are labelled R_n and S_n. Heterochiral oligopeptides are labelled R_hS_d in which $n=h+d$. The *labels represent the m/z range of the matrix signals that were omitted for clarity.
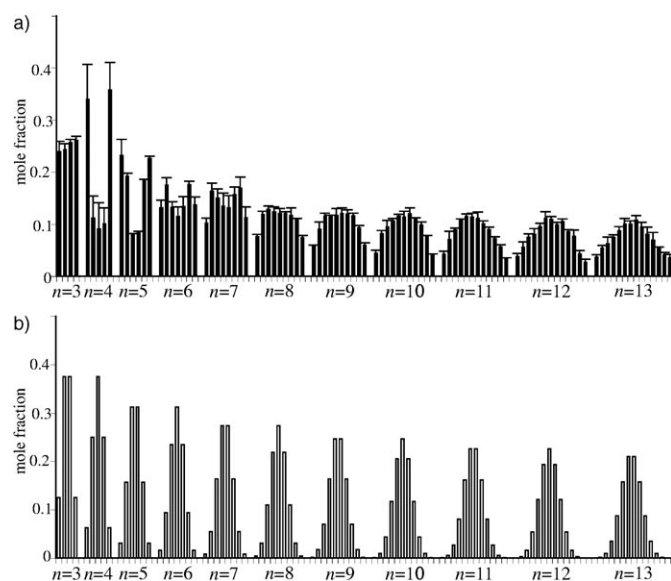
Figure 3. a) Histogram of the mole fraction of the diastereoisomeric oligopeptides of each length obtained in the polymerisation of (R,S)-ValNCA crystals suspended in hexane and initiated with n-butylamine. b) Histogram calculated by assuming a theoretical binomial distribution of the products.
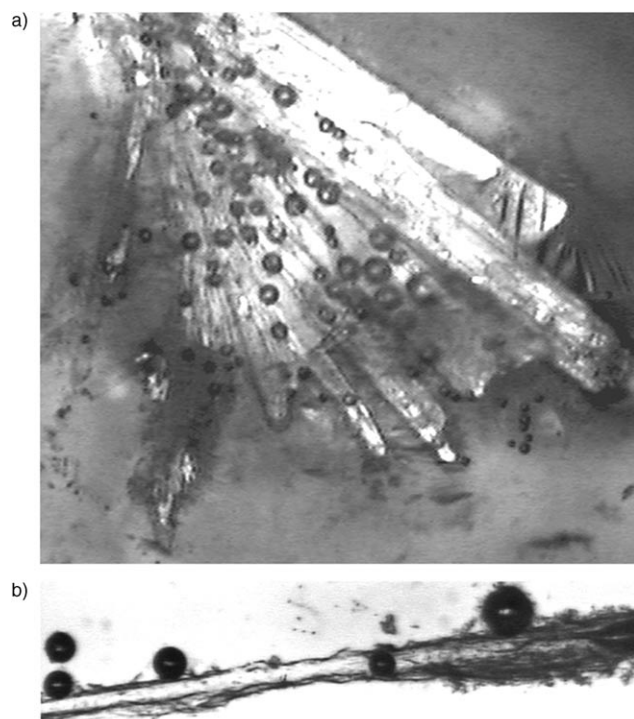

Figure 4. Optical photographs of (R,S)-ValNCA crystals suspended in an aqueous solution of n-butylamine initiator and showing CO_2 bubbles formed in the course of the reaction.

trum, are shown in Figure 2c, for the water-insoluble products.

In contrast to the polymerisation in hexane, when the reaction was performed on crystals suspended in an aqueous solution of n-butylamine, all homochiral di- and tripeptides

and heterochiral tetra- and pentapeptides were found to be soluble in water; the mass spectrum is shown in Figure 2b. The relative concentration of isotactic versus atactic water-insoluble tetra- and pentapeptides (Figure 2c) is lower than that for such peptides obtained in the hexane medium (Figure 2a). On the other hand, beyond hexa- or heptapeptides and up to the longest detectable ones of 15 repeat units, the isotactic peptides are dominant, as shown in the insets of Figure 2c in which these peptides are represented at the wings of the groups of signals corresponding to diastereoisomers of length $n = 11$–15 and labelled R_n and S_n.

On the basis of the geometrical considerations proposed above, we may envisage that if the solid-state polymerisation reaction takes place preferentially between homochiral monomer molecules along the c axis, the crystal lattice will force the short peptide chains to self-assemble into racemic parallel (p) β sheets composed of alternating oligo-R and oligo-S chains. The formation of β sheets is supported by FTIR spectra that display a characteristic stretching vibration at 1632 cm^{-1} and by X-ray diffraction patterns displaying d-spacings of 4.7 and 6.9 Å, a fingerprint for β-sheets and a 9.7 Å d-spacing corresponding to intersheet stacking. Further independent chemical evidence for the formation of racemic parallel β sheets is provided below.

Proposed polymerisation pathway: The formation of short isotactic oligopeptides as the dominant components, produced by suspending the monomer crystals both in hexane and in aqueous solutions of the initiator, implies that, in either of the suspending solvents, the early stages are controlled by the crystalline lattice of the monomer. There is a notable difference in the concentrations of the longer isotactic oligopeptides that are produced in hexane and in water, a difference that may be rationalised as follows. The smaller concentration of short isotactic peptides produced when the monomer crystals were suspended in aqueous solution might be due to either partial destruction of the monomer-crystal surfaces caused by minute NCA hydrolysis or by dissolution of the short atactic polymers in water. These two processes do not occur in hexane. On the other hand, the formation of increased concentrations of longer isotactic peptides when monomer crystals were suspended in aqueous solutions of the initiator implies that water induces the self-assembly of supramolecular templates that exert asymmetric control in the ensuing steps of the chain-elongation reaction. The solubility of the heterochiral short peptides in water increases the accumulation of isotactic oligopeptides that undergo self-assembly within the monomer crystal. A small amount of oligopeptides could arise from polymerisation of the solubilised monomer prior to the n-butylamine initiation reaction in the solid state. However, parallel studies of polymerisation in aqueous solutions to which the initiator was added five minutes after the complete monomer dissolution demonstrated the formation of racemic antiparallel (ap) β sheets.[4]

