Glycosylation pattern analysis of glycoprotein hormones and their receptors

Ricardo Núñez Miguel, Jane Sanders, Jadwiga Furmaniak and Bernard Rees Smith
FIRS Laboratories, RSR Ltd, Cardiff, UK

Abstract

We have studied glycosylation patterns in glycoprotein hormones (GPHs) and glycoprotein hormone receptor (GPHR) extracellular domains (ECD) from different species to identify areas not glycosylated that could be involved in intermolecular or intramolecular interactions. Comparative models of the structure of the TSHR ECD in complex with TSH and in complex with TSHR autoantibodies (M22, stimulating and K1-70, blocking) were obtained based on the crystal structures of the FSH-FSHR ECD, M22-TSHR leucine-rich repeat domain (LRD) and K1-70-TSHR LRD complexes. The glycosylation sites of the GPHRs and GPHs from all species studied were mapped on the model of the human TSH TSHR ECD complex. The areas on the surfaces of GPHs that are known to interact with their receptors are not glycosylated and two areas free from glycosylation, not involved in currently known interactions, have been identified. The concave faces of GPHRs leucine-rich repeats 3–7 are free from glycosylation, consistent with known interactions with the hormones. In addition, four other non-glycosylated areas have been identified, two located on the receptors’ convex surfaces, one in the long loop of the hinge regions and one at the C-terminus of the extracellular domains. Experimental evidence suggests that the non-glycosylated areas identified on the hormones and receptors are likely to be involved in forming intramolecular or intermolecular interactions.

Introduction

TSH belongs to the glycoprotein hormone (GPH) family, and the GPH family belongs to the cystine knot growth factor superfamily (McDonald & Hendrickson 1993). The carbohydrate moiety in GPHs constitute 15–35% by weight. The human (h) α subunit has two N-linked oligosaccharides, whereas the β subunits have one in TSH and LH and two in CG and FSH (Grossmann et al. 1997). In contrast, hCG has a C-terminal extension of ~25 residues in its β chain that shows O-linked glycosylation. In addition, hCG is composed of a group of three different molecules, glycosylated hCG, hyperglycosylated hCG and hyperglycosylated hCG-free β subunit, each having different physiological functions (Cole 2009, Fournier 2016). The O-linked sites are expendable for hCG function (Wu et al. 1994), whereas N-linked glycosylations play a role in signal transduction but do not affect receptor-binding affinity (Moyle et al. 1975). Deglycosylated hCG forms a tight complex with its receptor but fails to activate it, acting as an antagonist. In contrast, hypoglycosylated hFSH is more active in vitro.
than fully glycosylated hFSH (Bousfield et al. 2014). The crystal structures of hCG (Lapthorn et al. 1994, Wu et al. 1994) and hFSH (Fox et al. 2001) are available. The N-linked oligosaccharides of glycoproteins are necessary for proper folding, assembly, secretion, metabolic clearance, plasma half-life and biological activity (Benzaiger & Green 1988, Grossmann et al. 1997). In the endoplasmic reticulum (ER), oligosaccharides facilitate disulfide bond formation and are important for proper folding. Molecular chaperones retain glycoproteins in the ER until proper trimming of the carbohydrates before releasing to the next compartment. In the Golgi apparatus, the carbohydrates are further trimmed and processed by sequential addition of carbohydrate residues (Fiedler & Simons 1995).

The TSH receptor (TSHR) belongs to the glycoprotein hormone receptor (GPHR) subfamily of the leucine-rich repeat-containing family of the class A GPCRs (Hsu et al. 2000). GPHRs are characterised by a large amino-terminal extracellular domain (ECD), which forms the high-affinity binding site for their cognate hormones (Nagayama & Rapoport 1992, Segaloff & Ascoli 1993). GPHR extracellular domains are heavily N-glycosylated with oligosaccharides representing 30–40% of their molecular weight (Rapoport et al. 1996). The crystal structures of the leucine-rich repeat domain (LRD) of the hTSHR in complex with the TSH-stimulating human monoclonal autoantibody (hMAb) M22 (Sanders et al. 2007) and with the TSHR-blocking hMAb K1-70 (Sanders et al. 2011) are available. In addition, the crystal structure of the extracellular domain of the hFSHR bound to hFSH has been determined (Jiang et al. 2012).

