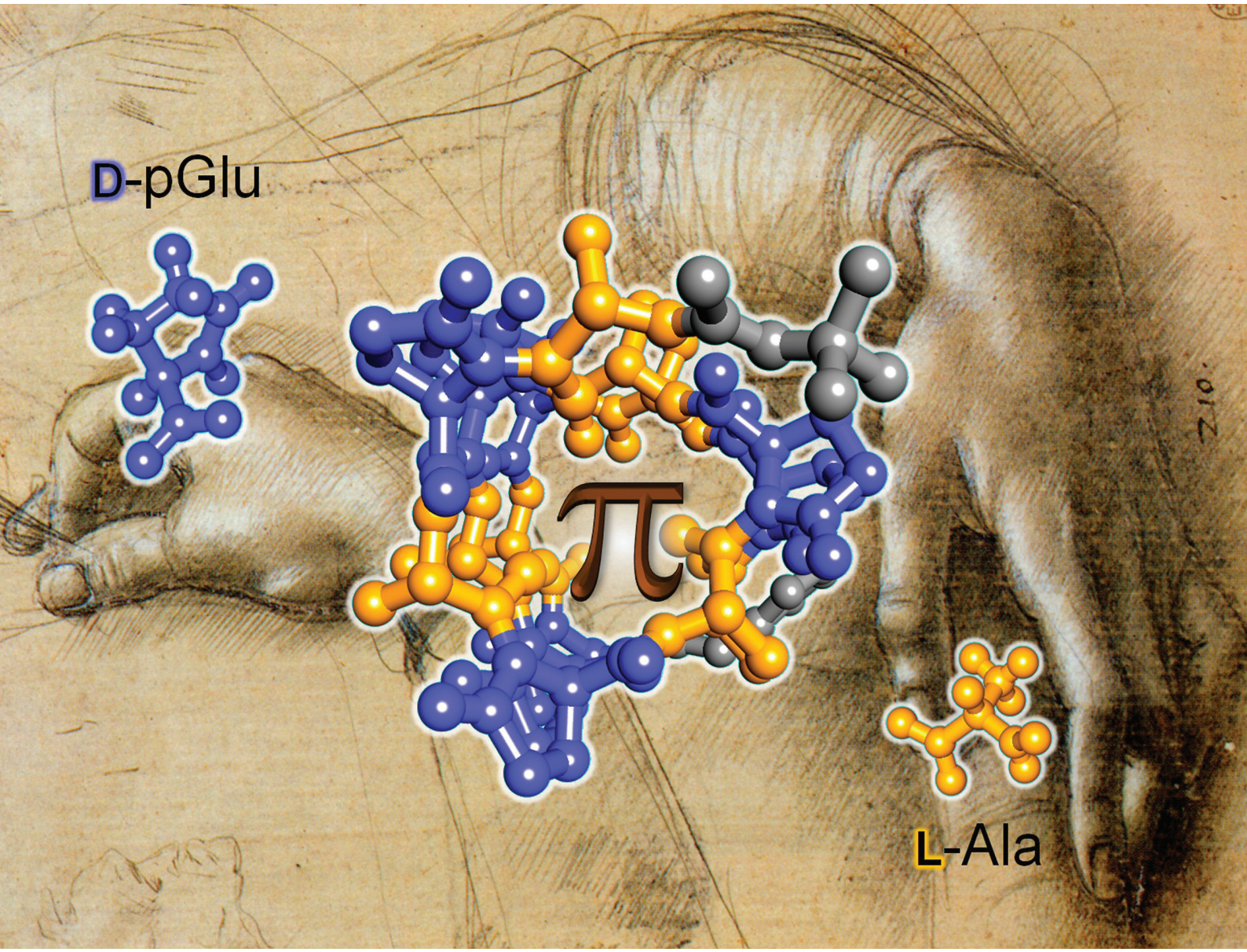
Volume 18
Number 5
7 February 2020
Pages 773-1014

Organic & Biomolecular Chemistry

rsc.li/obc



ISSN 1477-0520

ROYAL SOCIETY OF CHEMISTRY

PAPER
Claudia Tomasini, Gennaro Pescitelli *et al.*
Elusive π-helical peptide foldamers spotted by chiroptical studies

Organic & Biomolecular Chemistry

ROYAL SOCIETY OF CHEMISTRY

PAPER

Cite this: *Org. Biomol. Chem.*, 2020, **18**, 865

Elusive π-helical peptide foldamers spotted by chiroptical studies†

Sergio Di Silvio,[a] Fabio Bologna,[a] Lorenzo Milli,[a] Demetra Giuri,[a] Nicola Zanna,[a] Nicola Castellucci,[a] Magda Monari,[a] Matteo Calvaresi,[a] Marcin Górecki,[b,c] Gaetano Angelici,[b] Claudia Tomasini *[a] and Gennaro Pescitelli *[b]

Received 25th October 2019,
Accepted 10th December 2019

DOI: 10.1039/c9ob02313e

rsc.li/obc

A series of oligomers containing alternate L-Ala and pGlu (pyroglutamic acid) both in the L and D form have been prepared and conformationally investigated by X-ray, NMR, UV/ECD, IR/VCD and molecular modelling. X-ray diffraction analysis was possible for the shortest oligomers **LL-1** and **LD-1**. Molecular dynamics simulations of the oligomers demonstrated that the energy landscapes of the **LL**-series are broad. In contrast, the energy landscapes of the **LD**-series are characterized by well-defined minima corresponding to specific conformational structures. A single well-defined minimum exists in the energy landscape of the largest oligomer **LD-8**, corresponding to a precise conformation, characterized by $i + 5 \rightarrow i$ N–H···O=C hydrogen bonds, typical of a π-helix. ECD and VCD spectra were measured to identify the chiroptical profiles of the oligomers. The most striking element in the ECD spectra of the **LD**-series is their exceptionally strong intensity, which confirms that these polypeptides attain a high degree of helical order. VCD spectra for the **LD**-series are well reproduced by frequency calculations when π-helix folds are employed as input structures, suggesting that a symmetrical VCD couplet around 1720 cm^{-1} can be taken as the VCD signature of π-helices.

Introduction

Proteins and peptides owe their biological activity to their propensity to fold in well-defined secondary and tertiary structures.[1–6] The π-helices are extremely rare secondary structural elements in proteins,[7,8] as they display a repeating pattern in which the backbone C=O of residue i forms a hydrogen bond with the HN group of residue $i + 5$, thus closing a 16-atom membered pseudo-ring often termed C16 (where C stands for "cycle").[9] They have been observed by X-ray diffraction in protein crystal structures, such as FMN-dependent reductases,[10] cytochrome P450,[11] fumarase C (1FUO),[12] glycogen phosphorylase,[13] and lipoxygenases,[14] but they remain rare examples, as both characterization and comprehension of π-helices continue to be evanescent.[6,7,9,15–19]

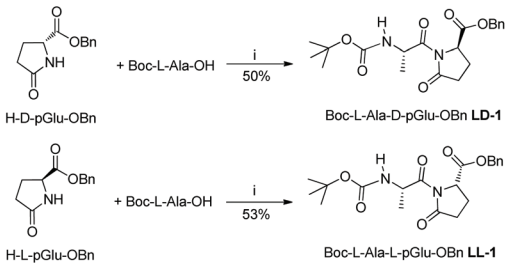
The scientific discussion about the very nature of π-helices started in the early 50s,[20] and continued until today.[15,16,21–26] Given the inherent difficulty to isolate a π-helix and its intrinsic fluxional nature in protein function, there are only few studies on the characterization of this structure in solution or solid phase. Electronic circular dichroism (ECD)[27,28] has been scarcely applied to proteins containing π-helix substructures. The latest version of the Protein Circular Dichroism Data Bank[29] includes only 2 items containing this structural element, but no further detail about π-helix can be found in the related citations.[30] An early theoretical ECD study by Manning and Woody[5] reported a simulated spectrum with a shallow, broad negative maximum at 225 nm and a positive band at 195 nm and with a band shape typical of β-sheets. Morgan *et al.*[31] found that the ECD spectra of a Zn^{2+}-binding amphiphilic polypeptide remained most consistent with that of α-helix, so the formation of a distorted π-helix with an $i \rightarrow i + 5$ structure was inferred, causing a reduction in the intensity of the helical signature. In 2008 Arora and co-workers succeeded to trap a π-helical structure[8] by synthesizing hydrogen bond surrogate (HBS)-based helices. Unfortunately, the existence of an equilibrium between α and π-helix prevented the use of ECD spectroscopy to endorse the peptide as a π-helix. In contrast, Kuczera *et al.*[32] suggested that the ECD profile associated to π-helices resembles that of unordered coils. Thus, the ECD signature of π-helices remains elusive; moreover, no

[a]*Dipartimento di Chimica "Giacomo Ciamician", Alma Mater Studiorum Università di Bologna, Via F. Selmi 2, 40126 Bologna, Italy. E-mail: claudia.tomasini@unibo.it*
[b]*Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via Moruzzi 13, 56124 Pisa, Italy. E-mail: gennaro.pescitelli@unipi.it*
[c]*Institute of Organic Chemistry, Polish Academy of Sciences, Warsaw 01-224, Poland*

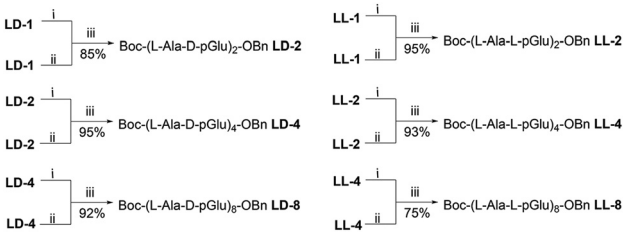
†Electronic supplementary information (ESI) available: Computational section, ^{1}H NMR and ^{13}C NMR spectra of oligomers, crystal data and structure refinement for **LD-1** and **LL-1**, comparison between the VCD spectra of the **LD**- and **LL**-series calculated ECD and VCD spectra. CCDC 1897071 and 1897072. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/c9ob02313e

characterization of π-helices through vibrational circular dichroism (VCD) has been reported so far, to the best of our knowledge.[33]

In addition to the fluxional nature of this structural element, a second reason for its poor characterization lies in the unavailability of peptides featuring large enough fraction of π helices. Synthetic procedures to form stable π helices have been hardly described. In this contribution, we aim to fill this gap reporting a tailored synthesis of a series of oligomers which fold assuming π-helical structures. L-Pyroglutamic acid (L-pGlu) is the dehydrated form of glutamic acid and contains a γ-lactam moiety. L-pGlu is a naturally occurring residue and is frequently found at the N-terminus (but not in an internal position) of bioactive peptides. It lacks the α-amino functionality of α-amino acids that is replaced by a much more difficult to acylate amide function.[34,35] We previously demonstrated that it may be derivatized and introduced into polypeptides chains, thus forming several structures.[36] Among them, homo-oligomers of pGlu tend to form stable helical structures,[37] similar to that adopted by poly-(L-Pro)$_n$ with *trans* tertiary peptide bonds (type-II).[38–41] Both repeating systems generate left-handed, ternary (3_1 symmetry) helices, but the (L-pGlu)$_n$ system is rigid, whereas the (L-Pro)$_n$ system is remarkably more flexible, due to the *cis–trans* isomerism about the tertiary amide bonds. This equilibrium is prevented by the (L-pGlu)$_n$ homo-oligomeric system, due to the αC–H···O═C ($i + 1 \rightarrow i$) intramolecular H-bonding and the dipole–dipole repulsion. The same effect was observed for homo-oligomers of 4-carboxy-5-methyl-oxazolidine-2-ones, that have a similar structure.[37,42]



Scheme 1 Reagents and conditions: (i) HBTU (1.1 equiv.), DBU (2 equiv.), dry CH_3CN, 1 h, r.t.



