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

Jose Luis Ortega-Roldan*, Felipe Ossa, Nader T. Amin, Jason R. Schnell*

Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

ARTICLE INFO

Article history:
Received 14 November 2014
Revised 19 January 2015
Accepted 22 January 2015
Available online 31 January 2015
Edited by Christian Griesinger

Keywords:
Sigma-1 receptor
Transmembrane domain
NMR

1. Introduction

The sigma-1 receptor (S1R) is a membrane chaperone protein present variably 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 number 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 SBDLI have been implicated in S1R oligomerisation, including a five amino acid sequence (GGWMG; residues 87–91) proposed to contain a glycoporphin 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 SBDLI 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

* Corresponding authors.

E-mail addresses: Jose.Ortega-Roldan@bioch.ox.ac.uk (J.L. Ortega-Roldan), Jason.Schnell@bioch.ox.ac.uk (J.R. Schnell).
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(cd)) 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(cd) is consistent with that determined previously for the chaperone domain [21], that TM2 consists of residues 91–107, 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.

2. Materials and methods

2.1. Protein sample production

An ACA-free gene construct (GeneArt) containing an N-terminal (His)_6-tag or a (His)_12-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 MHNKVVHHHHHHIEGRHM or MHNKWVHHHHHHHHHHHIEGRHM. The S1R(cd) plasmid and a pMAF plasmid containing the gene for the RNA interferase MazF (Takara) were transformed into C43(DE3) cells. Transformed cells were grown to an OD_600 of 0.8–0.9, cold shocked on ice, and incubated for a further 45–60 min at 15°C. The recombinant inclusion bodies or from the E. coli cell lysate or from the inclusion bodies. The effective size of the protein in DPC micelles was assessed by SEC-MALS and TRACt [27]. The S1R(cd) 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 leuminal, N-terminal residues. S1R(cd) was reconstituted for NMR studies either from inclusion bodies or from the Escherichia coli membranes (Fig. 1A). Well resolved backbone amide NMR spectra of S1R(cd) 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(cd) in DPC alone were weak or missing, indicating conformational exchange. The assigned backbone amide resonances are shown in Fig. 1B. No significant chemical shift differences were observed between reconstituted S1R(cd) purified from membranes or inclusion bodies. The effective size of the protein in DPC micelles was assessed by SEC-MALS and TRACT [27].

Table 1
Sequence-based predictions of S1R transmembrane domains.

| Algorithm  | TM1 (residues) | TM2 (residues) | Refs. |
|------------|---------------|---------------|-------|
| DAS        | 12–29         | 99–107        | [54]  |
| HMMTOP     | 13–37         | 83–107        | [35]  |
| Membrain   | 8–27          | 89–108        | [36]  |
| MEMSAT-SVM | 15–30         | 91–106        | [37]  |
| PHBBIUS    | 9–30          | 89–111        | [38]  |
| PRED-TMR2  | 9–30          | 81–100        | [39]  |
| SOSUI      | 13–34         | 86–108        | [40]  |
| SPECTOPUS  | 9–29          | 88–108        | [41]  |
| TMHMM      | 9–31          | 89–111*       | [42]  |
| TMMOD      | 12–32         | Not predicted | [43]  |
| TMpred     | 9–28          | 81–101        | [44]  |

* TMHMM indicated an increased transmembrane helix probability for these residues, but did not identify it as a transmembrane domain.
experiments, which indicated that S1R(D35) behaves in solution as a ~29kDa complex. This is consistent with an S1R(D35) monomer (theoretical mass of 23.3kDa) and several strongly bound detergent molecules.

The secondary structure of the cytosolic domain of S1R(D35) determined from chemical shifts consists of three helices: residues 41–49 (cH1), 55–67 (cH2), and 81–85 (cH3) (Fig. 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(D35), and the secondary structure of the chaperone domain in S1R(D35) 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(A35), the helix ends at about residue 130 and the remaining residues exhibited a mixed propensity for helical and extended conformations. In addition, a region (residues 160–164) in the chaperone domain of S1R(A35) 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 (Fig. 3). The CD spectrum of S1R(A35) 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(A35) and S1R(cd), respectively).