According to the packing arrangement of the (R,S)-ValNCA monomer crystal, the racemic isotactic peptides

should self-assemble into racemic *p* β sheets, although they are anticipated, on the basis of the results of polymerisation in aqueous solution,[4] to be energetically less stable than the ap β sheets. This self-assembly is illustrated in Figure 5, in
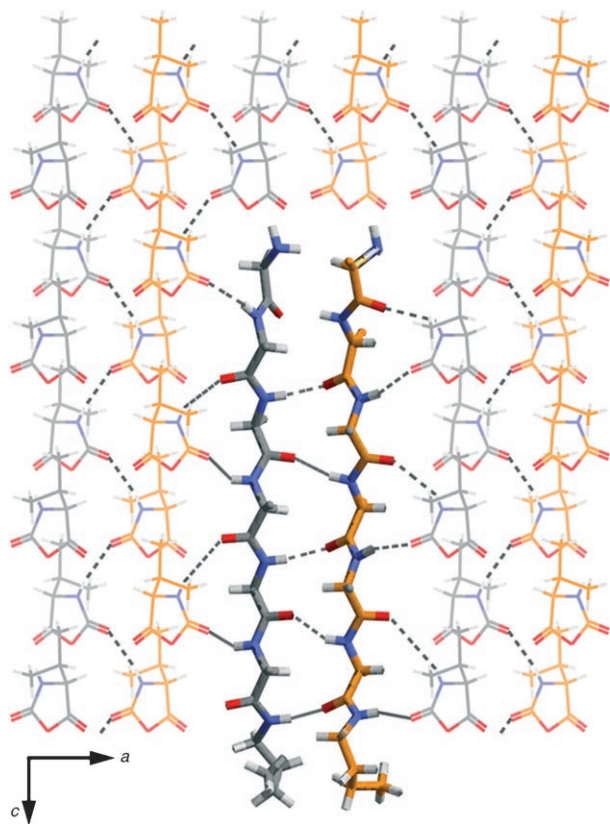

Figure 5. Computer model of two adjacent isotactic *R* and *S* hexapeptides assembled into a racemic parallel β sheet formed within the crystalline lattice of the (*R*,*S*)-ValNCA monomer. For clarity, the C atoms of the *R* monomer molecules and those of the *R* peptide are shown in orange; in addition, the isopropyl side chains of the two peptides are omitted.

which two parallel isotactic chains of opposite handedness can be accommodated in this architecture without introducing high strain along the *a* direction of the monomer lattice. The growing NH_2 groups, present at one edge of the sheet, are hydrated through hydrogen bonding with the water molecules. These groups react at the polymer/monomer interface with the hydrated chiral monomer molecules through the formation of diastereoisomeric transition states. In addition, at the locus of the propagation reaction, the released CO_2 molecules dissolve in water to form carbonic acid, which can interact with the NH_2 end groups to further modify the structure of the reaction site for chain elongation. On the other hand, when the reaction is performed with monomer crystals suspended in hexane, either the β sheets are not formed or, if they do form, the propagation reaction between a presumably nonsolvated peptide end and a nonsolvated NCA monomer molecule at the partially destroyed polymer/NCA crystal interface engenders a lower stereospecificity.

Polymerisation of (*R*,*S*)-ValNCA crystals suspended in aqueous solutions of enantiopure α-amino acid esters as chiral initiators: To further confirm the formation of racemic *p* β sheets, we performed the polymerisation reactions with such initiators that should differentiate between the formation of the various possible architectures, racemic *p* or *ap* β sheets and enantiomorphous pleated β sheets, as a result of their spontaneous separation. If racemic *p* β sheets are formed, an initiator resembling a peptide repeat unit that is linked at the C terminus of the chain should reside far away from the growing NH_2 end and should not interfere with chain elongation of the isotactic peptides. Therefore, we anticipate that the enantiomeric excess (*ee*) of the formed long oligopeptides should remain constant as a function of peptide length. Such behaviour should be at variance with that for *ap* β sheets, be they racemic or pleated. In these architectures, the initiator repeat unit should reside in close proximity to the growing NH_2 group of an adjacent chain and should induce steric hindrance in the growth process. Consequently, such an impediment on the growing chains will result in desymmetrisation of the racemic mixtures of the long peptides. In the case of racemic *ap* β sheets, isotactic chains of the same absolute configuration as the initiator will be formed in lower concentration.[17] If enantiomorphous pleated sheets are formed, those composed of isotactic oligopeptides of the same absolute configuration as the initiator will be formed in excess.[4]

The MALDI-TOF mass spectra of the water-soluble and water-insoluble diastereoisomeric peptides generated in the polymerisation reaction of (*R*,*S*)-ValNCA crystals suspended in aqueous solutions containing the (*S*)-Val-OMe initiator are shown in Figure 6a and b. Oligopeptides comprising up to 14 repeat units were detected and we noticed that the water-soluble isotactic di- and tripeptides formed in excess are of the same absolute configuration as that of the initiator (Figure 6a). On the other hand, beyond the tetrapeptides (Figure 6b), there is a reversal in sign of the *ee* value of the isotactic oligopeptides and this *ee* value remains almost unchanged for the longer peptides, as shown in Figure 7a. Polymerisation with the (*R*)-Val-OMe initiator yielded a mirror-image behaviour (Figure 7a). Similar results were obtained when the polymerisation reaction was performed with various initiators, including methyl esters of phenylalanine (Figures 7b and 8), tryptophan or the valine di- and tripeptides, (Val)₃-OMe and (Val)₃-OMe (Figure 9).

The following possible effects have been considered for the rationalisation of the observed asymmetric induction in the initiation steps of the reaction; these effects would result in a reversal in the *ee* value of the oligomers beyond the dimer stage: 1) Preferred recognition of the enantiopure initiator by one of the heterochiral monomer sites at the crystal/solution interface. 2) A steric effect induced by the enantiopure initiator at the crystal/solution interface that enantioselectively hampers the interactions with the monomer site of the same handedness. 3) Preferred solubility of the short diastereoisomers $I_S S_n$ relative to that of $I_S R_n$ in water, which would increase the concentration of the major diaste-
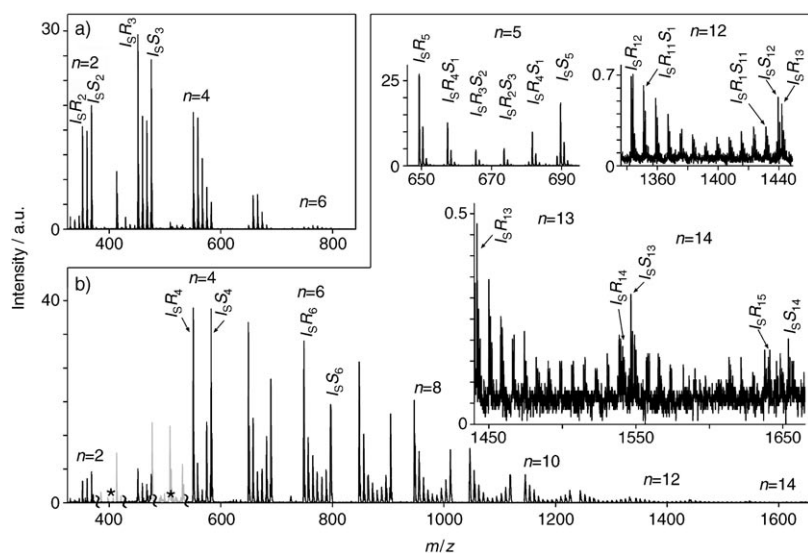
Figure 6. MALDI-TOF mass spectra of the product of polymerisation of (R,S)-ValNCA crystals suspended in an aqueous solution of (S)-Val-OMe: a) water-soluble material, b) water-insoluble material. Note that to highlight the signals that belong to the oligopeptides, those originating from the matrix are marked with ∗ labels.
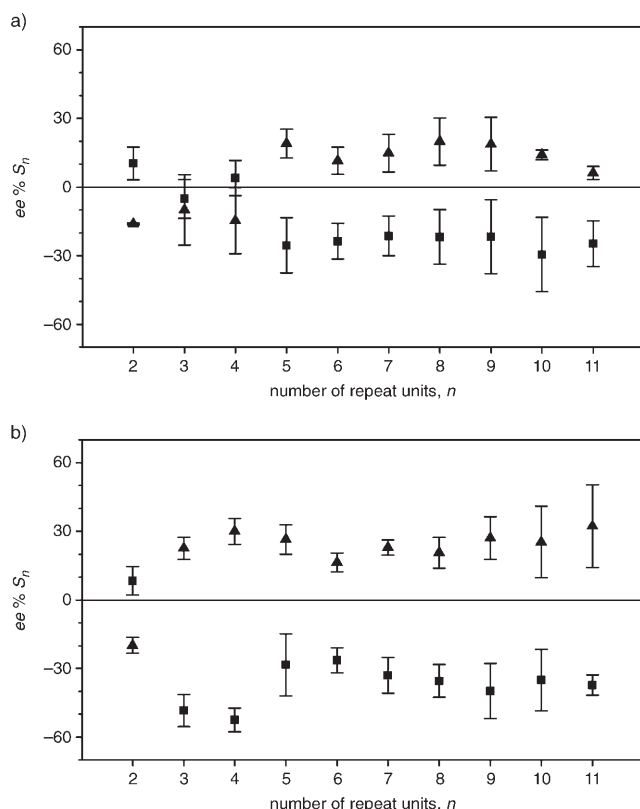

