Structural analysis of the starfish SALMFamide neuropeptides S1 and S2: the N-terminal region of S2 facilitates self-association

Claire B. Otara[a*], Christopher E. Jones[a, b*], Nadine D. Younan[a], John H. Viles[a] and Maurice R. Elphick[a]

[a.] The School of Biological & Chemical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, UK. [b.] The School of Science and Health, The University of Western Sydney, Locked bag 1797, Penrith, 2751, Sydney, Australia.

*CBO and CEJ contributed equally

Running title: *Self-association of the SALMFamide neuropeptide S2*

To whom correspondence should be addressed:

Maurice R. Elphick, School of Biological & Chemical Sciences, Queen Mary, University of London, Mile End Road, London, E1 4NS, UK. Tel: 0207 882 5290; Fax: 0207 882 7732; E-mail: m.r.elphick@qmul.ac.uk

or

John H. Viles, School of Biological & Chemical Sciences, Queen Mary, University of London, Mile End Road, London, E1 4NS, UK. Tel: 0207 882 8443; Fax: 0207 882 7732; E-mail: j.viles@qmul.ac.uk

ABSTRACT

The neuropeptides S1 (GFNSALMFamide) and S2 (SGPYSFNSGLTFamide), which share sequence similarity, were discovered in the starfish *Asterias rubens* and are prototypical members of the SALMFamide family of neuropeptides in echinoderms. SALMFamide neuropeptides act as muscle relaxants and both S1 and S2 cause relaxation of cardiac stomach and tube foot preparations *in vitro* but S2 is an order of magnitude more potent than S1. Here we investigated a structural basis for this difference in potency using spectroscopic techniques. Circular dichroism spectroscopy showed that S1 does not have a defined structure in aqueous solution and this was supported by 2D nuclear magnetic resonance experiments. In contrast, we found that S2 has a well-defined conformation in aqueous solution. However, the conformation of S2 was concentration dependent, with increasing concentration inducing a transition from an unstructured to a structured conformation. Interestingly, this property of S2 was not observed in an N-terminally truncated analog of S2 (short S2 or SS2; SFNSGLTFamide). Collectively, the data obtained indicate that the N-terminal region of S2 facilitates peptide self-association at high concentrations, which may have relevance to the biosynthesis and/or bioactivity of S2 *in vivo*.

Keywords: neuropeptide, SALMFamide, circular dichroism, NMR, muscle relaxant, self-association

1.0 Introduction

The SALMFamides are a family of neuropeptides that act as muscle relaxants in species belonging to the phylum Echinodermata (e.g. starfish, sea urchins, sea cucumbers). Two types of SALMFamides have been identified in echinoderms: L-type SALMFamides, which have the C-terminal motif SxLxFamide, and F-type SALMFamides, which have the C-terminal motif SxFxFamide [1, 2]. The first members of the SALMFamide neuropeptide family to be identified were two L-type SALMFamides, S1 and S2, which were both isolated from the nervous system of the starfish species *Asterias rubens* and *Asterias forbesi* [3]. S1 is an amidated octapeptide (Gly-Phe-Asn-Ser-Ala-Leu-Met-Phe-NH_2) and S2 is an amidated dodecapeptide (Ser-Gly-Pro-Tyr-Ser-Phe-Asn-Ser-Gly-Leu-Thr-Phe-NH_2), which share the C-terminal sequence Phe-Asn-Ser-X-Leu-X-Phe-NH_2 (where X is variable).

Investigation of the physiological roles of S1 and S2 has revealed that injection of these peptides into starfish (*Asterias rubens*) causes eversion of the cardiac stomach. Consistent with these *in vivo* effects, analysis of the effects of S1 and S2 *in vitro* revealed that both peptides cause relaxation of the cardiac stomach [4]. These findings are of interest from a physiological and behavioural perspective because starfish feed by everting their cardiac stomach over the digestible parts of prey such as mussels and oysters. Therefore, S1 and/or S2 may mediate neural control of stomach eversion in starfish *in vivo* by causing relaxation of muscle associated with the cardiac stomach. Consistent with this notion, immunocytochemical studies have revealed that S1 and S2 are present in nerve fibres in close proximity to the muscle layer of the cardiac stomach [5]. Furthermore, S1 and S2 appear to be expressed in different populations of

neurons in *Asterias rubens* [5, 6] and consistent with this observation recent analysis of the genome sequence of the starfish *Patiria miniata* has revealed that S1 and an S2-like peptide are encoded by different genes [7].

Comparison of the effectiveness of S1 and S2 as muscle relaxants in starfish has revealed that S2 is more effective than S1 when tested at the same concentration (10 μM) on three different preparations from *Asterias rubens*: the cardiac stomach, tube feet and the body wall apical muscle [8]. Furthermore, extrapolated dose-response data indicate that S2 may be approximately ten times more potent than S1 [1]. However, a direct comparison of the potency of S1 and S2 as muscle relaxants in starfish has as yet not been conducted. Therefore, here we have compared the relaxing effects of both S1 and S2 at a range of concentrations on two different preparations, the cardiac stomach and tube feet.

It is not known if S1 and S2 exert their relaxing effects on starfish muscle preparations by binding to a single receptor type or via different receptors, although it is assumed that these peptides act on G protein-coupled receptors [1]. Nevertheless, given the similarity of the C-terminal regions of S1 (FNSALMFamide) and S2 (FNSGLTFamide), we hypothesized that the presence of four additional N-terminal residues in S2 (SGPY) may influence its structure in solution and account, at least in part, for differences in the potency of S1 and S2. To address this issue, here we have used circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy to analyse and compare the conformations of S1 and S2 in aqueous solution.

Previous studies on other neuropeptides comparable in size to S1 and S2 have revealed that they exist in aqueous solution as an ensemble of rapidly interconverting structures lacking stable secondary or tertiary structural elements [9]. For example, the neuropeptide neurotensin, which comprises thirteen amino acid residues, exists mainly as a flexible “random coil”

polypeptide chain in water [10, 11]. Similarly, the mammalian tachykinin peptides substance P, neurokinin A and neurokinin B and the non-mammalian tachykinins physalaemin, eledoisin and kassinin all adopt a variety of interconverting conformational states in water [12-16]. Therefore, our expectation was that both S1 and S2 would similarly adopt multiple conformational states in aqueous solution. Surprisingly, however, we have discovered that S2, but not S1, has a remarkably well-defined conformation in aqueous solution, and the structuring is caused by self-association. The structural basis and functional relevance of this property of S2 is investigated and discussed.

2.0 Materials and Methods

2.1 Materials - Peptides were custom-synthesized by the Advanced Biotechnology Centre at Imperial College London and purified using high performance liquid chromatography. All other chemicals used were obtained from VWR (Poole, Dorset, UK) with the exception of D_2O, which was obtained from Goss Scientific Instruments Ltd (Great Baddow, Essex, UK).

2.2 In vitro *pharmacology* - Starfish (*Asterias rubens*) were obtained from the Menai Straits (UK) and then maintained in a circulating seawater aquarium in the School of Biological & Chemical Sciences at QMUL. Cardiac stomach preparations were dissected from animals ranging in diameter from 12 – 16 cm and then linked to an isotonic transducer (model 60-3001; Harvard, South Natick, MA, USA) in a 20 ml organ bath containing aerated seawater at 11°C, as described previously [1, 17]. Changes in cardiac stomach length were recorded on a chart recorder (Goerz Servogor 124). Sustained contracture of the cardiac stomach was induced and maintained using seawater with 30 mM added KCl, as described previously [4]. Solutions of S1 and S2 were then prepared in water and added (20 µl) to the organ bath to achieve final bath concentrations of 0.001, 0.01, 0.1 and 1 µM. S1 and S2 were tested on a total of fifteen cardiac stomach preparations, with both peptides being applied twice at each concentration tested. The order in which different peptides and concentrations were tested was random after having first tested 1 µM S2. Preparations were washed between tests by refilling the organ bath several times with seawater/KCl and then allowing a 15-20 min interval before the next test, during which time preparations regained basal resting length. The effects of each peptide concentration tested were

normalised to the effect of 1 μM S2 and mean % relaxation values were calculated for both S1 and S2 at each of the concentrations tested.

The relaxing effects of S1 and S2 on tube foot preparations were also compared at a range of concentrations, recording under isotonic conditions as described previously [8]. At concentrations below 0.1 μM, the magnitudes of the relaxing effects observed were too small to be consistently measurable and therefore comparison of the effects of S1 and S2 was limited to tests at 0.1, 1 and 10 μM.

An N-terminally truncated analog of S2 (short S2 or SS2; SFNSGLTFamide) was also tested on both cardiac stomach and tube foot preparations to enable comparison of its bioactivity with that of S1 and S2.

2.3 Circular Dichroism (CD) spectroscopy - CD spectra were recorded using a Chirascan CD spectrometer (Applied Photophysics Ltd.) equipped with a Peltier temperature controller. Spectra were the average of three scans recorded with a 1 nm bandwidth, a 0.5 nm step size, and a 5 s time constant. After background subtraction, the observed ellipticity (θ; mdeg) was converted to a molar ellipticity (ΔE; M^{-1} cm^{-1}), using the formula: $\Delta E = \theta/(33000\ lc)$ where l is the path length (cm) and c is the concentration (M). To investigate the effect of pH on CD spectra, the pH of the peptide solution was altered using 10 mM NaOH or 10 mM HCl to achieve four pH values in the range 5.3 – 8.4. Samples were then incubated for 20 min prior to recording CD spectra. To examine the effect of temperature, CD spectra were obtained at 10, 20, 30, 40 and 50°C. Thermal unfolding of S2 was also investigated up to 90°C using a 1 mm path length for a 0.1 mg/ml sample and a 1 cm path length for a 0.01 mg/ml sample.

2.4 NMR spectroscopy - Peptides were dissolved in 10% D_2O/90% H_2O to achieve a final concentration of 2 mM. The pH of the solutions was adjusted to pH 5.6 using 10 mM NaOH and 10 mM HCl. The peptide solutions were centrifuged to remove any suspended material and then the supernatant was transferred to 5 mm NMR tubes. Data were acquired using a Bruker Avance 600 MHz spectrometer and collected using Topspin software on a UNIX workstation. All experiments were performed using a 5 mm TXI, triple resonance probe equipped with a z-axis gradient. 2D NMR spectra were obtained at 293 K for S1, 303 K for S2 and 283 K for SS2. Water suppression was achieved using a Water Gradient Tailored Excitation (WATERGATE) technique. 2D-Total Correlation Spectroscopy (TOCSY) experiments employed a DIPSI2 sequence for isotropic mixing, with a 75 ms mixing time unless otherwise noted. A 300 ms mixing time was used for Nuclear Overhauser Effect Spectroscopy (NOESY) and Rotational Overhauser Effect Spectroscopy (ROESY) experiments. All 2D experiments used STATES-TPPI phase cycling, and a spectral width of 14 ppm was applied in both dimensions with 2048 x 512 complex data points in the t2 and t1 dimensions, respectively. Prior to Fourier Transformation (FT), the data were linear predicted in the F1 dimension to 512 real points and then zero filled to produce a final matrix size of 4k x 1k, with 90^o phase-shifted sine squared window function applied to both dimensions. The chemical shifts were referenced to water at 4.7 ppm at 303K, pH 5.6, based on the method of Cavanagh et al. [18]. Spin-systems were manually assigned using the TOCSY data and sequential assignments were completed using the 300 ms NOESY spectrum [19]. Chemical shift data for S2 have been deposited in the Biological Magnetic Resonance Data Bank (code 18940).

3.0 Results

3.1 S2 is an order of magnitude more potent than S1- Previous studies using extrapolated dose-response data indicated that S2 is a more potent muscle relaxant than S1 [1]. Here we directly compared the relaxing effects of S1 and S2 at a range of concentrations on starfish cardiac stomach and tube foot preparations. S1 and S2 were tested on cardiac stomach preparations at 1 nM, 0.01 μM, 0.1 μM and 1 μM and the relaxing effect of S2 was found to be larger than the effect of S1 at all concentrations tested (Fig. 1A). Furthermore, the effect of 1 μM S1 (68 ± 8 %) was not significantly different ($p = 0.71$; t-test) to the effect of one-tenth the concentration of S2 (0.1 μM) (66 ± 7 %). Similarly, the effect of 0.1 μM S1 (34 ± 6%) was not significantly different ($p = 0.89$; t-test) to the effect of 0.01 μM S2 (33 ± 8%). S1 and S2 were also tested on tube foot preparations at 0.1 μM, 1.0 μM and 10 μM and, as with cardiac stomach preparations, the relaxing effect of S2 was found to be larger than the effect of S1 at all concentrations tested (Fig. 1B). Moreover, the magnitude of the difference in the relaxing activity of S1 and S2 was greater with tube foot preparations than with cardiac stomach preparations. For example, at 1 μM the mean relaxing effect of S1 was 68% of the mean effect of S2 on cardiac stomach preparations, whereas on tube feet the mean relaxing effect of S1 was only 14% of the effect of S2. These data demonstrate that S2 is at least an order of magnitude more potent than S1 as a muscle relaxant in starfish. This difference in the potency/activity of S1 and S2 as muscle relaxants in starfish provided a rationale for comparative analysis of the solution conformations of these peptides using CD spectroscopy and NMR spectroscopy.

3.2 CD spectroscopy of S1 and S2- CD spectroscopy in the far UV range was used to investigate the secondary structure propensity of S1 and S2 in aqueous solution and the effect of temperature and pH on their main-chain conformations (Fig. 2). The CD spectra obtained for S1 (Fig. 2A, B) have a strong negative band at 200 nm along with a weak positive band at 218 nm. These CD bands suggest that S1 has predominantly an irregular random-coil conformation in aqueous solution [20]. Varying the temperature from 10°C to 50°C (at pH 5.5; Fig. 2A) and the pH from 5.3 to 8.3 (at 11^{o}C; Fig. 2B) had little effect on the appearance of the CD spectra. Both resulted in a slight reduction in the amplitude of the bands at 200 nm and 218 nm but no shift in the wavelength of each of the peaks. The lack of a shift in the wavelengths suggests that S1 remains unstructured over this range of pHs and temperatures.

