Solution NMR studies reveals the location of the second transmembrane domain of the human sigma-1 receptor.

Author Names:
Jose Luis Ortega-Roldan¹,*, Felipe Ossa¹, Nader T. Amin¹, and Jason R. Schnell¹,*

Author Addresses:
¹Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, U.K.

*Corresponding Authors

Corresponding E-mail Addresses:
Jose.Ortega-Roldan@bioch.ox.ac.uk
Jason.Schnell@bioch.ox.ac.uk

Abstract

The sigma-1 receptor (S1R) is a ligand-regulated, membrane chaperone protein associated with stress responses of the endoplasmic reticulum and modulation of ion channel activities at the plasma membrane. We report here solution NMR studies of a S1R construct (S1R(Δ35)) in which only the first transmembrane domain and the eight-residue N-terminus have been removed. The second transmembrane helix is found to be composed of residues 91-107, which corresponds closely to the first steroid binding domain-like region. The cytosolic domain is found to contain three helices, and the secondary structure and backbone dynamics of the chaperone domain are consistent with that determined previously for the chaperone domain alone. The position of TM2 provides a framework for ongoing studies of S1R ligand binding and oligomerisation.
Highlights:

- The second transmembrane domain (TM2) of S1R consists of residues 91-107.
- The S1R cytosolic region contains three α-helices.
- The chaperone domain structure is largely unperturbed by the addition of TM2.
Introduction

The sigma-1 receptor (S1R) is a membrane chaperone protein present varyingly in both the endoplasmic reticulum (ER) and plasma membranes, where it functions as an accessory protein to a number of ion channels and receptors [1-3]. S1R has been observed to modulate the activity of several ion channels including IP3 receptors [2, 4] and voltage-gated channels selective for potassium [5, 6], sodium [7], and calcium [8]. S1R has also been shown to interact with acid-sensing ion channels [9], glutamate receptors [10], and dopamine receptors [11]. S1R is highly expressed in the central nervous system (CNS), primarily in the cerebral cortex, hippocampus and cerebellar Purkinje cells [12, 13], and binds a large number of small molecules (opiates, antipsychotics, antidepressants, antihistamines, phencyclidine-like compounds, β-adrenergic receptor ligands, cocaine, dimethyltryptamine, progesterone, and sphingosine), many of which have been shown to modulate the effect of S1R on receptors and ion channels (reviewed in [14, 15]). Thus, S1R is a potential therapeutic target in the treatment of a range of diseases of the CNS, including schizophrenia, Alzheimer’s and Parkinson’s diseases, amnesia, depression, amyotrophic lateral sclerosis and addiction.

S1R contains two transmembrane domains connected by a cytosolic domain, and an ER-accessible C-terminal chaperone domain [1, 6]. Based on homology to the steroid binding regions of fungal sterol C7-C8 isomerases, the regions from residues 91-109 and 176-194 have been termed Steroid Binding Domain Like I and II (SBDLI and SBDLII, respectively), and shown by mutagenesis and chemical derivatisation studies to be located within the S1R ligand binding site [16, 17].

In addition, several residues immediately N-terminal to SBDL1 have been implicated in S1R oligomerisation, including a five amino acid sequence (GGWMG; residues 87-91) proposed to contain a glycoporin A-like GxxxG intramembrane dimerisation motif [18, 19]. S1R oligomerisation is ligand dependent and may provide a mechanism by which its activity is regulated [18, 20]. An oligomerisation interface at or near SBDL1 would provide an obvious structural link to drug binding.

Whereas SBDLII is centered on helix 3 of the membrane associated chaperone domain [17, 21], the structure of the region encompassing SBDLI and its relationship to the membrane is not known. Sequence-based predictions of transmembrane helices
indicate that SBDLI will have at least some overlap with the second transmembrane helix (TM2). However, these predictions do not converge on a single position for TM2 (Table 1). Of the algorithms tested, some indicate the presence of a transmembrane helix between residues ~80-100, whereas others locate it between residues ~90-110, and at least one algorithm fails to confidently identify any transmembrane helix in this region.