Scheme 2 Reagents and conditions: (i) H_2, Pd/C (10%), MeOH, r.t. 4 h; (ii) TFA (18 equiv.), dry CH_2Cl_2, r.t. 4 h; (iii) HBTU (1.1 equiv.), DIEA (3 equiv.), dry CH_3CN, r.t., 1 h.

Results and discussion

Synthesis

A series of oligomers containing alternate L-Ala and pGlu (pyroglutamic acid) both in the L and D form have been prepared to check their preferred conformations. The synthesis started from the commercially available Boc-L-Ala-OH, L-pGlu and D-pGlu. The two enantiomers L-pGlu and D-pGlu were transformed into the corresponding benzyl esters H-D-pGlu-OBn and H-L-pGlu-OBn,[36,43] that were coupled with Boc-L-Ala in the presence of HBTU (*N*,*N*,*N*′,*N*′-tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate) and DBU [1,5-diazabiciclo (5.4.0)undec-7-ene] (Scheme 1).

The reaction proved to be quite difficult due to the low reactivity of the amide nitrogen of the lactam ring. Several reaction conditions were tested to optimize the reaction yield that unfortunately never reached completeness. Scheme 1 reports the best reaction conditions found for the coupling reaction, with yields ranging between 50% and 53%. Both diastereoisomers **LD-1** and **LL-1** were purified by flash chromatography, and then oligomers of both series were prepared by coupling reactions in solution. Scheme 2 reports the synthetic steps that allowed us to prepare fully protected dimers **LD-2** and **LL-2**, tetramers **LD-4** and **LL-4** and octamers **LD-8** and **LL-8**.

All compounds were obtained in good to excellent yields, by repeating deprotection of the NH (with trifluoroacetic acid (TFA)) or of the COOH (with H_2, Pd/C) groups and coupled using HBTU and DIEA (*N*,*N*-diisopropylethylamine). The shortest oligomers **LD-2** and **LL-2** were purified by flash chromatography, while the longer oligomers **LD-4**, **LD-8**, **LL-4** and **LL-8** were purified by an ultrasound-assisted washing process using organic solvents as *n*-hexane (*n*-Hex). This method proved to furnish very pure compounds in high yield, as the compound lost during the purification is a negligible amount, due to the very low solubility in *n*-Hex. In contrast, purification of these compounds with flash chromatography was responsible of a significative reduction of the final yield due to the products high tendency to bind to the silica gel.

Characterization

We attempted to grow crystals suitable for single crystals X-ray diffraction with oligomers of both series and of any length. Unfortunately, only **LL-1** furnished crystals suitable for diffraction from ethyl acetate (Fig. 1 and Fig. S1, ESI†). So we prepared the enantiomers of all the oligomers, as it is known that the crystallization of enantiomerically pure chiral molecules is more difficult than the crystallization of the corresponding racemate.[44–46] It is still not clear whether more facile crystallization of racemic protein mixtures is a general phenomenon, as predicted by Wukovitz and Yeates and co-workers,[47,48] but this phenomenon has been extensively studied both for proteins and for short peptides and has been successfully applied to the determination of the unknown structure of snow flea

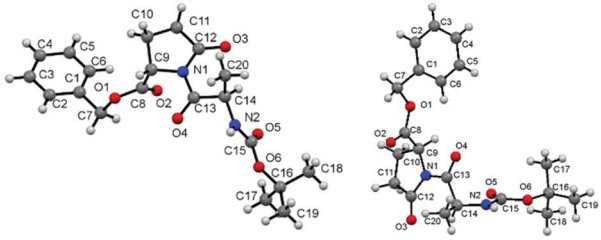


Fig. 1 Molecular structure of **LD-1** in the racemate (left) and of **LL-1** (right) as deduced from X-ray diffraction analysis.

antifreeze protein (sfAFP).[49] Balaram *et al.* have determined the crystal structures of seven tripeptides in enantiomeric and racemic forms.[50] Using this technique, we could grow crystals suitable for X-ray diffraction only from **LD-1** in ethyl acetate (Fig. 1 and Fig. S2†). Interestingly, racemic **LD-1** crystallizes as an antiparallel β-sheet structure (Fig. S3,† relevant distances and dihedral angles are listed in Tables S2 and S3†). The crystallization of longer oligomers always failed both as racemate and in the enantiomerically pure form from a wide range of solvents.

Oligomers of both series were analysed to check their tendency to form stable folded conformations. ^{1}H NMR analysis did not furnish any striking information; ROESY analysis of the longest oligomer **LD-8** did not show the presence of NH–NH cross peaks (see Fig. S4†). The FT-IR absorption spectra of the three sets were obtained as 3 mM solutions in methylene chloride: at this concentration, self-aggregation is usually unimportant.[51–56] In contrast, the FT-IR absorption analysis suggested that N–H···O═C bonds are formed in the longer oligomers of both sets. Fig. 2 shows the FT-IR absorption spectra (N–H and C═O stretching regions) for all the synthesized compounds and helped us to detect non-hydrogen-bonded amide N–H bands (above 3400 cm^{-1}) and hydrogen-bonded amide protons bands (below 3400 cm^{-1}). The shorter oligomers **LD-1**, **LD-2**, **LD-4**, **LL-1**, **LL-2** and **LL-4** give very similar spectra, with both peaks below and above 3400 cm^{-1}, thus suggesting that equilibria between folded and unfolded conformations take place. In contrast, the longest oligomers **LD-8** and **LL-8** feature a very different behaviour, as the FT-IR absorption spectrum of **LD-8** shows the presence of strong peaks at 3294 cm^{-1} and 1660 cm^{-1}, both typical of the formation of N–H···O═C hydrogen bonds, thus revealing a folded conformation. In contrast, the FT-IR spectrum of **LL-8** shows the appearance of a weak peak at 3364 cm^{-1}, together with a broad peak at 1683 cm^{-1}. These two signals both agree with the formation of a partially folded conformation in a rapid equilibrium with unfolded structures.

VCD and ECD characterization of oligomers of both series was pursued with the aim of obtaining more sensitive information on the preferred conformations. ECD spectra were recorded in acetonitrile using quartz cells of various path-lengths from 0.01 to 1 cm, in order to properly capture the signals appearing in the region between 190 and

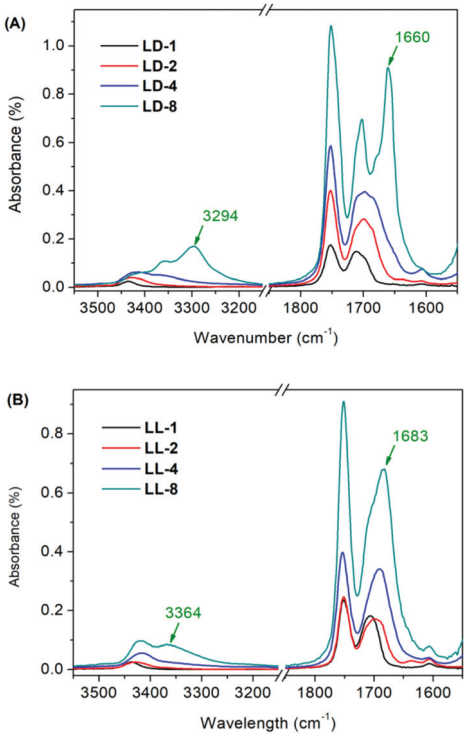


Fig. 2 Selected regions of FT-IR absorption spectra of oligomers of the **LD**-series (A) and of the **LL**-series (B). All the spectra have been recorded as 3 mM solutions in dichloromethane.

300 nm. The spectra for the **LD** and **LL** series are shown in Fig. 3 and 4.

At first sight, the spectra for the two series look like the mirror image of each other. This is an indication that the secondary structure handedness is mainly dictated by the chirality of the pGlu residues, while the Ala residues play a minor role. In particular, the 240 nm band preserves its whole intensity and changes sign from negative to positive when passing from the **LD** to the **LL** series. This is better appreciated by means of

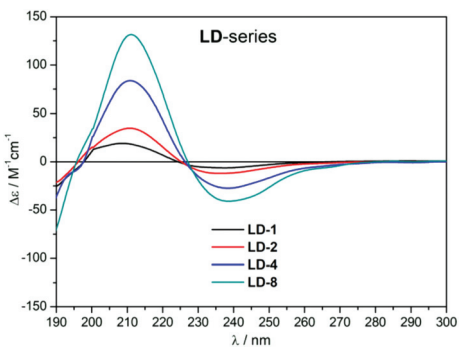


Fig. 3 ECD spectra of the **LD**-series recorded with cells having path lengths of 0.01 cm, 0.05 cm and 1 cm (from left to right), as acetonitrile solutions having the following concentrations: **LD-1** (7.4 × 10^{-4} M; **LD-2** (6.6 × 10^{-4} M); **LD-4** (2.5 × 10^{-4} M); **LD-8** (1.3 × 10^{-4} M).

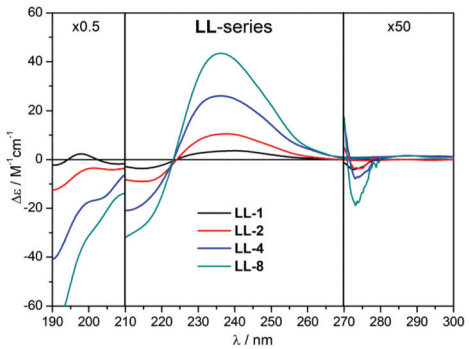


Fig. 4 ECD spectra of the **LL**-series recorded with cells having path lengths of 0.01 cm, 0.05 cm and 1 cm (from left to right), as acetonitrile solutions having the following concentrations: **LL-1** (8.4×10^{-4} M); **LL-2** (4.8×10^{-4} M); **LL-4** (2.5×10^{-4} M); **LL-8** (1.5×10^{-4} M).

a direct comparison between the ECD spectra recorded for the longest oligomers **LD-8** and **LL-8** (Fig. 5). The **LD** series display a second well-defined positive band with maximum at 210 nm. This band is negative and much weaker for the **LL** series (Fig. 3–5), which shows instead a relatively strong negative tail at shorter wavelengths which has no positive counterpart for the **LD** series. Additionally, the **LL**-series features a weak negative band above 270 nm which appears as a shoulder of the same sign for the **LD**-series.

