CHEMISTRY
A European Journal

A Journal of
ChemPubSoc Europe

Accepted Article

Title: A Lasso-Inspired Bicyclic Peptide: Synthesis, Structure and Properties

Authors: Helena Martin-Gómez, Fernando Albericio, and Judit Tulla-Puche

To be cited as: *Chem. Eur. J.* 10.1002/chem.201803899

Link to VoR: http://dx.doi.org/10.1002/chem.201803899

Supported by
ACES

WILEY-VCH

10.1002/chem.201803899

WILEY-VCH

FULL PAPER

A Lasso-Inspired Bicyclic Peptide: Synthesis, Structure and Properties

Helena Martin-Gómez, [a] Fernando Albericio [b][c][d] and Judit Tulla-Puche* [b][c][e]

Abstract: The chemical synthesis of a bicycle inspired by the natural lasso peptide sungsanpin using a combination of solid-phase and in-solution chemistries is described. The bicyclic-derived topoisomer was designed introducing a covalent linkage between the ring and the loop, which allowed the tying of these two parts of the peptide rendering the bicyclic structure. Several structural techniques, such as MS fragmentation, ion-mobility and NMR were used to characterize the bicycle. Ion-mobility spectroscopy studies revealed that it showed lasso-like behavior. Its 3D structure was predicted on the basis of the NMR restraints. In addition, the high proteolytic and thermal stability of the bicycle potentially make it a suitable scaffold for epitope grafting.

Introduction

Lasso peptides are natural products that exhibit a broad spectrum of bioactivities. They consist of 15–26 proteinogenic amino acids and share an *N*-terminal 7- to 9-residue macrolactam ring.[1] The *C*-terminal tail threads through an *N*-terminal macrolactam ring in a right-handed conformation, which provides them a unique 3D conformation. The rigid and folded structure of these peptides is the result of a steric lock, in which a bulky and often an aromatic side chain is unable to pass through the macrolactam ring. This constrained structure confers lasso peptides extraordinary stability against chemical, thermal and proteolytic degradation.[1–4] They show a well-defined secondary structure; sometimes presenting β-sheets associated with β-turns, forming a hairpin-like structure despite their relatively small length.[5] These distinctive features could make them promising candidates as molecular scaffolds for drug design.

Importantly, no successful chemical synthesis approaches to lasso peptides have been reported to date due to the difficulties encountered in building and maintaining the threaded lasso structure. Recently, peptide-based [1]rotaxanes were synthesized, which were used as molecular grafting scaffolds.[6] Here, we focused on the class II lasso peptides sungsanpin[7] **1** and chaxapeptin[8] **2**. Both peptides contain 15 residues, and the 7-amino acid tail is threaded through the ring via hydrophobic interactions. These peptides were isolated from two distinct *Streptomyces* strains, which share 16S rRNA gene sequence similarity of 94%.[8] An analysis of their sequences reveals that they share the same number and chemical character of residues (Figure 1). They differ in only four residues, three of which are located at the *C*-terminal tail. In this position, NFF residues are replaced by SWL, thus maintaining hydrophobicity, and there is also a conservative change in the ring, namely Leu instead of Ile. Regarding their activity, both peptides show inhibitory effect on the invasion of human lung cancer cells.[8]

Here, we describe the first chemical synthesis of a bicyclic peptide **5** in which the loop sequence is tied to the ring through a covalent bond taking as basis the natural lasso peptide sungsanpin **1** (Figure 2). For this purpose, we have designed a synthetic route in which the bicyclic structure was achieved first, through the ring and the loop, creating a bridged bicyclic peptide.[9] The *C*-terminal tail was then introduced. It was known that once the ring is closed the *C*-terminal tail cannot be trapped due to the steric hindrance in these positions.[10] The synthetic approach has been designed using the feasible and efficient solid-phase peptide synthesis (SPPS),[11] in combination with solution chemistry, which can facilitate the last synthetic steps that are more demanding due to the steric rigidity already present in the intermediates. Chaxapeptin **2**, which was obtained via homologous expression (see Supplementary information), was used as a control for the characterization studies and structural analysis.

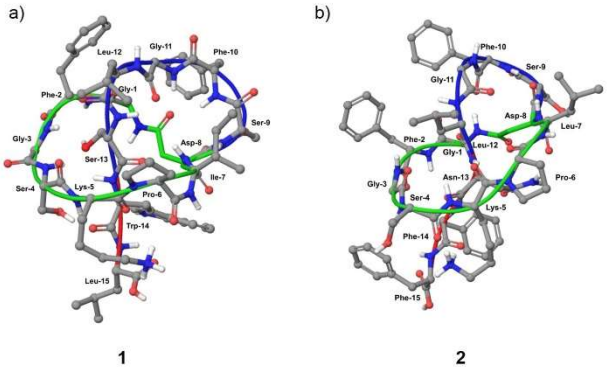


Figure 1. Structures of the natural lasso peptides (A) sungsanpin **1** and (B) chaxapeptin **2** (PDB code 2N5C). The ring residues are shown in green ribbon, the loop amino acids in blue and the tail amino acids in red.

[a] Dr. H. Martin-Gómez
Institute for Research in Biomedicine
Baldiri Reixac 10, 08028 Barcelona, Spain
[b] Prof. F. Albericio, Dr. J. Tulla-Puche
Department of Inorganic and Organic Chemistry – Organic Chemistry Section
University of Barcelona
Martí i Franquès 1-11, 08028 Barcelona, Spain
E-mail: judit.tulla@ub.edu
[c] Prof. F. Albericio, Dr. J. Tulla-Puche
CIBER-BBN, Networking Centre on Bioengineering, Biomaterials and Nanomedicine
Baldiri Reixac 10, 08028 Barcelona, Spain
[d] Prof. F. Albericio
School of Chemistry and Physics
University of KwaZulu-Natal
Durban 4001, South Africa
[e] Dr. J. Tulla-Puche
Institut de Biomedicina de la Universitat de Barcelona (IBUB)
08028 – Barcelona, Catalonia, Spain

Supporting information for this article is given via a link at the end of the document.