Figure 7. Plots of the *ee* value of $S_n = 100([S]-[R])/([S]+[R])$ for isotactic oligopeptides of each length *n* obtained in the polymerisation of (R,S)-ValNCA crystals suspended in an aqueous solution of enantiopure initiator (50 mol % mol^{-1}): a) Val-OMe and b) Phe-OMe of *S* absolute configuration (■) and *R* absolute configuration (▲). We represent the excess of oligo-*S* chains as an *ee* value instead of a diastereomeric excess, *de*, value by not considering the repeat unit of the initiator at the C terminus, for convenience. Error bars indicate the average of five samples.

reoisomer at the crystal surface. 4) A change in conformation of the *R* and *S* repeat units of the short diastereoisomeric peptides to favour different propagation pathways.

Some insight into the initiator recognition by the heterochiral crystal sites was obtained from a computational study of the energy required for an (S)-Val-OMe initiator molecule to dock at the two enantiomeric sites on various crystal surfaces. Single crystals of (R,S)-ValNCA are (010) plates elongated along the *c* axis and delineated by (111), ($\bar{1}$11) and (110) side faces, as shown in Figure 10. Although the (010) and (0$\bar{1}$0) faces are the most expressed, they are of lower relevance for our purpose as the initiator molecules that dock at these faces will not be able to initiate the polymerisation reaction along the polar $-c$ direction due to an inappropriate orientation of the NH$_2$ group of the initiator with regard to the αC=O of the ValNCA molecule. On the other hand, the two less-expressed enantiotopic side faces (111), (11$\bar{1}$), ($\bar{1}$11) and ($\bar{1}$1$\bar{1}$) are more relevant to the polymerisation process as they comprise sites of both handedness, and binding of an (S)-Val-OMe initiator, I_S, to such crystal sites followed by initiation of the polymerisation reaction leads to the formation of diastereomeric intermediates. The lowest docking energy of an I_S molecule to a ValNCA site of any handedness on the two cleaved crystal surfaces was obtained, as expected, for an *S* site, with docking being somewhat more favourable on the ($\bar{1}$11) face than on the (111) face. This result is in keeping with the formation of an excess of short peptides of the same absolute configuration as the initiator, for example, $[S_2I_S] > [R_2I_S]$ or, by analogy, $[R_2I_R] > [S_2I_R]$. However, the docking energy of an I_S molecule into an *R* site, in one out of four possible sites on the same crystal faces, is larger by only 1 kcal mol^{-1}, a fact implying that formation of both the SI_S and RI_S diastereoisomers is possible. The result of this computation can provide a rationale for the observed reversal in the *ee* value of longer isotactic peptides.

The possible steric effect of the initiator repeat unit in the two diastereoisomeric dipeptides could be ignored as a similar reversal in the *ee* value was obtained by using various initiators (Figures 7 and 9).

The analysis of the water-soluble short peptides by MALDI-TOF MS indicates that there is not a substantial difference, if any, in the degree of solubility of the diastereoisomeric peptides I_SS_2, I_SR_2 and I_SS_3, I_SR_3 as the relative compositions in the water-soluble product mixture (Figure 6a) and in the water-insoluble product mixture (Fig-
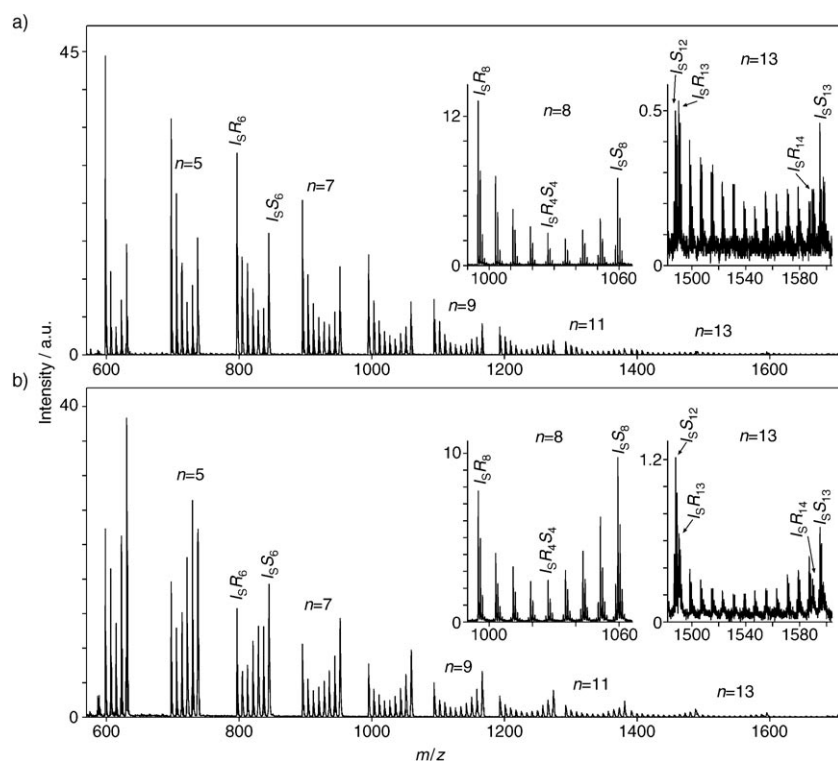
Figure 8. MALDI-TOF mass spectra of the diastereoisomeric peptides obtained from the polymerisation of (R,S)-ValNCA crystals suspended in aqueous solutions of a) (S)-Phe-OMe and b) (R)-Phe-OMe initiator.
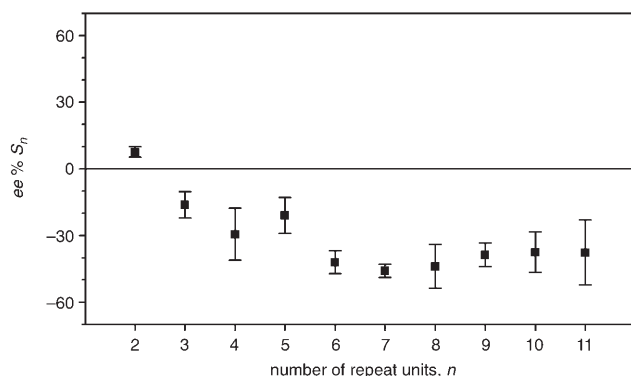