A second piece of information offered by ECD spectra is the nature of secondary structure and the related degree of helical order reached by the various oligomers. The ECD profile of **LD** series is reminiscent of a β-sheet folding of L-amino acid proteins and peptides,[27] although both maxima are shifted to longer wavelengths by 17–25 nm. This red-shift cannot be attributed solely to solvent effects[57] and needs a different explanation (*vide infra*). However, the most striking element in the ECD spectra of the **LD** series is their exceptionally strong intensity. The $\Delta\varepsilon$ for the 212 nm band of octamer **LD-8** reaches a value of 140 M^{-1} cm^{-1}, that is one order of magnitude larger than the typical values for the 195 nm band of β-sheet (5–10 M^{-1} cm^{-1}) and the 192 nm band of α-helix (15–20 M^{-1} cm^{-1}).[58] As the ECD intensity can be taken as a measure of the degree of polypeptide helical order, this observation suggests that the octamer is folded in a well-ordered secondary structure. The same is also true for the tetramer **LD-4** and, to a lesser extent, even for the dimer **LD-2**. Not surprisingly, the degree of order increases with the chain length (Fig. 3).

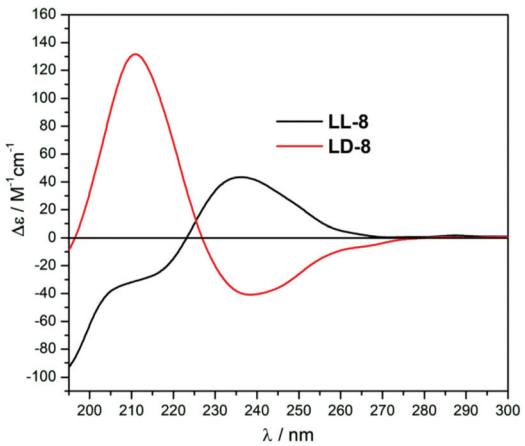


Fig. 5 ECD spectra of **LL-8** and **LD-8** respectively recorded as 1.5×10^{-4} M and 1.3×10^{-4} M solutions in acetonitrile with 0.05 cm cells.

The ECD profile of the **LL** series (Fig. 4) does not resemble any of the typical ECD signatures for standard secondary structure motifs of proteins and peptides based on L-amino acids. The ECD intensity is again quite large (40 M^{-1} cm^{-1} for the 240 nm band of **LL-8**) although less impressive than for the **LD** series. The reason for the extraordinary appearance of the ECD spectra of **LL** and especially **LD** series must be sought not only in conformational reasons (the unusual type of secondary folding), but also in electronic ones, because the presence of pGlu alters the nature of one chromophore (an imide group) responsible for the ECD bands. This latter expectation is verified by ECD calculations discussed later. Although ECD spectra of several chiral imide derivatives have been reported in the literature,[59–61] we are aware of only report concerning peptides carrying a single imide chromophore in the form of a L-aminosuccinyl (L-Asu) residue.[62] They feature ECD bands >240 nm ascribable to the imide chromophores, which were allied with $\Delta\varepsilon$ values around −1 M^{-1} cm^{-1}.

VCD spectra of all oligomers were recorded in d^3-acetonitrile using a BaF_2 cell with 100 μm pathlength. Although the region between 1100 and 1500 cm^{-1} displays several intense bands, we focused our attention on the region between 1600 and 1800 cm^{-1}, which corresponds to the amide I/II band region most often considered in the VCD analysis of peptides and proteins.[33] In this spectral window, all oligomers display a sequence of three bands with alternating −/+/− sign (Fig. 6 and 7). A closer inspection of the three bands reveals that they increase in intensity upon chain lengthening, albeit less regularly than the ECD profiles. Another marked difference with ECD is that the VCD spectra for the **LL**-series are consistently stronger than the **LD**-series for a given oligomer size. This fact can be easily appreciated looking at direct spectral comparisons reported in Fig. S5.† As it emerges from normal mode analysis (*vide infra*), the two bands in the range 1680–1720 cm^{-1} are due to C═O stretching vibrations located on the amide groups. Thus, they correspond to the amide I band, which is normally observed for peptides and proteins around 1650 cm^{-1}. The shift to higher frequencies (by ~50 cm^{-1}) is mostly due to the solvent effect passing from water to acetonitrile.[63] The third band centred at 1750 cm^{-1} is instead due to C═O stretching vibrations located on the imide groups. The region below 1600 cm^{-1}, where amide II bands are expected in acetonitrile,[63] is weak and little significant in our spectra.

If we restrict our analysis to the amide I band, we recognize a negative couplet (*i.e.*, a sequence of two bands with −/+, from shorter to longer wavenumbers) for both series. The

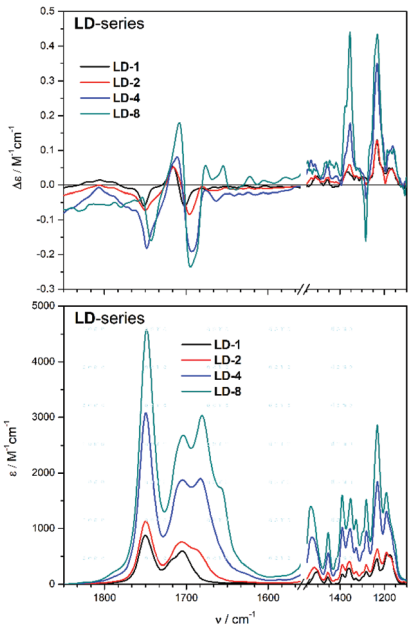


Fig. 6 VCD (top) and FT-IR absorption (bottom) spectra of the **LD**-series recorded with a 100 μm cell, as d^3-acetonitrile solutions having the following concentrations: **LD-1** (0.14 M); **LD-2** (0.09 M); **LD-4** (0.04 M); **LD-8** (0.02 M).

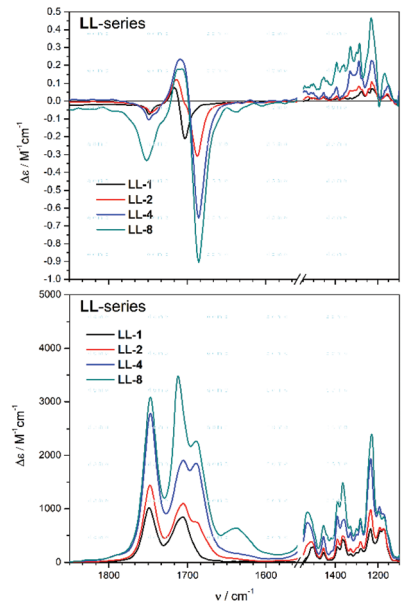


Fig. 7 VCD (top) and FT-IR absorption (bottom) spectra of the **LL**-series recorded with a 100 μm cell, as d^3-acetonitrile solutions having the following concentrations: **LL-1** (0.12 M); **LL-2** (0.09 M); **LL-4** (0.04 M); **LL-8** (0.02 M).

couplet is quite symmetrical for the **LD**-series (Fig. 6), while it is biased toward the negative component for the **LL**-series (Fig. 7). The couplet intensity (peak-to-trough amplitude) is ~1.1 M^{-1} cm^{-1} for **LL-8** and ~0.4 M^{-1} cm^{-1} for **LD-8**, while the *g*-factors ($\Delta\varepsilon/\varepsilon$ or $\Delta A/A$) for the most intense band around 1690 cm^{-1} attain values of 4×10^{-4} for **LL-8** and 9×10^{-5} for **LD-8**, respectively. An intense negative couplet with negative bias and *g*-factor $\geq 10^{-4}$ is commonly accepted as the VCD signature of random coil or extended helix conformation.[33] In our foldamers, the amide I band is also excitonically-coupled with the imide C═O stretching at higher frequency. A similar situation was observed for an oxindolyl *N*-acetyloxazolidinone derivative of flustramine B.[64] In this latter case, a VCD couplet was detected around 1700 cm^{-1} yielding a negative imide VCD band at 1730 cm^{-1} with intensity around −0.2 M^{-1} cm^{-1}.

The present qualitative VCD analysis thus suggests a preferred random coil or extended helix conformation for the **LL**-series, while it is not sufficient to infer the favoured folding of the **LD**-oligomers. As we shall see below, however, VCD spectra for this series are well reproduced by frequency calculations when π-helix folds are employed as input structures, suggesting that a symmetrical negative couplet with *g*-factor $<10^{-4}$ can be taken as the VCD signature of π-helices.

Molecular dynamics simulations

Molecular dynamics simulations for 100 ns of **LD-1**, **LL-1**, **LD-2**, **LL-2**, **LD-4**, **LL-4**, **LD-8** and **LL-8** were carried out in an explicit acetonitrile box, reproducing the conditions in which VCD and ECD spectra have been recorded. Principal component analysis (PCA) of the MD trajectories generated the free-energy landscape of the oligomers (see Computational section†). This approach proved to be effective in reproducing the NMR and crystallographic data of peptides and pseudopeptides and in addition provides insight into their free energy landscape.[65,66]

In the energy landscapes of the **LL**-series (Fig. 8a–d) the local minima are broad, and there is no significant basin of attraction indicative of a folded conformation. The most populated conformations represent non-folded extended conformations of the oligomers that undergo practically free rotation in the solvent (Fig. S6†). This finding is consistent with the results of the qualitative analysis of VCD spectra for the **LL** series, which suggested a preferred random coil or extended helix conformation. On the contrary, the energy landscapes of the **LD**-series (Fig. 8e–h) are characterized by well-defined minima corresponding to specific conformations (Fig. S7†). The energy landscapes of the oligomers **LD-1**, **LD-2** and **LD-4** suggest that an equilibrium between conformations characterized by a different folding takes place. In the energy landscape of the longest oligomers **LD-8**, a single, well-defined minimum exists, corresponding to a precise conformation, confirming the suggestions provided by the FT-IR absorption spectra, ECD and VCD spectra.

MD simulations show an almost completely folded structure of **LD-8** (Fig. 9) in acetonitrile at room temperature. Top view perspective (Fig. 9a) of **LD-8** draws attention on the structural role of pGlu rings and Ala methyl groups that are stacked vertically and are tightly packed and interlocked. This conformer is also characterized by a well-defined hydrogen-bond pattern, including two distinct types of hydrogen bonds. The formation of the $i + 5 \rightarrow i$ N–H···O═C hydrogen bonds is the

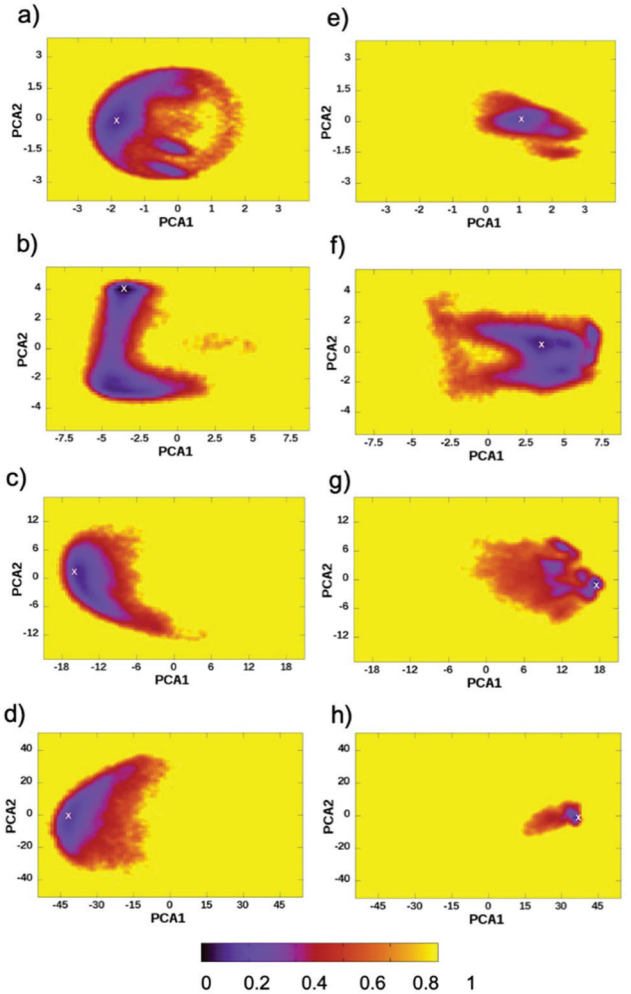


Fig. 8 Two-dimensional normalized free-energy landscape of the (a) **LL-1**, (b) **LL-2**, (c) **LL-4**, (d) **LL-8**, (e) **LD-1**, (f) **LD-2**, (g) **LD-4**, and (h) **LD-8** oligomers. PCA1 and PCA2 are the two eigenvectors with the lowest eigenvalues of the principal component analysis (PCA) calculated from the analysis of the MD trajectory.