The CD spectra for S2 were quite different to those obtained for S1, with negative bands at 207 nm and 196 nm and a weak positive band at 228 nm (Fig. 2C, 10°C spectrum). These spectra are different to those observed with peptides that have α-helical or β-sheet type conformations; however, they are also inconsistent with a random coil conformation, which are characterized by a single negative CD band at ~198 nm. It is possible that S2 adopts an extended conformation that does not contain formal β-strands; this is also supported by NMR chemical shift analysis and the absence of main-chain hydrogen bonding (discussed in section 3.5). The negative peak near 198 nm is suggestive of an unstructured peptide, but proline residues are known to have transitions at 198 nm, with the trans Xxx-Pro bond having a negative Cotton effect and the cis Xxx-Pro bond having a positive effect at 198 nm [21, 22]. It is likely that the CD band at 198 nm for S2 is a combination of signals from a structured, proline-containing peptide along with 'random coil' type signals. The slight intensity reduction at 207 nm and 228 nm as the temperature increases (Fig. 2C) could be due to some loss of overall structure.

Increasing the pH from 5.5 to 8.4 reduced the negative band at 196 nm, broadened negative bands at 207 nm and resulted in loss of the positive band at 228 nm (Fig. 2D). Change in pH can affect the Xxx-pro cis-trans equilibrium and it is plausible that the loss of the negative signal at 198 nm is a result of a greater proportion of the cis isoform at high pHs. Above neutrality the peak at 207 nm not only broadens but also moves to longer wavelengths with a concomitant decrease in the intensity of the 228 nm positive peak. Possibly, the increasing pH causes the structure to adopt some β-turns. Indeed, β-turns tend to have a positive signal near 200 nm and a slight negative signal near 220. The appearance of β-turn-like structure might also account for the loss of negative intensity around 198 nm. Overall, and over a range of pH values, it is clear that temperatures up to 30^{o}C do not cause S2 to unfold into an irregular structure. Based on these observations we chose to study both S1 and S2 further using NMR at pH 5.8 and 30 °C.

3.3 ^{1}H NMR comparison of S1 and S2- Two-dimensional NMR spectra were obtained for both S1 and S2 in order to determine if any unique structures existed. From the 2D TOCSY spectra we were able to assign complete spin systems for all but Gly_1 and Phe_2 of S1 (supplementary material Table S1A). The 1D spectrum (Fig. 3) shows considerable line broadening of Phe_2 suggesting that the N-terminus of S1 is exchange broadened, undergoing milli-to-micro second time-scale motions. Furthermore, the limited chemical shift dispersion of the amide protons suggests the S1 peptide is unstructured. The NOESY spectrum for S1 gave a few weak NOE cross-peaks for intra-residue NOEs suggesting that the size of the peptide and the resulting correlation time causes the NOE to be weak and close to the null. For this reason, a ROESY spectrum was obtained from which we were able to obtain sequential assignments as shown in Fig. 3 (H_α-H_N fingerprint region). It was possible to unambiguously assign all peaks present as

S1 contained six unique amino acids and two phenylalanines. However, only sequential connectivities were observed for S1 and no long range ROEs that would be indicative of stable structure were apparent. The fact that the ROESY spectrum of S1 provided no further information beyond a limited number of sequential assignments is entirely consistent with the CD data, which indicated that S1 exists in a random-coil conformation.

From the TOCSY spectrum of S2 (Fig. 4A) we were able to assign all but Ser_1 and, in contrast to S1, we were able to obtain a ^1H-^1H NOESY spectrum that contains a remarkable number of cross peaks (the complete 300 ms NOESY spectrum is shown as supplementary material Fig. S1). Sequential assignments for S2 were readily achieved using standard 2D methods [19] and are shown as supplementary material Table S1B. The H_α-H_N fingerprint region is shown in Fig. 4B, and from this a complete sequential path could be followed from Ser_1 through to the C-terminal amide on Phe_{12}, with only a break at Pro_3. In the spectrum there were two sets of peaks associated with the N-terminal residues Gly_2, Pro_3, Tyr_4 and Ser_5, and one set was more intense than the other (Fig. 4B). In the aliphatic region (see supplementary material Fig. S1), a strong $d_{\alpha\delta}(i, i+1)$ NOE was observed between the Gly_2 and Pro_3 for the most intense set of peaks indicating that these originated from a *trans* conformation around the Gly_2-Pro_3 peptide bond. The lower intensity peaks had a $d_{\alpha\alpha}(i, i+1)$ between Gly_2 and Pro_3 consistent with a *cis* conformation. The large chemical shift differences between residues in the *cis* and *trans* conformations, notably Tyr_4 and Ser_5 observed in Fig. 4B, is likely due to the different conformation of the aromatic side-chain of Tyr_4, and the consequently altered ring-current effects. On intensity measurements, the *trans*-conformation is about 80-85% prevalent. A weak $d_{NN}(i, i+1)$ NOE is apparent between Ser_{5Cis} and Phe_6 (data not shown), which suggests that residues beyond Ser_5 are not influenced by the *cis-trans* isomerism and they have the same

structure in each isoform. This is also supported by the presence of an NOE between the Ser_{5Cis} β-protons and the ring protons of Phe_6 (data not shown).

The H_α-H_N region shows several NOEs indicative of substantial order in S2. For instance, $d_{\alpha N}(i, i+2)$ connections are readily observed for residues from Gly_2 through to Thr_{11} and a longer range $d_{\alpha N}(i, i+3)$ NOE and a weak $d_{\alpha N}(i, i+4)$ are observed from Gly_2 to Ser_5 and Phe_6 respectively. Further, $d_{\beta N}(i, i+1,2,3)$ NOEs are apparent for the β-protons of Ser_5 and Ser_8. Along with the $d_{\alpha N}$ connections, there are also several, relatively weak, $d_{NN}(i, i+2)$ connections for Tyr_4, Ser_5, Phe_6 and Ser_8 (data not shown). Remarkably, in all there were more than 220 NOEs observed for the S2 peptide. We were surprised by this amount of structuring, as indicated by the large number of NOEs, for such a small (12 residue) peptide and were intrigued as to the source of the structuring in the absence of disulphide bonds.

In the absence of disulphide bonds, it is possible that extensive backbone hydrogen bonding hold S2 in a rigid conformation. To investigate this we explored the temperature dependence of amide proton NMR resonances. ^{1}H NMR amide temperature coefficients can be a reliable indicator of the presence of hydrogen bonding. In particular, temperature coefficients of between -1 ppb/K and -4 ppb/K indicate a high probability (93%) of the presence of a hydrogen-bonded amide [23, 24]. 1D spectra for S2 were obtained over a range of temperatures and the change in amide proton chemical shift as a function of temperature (283 K – 313 K) was determined. Table 1 shows that for S2 the smallest temperature coefficient observed is |-4.6| ppb/K for Leu_{10}, a value which can predict a hydrogen bond in more than 85% of cases. The temperature coefficients for the other amides in S2 suggest hydrogen bonding is unlikely to occur. Experiments with S1 did not show any temperature coefficient small enough to predict any hydrogen bonding. Furthermore, $^3J_{H\alpha NH}$ coupling constants for S2 are not indicative of any

regular secondary structure. So with no extensive hydrogen bonding network, why is there so much structuring of S2, as apparent from the large number of NOE contacts observed and also evident in the CD spectra?

3.4 Structuring of S2 is caused by self-association- Next we investigated the concentration dependence of the CD spectrum for the S2 peptide. It is clear from Fig. 5A that the structuring of the S2 peptide is highly concentration dependant. An increase in concentration from 0.01 mg/ml to 0.1 and 0.2 mg/ml (the concentration at which the NMR data was obtained) causes a profound increase in the structuring of the S2 peptide, as indicated by a strong negative CD band at 216 nm, typical of an extended-conformation. With the strong concentration dependence of the CD spectra it is clear that self-association is the source of the observed structuring of S2. At the lower concentration (0.01 mg/ml) there is little structuring of S2 with a single weak negative CD band at 198 nm, typical of irregular structure.

We went on and studied the thermal stability of the S2 peptide, at two different concentrations. At 0.1 mg/ml S2 undergoes a clear thermal unfolding transition. The unfolding curve has a sigmoidal appearance typical for cooperative, essentially two-state unfolding, with a mid-point at 55°C (Fig. 5B). The temperature dependence of the CD spectra for S2 at 0.01 mg/ml is very different; there is no cooperative unfolding with temperature. The increase in temperature causes a loss of some PP-II character in the spectra, but a co-operative unfolding transition is not apparent for the dilute S2 sample.

Given that the CD data suggests that S2 self-associates at the concentrations used here for NMR, NOESY peaks other than the intense sequential NOEs may arise from intra- or inter-peptide contacts. For instance, the β-protons of Ser_5 have clear NOEs to the amide protons of

Phe_6, Asn_7 and Ser_8 along with a reasonably intense NOE to the amide proton of Tyr_4 (Fig. 4B). Although the gradually weakening NOEs to residues 6, 7, and 8 could be expected to be intra-peptide, the NOE to Tyr_4 may arise from an intra-peptide contact or from an inter-peptide contact. Although there are a large number of NOEs (220 pairs), and particularly because of the small size of the peptide, determining which of these NOEs are intra-molecular and which are inter-molecular is difficult. Thus, standard approaches to generate a monomeric solution structure from the NOE data set are not appropriate. Nevertheless, the large number of NOEs and the ability to observe a well resolved set of resonances for residues Gly_2 to Phe_{12} indicates that self-association induces a single, soluble, low molecular mass structure in S2. The narrow line-widths of the amide NMR peaks (~4 Hz, at half-height) suggest that the size of the complex is small, perhaps only dimeric. Obtaining NMR spectra at lower concentrations in order to differentiate inter-molecular NOEs from intra-molecular ones is problematic as disrupting the oligomer also causes loss of structure. Therefore, we cannot assume that loss of a particular NOE necessarily indicates it is a inter-molecular contact as loss of structure means only sequential NOEs are observed, as seen for S1 (Fig. 3).

3.5 Evidence of a role for the N-terminal region of S2 in self-association- The clear difference between S1 and S2 in their solution conformations is intriguing given their sequence similarity. However, S2 has the N-terminal sequence SGPY that is absent in S1. To determine if this sequence impacted on the self-association property of S2, we constructed a modified S2 peptide that lacks this N-terminal sequence (short S2 or SS2; SFNSGLTFamide). The TOCSY and ROESY spectra of this peptide, at the same concentration as S2, are shown overlaid in Fig. 6A. The dispersion of amide peaks in the 1D spectrum is similar to that of S1 and suggests a lack of

regular structure. Moreover, the array of NOEs apparent in S2 (Fig. 4B) is missing in SS2 and only sequential NOEs are visible (Fig 6A). The lack of a large number of NOEs in the spectrum of SS2 suggests that the N-terminal sequence SGPY is an important factor in the self-association of S2 and subsequent generation of a defined structure. Aromatic amino acids can often associate via π-π interactions and it is plausible that Tyr_4 is required to support such interactions and therefore S2 self-association. Intriguingly, analysis of the H_α chemical shift index (i.e. deviation from 'random coil') of S2 and SS2 shows that they are reasonably similar (Fig. 6B), but dissimilar to S1 (Fig. 6C). Note that although trending negative from Phe_6, the CSI shifts for S2 and SS2 are not great enough to suggest any true β-sheet-like structure exists. Nevertheless, the similarity between S2 and SS2 shifts suggests that SS2 may be accessing the same conformational space as is S2 despite not forming a defined structure.

3.6 Analysis of the bioactivity of SS2- Having analysed the solution conformation of SS2 in comparison with S2, it was of interest to compare the bioactivity of SS2 with that of S1 and S2. Therefore, we tested SS2 on the two preparations that we originally used (Fig. 1) to compare the pharmacological activity of S1 and S2 – starfish cardiac stomach and tube foot preparations.

Consistent with relaxing effects of S1 and S2, SS2 caused dose-dependent relaxation of cardiac stomach preparations when tested at 0.1 μM, 1 μM and 10 μM (Fig. 7A). Comparison of the relaxing actions of S1, SS2 and S2 when tested at 1 μM revealed that SS2 was significantly more effective than S1 ($p = 0.0004$; t-test) and slightly, but significantly ($p = 0.03$; t-test), less effective than S2 (Fig. 7B). These findings indicate that the biological activity of S2 is largely attributable to its C-terminal octapeptide sequence (SFNSGLTFamide).

Next an analysis of the effects of SS2 on tube foot preparations yielded some interesting findings. As with the cardiac stomach preparations, SS2 caused dose dependent relaxation of tube-feet (Fig. 7C). However, it was noteworthy that the effects of SS2 at 10 μM were consistently larger than the effects of S2 at the same concentration. When the effects of S1, S2 and SS2 were compared at 1 μM, SS2 was significantly ($p = 0.00006$; t-test) more effective than S1 and significantly ($p = 0.000003$; t-test) less effective than S2 (Fig. 7D). Thus, for tube foot preparations conflicting conclusions can be drawn from the data obtained, with results from tests at 10 μM indicating that the presence of the N-terminal SGPY sequence actually impairs the bioactivity of S2, while results from tests at 1 μM indicate that the presence of the N-terminal SGPY sequence contributes to the bioactivity of S2.

4.0 Discussion

S2 is related to the SALMFamide neuropeptide S1 (GFNSALMFamide), which shares with S2 the C-terminal sequence motif FNSxLxFamide (where is x is variable). Both S1 and S2 cause muscle relaxation in starfish but S2 is more effective than S1 when tested at the same concentration [8]. Here we have specifically compared the potency of S1 and S2 by testing both peptides at a range of concentrations on two different preparations from the starfish *Asterias rubens*, the cardiac stomach and tube feet. These experiments revealed that S2 is at least an order of magnitude more potent than S1, which provided the rationale for a comparative investigation of the solution structures of S1 and S2 using CD and NMR spectroscopy.