The backbone flexibility of S1R(A35) was evaluated by measuring \( ^{1}H-^{15}N \) heteronuclear NOEs and the ratio of the \( ^{15}N \) transverse (\( R_2 \)) and longitudinal relaxation rates (\( R_1 \)), both of which decrease with increased internal dynamics (Fig. 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(A35) sample containing only DPC micelles to a \( q \) ratio of 1 (Fig. 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.

Fig. 1. (A) SDS–PAGE of a typical sample of S1R(A35) used for NMR studies. (B) Assigned NMR spectrum of S1R(A35). \(^{1}H,^{15}N\) TROSY-HSQC spectrum (600 MHz, \(^{1}H\)) of S1R(A35) 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.
4. 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(D35) 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 solubilized 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
SBDL1, 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 $\alpha$-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 ("LHASLEIV") 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(A35) with that determined previously for the chaperone domain on its own [21] indicates that the domain structure is largely maintained in the longer S1R(A35) construct. The largest
changes occur in the C-terminal end of H1, which becomes more dynamic in S1R(A35), and residues 160–164, which have a stronger propensity for an extended conformation in S1R(A35) (Fig. 2). A comparison of the CD spectra for S1R(A35) 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(A35) are stabilized with the presence of lipid molecules [21].

In conclusion, we have characterized 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 (Fig. 6). The loop preceding TM2 contains G87, which is the first residue in the GGWGQ 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 (colored gray) is available, although the first transmembrane domain is robustly predicted to be composed of the residues from approximately 9 to 30.

Acknowledgements

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

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2015.01.033.