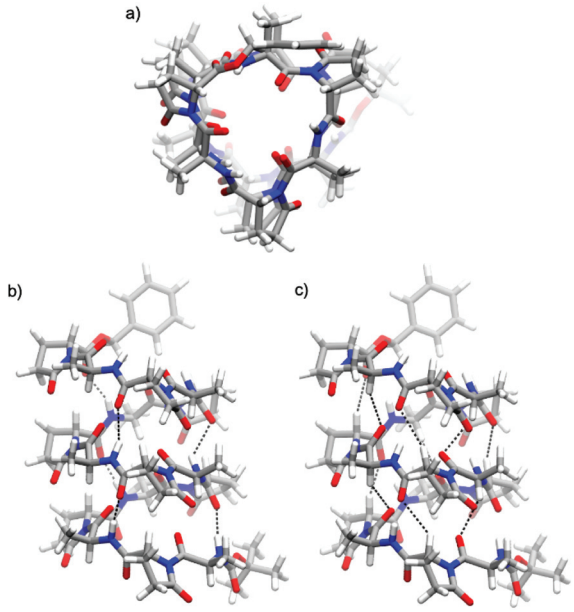


Fig. 9 Representative structure of **LD-8** corresponding to the absolute minimum of the free energy landscape in Fig. 8h seen from three viewpoints. (a) Top view perspective, (b and c) front view perspectives. In black (b) the $i + 5 \rightarrow i$ N–H···O=C hydrogen bonds, (c) the non-standard $i + 6 \rightarrow$ C=O···H–C hydrogen bonds.

driving force for the folding (Fig. 9b). Amide carbonyl of Ala are always the hydrogen bond acceptors of the $i + 5 \rightarrow i$ N–H···O=C hydrogen bonds, while the endocyclic carbonyl of pGlu is not involved in the hydrogen bond network. This well-defined hydrogen-bond pattern is typical of a π-helix. An additional stabilization element of this folded structure derives also from non-standard $i + 6 \rightarrow i$ C=O···H–C hydrogen bonds between the carbonyl group of Ala and the Cα–H of pGlu (Fig. 9c). The important role of this latter class of hydrogen bonds was recently highlighted on protein and nucleic acid structure, molecular recognition, and enzyme catalysis.[67] A careful analysis of the values of the torsion angles Φ and Ψ (Table S4†) of the **LD-8** structure allowed the precise assignment of the identified conformer as a π_{LD} helix. The π_{LD} helices are formed by repeating **LD** dipeptides, in which the two residues populate the extended regions of the Φ–Ψ map, that lie on opposite sides of the Ramachandran plot, although the overall structure formed is helical (Fig. S8†), in analogy to Gramicidin or Feglymycin helices.[68] The alternating pattern of L- and D-amino acids in **LD-8** is crucial for the formation of these structures.

ECD and VCD calculations

DFT-based (density functional theory) calculations of ECD and VCD spectra are nowadays routine procedures to study the configuration and conformation of several types of molecules.[69] We run ECD and VCD calculations on all compounds for which experimental ECD/VCD spectra were measured, namely **LD-1**, **LL-1**, **LD-2**, **LL-2**, **LD-4**, **LL-4**, **LD-8** and **LL-8**. In all cases, we used a single input structure selected from the global energy minima obtained from MD simulations (see white marks on Fig. 8). Especially for the oligomers characterized by broad energy landscapes, we did not expect such a procedure to reproduce the experimental spectra perfectly, for which a thorough conformational sampling would be needed. Our main purpose was in fact to interpret ECD and VCD spectra in terms of the most significant electronic and vibrational transitions, respectively, and to rationalize the observed trends along each series, and the observed differences among the two series. On the other hand, the sharp distinct energy minima found for some oligomers might be validated by ECD and VCD calculations. All calculated spectra are shown in Fig. S9A and S9B,† while the most representative results are summarized in Fig. 10 and 11.

The input structures for all calculations were obtained by DFT re-optimization, at B3LYP/6-311G(d,p) level of theory, of the MD energy minima. Excited states calculations were run

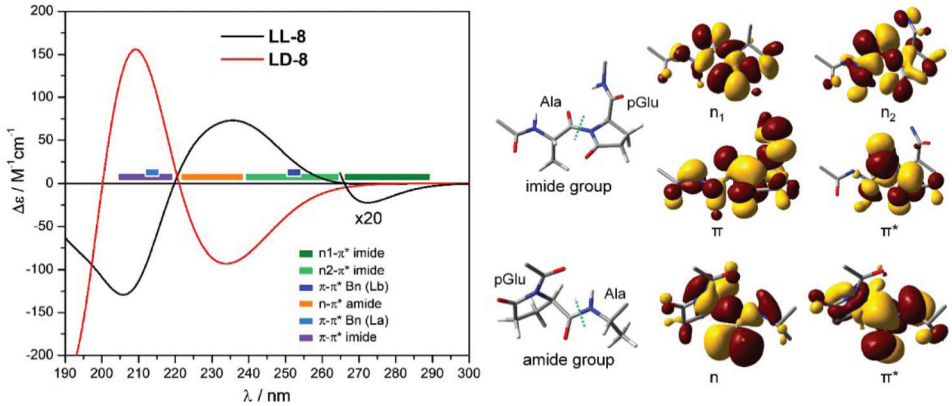


Fig. 10 Left: Calculated ECD spectra of **LL-8** and **LD-8** at the CAM-B3LYP/def2-TZVP//B3LYP/6-311G(d,p) level of theory with PCM for acetonitrile. The horizontal bars define the spectral regions associated to the transitions indicated in the legend (with reference to **LD-8**). Right: Main orbitals involved in the amide- and imide-centred transitions.

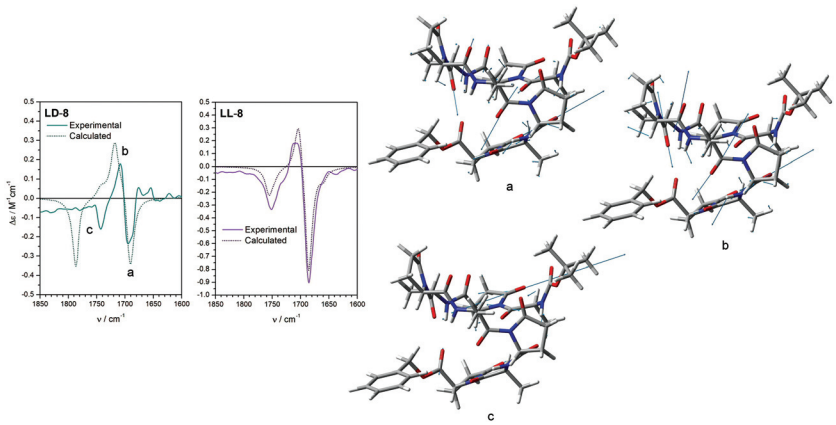


Fig. 11 (Left) Comparison between experimental and calculated VCD spectra of **LL-8** and **LD-8**. Calculations run at the B3LYP/6-311G(d,p) level of theory; frequency scaling factor 0.966 and 0.985, respectively; spectra plotted as sums of Lorentzians with 8 cm^{-1} full width at half height. (Right) Main normal modes responsible for the major calculated VCD bands a–c, plotted for **LD-4**; the arrows indicate the displacement vectors.

with time-dependent DFT (TDDFT) at CAM-B3LYP/def2-TZVP level of theory including a continuum solvent model (IEF-PCM) for acetonitrile. The calculated ECD spectra for both series (Fig. S9A and S9B†) compare well with the observed ECD profiles and the main trends. For all compounds except **LD-2**, the sign of the first ECD band is correctly predicted by calculations (negative for the **LD** series, positive for the **LL**). For most compounds, the overall couplet-like feature is also reproduced (negative couplet for **LD**, positive couplet for **LL**). More specifically, for the two largest members of each series (**LD-4**, **LL-4**, **LD-8** and **LL-8**), the agreement between calculated and experimental spectra is very good, and replicates the differences observed between the two series (different sign and intensity) and along each series (increasing intensity). The ECD spectra calculated for octamers **LL-8** and **LD-8** are shown in Fig. 10. If we compare them with Fig. 5, we may appreciate how even some fine details of the experimental spectra are reproduced, like the small negative band at 270 nm for **LL-8**. More importantly, the calculations confirm the exceptionally strong ECD spectra associated with our foldamers. The agreement between experimental and calculated spectra not only substantiates the outcome of MD simulations, that predict better defined conformational ensembles for the longer oligomers, but also signifies that the MD-predicted energy minima are well representative of the conformational population.

The calculated ECD spectra of **LD-4** and **LD-8** were analysed in detail to ascertain if a band assignment was possible. In fact, every observed band originates from the superposition of several different transitions localized on the various chromophores. Following the coloured horizontal bars in Fig. 10, the negative ECD band in the spectrum of **LD-8** is due to the convolution of all possible n–π* transitions localized on the amide and imide groups, while the positive ECD band is mainly due to π–π* transitions of the imide groups. While the main nature of the two bands is typical of standard peptides, the substitution of an amide bond with an imide one strongly affects the character of the various transitions, as can be appreciated by looking at the natural transition orbitals (NTO) in Fig. 10.