CD spectroscopy indicates that S1 is largely unstructured in aqueous solution and, consistent with the CD data, the NMR NOESY spectrum of S1 showed only sequential NOEs, which implies a lack of a unique, stable conformation. This is in line with other neuropeptide studies showing that small peptides like S1 are usually unstructured in aqueous solutions [9-16]. In contrast, the NMR NOESY spectrum for S2 showed a remarkable array of NOE cross-peaks (220 NOE pairs) indicative of a well-structured peptide. The H_α chemical shift analysis, the lack of regular hydrogen bonding and the pattern of NOEs, all indicate that S2 is not composed of a regular α-helical or β-sheet like structure. However, the sequence of S2 has polar, hydrophilic amino acids interspersed with hydrophobic and aromatic amino acids, which may support the generation of some structure. Keire *et al*. [9] have shown that an eight residue cholecystokinin peptide CCK8 (DYMGWMDF-NH_2) in aqueous solution has a preference for γ-and β-turns around GWMD and MDF-NH_2. In the case of S2, the structure adopted is dependent on the peptide concentration. At lower concentrations S2 is unstructured, and presumably monomeric, whereas increasing the concentration induces structure, which is likely due to the formation of

low molecular weight oligomers, perhaps dimers. However, because of the small size of the peptide it is not possible to ascertain, from the numerous NOEs observed, the type of contacts between S2 peptides (i.e. head-to-head or head-to-tail).

Many neuropeptides that are unstructured and flexible in aqueous solvents tend to adopt structure when in the presence of stabilising solvents such as alcohols (trifluoroethanol, for instance), as well as in the presence of membranes. This finding has led to the proposition of a two-step theory of neuropeptide/receptor interaction wherein the neuropeptide binds first to the membrane and then undergoes diffusion to the appropriate receptor [25]. It is envisaged that interaction with the membrane not only reduces the degrees of freedom but also stabilises conformations that promote receptor binding.

Our discovery that S2 has a well defined, but concentration dependent conformation in aqueous solution may have physiological implications. The concentration of neuropeptides in vesicles prior to release is high and structuring may support a higher degree of packing in vesicles. Furthermore, high peptide concentrations may persist locally after release and therefore analysis of the self-association structure and process may be relevant to function. Indeed, studies with neuropeptide Y, whose physiological roles include regulation of food intake and pain perception, have shown that the hydrophobic residues that support self-association in this peptide also allow binding to cell membranes [26].

The conformation of S2 in aqueous solution may explain why S2 is more potent than S1 as a muscle relaxant in starfish. We speculate that S2 might bind to its receptor as an oligomer (dimer) and that the solution conformation of S2 is the same or very similar to its conformation when bound to a receptor. Therefore the probability of receptor-binding and receptor-activation may be greater than for a peptide (S1) that lacks a well-defined conformation in aqueous solution

and that will have an entropic penalty when binding to its receptor. Earlier studies on FMRFamide-like peptides have shown that, although not adopting a unique conformation, some can contain significant populations of turns and the percentage of turns is highly correlated with IC_{50} values of receptor binding [27]. This suggests that the presence of structure in neuropeptides confers an added level of 'fine-tuning' of receptor interactions.

Clearly the key structural difference between S1 and S2 is the presence of four N-terminal residues (SGPY) in S2 that are lacking in S1. It appears, therefore, that the presence of these residues confers on S2 an ability to adopt a stable conformation in solution upon self-association. Our finding that S2 and SS2 (short S2) have similar H_α chemical shift indexes suggests that SS2 may explore a conformation similar to that adopted by S2 at the same concentration, but a stable structure is not generated. Accordingly, differences in the C-terminal regions of S1 and S2 clearly contribute to the differing potencies of S1 and S2 because tests with SS2 revealed that it is significantly more active as a muscle relaxant than S1. It is the N-terminal SGPY motif in S2 that appears to be the key to its conformational stability in aqueous solution and this motif contributes to superior bioactivity of S2 compared to both SS2 and S1 when tested at 1 μM and 10 μM on cardiac stomach preparations and when tested at 1 μM on tube preparations. However, our finding that SS2 is more active as a muscle relaxant than S2 when tested on tube feet at 10 μM suggests complex structure-activity relationships, which may be influenced by the combinations of receptor types that are expressed in different organs. Further investigation of this issue will be possible if the receptors that mediate the effects of SALMFamides are identified (see [28] for a discussion of a potential strategy for identification of SALMFamide receptors). If there are multiple receptor types that mediate the effects of SALMFamides, then tissue-specific differential expression of receptor types may explain why

dose-dependent differences in the relative activity S1, S2 and a synthetic analog of S2 (SS2) are observed.

From a physiological perspective it is of interest to consider why two peptides (S1 and S2) with the same effect (muscle relaxation) but with different potencies both exist in starfish. Immunocytochemical studies have revealed that S1 and S2 are expressed by different populations of neurons in the starfish *Asterias rubens* and this finding is supported by analysis of the genome sequence of the starfish *Patiria miniata*, which has revealed that S1 and an S2-like peptide are derived from precursor proteins encoded by different genes [7]. Moreover, in *Asterias rubens* S2 is detected by radioimmunoassay in the perivisceral coelomic fluid whereas S1 is not [17]. This suggests that S2 may act as a hormonal signalling molecule in starfish and therefore it would be necessary for its potency to be greater than that of S1, which may only act more locally as a neurotransmitter or paracrine signalling molecule. Thus, we speculate that the structural features of S2 that confer on it greater potency compared to S1 may be molecular adaptations for a hormonal signalling role.

Finally, it has recently been discovered that in the starfish *Patiria miniata*, S1 is derived from a precursor protein that comprises six other putative L-type or L-type-like SALMFamides [7]. On the other hand, the S2-like neuropeptide in this species is derived from a different precursor protein that comprises eight F-type or F-type-like SALMFamides [7]. Our investigation of a structural basis for differences in the activity of S1 and S2, as reported here, therefore provides the foundations for further studies on more complex "cocktails" of SALMFamides that occur *in vivo*.

Acknowledgements - We are grateful to Harold Toms (QMUL) for technical assistance with the NMR studies, Richard Melarange (QMUL) for technical assistance with *in vitro* pharmacology and Urooj Khan (UWS) for analysis of some of the NMR data.

REFERENCES

[1] M.R. Elphick, R. Melarange, Neural control of muscle relaxation in echinoderms, J Exp Biol, 204 (2001) 875-885.
[2] M.L. Rowe, M.R. Elphick, Discovery of a second SALMFamide gene in the sea urchin *Strongylocentrotus purpuratus* reveals that L-type and F-type SALMFamide neuropeptides coexist in an echinoderm species, Mar Genomics, 3 (2010) 91-97.
[3] M.R. Elphick, D.A. Price, T.D. Lee, M.C. Thorndyke, The SALMFamides: a new family of neuropeptides isolated from an echinoderm, Proc Biol Sci, 243 (1991) 121-127.
[4] R. Melarange, D.J. Potton, M.C. Thorndyke, M.R. Elphick, SALMFamide neuropeptides cause relaxation and eversion of the cardiac stomach in starfish., Proc Biol Sci, 266 (1999) 1785-1789.
[5] S.J. Newman, M.R. Elphick, M.C. Thorndyke, Tissue distribution of the SALMFamide neuropeptides S1 and S2 in the starfish *Asterias rubens* using novel monoclonal and polyclonal antibodies. 2. Digestive system, Proc Biol Sci, 261 (1995) 187-192.
[6] S.J. Newman, M.R. Elphick, M.C. Thorndyke, Tissue distribution of the SALMFamide neuropeptides S1 and S2 in the starfish *Asterias rubens* using novel monoclonal and polyclonal antibodies. I. Nervous and locomotory systems, Proc Biol Sci, 261 (1995) 139-145.
[7] M.R. Elphick, S. Achhala, N. Martynyuk, The evolution and diversity of SALMFamide neuropeptides, PLoS One, 8 (2013) e59076.
[8] R. Melarange, M.R. Elphick, Comparative analysis of nitric oxide and SALMFamide neuropeptides as general muscle relaxants in starfish, J Exp Biol, 206 (2003) 893-899.
[9] D.A. Keire, T.E. Solomon, J.R. Reeve, Jr., NMR evidence for different conformations of the bioactive region of rat CCK-8 and CCK-58, Biochem Biophys Res Commun, 293 (2002) 1014-1020.
[10] G.Y. Xu, C.M. Deber, Conformations of neurotensin in solution and in membrane environments studied by 2-D NMR spectroscopy, Int J Pept Protein Res, 37 (1991) 528-535.
[11] J.L. Nieto, M. Rico, J. Santoro, J. Herranz, F.J. Bermejo, Assignment and conformation of neurotensin in aqueous solution by 1H NMR, Int J Pept Protein Res, 28 (1986) 315-323.
[12] S.C. Sumner, K.S. Gallagher, D.G. Davis, D.G. Covell, R.L. Jernigan, J.A. Ferretti, Conformational analysis of the tachykinins in solution: substance P and physalaemin, J Biomol Struct Dyn, 8 (1990) 687-707.
[13] G. Chassaing, O. Convert, S. Lavielle, Preferential conformation of substance P in solution, Eur J Biochem, 154 (1986) 77-85.
[14] S.M. Cowsik, C. Lucke, H. Ruterjans, Lipid-induced conformation of substance P, J Biomol Struct Dyn, 15 (1997) 27-36.
[15] R.C. Grace, A.M. Lynn, S.M. Cowsik, Lipid induced conformation of the tachykinin peptide Kassinin, J Biomol Struct Dyn, 18 (2001) 611-621, 623-615.
[16] R.C. Grace, I.R. Chandrashekar, S.M. Cowsik, Solution structure of the tachykinin peptide eledoisin, Biophys J, 84 (2003) 655-664.
[17] M.R. Elphick, S.J. Newman, M.C. Thorndyke, Distribution and action of SALMFamide neuropeptides in the starfish Asterias rubens, J Exp Biol, 198 (1995) 2519-2525.
[18] F.W. Cavangh J, Palmer AG, Skelton NJ, Protein NMR Spectroscopy: Principles and Practice, Academic Press Inc., San Diego, Ca, USA, 1996.
[19] K. Wuthrich, NMR of proteins and nucleic acids, Wiley, New York, USA, 1986.

[20] N.J. Greenfield, Applications of circular dichroism in protein and peptide analysis, Trac-Trend Anal Chem, 18 (1999) 236-244.
[21] J.D. Fontenot, N. Tjandra, D. Bu, C. Ho, R.C. Montelaro, O.J. Finn, Biophysical Characterization of One-Tandem, 2-Tandem, and 3-Tandem Repeats of Human Mucin (Muc-1) Protein Core, Cancer Res, 53 (1993) 5386-5394.
[22] K. Ma, L.S. Kan, K. Wang, Polyproline II helix is a key structural motif of the elastic PEVK segment of titin, Biochemistry-Us, 40 (2001) 3427-3438.
[23] T. Cierpicki, J. Otlewski, Amide proton temperature coefficients as hydrogen bond indicators in proteins, J Biomol Nmr, 21 (2001) 249-261.
[24] J. Feeney, The conformation of hormonal peptides, in: G. Roberts (Ed.) Drug action at the molecular level [Proceedings of a Symposium, 1976], Univ. Park Press, Baltimore, 1977, pp. 55-76.
[25] H. Inooka, T. Ohtaki, O. Kitahara, T. Ikegami, S. Endo, C. Kitada, K. Ogi, H. Onda, M. Fujino, M. Shirakawa, Conformation of a peptide ligand bound to its G-protein coupled receptor, Nat Struct Biol, 8 (2001) 161-165.
[26] R. Bader, A. Bettio, A.G. Beck-Sickinger, O. Zerbe, Structure and dynamics of micelle-bound neuropeptide Y: comparison with unligated NPY and implications for receptor selection, Journal of molecular biology, 305 (2001) 307-329.
[27] A.S. Edison, E. Espinoza, C. Zachariah, Conformational ensembles: the role of neuropeptide structures in receptor binding, J Neurosci, 19 (1999) 6318-6326.
[28] M.R. Elphick, From gonadotropin-inhibitory hormone to SIFamides: Are echinoderm SALMFamides the "missing link" in a bilaterian family of neuropeptides that regulate reproductive processes?, Gen Comp Endocrinol, (2013).

FIGURE LEGENDS

Fig. 1. Comparison of the potency of S1 and S2 as muscle relaxants. A, Graph comparing the relaxing effects of S1 (filled bar) and S2 (empty bar) (0.001 – 1 μM) on starfish cardiac stomach preparations, expressed as a mean percentage (± SEM; n = 15) of the effect of 1 μM S2. The effect of 1 μM S1 is not significantly different (p = 0.71; t-test) to the effect of 0.1 μM S2 and the effect of 0.1 μM S1 is not significantly different (p = 0.89; t-test) to the effect of 0.01 μM S2. B, Graph comparing the relaxing effects of S1 and S2 (0.1 – 10 μM) on starfish tube foot preparations, expressed as a mean percentage (± SEM; n = 4 for S1; n = 18 for S2) of the effect of 10 μM S2.

Fig. 2. CD spectra of S1 and S2 peptides in water. A, CD of S1 as the temperature changes at constant pH (5.3). Varying the temperature from 10°C to 50°C had little effect on the CD spectra, with just a slight reduction in the amplitude of the CD signal. B, CD of S1 as the pH changes from pH 5.3 to pH 8.3 at constant temperature (11°C). Similarly, there is a slight reduction in the amplitude of the CD signal as the pH increases. In both A and B there is a strong negative band at 200 nm characteristic of a random coil conformation. C. CD of S2 at different temperatures and constant pH (5.5). Increasing the temperature from 10°C to 50°C results in a slight reduction in the amplitude of the negative bands at 196 nm and 207 nm. D, Increasing the pH (at a constant temperature of 11°C) from 5.5 to 8.4 reduces the negative band at 196 nm and causes the negative band at 207 nm to shift to longer wavelengths.

Fig. 3. ROESY spectrum of S1. The H_α-H_N fingerprint region of the S1 neuropeptide ROESY (τ_m 300 ms) spectrum at 293K, pH 5.6, showing the $d_{\alpha N}(i, i+1)$ assignments of S1. The box

around F2 denotes the H_α-H_N peak that is observed in the TOCSY spectrum but not in the ROESY spectrum.