Because of the importance of SBDLI and the adjacent region in ligand binding and receptor oligomerisation, defining the residues that constitute TM2 can shed light on the structural link between these two features of S1R function. We previously showed that the S1R chaperone domain could be reconstituted into DPC detergent micelles as a monomeric species capable of binding the ER chaperone BiP [21]. The S1R chaperone domain was found to contain five helices (H1-H5) and a flexible internal region of ~30 amino acids containing at least two regions with propensity to adopt an extended conformation. The fourth helix in the chaperone domain is amphipathic and likely drives its association with membranes. Here we report solution NMR studies of a novel S1R construct (S1R(Δ35)) in which the first transmembrane domain has been removed, enabling a description of the residues within the TM2 helix and the secondary structure of the cytosolic domain. We show that the topology and backbone dynamics of natively purified S1R(Δ35) is consistent with that determined previously for the chaperone domain [21], that TM2 consists of residues 91 to 107, and that the cytosolic region contains three helices (cH1-cH3). Identification of the TM2 residues provides a framework for further studies of S1R ligand binding and oligomerisation.

Materials and Methods

Protein sample production.

An ACA-free gene construct (GeneArt) containing an N-terminal (His)_6-tag or a (His)_9-tag, a Factor Xa cleavage site, and residues 36-223 of human S1R was subcloned into the pCOLD-I vector (Takara) and confirmed by sequencing. The substitution C94A, shown to have no effect on ligand binding in full-length guinea pig S1R [17], was introduced to prevent intermolecular disulfide bond formation during purification. The N-terminal sequence preceding the S1R residues 36-223 was
MNHKVHHHHHHIEGRHM or MNHKVHHHHHHHHHIEGRHM. The S1R(Δ35) plasmid and a pMazF plasmid containing the gene for the RNA interferase MazF (Takara) were transformed into C43(DE3) cells. Transformed cells were grown to an OD₆₀₀ of 0.8-0.9, cold shocked on ice, and incubated for 45 minutes at 15 °C. Cells were pelleted and washed with M9 salt solution, pelleted a second time, and then resuspended into a ten-fold lower volume of isotopically labeled media. Cells were incubated for a further 45-60 minutes at 15 °C before induction with 2 mM IPTG. Expression proceeded for 16 hours at 15 °C. Membranes were collected by centrifugation at 200,000 g, and incubated overnight in a solution containing 6 M guanidine, 200 mM NaCl, 1% Triton, and 20 mM Tris at pH 8.0. S1R(Δ35) was separated by Ni affinity chromatography and dialyzed against water to remove guanidine. The precipitated protein was resolubilized in hexafluoro-2-propanol and purified by HPLC on a C3 reverse phase column over a gradient from buffer A (95% water, 5% acetonitrile, 0.1% trifluoroacetic acid) to buffer B (57% 2-propanol, 38% acetonitrile, 5% water, 0.1% trifluoroacetic acid). Fractions containing S1R(Δ35) were pooled and lyophilized. For the native purification, the resuspended membranes were incubated overnight with 1% dodecylphosphocholine (DPC). The solubilized membranes were purified by Ni affinity chromatography in 20 mM DPC and subsequently gel filtrated into a low DPC concentration (3.3 mM) in a Superdex200 column (GE). Both methods yielded approximately 40 mg of pure protein per liter of labeled media. The lipid 1,2- dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) was added from powder to the sample to a q ratio of 0.1, where q = [lipid]/[detergent].

Circular Dichroism

Samples for circular dichroism (CD) contained 12.6 or 7.3 µM protein of S1R(cd) or S1R(Δ35), respectively, in DPC:DPPC mixed micelles at a q ratio of 0.1, and 20 mM potassium phosphate at pH 6.5. CD spectra were collected at room temperature on a Jasco J-815 circular dichroism spectropolarimeter from 250 to 200 nm with 10 accumulations. Data were smoothed with a Savitsky-Golay filter [22] using a window of 11 points and then corrected for concentration and number of residues.