Frequency calculations were run at B3LYP/6-311G(d,p) level of theory. Limiting again our analysis to the region between 1600 and 1800 cm^{-1}, we notice also in this case a qualitative correspondence between experimental and calculated IR/VCD spectra (Fig. S10†) for almost all compounds except **LD-2** and **LD-4**, for which the single conformer picture is evidently too rough. In Fig. 11, the direct comparison of VCD spectra for the longest oligomers **LL-8** and **LD-8** is displayed. In both cases, the calculations correctly reproduce the pattern of signs −/+/− and the relative signal intensity of the three major bands in the significant region. The agreement for **LL-8** is almost perfect; for **LD-8**, instead, the frequency separation between the three bands is overestimated by calculations. The discrepancy is most likely due to a somewhat inadequate treatment of hydrogen bonding at the present level of calculations; inclusion of an empirical dispersion of Grimme's D3 type had no or little effect on the result and did not improve the overall outcome (Fig. S11†). However, the agreement is sufficient to interpret the observed bands in terms of underlying normal modes. Normal modes analysis reveals that the three major VCD bands are due to the combination of all the possible carbonyl stretching modes. The negative couplet (−/+ signs from low to high frequency) around 1700 cm^{-1} is due to the stretching vibrations of amide carbonyl groups and of imide carbonyl groups on the Ala sides, that is, the carbonyl moieties one will find in a "standard" peptide devoid of the additional pGlu C═O groups. We can therefore identify these two bands as those corresponding to amide I vibrations. The higher frequency negative band is instead associated with the stretching vibration of imide carbonyl groups on the pGlu sides, that is, the "additional" carbonyl moieties with respect to a "standard" peptide.

The normal modes most contributing to the major calculated VCD bands for **LD-4** are displayed in Fig. 11; the assignment is consistent for **LD-8** and for the **LL**-series. Our analysis let us conclude that the couplet of bands around 1700 cm^{-1} can be taken as the VCD signature of π-helix structure (in CD_3CN).

Experimental section

General remarks

Solvents were dried by distillation before use. All reactions were carried out in dried glassware. The melting points of the compounds were determined in open capillaries and are uncorrected. High quality infrared spectra (64 scans) were obtained at 2 cm^{-1} resolution with an ATR-FT-IR Bruker Alpha System spectrometer. All spectra were obtained in 3 mM solutions in CH_2Cl_2. All compounds were dried *in vacuo* and all the sample preparations were performed in a nitrogen atmosphere. NMR spectra were recorded with a Varian Inova 400 spectrometer at 400 MHz (^{1}H NMR) and at 100 MHz (^{13}C NMR). Chemical shifts are reported in δ values relative to the solvent peak. HPLC-MS was used to check the purity of compounds. ECD spectra were measured with a Jasco J-715 spectropolarimeter with the following conditions: scan speed 100 nm min^{-1}; response 0.5 s; data pitch 0.2 nm; bandwidth 1.0 nm; 4 accumulations. VCD and IR spectra were recorded using a Jasco FVS-6000 VCD spectrometer with the following conditions: resolution 4 cm^{-1}; range 2000–900 cm^{-1}; 4000 accumulations.

Boc-L-Ala-D-pGlu-OBn, LD-1. To a stirred solution of Boc-L-Ala-OH (0.86 g, 4.56 mmol) and HBTU (1.90 g, 5.02 mmol) in dry acetonitrile (12 mL), under inert atmosphere, D-pGlu-OBn (1.00 g, 4.56 mmol) in dry acetonitrile (12 mL) was added at room temperature, followed by a solution of DBU (1.36 mL, 9.13 mmol). The solution was stirred for 1 hour under inert atmosphere, then acetonitrile was removed under reduced pressure and replaced with dichloromethane (100 mL). The mixture was washed with brine (70 mL), 1 N aqueous HCl (70 mL) and with 5% aqueous $NaHCO_3$ (70 mL), dried over sodium sulphate and concentrated *in vacuo*. The product was obtained pure after silica gel chromatography (7 : 3 → 1 : 1 → 1 : 9 cyclohexane/AcOEt) in 50% of yield (0.89 g, 2.28 mmol). M.p. = 68–70 °C; IR (3 mM, CH_2Cl_2) = 3434, 2933, 1753, 1709, 1501 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 1.37 (d, 3H, J = 6.7 Hz, CH_3 Ala), 1.44 (s, 9H, 3 × CH_3 Boc), 2.09 (ddt, 1H, J = 13.6 Hz, J = 9.4 Hz, J = 2.3 Hz, C*H*H β pGlu), 2.34 (ddt, 1H, J = 13.6 Hz, J = 11.3 Hz, J = 9.4 Hz, C*H*H β pGlu), 2.55 (ddd, 1H, J = 17.7 Hz, J = 9.4 Hz, J = 2.3 Hz, C*H*H γ pGlu), 2.66–2.79 (m, 1H, C*H*H γ pGlu), 4.74 (dd, 1H, J = 9.4 Hz, J = 2.3 Hz, CH α pGlu), 5.20 (d, 2H, J = 3.0 Hz, O–CH_2–Ph), 5.31–5.38 (m, 1H, NH), 5.39–5.47 (m, 1H, CH α Ala), 7.31–7.41 (m, 5H, Ar H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 19.3, 21.6, 28.4, 31.8, 50.2, 58.3, 67.5, 79.6, 128.3, 128.4, 128.6, 128.7, 128.8, 135.2, 154.8, 170.5, 173.8, 174.5. Anal. calcd for $C_{20}H_{26}N_2O_6$: C, 61.53; H, 6.71; N, 7.18. Found: C, 61.58; H, 6.69; N, 7.13.

Boc-(L-Ala-D-pGlu)$_2$-OBn, LD-2. TFA (2.72 mL, 16.10 mmol) was added under nitrogen atmosphere to a solution of Boc-L-Ala-D-pGlu-OBn **LD-1** (0.35 g, 0.90 mmol) in dry dichloromethane (10 mL). The mixture was stirred at room temperature for 4 hours and concentrated *in vacuo* after a control *via* TLC. H-L-Ala-D-pGlu-OBn·CF_3CO_2H was obtained in quantitative yield.

Pd/C (10% w/w, 35 mg) was added under inert atmosphere to a stirred solution of Boc-L-Ala-D-pGlu-OBn **LD-1** (0.35 g, 0.90 mmol) in methanol (6 mL). Vacuum was created inside the flask using the vacuum line. The flask was then filled with hydrogen using a balloon (1 atm). The solution was stirred for 4 hours under a hydrogen atmosphere. Boc-L-Ala-D-pGlu-OH was obtained pure after filtration through Celite filter and concentrated *in vacuo* in quantitative yield. A solution of Boc-L-Ala-D-pGlu-OH and HBTU (0.37 g, 0.99 mmol) in dry acetonitrile (6 mL) was stirred under nitrogen atmosphere for 10 min at room temperature, then a solution of the H-L-Ala-D-pGlu-OBn·CF_3CO_2H and DIEA (0.27 mL, 2.69 mmol) in dry acetonitrile (6 mL) was added dropwise. The solution was stirred for 1 h under nitrogen atmosphere, then acetonitrile was removed under reduced pressure and replaced with dichloromethane (50 mL). The mixture was washed with brine (30 mL), 1 N aqueous HCl (30 mL) and with saturated aqueous $NaHCO_3$

(30 mL), dried over sodium sulphate and concentrated *in vacuo*. The product was obtained pure after silica gel chromatography (cyclohexane/ethyl acetate 3 : 7 → ethyl acetate) in 85% of yield (0.44 g, 0.76 mmol). M.p. = 89–92 °C; IR (3 mM, CH_2Cl_2) = 3427, 2934, 1752, 1698, 1500 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 1.33 (d, 3H, J = 6.1 Hz, CH_3 Ala), 1.38 (s, 9H, 3 × CH_3 Boc), 1.41 (d, 3H, J = 6.8 Hz, CH_3 Ala), 2.01–2.12 (m, 1H, C*H*H β pGlu), 2.13–2.40 (m, 3H, 3 × C*H*H β pGlu), 2.42–2.59 (m, 2H, 2 × C*H*H γ pGlu), 2.63–2.76 (m, 1H, C*H*H γ pGlu), 2.79–2.89 (m, 1H, C*H*H γ pGlu), 4.63–4.73 (m, 2H, 2 × CH α pGlu), 5.17 (d, 2H, J = 8.2 Hz, O–CH_2–Ph), 5.28–5.35 (m, 1H, CH α Ala), 5.39–5.57 (m, 2H, CH α Ala + NH), 7.05–7.13 (m, 1H, NH), 7.28–7.39 (m, 5H, Ar H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 17.6, 18.2, 21.8, 23.1, 28.4, 31.8, 32.2, 49.9, 50.3, 58.5, 59.4, 67.5, 80.0, 128.3, 128.6, 128.7, 135.2, 155.4, 169.9, 170.5, 173.6, 174.1, 174.8, 174.9. Anal. calcd for $C_{28}H_{36}N_4O_9$: C, 58.73; H, 6.34; N, 9.78. Found: C, 58.67; H, 6.32; N, 9.81.

Boc-(L-Ala-D-pGlu)$_4$-OBn, LD-4. TFA (0.44 mL, 5.66 mmol) was added under nitrogen atmosphere to a solution of Boc-(L-Ala-D-pGlu)$_2$-OBn **LD-2** (0.18 g, 0.31 mmol) in dry dichloromethane (4 mL). The mixture was stirred at room temperature for 4 hours and concentrated *in vacuo* after a control *via* TLC. H-(L-Ala-D-pGlu)$_2$-OBn·CF_3CO_2H was obtained in quantitative yield. Pd/C (10% w/w, 18 mg) was added under inert atmosphere to a stirred solution of Boc-(L-Ala-D-pGlu)$_2$-OBn **LD-2** (0.18 g, 0.31 mmol) in methanol (5 mL). Vacuum was created inside the flask using the vacuum line. The flask was then filled with hydrogen using a balloon (1 atm). The solution was stirred for 4 hours under a hydrogen atmosphere. Boc-(L-Ala-D-pGlu)$_2$-OH was obtained pure after filtration through Celite filter and concentrated *in vacuo* in quantitative yield. A solution of Boc-(L-Ala-D-pGlu)$_2$-OH and HBTU (0.13 g, 0.35 mmol) in dry acetonitrile (6 mL) was stirred under nitrogen atmosphere for 10 min at room temperature, then a solution of the H-(L-Ala-D-pGlu)$_2$-OBn·CF_3CO_2H and DIEA (0.16 mL, 0.94 mmol) in dry acetonitrile (6 mL) was added dropwise. The solution was stirred for 1 h under nitrogen atmosphere, then acetonitrile was removed under reduced pressure and replaced with dichloromethane (40 mL). The mixture was washed with brine (30 mL), 1 N aqueous HCl (30 mL) and with saturated aqueous $NaHCO_3$ (30 mL), dried over sodium sulphate and concentrated *in vacuo*. The purification occurred with an ultrasound-assisted washing process using organic solvents as *n*-Hex achieving product in 95% of yield (279.8 mg, 0.30 mmol). M.p. = 135–137 °C; IR (3 mM, CH_2Cl_2) = 3416, 3367, 2934, 1752, 1697, 1605, 1525 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 1.35 (d, 3H, J = 6.9 Hz, CH_3 Ala), 1.41 (s, 12H, 3 × CH_3 Boc + CH_3 Ala), 1.42–1.44 (m, 3H, CH_3 Ala), 1.45–1.47 (m, 3H, CH_3 Ala), 2.14–2.40 (m, 8H, 4 × CH_2 β pGlu), 2.41–2.61 (m, 4H, 4 × C*H*H γ pGlu), 2.65–2.77 (m, 2H, 2 × C*H*H γ pGlu), 2.82–2.98 (m, 2H, 2 × C*H*H γ pGlu), 4.65–4.82 (m, 4H, 4 × CH α pGlu), 5.11–5.29 (m, 4H, O–CH_2–Ph + CH α Ala + NH), 5.34–5.41 (m, 1H, CH α Ala), 5.42–5.48 (m, 1H, CH α Ala), 5.55–5.65 (m, 1H, CH α Ala), 7.14–7.18 (m, 1H, NH) 7.30–7.38 (m, 6H, Ar H + NH), 7.40–7.45 (m, 1H, NH); ^{13}C NMR (100 MHz, $CDCl_3$) δ 16.4, 17.5, 21.7, 23.0, 28.4, 31.8, 32.2, 49.8, 49.9, 50.4, 58.5, 59.2, 59.4, 67.6, 80.4, 128.3, 128.5, 128.6, 128.8, 135.3, 155.8, 170.1, 170.8, 173.7, 173.9, 174.2, 175.0, 175.2. Anal. calcd for $C_{44}H_{56}N_8O_{15}$: C, 56.40; H, 6.02; N, 11.96. Found: C, 56.35; H, 6.08; N, 12.02.