Fig. 4. NMR data for S2 highlights the presence of structure**.** A, The H_α-H_N finger print region of the S2 TOCSY (τ_m 75 ms) spectrum in water at 303K, pH 5.6 showing spin-system assignment. B, The H_α-H_N fingerprint region of the S2 NOESY (τ_m 300 ms) spectrum in water at 303K, pH 5.6. Solid lines show the $d_{\alpha N}(i, i+1)$ sequential assignments of S2. Dashed lines identify residues from the *cis*-Pro_3 isoform.

Fig. 5: Concentration dependence of S2 peptide structure and folding stability. Far UV-CD spectra of S2, showing concentration dependent secondary structure of S2 at 0.01 mg/ml (7.8 μM) (red), 0.1 mg/ml (78 μM) (black) and 0.2 mg/ml (156 μM) (blue) (A). Thermal denaturation of S2, between 10-90^{o}C, blue spectra are from 10^{o}C and red are at 80^{o}C. Spectra were obtained at two concentrations; 0.1 mg/ml (B), and 0.01 mg/ml of S2 (C). Inserts are thermal denaturation curves at 210 nm and 228 nm, respectively.

Fig. 6. NMR analysis of SS2 (short S2). A, the H_α-H_N fingerprint region of the TOCSY (blue, τ_m 65 ms) and ROESY (red, (τ_m 300 ms) spectra of SS2 shown overlaid. The sequential $d_{\alpha N}(i, i+1)$ assignments are shown with solid lines. B, H_α chemical shift analysis of S2 (black bars) and SS2 (white bars). C, H_α chemical shift analysis of S1.

Fig. 7. Analysis of the effects of SS2 as a muscle relaxant and comparison of the activity of SS2 with S1 and S2. A, SS2 causes dose-dependent relaxation of starfish cardiac stomach

preparations; error bars represent mean ± SEM (n=5) and data are normalised to the effect of S2 at 10 μM. B, Comparison of the relaxing effect of SS2 to that of S2 and S1 on cardiac stomach (all at 1 μM). Data are normalised to relaxation effect of S2 (at 1 μM); error bars represent mean ± SEM (n=12). Significance *** = $p < 0.001$ (students t-test). C, SS2 causes dose-dependent relaxation of starfish tube foot preparations; error bars represent mean ± SEM (n=9) and data are normalised to the effect of S2 at 10 μM. D, Comparison of the relaxing effect of SS2 to that of S2 and S1 on tube feet. Data are normalised to relaxation effect of S2 (1 μM); error bars represent mean ± SEM (n=13). Significance *** = $p < 0.001$, * = $p<0.05$.

REVISIONS MARKED

Structural analysis of the starfish SALMFamide neuropeptides S1 and S2: the N-terminal region of S2 facilitates self-association

Claire B. Otara[a*], Christopher E. Jones[a, b*], Nadine D. Younan[a], John H. Viles[a] and Maurice R. Elphick[a]

[a.] The School of Biological & Chemical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, UK. [b.] The School of Science and Health, The University of Western Sydney, Locked bag 1797, Penrith, 2751, Sydney, Australia.

*CBO and CEJ contributed equally

Running title: *Self-association of the SALMFamide neuropeptide S2*

To whom correspondence should be addressed:

Maurice R. Elphick, School of Biological & Chemical Sciences, Queen Mary, University of London, Mile End Road, London, E1 4NS, UK. Tel: 0207 882 5290; Fax: 0207 882 7732; E-mail: m.r.elphick@qmul.ac.uk

or

John H. Viles, School of Biological & Chemical Sciences, Queen Mary, University of London, Mile End Road, London, E1 4NS, UK. Tel: 0207 882 8443; Fax: 0207 882 7732; E-mail: j.viles@qmul.ac.uk

ABSTRACT

The neuropeptides S1 (GFNSALMFamide) and S2 (SGPYSFNSGLTFamide), which share sequence similarity, were discovered in the starfish *Asterias rubens* and are prototypical members of the SALMFamide family of neuropeptides in echinoderms. SALMFamide neuropeptides act as muscle relaxants and both S1 and S2 cause relaxation of cardiac stomach and tube foot preparations *in vitro* but S2 is an order of magnitude more potent than S1. Here we investigated a structural basis for this difference in potency using spectroscopic techniques. Circular dichroism spectroscopy showed that S1 does not have a defined structure in aqueous solution and this was supported by 2D nuclear magnetic resonance experiments. In contrast, we found that S2 has a well-defined conformation in aqueous solution. However, the conformation of S2 was concentration dependent, with increasing concentration inducing a transition from an unstructured to a structured conformation. Interestingly, this property of S2 was not observed in an N-terminally truncated analog of S2 (short S2 or SS2; SFNSGLTFamide). Collectively, the data obtained indicate that the N-terminal region of S2 facilitates peptide self-association at high concentrations, which may have relevance to the biosynthesis and/or bioactivity of S2 *in vivo*.

Keywords: neuropeptide, SALMFamide, circular dichroism, NMR, muscle relaxant, self-association

1.0 Introduction

The SALMFamides are a family of neuropeptides that act as muscle relaxants in species belonging to the phylum Echinodermata (e.g. starfish, sea urchins, sea cucumbers). Two types of SALMFamides have been identified in echinoderms: L-type SALMFamides, which have the C-terminal motif SxLxFamide, and F-type SALMFamides, which have the C-terminal motif SxFxFamide [1, 2]. The first members of the SALMFamide neuropeptide family to be identified were two L-type SALMFamides, S1 and S2, which were both isolated from the nervous system of the starfish species *Asterias rubens* and *Asterias forbesi* [3]. S1 is an amidated octapeptide (Gly-Phe-Asn-Ser-Ala-Leu-Met-Phe-NH_2) and S2 is an amidated dodecapeptide (Ser-Gly-Pro-Tyr-Ser-Phe-Asn-Ser-Gly-Leu-Thr-Phe-NH_2), which share the C-terminal sequence Phe-Asn-Ser-X-Leu-X-Phe-NH_2 (where X is variable).

Investigation of the physiological roles of S1 and S2 has revealed that injection of these peptides into starfish (*Asterias rubens*) causes eversion of the cardiac stomach. Consistent with these *in vivo* effects, analysis of the effects of S1 and S2 *in vitro* revealed that both peptides cause relaxation of the cardiac stomach [4]. These findings are of interest from a physiological and behavioural perspective because starfish feed by everting their cardiac stomach over the digestible parts of prey such as mussels and oysters. Therefore, S1 and/or S2 may mediate neural control of stomach eversion in starfish *in vivo* by causing relaxation of muscle associated with the cardiac stomach. Consistent with this notion, immunocytochemical studies have revealed that S1 and S2 are present in nerve fibres in close proximity to the muscle layer of the cardiac stomach [5]. Furthermore, S1 and S2 appear to be expressed in different populations of

neurons in *Asterias rubens* [5, 6] and consistent with this observation recent analysis of the genome sequence of the starfish *Patiria miniata* has revealed that S1 and an S2-like peptide are encoded by different genes [7].

Comparison of the effectiveness of S1 and S2 as muscle relaxants in starfish has revealed that S2 is more effective than S1 when tested at the same concentration (10 μM) on three different preparations from *Asterias rubens*: the cardiac stomach, tube feet and the body wall apical muscle [8]. Furthermore, extrapolated dose-response data indicate that S2 may be approximately ten times more potent than S1 [1]. However, a direct comparison of the potency of S1 and S2 as muscle relaxants in starfish has as yet not been conducted. Therefore, here we have compared the relaxing effects of both S1 and S2 at a range of concentrations on two different preparations, the cardiac stomach and tube feet.

It is not known if S1 and S2 exert their relaxing effects on starfish muscle preparations by binding to a single receptor type or via different receptors, although it is assumed that these peptides act on G protein-coupled receptors [1]. Nevertheless, given the similarity of the C-terminal regions of S1 (FNSALMFamide) and S2 (FNSGLTFamide), we hypothesized that the presence of four additional N-terminal residues in S2 (SGPY) may influence its structure in solution and account, at least in part, for differences in the potency of S1 and S2. To address this issue, here we have used circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy to analyse and compare the conformations of S1 and S2 in aqueous solution.

Previous studies on other neuropeptides comparable in size to S1 and S2 have revealed that they exist in aqueous solution as an ensemble of rapidly interconverting structures lacking stable secondary or tertiary structural elements [9]. For example, the neuropeptide neurotensin, which comprises thirteen amino acid residues, exists mainly as a flexible “random coil”

polypeptide chain in water [10, 11]. Similarly, the mammalian tachykinin peptides substance P, neurokinin A and neurokinin B and the non-mammalian tachykinins physalaemin, eledoisin and kassinin all adopt a variety of interconverting conformational states in water [12-16]. Therefore, our expectation was that both S1 and S2 would similarly adopt multiple conformational states in aqueous solution. Surprisingly, however, we have discovered that S2, but not S1, has a remarkably well-defined conformation in aqueous solution, and the structuring is caused by self-association. The structural basis and functional relevance of this property of S2 is investigated and discussed.

2.0 Materials and Methods

2.1 Materials - Peptides were custom-synthesized by the Advanced Biotechnology Centre at Imperial College London and purified using high performance liquid chromatography. All other chemicals used were obtained from VWR (Poole, Dorset, UK) with the exception of D_2O, which was obtained from Goss Scientific Instruments Ltd (Great Baddow, Essex, UK).

2.2 In vitro *pharmacology* - Starfish (*Asterias rubens*) were obtained from the Menai Straits (UK) and then maintained in a circulating seawater aquarium in the School of Biological & Chemical Sciences at QMUL. Cardiac stomach preparations were dissected from animals ranging in diameter from 12 – 16 cm and then linked to an isotonic transducer (model 60-3001; Harvard, South Natick, MA, USA) in a 20 ml organ bath containing aerated seawater at 11°C, as described previously [1, 17]. Changes in cardiac stomach length were recorded on a chart recorder (Goerz Servogor 124). Sustained contracture of the cardiac stomach was induced and maintained using seawater with 30 mM added KCl, as described previously [4]. Solutions of S1 and S2 were then prepared in water and added (20 µl) to the organ bath to achieve final bath concentrations of 0.001, 0.01, 0.1 and 1 µM. S1 and S2 were tested on a total of fifteen cardiac stomach preparations, with both peptides being applied twice at each concentration tested. The order in which different peptides and concentrations were tested was random after having first tested 1 µM S2. Preparations were washed between tests by refilling the organ bath several times with seawater/KCl and then allowing a 15-20 min interval before the next test, during which time preparations regained basal resting length. The effects of each peptide concentration tested were

normalised to the effect of 1 μM S2 and mean % relaxation values were calculated for both S1 and S2 at each of the concentrations tested.

The relaxing effects of S1 and S2 on tube foot preparations were also compared at a range of concentrations, recording under isotonic conditions as described previously [8]. At concentrations below 0.1 μM, the magnitudes of the relaxing effects observed were too small to be consistently measurable and therefore comparison of the effects of S1 and S2 was limited to tests at 0.1, 1 and 10 μM.

An N-terminally truncated analog of S2 (short S2 or SS2; SFNSGLTFamide) was also tested on both cardiac stomach and tube foot preparations to enable comparison of its bioactivity with that of S1 and S2.

2.3 Circular Dichroism (CD) spectroscopy - CD spectra were recorded using a Chirascan CD spectrometer (Applied Photophysics Ltd.) equipped with a Peltier temperature controller. Spectra were the average of three scans recorded with a 1 nm bandwidth, a 0.5 nm step size, and a 5 s time constant. After background subtraction, the observed ellipticity (θ; mdeg) was converted to a molar ellipticity (ΔE; M^{-1} cm^{-1}), using the formula: $\Delta E = \theta/(33000\ lc)$ where l is the path length (cm) and c is the concentration (M). To investigate the effect of pH on CD spectra, the pH of the peptide solution was altered using 10 mM NaOH or 10 mM HCl to achieve four pH values in the range 5.3 – 8.4. Samples were then incubated for 20 min prior to recording CD spectra. To examine the effect of temperature, CD spectra were obtained at 10, 20, 30, 40 and 50°C. Thermal unfolding of S2 was also investigated up to 90°C using a 1 mm path length for a 0.1 mg/ml sample and a 1 cm path length for a 0.01 mg/ml sample.

2.4 NMR spectroscopy - Peptides were dissolved in 10% D_2O/90% H_2O to achieve a final concentration of 2 mM. The pH of the solutions was adjusted to pH 5.6 using 10 mM NaOH and 10 mM HCl. The peptide solutions were centrifuged to remove any suspended material and then the supernatant was transferred to 5 mm NMR tubes. Data were acquired using a Bruker Avance 600 MHz spectrometer and collected using Topspin software on a UNIX workstation. All experiments were performed using a 5 mm TXI, triple resonance probe equipped with a z-axis gradient. 2D NMR spectra were obtained at 293 K for S1, 303 K for S2 and 283 K for SS2. Water suppression was achieved using a Water Gradient Tailored Excitation (WATERGATE) technique. 2D-Total Correlation Spectroscopy (TOCSY) experiments employed a DIPSI2 sequence for isotropic mixing, with a 75 ms mixing time unless otherwise noted. A 300 ms mixing time was used for Nuclear Overhauser Effect Spectroscopy (NOESY) and Rotational Overhauser Effect Spectroscopy (ROESY) experiments. All 2D experiments used STATES-TPPI phase cycling, and a spectral width of 14 ppm was applied in both dimensions with 2048 x 512 complex data points in the t2 and t1 dimensions, respectively. Prior to Fourier Transformation (FT), the data were linear predicted in the F1 dimension to 512 real points and then zero filled to produce a final matrix size of 4k x 1k, with 90^o phase-shifted sine squared window function applied to both dimensions. The chemical shifts were referenced to water at 4.7 ppm at 303K, pH 5.6, based on the method of Cavanagh et al. [18]. Spin-systems were manually assigned using the TOCSY data and sequential assignments were completed using the 300 ms NOESY spectrum [19]. Chemical shift data for S2 have been deposited in the Biological Magnetic Resonance Data Bank (code 18940).

