Journal of Peptide Science
J. Pept. Sci. 2007; **13**: 728–741
Published online 10 September 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psc.880

JOURNAL OF PEPTIDE SCIENCE

Structural analysis of substance P using molecular dynamics and NMR spectroscopy

FRANCESC J. CORCHO,[a]* XAVIER SALVATELLA,[b] JOSEP CANTO,[a] ERNEST GIRALT[b] and JUAN J. PEREZ[a]

[a] Departament d'Enginyeria Química (UPC), ETS d'Enginyeria Industrial de Barcelona, Diagonal 647, 08028 Barcelona, Spain
[b] Institut de Recerca Biomèdica, Parc Científic de Barcelona and Departament de Química Orgànica, Facultat de Química, Universitat de Barcelona, Martí i Franqués, 1, 08028 Barcelona, Spain

Received 23 November 2006; Revised 2 May 2007; Accepted 4 May 2007

Abstract: The present work is a combined structural study, using Nuclear Magnetic Resonance (NMR) and Molecular Dynamics(MD), of the amidated and the free acid forms of substance P in water and methanol. The results obtained using both approaches were compared in order to characterize the structural features of both peptides in solution. From the NMR experiments it was derived that the free acid form adopts an extended conformation at the *N*-terminus and a helical conformation at the *C*-terminal segment of the peptide in both water and methanol; these structural features are in qualitative agreement with the results of the MD simulations. No significant differences in behavior were observed between the amidated and the free acid forms of the peptide in the simulations and in the experiments carried out in water, suggesting that the different activities of these analogs are due to their different mode of interaction with the receptor rather than to their structural preferences. Finally, we propose that the structure of substance P can be partially inferred from its sequence due to the presence of a Pro-X-Pro motif on the *N*-terminus and a Gly–Leu sequence on the *C*-terminus. Copyright © 2007 European Peptide Society and John Wiley & Sons, Ltd.

Supplementary electronic material for this paper is available in Wiley InterScience at http://www.interscience.wiley.com/jpages/1075-2617/suppmat/

Keywords: AMBER; substance P; molecular dynamics; NMR; folding; clustering techniques

INTRODUCTION

The characterization of the folding mechanism of peptides and mini-proteins by computational methods is a topic that has received much attention in recent years [1–3]. The study of folding processes of peptides and so-called mini-proteins using simulation is within the reach of currently available computational power [4–8]. If helix nucleation occurs in the nanosecond timescale as has been suggested [9], a 10-residue peptide would be expected to fold to its typically structurally heterogeneous and dynamic native state within a 10-ns trajectory. Substance P (SP) is a 11-residue neuropeptide of sequence Arg^1-Pro^2-Lys^3-Pro^4-Gln^5-Gln^6-Phe^7-Phe^8-Gly^9-Leu^{10}-Met^{11} involved in smooth muscle contraction, hypotension, salivation and depression. This peptide is generated through the cleavage of its precursor leading predominantly to the liberation of a dodecapeptide, that presents a terminal glycine (SP-Gly) [10]. This glycine-extended prohormone is converted to the undecapeptide SP with the amidated *C*-terminus by the sequential action of two enzymes, peptidylglycine α-hydroxylating mono-oxygenase and peptidylamindoglycolate lyase [11]. Since the acid form of substance P (SPOH), where the amidated methionine has been substituted by methionine, has no biological activity [12,13] it has been proposed that amidation confers to the peptide the structural or dynamic properties required for activity. Different spectroscopic studies of substance P have been carried out in water and organic solvents using diverse techniques, including NMR [14–17] and circular dichroism, Raman and IR spectroscopy [18,19] with the aim of assessing its structural and dynamic properties. Other spectroscopic studies were carried out in model membranes and surfactants using NMR spectroscopy [20–25]. Moreover, substance P was also studied by energy calculations [26,27] and, more recently, by simulations using MD [28–30]. The abovementioned studies clearly point out the dependence of the conformation of the peptide on its environment; indeed, experimental and simulation studies have shown that the structural and dynamical behavior of peptides is significantly affected by the solvent. Thus, organic solvents rigidify the peptide and favor secondary structure, whereas water allows for greater flexibility [31,32]. Regarding SP, NMR studies carried in aqueous solution suggest some degree of persistent structure in the region between residues 5 and 11 [14] whereas the peptide adopts an extended conformation in DMSO and pyridine. Furthermore, the same authors proposed that this peptide, when dissolved in methanol, showed a structure consisting of either a 3_{10}-helix along

*Correspondence to: Francesc J. Corcho, Departament d'Enginyeria Química (UPC), ETS d'Enginyeria Industrial de Barcelona, Diagonal 647, 08028 Barcelona, Spain; e-mail: francesc.corcho@upc.edu

the segment Gln^5-Gln^6-Phe^7-Phe^8 or a α-helix along the segment Pro^4-Gln^5-Gln^6-Phe^7-Phe^8, where the *N*-terminus region segment Arg^1-Pro^2-Lys^3 is a flexible region and the C-terminal segment Gly^9-Leu^{10}-Met^{11} is oriented towards the side chains of Gln^5 and Gln^6. Other studies [20–23] suggest that in lipid membranes and sodium dodecylsulfate, SP adopts an α-helical structure. Interestingly, a NMR study reported the comparison of the structures of a segment of the β-amyloid peptide (25–35) (Aβ25–35) and SP showing a similar α-helical structure in their *C*-terminal regions in a trifluoroethanol (TFE)/water solution [33].

The abovementioned studies highlight the flexibility and adaptability of the structure of this peptide but have in most cases not considered the impact that the post-translational modification has on its structural preferences. In order to test whether the differences in biological activity between the acid and amidated forms of SP are due to a change of structure or rather to a change in the specific interactions established between the active form of the peptide and its receptor we have carried out a detailed study of the structural properties of substance P with its *C*-terminal end amidated SP and itsSPOH using a combination of MD simulations and NMR spectroscopy. The aim is to characterize the structural properties of both peptides and analyze whether differences in structure could give rise to differences in activity.

MATERIALS AND METHODS

NMR Experiments

NMR structural analysis in water and in methanol. NMR Samples: 0.75 mg of substance P in its free acidic form (from SIGMA-ALDRICH S2136, purity >99%) were dissolved in 0.5 ml of H_2O/D_2O (9 : 1) (sample 1) and in 0.7 ml of CD_3OH (from SDS with isotopic enrichment >99.8) (sample 2). The pH of sample 1 was adjusted to 5.05 by additions of acetate buffer; sodium azide (0.02% in weight) was added to ensure conservation of sample 1. One TOCSY (mixing time 70 ms) and a series of NOESY spectra at increasing mixing times (100, 200 and 400 ms) were acquired on a Bruker Avance 600 spectrometer, at a temperature of 298 K. The spectra were processed with NMR pipe, whereas the assignment of the resonances was carried out in NMR VIEW.

MD

All the calculations were carried out within the molecular mechanics framework using the AMBER 5 package [34] and the Cornell force field [35]. The RESP charges of the *C*-terminus amidated methionine were obtained by *Ab initio* calculations performed with the GAUSSIAN94 package [36].

MD of SP and SPOH in water. In order to simulate SP in its native form, the peptide was generated with an amidated methionine (Men) using the PREP module of the AMBER 5 package. The software GAUSSIAN94 was used to compute the molecular electrostatic potential (MEP) at Hartree-Fock level with a 6–31G* basis set and to fit the MEP to point charges located on the nuclei. The RESP module of AMBER was used to adjust the electrostatic potential with the addition of hyperbolic restraints on charges on nonhydrogen atoms in two steps. In the second step, equivalent hydrogen atoms are assigned the same charge [35]. The charges have been published elsewhere [37].

The extended conformations of SP and SPOH were generated using the AMBER program. SP was studied with a positively charged *N*-terminus whereas SPOH was studied with charged *N*- and *C*-termini. Both peptides were minimized with the module SANDER *in vacuo* using a dielectric constant of 80. One thousand five hundred cycles of steepest descent followed by the conjugated gradient method were applied until the RMS derivative between two consecutive structures was smaller than 0.001 $kcal \cdot mol^{-1} \cdot Å^{-1}$. These peptides were used as the starting point for the MD trajectories in explicit water and methanol. Subsequently, SP and SPOH were soaked in a rectangular box containing 1233 and 1150 TIP3P water molecules, respectively. The dimensions of the boxes were chosen to ensure that the minimum distance from the peptides in their extended conformation to the box walls was larger than 8 Å; the energy of the system was then minimized by fixing the atomic coordinates of the peptides and allowing the water molecules to move; a relative dielectric constant $\varepsilon = 1$ and a cutoff of 10 Å were used. A trajectory of 100 ps was simulated using periodic boundary conditions at constant temperature (300 K) and constant pressure (1 atm), with a cutoff of 10 Å. An integration step of 2 fs in conjunction with the use of the SHAKE algorithm to constrain the stretching of bonds involving hydrogen atoms was applied. The translation and rotational motion was removed after each 1000 steps and at the beginning of each 500 ps fragment of MD. Then, the system was changed to Particle Mesh Ewald conditions with a grid spacing of 1 Å and a tolerance for the Ewald sum of 10^{-5} Å snapshot was recorded after each ps and the MD trajectory was carried out for 40 ns.