NMR spectroscopy and data analysis
NMR experiments for resonance assignments were recorded on $^2$H, $^{15}$N, and $^{13}$C labeled samples in DPC:DPPC mixed micelles at a q ratio of 0.1, in 20 mM potassium phosphate at pH 6.5. Spectra were recorded at 600 MHz ($^1$H) on a Bruker Avance III HD spectrometer equipped with a TCI CryoProbe at 41˚ C. Backbone resonance assignments were obtained for 160 of the 181 non-proline S1R residues using a conventional set of TROSY-based experiments (HNCA, HNCACB, CBCA(CO)NH and HNCO) collected with nonuniform sampling (NUS). For NUS, random sampling schedules were used, with typically 15% of the total number of points collected. Spectra were reconstructed with compressed sensing using qMDD [23], processed using NMRPipe [24], and analysed using NMRView [25]. The assigned chemical shifts have been deposited to the BMRB with accession code 25410.

The secondary structure of S1R(Δ35) was predicted from backbone $^1$H$_N$, $^{15}$N, $^{13}$C, $^{13}$Cα and sidechain $^{13}$Cβ chemical shifts using TALOS-N [26]. $^{15}$N R$_1$, R$_2$ and $^1$H-$^{15}$N heteronuclear NOE values were measured with TROSY-based sequences collected on a 0.4 mM $^{15}$N, $^{13}$C-labeled S1R(Δ35) sample at 600 MHz using the following relaxation delays: 10, 175, 350 and 525 ms (R$_1$), and 0, 17, 51 and 85 ms (R$_2$). The recycle delays for R$_1$, R$_2$ and heteronuclear NOE experiments were 1.5 s.

**Results**

The S1R(Δ35) construct contains the cytosolic domain, the region of the second transmembrane domain, and the chaperone domain, but lacks the first transmembrane domain (predicted in residues 9-30) and approximately eight lumenal, N-terminal residues. S1R(Δ35) was reconstituted for NMR studies either from inclusion bodies or from the *E. coli* membranes **Figure 1A**. Well resolved backbone amide NMR spectra of S1R(Δ35) could be obtained in mixed micelles containing the detergent DPC and relatively small amounts of the lipid DPPC. By contrast, a subset of the resonances for S1R(Δ35) in DPC alone were weak or missing, indicating conformational exchange. The assigned backbone amide resonances are shown in **Figure 1B**. No significant chemical shift differences were observed between reconstituted S1R(Δ35) purified from membranes or inclusion bodies. The effective size of the protein in DPC micelles was assessed by SEC-MALS and TRACT [27] experiments, which indicated
that S1R(Δ35) behaves in solution as a ~29 kDa complex. This is consistent with an S1R(Δ35) monomer (theoretical mass of 23.3 kDa) and several strongly bound detergent molecules.

The secondary structure of the cytosolic domain of S1R(Δ35) determined from chemical shifts consists of three helices: residues 41-49 (cH1), 55-67 (cH2), and 81-85 (cH3) (Figure 2). The cytosolic region is followed by a mostly hydrophobic helix between residues 91-107, which is long enough to traverse the ER membrane, and therefore likely to be TM2. Residues 91-107 were predicted to be in a transmembrane helix by more than half of the algorithms tested (Table 1).

The chemical shifts for residues in common with the isolated chaperone domain are largely unchanged in S1R(Δ35), and the secondary structure of the chaperone domain in S1R(Δ35) is similar to that of the isolated chaperone domain construct [21]. The region that was most affected by the addition of the 78 N-terminal residues of the native sequence is H1 of the chaperone domain. In S1R(cd), this helix is longer (residues 121-137), but conformational exchange was observed in the C-terminal region. In S1R(Δ35), the helix ends at about residue 130 and the remaining residues exhibited a mixed propensity for helical and extended conformations. In addition, a third region (residues 160-164) in the chaperone domain of S1R(Δ35) was found to have a high propensity for an extended conformation, in addition to the two regions previously indicated by studies of the isolated chaperone domain [21].