3.0 Results

3.1 S2 is an order of magnitude more potent than S1- Previous studies using extrapolated dose-response data indicated that S2 is a more potent muscle relaxant than S1 [1]. Here we directly compared the relaxing effects of S1 and S2 at a range of concentrations on starfish cardiac stomach and tube foot preparations. S1 and S2 were tested on cardiac stomach preparations at 1 nM, 0.01 μM, 0.1 μM and 1 μM and the relaxing effect of S2 was found to be larger than the effect of S1 at all concentrations tested (Fig. 1A). Furthermore, the effect of 1 μM S1 (68 ± 8 %) was not significantly different ($p = 0.71$; t-test) to the effect of one-tenth the concentration of S2 (0.1 μM) (66 ± 7 %). Similarly, the effect of 0.1 μM S1 (34 ± 6%) was not significantly different ($p = 0.89$; t-test) to the effect of 0.01 μM S2 (33 ± 8%). S1 and S2 were also tested on tube foot preparations at 0.1 μM, 1.0 μM and 10 μM and, as with cardiac stomach preparations, the relaxing effect of S2 was found to be larger than the effect of S1 at all concentrations tested (Fig. 1B). Moreover, the magnitude of the difference in the relaxing activity of S1 and S2 was greater with tube foot preparations than with cardiac stomach preparations. For example, at 1 μM the mean relaxing effect of S1 was 68% of the mean effect of S2 on cardiac stomach preparations, whereas on tube feet the mean relaxing effect of S1 was only 14% of the effect of S2. These data demonstrate that S2 is at least an order of magnitude more potent than S1 as a muscle relaxant in starfish. This difference in the potency/activity of S1 and S2 as muscle relaxants in starfish provided a rationale for comparative analysis of the solution conformations of these peptides using CD spectroscopy and NMR spectroscopy.

3.2 CD spectroscopy of S1 and S2- CD spectroscopy in the far UV range was used to investigate the secondary structure propensity of S1 and S2 in aqueous solution and the effect of temperature and pH on their main-chain conformations (Fig. 2). The CD spectra obtained for S1 (Fig. 2A, B) have a strong negative band at 200 nm along with a weak positive band at 218 nm. These CD bands suggest that S1 has predominantly an irregular random-coil conformation in aqueous solution [20]. Varying the temperature from 10°C to 50°C (at pH 5.5; Fig. 2A) and the pH from 5.3 to 8.3 (at 11^{o}C; Fig. 2B) had little effect on the appearance of the CD spectra. Both resulted in a slight reduction in the amplitude of the bands at 200 nm and 218 nm but no shift in the wavelength of each of the peaks. The lack of a shift in the wavelengths suggests that S1 remains unstructured over this range of pHs and temperatures.

The CD spectra for S2 were quite different to those obtained for S1, with negative bands at 207 nm and 196 nm and a weak positive band at 228 nm (Fig. 2C, 10°C spectrum). These spectra are different to those observed with peptides that have α-helical or β-sheet type conformations; however, they are also inconsistent with a random coil conformation, which are characterized by a single negative CD band at ~198 nm. It is possible that S2 adopts an extended conformation that does not contain formal β-strands; this is also supported by NMR chemical shift analysis and the absence of main-chain hydrogen bonding (discussed in section 3.5). The negative peak near 198 nm is suggestive of an unstructured peptide, but proline residues are known to have transitions at 198 nm, with the trans Xxx-Pro bond having a negative Cotton effect and the cis Xxx-Pro bond having a positive effect at 198 nm [21, 22]. It is likely that the CD band at 198 nm for S2 is a combination of signals from a structured, proline-containing peptide along with 'random coil' type signals. The slight intensity reduction at 207 nm and 228 nm as the temperature increases (Fig. 2C) could be due to some loss of overall structure.

Increasing the pH from 5.5 to 8.4 reduced the negative band at 196 nm, broadened negative bands at 207 nm and resulted in loss of the positive band at 228 nm (Fig. 2D). Change in pH can affect the Xxx-pro cis-trans equilibrium and it is plausible that the loss of the negative signal at 198 nm is a result of a greater proportion of the cis isoform at high pHs. Above neutrality the peak at 207 nm not only broadens but also moves to longer wavelengths with a concomitant decrease in the intensity of the 228 nm positive peak. Possibly, the increasing pH causes the structure to adopt some β-turns. Indeed, β-turns tend to have a positive signal near 200 nm and a slight negative signal near 220. The appearance of β-turn-like structure might also account for the loss of negative intensity around 198 nm. Overall, and over a range of pH values, it is clear that temperatures up to 30^{o}C do not cause S2 to unfold into an irregular structure. Based on these observations we chose to study both S1 and S2 further using NMR at pH 5.8 and 30 °C.

3.3 ^{1}H NMR comparison of S1 and S2- Two-dimensional NMR spectra were obtained for both S1 and S2 in order to determine if any unique structures existed. From the 2D TOCSY spectra we were able to assign complete spin systems for all but Gly_1 and Phe_2 of S1 (supplementary material Table S1A). The 1D spectrum (Fig. 3) shows considerable line broadening of Phe_2 suggesting that the N-terminus of S1 is exchange broadened, undergoing milli-to-micro second time-scale motions. Furthermore, the limited chemical shift dispersion of the amide protons suggests the S1 peptide is unstructured. The NOESY spectrum for S1 gave a few weak NOE cross-peaks for intra-residue NOEs suggesting that the size of the peptide and the resulting correlation time causes the NOE to be weak and close to the null. For this reason, a ROESY spectrum was obtained from which we were able to obtain sequential assignments as shown in Fig. 3 (H_α-H_N fingerprint region). It was possible to unambiguously assign all peaks present as

S1 contained six unique amino acids and two phenylalanines. However, only sequential connectivities were observed for S1 and no long range ROEs that would be indicative of stable structure were apparent. The fact that the ROESY spectrum of S1 provided no further information beyond a limited number of sequential assignments is entirely consistent with the CD data, which indicated that S1 exists in a random-coil conformation.

From the TOCSY spectrum of S2 (Fig. 4A) we were able to assign all but Ser_1 and, in contrast to S1, we were able to obtain a ^1H-^1H NOESY spectrum that contains a remarkable number of cross peaks (the complete 300 ms NOESY spectrum is shown as supplementary material Fig. S1). Sequential assignments for S2 were readily achieved using standard 2D methods [19] and are shown as supplementary material Table S1B. The H_α-H_N fingerprint region is shown in Fig. 4B, and from this a complete sequential path could be followed from Ser_1 through to the C-terminal amide on Phe_{12}, with only a break at Pro_3. In the spectrum there were two sets of peaks associated with the N-terminal residues Gly_2, Pro_3, Tyr_4 and Ser_5, and one set was more intense than the other (Fig. 4B). In the aliphatic region (see supplementary material Fig. S1), a strong $d_{\alpha\delta}(i, i+1)$ NOE was observed between the Gly_2 and Pro_3 for the most intense set of peaks indicating that these originated from a *trans* conformation around the Gly_2-Pro_3 peptide bond. The lower intensity peaks had a $d_{\alpha\alpha}(i, i+1)$ between Gly_2 and Pro_3 consistent with a *cis* conformation. The large chemical shift differences between residues in the *cis* and *trans* conformations, notably Tyr_4 and Ser_5 observed in Fig. 4B, is likely due to the different conformation of the aromatic side-chain of Tyr_4, and the consequently altered ring-current effects. On intensity measurements, the *trans*-conformation is about 80-85% prevalent. A weak $d_{NN}(i, i+1)$ NOE is apparent between Ser_{5Cis} and Phe_6 (data not shown), which suggests that residues beyond Ser_5 are not influenced by the *cis-trans* isomerism and they have the same

structure in each isoform. This is also supported by the presence of an NOE between the Ser_{5Cis} β-protons and the ring protons of Phe_6 (data not shown).

The H_α-H_N region shows several NOEs indicative of substantial order in S2. For instance, $d_{\alpha N}(i, i+2)$ connections are readily observed for residues from Gly_2 through to Thr_{11} and a longer range $d_{\alpha N}(i, i+3)$ NOE and a weak $d_{\alpha N}(i, i+4)$ are observed from Gly_2 to Ser_5 and Phe_6 respectively. Further, $d_{\beta N}(i, i+1,2,3)$ NOEs are apparent for the β-protons of Ser_5 and Ser_8. Along with the $d_{\alpha N}$ connections, there are also several, relatively weak, $d_{NN}(i, i+2)$ connections for Tyr_4, Ser_5, Phe_6 and Ser_8 (data not shown). Remarkably, in all there were more than 220 NOEs observed for the S2 peptide. We were surprised by this amount of structuring, as indicated by the large number of NOEs, for such a small (12 residue) peptide and were intrigued as to the source of the structuring in the absence of disulphide bonds.

In the absence of disulphide bonds, it is possible that extensive backbone hydrogen bonding hold S2 in a rigid conformation. To investigate this we explored the temperature dependence of amide proton NMR resonances. ^{1}H NMR amide temperature coefficients can be a reliable indicator of the presence of hydrogen bonding. In particular, temperature coefficients of between -1 ppb/K and -4 ppb/K indicate a high probability (93%) of the presence of a hydrogen-bonded amide [23, 24]. 1D spectra for S2 were obtained over a range of temperatures and the change in amide proton chemical shift as a function of temperature (283 K – 313 K) was determined. Table 1 shows that for S2 the smallest temperature coefficient observed is |-4.6| ppb/K for Leu_{10}, a value which can predict a hydrogen bond in more than 85% of cases. The temperature coefficients for the other amides in S2 suggest hydrogen bonding is unlikely to occur. Experiments with S1 did not show any temperature coefficient small enough to predict any hydrogen bonding. Furthermore, $^3J_{H\alpha NH}$ coupling constants for S2 are not indicative of any

regular secondary structure. So with no extensive hydrogen bonding network, why is there so much structuring of S2, as apparent from the large number of NOE contacts observed and also evident in the CD spectra?

3.4 Structuring of S2 is caused by self-association- Next we investigated the concentration dependence of the CD spectrum for the S2 peptide. It is clear from Fig. 5A that the structuring of the S2 peptide is highly concentration dependant. An increase in concentration from 0.01 mg/ml to 0.1 and 0.2 mg/ml (the concentration at which the NMR data was obtained) causes a profound increase in the structuring of the S2 peptide, as indicated by a strong negative CD band at 216 nm, typical of an extended-conformation. With the strong concentration dependence of the CD spectra it is clear that self-association is the source of the observed structuring of S2. At the lower concentration (0.01 mg/ml) there is little structuring of S2 with a single weak negative CD band at 198 nm, typical of irregular structure.

We went on and studied the thermal stability of the S2 peptide, at two different concentrations. At 0.1 mg/ml S2 undergoes a clear thermal unfolding transition. The unfolding curve has a sigmoidal appearance typical for cooperative, essentially two-state unfolding, with a mid-point at 55°C (Fig. 5B). The temperature dependence of the CD spectra for S2 at 0.01 mg/ml is very different; there is no cooperative unfolding with temperature. The increase in temperature causes a loss of some PP-II character in the spectra, but a co-operative unfolding transition is not apparent for the dilute S2 sample.

Given that the CD data suggests that S2 self-associates at the concentrations used here for NMR, NOESY peaks other than the intense sequential NOEs may arise from intra- or inter-peptide contacts. For instance, the β-protons of Ser_5 have clear NOEs to the amide protons of

Phe_6, Asn_7 and Ser_8 along with a reasonably intense NOE to the amide proton of Tyr_4 (Fig. 4B). Although the gradually weakening NOEs to residues 6, 7, and 8 could be expected to be intra-peptide, the NOE to Tyr_4 may arise from an intra-peptide contact or from an inter-peptide contact. Although there are a large number of NOEs (220 pairs), and particularly because of the small size of the peptide, determining which of these NOEs are intra-molecular and which are inter-molecular is difficult. Thus, standard approaches to generate a monomeric solution structure from the NOE data set are not appropriate. Nevertheless, the large number of NOEs and the ability to observe a well resolved set of resonances for residues Gly_2 to Phe_{12} indicates that self-association induces a single, soluble, low molecular mass structure in S2. The narrow line-widths of the amide NMR peaks (~4 Hz, at half-height) suggest that the size of the complex is small, perhaps only dimeric. Obtaining NMR spectra at lower concentrations in order to differentiate inter-molecular NOEs from intra-molecular ones is problematic as disrupting the oligomer also causes loss of structure. Therefore, we cannot assume that loss of a particular NOE necessarily indicates it is a inter-molecular contact as loss of structure means only sequential NOEs are observed, as seen for S1 (Fig. 3).

3.5 Evidence of a role for the N-terminal region of S2 in self-association- The clear difference between S1 and S2 in their solution conformations is intriguing given their sequence similarity. However, S2 has the N-terminal sequence SGPY that is absent in S1. To determine if this sequence impacted on the self-association property of S2, we constructed a modified S2 peptide that lacks this N-terminal sequence (short S2 or SS2; SFNSGLTFamide). The TOCSY and ROESY spectra of this peptide, at the same concentration as S2, are shown overlaid in Fig. 6A. The dispersion of amide peaks in the 1D spectrum is similar to that of S1 and suggests a lack of

regular structure. Moreover, the array of NOEs apparent in S2 (Fig. 4B) is missing in SS2 and only sequential NOEs are visible (Fig 6A). The lack of a large number of NOEs in the spectrum of SS2 suggests that the N-terminal sequence SGPY is an important factor in the self-association of S2 and subsequent generation of a defined structure. Aromatic amino acids can often associate via π-π interactions and it is plausible that Tyr_4 is required to support such interactions and therefore S2 self-association. Intriguingly, analysis of the H_α chemical shift index (i.e. deviation from 'random coil') of S2 and SS2 shows that they are reasonably similar (Fig. 6B), but dissimilar to S1 (Fig. 6C). Note that although trending negative from Phe_6, the CSI shifts for S2 and SS2 are not great enough to suggest any true β-sheet-like structure exists. Nevertheless, the similarity between S2 and SS2 shifts suggests that SS2 may be accessing the same conformational space as is S2 despite not forming a defined structure.