MD of SPOH in methanol. The methanol OPLS-AA parameters [38] derived from a gas phase optimization carried out with BOSS Version 4.2 [39] were kindly provided by Prof. Bill Jorgensen. The rotational barrier used for methanol was 1.36 $kcal \cdot mol^{-1}$ [38]. A methanol molecule was generated with the AMBER package and was minimized *in vacuo* with the SANDER module in order to obtain a good starting geometry. A relative dielectric constant $\varepsilon = 32.63$, corresponding to the experimental value for methanol, was used [40]. A box containing 65 methanol molecules was generated. In order to obtain a good starting density for the box the maximum distance from the central methanol molecule to the boundaries of the box was estimated to be 8.2 Å; this estimation was carried out by assigning to the density the value obtained from tables at 20 °C, 0.7914 $g \cdot cm^{-3}$ [41]. The equilibration procedure followed for the methanol box is described in detail in the supplementary Protocol 1 online.

The equilibrated box of methanol molecules was used to carry out a MD trajectory with the peptide in its zwitterionic form. The peptide was minimized *in vacuo* with a dielectric constant of 80 as described above. Subsequently, the minimized peptide was soaked in a rectangular box containing 1796 methanol molecules. The system was minimized by fixing the coordinates of the peptide and allowing the methanol

DOI: 10.1002/psc

molecules to move. The dielectric constant, the cutoff, the equilibration protocol, the SHAKE algorithm and the Ewald sum conditions were identical to the ones used for the MD simulation of SPOH in water. A snapshot was recorded after each ps and the trajectories were extended for 40 ns.

Conformation classification. Conformations were classified using the computer program CLASICO developed in this laboratory and designed to group tens of thousands of structures obtained from the MD trajectories into hundreds of patterns based on the presence of secondary structure motifs. Specifically, for a given conformation, the procedure starts by assigning a letter to each residue of the peptide, excluding both termini, according to the values of the backbone dihedral angles of the residue, following a partition of the space proposed by Srinivasan and Rose [42]. This procedure then divides the conformational space in regions characterized by the letter assigned to each residue (Figure 1) and where, as a consequence, each conformation is represented by a string of letters. In a second step, for each conformation, windows containing two or three successive letters are created and motif codes assigned to each residue following the rules shown in Table 1. This new set of conformational motifs assigned to each conformation is named a *pattern*. This procedure allows to identify the different patterns displayed by the peptide, as well as to assess the diversity of the conformational space sampled using different computational methods or solvent conditions. The CLASICO procedure then numbers and plots the secondary structure patterns. The evolution of the sampling procedure and the frequency of the different patterns can be graphically assessed by following the distribution of patterns with time. The CLASICO algorithm can therefore be used to group tens of thousands of structures into hundreds of patterns, easing the subsequent treatment of the information contained in the trajectory (supplementary Figure S1 online).

Helical content and helicity. In order to compare values of the helical content of the peptide and the helicity at each residue obtained from the simulation with those obtained experimentally, a modified version of the CLASICO algorithm that assigned the α- or 3_{10}-helical code at a residue level was used. Data were analyzed from $t = 25$ ns to $t = 35$ ns, when it was considered that the peptide had reached its native state. The helical content and the helicity were computed for this interval of the simulation as the percentage of snapshots containing a helical motif in the peptide or at a residue level, respectively.

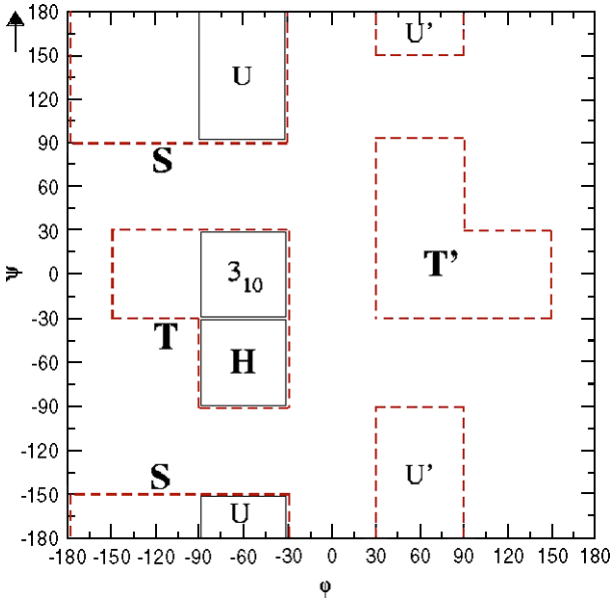

Figure 1 The conformational space is partitioned into several regions based on Srinivasan and Rose [42]. These regions are used to assign the secondary structure motifs following the rules described in Table 1. The regions are named after the secondary structure motiff that each encompasses (3_{10} : 3_{10}-helix; H: α-helix; S: β-strand; T: type I β-turn, T′: type I′ β-turn; U: type II β-turn; U′: type II′ β-turn). This figure is available in colour online at www.interscience.wiley.com/journal/jpepsci.

Clustering of the patterns. The patterns displayed by the peptide during the trajectories are numerous and diverse. In order to address such diversity, we grouped the conformations into clusters according to similarity. Clusters were computed using an algorithm based on information theory, CLUSTERIT. The algorithm considers that a pattern can be categorized unambiguously by the presence (1) or absence (0) of each of the secondary structure motifs. Thus, the pattern can

Table 1 Conditions for secondary structure definition of three consecutive residues

Motif	Condition	Definition	Code
3_{10}–helix	j and j + 1 and j + 2 ∈ 3_{10}	j,j + 1,j + 2 = 3_{10}-helix	310
α-helix	j and j + 1 and j + 2 ∈ H	j,j + 1,j + 2 = α-helix	H
β-strand	j and j + 1 and j + 2 ∈ S	j,j + 1,j + 2 = β-strand	S
Type I β-turn	j + 1 ∈ T and j + 2 ∈ T	j + 1 = type I β-turn (residue i + 1)	I1
	j ∈ T and j + 1 ∈ T	j + 1 = type I β-turn (residue i + 2)	I2
Type I′ β-turn	j + 1 ∈ T′ and j + 2 ∈ T′	j + 1 = type I′ β-turn (residue i + 1)	i1
	j ∈ T′ and j + 1 ∈ T′	j + 1 = type I′ β-turn (residue i + 2)	i2
Type II β-turn	j + 1 ∈ U and j + 2 ∈ T′	*j* + 1 = type II β-turn (residue i + 1)	II1
	j ∈ U and j + 1 ∈ T′	j + 1 = type II β-turn (residue i + 2)	II2
Type II′ β-turn	j + 1 ∈ U′ and j + 2 ∈ T	j + 1 = type II′ β-turn (residue i + 1)	ii1
	j ∈ U′ and j + 1 ∈ T	j + 1 = type II′ β-turn (residue i + 2)	ii2
Coil	none of the above	j + 1 = coil	—

DOI: 10.1002/psc

be translated into a binary code and the number of bits of information contained in the pattern can be computed using Shannon's equation [43] (for a thorough description of the method see Refs. 7,44). Thus, the different patterns are classified into several clusters, eliminating those that do not affect the amount of information of the total system, that is, grouping one pattern into a given cluster does not decrease the amount of information. This procedure results in grouping in the same cluster those conformations that share most of their secondary structure motifs, i.e. that share most of their binary positions or that have a smaller Hamming distance. The classification proceeds until a threshold value was reached. The clusters that were kept at this point were considered to be the minimum number of clusters necessary to describe the conformational diversity of the total system. In order to characterize each cluster, the probability of each motif at each residue within a cluster was computed by adding the number of times that a pattern with such motif at such position had appeared for the structures contained in the cluster. For the discussion of results, values below 10% are generally discarded as they are considered nonrepresentative.

Using the CLASICO algorithm described above, the ensemble of structures obtained from the MD trajectory can be classified into patterns based in the presence of different secondary structure motifs. This reduces the initial set of around 50 000 structures to 200–500 patterns. Finally, the patterns are grouped with the CLUSTERIT program as described above, into 10–30 clusters. This procedure allows for the characterization of the complexity of the group of structures obtained in the different MD trajectories and for the interpretation of the folding process.

Transition analysis. Using the clustering procedure described above, each configuration in the trajectory was assigned a pattern and thus, a cluster. In order to characterize the transitions between clusters which can be observed during the simulations of the peptide, SP transition analysis were carried out. A *transition* was defined when two consecutive structures belong to different clusters. The opposite can be considered as a *self-transition*. If the transitions among clusters occurred randomly it could be expected that the probability of self-transitions for a cluster will equal the percentage of structures contained by that cluster. Therefore, a cluster is defined as *stable* when the probability of suffering a self-transition is significantly higher than the probability of falling in that cluster in a random move. The percentage of transitions between two clusters (the first being the origin and the second being the end) can be computed dividing the number of transitions by the total number of transitions for the origin cluster (*local transitions*). Similarly, the percentage of transitions in regard to the total number of transitions is termed *global transitions* and provides a measure of the statistical relevance of such transition.