The secondary structure composition was probed by far UV CD (Figure 3). The CD spectrum of S1R(Δ35) exhibited a similar helical content per residue as S1R(cd), consistent with their similar percent-wise composition calculated from NMR chemical shifts (42% and 46% for S1R(Δ35) and S1R(cd), respectively).

The backbone flexibility of S1R(Δ35) was evaluated by measuring $^{1}\text{H}$-$^{15}\text{N}$ heteronuclear NOEs and the ratio of the $^{15}\text{N}$ transverse ($R_2$) and longitudinal relaxation rates ($R_1$), both of which decrease with increased internal dynamics (Figure 4). Small heteronuclear NOEs and low $R_2/R_1$ values were observed for a large number of residues within the region 140-165, indicating that the fast, picosecond to nanosecond time scale motions that were previously observed here for the isolated chaperone domain [21] are maintained in the longer construct. In addition,
lower than average $R_2/R_1$ values were found in the loops between the cytosolic helices cH1 and cH2, and between cH2 and cH3, indicating flexibility.

To further validate the secondary structure-based delineation of TM2, DPPC was titrated into a S1R($\Delta$35) sample containing only DPC micelles to a q ratio of 1 (Figure 5). Assuming that lipids preferentially solvate the natively transmembrane regions of the protein, the predominant chemical shift perturbations are expected to occur at the edges of the transmembrane domain, where the phosphocholine headgroup of DPC is replaced by the phosphatidylcholine headgroup of DPPC. Similar to regions with no lipid contact, few changes would be expected in the center of the transmembrane domain because the hydrocarbon chains of the lipid and the detergent are chemically similar. Upon addition of DPPC, the largest clusters of perturbations were centered on residues 89 and 109, consistent with residues 91-107 forming the hydrophobic core of TM2.

**Discussion**

S1R functions by regulating a wide range of ion channels and receptors [15]. The activity of S1R itself depends on binding of endogenous and exogenous compounds, and ligand binding has been linked to receptor oligomerisation [18, 20]. Possibly reflecting the diversity of S1R protein-protein interactions, S1R ligands are in clinical use or preclinical studies for a wide variety of diseases [28]. Despite the pharmacological interest in S1R, however, to date structural information has only been available for the C-terminal chaperone domain [21].

Of particular interest is the location and structure of the second transmembrane domain because of its overlap with SBDLI (residues 91-109) [16, 17]. Sequence based predictions for TM2 are inconsistent (Table 1), however, predicting either residues ~80-100 or ~90-110, with some predictions being of very low confidence. Glycine residues at positions 87, 88 and 91 have been implicated in ligand binding and receptor oligomerisation [18]. Therefore an accurate definition of TM2 has implications for understanding how these residues participate in ligand binding and possibly also S1R oligomerisation.
We have studied the N-terminal truncation mutant S1R(Δ35) consisting of all residues from the start of the cytosolic domain through to the chaperone domain, and therefore containing both SBDLI and SBDLII and the glycines at positions 87, 88, and 91. Chemical shift-based secondary structure determination and lipid-induced chemical shift perturbations have been used to infer the location of TM2. Although truncation of the first transmembrane domain or S1R oligomerisation may lead to structural changes, isolated α-helical transmembrane domains solubilised in detergents are known to correlate strongly with that observed in X-ray crystallographic structures of the intact proteins [29]. Our results indicate that the S1R TM2 overlaps nearly completely with SBDLI, and that the adjacent G87 and G88 are not helical. Instead, G87 and G88 are in a linker connecting cH3 to TM2, although they do not exhibit a large increase in flexibility relative to the helical regions. One possibility consistent with these results is that the GGW sequence adopts a structured turn [30].