10.1002/chem.201803899

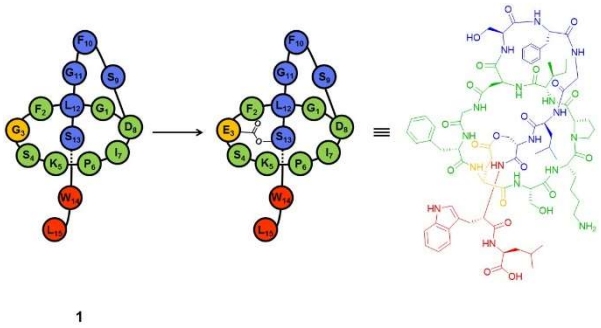
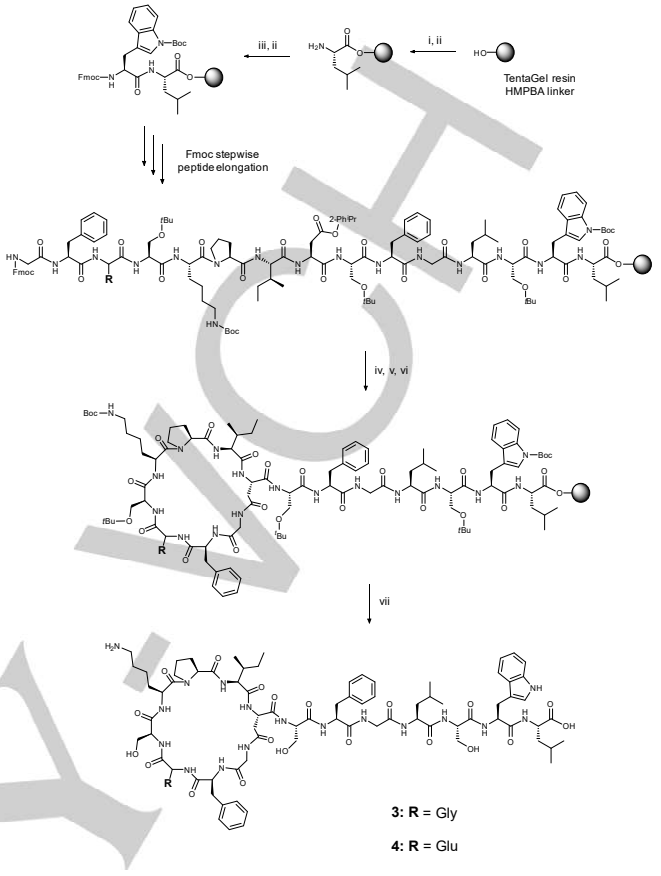


Figure 2. Schematic representation of the target peptide. Modification between sungsanpin **1** and the target peptide. The ring residues are shown in green, the loop amino in blue, the tail amino acids in red and Glu3 in yellow.

Results and Discussion

Towards the synthesis of sungsanpin: design and synthetic approach

As previously mentioned, once the macrolactam bond is closed, the *C*-terminal tail cannot be trapped inside. This was confirmed with the tentative synthesis of sungsanpin **3** using stepwise Fmoc-SPPS (Scheme 1). The synthesis was performed on Wang-type resins (TentaGel S-NH_2) with the 3-(4-hydroxymethylphenoxy)propionic acid (HMPBA) linker. In addition to the *t*Bu based protecting group, the β-carboxylic acid of Asp8 was protected in form of 2-phenylisopropyl ester (2-Ph^{i}Pr). This hindered protecting group can be removed with mild acid conditions (5% TFA in DCM), and is compatible with Boc and *t*Bu. 2-Ph^{i}Pr was used to prevent aspartimide formation between Asp8 and Ser9.[12] The cyclization was carried out on resin between the N^{α}-amino of the Gly1 and the side-chain carboxyl group of Asp8 released from its 2-Ph^{i}Pr ester by treatment with a low percentage of TFA. The main goal of this synthesis was to observe whether the tail was able to thread into the ring unaided. Once the cyclization was carried out, a treatment with TFA-TIS-H_2O (95:2.5:2.5) rendered the unprotected peptide. After several attempts of folding only the structure corresponding to the branched-cyclic sungsanpin **3** was obtained. Thus, we can conclude that using a straightforward strategy, which involved SPPS peptide elongation and cyclization, is not enough for the obtention of this intriguing peptide.



Scheme 1. Solid-phase synthesis toward peptides **3** and **4**. (i) Fmoc-L-Leu-OH, DIPCDI, DMAP, CH_2Cl_2-DMF (9:1), 2 h + 3 h; (ii) piperidine-DMF (1:4) (1 x 1 min, 1 x 3 min); (iii) Fmoc-L-Trp(Boc)-OH, DIPCDI, OxymaPure, DMF, 30 min; (iv) TFA-TIS-CH_2Cl_2 (5:1:94), 4 x 10 min; (v) piperidine-DMF (1:4) (1 x 1 min, 1 x 3 min, 1 x 5 min, 1 x 30 min); (vi) PyBOP, HOAt, DIEA, DMF, 1 h; and (vii) TFA-TIS-H_2O (95:2.5:2.5), 1 h.

Design and synthesis of the lasso-inspired bicyclic peptide via ester bond

The synthesis of peptide **5** was designed to ensure that the *C*-terminal tail was trapped inside the ring. We hypothesized that this could be done by tying the peptide loop to the ring through a covalent linkage, creating a bicycle. An ester bond was chosen because it could be later in principle easily removed while maintaining the lasso structure. The Gly at position 3 was replaced by Glu to form the ester bond between the side chain of Glu3 and the hydroxyl group of Ser13. Two branched peptide chains—linked by the ester bond—were elongated. One of these chains corresponded to the ring and the other to the loop. The linkage of the other extreme of the loop with the ring was then built, followed by a final cyclization step to close the ring (Figure 3). For the next sets of experiments, when the bicyclic peptide was planned to be synthesized having an extra ester bond between the hydroxyl group of Ser13 and the γ-carboxylic acid of Glu at position 3 instead of Gly, the corresponding branched-cyclic analog **4**, with the same topology of **3**, was synthesized as a control, following the same procedure described in Scheme 1.