Figure 9. Plot of the S_n ee value for isotactic oligopeptides of each length n obtained in the polymerisation of (R,S)-ValNCA crystals suspended in an aqueous solution (with 85 mol%) of (S)-(Val)$_3$-OMe initiator.

groups of the adjacent S or R monomer molecules in the crystal will be different. To obtain additional insight into the reaction pathways at the early propagation stages, we analysed the sequences of the heterochiral oligopeptides by MALDI-TOF-TOF MS/MS.

MALDI-TOF-TOF MS/MS analysis of the heterochiral peptides: The sequence of the various diastereoisomeric heterochiral peptides obtained from the polymerisation of (R,S)-ValNCA crystals suspended in an aqueous solution of the initiator (S)-PheOMe were studied by MS/MS (see the Experimental Section). The fragmentation analysis allowed us to determine the most abundant sequences (MAS) of the obtained diastereoisomers, as summarised in Table 1. From these results, we noticed that all heterochiral peptides are composed from blocks of repeat units of the same handedness. When I_S is used as the initiator, the MAS in the family (h,1) (in which h represents the number of repeat units of R configuration (protonated) and 1 is the deuterated S repeat unit) is either $(R)_hSI_S$ or $S(R)_hI_S$, in which the single S unit resides either at the N terminus and operates as a terminator or is linked directly to the I_S initiator. By contrast, for the family (1,d) (in which d represents the number of repeat units of S con-

ure 6b) are the same. Furthermore, the solubility of I_SS_4 is lower than that of I_SR_4; the concentration of the former is higher in the solid phase. If differences in solubility were responsible for the reversal in sign of the ee value, the longer peptides would be enriched with I_SS_n, which is at variance with the experimental result (Figure 7).

We are thus left with the fourth possibility mentioned above, which considers the removal of the mirror symmetry of the R and S monomer molecules located at the crystal/solution interface after formation of the diastereoisomers I_SS and I_SR. As a result, the contact of the growing NH$_2$ moieties of these two diastereoisomers with the reactant carbonyl
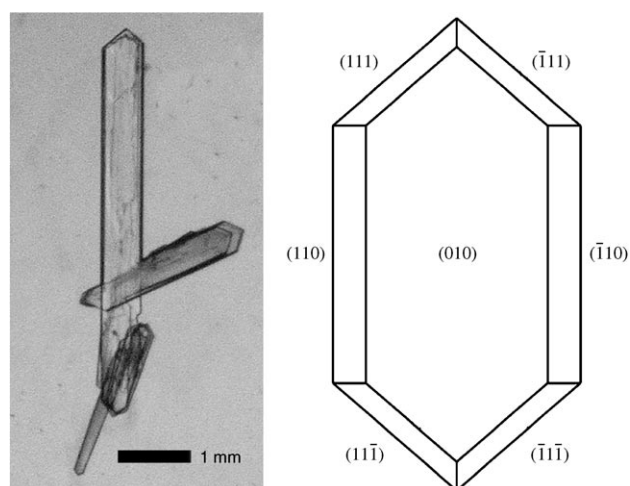

Figure 10. Photograph of crystals grown from THF/hexane and the computer-calculated growth morphology of (R,S)-ValNCA crystals.

Table 1. The most abundant sequences (MAS) for the heterochiral oligo-peptides obtained from the polymerisation of (R,S)-ValNCA suspended in an aqueous solution of (S)-Phe-OMe (I_S) initiator as determined by MALDI-TOF-TOF MS/MS. The most abundant sequences appear in bold.

(h,d)	MAS	(h,d)	MAS
(3,1)	$SRRRI_S$[a]	(7,1)	$RRRRRRSI_S$, $SRRRRRRRI_S$
(2,2)	$RRSSI_S$	(6,2)	**$RRRRRRSSI_S$**, **$SSSRRRRRRI_S$**, $RSRRRRRRI_S$
(1,3)	$RSSSI_S$, $SRSSI_S$	(5,3)	**$RRRRRSSSI_S$**, $SSSRRRRRRI_S$, $SRRRRRSSSI_S$
(4,1)	**$RRRRSI_S$**, $SRRRRI_S$	(4,4)	**$RRRRRSSSI_S$**, $RRRSRSSSI_S$
(3,2)	$RRRSSI_S$	(3,5)	$RRRSSSSSI_S$, $RRSRSSSSI_S$, **$SSSSSRRRI_S$**
(2,3)	**$RRSSSI_S$**, $SSSRRI_S$	(2,6)	$RSSSSSSSI_S$, **$SSSSSSRRI_S$**
(1,4)	$RSSSSI_S'$	(1,7)	$RSSSSSSSI_S$
(5,1)	**$RRRRRSI_S$**, $SRRRRRI_S$	(8,1)	**$RRRRRRRSI_S$**, $SRRRRRRRRI_S$
(4,2)	**$RRRRSSI_S$**, $SSRRRRI_S$	(7,2)	**$RRRRRRRSSI_S$**, $SSRRRRRRRI_S$
(3,3)	**$RRRSSSI_S$**, $SSSRRRI_S$	(6,3)	**$RRRRRRSSSI_S$**, $SRRRRRRSSI_S$, $SSRSRRRRRI_S$, $SSSRRRRRRI_S$
(2,4)	**$RRSSSSI_S$**, $SSSSRRI_S$	(5,4)	**$RRRRRSSSSI_S$**, $RSSSSRRRRI_S$, **$SSSSRRRRI_S$**
(1,5)	$RSSSSSI_S$	(4,5)	**$RRRRSSSSSI_S$**, $SRRRSRSSSI_S$
(6,1)	$RRRRRRSI_S$, $SRRRRRRI_S$	(3,6)	**$RRSSSSSSI_S$**, $SRRRSSSSSI_S$, **$SSSSSSRRRI_S$**
(5,2)	$RRRRRSSI_S$	(2,7)	**$RRSSSSSSSI_S$**, $SSSSSSSRRI_S$
(4,3)	**$RRRRSSSI_S$**, $SSSRRRRI_S$	(1,8)	$RSSSSSSSSI_S$
(3,4)	**$RRRSSSSI_S$**, $SRRRSSSI_S$, $RRSRSSSI_S$		
(2,5)	**$RRSSSSSI_S$**, $SRRSSSSI_S$, $SSSSSRRI_S$		
(1,6)	$RSSSSSSI_S$		

[a] For example, the sequence $SRRRI_S$ stands for NH_2-(S)-Val-(R)-Val-(R)-Val-(R)-Val-I_S.

of either handedness. On the other hand, RI_S reacts with monomers of the same handedness along the translation c axis.