The position of TM2 determined here is in agreement with the larger subset of transmembrane helix prediction algorithms (in particular, Membrain, MEMSAT-SVM, Phobius, SOSUI, SPOCTOPUS, and TMHMM). The positioning of the helix is also consistent with commonly used non-membrane predictors of α-helices: PSIPRED predicts residues 89-107 to form a helix, and JPRED predicts residues 91-104 to form a helix. However, the positioning is in disagreement with a subset of the tested transmembrane helix prediction algorithms (Table 1). The lack of consensus in predicting the position of TM2 is likely due to the presence of several polar residues within the transmembrane domain. The center of TM2 is highly amphipathic, with residues 96-104 (“LHASLSEYV”) having a large, hydrophobic moment [31]. Interestingly, several of these residues, including H97 [18], S99 and Y103 [32] have been implicated in drug binding. In addition, the substitution E102Q has been implicated in a form of amyotrophic lateral sclerosis and leads to membrane mislocalisation of S1R in neuronal cells [33].

A comparison of the chaperone domain secondary structure in S1R(Δ35) with that determined previously for the chaperone domain on its own [21] indicates that the domain structure is largely maintained in the longer S1R(Δ35) construct. The largest changes occur in the C-terminal end of H1, which becomes more dynamic in S1R(Δ35), and residues 160-164, which have a stronger propensity for an extended
conformation in S1R(Δ35) (Figure 2). A comparison of the CD spectra for S1R(Δ35) and S1R(cd) supports a similar proportion of α-helical structure in the two constructs, consistent with that determined from the NMR chemical shifts. In addition, comparison of CD spectra collected in DPC micelles and DPC:DPPC mixed micelles indicated that the helical regions of both S1R(cd) and S1R(Δ35) are stabilised with the presence of lipid molecules [21].

In conclusion, we have characterised the secondary structure of a construct of S1R missing only the eight residue N-terminus and the first transmembrane domain. Because the first transmembrane helix is robustly predicted (Table 1), a secondary structure schematic of essentially the entire receptor can be made (Figure 6). The second transmembrane helix, TM2, was found to be composed of residues 91-107, which corresponds closely to the SBDLI. The cytosolic domain contains three α-helices, and the secondary structure of the chaperone domain was largely unchanged by the presence of additional native S1R sequence. The third cytosolic helix (residues 81-85) was mobile compared with cH1 and cH2, and an only modest increase in mobility of the GGW motif that joins cH3 to TM2 was observed. These findings facilitate interpretation of previous results and will guide design of subsequent studies of the role of TM2 in ligand binding and oligomerisation. These results also constitute a significant step towards a complete structure determination of S1R that will lead to a better atomic-level understanding of S1R function.

Acknowledgements

This work was funded by the UK Medical Research Council (K018590) and a FEBS Long-Term Fellowship (J.L.O.R.).
Figure Legends

Figure 1: (A) SDS-PAGE of a typical sample of S1R(Δ35) used for NMR studies. (B) Assigned NMR spectrum of S1R(Δ35). \(^{1}H,^{15}N\) TROSY-HSQC spectrum (600MHz, \(^{1}H\)) of S1R(Δ35) in 50 mM DPC, 5 mM DPPC at 41°C. Backbone amide resonance assignments are indicated. A table of chemical shifts is provided in the supplementary information.

Figure 2: Chemical shift-based secondary structure analysis of S1R(Δ35). In the top plot, the TALOS-N [26] outputs are shown, in which the magnitude of the black bars above or red bars below zero on the y-axis indicate the propensities for α-helical or extended conformations, respectively, and the propensities for an unstructured coil conformation are shown as blue circles. In the bottom plot, the \(^{13}C\alpha\) chemical shift index is shown [45]. Above the top plot is shown schematically a summary of the α-helical (cylinders) or extended (arrow) regions.

Figure 3: Far UV circular dichroism spectra of S1R(Δ35) (continuous line) and S1R(cd) (dashed line) in 50 mM DPC, 5 mM DPPC in 5 mM potassium phosphate, pH 6.5, plotted as the mean residue molar ellipticity per residue.