10.1002/chem.201803899

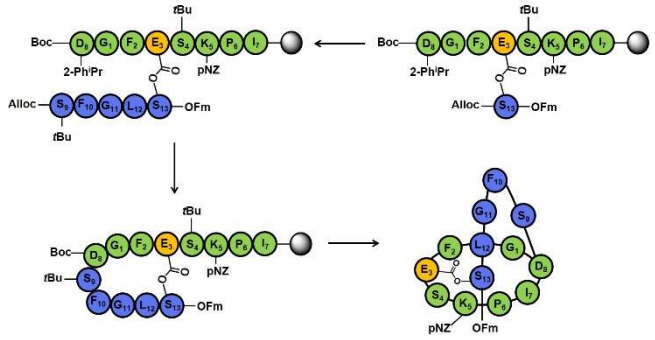
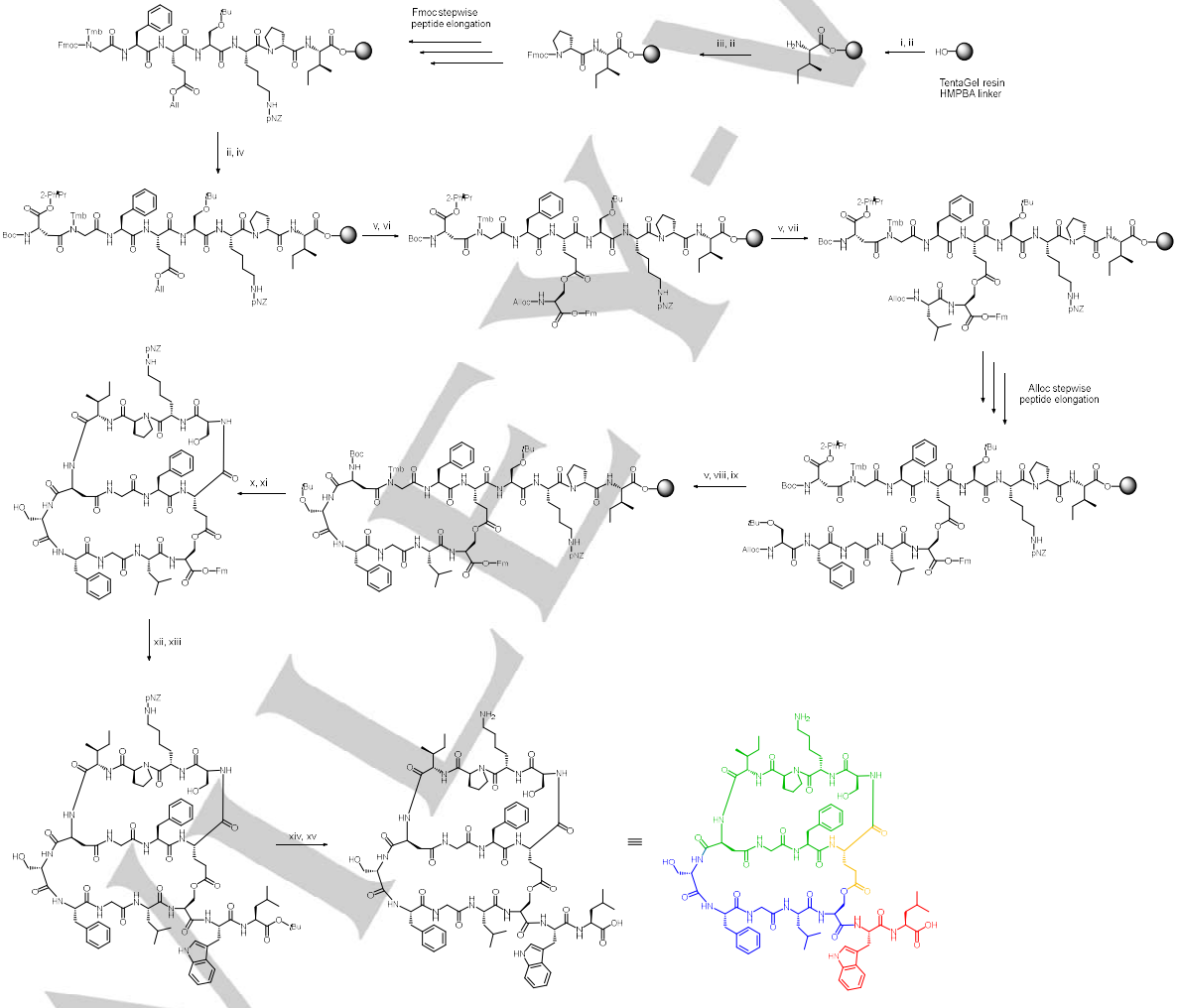


Figure 3. Schematic representation of the main important synthetic steps toward the bicycle formation of the peptide **5**. (i) Elongation of the second branched peptide chain using Alloc-chemistry, (ii) First cyclization between the loop and the ring and (iii) cleavage and macrolactamization to close the ring. The ring residues are shown in green, the loop amino in blue and Glu3 in yellow.

The synthesis of this bicyclic peptide **5** represents a "tour de force" as it will imply the use of several orthogonal protecting groups (Scheme 2).[13] Furthermore, the sequence Asp-Gly present in the sequence is very prone to give aspartimide side-reaction.[14] The synthesis was carried out on HMPBA-Wang-TentaGel resin with low loading to facilitate both the elongation of the two peptide chains and the on-resin cyclization. The first peptide chain was starting from the *C*-terminal Ile7 and finished at Asp8. Asp is a key residue in the peptide sequence because it is the only residue involved in the ring and the loop. For this reason, a unique protection strategy was designed for this residue: (i) Boc was used as protecting group of the amino group, because Asp8 was located at the *N*-terminal of the first peptide chain elongated. This position will pair with the carboxyl group of the *C*-terminal Ile7 which is anchored to the HMPBA linker. Both functions will be released during the last TFA treatment; (ii) the 2-Ph^{i}Pr was used



Scheme 2. Synthetic strategy toward the bicyclic peptide 5. (i) Fmoc-L-Ile-OH, DIPCDI, DMAP, CH_2Cl_2-DMF (9:1), 2 h + 3 h; (ii) piperidine-DMF (1:4) (1 x 1 min, 1 x 3 min); (iii) Fmoc-L-Pro-OH, DIPCDI, OxymaPure, DMF, 30 min; (iv) Boc-L-Asp-O(2-Ph^{i}Pr) **8**, HATU, DIEA, DMF, 1 h; (v) $Pd(PPh_3)_4$, $PhSiH_3$, CH_2Cl_2, 3 x 15 min; (vi) Alloc-L-Ser-OFm **10**, DIPCDI, DMAP, CH_2Cl_2-DMF (9:1), 2 h + 3 h; (vii) Alloc-L-Leu-OH 11, DIPCDI, OxymaPure, DMF, 30 min; (viii) TFA-TIS-H_2O (5:1:94), 4 x 10 min; (ix) PyBOP, HOAt, DIEA, DMF, 2 h; (x) TFA-TIS-H_2O (95:2.5:2.5), 1 h; (xi) PyBOP, HOAt, DIEA, CH_2Cl_2-DMF (1:1), 0 ºC → rT, 2 h; (xii) DEA, DCM, 1 h; (xiii) H-Trp-Leu-O*t*Bu **15**, DIPCDI, OxymaPure, DMF, 3 h; (xiv) TFA-DCM (1:1), 30 min; and (xv) 6 M $SnCl_2$ in DMF, 1.6 mM HCl/dioxane, 2 x 45 min. The ring residues are shown in green, the loop amino in blue, the tail amino acids in red and Glu3 in yellow.