Table 2 Chemical shifts (ppm) of SPOH in water

	NH	αH	β_1H	β_2H	γ_1H	γ_2H	δ_1H	δ_2H	εH
R1	7.26	4.42	1.98	1.98	1.75	1.75	3.27	3.27	—
P2	—	4.52	2.05	2.36	2.06	2.06	3.78	3.60	—
K3	8.57	4.57	1.85	1.85	1.54	1.54	1.72	1.72	3.02
P4	—	4.42	2.05	2.33	2.06	2.06	3.78	3.60	—
Q5	8.50	4.22	1.97	1.97	2.32	2.32	—	—	—
Q6	8.32	4.24	1.87	1.87	2.16	2.16	—	—	—
F7	8.24	4.62	3.08	2.95	—	—	—	—	—
F8	8.27	4.62	3.18	2.98	—	—	—	—	—
G9	7.92	3.71, 3.83	—	—	—	—	—	—	—
L10	8.07	4.44	1.66	1.66	1.66	—	0.91	0.96	—
M11	7.96	4.31	1.96	2.12	2.55	2.55	—	—	

Table 3 Chemical shifts (ppm) of SPOH in methanol

	NH	αH	β_1H	β_2H	γ_1H	γ_2H	δ_1H	δ_2H
R1	7.62	4.25	1.92	1.92	1.76	1.76	3.21	3.21
P2	—	4.53	—	—	—	—	3.73	3.60
K3	8.52	4.57	1.96	1.96	1.72	1.72	1.53	1.53
P4	—	4.33	—	—	2.31	2.31	3.80	3.73
Q5	8.74	4.14	2.37	2.37	2.02	2.02	—	—
Q6	8.14	4.23	2.22	2.22	1.99	1.99	—	—
F7	7.92	4.41	2.99	2.99	2.86	2.86	—	—
F8	7.93	4.46	3.26	3.26	3.00	3.00	—	—
G9	7.97	3.69,3.96	—	—	—	—	—	—
L10	7.80	4.46	1.62	1.62	1.62	—	0.92	0.92
M11	8.08	4.40	1.97	2.10	2.53	2.53	—	

DOI: 10.1002/psc

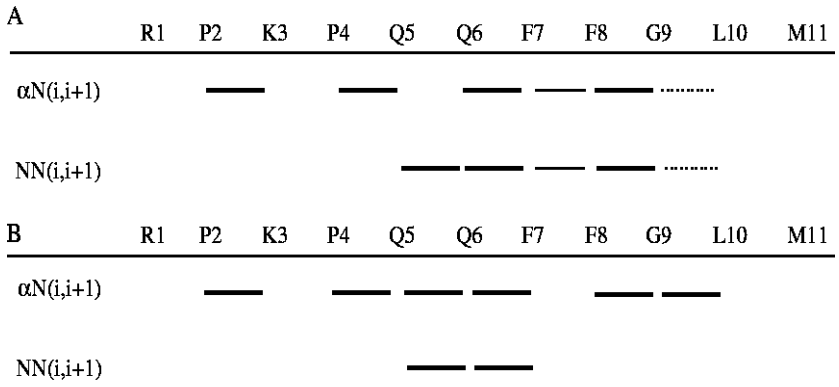

Figure 2 Summary of the observed NOEs for SPOH in (A) water and (B) methanol. Increasing line thickness represents increasing NOE intensity (weak, medium and strong). The mixing time used for these NOESY experiments was 100 ms.

RESULTS

NMR Experiments

In order to assess the structural differences in solution, substance P has been studied in two different solvents, water and methanol. Different conformations arising due to slow isomerization around either the Arg–Pro or the Lys–Pro peptide bonds could be observed in the NMR spectra of SP. Unfortunately, the low intensity and similarity in chemical shifts of these resonances, which could potentially include signals stemming from three different isomers (*cis/cis*, *trans/cis*, *cis/trans*) precluded the accurate determination of their total concentration or rate of exchange with the major, *trans/trans* isomer. The ^{1}H chemical shifts of the major isomer were assigned using the TOCSY and NOESY experiments [45]. The chemical shifts assignments are shown in Tables 2 and 3, and a summary of the observed NOEs is shown in Figure 2.

Chemical shifts are a valuable tool for the conformational analysis of peptides devoid of tertiary structure, where NOE analysis is of limited use due to the lack of persistent long-range contacts. Contrary to NOEs, chemical shifts are affected by a variety of structural features such as backbone torsion angles, hydrogen bonding and mid-range tertiary contacts such as ring current effects; this convolution of effects renders their interpretation challenging and, as a consequence, chemical shifts are seldom used as restraints for structure determination. In spite of these limitations ^{1}H chemical shifts, in particular Hα shifts, have been extensively used for the qualitative assessment of the extent of formation of secondary structure in unstructured peptides and chemically, thermally and natively unfolded proteins, where the absence of long-range hydrophobic contacts simplifies the analysis of the shifts, which can be interpreted as depending on short-range, local interactions.

The differences between the observed Hα chemical shifts and the random coil values described in Ref. 45 are known as Hα conformational shifts (HαCS), where negative values of HαCS are qualitative indicators of fractional helicity; the results of this analysis for SP are shown in Figure 3. The fractional helicity of each residue was calculated by dividing its HαCS by 0.38 ppm, which is the consensus shift of a residue in the fully helical context [46,47]. In the case of Gly^{9} the HαCS was computed as the average value of the chemical shift of the two Hαs. Using this qualitative approach, the average helical content of the peptide was estimated to be 10% in water and 23% in Methanol; the helical content per residue is plotted in Figure 4.

MD

Simulations amounting to 40 ns MD of SP and SPOH at 300 K were carried out starting from their extended

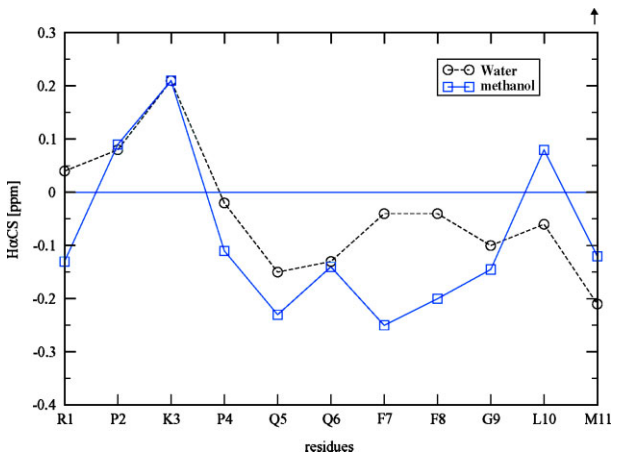

Figure 3 Hα conformational shifts SPOH in water (○) and methanol (□). This figure is available in colour online at www.interscience.wiley.com/journal/jpepsci.

DOI: 10.1002/psc

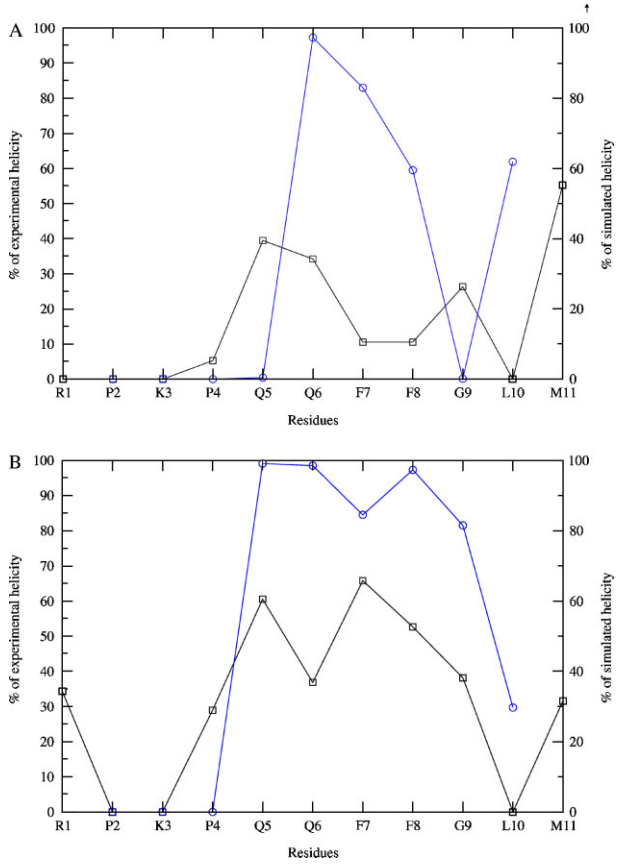

Figure 4 Percentage of computed helicity of SPOH in water (A) and methanol (B) obtained from MD simulations (○) and NMR experiments (□). The helicity for the end residues Arg[1] and Met[11] in the MD trajectories is not computed as they only have a unique dihedral angle each, and a pair of dihedrals is required for assigning a region of the Ramachandran plot. This figure is available in colour online at www.interscience.wiley.com/journal/jpepsci.

conformation soaked in a box of water molecules, using PBC and PME conditions as described in the methods section. The total energy and the density were constant in both the SP and SPOH trajectories at $-9817\ \text{kcal}\cdot\text{mol}^{-1}$ and $-9102\ \text{kcal}\cdot\text{mol}^{-1}$, respectively, for the total energy and $0.962\ \text{g}\cdot\text{cm}^{-3}$ and $0.945\ \text{g}\cdot\text{cm}^{-3}$, respectively for the density. Regarding SPOH in methanol, 40 ns of trajectory at 300 K were carried out with the total energy and the density kept constant at $7830\ \text{kcal}\cdot\text{mol}^{-1}$ and $0.756\ \text{g}\cdot\text{cm}^{-3}$, respectively.