3.6 Analysis of the bioactivity of SS2- Having analysed the solution conformation of SS2 in comparison with S2, it was of interest to compare the bioactivity of SS2 with that of S1 and S2. Therefore, we tested SS2 on the two preparations that we originally used (Fig. 1) to compare the pharmacological activity of S1 and S2 – starfish cardiac stomach and tube foot preparations.

Consistent with relaxing effects of S1 and S2, SS2 caused dose-dependent relaxation of cardiac stomach preparations when tested at 0.1 μM, 1 μM and 10 μM (Fig. 7A). Comparison of the relaxing actions of S1, SS2 and S2 when tested at 1 μM revealed that SS2 was significantly more effective than S1 ($p = 0.0004$; t-test) and slightly, but significantly ($p = 0.03$; t-test), less effective than S2 (Fig. 7B). These findings indicate that the biological activity of S2 is largely attributable to its C-terminal octapeptide sequence (SFNSGLTFamide).

Next an analysis of the effects of SS2 on tube foot preparations yielded some interesting findings. As with the cardiac stomach preparations, SS2 caused dose dependent relaxation of tube-feet (Fig. 7C). However, it was noteworthy that the effects of SS2 at 10 μM were consistently larger than the effects of S2 at the same concentration. When the effects of S1, S2 and SS2 were compared at 1 μM, SS2 was significantly ($p = 0.00006$; t-test) more effective than S1 and significantly ($p = 0.000003$; t-test) less effective than S2 (Fig. 7D). Thus, for tube foot preparations conflicting conclusions can be drawn from the data obtained, with results from tests at 10 μM indicating that the presence of the N-terminal SGPY sequence actually impairs the bioactivity of S2, while results from tests at 1 μM indicate that the presence of the N-terminal SGPY sequence contributes to the bioactivity of S2.

4.0 Discussion

S2 is related to the SALMFamide neuropeptide S1 (GFNSALMFamide), which shares with S2 the C-terminal sequence motif FNSxLxFamide (where is x is variable). Both S1 and S2 cause muscle relaxation in starfish but S2 is more effective than S1 when tested at the same concentration [8]. Here we have specifically compared the potency of S1 and S2 by testing both peptides at a range of concentrations on two different preparations from the starfish *Asterias rubens*, the cardiac stomach and tube feet. These experiments revealed that S2 is at least an order of magnitude more potent than S1, which provided the rationale for a comparative investigation of the solution structures of S1 and S2 using CD and NMR spectroscopy.

CD spectroscopy indicates that S1 is largely unstructured in aqueous solution and, consistent with the CD data, the NMR NOESY spectrum of S1 showed only sequential NOEs, which implies a lack of a unique, stable conformation. This is in line with other neuropeptide studies showing that small peptides like S1 are usually unstructured in aqueous solutions [9-16]. In contrast, the NMR NOESY spectrum for S2 showed a remarkable array of NOE cross-peaks (220 NOE pairs) indicative of a well-structured peptide. The H_α chemical shift analysis, the lack of regular hydrogen bonding and the pattern of NOEs, all indicate that S2 is not composed of a regular α-helical or β-sheet like structure. However, the sequence of S2 has polar, hydrophilic amino acids interspersed with hydrophobic and aromatic amino acids, which may support the generation of some structure. Keire *et al*. [9] have shown that an eight residue cholecystokinin peptide CCK8 (DYMGWMDF-NH_2) in aqueous solution has a preference for γ-and β-turns around GWMD and MDF-NH_2. In the case of S2, the structure adopted is dependent on the peptide concentration. At lower concentrations S2 is unstructured, and presumably monomeric, whereas increasing the concentration induces structure, which is likely due to the formation of

low molecular weight oligomers, perhaps dimers. However, because of the small size of the peptide it is not possible to ascertain, from the numerous NOEs observed, the type of contacts between S2 peptides (i.e. head-to-head or head-to-tail).

Many neuropeptides that are unstructured and flexible in aqueous solvents tend to adopt structure when in the presence of stabilising solvents such as alcohols (trifluoroethanol, for instance), as well as in the presence of membranes. This finding has led to the proposition of a two-step theory of neuropeptide/receptor interaction wherein the neuropeptide binds first to the membrane and then undergoes diffusion to the appropriate receptor [25]. It is envisaged that interaction with the membrane not only reduces the degrees of freedom but also stabilises conformations that promote receptor binding.

Our discovery that S2 has a well defined, but concentration dependent conformation in aqueous solution may have physiological implications. The concentration of neuropeptides in vesicles prior to release is high and structuring may support a higher degree of packing in vesicles. Furthermore, high peptide concentrations may persist locally after release and therefore analysis of the self-association structure and process may be relevant to function. Indeed, studies with neuropeptide Y, whose physiological roles include regulation of food intake and pain perception, have shown that the hydrophobic residues that support self-association in this peptide also allow binding to cell membranes [26].

The conformation of S2 in aqueous solution may explain why S2 is more potent than S1 as a muscle relaxant in starfish. We speculate that S2 might bind to its receptor as an oligomer (dimer) and that the solution conformation of S2 is the same or very similar to its conformation when bound to a receptor. Therefore the probability of receptor-binding and receptor-activation may be greater than for a peptide (S1) that lacks a well-defined conformation in aqueous solution

and that will have an entropic penalty when binding to its receptor. Earlier studies on FMRFamide-like peptides have shown that, although not adopting a unique conformation, some can contain significant populations of turns and the percentage of turns is highly correlated with IC_{50} values of receptor binding [27]. This suggests that the presence of structure in neuropeptides confers an added level of 'fine-tuning' of receptor interactions.

Clearly the key structural difference between S1 and S2 is the presence of four N-terminal residues (SGPY) in S2 that are lacking in S1. It appears, therefore, that the presence of these residues confers on S2 an ability to adopt a stable conformation in solution upon self-association. Our finding that S2 and SS2 (short S2) have similar H_α chemical shift indexes suggests that SS2 may explore a conformation similar to that adopted by S2 at the same concentration, but a stable structure is not generated. Accordingly, differences in the C-terminal regions of S1 and S2 clearly contribute to the differing potencies of S1 and S2 because tests with SS2 revealed that it is significantly more active as a muscle relaxant than S1. It is the N-terminal SGPY motif in S2 that appears to be the key to its conformational stability in aqueous solution and this motif contributes to superior bioactivity of S2 compared to both SS2 and S1 when tested at 1 µM and 10 µM on cardiac stomach preparations and when tested at 1 µM on tube preparations. However, our finding that SS2 is more active as a muscle relaxant than S2 when tested on tube feet at 10 µM suggests complex structure-activity relationships, which may be influenced by the combinations of receptor types that are expressed in different organs. Further investigation of this issue will be possible if the receptors that mediate the effects of SALMFamides are identified (see [28] for a discussion of a potential strategy for identification of SALMFamide receptors). If there are multiple receptor types that mediate the effects of SALMFamides, then tissue-specific differential expression of receptor types may explain why

dose-dependent differences in the relative activity S1, S2 and a synthetic analog of S2 (SS2) are observed.

From a physiological perspective it is of interest to consider why two peptides (S1 and S2) with the same effect (muscle relaxation) but with different potencies both exist in starfish. Immunocytochemical studies have revealed that S1 and S2 are expressed by different populations of neurons in the starfish *Asterias rubens* and this finding is supported by analysis of the genome sequence of the starfish *Patiria miniata*, which has revealed that S1 and an S2-like peptide are derived from precursor proteins encoded by different genes [7]. Moreover, in *Asterias rubens* S2 is detected by radioimmunoassay in the perivisceral coelomic fluid whereas S1 is not [17]. This suggests that S2 may act as a hormonal signalling molecule in starfish and therefore it would be necessary for its potency to be greater than that of S1, which may only act more locally as a neurotransmitter or paracrine signalling molecule. Thus, we speculate that the structural features of S2 that confer on it greater potency compared to S1 may be molecular adaptations for a hormonal signalling role.

Finally, it has recently been discovered that in the starfish *Patiria miniata*, S1 is derived from a precursor protein that comprises six other putative L-type or L-type-like SALMFamides [7]. On the other hand, the S2-like neuropeptide in this species is derived from a different precursor protein that comprises eight F-type or F-type-like SALMFamides [7]. Our investigation of a structural basis for differences in the activity of S1 and S2, as reported here, therefore provides the foundations for further studies on more complex "cocktails" of SALMFamides that occur *in vivo*.

Acknowledgements - We are grateful to Harold Toms (QMUL) for technical assistance with the NMR studies, Richard Melarange (QMUL) for technical assistance with *in vitro* pharmacology and Urooj Khan (UWS) for analysis of some of the NMR data.

REFERENCES

[1] M.R. Elphick, R. Melarange, Neural control of muscle relaxation in echinoderms, J Exp Biol, 204 (2001) 875-885.
[2] M.L. Rowe, M.R. Elphick, Discovery of a second SALMFamide gene in the sea urchin *Strongylocentrotus purpuratus* reveals that L-type and F-type SALMFamide neuropeptides coexist in an echinoderm species, Mar Genomics, 3 (2010) 91-97.
[3] M.R. Elphick, D.A. Price, T.D. Lee, M.C. Thorndyke, The SALMFamides: a new family of neuropeptides isolated from an echinoderm, Proc Biol Sci, 243 (1991) 121-127.
[4] R. Melarange, D.J. Potton, M.C. Thorndyke, M.R. Elphick, SALMFamide neuropeptides cause relaxation and eversion of the cardiac stomach in starfish., Proc Biol Sci, 266 (1999) 1785-1789.
[5] S.J. Newman, M.R. Elphick, M.C. Thorndyke, Tissue distribution of the SALMFamide neuropeptides S1 and S2 in the starfish *Asterias rubens* using novel monoclonal and polyclonal antibodies. 2. Digestive system, Proc Biol Sci, 261 (1995) 187-192.
[6] S.J. Newman, M.R. Elphick, M.C. Thorndyke, Tissue distribution of the SALMFamide neuropeptides S1 and S2 in the starfish *Asterias rubens* using novel monoclonal and polyclonal antibodies. I. Nervous and locomotory systems, Proc Biol Sci, 261 (1995) 139-145.
[7] M.R. Elphick, S. Achhala, N. Martynyuk, The evolution and diversity of SALMFamide neuropeptides, PLoS One, 8 (2013) e59076.
[8] R. Melarange, M.R. Elphick, Comparative analysis of nitric oxide and SALMFamide neuropeptides as general muscle relaxants in starfish, J Exp Biol, 206 (2003) 893-899.
[9] D.A. Keire, T.E. Solomon, J.R. Reeve, Jr., NMR evidence for different conformations of the bioactive region of rat CCK-8 and CCK-58, Biochem Biophys Res Commun, 293 (2002) 1014-1020.
[10] G.Y. Xu, C.M. Deber, Conformations of neurotensin in solution and in membrane environments studied by 2-D NMR spectroscopy, Int J Pept Protein Res, 37 (1991) 528-535.
[11] J.L. Nieto, M. Rico, J. Santoro, J. Herranz, F.J. Bermejo, Assignment and conformation of neurotensin in aqueous solution by 1H NMR, Int J Pept Protein Res, 28 (1986) 315-323.
[12] S.C. Sumner, K.S. Gallagher, D.G. Davis, D.G. Covell, R.L. Jernigan, J.A. Ferretti, Conformational analysis of the tachykinins in solution: substance P and physalaemin, J Biomol Struct Dyn, 8 (1990) 687-707.
[13] G. Chassaing, O. Convert, S. Lavielle, Preferential conformation of substance P in solution, Eur J Biochem, 154 (1986) 77-85.
[14] S.M. Cowsik, C. Lucke, H. Ruterjans, Lipid-induced conformation of substance P, J Biomol Struct Dyn, 15 (1997) 27-36.
[15] R.C. Grace, A.M. Lynn, S.M. Cowsik, Lipid induced conformation of the tachykinin peptide Kassinin, J Biomol Struct Dyn, 18 (2001) 611-621, 623-615.
[16] R.C. Grace, I.R. Chandrashekar, S.M. Cowsik, Solution structure of the tachykinin peptide eledoisin, Biophys J, 84 (2003) 655-664.
[17] M.R. Elphick, S.J. Newman, M.C. Thorndyke, Distribution and action of SALMFamide neuropeptides in the starfish Asterias rubens, J Exp Biol, 198 (1995) 2519-2525.
[18] F.W. Cavangh J, Palmer AG, Skelton NJ, Protein NMR Spectroscopy: Principles and Practice, Academic Press Inc., San Diego, Ca, USA, 1996.
[19] K. Wuthrich, NMR of proteins and nucleic acids, Wiley, New York, USA, 1986.

[20] N.J. Greenfield, Applications of circular dichroism in protein and peptide analysis, Trac-Trend Anal Chem, 18 (1999) 236-244.
[21] J.D. Fontenot, N. Tjandra, D. Bu, C. Ho, R.C. Montelaro, O.J. Finn, Biophysical Characterization of One-Tandem, 2-Tandem, and 3-Tandem Repeats of Human Mucin (Muc-1) Protein Core, Cancer Res, 53 (1993) 5386-5394.
[22] K. Ma, L.S. Kan, K. Wang, Polyproline II helix is a key structural motif of the elastic PEVK segment of titin, Biochemistry-Us, 40 (2001) 3427-3438.
[23] T. Cierpicki, J. Otlewski, Amide proton temperature coefficients as hydrogen bond indicators in proteins, J Biomol Nmr, 21 (2001) 249-261.
[24] J. Feeney, The conformation of hormonal peptides, in: G. Roberts (Ed.) Drug action at the molecular level [Proceedings of a Symposium, 1976], Univ. Park Press, Baltimore, 1977, pp. 55-76.
[25] H. Inooka, T. Ohtaki, O. Kitahara, T. Ikegami, S. Endo, C. Kitada, K. Ogi, H. Onda, M. Fujino, M. Shirakawa, Conformation of a peptide ligand bound to its G-protein coupled receptor, Nat Struct Biol, 8 (2001) 161-165.
[26] R. Bader, A. Bettio, A.G. Beck-Sickinger, O. Zerbe, Structure and dynamics of micelle-bound neuropeptide Y: comparison with unligated NPY and implications for receptor selection, Journal of molecular biology, 305 (2001) 307-329.
[27] A.S. Edison, E. Espinoza, C. Zachariah, Conformational ensembles: the role of neuropeptide structures in receptor binding, J Neurosci, 19 (1999) 6318-6326.
[28] M.R. Elphick, From gonadotropin-inhibitory hormone to SIFamides: Are echinoderm SALMFamides the "missing link" in a bilaterian family of neuropeptides that regulate reproductive processes?, Gen Comp Endocrinol, (2013).