References

[1] Hayashi, T. and Su, T.P. (2007) Sigma-1 receptor chaperones at the ER-mitochondrion interface regulate Ca(2+) signaling and cell survival. Cell 131, 596–610.
[2] Hayashi, T. and Su, T.P. (2001) Regulating ankyrin dynamics: roles of sigma-1 receptors. Proc. Natl. Acad. Sci. U.S.A. 98, 491–496.
[3] Pal, A., Fontanilla, D., Gopalakrishnan, A., Chae, Y.K., Markley, J.L. and Ruoho, A.E. (2012) The sigma-1 receptor protects against cellular oxidative stress and activates antioxidant response elements. Eur. J. Pharmacol. 682, 12–20.
[4] Wu, Z. and Bowen, W.D. (2008) Role of sigma-1 receptor C-terminal segment in inositol 1,4,5-trisphosphate receptor activation: constitutive enhancement of Ca(2+) signaling in MCF-7 tumor cells. J. Biol. Chem. 283, 28198–28205.
[5] Balasuriya, D., D’Sa, L., Talker, R., Dupuis, E., Maurin, F., Martin, P., Borgese, F., Soriani, O. and Edwardsion, J.M. (2014) A direct interaction between the sigma-1 receptor and the hERG voltage-gated K+ channel revealed by atomic force microscopy and homogeneous time-resolved fluorescence (HTRF(R)). J. Biol. Chem. 289, 32353–32363.
[6] Kourrich, S., Hayashi, T., Chuang, J.Y., Tsai, S.Y., Su, T.P. and Bonci, A. (2013) Dynamic interaction between sigma-1 receptor and Kv1.2 shapes neuronal and behavioral responses to cocaine. Cell 152, 236–247.
[7] Balasuriya, D., Stewart, A.P., Crottes, D., Borgese, F., Soriani, O. and Edwardsion, J.M. (2012) The sigma-1 receptor binds to the Nav1.5 voltage-gated Na+ channel with 4-fold symmetry. J. Biol. Chem. 287, 37021–37029.
[8] Tcheder, K.T., Huang, R.Q., Ding, A., Krishnannothriy, R.R., Dillon, G.H. and Yorio, T. (2008) Sigma-1 receptor regulation of voltage-gated calcium channels involves a direct interaction. Invest. Ophthalmol. Vis. Sci. 49, 4993–5002.
[9] Curny, M., Johansson, R., Jackson, M. and Edwardsion, J.M. (2010) Demonstration of a direct interaction between sigma-1 receptors and acid-sensing ion channels. Biophys. J. 98, 1182–1191.
[10] Balasuriya, D., Stewart, A.P. and Edwardsion, J.M. (2013) The Sigma-1 receptor interacts directly with GluN1 but not GluN2A in the GluN1/GluN2A NMDA receptor. J. Neurosci. 33, 18219–18224.
[11] Navarro, G., Moreno, E., Aymerich, M., Marcello, D., McCormick, P.J., Maillo, J., Cortes, A., Casado, V., Canela, E.L., Ortiz, J., Fuxe, K., Lluis, C., Ferrer, S. and Franco, R. (2010) Direct involvement of sigma-1 receptors in the dopamine D1 receptor-mediated effects of cocaine. Proc. Natl. Acad. Sci. U.S.A. 107, 18676–18681.
[12] Weissman, A.D., Su, T.P., Hedreen, J.C. and London, E.D. (1988) Sigma receptors in post-mortal human brains. J. Pharmacol. Exp. Ther. 247, 93–99.
[13] Seth, P., Ganapathy, M.E., Conway, S.J., Bridges, C.D., Smith, S.B., Casellas, P. and Ganapathy, V. (2001) Expression pattern of the type 1 sigma receptor in the brain and identity of critical anionic amino acid residues in the ligand-binding domain of the receptor. Biochim. Biophys. Acta 1540, 59–67.
[14] Maurice, T. and Su, T.P. (2009) The pharmacology of sigma-1 receptors. Pharmacol. Ther. 124, 195–206.
[15] Kourrich, S., Su, T.P., Fujimoto, M. and Bonci, A. (2012) The sigma-1 receptor: roles in neuronal plasticity and disease. Trends Neurosci. 35, 762–771.
[16] Pal, A., Hajipour, A.R., Fontanilla, D., Ramachandran, S., Chu, U.B., Mavlyutov, T. and Ruoho, A.E. (2007) Identification of regions of the sigma-1 receptor ligand binding site using a novel photoprobe. Mol. Pharmacol. 72, 921–933.
[17] Pal, A., Chu, U.B., Ramachandran, S., Grawio, D., Guo, L.W., Hajipour, A.R. and Ruoho, A.E. (2008) Juxtaposition of the steroid binding domain-like 1 and II regions constitutes a ligand binding site in the sigma-1 receptor. J. Biol. Chem. 283, 19564–19566.
[18] Crome, K.A., Suchy, F.P., Meddahah, H.R., Wrobel, R.L., LaPointe, L.M., Chu, U.B., Primm, J.G., Ruoho, A.E., Senes, A. and Fox, B.G. (2014) The oligomeric states of the purified sigma-1 receptor are stabilized by ligands. J. Biol. Chem. 289, 20333–20344.
[19] MacKenzie, K.K., Prestegard, J.H. and Engelman, D.M. (1997) A transmembrane helix dimer: structure and implications. Science 276, 131–133.
[20] Chu, U.B., Ramachandran, S., Hajipour, A.R. and Ruoho, A.E. (2013) Photofluorophy labeling of the sigma-1 receptor with N-[3-(4-nitrophenyl)propyl]-N-dodecylamine: evidence of receptor dimers. Biochemistry 52, 859–868.
[21] Ortega-Roldan, J.L., Ossa, F. and Schnell, J.R. (2013) Characterization of the human sigma-1 receptor chaperone domain structure and binding immunoglobulin protein (BIP) interactions. J. Biol. Chem. 288, 21448–21457.
[22] Savitzky, A. and Golay, M.J.E. (1964) Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36, 1627–1639.
[23] Kourrich, S., Cortes, A., Casado, V., Fuxe, K., Lluis, C., Ferre, S. and Fuxe, K. (2011) Inositol 1,4,5-trisphosphate receptor activation: constitutive enhancement in post-mortem human brains. J. Pharmacol. Exp. Ther. 247, 93–99.
[24] Savitzky, A. and Golay, M.J.E. (1964) Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36, 1627–1639.
[25] Shafizadeh, K., Ossal, I., Somogyi, M., Cukor, O., Sarkadi, G. and Su, T.P. (2001) Characterization of the human sigma-1 receptor chaperone domain structure and binding immunoglobulin protein (BIP) interactions. J. Biol. Chem. 288, 21448–21457.
[26] Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRpipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6, 277–293.
[27] Johnson, B.A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol. Biol. 278, 313–352.
[28] Shen, Y. and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J. Biomol. NMR 56, 227–241.
[29] Lee, D., Hilty, C., Wider, G. and Wuthrich, K. (2006) Effective rotational correlation times of proteins from NMR relaxation interference. J. Magn. Reson. 178, 72–76.
[30] Hayashi, T., Tsai, S.Y., Mori, T., Fujimoto, M. and Su, T.P. (2011) Targeting ligand-operated chaperone sigma-1 receptors in the treatment of neuropsychiatric disorders. Expert Opin Ther Targets 15 (5), 557–577.
[29] Tulumello, D.V. and Deber, C.M. (2012) Efficiency of detergents at maintaining membrane protein structures in their biologically relevant forms. Biochim. Biophys. Acta 1818, 1351–1358.