to protect the α-carboxylic acid, thus allowing the β-carboxylic acid to remain free. To prevent aspartimide formation at the Asp8-Gly1 sequence, 2,4,6-trimethoxybenzyl (Tmb) backbone protection was used for blocking the NH Gly1. In addition, to ensure orthogonal deprotection before the formation of the ester bond, allyl was selected as protecting group of the α-carboxylic acid group of Glu3. Moreover, the standard side-chain Boc protecting group of Lys5 was substituted by *p*-nitrobenzyloxycarbonyl (pNZ), yielding compound **9**.[15] This modification was performed because the free amino group of the Lys5 side chain may compete during the second cyclization in solution, thus producing a non-desired cyclization. The γ-carboxylic acid of Glu3 was protected with allyl, which was removed before the esterification of that carboxylic acid with the hydroxyl group of Ser13. Alloc-chemistry was then used for the elongation of the second peptide chain to avoid the use of piperidine treatment that can produce aminolysis of the ester bond, serine *N* → *O* transacylation, and DKP formation. For this reason, new protecting scheme for Ser13 was designed. 9-Fluorenylmethyl ester (Fm) was used as protecting group of the α-carboxylic acid rather than of the amino group, which this time was protected with Alloc, yielding compound **10**.[16]

The formation of the ester bond proceeded via the *O*-acylisourea activation. A high excess of DIPCDI and Alloc-L-Ser-OFm **10** was used to ensure its fast and efficient formation. Short reaction times and high excesses were needed to prevent the *N* → *O* transacylated product.

After elongation of the second peptide chain, selective deprotection of the Alloc and 2-Ph^{i}Pr groups and addition of PyBOP-HOAt-DIEA provided the first cyclization. The peptide was then cleaved, and the second cyclization, corresponding to the ring closure, was performed in solution in order to provide a more flexible environment for the peptide backbone during cyclization. At this point the ring and the loop sequences, linked by the ester bond, set up a bicycle. The Fm from the α-carboxylic acid of Ser13 was then removed and the *C*-terminal dipeptide **15** was introduced in a straightforward way after the bicycle formation. The main advantage of this last coupling was that the steric hindrance of the ring was avoided, and the tail could be introduced through the middle of the ring. Finally, *t*Bu and pNZ protecting groups were removed using TFA and $SnCl_2$, respectively. A special beauty of this synthetic approach is that the integrity of the ester bond during the synthesis was preserved by avoiding the use of harsh basic conditions, thus using Alloc and pNZ chemistries, which can be removed smoothly using Pd(0) and $SnCl_2$.

Thermal and proteolytic stability

Peptides **2-5** were subjected to proteases and thermal stability assays. The bicyclic peptide **5** remained intact after both thermal and carboxypeptidase Y treatment, and only the *C*-terminal Leu15 was digested. In contrast, the corresponding branched-cyclic topoisomer **4** and **3** were readily degraded, showing cleavage between Ser9 and Phe10. Regarding this result, **5** was tested against a stronger protease, pepsin. After 2 h, the peptide remained intact and after 24 h only 12% had been degraded (Figure S1-S3). The peptide was digested between Asp8-Ser9 and Phe10-Gly11, with the concomitant loss of the dipeptide Ser9-Phe10 (*m/z* 1407.5). This result indicated the extraordinary stability of the bicyclic peptide **5**, in which long times were required before degradation was observed. As expected, the natural lasso peptide, chaxapeptin **2**, remained intact after proteases treatment.[8] However, there are several thermal sensitive lasso peptides and also carboxypeptidase Y sensitive in which the *C*-terminal residues are degraded. One such example is sphingopyxin I,[17] in which the last two residues were digested and the structure unfolds after thermal assay.

MS/MS analysis

Peptides **2-5** were extensively analyzed by tandem mass spectrometry (Table 1). The fragmentation pattern of **3** and 4 was very similar (Figure S4-S5), showing only fragmentation around the tail, while the cycle remained intact. On the other hand, chaxapeptin **2** showed a very similar fragmentation pattern to that of the two aforementioned peptides, where b_{12} was also the most intense signal (*m/z* 1188.60) (Figure S6). However, the signal intensities are different, which can be attributed to the constrained structures of native lasso peptides.[5],[18],[19]

Table 1. Main MS^2 fragmentation of peptides **2-5**, being the first the most abundant.

Peptide	2	3 and 4	5
Fragmentations	**b_{12}** Leu12-Asn13	**b_{12}** Leu12-Ser13	**b_{13}** Ser13-Trp14
	b_{11} Gly11-Leu12	**b_{13}** Ser13-Trp14	**$b_2 + y_{12}$** Ser4-Lys5/Lys5-Pro6
	b_{13} Asn13-Phe14	**b_{10}** Phe10-Gly11	**$b_2 + y_{11}$** Glu3-Ser4/Lys5-Pro6
	b_{10} Phe10-Gly11		**b_{12}** Leu12-Ser13

A totally different fragmentation spectrum was obtained for bicyclic peptide **5**. In this case, fragmentation was observed in the macrolactam ring (Figure S7). The fragmentation pattern observed was more similar to class I and III lasso peptides, than to class II. In this regard, class I and III show a weaker fragmentation behavior than class II and also fragmentation in the ring, thus hindering the distinction between a bicyclic and a lasso peptide.[2],[20]

In conclusion, we were unable to distinguish between our bicyclic peptide and a lasso-structured peptide based on the MS^2 studies, as the weak fragmentation behavior could also be a consequence of the fact that more than one bond breakage was required for the fragmentation to be detected.