Boc-(L-Ala-D-pGlu)$_8$-OBn, LD-8. TFA (0.15 mL, 1.92 mmol) was added under nitrogen atmosphere to a solution of Boc-(L-Ala-D-pGlu)$_4$-OBn **LD-4** (0.10 g, 0.11 mmol) in dry dichloromethane (2 mL). The mixture was stirred at room temperature for 4 hours and concentrated *in vacuo* after a control *via* TLC. H-(L-Ala-D-pGlu)$_4$-OBn·CF_3CO_2H was obtained in quantitative yield. Pd/C (10% w/w, 10 mg) was added under inert atmosphere to a stirred solution of Boc-(L-Ala-D-pGlu)$_4$-OBn **LD-4** (0.10 g, 0.11 mmol) in methanol (3 mL). Vacuum was created inside the flask using the vacuum line. The flask was then filled with hydrogen using a balloon (1 atm). The solution was stirred for 4 hours under a hydrogen atmosphere. Boc-(L-Ala-D-pGlu)$_4$-OH was obtained pure after filtration through Celite filter and concentrated *in vacuo* in quantitative yield. A solution of Boc-(L-Ala-D-pGlu)$_4$-OH and HBTU (0.45 g, 0.12 mmol) in dry acetonitrile (3 mL) was stirred under nitrogen atmosphere for 10 min at room temperature, then a solution of the H-(L-Ala-D-pGlu)$_4$-OBn·CF_3CO_2H and DIEA (0.55 mL, 0.32 mmol) in dry acetonitrile (3 mL) was added dropwise. The solution was stirred for 1 h under nitrogen atmosphere, then acetonitrile was removed under reduced pressure and replaced with dichloromethane (30 mL). The mixture was washed with brine (20 mL), 1 N aqueous HCl (20 mL) and with saturated aqueous $NaHCO_3$ (20 mL), dried over sodium sulphate and concentrated *in vacuo*. The purification occurred with an ultrasound-assisted washing process using organic solvents as *n*-Hex achieving product in 92% of yield (0.16 g, 0.098 mmol). M.p. = 220–222 °C; IR (3 mM, CH_2Cl_2) = 3358, 3294, 2937, 2882, 1751, 1702, 1660, 1534 cm^{-1}; ^1H NMR (400 MHz, DMSO, d_6) δ 1.07–1.28 (m, 33H, 8 × CH_3 Ala + 3 × CH_3 Boc), 1.68–2.70 (m, 32H, 8 × CH_2 β pGlu + 8 × CH_2 γ pGlu), 4.55–4.88 (m, 8H, 8 × CH α pGlu), 5.02–5.26 (m, 3H, O–CH_2–Ph + CH α Ala), 5.33–5.62 (m, 7H, 7 × CH α Ala), 6.92 (d, J = 6 Hz, 1H, NH), 7.24–7.39 (m, 5H, Ar), 8.30–8.77 (m, 8H, NH); ^{13}C NMR (100 MHz, $CDCl_3$) δ 17.9, 20.8, 21.1, 28.2, 28.4, 28.5, 29.7, 30.9, 49.0, 53.4, 61.9, 74.8, 128.3, 128.7, 128.8, 151.6, 168.6, 171.2, 171.9. Anal. calcd for $C_{76}H_{96}N_{16}O_{27}$: C, 54.80; H, 5.81; N, 13.45. Found: C, 54.87; H, 5.85; N, 13.50.

Boc-L-Ala-L-pGlu-OBn, LL-1. To a stirred solution of Boc-L-Ala-OH (0.86 g, 4.56 mmol) and HBTU (1.90 g, 5.02 mmol) in dry acetonitrile (12 mL), under inert atmosphere, L-pGlu-OBn (1.00 g, 4.56 mmol) in dry acetonitrile (12 mL) was added at room temperature, followed by a solution of DBU (1.36 mL, 9.13 mmol). The solution was stirred for 1 hour under inert atmosphere, then acetonitrile was removed under reduced pressure and replaced with dichloromethane (100 mL). The mixture was washed with brine (70 mL), 1 N aqueous HCl (70 mL) and with 5% aqueous $NaHCO_3$ (70 mL), dried over sodium sulphate and concentrated *in vacuo*. The product was obtained pure after silica gel chromatography (7 : 3 → 1 : 1 → 1 : 9 ciclohexane/AcOEt) in 53% of yield (0.94 g, 2.42 mmol). M.p. = 104–105 °C; IR (3 mM, CH_2Cl_2) = 3437, 2934, 1750,

1706, 1500 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 1.29 (d, 3H, J = 7.0 Hz, CH_3 Ala), 1.41 (s, 9H, 3 × CH_3 Boc), 2.00–2.16 (m, 1H, C*H*H β pGlu), 2.51–2.71 (m, 1H, C*H*H β pGlu + C*H*H γ pGlu), 4.84 (dd, 1H, J = 9.5 Hz, J = 3.1 Hz, CH α pGlu), 5.06 (d, 1H J = 8.0 Hz, NH), 5.18 (AB, 2H, J = 12.0 Hz, O–C*H*$_2$–Ph), 5.34 (dq, 1H, J = 6.0, 6.8 Hz, CH α Ala), 7.28–7.39 (m, 5H, Ar H Conf A + B); ^{13}C NMR (100 MHz, $CDCl_3$) δ 17.8, 21.4, 28.3, 31.8, 49.7, 57.8, 67.5, 79.8, 128.3, 128.6, 128.7, 134.9, 155.2, 170.6, 173.8, 174.7. Anal. calcd for $C_{20}H_{26}N_2O_6$: C, 61.53; H, 6.71; N, 7.18. Found: C, 61.50; H, 6.72; N, 7.11.

Boc-(L-Ala-L-pGlu)$_2$-OBn, LL-2. TFA (1.59 mL, 20.7 mmol) was added under nitrogen atmosphere to a solution of Boc-L-Ala-L-pGlu-OBn **LL-2** (0.45 g, 1.148 mmol) in dry dichloromethane (11 mL). The mixture was stirred at room temperature for 4 hours and concentrated *in vacuo* after a control *via* TLC. H-L-Ala-L-pGlu-OBn·CF_3CO_2H was obtained in quantitative yield. Pd/C (10% w/w, 48 mg) was added under inert atmosphere to a stirred solution of Boc-L-Ala-L-pGlu-OBn **LL-2** (0.45 g, 1.15 mmol) in methanol (10 mL). Vacuum was created inside the flask using the vacuum line. The flask was then filled with hydrogen using a balloon (1 atm). The solution was stirred for 4 hours under a hydrogen atmosphere. Boc-L-Ala-L-pGlu-OH was obtained pure after filtration through Celite filter and concentrated *in vacuo* in quantitative yield. A solution of Boc-L-Ala-L-pGlu-OH and HBTU (0.48 g, 1.26 mmol) in dry acetonitrile (6 mL) was stirred under nitrogen atmosphere for 10 min at room temperature, then a solution of the H-L-Ala-L-pGlu-OBn·CF_3CO_2H and DIEA (0.59 mL, 3.44 mmol) in dry acetonitrile (6 mL) was added dropwise. The solution was stirred for 1 h under nitrogen atmosphere, then acetonitrile was removed under reduced pressure and replaced with dichloromethane (60 mL). The mixture was washed with brine (40 mL), 1 N aqueous HCl (40 mL) and with saturated aqueous $NaHCO_3$ (40 mL), dried over sodium sulphate and concentrated *in vacuo*. The product was obtained pure after silica gel chromatography (cyclohexane/ethyl acetate 3 : 7 → ethyl acetate) in 95% of yield (0.62 g, 1.090 mmol). M.p. = 88–90 °C; IR (3 mM, CH_2Cl_2) = 3427, 2934, 1750, 1698, 1501 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 1.35 (d, 3H, J = 5.5 Hz, CH_3 Ala), 1.40 (d, 3H, J = 2.7 Hz, CH_3 Ala), 1.42 (s, 9H, 3 × CH_3 Boc), 2.04–2.14 (m, 1H, C*H*H β pGlu), 2.15–2.31 (m, 2H, 2 × C*H*H β pGlu), 2.32–2.44 (m, 1H, C*H*H β pGlu), 2.45–2.54 (m, 2H, 2 × C*H*H γ pGlu), 2.65–2.87 (m, 2H, 2 × C*H*H γ pGlu), 4.59–4.78 (m, 3H, 2 × CH α pGlu Conf A + CH α pGlu Conf B), 4.79–4.88 (m, 1H, CH α pGlu Conf B), 5.05–5.15 (m, 2H, NH Conf A + NH Conf B) 5.21 (s, 2H, O–CH_2–Ph), 5.27–5.40 (m, 2H, CH α Ala Conf A + CH α Ala Conf B), 5.46–5.57 (m, 1H, CH α Ala Conf A), 5.58–5.68 (m, 1H, CH α Ala Conf B), 6.42 (d, 1H, J = 6.3 Hz, NH Conf A), 6.66 (d, 1H, J = 7.8 Hz, NH Conf B), 7.30–7.40 (m, 5H, Ar H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 18.0, 18.8, 21.6, 21.9, 28.5, 31.9, 32.2, 49.9, 50.4, 58.0, 58.8, 67.6, 80.0, 128.3, 128.4, 128.8, 135.0, 155.4, 169.4, 170.0, 170.5, 173.6, 174.0, 174.8. Anal. calcd for $C_{28}H_{36}N_4O_9$: C, 58.73; H, 6.34; N, 9.78. Found: C, 58.72; H, 6.38; N, 9.83.