On the basis of these results, we infer that the isotactic oligo-R chains should be formed in excess as the SI_S dipep-tide proceeds along two propagation pathways and reacts with either S or R monomers, thus decreasing the concentra-tion of the oligo-S chains (Figure 11). This deduction sup-ports the observed reversal in the ee value at the early stages of the polymerisation.

To summarise the reaction mechanism with the enantio-pure initiator in these polymerisation experiments, we may trace the following sequential steps. First is an asymmetric induction of the enantiopure initiator with either an R or an S monomer molecule at the crystal/liquid interface to form two diastereoisomeric dipeptides that react differently with their adjacent neighbours in the crystal to yield a nonrace-mic composition of the short dipeptides. Once short isotactic oligopeptides are formed, they assemble into racemic p β sheets, as dictated by the crystalline lattice of the monomer. The operation of asymmetric induction in the initiation step induces a desymmetrisation of the racemic mixtures of the formed isotactic oligopeptides, in spite of their parallel ori-entation in which the enantiopure initiator is located far away from the NH_2 growing groups. The ensuing chain elon-gation proceeds with almost equal rates for the two enantio-meric chains, thereby leaving the ee value unchanged as function of length. This process occurs at the growing sheets/crystal interface presumably through a diffusion mechanism of the monomers towards the growing rims of

figuration (deuterated) and 1 is the protonated R repeat unit), the MAS with $d>5$ are $R(S)_dI_S$ in which the single R repeat unit is found only at the N terminus. The diastereoisomer with the R repeat unit at-tached to the initiator was not detected. This result suggests that, whereas the RI_S dipeptide reacts exclusively with R repeat units, the SI_S dipeptide reacts with both S and R mo-nomer molecules; this implies that, once the latter dipeptide is formed, it assumes a confor-mation in which the S repeat unit is dislocated from its origi-nal site in the crystal, such that its NH_2 group can react with adjacent monomer molecules

Figure 11. Computer models of the propagation steps during polymerisation of (R,S)-ValNCA crystals initiated with (S)-Val-OMe, I_S: a) Diastereomeric RI_S and SI_S dipeptides, the former with only one possible propagation pathway along the R monomer columns and the latter with two possibilities, yielding b) diastereomeric RRI_S, SSI_S and RSI_S tripeptides. For clarity, the C atoms of the R monomers are coloured in orange.

the peptides. From this stage on, the stereospecificity of the reaction is controlled primarily by the structure of the growing rims of the oligopeptides. As the composition of the long peptides is very different when the reaction is performed by suspending the crystals in hexane, the hydration of both the monomer and the growing rims is imperative for the formation of the isotactic peptides.

To cast additional insight on the role played by the crystalline lattice in the distribution of the diastereoisomeric oligopeptides, we investigated the solid-state polymerisation of racemic crystals of LeuNCA.

Polymerisation of (R,S)-LeuNCA crystals suspended in water and hexane: The packing arrangement of (R,S)-LeuNCA crystals reported by Kanazawa[22] (monoclinic, $a = 9.878$, $b = 5.640$, $c = 15.448$ Å, $\beta = 105.844°$, space group $P2_1/c$, $Z = 4$) is shown in Figure 12 as viewed along the b and a axes.

The molecules form bilayer structures that are parallel to the (001) plane of the crystal. Each bilayer is composed of antiparallel ordered alternating rows of R and S molecules along the unique b axis. Rows of R or S molecules, related by a twofold axis parallel to the b axis, form the bilayers. The S molecules are related to the R molecules by a c glide
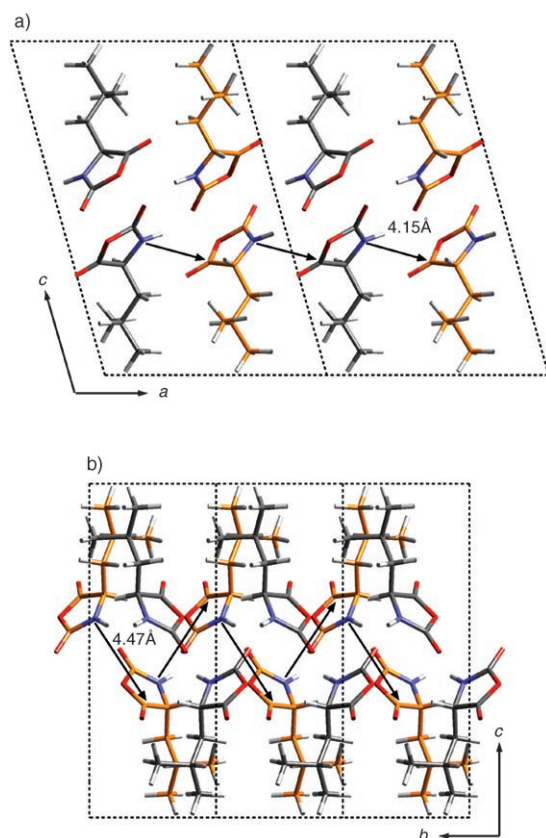

Figure 12. Packing arrangement of (R,S)-LeuNCA viewed along a) the b axis and b) the a axis. Black arrows indicate feasible reaction pathways in the early stages. The short contacts between the reactive atoms are marked. All of the C atoms of molecules of R absolute configuration are coloured in orange.

perpendicular to the b axis. Should the reaction proceed along the b axis, the S molecules will react in the $+b$ direction and the R molecules in the opposite $-b$ direction. The spacing between a nitrogen atom of one molecule and the αC(carbonyl) atom of a neighbouring molecule of the same handedness along the b axis is 4.47 Å. Reactivity along this direction should produce racemic ap β sheets. However, according to the packing arrangement of this crystal, there is another polymerisation pathway along the a axis to produce a syndiotactic polymer composed from alternating R and S repeat units. Furthermore, during the chain elongation process, the reaction may proceed along either pathway to yield atactic polymers.