IM-MS analysis

Three characteristic IM-MS trends have been established for class II lasso peptides.[21] Regarding the first trend when the charge state was increased, the CCS remained relatively

constant for peptides **4** and **5** (Table 2). The $[M + 3H]^{3+}$ was the most abundant charge state, as it was present for all the peptides, while the $[M + 4H]^{4+}$, which is characteristic of unfolded branched cyclic forms, was present only for **3**.[21–23] The similar behavior of **4** and **5** can be related to the stabilization through hydrogen bonds in the gas-phase of peptide **4**.

The low range of CCS $\Delta\Omega/\Omega$ covered by all protonated molecules of bicycle **5** is consistent with a rigid and constrained structure, even the value (0.6 %) is lower value than the natural lasso peptide **2** (3.9 %) (Table S1).[23]

Considering the second trend, low intensity of the highly protonated species is expected for lasso peptides (Figure S8). The bicyclic peptide **5** showed a slightly different spectrum after the addition of sulfolane compared to **2**, in which the intensity of the $[M + H]^+$ peak decreased, being the doubly protonated ion the most abundant. This decrease was more readily detected for **3**, in which the peak of $[M + H]^+$ dropped considerably after the addition of sulfolane. Regarding the parameter (ζ),[21],[23]] which provides a comparison of both the intensity and charge state, the bicycle **5** showed the lowest values, indicating a rigid and constrained structures (Table S2).

Regarding the width of the ion-mobility peak on the $[M + 3H]^{3+}$ charge state, peptides **4** and **5** showed a similar peak profile than chaxapeptin **2** with narrow IM peaks (Figure S9). Although the peptide sequence of **4** differed from that of **3** in only one residue, the shape of the IM peak was more similar to that of the natural lasso peptide **2** and to the bicyclic peptide **5**. This observation again reflects the importance of the structure stabilization effect caused by the replacement of Gly3 by Glu3, which produces a structure that resembles a bicyclic peptide (Figure S10).

NMR analysis

On the basis of the previous results, an NMR analysis was carried out to assess the 3D structure of peptides **3** and **5**. The NMR data for structural characterization of **3** was recorded in pyriridne-d_5 and was compared with those of natural sungsanpin (Table S3).[7] The general broadness of the spectral lines and the poor dispersion of amide signals led us to conclude that **3** did not show a lasso structure (Figure S11-S14). This result was also confirmed by the NMR-derived structure (Figure S15). The long-range NOEs between the ring and the loop revealed that the linear part was folded over the ring, but with no threading.

To achieve better resolution, we tested two solvents. To this end, the NMR data of the bicyclic peptide **5** was recorded in DMSO-d_6 (Figure S16-S22) and CD_3CN-H_2O (1:5) (Figure S23-S25). Duplicity in the residues Trp14, Leu15 and Pro6 was observed, indicating that two conformations were present. On the one hand, this duplicity could be assigned to the *cis/trans* isomerization. [24] On the other hand, the duplicity of the *C*-terminal residues (Trp14 and Leu15) pointed to high flexibility on this part, which is not constrained by the bicyclic structure.

Exhaustive analysis of the fingerprint region of bicyclic peptide **5** revealed poor ^1H signal dispersion in the amide proton region compared to lasso peptides.[25] In this regard, the chemical shift dispersion ($\Delta\delta$) of the bicyclic peptide **5** was low compared to reported lasso peptides (Figure S26). The chemical shift deviation values of αH of **5** did not show a clear tendency (Figure 4). The negative values for the Gly1-Ser4 segment could indicate an α-helical fold. These findings may indicate the presence of α-turns.

Table 2. Number of conformations and CCS of the four peptides with 400 mM of sulfolane.

Peptide	Charge state[a]	no. of conformations	CCS (Å) (Ratio(%))[b]
2	$[M + 2H]^{2+}$	4	290 (10)/**307 (53)**/ 330 (33)/367 (4)
	$[M + 3H]^{3+}$	4	281 (18)/**295 (52)**/ 312 (23)/337 (8)
3	$[M + 2H]^{2+}$	3	346 (39)/328 (18)/ **312 (43)**
	$[M + 3H]^{3+}$	3	388 (22)/**352 (63)**/ 310 (15)
	$[M + 4H]^{4+}$	2	391 (25)/**426 (75)**
4	$[M + 2H]^{2+}$	4	297 (24)/**321 (43)**/ 346 (29)/372 (4)
	$[M + 3H]^{3+}$	4	305 (16)/**323 (52)**/ 337 (23)/357 (9)
5	$[M + 2H]^{2+}$	4	306 (10)/**319 (61)**/ 340 (27)/373 (1)
	$[M + 3H]^{3+}$	4	296 (27)/**317 (51)**/ 337 (18)/350 (5)

[a] Charge state $[M + H]^+$ is not shown because it is an unfolded structure with the highest drift time values and is not relevant for the comparison. [b] Each CCS was calculated based on the drift times.[26] When several conformations are present, the major species is highlighted in bold.

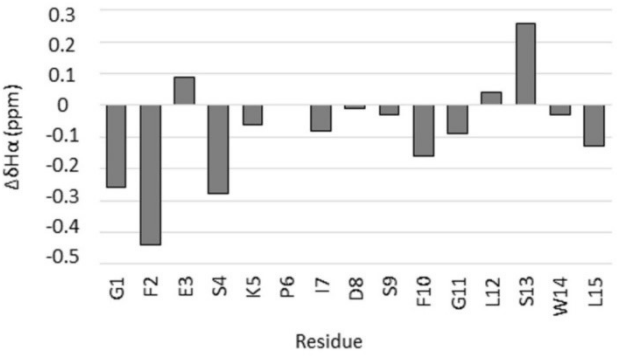


Figure 4. Chemical shift deviations histogram from random coil for ΔHα in the bicyclic peptide **5** in DMSO-d_6. Random coil values were extracted from Wüthrich et *al*.[27,28]

10.1002/chem.201803899

Regarding the long-range NOEs between the C-terminal tail and the N-terminal macrolactam ring, bicyclic peptide **5** show quite sequential NOEs and few of them were located between the tail and the ring. A better dispersion was achieved in CD_3CN-H_2O (1:5) than in DMSO-d_6, thus allowing the measurement of some long-range NOEs (Table 3). Contrary to what we expected, the rigid structure of the bicycle was not enough to provide an organized and structured spectrum**.**

Table 1. Long-range NOEs between the *C*-terminal tail and the *N*-terminal ring of the bicyclic peptide **5**.