Boc-(L-Ala-L-pGlu)$_4$-OBn, LL-4. TFA (0.69 mL, 8.80 mmol) was added under nitrogen atmosphere to a solution of Boc-(L-Ala-L-pGlu)$_2$-OBn **LL-2** (0.28 g, 0.49 mmol) in dry dichloromethane (5 mL). The mixture was stirred at room temperature for 4 hours and concentrated *in vacuo* after a control *via* TLC. H-(L-Ala-L-pGlu)$_2$-OBn·CF_3CO_2H was obtained in quantitative yield. Pd/C (10% w/w, 18 mg) was added under inert atmosphere to a stirred solution of Boc-(L-Ala-L-pGlu)$_2$-OBn **LL-2** (0.28 g, 0.49 mmol) in methanol (5 mL). Vacuum was created inside the flask using the vacuum line. The flask was then filled with hydrogen using a balloon (1 atm). The solution was stirred for 4 hours under a hydrogen atmosphere. Boc-(L-Ala-L-pGlu)$_2$-OH was obtained pure after filtration through Celite filter and concentrated *in vacuo* in quantitative yield. A solution of Boc-(L-Ala-L-pGlu)$_2$-OH and HBTU (0.20 g, 0.54 mmol) in dry acetonitrile (6 mL) was stirred under nitrogen atmosphere for 10 min at room temperature, then a solution of the H-(L-Ala-L-pGlu)$_2$-OBn·CF_3CO_2H and DIEA (0.25 mL, 1.47 mmol) in dry acetonitrile (6 mL) was added dropwise. The solution was stirred for 1 h under nitrogen atmosphere, then acetonitrile was removed under reduced pressure and replaced with dichloromethane (50 mL). The mixture was washed with brine (30 mL), 1 N aqueous HCl (30 mL) and with saturated aqueous $NaHCO_3$ (30 mL), dried over sodium sulphate and concentrated *in vacuo*. The purification occurred with an ultrasound-assisted washing process using organic solvents as *n*-Hex achieving product in 93% of yield (0.43 g, 0.45 mmol). M.p. = 144–145 °C; IR (3 mM, CH_2Cl_2) = 3416, 2935, 1753, 1690, 1516 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 1.30–1.54 (m, 21H, 4 × CH_3 Ala + 3 × CH_3 Boc), 2.19–2.46 (m, 8H, 4 × CH_2 β pGlu), 2.47–2.66 (m, 5H, 5 × C*H*H γ pGlu), 2.68–2.89 (m, 3H, 3 × C*H*H γ pGlu), 4.59–4.89 (m, 4H, 4 × CH α pGlu), 5.20 (s, 2H, O–CH_2–Ph), 5.28–5.39 (m, 2H, CH α Ala + NH), 5.40–5.66 (m, 3H, 3 × CH α Ala), 6.39–6.55 (m, 1H, NH), 6.57–6.72 (m, 1H, NH), 6.86–7.02 (m, 1H, NH), 7.29–7.46 (m, 5H, Ar H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 14.1, 17.1, 17.4, 17.8, 21.0, 21.3, 21.9, 28.2, 29.6, 31.7, 31.9, 31.9, 32.0, 36.5, 38.6, 49.0, 49.7, 57.8, 58.6, 60.3, 67.5, 79.6, 128.2, 128.6, 155.2, 169.9, 135.0, 170.4, 173.5, 173.8, 174.7, 175.1. Anal. calcd for $C_{44}H_{56}N_8O_{15}$: C, 56.40; H, 6.02; N, 11.96. Found: C, 56.44; H, 5.98; N, 12.00.

Boc-(L-Ala-L-pGlu)$_8$-OBn, LL-8. TFA (0.15 mL, 1.92 mmol) was added under nitrogen atmosphere to a solution of Boc-(L-Ala-L-pGlu)$_4$-OBn **LL-4** (0.10 g, 0.11 mmol) in dry dichloromethane (2 mL). The mixture was stirred at room temperature for 4 hours and concentrated *in vacuo* after a control *via* TLC. H-(L-Ala-D-pGlu)$_4$-OBn·CF_3CO_2H was obtained in quantitative yield. Pd/C (10% w/w, 10 mg) was added under inert atmosphere to a stirred solution of Boc-(L-Ala-L-pGlu)$_4$-OBn **LL-4** (0.10 g, 0.11 mmol) in methanol (3 mL). Vacuum was created inside the flask using the vacuum line. The flask was then filled with hydrogen using a balloon (1 atm). The solution was stirred for 4 hours under a hydrogen atmosphere. Boc-(L-Ala-L-pGlu)$_4$-OH was obtained pure after filtration through Celite filter and concentrated *in vacuo* in quantitative yield. A solution of Boc-(L-Ala-L-pGlu)$_4$-OH and HBTU (0.45 g, 0.12 mmol) in dry acetonitrile (3 mL) was stirred under nitrogen atmosphere for 10 min at room temperature, then a solution of the H-(L-Ala-L-pGlu)$_4$-OBn·CF_3CO_2H and DIEA (0.55 mL, 0.32 mmol) in dry

acetonitrile (3 mL) was added dropwise. The solution was stirred for 1 h under nitrogen atmosphere, then acetonitrile was removed under reduced pressure and replaced with dichloromethane (30 mL). The mixture was washed with brine (20 mL), 1 N aqueous HCl (20 mL) and with saturated aqueous $NaHCO_3$ (20 mL), dried over sodium sulphate and concentrated *in vacuo*. The purification occurred with an ultrasound-assisted washing process using organic solvents as *n*-Hex achieving product in 75% of yield (0.13 g, 0.08 mmol). M.p. = 166–168 °C; IR (3 mM, CH_2Cl_2) = 3418, 3368, 2935, 1751, 1683, 1515 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 1.02–1.50 (m, 41H, 8 × CH_3 Ala + 3 × CH_3 Boc + 8 × C*H*H β pGlu), 1.55–1.85 (m, 8H, 8 × CH*H* β pGlu), 1.92–2.27 (m, 8H, 8 × C*H*H γ pGlu), 2.33–2.77 (m, 8H, 8 × CH*H* γ pGlu), 4.45–4.78 (m, 8H, 8 × CH α pGlu), 5.03 (bs, 3H, O–CH_2–Ph + NH-Boc), 5.17–5.57 (m, 8H, 8 × CH α Ala), 6.28–6.64 (m, 4H, 4 × NH), 7.06–7.32 (m, 8H, Ar H + 3 × NH); ^{13}C NMR (100 MHz, $CDCl_3$) δ 17.5, 21.4, 22.0, 22.7, 28.3, 29.7, 31.7, 31.9, 49.1, 49.7, 57.7, 57.9, 58.7, 65.8, 79.8, 128.3, 128.7, 134.9, 153.1, 155.0, 169.9, 173.6, 174.0, 174.7. Anal. calcd for: $C_{76}H_{96}N_{16}$: C, 54.80; H, 5.81; N, 13.45. Found: C, 54.84; H, 5.77; N, 13.49.

Single-crystal X-ray crystallography

The X-ray intensity data for **LD-1** and **LL-1** were measured on a Bruker ApexII CCD diffractometer. The Bruker Apex2 package[70] was used for collecting frames of data, indexing reflections, and determining lattice parameters. The collected frames were then processed for integration by the SAINT program,[70] and an empirical absorption correction was applied using SADABS.[71] The structure was solved by direct methods (SIR 2014)[72] and subsequent Fourier syntheses and refined by full-matrix least squares on F^2 (SHELXL),[73] using anisotropic thermal parameters for all non-hydrogen atoms. All hydrogen atoms except the amidic ones were added in calculated positions included in the final stage of refinement with isotropic thermal parameters, $U(H) = 1.2[U_{eq}(C)]$ ($U(H) = 1.5[U_{eq}(C–Me)]$), and allowed to ride on their carrier carbons. Crystal data and details of the data collection for **LD-1** and **LL-1** are reported in Table S1.† Molecular drawings were generated using Mercury.[74]

CCDC 1897071 and 1897072† contains the supplementary crystallographic data for this paper.

Computational section

See ESI.†

Conclusions

The oligomers Boc-(L-Ala-D-pGlu)$_n$-OBn and Boc-(L-Ala-L-pGlu)$_n$-OBn (n = 1, 2, 4 and 8) have been prepared by liquid-phase synthesis and their preferred conformation was analysed by X-ray, NMR, IR absorption and molecular dynamics. With these compounds in our hands, we measured ECD and VCD spectra to identify the respective spectral profiles, and substantiated our investigation by ECD and VCD calculations. As the ECD spectra for the two series look like the mirror image of each other, the secondary structure handedness is mainly dictated by the chirality of the pGlu residues, while the Ala residues play a minor role. The most striking element in the ECD spectra of the **LD** series is their exceptionally strong intensity, that can be taken as a measure of the polypeptides helical order degree and suggests that the oligomers are folded in well-ordered secondary structures. The intensity of the ECD spectra of the **LL** series is again quite large although less impressive than the **LD** series.

In VCD spectra, the region between 1600 and 1800 cm^{-1}, corresponding to the amide I/II band region, is especially significant in the analysis of peptides and proteins. In this spectral window, all oligomers display a sequence of three bands with alternating –/+/– sign, increasing in intensity upon chain lengthening. Molecular dynamics simulations for 100 ns of oligomers of both series were carried out in an explicit acetonitrile box, reproducing the conditions in which VCD and ECD spectra have been recorded. In the energy landscapes of the **LL**-series the local minima are broad, and there is no significant basin of attraction indicative of a folded conformation. In contrast, the energy landscapes of the **LD**-series are characterized by well-defined minima corresponding to specific conformational structures. In particular, in the energy landscape of the longer oligomer **LD-8**, a single, well-defined minimum exists, corresponding to a precise conformation, characterized by well-defined $i + 5 \rightarrow i$ N–H···O═C hydrogen bonds, typical of a π-helix.

Starting from MD local minima, we run ECD and VCD calculations on all compounds for which experimental ECD/VCD spectra were measured. The calculations reproduced satisfactorily the observed trends for the two series of ECD and VCD spectra. Some discrepancy was found for the VCD spectra of the **LD-8** series, most likely due to a somewhat inadequate treatment of hydrogen bonding at the present level of calculations. However, the agreement is good enough to interpret the observed bands in terms of underlying normal modes.

Using this technique, we could identify two bands as those corresponding to amide I vibrations. The higher frequency negative band is instead associated with the stretching vibration of imide carbonyl groups on the pGlu sides, that is, the "additional" carbonyl moieties with respect to a "standard" peptide. Our analysis let us conclude that the couplet of bands around 1700 cm^{-1} can be taken as the VCD signature of π-helix structure (in CD_3CN). Because of the elusive nature of the π-helix structural motif, no VCD fingerprint had been so far identified in the literature.

Further work is in progress to prepare crystals of the longer oligomer **LD-8**. Their X-ray diffraction analysis will help to obtain a definitive crystallographic evidence, with precise backbone torsion angles, for the π-helix in peptides.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

C. T. gratefully acknowledges the University of Bologna and Ministero dell'Università e della Ricerca (PRIN 2015 project 20157WW5EH). G. P. acknowledges the CINECA award under the ISCRA initiative, for the availability of high-performance computing resources and support. M. G. thanks the program Bekker of the Polish National Agency for Academic Exchange (NAWA).