Polymerisation experiments were performed on racemic crystals, in which the S enantiomer was enantioselectively tagged with ten deuterium atoms, suspended both in hexane and in water containing the n-butylamine initiator. The diastereoisomeric distributions of the oligopeptides for both suspending liquids are shown in Figure 13. (R,S)-LeuNCA crystals, which are insoluble in hexane, yield a complex library dominated by the long heterochiral peptides (Figure 13b). When the same reaction was performed in water containing the n-butylamine initiator, the isotactic isomers were more abundant in the formed oligopeptides (Figure 13a) in comparison to those obtained in hexane but were less abundant than those obtained from the (R,S)-ValNCA crystals in water. Such a distribution is in keeping with the double reaction pathway that is possible in the (R,S)-LeuNCA crystal, which makes it difficult to engender the β-sheet templates required for the formation of the long homochiral peptides. However, water enhances, in comparison to hexane, the formation of the isotactic peptides, as a result of the hydration of the monomer and of the growing ends of the chains, as well as due to the higher solubility of the atactic short peptides than that of the isotactic ones. In this system, we cannot ignore the possibility that a small number of the isotactic peptides were formed in the aqueous solution due to some dissolution of the monomer.[4]

Conclusion

In this study, we demonstrate the formation of isotactic oligopeptides of valine containing up to 14–15 repeat units of the same handedness by starting with (R,S)-ValNCA crystals suspended in aqueous solutions containing an amine initiator. This reaction is stereospecific, in contrast with previously reported nonstereospecific solid-state polymerisation reactions,[23,24] although the polymeric phase is denser than the monomer crystal. This different behaviour can be rationalised by assuming a multistep mechanism comprising 1) lattice control in the early stages of the reaction to yield short racemic oligopeptides enriched with the isotactic ones and 2) self-assembly of the latter into supramolecular p β sheets followed by 3) a chain-elongation step in which the monomer molecules diffuse towards the structured rims of the newly formed polymeric phase. The stereospecificity of the
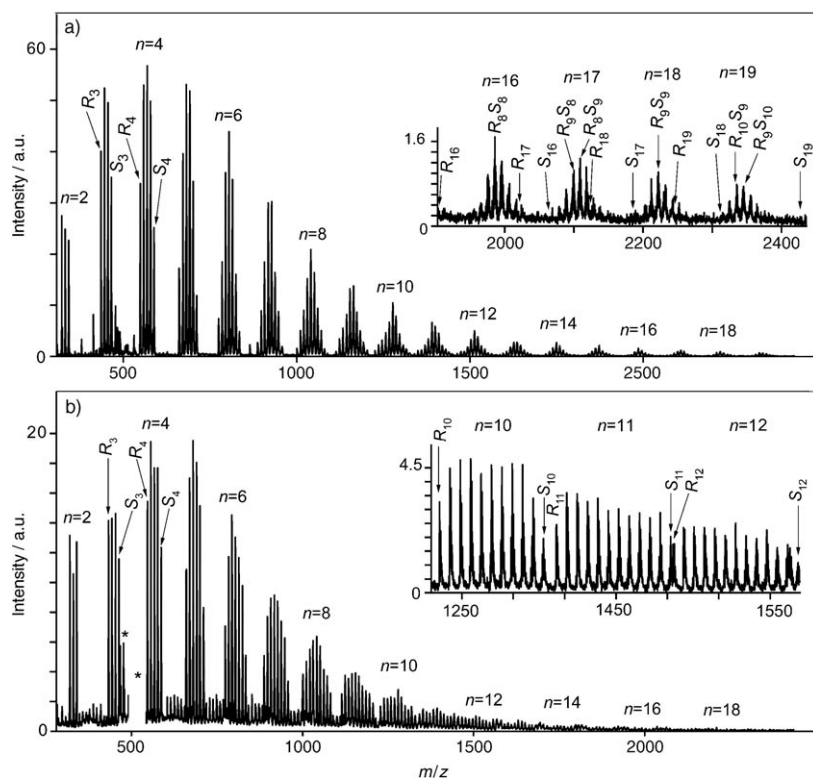
Figure 13. MALDI-TOF mass spectra of diastereoisomeric oligopeptides obtained from the polymerisation of (R,S)-LeuNCA suspended in a) aqueous and b) hexane solutions of n-butylamine.

chain-elongation process is defined by the structures of the rims of the supramolecular p β sheets that operate as templates. The monomer crystal structure also defines the mode of organisation of the peptide chains as racemic β sheets. The templates may be either parallel, as described here for ValNCA, or antiparallel as reported in a previous study of the polymerisation of (R,S)-PheNCA.[16,17] The packing arrangement of (R,S)-LeuNCA presumably does not permit the formation of either architecture as the isotactic peptides are not formed in substantial quantities from this starting material.

The present study also demonstrates that the suspending solvent plays an important role in determining the diastereoisomeric distribution of the oligopeptides. At least two effects played by water have to be taken into consideration: firstly, the ability of the hydrophobic peptide chains to self-assemble preferentially in an aqueous millieux as compared to hexane and, secondly, the formation of chiral hydrated regions at the growing NH_2 groups that increase the enantiomeric recognition for the hydrated monomers.

The operation of the multistep mechanism proposed here was instrumental in the generation of homochiral peptides from amphiphilic activated α-amino acids polymerised in aqueous solutions.[4] These studies might be of relevance to the formation of homochiral primeval peptides as NCA activation of α-amino acids is currently considered as a relevant precursor step.[25–28]

Experimental Section

Materials: ValNCA and LeuNCA monomers were prepared according to the procedure of Daly and Poché,[29] with minor adjustments.

(R,S[D₈])-ValNCA: Finely ground (R)-Val (117.2 mg, 1 mmol, 0.5 equiv) and (S[D_8])-Val (98%, Cambridge Isotope Laboratories; 125.2 mg, 1 mmol, 0.5 equiv) were suspended in dry THF and heated to reflux under argon. Solid bis(trichloromethyl)carbonate (triphosgene; 237.4 mg, 0.8 mmol, 1.2 equiv) was added and reaction was continued for 0.5 h. Triphosgene portions (0.4 mmol, 0.6 equiv) were consecutively added every 0.5 h until a clear solution was obtained (usually two or three portions were needed). The reaction was then continued for one additional hour. The clear solution was allowed to cool to room temperature and the THF was evaporated under reduced pressure. CH_2Cl_2 (≈ 5 mL) was added, followed by hexane (≈ 50 mL), with stirring. A white precipitate appeared almost immediately and the suspension was cooled to 4°C for 2–3 h, after which the crystals were filtered off. The small needle-like (R,S[D_8])-ValNCA crystals (272.9 mg, 93% yield) were analysed by IR spectroscopy. All other NCAs were prepared in a similar manner.

Methyl esters of amino acids and peptides: (R)- and (S)-Valine methyl ester (Val-OMe) hydrochloride, (R)- and (S)-leucine methyl ester (Leu-OMe) hydrochloride, (R)- and (S)-phenylalanine methyl ester (Phe-OMe) hydrochloride and (R)- and (S)-tryptophan methyl ester (Trp-OMe) hydrochloride were purchased from Sigma and the corresponding free amines were prepared by suspending the material in dry CH_2Cl_2 and bubbling ammonia gas through the suspension for a few minutes. The resulting ammonium chloride was filtered off and the CH_2Cl_2 was removed by evaporation.