FIGURE LEGENDS

Fig. 1. Comparison of the potency of S1 and S2 as muscle relaxants. A, Graph comparing the relaxing effects of S1 (filled bar) and S2 (empty bar) (0.001 – 1 μM) on starfish cardiac stomach preparations, expressed as a mean percentage (± SEM; n = 15) of the effect of 1 μM S2. The effect of 1 μM S1 is not significantly different (p = 0.71; t-test) to the effect of 0.1 μM S2 and the effect of 0.1 μM S1 is not significantly different (p = 0.89; t-test) to the effect of 0.01 μM S2. B, Graph comparing the relaxing effects of S1 and S2 (0.1 – 10 μM) on starfish tube foot preparations, expressed as a mean percentage (± SEM; n = 4 for S1; n = 18 for S2) of the effect of 10 μM S2.

Fig. 2. CD spectra of S1 and S2 peptides in water. A, CD of S1 as the temperature changes at constant pH (5.3). Varying the temperature from 10°C to 50°C had little effect on the CD spectra, with just a slight reduction in the amplitude of the CD signal. B, CD of S1 as the pH changes from pH 5.3 to pH 8.3 at constant temperature (11°C). Similarly, there is a slight reduction in the amplitude of the CD signal as the pH increases. In both A and B there is a strong negative band at 200 nm characteristic of a random coil conformation. C. CD of S2 at different temperatures and constant pH (5.5). Increasing the temperature from 10°C to 50°C results in a slight reduction in the amplitude of the negative bands at 196 nm and 207 nm. D, Increasing the pH (at a constant temperature of 11°C) from 5.5 to 8.4 reduces the negative band at 196 nm and causes the negative band at 207 nm to shift to longer wavelengths.

Fig. 3. ROESY spectrum of S1. The H_α-H_N fingerprint region of the S1 neuropeptide ROESY (τ_m 300 ms) spectrum at 293K, pH 5.6, showing the $d_{\alpha N}(i, i+1)$ assignments of S1. The box

around F2 denotes the H_α-H_N peak that is observed in the TOCSY spectrum but not in the ROESY spectrum.

Fig. 4. NMR data for S2 highlights the presence of structure**.** A, The H_α-H_N finger print region of the S2 TOCSY (τ_m 75 ms) spectrum in water at 303K, pH 5.6 showing spin-system assignment. B, The H_α-H_N fingerprint region of the S2 NOESY (τ_m 300 ms) spectrum in water at 303K, pH 5.6. Solid lines show the $d_{\alpha N}(i, i+1)$ sequential assignments of S2. Dashed lines identify residues from the *cis*-Pro_3 isoform.

Fig. 5: Concentration dependence of S2 peptide structure and folding stability. Far UV-CD spectra of S2, showing concentration dependent secondary structure of S2 at 0.01 mg/ml (7.8 μM) (red), 0.1 mg/ml (78 μM) (black) and 0.2 mg/ml (156 μM) (blue) (A). Thermal denaturation of S2, between 10-90^{o}C, blue spectra are from 10^{o}C and red are at 80^{o}C. Spectra were obtained at two concentrations; 0.1 mg/ml (B), and 0.01 mg/ml of S2 (C). Inserts are thermal denaturation curves at 210 nm and 228 nm, respectively.

Fig. 6. NMR analysis of SS2 (short S2). A, the H_α-H_N fingerprint region of the TOCSY (blue, τ_m 65 ms) and ROESY (red, (τ_m 300 ms) spectra of SS2 shown overlaid. The sequential $d_{\alpha N}(i, i+1)$ assignments are shown with solid lines. B, H_α chemical shift analysis of S2 (black bars) and SS2 (white bars). C, H_α chemical shift analysis of S1.

Fig. 7. Analysis of the effects of SS2 as a muscle relaxant and comparison of the activity of SS2 with S1 and S2. A, SS2 causes dose-dependent relaxation of starfish cardiac stomach

preparations; error bars represent mean ± SEM (n=5) and data are normalised to the effect of S2 at 10 μM. B, Comparison of the relaxing effect of SS2 to that of S2 and S1 on cardiac stomach (all at 1 μM). Data are normalised to relaxation effect of S2 (at 1 μM); error bars represent mean ± SEM (n=12). Significance *** = $p < 0.001$ (students t-test). C, SS2 causes dose-dependent relaxation of starfish tube foot preparations; error bars represent mean ± SEM (n=9) and data are normalised to the effect of S2 at 10 μM. D, Comparison of the relaxing effect of SS2 to that of S2 and S1 on tube feet. Data are normalised to relaxation effect of S2 (1 μM); error bars represent mean ± SEM (n=13). Significance *** = $p < 0.001$, * = $p<0.05$.

Table 1

Residue	$\Delta \delta_{NH}$ (ppb)/ΔK
S_1	-
G_2	-8.7
P_3	-
Y_4	-7.6
S_5	-5.4
F_6	-6.6
N_7	-4.9
S_8	-5.7
G_9	-5.0
L_{10}	-4.6
T_{11}	-5.3
F_{12}	-6.7

Figure 1

Figure 1.