Notes and references

1 M. A. Schmitt, S. H. Choi, I. A. Guzei and S. H. Gellman, *J. Am. Chem. Soc.*, 2005, **127**, 13130–13131.
2 C. Toniolo, F. Formaggio, S. Tognon, Q. B. Broxterman, B. Kaptein, R. Huang, V. Setnicka, T. A. Keiderling, I. H. McColl, L. Hecht and L. D. Barron, *Biopolymers*, 2004, **75**, 32–45.
3 W. C. Pomerantz, T. L. R. Grygiel, J. R. Lai and S. H. Gellman, *Org. Lett.*, 2008, **10**, 1799–1802.
4 B. Bochicchio and A. M. Tamburro, *Chirality*, 2002, **14**, 782–792.
5 M. C. Manning and R. W. Woody, *Biopolymers*, 1991, **31**, 569–586.
6 K. H. Lee, D. R. Benson and K. Kuczera, *Biochemistry*, 2000, **39**, 13737–13747.
7 T. M. Weaver, *Protein Sci.*, 2000, **9**, 201–206.
8 R. Chapman, J. L. Kulp III, A. Patgiri, N. R. Kallenbach, C. Bracken and P. S. Arora, *Biochemistry*, 2008, **47**, 4189–4195.
9 R. P. Riek and R. M. Graham, *J. Struct. Biol.*, 2011, **173**, 153–160.
10 J. S. McFarlane, R. A. Hagen, A. S. Chilton, D. L. Forbes, A. L. Lamb and H. R. Ellis, *Protein Sci.*, 2019, **28**, 123–134.
11 C. A. Hasemann, R. G. Kurumbail, S. S. Boddupalli, J. A. Peterson and J. Deisenhofer, *Structure*, 1995, **3**, 41–62.
12 T. Weaver and L. Banaszak, *Biochemistry*, 1996, **35**, 13955–13965.
13 S. G. Withers, N. B. Madsen, S. R. Sprang and R. J. Fletterick, *Biochemistry*, 1982, **21**, 5372–5382.
14 J. C. Boyington, B. J. Gaffney and L. M. Amzel, *Adv. Exp. Med. Biol.*, 1997, **400**, 133–138.
15 B. W. Low and H. J. Grenville-Wells, *Proc. Natl. Acad. Sci. U. S. A.*, 1953, **39**, 785–801.
16 C. A. Rohl and A. J. Doig, *Protein Sci.*, 1996, **5**, 1687–1696.
17 R. B. Cooley, D. J. Arp and P. A. Karplus, *J. Mol. Biol.*, 2010, **404**, 232–246.
18 M. N. Fodje and S. Al-Karadaghi, *Protein Eng., Des. Sel.*, 2002, **15**, 353–358.
19 J. M. Musila, D. Forbes and H. R. Ellis, *Biochemistry*, 2018, **57**, 4469–4477.
20 R. B. Corey and J. Donohue, *J. Am. Chem. Soc.*, 1950, **72**, 2899–2900.
21 M. L. Huggins, *J. Am. Chem. Soc.*, 1952, **74**, 3963–3964.
22 L. Pauling and R. B. Corey, *J. Am. Chem. Soc.*, 1952, **74**, 3964.
23 B. W. Low and R. B. Baybutt, *J. Am. Chem. Soc.*, 1952, **74**, 5806–5807.
24 J. Donohue, *Proc. Natl. Acad. Sci. U. S. A.*, 1953, **39**, 470–478.
25 G. N. Ramachandran and V. Sasisekharan, *Adv. Protein Chem.*, 1968, **23**, 283–437.
26 Y. Wang and K. Kuczera, *J. Phys. Chem. B*, 1997, **101**, 5205–5213.
27 N. Sreerama, S. Y. U. Venyaminov and R. W. Woody, *Protein Sci.*, 2008, **8**, 370–380.
28 R. W. Janes, *Bioinformatics*, 2005, **21**, 4230–4238.
29 L. Whitmore, A. J. Miles, L. Mavridis, R. W. Janes and B. A. Wallace, *Nucleic Acids Res.*, 2017, **45**, D303–D307.
30 A. J. Miles, N. U. Fedosova, S. V. Hoffmann, B. A. Wallace and M. Esmann, *Biochem. Biophys. Res. Commun.*, 2013, **435**, 300–305.
31 D. M. Morgan, D. G. Lynn, H. Miller-Auer and S. C. Meredith, *Biochemistry*, 2001, **40**, 14020–14029.
32 K.-H. Lee, D. R. Benson and K. Kuczera, *Biochemistry*, 2000, **39**, 13737–13747.
33 T. A. Keiderling and A. Lakhani, in *Comprehensive Chiroptical Spectroscopy*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2012, vol. 2, pp. 707–758.
34 G. A. N. Gowda, Y. N. Gowda and D. Raftery, *Anal. Chem.*, 2015, **87**, 3800–3805.
35 B. Rigo, B. Herb, S. El Ghammarti, P. Gautret and D. Couturier, *J. Heterocycl. Chem.*, 1995, **32**, 1599–1604.
36 C. Tomasini and M. Villa, *Tetrahedron Lett.*, 2001, **42**, 5211–5214.
37 C. Tomasini, V. Trigari, S. Lucarini, F. Bernardi, M. Garavelli, C. Peggion, F. Formaggio and C. Toniolo, *Eur. J. Org. Chem.*, 2003, 259–267.
38 A. A. Adzhubei, M. J. E. Sternberg and A. A. Makarov, *J. Mol. Biol.*, 2013, **425**, 2100–2132.
39 L. Garbuio, B. Lewandowski, P. Wilhelm, L. Ziegler, M. Yulikov, H. Wennemers and G. Jeschke, *Chem. – Eur. J.*, 2015, **21**, 10747–10753.
40 P. M. Cowan and S. McGavin, *Nature*, 1955, **175**, 642–643.
41 N. Zanna, L. Milli, B. Del Secco and C. Tomasini, *Org. Lett.*, 2016, **18**, 1662–1665.
42 S. Lucarini and C. Tomasini, *J. Org. Chem.*, 2001, **66**, 727–732.
43 F. Bernardi, M. Garavelli, M. Scatizzi, C. Tomasini, V. Trigari, M. Crisma, F. Formaggio, C. Peggion and C. Toniolo, *Chem. – Eur. J.*, 2002, **8**, 2516–2525.
44 W. Cai, J. Marciniak, M. Andrzejewski and A. Katrusiak, *J. Phys. Chem. C*, 2013, **117**, 7279–7285.
45 C. Toniolo, C. Peggion, M. Crisma, F. Formaggio, X. Shui and D. S. Eggleston, *Nat. Struct. Biol.*, 1994, **1**, 908–914.
46 L. E. Zawadzke and J. M. Berg, *Proteins: Struct., Funct., Genet.*, 1993, **16**, 301–305.
47 T. O. Yeates and S. B. H. Kent, *Annu. Rev. Biophys.*, 2012, **41**, 41–61.
48 M. Pellegrini, S. W. Wukovitz and T. O. Yeates, *Proteins*, 1997, **28**, 515–521.

49 B. L. Pentelute, Z. P. Gates, V. Tereshko, J. L. Dashnau, J. M. Vanderkooi, A. A. Kossiakoff and S. B. H. Kent, *J. Am. Chem. Soc.*, 2008, **130**, 9695–9701.

50 I. Saha, B. Chatterjee, N. Shamala and P. Balaram, *Pept. Sci.*, 2008, **90**, 537–543.

51 L. Belvisi, C. Gennari, A. Mielgo, D. Potenza and C. Scolastico, *Eur. J. Org. Chem.*, 1999, 389–400.

52 B. Imperiali, R. A. Moats, S. L. Fisher and T. J. Prins, *J. Am. Chem. Soc.*, 1992, **114**, 3182–3188.

53 F. M. Mir, M. Crisma, C. Toniolo and W. D. Lubell, *Chem. Sci.*, 2019, **10**, 6908–6914.

54 E. S. Stevens, N. Sugawara, G. M. Bonora and C. Toniolo, *J. Am. Chem. Soc.*, 1980, **102**, 7048–7050.

55 I. G. Jones, W. Jones and M. North, *J. Org. Chem.*, 1998, **63**, 1505–1513.

56 J. Yang and S. H. Gellman, *J. Am. Chem. Soc.*, 1998, **120**, 9090–9091.

57 E. B. Nielsen and J. A. Schellman, *J. Phys. Chem.*, 1967, **71**, 2297–2304.

58 N. Greenfield, *Nat. Protocol.*, 2007, **1**, 2876–2890.

59 J. Gawroński, M. Brzostowska, K. Kacprzak, H. Kołbon and P. Skowronek, *Chirality*, 2000, **12**, 263–268.

60 T. Połoński, *J. Chem. Soc., Perkin Trans. 1*, 1988, 629–637.

61 T. Połoński, *J. Chem. Soc., Perkin Trans. 1*, 1988, 639–648.

62 S. Capasso, L. Mazzarella, F. Sica and A. Zagari, *Int. J. Pept. Protein Res.*, 1989, **33**, 124–132.

63 J. Kubelka and T. A. Keiderling, *J. Phys. Chem. A*, 2001, **105**, 10922–10928.

64 R. E. Cordero-Rivera, M. Meléndez-Rodríguez, O. R. Suárez-Castillo, C. I. Bautista-Hernández, N. Trejo-Carbajal, J. Cruz-Borbolla, L. E. Castelán-Duarte, M. S. Morales-Ríos and P. Joseph-Nathan, *Tetrahedron: Asymmetry*, 2015, **26**, 710–720.

65 L. Milli, M. Larocca, M. Tedesco, N. Castellucci, E. Ghibaudi, A. Cornia, M. Calvaresi, F. Zerbetto and C. Tomasini, *J. Org. Chem.*, 2014, **79**, 5958–5969.

66 S. Fermani, X. Trivelli, F. Sparla, A. Thumiger, M. Calvaresi, L. Marri, G. Falini, F. Zerbetto and P. Trost, *J. Biol. Chem.*, 2012, **287**, 21372–21383.

67 S. Horowitz and R. C. Trievel, *J. Biol. Chem.*, 2012, **287**, 41576–41582.

68 R. Mahalakshmi and P. Balaram, in *D-Amino Acids: A New Frontier in Amino Acid and Protein Research*, Nova Science Publishers, Inc., Hauppauge, NY, USA, 2006, pp. 415–428.

69 M. Srebro-Hooper and J. Autschbach, *Annu. Rev. Phys. Chem.*, 2017, **68**, 399–420.

70 *SMART & SAINT, Softw. Ref. Manuals, version 5.051 (Windows NT Version)*, Bruker Anal. X-ray Instruments Inc., Madison, WI, 1998.

71 G. M. Sheldrick, *SADABS-2008/1 - Bruker AXS Area Detect. Scaling Absorpt. Correct*, Bruker AXS Madison, WI, USA, 2008.

72 M. C. Burla, R. Caliandro, B. Carrozzini, G. L. Cascarano, C. Cuocci, C. Giacovazzo, M. Mallamo, A. Mazzone and G. Polidori, *J. Appl. Crystallogr.*, 2015, **48**, 306–309.

73 G. M. Sheldrick, *Acta Crystallogr., Sect. C: Struct. Chem.*, 2015, **71**, 3–8.

74 C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek and P. A. Wood, *J. Appl. Crystallogr.*, 2008, **41**, 466–470.