Conformation classification. The CLASICO algorithm was used to analyze the structures of the MD simulations in water and in methanol at 300 K in order to determine the conformational motifs displayed by SP and SPOH that are summarized in supplementary Table S1 online. The average helical content for SP and SPOH in water and methanol were calculated as the sum of all the percentages of structures exhibiting α- or 3_{10}-helical motifs and divided by nine, the number of residues that contain dihedral angle pairs. The helical content obtained was 6.1% for SP in water, 5.7% for SPOH in water and 12.6% for SPOH in methanol. For the MD trajectories of SPOH in water and methanol the helicity of each residue was computed using a modified version of the CLASICO method as described in the methods section and the results compared with the experimental values are shown in Figure 4. The different conformations were labeled according to the conformational motifs exhibited and classified into different patterns. The 40 ns MD simulation of SP and SPOH in water yielded 196 and 231 patterns, respectively. In methanol, SPOH exhibits a substantially more flexible behavior than in water, adopting 264 patterns in 40 ns of MD trajectory. Figure 5 shows the evolution of new patterns sampled along time for the three simulations and in all instances a plateau is reached. This plot can be of help in assessing the thoroughness of the conformational space search.

In the MD study of SP in water, at the beginning of the trajectory the profile grows linearly until $t = 9$ ns, after which there is a significant increase in the number of new patterns. Then, from nanosecond 12 the pace of new pattern appearance decreases and the profile converges (Figure 5(a)). Thus for the first 12 ns, 160 patterns are obtained, whereas for the rest of the 28 ns of the MD trajectory only 36 new patterns are observed, thus indicating that the simulation has thoroughly sampled the conformation space available to the peptide under this set of conditions. The first patterns to appear correspond to β-strand motifs spanning all the molecules that are partially replaced by type I β-turns at residues 5, 6, 9 and 10. This presence of beta secondary structure is simply due to the relaxation from the initial, extended, conformation. The increase in the number of new patterns observed at $t = 9$ ns is due to the appearance of helical motifs and the type I β-turns spanning from residues 5–10. The patterns that are obtained from that moment on are combinations of structures exhibiting a β-strand or no structure for residues 2–4 and for the segment spanning from residues 5 to 10 exhibiting type I β-turns and in some cases α- or 3_{10}-helices. A II2 motif in residue number 3 appears at pattern 185 giving rise to a small increase in the number of patterns, until pattern 194. These patterns are present between $t = 33.7$ and $t = 34.2$. In summary, the simulation shows the propensity of SP to form a helical turn in the center of the molecule, however this turn is not stable and the molecule fluctuates around an ensemble of folded conformations characterized by the presence of two consecutive turns spanning residues 5 to 10.

In both MD trajectories of SPOH in water and methanol there are different rates of new patterns appearance along the process. In the case of water, at the beginning of the trajectory, the profile exhibits a plateau at 5 ns after an initial steady increase of the

DOI: 10.1002/psc

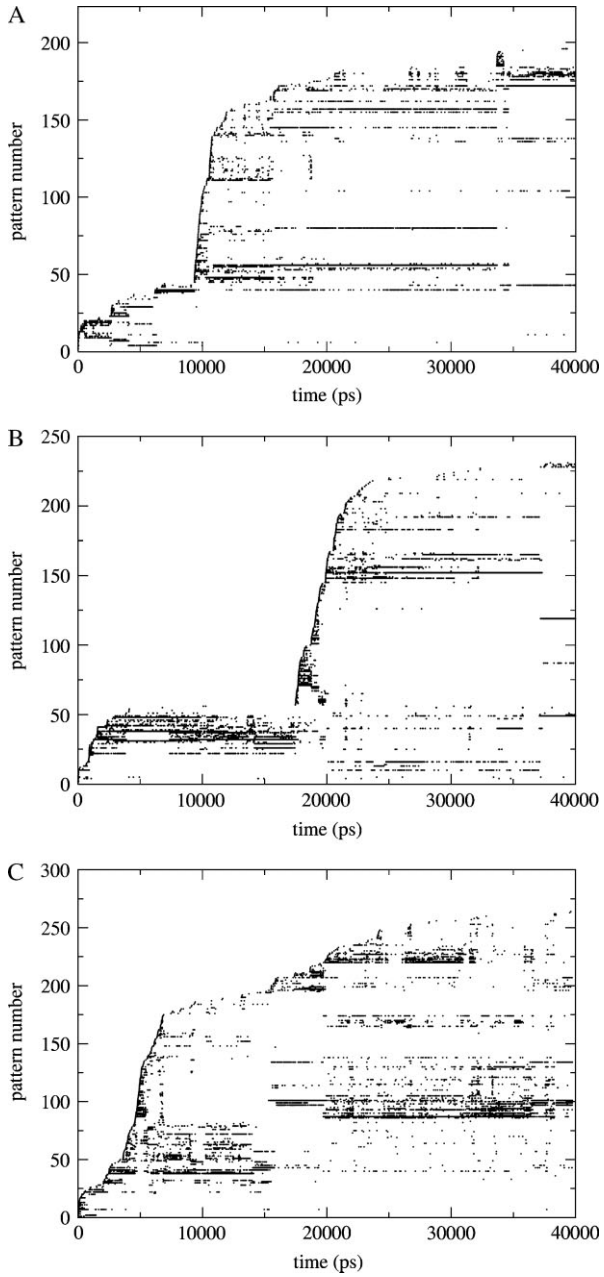

Figure 5 Evolution of patterns for the MD of SP in water (A) and SPOH in water (B) and in methanol (C).

number of patterns. This behavior can be associated with the sampling of new secondary structure from the extended conformation. The number of new patterns is very low until $t = 17$ ns when the peptide samples an II2 motif located on the third residue, generating a wealth of new patterns from this point, until a second plateau is reached after 25 ns. Since all the simulations started from the extended conformation, the first patterns sampled correspond in all cases to structures with no conformational motifs or adopting dihedral angles of a β-strand. Subsequently, SPOH adopts in both cases a type I β-turn between residues 9 and 10 that unfolds, falling into a type I β-turn between residues 6 and 7. This conformation appears to act as a *nucleation site* of the region that extends from residues 5–9 in water and residues 6 to 10 in methanol that folds later to form two consecutive turns that are kept for the main part of both MD trajectories. These central turns adopt occasionally the form of a α- or 3_{10}-helix between residues 5–7 or between residues 6–8. In conclusion, the results obtained for the simulations of SPOH in water and methanol are therefore very similar to those obtained for SP as SPOH also adopts after 2 ns a helical turn in the center of the molecule that fluctuates around an ensemble of folded conformations characterized by the presence of two consecutive turns between residues from 5 to 9.

Comparison of the structural properties of peptides SP and SPOH in different solvent conditions. Regarding the conformational motifs SP and SPOH present under the different solvent conditions that have been summarized in the supplementary Table S1 online, several main differences arouse. In water, SPOH has a greater propensity to form the II2 motif than SP (9.2% *vs* 1.1%). SP in water has a greater propensity to form type I β-turn at residue 5 (81.1%) than SPOH within the same conditions (46.5%). On the contrary, β-strand propensity is larger for SPOH (40.3%) than SP (14.5%) for the same residue. Furthermore, residues 6 to 10 show a greater tendency to attain 3_{10}-helical turns for SPOH and on the opposite SP exhibits a greater proportion of α-helices for the same residues.

On the other hand, when comparing the behavior of SPOH in water and methanol the molecule shares the general structural traits, although several differences can be pointed out. First of all, β-strand dihedral angle values for residues 2–4 show a marked greater propensity in methanol than in water (95% *vs* 69%, respectively). For residues 5, 6 and 7, the molecule in methanol has a clear preference (two times) to form 3_{10}-helical turns than to form α-helical turns, whereas in water the difference is lower (1.5 times greater preference for 3_{10}- than α-helical turns). For residues 8 to 10 the tendency to form helical turns is greater in methanol being the 3_{10}-helical turn the preferred type of helix in water and methanol. Residues 9 and 10 adopt type I β-turn in a greater proportion than in water (more than 4 times).

Clustering of the patterns. The structures of the MD trajectories of SP in water, of SPOH in water and in methanol were classified into 13, 16 and 14 clusters, respectively using the CLUSTERIT algorithm. Supplementary Table S2 online contains a summary of the main structural features of the nine largest clusters of the trajectories studied in the present work. For more detailed information on all the clusters found from the analysis of the MD trajectory of SP in water, SPOH in water and in methanol see the supplementary Tables S3–5 online, respectively.