Figure 4: Dynamic properties of S1R(Δ35) from relaxation and heteronuclear NOE collected at 600 MHz (\(^{1}H\)) and 41°C. Top plot: the ratio of the \(^{15}N\) R\(_2\) and R\(_1\), as a function of residue number for S1R(Δ35). Bottom plot: the \(^{1}H,^{15}N\) heteronuclear NOEs. Residues K142 and S143 exhibited unusually large negative heteronuclear NOEs but were excluded from the plot due to the large uncertainty in the data.

Figure 5: Chemical shift perturbations in the backbone amide \(^{1}H\) and \(^{15}N\) resonances of S1R(Δ35) in DPC upon addition of DPPC to a DPC:DPPC ratio of 1. The solid line indicates the sum of the absolute chemical shift perturbations (CSP) over a 7-residue window, with the magnitude of the CSP for each residue (blue bars) calculated as \(\sqrt{\left(\frac{1}{6.5} \times \Delta \delta_N\right)^2 + (\Delta \delta_H)^2}\). A grey bar indicates prolines or residues for which the chemical shift assignments could not be determined.

Figure 6: A schematic diagram of full-length S1R in a lipid bilayer showing the secondary structure determined here for S1R(Δ35) construct. The loop preceding
TM2 contains G87, which is the first residue in the GGWMG sequence, and is indicated with an arrow. SBDLI corresponds closely to TM2 and SBDLII corresponds to H3 and the adjacent coil regions. No experimental data for the first transmembrane domain and N-terminus (coloured grey) is available, although the first transmembrane domain is robustly predicted to be composed of the residues from approximately 9 to 30.
References


<table>
<thead>
<tr>
<th>Algorithm</th>
<th>TM1 (residues)</th>
<th>TM2 (residues)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAS</td>
<td>12-29</td>
<td>99-107</td>
<td>[34]</td>
</tr>
<tr>
<td>HMMTOP</td>
<td>13-37</td>
<td>83-107</td>
<td>[35]</td>
</tr>
<tr>
<td>Membrain</td>
<td>8-27</td>
<td>89-108</td>
<td>[36]</td>
</tr>
<tr>
<td>MEMSAT-SVM</td>
<td>15-30</td>
<td>91-106</td>
<td>[37]</td>
</tr>
<tr>
<td>PHOBIUS</td>
<td>9-30</td>
<td>89-111</td>
<td>[38]</td>
</tr>
<tr>
<td>PRED-TMR2</td>
<td>9-30</td>
<td>81-100</td>
<td>[39]</td>
</tr>
<tr>
<td>SOSUI</td>
<td>13-34</td>
<td>86-108</td>
<td>[40]</td>
</tr>
<tr>
<td>SPOCTOPUS</td>
<td>9-29</td>
<td>88-108</td>
<td>[41]</td>
</tr>
<tr>
<td>TMHMM</td>
<td>9-31</td>
<td>89-111*</td>
<td>[42]</td>
</tr>
<tr>
<td>TMMOD</td>
<td>12-32</td>
<td>not predicted</td>
<td>[43]</td>
</tr>
<tr>
<td>TMpred</td>
<td>9-28</td>
<td>81-101</td>
<td>[44]</td>
</tr>
</tbody>
</table>

*TMHMM indicated an increased transmembrane helix probability for these residues, but did not identify it as a transmembrane domain.
Figure 3

Click here to download Figure: Fig3-CD.pdf

Molar Ellipticity, \( \theta \) (mol cm\(^2\) dmol\(^{-1}\))

S1R(\(\Delta 35\))

S1R(cd)
Figure 4

$^{15}$N $R_2/R_1$

Residue number

$\{^{\text{H},^{15}\text{N}}\text{HetNOE}\}$
Figure 5

Click here to download Figure: Fig5-DPPCtitration-v2.pdf