The methyl ester hydrochlorides of di- and tripeptides ($(Val)_2$, $(Val)_3$, and $(Leu)_2$; Sigma or Bachem) were prepared as follows: Dry methanol (20 mL) was placed in a 50 mL round-bottomed flask equipped with a drying tube and a magnetic stirring bar and acetyl chloride (3.2 mL) was added dropwise. The solution was allowed to cool to room temperature and, for example, the $(Val)_3$-OMe, $(Val)_3$-OH tripeptide (500 mg, 1.58 mmol, 1 equiv) was added. The reaction was monitored by TLC (30% aqueous NH_3:isopropyl alcohol 3:7) and was continued for two to three days. After completion of the reaction, the MeOH was evaporated under vacuum, then CH_2Cl_2 was added and evaporated; this procedure was repeated until the remaining material was transformed from an oil into a white solid. Hexane was added and evaporated several times until a loose white precipitate appeared. The material was filtered off to give $(Val)_3$-OMe·HCl (538 mg, 92% yield). The hydrochloride salts were recrystallised from CH_2Cl_2/hexane at 4°C overnight.

Polymerisation procedure: The appropriate NCA material was weighed into a glass vial or polypropylene centrifuge tube (1–2 mg, or 10 mg when more material was needed for analyses). Stock solution (0.75 mL for each milligram of NCA) of the appropriate free amine in water or hexane was added and the reaction was stirred for 24–72 h. The stock solutions were prepared so that the amount of free amine added will equal 0.85 mol equivalents of the amount of NCA. Some of the peptide methyl esters did not dissolve completely in the hexane and a well-stirred sus-

pension was used. After 24–72 h, the vessels were subjected to centrifuge and the solvent was carefully decanted and filtered to remove any residual precipitate. The remaining precipitate and the solution were both dried, hexane was removed by a stream of dry nitrogen and water was removed by lyophilisation.

MALDI-TOF mass spectroscopy measurements and analyses:

Sample preparation: A small amount of the polymerisation product (approximately 0.05–0.1 mg) was placed in a 500 μL polypropylene microcentrifuge tube and trifluoroacetic acid (TFA, 20 μL) and THF (80 μL) were added in succession to obtain clear solution I. A matrix solution (0.5 mL, dithranol solution in chloroform/saturated NaI solution in THF 1:1) was placed on the target plate and left to dry; on top of this, solution I (0.5 μL) was deposited (double deposition).

The MALDI-TOF positive-ion mass spectra were obtained in reflector mode from a Bruker Reflex III instrument equipped with an N_2 laser. Only singly charged ions, $[M+Na]^+$, with the expected isotopic pattern were observed. The final mass spectra resulted from the signal average of at least a few hundred laser shots in different spots of the target, to ensure reliable statistics about the ion peak. Good statistics were obtained when the isotopic distribution of an ion species corresponded to that expected from calculation. Mass assignments were made by using both the m/z value and the isotopic distribution.

The following notation code is used for the peptide molecules: (h,d) designates all molecules composed of h repeat units of R configuration (protonated) and d repeat units of S configuration (deuterated), with $n = h + d$ being the total number of repeat units. The relative abundance (ra) of each type of oligopeptide (h,d) was obtained by dividing the intensity of the signals from a particular molecule $I(h,d)$ by the total intensity of the ions from all of the molecules of the same length, n. For example, the relative abundance of the tetrapeptide (4,0), composed of four (R)-Val repeat units, was calculated from Equation (1).

$$ra(4,0) = I(4,0)/\{I(4,0)+I(3,1)+I(2,2)+I(1,3)+I(0,4)\} \quad (1)$$

The relative abundance of the tetrapeptide (1,3), composed of one R (protonated) and three S (deuterated) Val repeat units, is calculated from Equation (2).

$$ra(1,3) = I(1,3)/\{I(4,0)+I(3,1)+I(2,2)+I(1,3)+I(0,4)\} \quad (2)$$

The ionisation yield is expected to be very similar for (h,d) oligopeptides of the same length $n = h + d$ due to their similar chemical properties and detection efficiency. Thus, the ion intensities of the different (h,d) diastereoisomeric oligopeptides of the same length n are directly and reliably comparable.

MALDI-TOF-TOF mass spectra: The spectra were obtained with an Applied Biosystems 4700 Proteomics MALDI-TOF-TOF mass spectrometer fitted with an NdYag laser operating at 355 nm and at a frequency of 200 Hz. The window for selecting ions in MS/MS was set to 6 U (mass unit) The spectra were collected automatically for each original (h,d) peak. Only y-type and a-type fragmentations with masses consistent with sodium ionisation $[M+Na]^+$ were observed. The y-type fragmentation is when the peptide bond is cleaved and the positively charged Na^+ ion remains attached to the C terminus and an a-type fragmentation is when the Cα−C(carbonyl) bond is cleaved and the charge remains on the N terminus. From the fragmentation of the homochiral peptides $((n,0)$ and $0,n))$ in which only one sequence is present, it is clear that the experimental abundance of each fragment is not equal to the expected occurrence, which should be identical for all fragments. However, it is clear that the fragmentation is not affected by an isotopic effect or by the nature (R or S) of the residues since the fragment abundances of $(n,0)$ and $(0,n)$ isotactic peptides are very similar with respect to their relative masses (Figure 14).

A program was written to process the MS/MS spectral data (mass, intensity) and to compare these data to possible sequences for each (h,d)
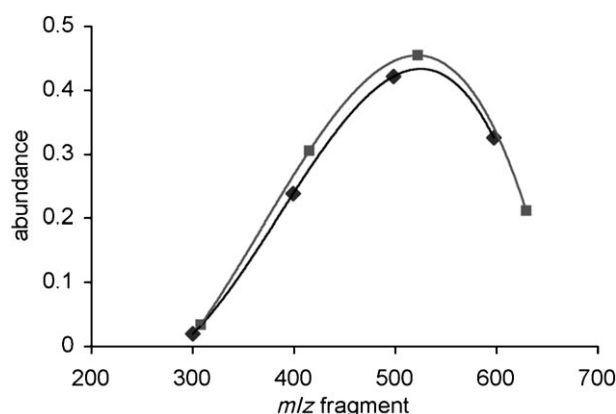

Figure 14. MALDI-TOF-TOF mass spectra data for the (6,0) (◆) and (0,6) (■) oligopeptides of (R,S)-ValNCA initiated with (S)-Phe-OMe. The observed abundance (the peak intensity divided by the sum of all the peak intensities) of the y-type fragments is plotted against the mass.