DOI: 10.1002/psc

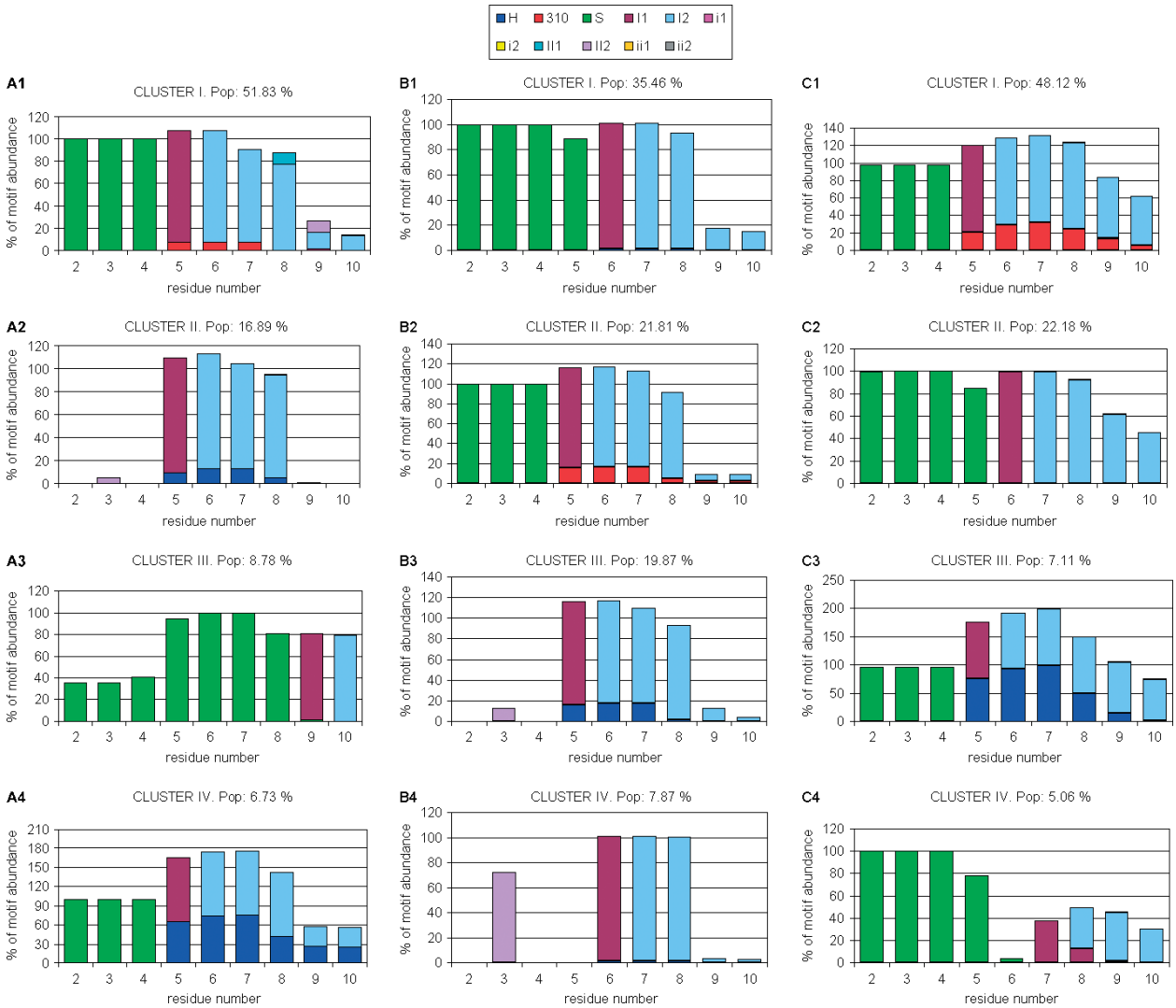

Figure 6 The four most abundant clusters for the MD trajectories of SP in water (A1–A4), of SPOH in water (B1–B4) and of SPOH in methanol (C1–C4).

The four most abundant clusters for each trajectory (Figure 6) represent 84, 85 and 82% of the total number of structures respectively, thus suggesting that the bulk of structures is grouped in a relatively small number of clusters. In all cases the most abundant clusters correspond to the clusters containing two consecutive type I β-turns in the region spanning residues 5 to 9 or 6 to 10 and with residues 2–4 or 2–5 exhibiting a β-strand or an II2 motif on residue 3. For residues 5–7 or 5–8 a α-helical turn or a 3_{10}-helical turn are the most probable structural features observed.

Transition analysis. It is interesting to analyze the visiting sequence of the different clusters along the MD trajectory. This analysis can provide information about the intrinsic characteristics of the folding process. Transitions between clusters for the MD trajectory of SP in water, for SPOH in water and methanol were analyzed as described in the methods section, and the data is presented in (supplementary Figures S2–4 online). The largest nine clusters for SP and SPOH in water and the largest 10 clusters for SPOH in methanol are shown. These clusters account for 98.6, 98 and 98.6% of the structures, respectively. Only local transitions above 5% have been depicted for conciseness, although some particular exceptions were made. For each snapshot in these three trajectories, its corresponding cluster has been assigned. The cluster number of each conformation has been plotted in supplementary Figure S5 online.

The study of transitions between consecutive conformations provides a detailed analysis of the relations of vicinity in the conformational space among the different groups of structures. It also provides the probability with which different groups of conformations can interconvert. This can, in some instances be associated with the existence of free energy barriers amongst different regions of the conformational space. From the study of the most frequent transitions between the different

DOI: 10.1002/psc

clusters in the MD trajectories of SP in water, and the MD trajectories of SPOH in water and methanol it can be inferred that clusters can be grouped into 2 or 3 different classes (supplementary Figures S2–4 and supplementary Table S2 online). The clusters are grouped into a class when the transitions take place mainly amongst them. The existence of these classes of clusters is confirmed by looking at the temporal evolution of the cluster transitions where it is evident that clusters within a class coexist in time in the three MD trajectories (supplementary Figure S5 online).

DISCUSSION

Comparison between MD and NMR Results

Both the MD simulations and the NMR experiments were carried out at 298 K, and from the analysis of the NOESY spectrum obtained in water and methanol (Figure 2), we can infer that SPOH does not present any persistent interaction spanning 3 or 4 residues that could lead to the observation of αN(i,i + 3) or αN(i,i + 4) NOEs, indicative of the presence of stable 3_{10}- or α-helical turns at room temperature. Indeed NN(i,i + 1) NOEs typical of nascent helices or turns can only be detected in the central region of the sequence, from Gln 5 to Gly 9. Sagan and co-workers, studied the conformation of SP and other analogues in methanol and were able to detect medium range NOEs, αN(i,i + 3), αN(i,i + 4) and $\alpha\beta$(i,i + 3), indicative of stable helical conformations at 273 K; these differences in the number and range of the NOEs that are observed are in agreement with the common observation that flexible peptides appear as structured at low temperatures when studied by NMR spectroscopy, especially in aqueous solution. This is due, on the one hand, the stabilization of inter-residue interactions that lead to the compaction of the peptide and are entropically disfavored a higher temperature, and, on the other hand, to the decrease of the tumbling time of the molecule, that enhances the NOE effect. In this work, since the aim was to compare the results obtained in the MD simulations and the NMR experiments, the same temperature, 298 K, was used for both studies. Besides, the use of low and relatively low temperatures, another common approach to the experimental characterization of the secondary structure of flexible peptides is the use of fluorinated co-solvents. These are known to reduce the flexibility of the peptide by inducing a secondary structure by a so far unclarified mechanism which leads to the stabilization of intra-peptides relative to solvent-peptide hydrogen bonds [31]; this approach is based on the assumption that presence of the co-solvent allows the detection of the secondary structure which is only nascent, as we observe, in its absence. The results that we have obtained with SPOH are, thus, also in agreement with those described by Lee and co-workers [33] on SP under a TFE/water mix, which showed a marked tendency of SP to adopt an extended conformation for the *N*-terminus and an α-turn for residues 5 to 9.

A very useful NMR parameter that can be used to detect the fractional formation of helices in situations where NOEs are of low intensity is the Hα chemical shift. By comparing these to those of the same residue in a completely unfolded peptide, thus computing the conformational shift (HαCS), it is possible to determine in a semi-quantitative fashion the helicity at a residue level i.e. negative values of HαCS are indicative of fractional helicity. From the results obtained in the present study for SPOH in water it can be inferred that the four first residues of this peptide are not sampling helical conformations, as expected due to the presence of two proline residues at positions 2 and 4, and that from residue 5 the peptide adopts a structure with a clearly measurable degree of helicity (Figure 4). When the same peptide is dissolved in methanol the same pattern is observed with some local deviations and with a significant increase in the average degree of helicity, which rises from 10 to 24%; in agreement with this result the helical content of the MD trajectories was also higher in methanol (12.6%) than in water (5.7%). Interestingly, the average helical content was 6.1% for SP in aqueous solution, showing no significant difference in respect to the MD trajectory of SPOH in water. The average fractional helicities of SPOH obtained using these two approaches are in disagreement by a factor of approximately two, almost certainly due to systematic errors, to differences in definition of helicity and, possibly, to the presence of two consecutive aromatic side chains in the sequence of SPOH, that may also affect the chemical shift of the α protons irrespective of the conformation of the peptide backbone (Figure 4). The fact that the *increase* in helicity that is observed upon dissolving the peptide in Methanol (+140%) is very well predicted by the MD simulations (+121%) highlights that the ensembles of structures obtained using this method are in agreement with the experiment.

The simulations that have been carried out in this work are, therefore, able to reproduce the structural properties of the system under study and provide further insights into the ensemble of conformations adopted by peptides that in the absence of strong experimental evidence is often concluded to be random. The NMR experiments constitute, thus, a preliminary validation tool for our MD simulations and both techniques are demonstrated to be complementary for the study of this type of problems; work is now under way in our laboratories to further validate this approach by using a more structured system, where the presence of a large number of mid- and long-range NOEs will allow us to carry out a more stringent validation of the approach that we have used.