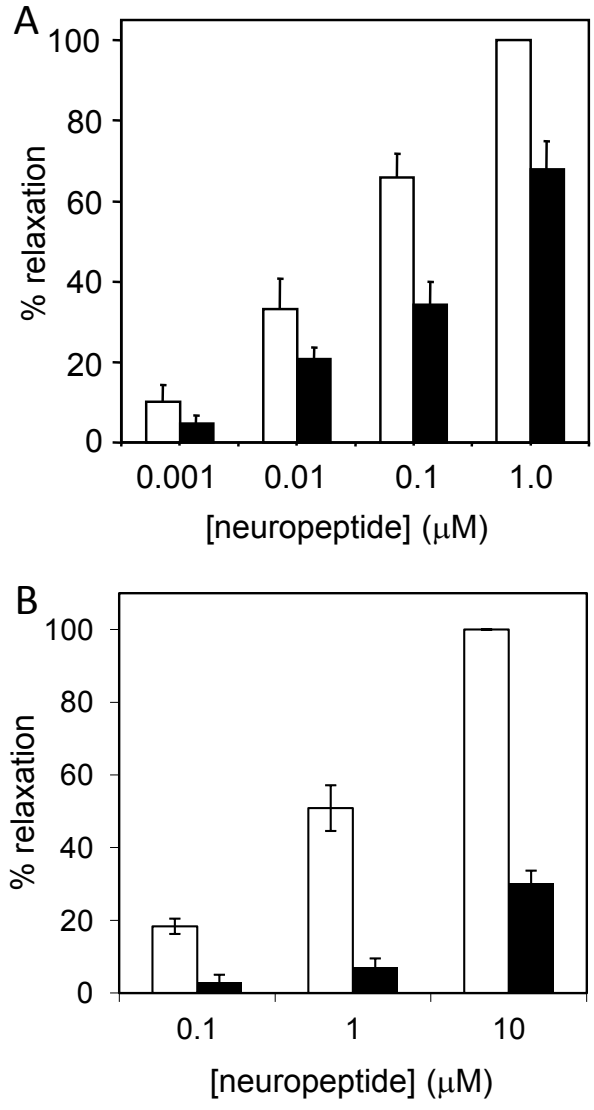


Figure 2

Figure 2

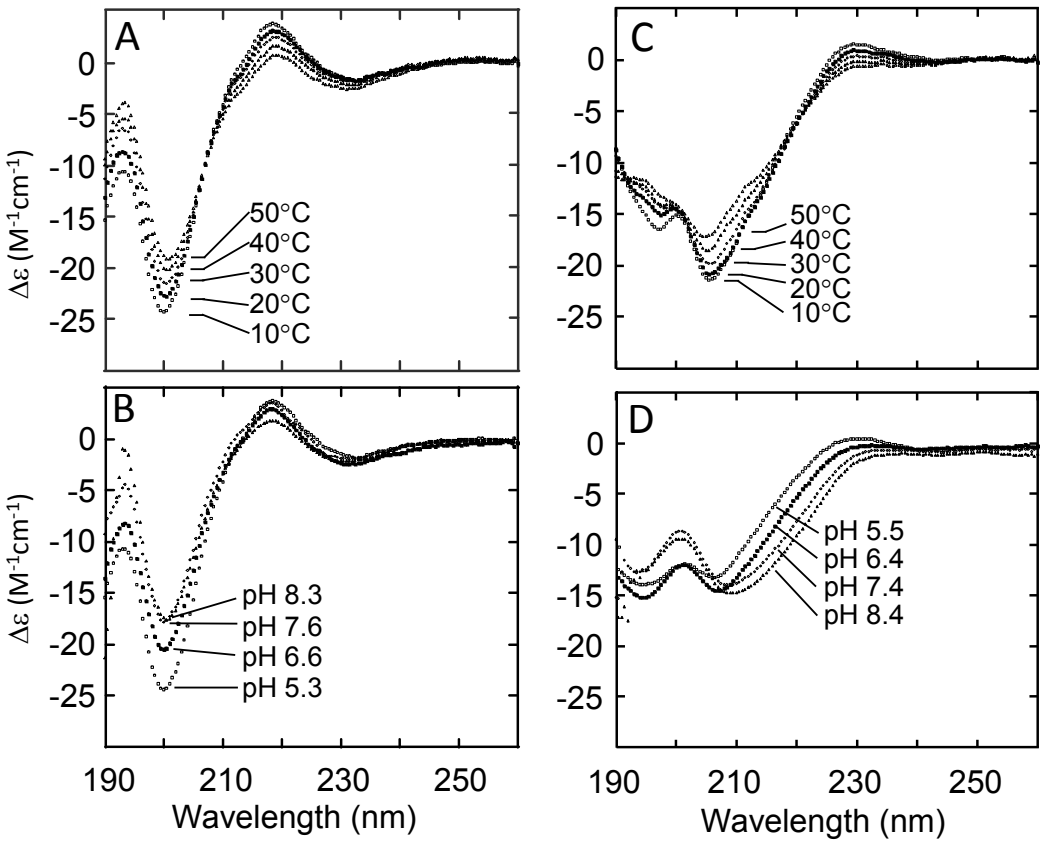


Figure 3

Figure 3

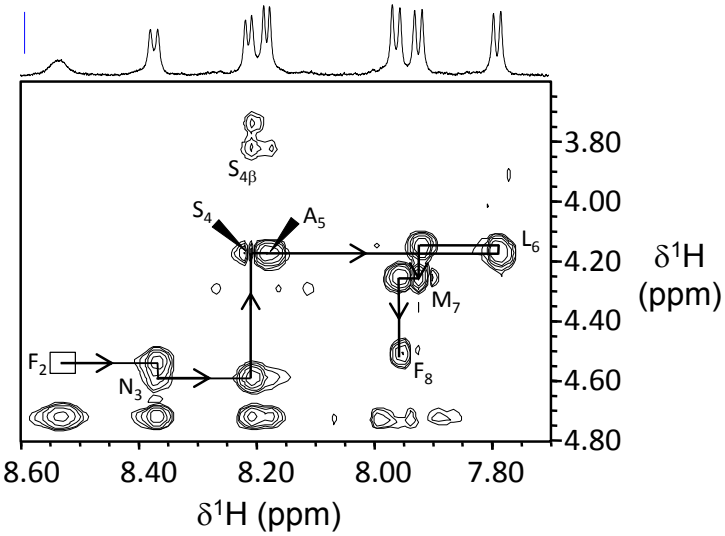


Figure 4

Figure 4

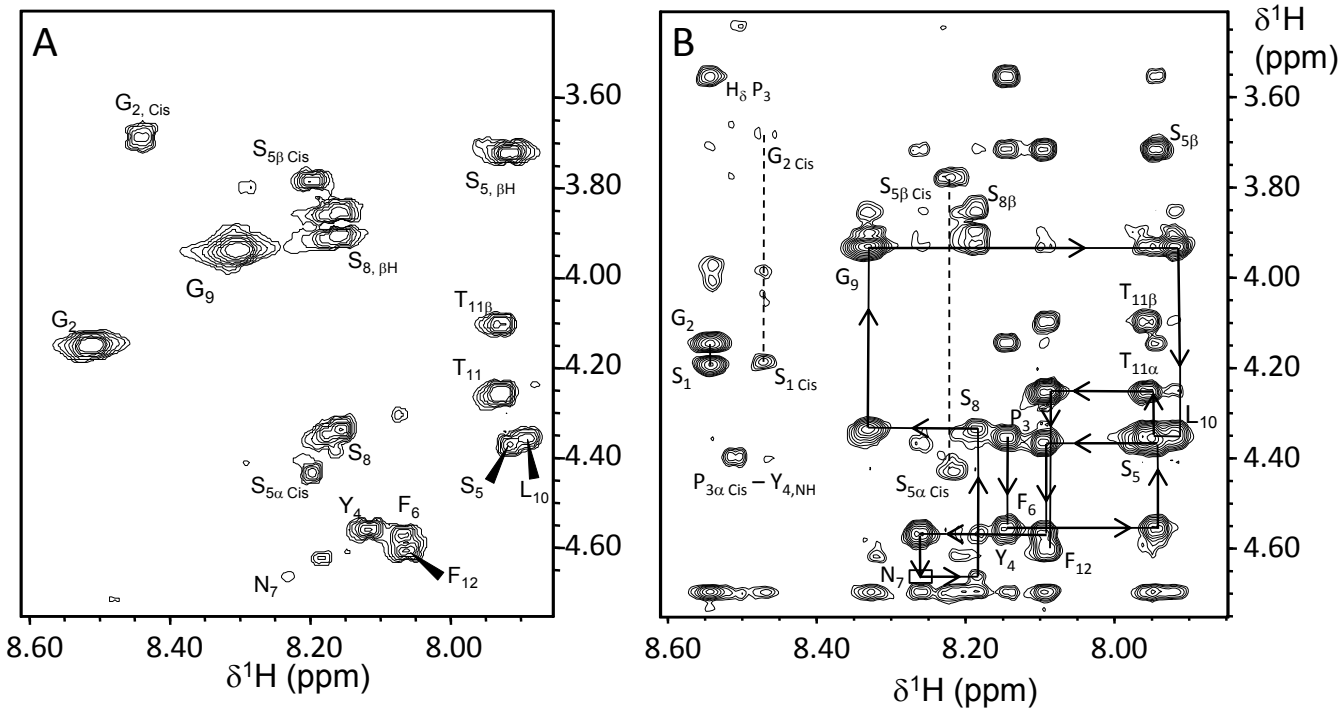


Figure 5

Figure 5

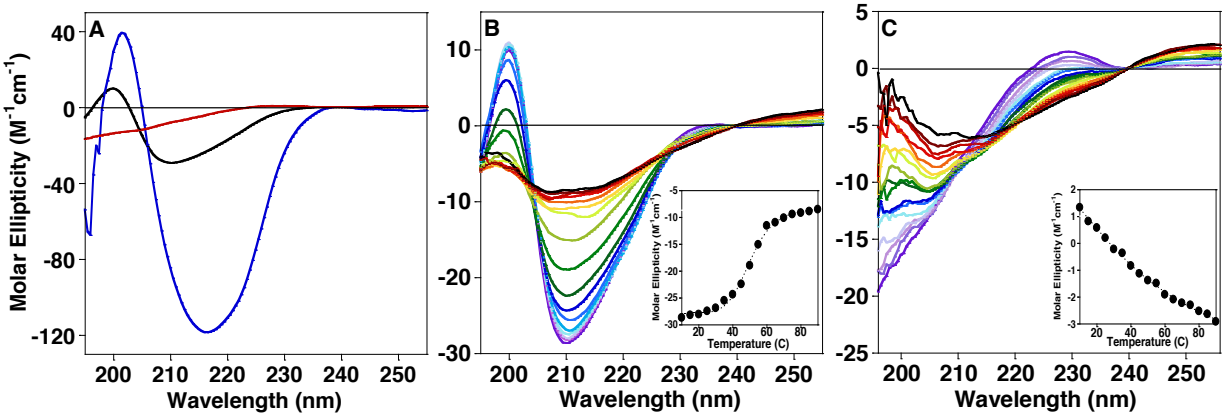


Figure 6

Figure 6

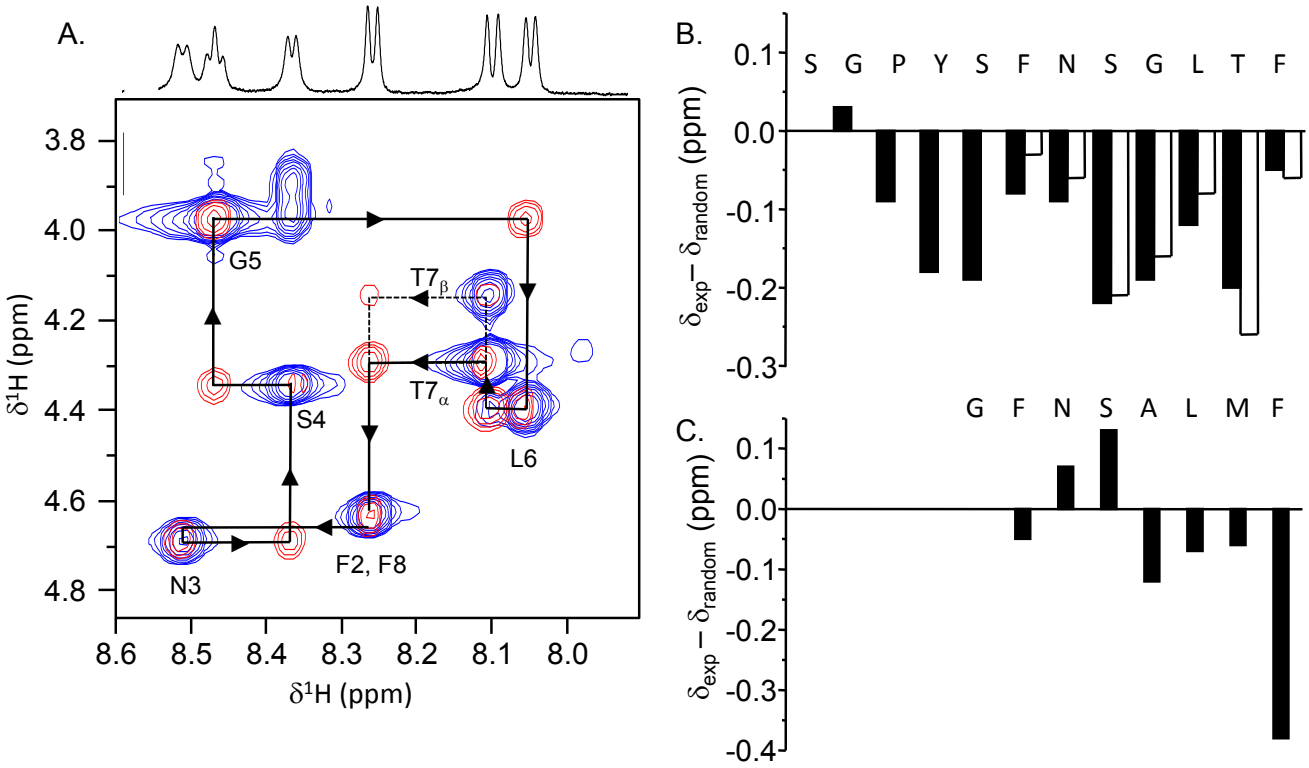
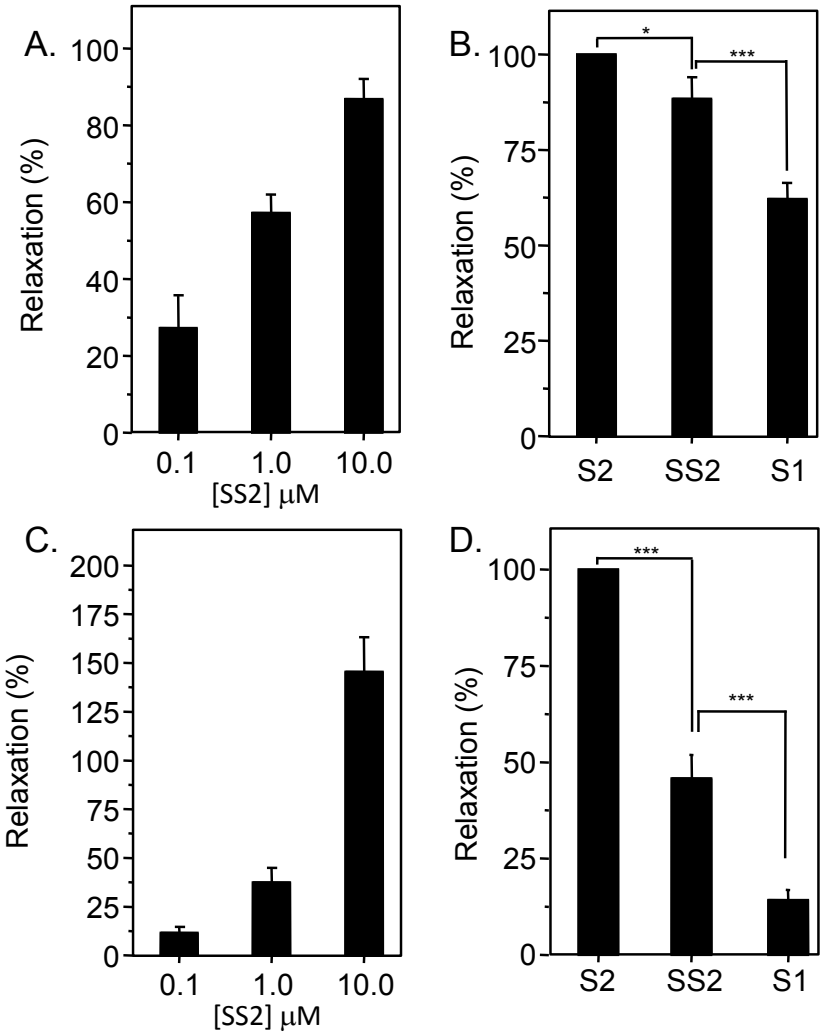


Figure 7

Figure 7



Supplementary Material (for online publication)

Click here to download Supplementary Material (for online publication): Otara et al. (supplementary material).pptx

A.

Residue	Chemical shift (δ) ppm			
	NH	αH	βH	others
Gly1	-	-		
Phe2	8.53	4.54	-	7.15 (Hδ), 7.22 (Hε) *
Asn3	8.37	4.59	2.98, 2.92	7.48 , 6.82
Ser4	8.21	4.16	2.68, 2.57	
Ala5	8.18	4.16	1.26	
Leu6	7.79	4.13	1.43	0.80, 0.73
Met7	7.92	4.24	1.78	2.30 (Hγ)
Phe8	7.96	4.51	3.82, 3.73	7.15 (Hδ), 7.22 (Hε)

* Ambiguous. Overlap with Phe8

B.

Residue		Chemical shift (δ) ppm			
		NH	αH	βH	others
Ser1	trans	-	4.19	4.00, 3.98	
	cis	-	4.19	4.04, 3.99	
Gly2	trans	8.54	4.15		
	cis	8.47	3.68		
Pro3	trans	-	4.35	1.94, 2.15	1.79, 1.72 (Hγ); 3.56(Hδ)
	cis	-	4.40	2.30, 1.96	1.83, 1.65 (Hγ); 3.45(Hδ)
Tyr4	trans	8.15	4.55	2.90, 2.98	6.81, 7.08
	cis	8.50	-	2.73, 3.03*	6.77 ,7.12
Ser5	trans	7.95	4.36	3.72	
	cis	8.22	4.31	3.78	
Phe6		8.10	4.57	3.10, 3.36	7.34, 7.25, 7.28
Asn7		8.26	4.66	2.77, 2.67	7.50, 6.83
Ser8		8.19	4.33	3.90, 3.86	
Gly9		8.33	3.93		
Leu10		7.92	4.35	1.58, 1.54	1.49, 0.89, 0.84
Thr11		7.96	4.35	4.10	1.08
Phe12		8.09	4.60	3.15, 2.98	7.33, 7.25, 7.28
$-NH_2$					7.46, 7.02

Table S1. Chemical shifts for (A) S1 and (B) S2. Chemical shifts for the S2trans isoform are also available from the BMRB (Accession code 18940)

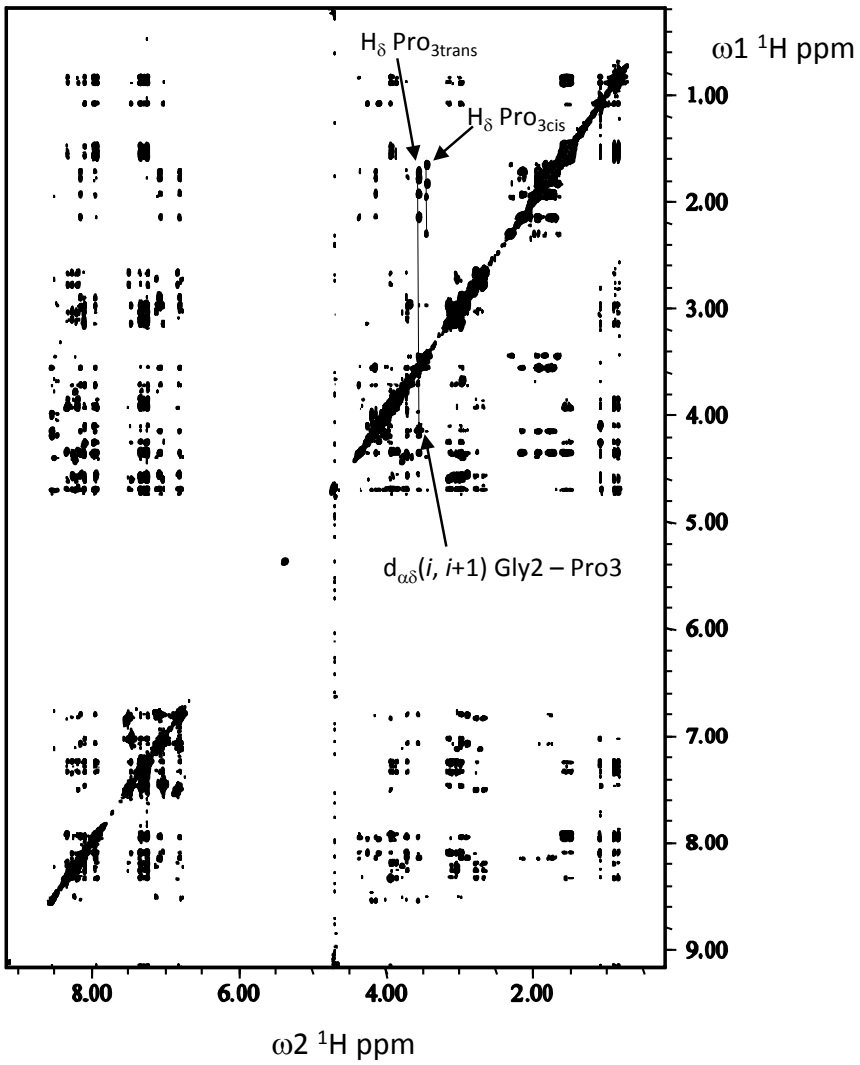


Figure S1. The complete S2 600 MHz NOESY (t_m = 300 ms) spectrum collected in water at 303K, pH 5.6. The NOEs between the H_δ, H_γ and H_β protons of both the *cis* and *trans* isoforms of Pro_3 are highlighted. The strong $d_{\alpha\delta}(i, i+1)$ crosspeak between Gly_2 and Pro_3 in the trans isoform is also highlighted.