[30] Lorieau, J.L., Louis, J.M. and Bax, A. (2011) Helical hairpin structure of influenza hemagglutinin fusion peptide stabilized by charge-dipole interactions between the N-terminal amino group and the second helix. J. Am. Chem. Soc. 133, 2824–2827.

[31] Gautier, R., Douquett, D., Antonny, B. and Drin, G. (2008) HELIQUEST: a web server to screen sequences with specific alpha-helical properties. Bioinformatics 24, 2101–2102.

[32] Yamamoto, H., Miura, R., Yamamoto, T., Shinozawa, K., Watanabe, M., Okuyama, S., Nakazato, A. and Nukada, T. (1999) Amino acid residues in the transmembrane domain of the type 1 sigma receptor critical for ligand binding. FEBS Lett. 445, 19–22.

[33] Al-Saif, A., Al-Mohanna, F. and Bohlega, S. (2011) A mutation in sigma-1 receptor causes juvenile amyotrophic lateral sclerosis. Ann. Neurol. 70, 913–919.

[34] Cserzo, M., Wallin, E., Simon, L., von Heijne, G. and Elofsson, A. (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng. 10, 673–676.

[35] Tusnady, G.E. and Simon, I. (2001) The HMMTOP transmembrane topology prediction server. Bioinformatics 17, 849–850.

[36] Jones, D.T. (2007) Improving the accuracy of transmembrane protein topology prediction using evolutionary information. Bioinformatics 23, 538–544.

[37] Krogh, A., Lorentzon, A. and Sonnhammer, E.L. (2004) A combined transmembrane topology and signal peptide prediction method. J. Mol. Biol. 338, 1027–1036.

[38]Pasquier, C. and Hamodrakas, S.J. (1999) An hierarchical artificial neural network system for the classification of transmembrane proteins. Protein Eng. 12, 631–634.

[39] Hirokawa, T., Boon-Chieng, S. and Mitaku, S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14, 378–379.

[40] Viklund, H., Bernsel, A., Skwark, M. and Elofsson, A. (2008) SPOCOPUS: a combined predictor of signal peptides and membrane protein topology. Bioinformatics 24, 2928–2929.

[41] Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L. (2001) Predicting transmembrane topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305, 567–580.

[42] Kahsay, R.Y., Gao, G. and Liao, L. (2005) An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. Bioinformatics 21, 1853–1858.

[43] Hoffmann, K. and Stoffel, W. (1993) TMbase – a database of membrane spanning proteins segments. Biol. Chem. 374, 166.

[44] Sper, S. and Bax, A. (1991) Empirical correlation between protein backbone conformation and Calpha and Cbeta 13C nuclear magnetic resonance chemical shifts. J. Am. Chem. Soc. 113, 5490–5492.