DOI: 10.1002/psc

Comparison of the Structural Properties of Peptides SP and SPOH in Different Solvent Conditions

From the results obtained in the different MD trajectories it can be concluded that there is a common structural behavior with no marked differences between the MD trajectories of SP in water and for SPOH in water and methanol. Both peptides exhibit a structure characterized by two consecutive type I β-turns in the region spanning residues 5 to 9 or 6 to 10 with residues 2 to 4 or 2 to 5 exhibiting a β-strand or an II2 motif on residue 3. For residues 5 to 7 or 5 to 8, an α-helical turn or a 3_{10}-helical turn are the most probable structural features observed. In particular, SP and SPOH show a balanced equilibrium between 3_{10}- and α-helical forms that is skewed towards the 3_{10}–helical form in the case of SPOH in methanol. The latter is in agreement with the results on the study of α- and 3_{10}-helical transition through the use of potential of mean force reported by Smythe and co-workers [48]. The authors suggest that from thermodynamic data it can be deduced that α-helices are enthalpically stabilized and 3_{10}-helices are entropically favored However, shorter 3_{10}-helices will also be favored enthalpically in nonpolar environments due to the additional intrahelical hydrogen bond that is present in 3_{10}-helices. This seems to be the case for SPOH when methanol is used as solvent. The fact that helices adopted by SPOH are favored in alcohol solvents seems to be in agreement with other studies indicating that helical turns are stabilized in peptides when the solvent is changed from water to methanol. This has been rationalized by Kinoshita *et al.* [49] on the basis of a greater work required for the cavity formation in methanol than in water due to the lower density of the alcohol, and as a consequence solvophobic atoms of a peptide can be exposed to the solvent more than in water and exposure of solvophilic atoms become less important. The solvation-free energy in alcohol becomes less variable against conformational changes. The peptide molecule has a tendency to adopt the conformation with the lowest conformational energy with intramolecular hydrogen bonds such as the helical motifs. From experimental results [50–52] it has been observed that alcohol induces peptides to form helical structures, and the degree of the induction increases with the increase in the size of the hydrocarbon group in alcohol [53].

In order to compare the results of the present study with the NMR structure in TFE/water [33], the pattern corresponding to this structure was calculated and compared with the three more abundant patterns obtained in the MD trajectories of SP in water and SPOH in water and methanol (supplementary Table S6 online). Regarding the structure derived from the NMR study in TFE/water, the conformation proposed agrees well with the structural motifs sampled in the MD simulations in all cases. The main differences consist in the number of consecutive type I β-turns present, running from residues 5 to 9 in the NMR structure and from residues 5 to 8 (for the two most abundant patterns for SP in water) and 5 to 10 (for the third most abundant pattern for SP in water). For the MD of SPOH in water and methanol the same patterns are observed only differing in that residues 5 or 6 can adopt either β-strand or type I β-turn conformations.

Use of CLASICO and CLUSTERIT for the Study of Peptide MD Trajectories

The use of MD trajectories for the characterization of the structural properties of peptides and proteins has been shown useful and it is widely described in the literature. This methodology provides highly detailed information on the structural motifs present in the molecule of study. However, a great amount of data is obtained from this type of studies. Here, we make use of CLASICO and CLUSTERIT as techniques that can help in the reduction and simplification of the data, thus converting the data into useful information. In the case of CLASICO, the different conformations obtained in the MD trajectory are translated into a string of letters corresponding to the different conformational motifs, and each different sequence, called here pattern, allow classifying similar conformations. In the present case around 50 000 snapshots have been grouped into around 500 patterns. In the case of CLUSTERIT, the patterns are grouped into around 10 different clusters based on the similarity amongst them. This procedure allows, then, for a great reduction in the amount of data to study without losing the global information of our system of study. If the clustering technique is effective, conformations that are close in time should belong to the same cluster. We have shown in this study, that stable clusters tend to have a large proportion (above 80%) of autotransitions, i.e. two consecutive snapshots belong to the same cluster, thus confirming the effectiveness of this methodology. Furthermore, we have seen that we can follow what are the predominant clusters at each stage of the MD trajectory and we have defined a cluster class, as the set of clusters that mainly show transitions occurring amongst them. With the use of the cluster classes we can describe the major changes occurring along the MD trajectories.

Relating Structure to Peptide Sequence

The existence of a stereochemical code [54] that could lead from the amino acid sequence to a well defined 3D-conformation has led to a wealth of studies attempting to obtain specific rules for secondary structure prediction from protein sequence ([55] for a list). Peptides have been described as flexible molecules in aqueous solvent based on the difficulty to detect any structural motifs by the use of NMR spectroscopic techniques. Thus, in order to determine the bioactive conformation, i.e. the conformation that

DOI: 10.1002/psc

will bind to the receptor and exert its biological function, authors have used solvents that are known to stabilize structural motifs, as TFE/water mixes. The analysis of the sequence could give further insight into whether the conformation stabilized by the use of less polar solvents is relevant for the biological activity by looking if the structure of the peptide could be *stereochemically codified* in its sequence. Gunasekaran reported a study on a set of 250, largely nonhomologous, high-resolution protein structures from the Brookhaven Data Bank in order to determine the propensities of different amino acids at the different positions flanking and contained within the helices [55]. From the analysis of the dihedral angles of the residue ending the helix, they group helices into two types: left-handed (α_L)-terminated and extended (E)-terminated. Their study reinforces the hypothesis that Gly acts as helix terminator in the case of α_L-terminated helices presenting the motif α_R-α_R-α_R-α_L motif where a succession of three residues in right-handed α-helical conformations is followed by a screw sense reversal at the *C*-terminal end of the helix. When looking at what residues are found at the T + 1 position, T being the helix terminator residue Gly they found that His, Asn, Leu and Phe had the greatest propensity. Given the fact, that SP has the Gly^9-Leu^{10} sequence at the *C*-terminus we have investigated whether the sequence could act as a C-capping motif. The helical capping has also been studied in peptides, and the propensities that are found in peptides are the same as those found in proteins and the effect of Gly as a helix terminator is a common property in both types of molecules [56,57]. Thus, we could infer that the Gly^9-Leu^{10} sequence could act as an ending for the helix. From the study of the probability distribution function of dihedral angles [37] and from the summary of secondary structure motifs (supplementary Table S1 online) we can conclude that the helical motifs, which are subsets of the type I β-turn that apply to three consecutive residues (Figure 1 and Table 1), are substantially disrupted by the presence of Gly^9-Leu^{10} because in all cases there is a clear decrease in the percentage of structures presenting a type I β-turn from Phe^8 to Gly^9 (70.1to 27.9% for SP in water, 87.6 to 15.3% for SPOH in water and 83.3 to 50.1% for SPOH in methanol). Serrano and co-workers [58] proposed that Gly in the terminator position (T) and with a left-handed conformation is largely exposed to solvent. Thus, the introduction of a less polar solvent, like methanol as has been argued above, favors the intramolecular interactions, as in helices, and unfavors the conformation that exposes the polar groups of the backbone to the solvent, left-handed conformation. This would explain the smaller decrease in type I β-turn motif for SP in methanol and also the differences observed in the analysis of the study of the probability distribution function of dihedral angles for the ϕ_9 dihedral angle that presents as the predominant peak the one at -70° instead of the one at 80° as it is the case for SP in water. This environment-dependent conformation of Gly^9 has been observed in other peptides like alamethicin. Indeed, Gibbs *et al.* carried out a MD trajectory of alamethicin in methanol [59]. Whereas the peptide adopts a regular α-helix in the crystal structure, in methanol, alamethicin establishes 3_{10}-helix hydrogen bonds at a motif containing Gly-Leu and this confers some flexibility and allow for the bending of the peptide around this position. This would be paralleled by SP and SPOH Gly^9-Leu^{10} motif and suggest the existence of a hinge region at the *C*-terminus of these peptides that could play a critical role in the induced fit of the peptide to its receptor providing flexibility to their structure.

The α_L terminated helix, also known as Schellman motif [60] is further characterized by the existence of two hydrogen bonds: a $6 \rightarrow 1$ hydrogen bond between the N–H group of residue T + 1 (Leu^{10} in SP) and the C = O group of residue T − 4 (Gln^5) and a $4 \rightarrow 1$ hydrogen bond between the N–H group of residue T (Gly^9)and the C = O group of residue T-3 (Gln^6) [61]. We have also calculated the probability of having these two additional interactions in the MD trajectory of SP in water. We have found that both interactions are highly abundant (44.5% and 41.4%, respectively). Hydrogen bonds were considered when fulfilling the criteria: O$\cdots$H–N angle lower than 60° and less than 4 Å for the distance between the O and N atoms of the carbonyl and the amide groups of the backbone. The existence of these interactions between backbone atoms of Leu^{10} and Gln^5 and of Gly^9 and Gln^6 directs the *C*-terminus of SP, residues Gly^9-Leu^{10}-Met^{11}, to residues Gln^5-Gln^6. This fact has been suggested by Chassaing *et al.* [14] for SP in methanol.