NOE cross-peaks[a]
NH Phe2 - Hβ, Hγ, Hδ Leu12
Hα Pro6 - Hγ, Hδ Leu12
NH_2 Lys5 – Hδ Leu12
H_2 Trp14 – Hβ Ser9; Hα, Hβ Phe2; Hγ Leu12

[a] From the NOESY spectrum in CD_3CN-H_2O (1:5) at 298 K.

Computational studies

To determine the 3D structure of the bicyclic peptide **5**, a conformational sampling algorithm was performed with the NMR-derived distance restraints (see Experimental Section for more details). To verify the compatibility of the NOE restraints, two possible structural conformations (threaded and unthreaded) were designed as starting point (Figure S27). Although the long-range NOEs were not conclusive, they were enough to guide the simulation of the bicyclic 3D structure. The initial and final structures were different for the threaded bicycle (RMSD = 6.3 Å), thereby revealing the incompatibility of the NOE-derived restraints with the threaded structure. On the other hand, the low deviation value (RMSD = 2.7 Å) showed that the starting structure with the unthreaded tail was closer to the one modeled with restraints. From the two starting conformations, the resulting minimal-energy structures displayed similar features and high similarity (RMSD = 4.2 Å) (Figure 5 b). Moreover, all the calculated minimal-energy conformations led to the same bicyclic topology, with the *C*-terminal tail outside the macrocyclic ring (Figure 5 a).

Conclusions

In summary, a combination of a solid-phase/solution chemistry hybrid approach allowed the synthesis of a bicyclic peptide inspired by the natural lasso peptide sungsanpin. The key of the synthetic approach relies in a proper selection of large number of protecting groups: (i) Alloc, Boc, Fmoc and pNZ for the amino function; (ii) allyl, Fm, 2-Ph^{i}Pr, *t*Bu and the oxymethyl from the HMPBA linker for the carboxylic function and (iii) Tmb for the amide of Gly. The combined use of these 10 protecting groups, which removals require mostly orthogonal conditions (low and high percentage of TFA, piperidine, Pd(0) and $SnCl_2$), has allowed the synthesis of this bicycle. The NMR-based experiments and the structural calculations, together with the results of the MS^2 fragmentation, ion-mobility and protease assays, confirmed that bicyclic peptide

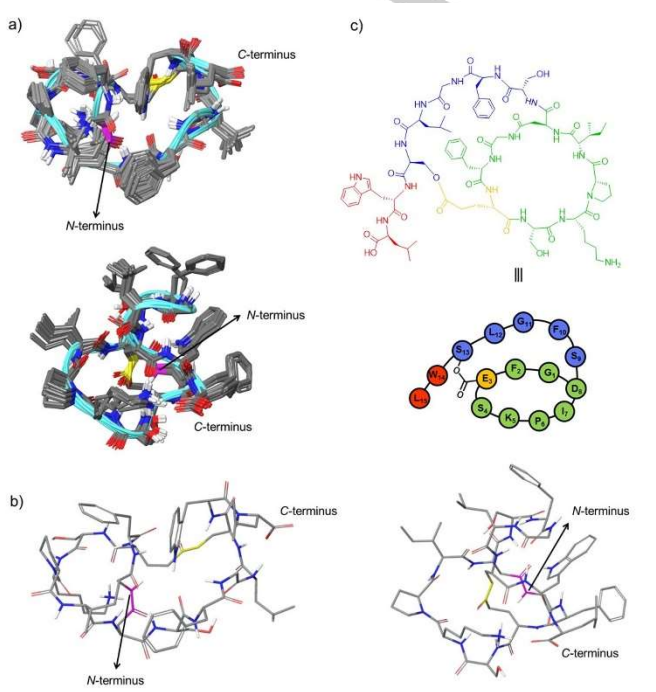


Figure 5. 3D structure of the bicyclic peptide **5**. (A) Ensemble of the 20-lowest energy structures derived from the NMR restraints in CD_3CN-H_2O (1:5): Above - Starting structure with the threaded tail and below - with the unthreaded tail. The maximum RMSD values between structures 1-20 are 0.6 and 0.4 Å, respectively. (B) Lowest-energy structures for the threaded and unthreaded starting conformation, from left to right. The ester bond is shown in yellow and the macrolactam bond in pink. (C) Chemical structure and schematic representation.

5 had the *C*-terminal tail outside the ring. The covalent linkage between the tail and the ring is responsible for providing proteolytic stability but was not sufficient to provide a well-defined structure. Despite the topology, the low collision cross section values and the narrow peak width observed for bicyclic peptide **5** indicated that it showed lasso-like behavior. Moreover, preliminary biological assays demonstrated that bicycle **5** showed similar cytotoxicity and invasion activity than chaxapeptin **2** (Figure S28-S29). On the other hand, the branched-cyclic peptide **4** displayed similar ion-mobility behavior as **5** with the electrostatic interaction Glu3-Lys5 still present in the gas phase, thus stabilizing the bicyclic structure. Despite this conformation, peptide **4** did not show proteolytic stability. Given the singular characteristics of **5**, it may be a good candidate as molecular scaffold for epitope grafting toward the generation of more potent and more selective bioactive compounds. Finally, this synthesis is expected to facilitate the future development of peptides as therapeutics tools

Experimental Section

Materials and reagents

Dichloromethane, dimethylformamide, methanol, diethyl ether were purchased from Scharlau; while acetone and acetonitrile were obtained from SDS. All the reagents used for the synthesis were supplied by the following: Bachem AG, Iris Biotech, Luxemburg Industries, Novabiochem, Panreac and Sigma-Aldrich. Carboxypeptidase Y was purchased from Alfa Aesar.

Analytics

Analytical RP-HPLC was performed on a Waters Alliance 2695 (Waters, Milford, MA) with an automatic injector and a photodiode array detector (Waters 2998 or Waters 996) and software Empower2. The columns used were a Xbridge™ C18 2.5 µm (4.6 mm x 75 mm) reversed-phase analytical column and a Sunfire™ C18 3.5 µm (4.6 mm x 100 mm) reversed-phase analytical column run with linear gradients of ACN (0.036% TFA) into H_2O (0.045% TFA) over 8 min, and a Phenomenex Luna C18 5 µm (4.6 mm x 250 mm) reversed-phase analytical column run with linear gradients over 40 min. UV detection was at 220 nm and the system was run at a flow rate 1 mL/min. For semi-analytical HPLC, the columns used were a Xbridge™ Prep BEH130 C18 5 µm (10 mm x 100 mm) reversed-phase analytical column and a Sunfire® C18 OBD™ Prep Column 5 µm (10 x 150 mm) reversed-phase column. Linear gradients of ACN (0.036% TFA) into H_2O (0.045% TFA) were run over 20 min. UV detection was at 220 nm and the system was run at a flow rate 3 mL/min. Semi-preparative RP-HPLC was carried on Waters instrument comprising a 2545 gradient module equipped with a sample manager module with automatic injector and fraction collector (Waters Alliance 2767), a Waters Alliance 2487 dual wavelength absorbance detector and a MassLynx v3.5 system controller. The column used was a XBridge® Prep C18 OBD™ 5 µm (19 x 100 mm) reversed-phase column. UV detection was at 220 and 254 nm and linear gradients of ACN (0.1% TFA) into H_2O (0.1% TFA) were run at a flow rate of 16 mL/min over 30 min.