Table 2. Table of the experimental y-type and a-type fragmentations of the (3,1) oligopeptides obtained from the polymerisation of (R,S)-ValNCA suspended in an aqueous solution of (S)-Phe-OMe (I_S) initiator measured in the MALDI-TOF-TOF MS/MS, as compared with a random polymerisation and the corresponding sequences. The resulting most abundant sequences (MAS) are shown in bold.

	y-Type fragmentations						
no.	$F(m/z)$[a]	m/z[b]	I[c]	O[d]	I_n[e]	O_n[f]	I_n/O_n
1	507.339	507.357	1279	3	0.23	0.75	0.31
2	**499.289**	499.292	4276	1	0.77	0.25	**3.08**
3	408.27	408.268	2284	2	0.15	0.5	0.31
4	**400.221**	400.233	12572	2	0.85	0.5	**1.69**
5	309.202	309.177	848	1	0.14	0.25	0.57
6	301.152	301.139	5129	3	0.86	0.75	1.14
7	202.084	0	0	4	0	1	0

	y-Type fragmentations			
Sequence	1st	2nd	3rd	4th
$RRRSI_S$	507.339	408.27	309.202	202.084
$RRSRI_S$	507.339	408.27	301.152	202.084
$RSRRI_S$	507.339	**400.221**	301.152	202.084
$SRRRI_S$	499.289	**400.221**	301.152	202.084

	a-Type fragmentations						
no.	$F(m/z)$	m/z	I	O	I_n	O_n	I_n/O_n
1	399.318	399.308	14373	4	1	1	1
2	300.249	300.233	7195	3	0.77	0.75	1.03
3	292.2	292.193	2123	1	0.23	0.25	0.91
4	**201.181**	201.175	2408	2	0.68	0.5	**1.36**
5	193.131	193.123	1138	2	0.32	0.5	0.64
6	102.112	0	0	1	0	0.25	0
7	94.063	0	0	3	0	0.75	0

	a-Type fragmentations			
Sequence	1st	2nd	3rd	4th
$RRRSI_S$	94.063	193.131	292.2	399.318
$RRSRI_S$	94.063	193.131	300.249	399.318
$RSRRI_S$	94.063	**201.181**	300.249	399.318
$SRRRI_S$	102.112	**201.181**	300.249	399.318

[a] $F(m/z)$ is the calculated fragment mass. [b] m/z is the measured peak mass. [c] I is the integrated intensity of the peak. [d] O is the occurrence of the fragment for random distribution; it is the number of times this fragment appears in the adjacent table. [e] I_n is the normalised intensity. [f] O_n is the normalised occurrence. Note that fragments with high I_n/O_n values and the most abundant sequences are shown in bold font.

oligopeptide expected from a random polymerisation. The comparison of the experimental fragment abundance to the normalised occurrence for a given sequence allows the extraction of information on the different possible sequences of the (h,d) oligopeptides, as shown in Table 2. For the (3,1) oligopeptide, the y-type fragment at m/z 499 is three times more abundant than the normalised occurrence of the sequences giving the same fragment (m/z 499) and expected for a random process (Table 2). This fragment indicates that the three R repeat units are mainly bound to the initiator I_S whereas the y-type fragment at m/z 507, which has a very low abundance, indicates that the single S repeat unit is less likely to be bound to the initiator. From the data of both fragmentation types, the most abundant sequences (MAS) can be defined (see example for the (3,1) oligopeptide in Table 2) for oligopeptides of length $n = 4$–9 (see Table 1).

Acknowledgements

[1] I. Weissbuch, L. Leiserowitz, M. Lahav, *Top. Curr. Chem.* **2005**, *259*, 123.

[2] P. Franck, W. A. Bonner, R. N. Zare in *Chemistry for the 21th Century* (Eds.: E. Keinan, I. Schecter), Wiley-VCH, Weinheim, **2000**, pp. 175–208.

[3] G. F. Joyce, G. M. Visser, C. A. van Boeckel, J. H. van Boom, L. E. Orgel, J. van Westrenen, *Nature* **1984**, *310*, 602.

[4] I. Rubinstein, R. Eliash, G. Bolbach, I. Weissbuch, M. Lahav, *Angew. Chem.* **2007**, *119*, 3784; *Angew. Chem. Int. Ed.* **2007**, *46*, 3710.

[5] E. R. Blout, P. Doty, J. T. Joung, *J. Am. Chem. Soc.* **1957**, *79*, 749.

[6] R. D. Lundberg, P. Doty, *J. Am. Chem. Soc.* **1957**, *79*, 3961.

[7] A. Brack, G. Spach, *Origins Life* **1981**, *11*, 135.

[8] H. Zepik, E. Shavit, M. Tang, T. R. Jensen, K. Kjaer, G. Bolbach, L. Leiserowitz, I. Weissbuch, M. Lahav, *Science* **2002**, *295*, 1266.

[9] I. Weissbuch, H. Zepik, G. Bolbach, E. Shavit, M. Tang, T. R. Jensen, K. Kjaer, L. Leiserowitz, M. Lahav, *Chem. Eur. J.* **2003**, *9*, 1782.

[10] M. Hasegawa, *Chem. Rev.* **1983**, *83*, 507.

[11] A. Matsumoto, *Top. Curr. Chem.* **2005**, *254*, 263.

[12] G. Wegner, *Makromol. Chem.* **1971**, *145*, 85.

[13] K. Cheng, B. M. Foxman, *J. Am. Chem. Soc.* **1977**, *99*, 8102.

[14] M. J. Vela, V. Buchholz, V. Enkelmann, B. B. Snider, B. M. Foxman, *Chem. Commun.* **2000**, 2225.

[15] J. G. Nery, G. Bolbach, I. Weissbuch, M. Lahav, *Angew. Chem.* **2003**, *115*, 2207; *Angew. Chem. Int. Ed.* **2003**, *42*, 2157.

[16] J. G. Nery, G. Bolbach, I. Weissbuch, M. Lahav, *Chem. Eur. J.* **2005**, *11*, 3039.

[17] G. J. Nery, R. Eliash, G. Bolbach, I. Weissbuch, M. Lahav, *Chirality* **2007**, *19*, 612.

[18] H. Kanazawa, *Polymer* **1992**, *33*, 2557.

[19] H. Kanazawa, Y. Ohashi, *Mol. Cryst. Liq. Cryst.* **1996**, *277*, 45.

[20] H. Kanazawa, T. Hayakawa, *Science Reports of the Faculty of Education Fukushima University* **1999**, *63*, 27.

[21] Y. Takenaka, Y. Ohashi, H. Kanazawa, *Acta Crystallogr. Sect. C* **1994**, *50*, 1950.

[22] H. Kanazawa, *Acta Crystallogr. Sect. E* **2003**, *59*, o1309.

[23] G. Adler, *J. Chem. Phys.* **1959**, *31*, 848.

[24] G. Adler, B. Baysal, *Mol. Cryst. Liq. Cryst.* **1969**, *9*, 361.

[25] C. Huber, W. Eisenreich, S. Hecht, G. Wächtershäuser, *Science* **2003**, *301*, 938.

[26] L. Leman, L. Orgel, M. R. Ghadiri, *Science* **2004**, *306*, 283.

[27] R. Pascal, L. Boiteau, A. Commeyras, *Top. Curr. Chem.* **2005**, *259*, 69.

[28] H. R. Kricheldorf, *Angew. Chem.* **2006**, *118*, 5884; *Angew. Chem. Int. Ed.* **2006**, *45*, 5752.

[29] W. H. Daly, D. Poché, *Tetrahedron Lett.* **1988**, *29*, 5859.