Regarding the *N*-terminus of SP, the existence of two Pro, a highly conformationally restricted peptide, is remarkable. MacArthur and Thornton [62] described the Pro-X-Pro sequence as a sequence favoring an extended conformation. The first Pro and the X residue, where X can be any of the natural amino acids except for Pro, are predominantly adopting a β-strand conformation, whereas the second Pro populates indistinctly the α and β conformations. They also found this sequence never occurring in the centre of α-helices but at the beginning or just before the turn of the helix. In addition, other authors have postulated that Pro residues located in the first turn of helices act as N-capping residues thus stabilizing helices [63]. In the case of SP and SPOH, Pro^2–Lys^3 adopts an extended or β-strand conformation as the predominant motif and Pro^4 can adopt both type I β-turn and β-strand dihedral angles. This is in total agreement with the results obtained by MacArthur and Thornton, thus further suggesting that the peptide has its predominant conformation *stereochemically* determined at both the

DOI: 10.1002/psc

extended conformation at the *N*-terminus and the ends of the helical motif at the *C*-terminus.

CONCLUSIONS

The present study is a combined approach using NMR and MD studies on the amidated (SP) and the free acid form (SPOH) of substance P. The aim is to characterize, structurally, both peptides under different solvent conditions. The ^{1}H-NMR experiments were carried out on SPOH in water and methanol. From these experiments, it is suggested that the peptide does not adopt a rigid conformation. The MD studies suggest that both SP in water and SPOH in water and methanol, adopt an ensemble of conformations characterized by the presence of an extended *N*-terminus affecting the first 3 residues, and from that point onwards the peptide would adopt a flexible helix that would interconvert between α- and 3_{10}-helical conformations. These results are in agreement with the results that had been previously obtained for SP in a TFE/water mix and for SP in methanol. Some authors have suggested that helical peptides are plastic structures and could be in rapid interconversion of a wide range of helix-like conformers as a function of their molecular environment [64]. Results obtained in the present study agree with this picture for SP and SPOH.

Furthermore, the overall structure of both SP and SPOH being highly similar, we must conclude that the differences in the biological activity between SP and its free acid form are not based on both peptides exhibiting different structural features, but rather, on the specific interactions established between the active form of SP and its receptor, that must be absent or less strong in the case of SPOH. Being the only difference between both peptides the presence of a negatively charged carboxylate group in SP and a neutral amide group in SPOH, the interaction of this functional moiety with the receptor is key to explain the differences in the biological activity that both peptides present.

On the other hand, Gly^9 would induce a left-handed helix termination at the *C*-terminus. The end of the helix at the *C*-terminus would be folded in a Schellman motif pointing the *C*-terminus to the central part of the peptide (Gln^5-Gln^6). The helix would fry from that point onwards (Gly^9), being methanol the solvent where SPOH exhibits a greater propensity for helix continuation at Gly^9. Gly residue terminating a left-handed helix is largely exposed to solvent. The introduction of a less polar solvent, like methanol would favor intramolecular interactions, as in helices, and unfavors the conformation that exposes the polar groups of the backbone to the solvent, i.e. left-handed conformation. In addition, we suggest that the Gly^9-Leu^{10} sequence in SP could act in the same manner as in alamethicin introducing flexibility at the *C*-terminus of the peptide. This could play a critical role in the induced fit of the peptide to its receptor allowing for the unfolding of the Schellman motif directing the *C*-terminus to the receptor.

Regarding the *N*-terminus of SP, the existence of the Pro-X-Pro sequence favors an extended conformation. Pro^2 and Lys^3 residues are predominantly adopting a β-strand conformation, whereas Pro^4 can adopt α-helical and β-strand conformations. This is in total agreement with what had been described for the Pro-X-Pro motif. Furthermore, this motif had been proposed to be at the beginning or just before the turn of the helix and given the known function of Pro as N-capping residue this suggest that the presence of this sequence has a clear structural function. Thus, Pro^2, Pro^4, Gly^9 and Leu^{10} would play a critical role in conferring the structural properties of SP. This is in agreement of the existence of a stereochemical code formulated by Anfinsen [54]. Thus, the absence of a unique conformation in water of the peptide cannot be thought of it as not having structure. On the contrary, the peptide would adopt a restricted group of conformations presenting flexible motifs that would interconvert easily.

We have used, in the present work, a methodology developed *in house* for classification (CLASICO) and clustering of structures based on information theory (CLUSTERIT) that efficiently group structures upon their structural similarity. This method has proven to be a useful tool in the analysis of the evolution of a peptide through an extended MD. The use of this tool has also allowed for the study of the transitions that take place along the MD trajectories. The peptides selected for the present study, SP and SPOH, have been used as a proof of concept for the methodology here presented, being peptides that exhibit secondary structure with a certain degree of flexibility. However, in order to validate the use of this methodology in peptides and proteins presenting a higher degree of structure, similar studies are being carried out in the group with peptides in the size range of 30–40 residues. The first results suggest that the same methodology could be used in peptides or proteins without any size restriction and could be easily automated.

Supplementary Material

Supplementary electronic material for this paper is available in Wiley InterScience at: http://www.interscience.wiley.com/jpages/1075-2617/suppmat/

Acknowledgements

FJC and XS wish to express his gratitude to the Generalitat de Catalunya for a fellowship to undertake doctoral studies. We are grateful to CEPBA for a generous allocation of computer time. This research was supported by the Spanish Ministry of Science and Technology through grant SAF2002-04325-C03-01.