MS-MS fragmentation

MS^2 was performed on LTQ-FT Ultra mass spectrometer (Thermo Scientific). 0.1 mg was reconstituted with 500 µL of water (100 µM). 10 µL (100 µM) of sample was diluted 1/1 with ACN (1% FA) to obtain 20 µL of 50 µM. Sample solutions were directly injected and were infused by automated chip-based nanoelectrospray using a Nanomate system (AdvionBioSciences, Ithaca, NY, USA) as the interface. The ionization was performed in positive mode using a spray voltage and a gas pressure of 1.75 kV and 0.5 psi, respectively. Capillary temperature and voltage were set up to 200 ºC and 44 V, respectively. The fragmentation technique used was CID (collision induced dissociation). The instrument was calibrated over the *m/z* range 200-2000. Data were acquired with Xcalibur software, vs.2.0SR2 (ThermoScientific).

IM-MS

IM-MS experiments were performed using a Synapt G1-HDMS mass spectrometer (Waters, Manchester, UK). 0.1 mg was reconstituted with 500 µL of water (100 µM). Sample were diluted 1/10 in H_2O-ACN (1:1) with 4% (FA) to obtain 10 µM, 1/10 in H_2O-ACN (1:1) with 4% FA and 200 mM sulfolane or 1/10 in H_2O/ACN (1:1) with 4% FA and 200 mM or 400 mM sulfolane. Sample solutions were directly injected and were infused by automated chip-based nanoelectrospray using a Triversa Nanomate system (AdvionBioSciences, Ithaca, NY, USA) as the interface. The ionization was performed in positive mode using a spray voltage and a gas pressure of 1.75 kV and 0.5 psi, respectively. Cone voltage, extraction cone and source temperature were set to 40 V, 3 V and 20 ºC, respectively. Trap and transfer collision energies were set to 6 V and 4 V, respectively. The pressure in the Trap and Transfer T-Wave regions were $5.99 \cdot 10^2$ mbar of Ar and the pressure in the IMS T-Wave was 0.477 mbar of N_2. Trap gas and IMS gas flows were 8 and 25 mL/sec, respectively. The travelling wave used in the IMS T-Wave for mobility separation was operated at a velocity of 250 m/sec. The wave amplitude was a linear ramp from 1.0 to 28.0 V. The bias voltage for entering in the T-wave cell was 15 V. The instrument was calibrated over the m/z range 200-3000 Da using a solution of cesium iodide. MassLynx version 4.1 SCN 704 and Drift scope version 2.4 software were used for data processing. Ion mobility data analysis was performed with Driftscope software vs. 2.4 integrated in MassLynx software.

NMR studies

^1H and ^{13}C NMR spectra were recorded on a Varian MERCURY 400 (400 MHz for ^1H NMR, 100 MHz for ^{13}C NMR) spectrometer for organic small molecules and on a Bruker 600 Avance II Ultrashield, provided with a cryoprobe TCI (600 MHz for ^1H NMR, 150 MHz for ^{13}C NMR) spectrometer for peptidic samples. Chemical shifts (δ) are expressed in parts per million (ppm) downfield tetramethylsilane (TMS). Coupling constants are expressed in Hertz (Hz). Xaa-Pro bond conformation was determined using a previously described method based on the ^{13}C chemical shift statistics.[24]

Computational studies

They were carried out using the Macrocyclic Conformational Sampling tool.[29] The coordinates for the peptides were extracted from the crystal structure of chaxapeptin **2** (PDB code: 2N5C). Then, Leu7, Asn13, Phe14 and Phe15 were manually mutated to Ile, Ser, Trp and Leu, respectively for all peptides. One additional mutation for **3** and **5** was required, the exchange of Gly3 for Glu3. Bicyclic peptide **5** was manually cyclized through the side chain of Asp8 and the ester bond through the side chains of Glu3 and Ser13. Additionally, the tail was manually unthreaded from the ring in the other bicyclic conformation, used as a starting structure. On the other hand, for peptides **3** and **4**, the tail was unthreaded and was cyclized only through the side chain of Asp8. The resulting structures were used as input for the Macrocycle Conformational Sampling tool of Schrödinger. Constraints from the NOESY cross-peaks were applied to each structure (Table S4). NOEs were classified as strong, medium and weak (upper limits for structure calculation were set as 2.2 Å, 3.0 Å and 4.0 Å, respectively). Conformations were generated using the OPLS-2005 force field and GBSA for electrostatic treatment. Conformations were kept when energies were below 10 kcal/mol, and redundant conformations (RMSD > 0.75 Å) were removed. For each peptide, 5000 simulation cycles were used with 5000 LLMOD search steps.

Acknowledgements

The authors wish to thank Dr. Margarida Gairí and Dr. Mireia Díaz-Lobo for helpful discussions on NMR and Ion Mobility data analysis, respectively. This study was partially funded by the CICYT (CTQ2015-67870-P to F.A., and CTQ2015-68677-R and EUIN2017-88320 to J.T.-P, and the Institute for Research in Biomedicine Barcelona (IRB Barcelona). H.M.-G. and J.T.-P. thank a MINECO for a Severo Ochoa Predoctoral fellowship and a Ramon y Cajal contract, respectively.