REFERENCES

1. Kubelka J, Hofrichter J, Eaton WA. The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* 2004; **14**(1): 76–88.
2. Gnanakaran S, Nymeyer H, Portman J, Sanbonmatsu KY, García AE. Peptide folding simulations. *Curr. Opin. Struct. Biol.* 2003; **13**(2): 168–174.
3. Shea JE, Brooks CL. From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. *Annu. Rev. Phys. Chem.* 2001; **52**: 499–535.
4. Daura X, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Reversible peptide folding in solution by molecular dynamics simulation. *J. Mol. Biol.* 1998; **280**(5): 925–932.
5. Daura X, van Gunsteren WF, Mark AE. Folding-unfolding thermodynamics of a beta-heptapeptide from equilibrium simulations. *Proteins* 1999; **34**(3): 269–280.
6. Daura X, Gademann K, Schäfer H, Jaun B, Seebach D, van Gunsteren W. The beta-peptide hairpin in solution: conformational study of a beta-hexapeptide in methanol by NMR spectroscopy and MD simulation. *J. Am. Chem. Soc.* 2001; **123**(10): 2393–2404.
7. Simmerling C, Strockbine B, Roitberg AE. All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.* 2002; **124**(38): 11258–11259.
8. Snow CD, Nguyen H, Pande VS, Gruebele M. Absolute comparison of simulated and experimental protein-folding dynamics. *Nature* 2002; **420**(6911): 102–106.
9. Hummer G, García A, Garde S. Helix nucleation kinetics from molecular simulations in explicit solvent. *Proteins* 2001; **42**(1): 77–84.
10. Harmar A, Schofield JG, Keen P. Substance P biosynthesis in dorsal root ganglia – an immunochemical study of [35S]methionine and [3H]proline incorporation in vitro. *Neuroscience* 1981; **6**(10): 1917–1922.
11. Bolkenius FN, Ganzhorn AJ. Peptidylglycine alpha-amidating mono-oxygenase: neuropeptide amidation as a target for drug design. *Gen. Pharmacol.* 1998; **31**(5): 655–659.
12. Fournier A, Couture R, Regoli D, Gendreau M, St-Pierre S. Synthesis of peptides by the solid-phase method. 7. Substance P and analogues. *J. Med. Chem.* 1982; **25**(1): 64–68.
13. Murakoshi T, Yanagisawa M, Kitada C, Fujino M, Otsuka M. The role of the N-terminus in the active conformation of the substance P analogues. *Eur. J. Biochem.* 1983; **90**(1): 133–137.
14. Chassaing G, Convert O, Lavielle S. Preferential conformation of substance P in solution. *Eur. J. Biochem.* 1986; **154**(1): 77–85.
15. Sumner SC, Gallagher KS, Davis DG, Covell DG, Jernigan RL, Ferretti JA. Conformational analysis of the tachykinins in solution: substance P and physalaemin. *J. Biomol. Struct. Dyn.* 1990; **8**(3): 687–707.
16. Patel AB, Srivastava S, Phadke RS. Substance P (free acid) adopts different conformation than native peptide in DMSO, water and DPPC bilayers. *J. Biomol. Struct. Dyn.* 2001; **19**(1): 129–138.
17. Sagan S, Lequin O, Frank F, Convert O, Ayoub M, Lavielle S, Chassaing G. C-alpha methylation in molecular recognition – Application to substance P and the two neurokinin-1 receptor binding sites. *Eur. J. Biochem.* 2001; **268**(10): 2997–3005.
18. Mehlis B, Bohm S, Becker M, Bienert M. Circular dichroism and infrared studies of substance P and C-terminal analogs. *Biochem. Biophys. Res. Commun.* 1975; **66**(4): 1447–1453.
19. Mehlis B, Rueger M, Becker M, Bienert M, Niedrich H, Oehme P. Circular dichroism studies of substance P and its C-terminal sequences. CD spectra in aqueous solution and effects of hydrogen ion concentration. *Int. J. Pept. Protein Res.* 1980; **15**(1): 20–28.
20. Erne D, Rolka K, Schwyzer R. Membrane-structure of substance-P. 3. Secondary structure of substance-P in 2,2,2-trifluoroethanol, methanol, and on flat lipid-membranes studied by infrared-spectroscopy. *Helv. Chim. Acta* 1986; **69**(8): 1807–1816.
21. Schwyzer R, Erne D, Rolka K. Membrane-structure of substance-P. 1. Prediction of preferred conformation, orientation, and accumulation of substance-P on lipid-membranes. *Helv. Chim. Acta* 1986; **69**(8): 1789–1797.
22. Rolka K, Erne D, Schwyzer R. Membrane-structure of substance-P. 2. Secondary structure of substance-P, [9-Leucine]substance-P, and shorter segments in 2,2,2-trifluoroethanol, methanol, and on liposomes studied by circular-dichroism. *Helv. Chim. Acta* 1986; **69**(8): 1798–1806.
23. Williams RW, Weaver JL. Secondary structure of substance P bound to liposomes in organic solvents and in solution from Raman and CD spectroscopy. *J. Biol. Chem.* 1990; **265**(5): 2505–2513.
24. Young JK, Anklin C, Hicks RP. NMR and molecular modeling investigations of the neuropeptide substance P in the presence of 15 mM sodium dodecyl sulfate micelles. *Biopolymers* 1994; **34**(11): 1449–1462.
25. Keire DA, Fletcher TG. The conformation of substance P in lipid environments. *Biophys. J.* 1996; **70**(4): 1716–1727.
26. Manavalan P, Momany FA. Conformational energy calculations on substance P. *Int. J. Pept. Protein Res.* 1982; **20**(4): 351–365.
27. Nikiforovich GV, Balodis YY, Chipens GI. Conformation-function relationships for substance-P. *Bioorg. Khim.* 1981; **7**(5): 645–654.
28. Wymore T, Wong TC. Molecular dynamics study of substance P peptides in a biphasic membrane mimic. *Biophys. J.* 1999; **76**(3): 1199–1212.
29. Wymore T, Wong TC. Molecular dynamics study of substance P peptides partitioned in a sodium dodecylsulfate micelle. *Biophys. J.* 1999; **76**(3): 1213–1227.
30. Coutinho E, Kamath S, Saran A, Srivastava S. NMR and molecular dynamics studies of tachykinins: conformation of substance P fragment 4–11. *J. Biomol. Struct. Dyn.* 1998; **16**(3): 747–755.
31. Bodkin MJ, Goodfellow JM. Competing interactions contributing to alpha-helical stability in aqueous solution. *Protein Sci.* 1995; **4**(4): 603–612.
32. Brooks CL, Nilsson L. Promotion of helix formation in peptides dissolved in alcohol and water-alcohol mixtures. *J. Am. Chem. Soc.* 1993; **115**(23): 11034–11035.
33. Lee S, Suh YH, Kim S, Kim Y. Comparison of the structures of beta amyloid peptide (25–35) and substance P in trifluoroethanol/water solution. *J. Biomol. Struct. Dyn.* 1999; **17**(2): 381–391.
34. Case D, Pearlman D, Cadwell J, Cheatham T, Ross W, Simmerling C, Darden T, Merz K, Stanton R, Cheng A, Vincent J, Crowley M, Ferguson D, Radmer R, Seibel G, Singh U, Weiner P, Kollman P. *AMBER 5*. University of California: San Francisco, 1997.
35. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 1995; **117**: 5179–5197.
36. Frisch M, Trucks G, Schlegel H, Gill P, Johnson B, Robb M, Cheeseman J, Kieth T, Petersson G, Montgomery J, Raghavachari K, Al-Laham M, Zakrzewski V, Ortiz J, Foresman J, Cioslowski J, Stefanov B, Nanayakkara A, Challacombe M, Peng C, Ayala P, Chen W, Wong M, Andres J, Replogle E, Gomperts R, Martin R, Fox D, Binkley J, Defrees D, Baker J, Stewart J, Head-Gordon M, Gonzalez C, Pople J. *GAUSSIAN94*. Gaussian: Pittsburgh, PA, 1995.
37. Corcho FJ. *Computational studies on the structure and dynamics of bioactive peptides*. PhD Thesis, Universitat Politecnica de Catalunya, 2004.
38. Jorgensen WL, Maxwell DS, Rives T. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 1996; **118**(45): 11225–11236.
39. Jorgensen WL. *BOSS, version 4.2*. Yale University: New Haven, CT, 2000.
40. Lide DR. *In Handbook of Chemistry and Physics*, 76th edn. CRC Press: Boca Raton, 1995; 8–64.
41. Lide DR. *In Handbook of Chemistry and Physics*, 76th edn. CRC Press: Boca Raton, 1995; 14–15.
42. Srinivasan R, Rose G. A physical basis for protein secondary structure. *Proc. Natl. Acad. Sci. U.S.A.* 1999; **96**(25): 14258–14263.

43. Shannon C. A mathematical theory of communication. *Bell Syst. Tech. J.* 1948; **27**: 379–423 and 623–656.
44. Corcho FJ, Canto J, Perez JJ. Comparative analysis of the conformational profile of substance P using simulated annealing and molecular dynamics. *J. Comput. Chem.* 2004; **25**(16): 1937–1952.
45. Wüthrich K. *NMR of Proteins and Nucleic Acids*. John Wiley & Sons: New York, 1986.
46. Rizo J, Blanco FJ, Kobe B, Bruch MD, Gierasch LM. Conformational behavior of Escherichia coli OmpA signal peptides in membrane mimetic environments. *Biochemistry* 1993; **32**(18): 4881–4894.
47. Wishart DS, Sykes BD. Chemical shifts as a tool for structure determination. *Meth. Enzymol.* 1994; **239**: 363–392.
48. Smythe ML, Huston SE, Marshall GR. The molten helix – effects of solvation on the alpha-helical to 3(10)-helical transition. *J. Am. Chem. Soc.* 1995; **117**(20): 5445–5452.
49. Kinoshita M, Okamoto Y, Hirata F. Peptide conformations in alcohol and water: analyses by the reference interaction site model theory. *J. Am. Chem. Soc.* 2000; **122**(12): 2773–2779.
50. Hamada D, Goto Y. The equilibrium intermediate of beta-lactoglobulin with non-native alpha-helical structure. *J. Mol. Biol.* 1997; **269**(4): 479–487.
51. Hamada D, Kuroda Y, Tanaka T, Goto Y. High helical propensity of the peptide fragments derived from beta-lactoglobulin, a predominantly beta-sheet protein. *J. Mol. Biol.* 1995; **254**(4): 737–746.
52. Hirota N, Mizuno K, Goto Y. Group additive contributions to the alcohol-induced alpha-helix formation of melittin: implication for the mechanism of the alcohol effects on proteins. *J. Mol. Biol.* 1998; **275**(2): 365–378.
53. Ramírez-Alvarado M, Blanco FJ, Serrano L. De novo design and structural analysis of a model beta-hairpin peptide system. *Nat. Struct. Biol.* 1996; **3**(7): 604–612.
54. Anfinsen CB. Principles that govern the folding of protein chains. *Science* 1973; **181**(96): 223–230.
55. Gunasekaran K, Nagarajaram HA, Ramakrishnan C, Balaram P. Stereochemical punctuation marks in protein structures: glycine and proline containing helix stop signals. *J. Mol. Biol.* 1998; **275**(5): 917–932.
56. Chakrabartty A, Doig AJ, Baldwin RL. Helix capping propensities in peptides parallel those in proteins. *Proc. Natl. Acad. Sci. U.S.A.* 1993; **90**(23): 11332–11336.
57. Thomas ST, Loladze VV, Makhatadze GI. Hydration of the peptide backbone largely defines the thermodynamic propensity scale of residues at the C′ position of the C-capping box of alpha-helices. *Proc. Natl. Acad. Sci. U.S.A.* 2001; **98**(19): 10670–10675.
58. Serrano L, Sancho J, Hirshberg M, Fersht AR. Alpha-helix stability in proteins. I. empirical correlations concerning substitution of side-chains at the N and C-caps and the replacement of alanine by glycine or serine at solvent-exposed surfaces. *J. Mol. Biol.* 1992; **227**(2): 544–559.
59. Gibbs N, Sessions RB, Williams PB, Dempsey CE. Helix bending in alamethicin: molecular dynamics simulations and amide hydrogen exchange in methanol. *Biophys. J.* 1997; **72**(6): 2490–2495.
60. Schellman C. In *Protein Folding*, Jaenicke R. (ed.). Elsevier/North-Holland: New York, 1980; 53–61.
61. Aurora R, Srinivasan R, Rose GD. Rules for alpha-helix termination by glycine. *Science* 1994; **264**(5162): 1126–1130.
62. MacArthur MW, Thornton JM. Influence of proline residues on protein conformation. *J. Mol. Biol.* 1991; **218**(2): 397–412.
63. Richardson JS, Richardson DC. Amino acid preferences for specific locations at the ends of alpha helices. *Science* 1988; **240**(4859): 1648–1652.
64. Shimizu S, Shimizu K. Alcohol denaturation: thermodynamic theory of peptide unit solvation. *J. Am. Chem. Soc.* 1999; **121**(11): 2387–2394.

DOI: 10.1002/psc