Conflicts of interest

There are no conflicts to declare

Keywords: lasso peptides • bicyclic peptides • solid-phase peptide synthesis • anticancer • orthogonal protection

[1] H. Martin-Gómez, J. Tulla-Puche, *Org. Biomol. Chem.* **2018**, *16*, 5065–5080.
[2] Y. Li, R. Ducasse, S. Zirah, A. Blond, C. Goulard, E. Lescop, C. Giraud, A. Hartke, E. Guittet, J. L. Pernodet, et al., *ACS Chem. Biol.* **2015**, *10*, 2641–2649.
[3] A. Blond, J. Péduzzi, C. Goulard, M. J. Chiuchiolo, M. Barthélémy, Y. Prigent, R. A. Salomón, R. N. Farías, F. Moreno, S. Rebuffat, *Eur. J. Biochem.* **1999**, *259*, 747–755.
[4] M. Zimmermann, J. D. Hegemann, X. Xie, M. A. Marahiel, *Chem. Sci.* **2014**, *5*, 4032–4043.
[5] K. J. Rosengren, R. J. Clark, N. L. Daly, U. Göransson, A. Jones, D. J. Craik, *J. Am. Chem. Soc.* **2003**, *125*, 12464–12474.
[6] F. Saito, J. W. Bode, *Chem. Sci.* **2017**, *8*, 2878–2884.
[7] S. Um, Y.-J. J. Kim, H. H. C. Kwon, H. Wen, S.-H. H. Kim, H. H. C. Kwon, S. Park, J. Shin, D.-C. C. Oh, *J. Nat. Prod.* **2013**, *76*, 873–879.
[8] S. S. Elsayed, F. Trusch, H. Deng, A. Raab, I. Prokes, K. Busarakam, J. A. Asenjo, B. A. Andrews, P. Van West, A. T. Bull, et al., *J. Org. Chem.* **2015**, *80*, 10252–10260.
[9] I. Di Bonaventura, X. Jin, R. Visini, D. Probst, S. Javor, B.-H. Gan, G. Michaud, A. Natalello, S. M. Doglia, T. Köhler, et al., *Chem. Sci.* **2017**, *8*, 6784–6798.
[10] A. L. Ferguson, S. Zhang, I. Dikiy, A. Z. Panagiotopoulos, P. G. Debenedetti, A. James Link, A. J. Link, *Biophys. J.* **2010**, *99*, 3056–65.
[11] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
[12] J. L. Lauer, C. G. Fields, G. B. Fields, *Lett. Pept. Sci.* **1995**, *1*, 197–205.
[13] F. Albericio, *Biopolymers* **2000**, *55*, 123–139.
[14] E. Nicolás, E. Pedroso, E. Girald, *Tetrahedron Lett.* **1989**, *30*, 497–500.
[15] A. Isidro-Llobet, J. Guasch-Camell, M. Álvarez, F. Albericio, *European J. Org. Chem.* **2005**, *2005*, 3031–3039.
[16] L. J. Cruz, N. G. Beteta, A. Ewenson, F. Albericio, *Org. Process Res. Dev.* **2004**, *8*, 920–924.
[17] J. D. Hegemann, M. Zimmermann, S. Zhu, D. Klug, M. A. Marahiel, *Biopolymers* **2013**, *100*, 527–542.
[18] J. D. Hegemann, M. Zimmermann, X. Xie, M. A. Marahiel, *J. Am. Chem. Soc.* **2013**, *135*, 210–222.
[19] T. a. Knappe, U. Linne, S. Zirah, S. Rebuffat, X. Xie, M. a. Marahiel, *J. Am. Chem. Soc.* **2008**, *130*, 11446–11454.
[20] T. A. Knappe, U. Linne, X. Xie, M. A. Marahiel, *FEBS Lett.* **2010**, *584*, 785–789.
[21] K. Jeanne Dit Fouque, C. Afonso, S. Zirah, J. D. Hegemann, M. Zimmermann, M. A. Marahiel, S. Rebuffat, H. Lavanant, *Anal. Chem.* **2015**, *87*, 1166–1172.
[22] K. Jeanne, D. Fouque, H. Lavanant, S. Zirah, J. Lemoine, S. Rebuffat, J. Tabet, A. Kulesza, C. Afonso, P. Dugourd, et al., *Rapid Commun. Mass Spectrom.* **2015**, 1411–1419.
[23] K. J. D. Fouque, H. Lavanant, S. Zirah, J. D. Hegemann, M. Zimmermann, M. A. Marahiel, S. Rebuffat, C. Afonso, *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 315–322.
[24] M. Schubert, D. Labudde, H. Oschkinat, P. Schmieder, *J. Biomol. NMR* **2002**, *24*, 149–154.
[25] X. Xie, M. A. Marahiel, *ChemBioChem* **2012**, *13*, 621–625.
[26] D. Smith, T. Knapman, I. Campuzano, R. Malham, J. Berryman, S. Radford, A. Ashcroft, *Eur. J. Mass Spectrom.* **2009**, *15*, 113.
[27] C. Grathwohl, K. Wüthrich, *J. Magn. Reson.* **1974**, *13*, 217–225.
[28] A. Bundi, C. Grathwohl, J. Hochmann, R. M. Keller, G. Wagner, K. Wüthrich, *J. Magn. Reson.* **1975**, *18*, 191–198.
[29] K. S. Watts, P. Dalal, A. J. Tebben, D. L. Cheney, J. C. Shelley, *J. Chem. Inf. Model.* **2014**, *54*, 2680–2696.

10.1002/chem.201803899

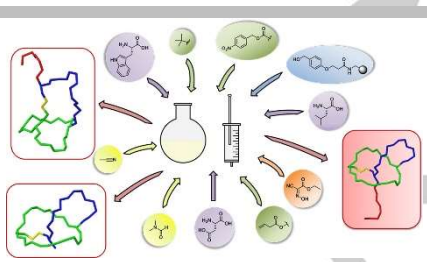
WILEY-VCH

Entry for the Table of Contents (Please choose one layout)

Layout 1:

FULL PAPER

The proper selection of large number of protecting groups together with the combination of solid-phase/solution chemistry hybrid approach, allowed the chemical synthesis of a bicyclic peptide inspired by the natural lasso peptide sungsanpin.



Helena Martin-Gómez, Fernando Albericio, Judit Tulla-Puche[]*

Page No. – Page No.

A Lasso-Inspired Bicyclic Peptide: Synthesis, Structure and Properties

[a] Dr. H. Martin-Gómez
Institute for Research in Biomedicine
Baldiri Reixac 10, 08028 Barcelona, Sp
[b] Prof. F. Albericio, Dr. J. Tulla-Puche
Department of Inorganic and Organic C
Chemistry Section
University of Barcelona
Martí i Franquès 1-11, 08028 Barcelona
E-mail: judit.tulla@ub.edu