It is likely that the interactions that occur during hormone binding and receptor activation are similar for all members of the glycoprotein hormone and glycoprotein hormone receptor families and for all species. This is because of the high amino acid sequence identity, ranging from 33% to 99%. Additional potential protein–protein interaction sites on the surfaces of glycosylated proteins should be identifiable by studying glycosylation of a large number of homologous proteins from different species. Here, we describe a study of glycosylation patterns in GPHs and GPHR extracellular domains for a large set of amino acid sequences from different species, taken from the UniProt database, to identify the areas on the surfaces of the GPHs and their receptors that are not glycosylated in any of the species studied. Such glycosylation-free areas may be functionally important by being involved in various protein–protein interactions. Although sugar–sugar and sugar–protein interactions are possible between glycoproteins, this study is limited to protein–protein interactions. We looked for glycosylation site candidates in the amino acid sequences of GPH and GPHR extracellular domains from different species as those complying with the consensus sequence Asn-Xxx-Ser/Thr, where Xxx is any amino acid except proline. We mapped the location of glycosylation site candidates from the three receptors and the four hormones in all species studied (77 receptor sequences, 62 α chain hormone sequences and 202 β chain hormone sequences) on the three-dimensional structures of hTSH and hTSHR ECD. To obtain the glycosylation maps, comparative models of the structure of the TSHR ECD in complex with TSH, M22 and K1-70 were obtained based on the crystal structures of M22-TSHR LRD (Sanders et al. 2007), K1-70-TSHR LRD (Sanders et al. 2011) and FSH-FSHR ECD (Jiang et al. 2012).

Materials and methods

Comparative modelling of the TSHR ECD in complex with TSH, M22 and K1-70

Initial alignments between the amino acid sequences of the TSH and FSH and between the TSHR ECD and FSHR ECD were obtained using the program ClustalW (Thompson et al. 1994). The initial sequence alignment between the TSHR ECD and FSHR ECD was then further modified based on a structural alignment between the crystal structures of the TSHR LRD (Sanders et al. 2011) and FSHR LRD (Fan & Hendrickson 2005), using the Discovery Studio 3.5 suite of software (http://accelrys.com/products/discovery-studio/). An initial model of the structure of the TSH-TSHR ECD complex was predicted by comparative modelling based on the crystal structure of the FSH-FSHR-ECD complex (Jiang et al. 2012) solved at 2.5Å resolution (PDB-Id: 4AY9). The geometry and conformation of the modelled structure were validated by validation programs (see below), problematic sequence segments were identified, the alignment was manually modified to improve validation results and the modelling and validation processes were repeated until a model with good geometry and conformation defined by the validation program was obtained. The final sequence alignment is shown in Supplementary Alignment 1 (see section on supplementary data given at the end of this article). Then, the atomic coordinates of the TSHR residues Cys24-Arg255 from the initial model of the TSH-TSHR ECD complex were replaced by the atomic coordinates of the same residues from the crystal structure of the TSHR LRD from the K1-70-TSHR LRD complex (Sanders et al. 2011) at 1.9Å resolution (PDB-Id: 2XWT).
after a coordinate superimposition that gave an RMSD of 0.36 Å for all backbone atoms. The model of the structure of the TSH-TSHR ECD complex was energy minimised with Discovery Studio 3.5 (see below) to optimise the interactions between the two components of the complex.

Models of the TSHR ECD bound to the thyroid-stimulating hM Ab M22 and the TSHR-blocking hM Ab K1-70 were obtained based on the structure of the comparative model of the TSH-TSHR ECD complex and the crystal structures of the M22-TSHR LRD complex (Sanders et al. 2007) at 2.55 Å resolution (PDB-Id: 3G04) and the K1-70-TSHR LRD complex (Sanders et al. 2011) at 1.9 Å resolution (PDB-Id: 2XWT), respectively. In the case of the M22-TSHR ECD model, after a coordinate superimposition of the structures of the TSH-TSHR ECD and M22-TSHR LRD complexes, the atomic coordinates of the TSHR residues Glu30-Arg255 and TSH from the model of the TSH-TSHR ECD complex were substituted for the atomic coordinates of the same TSH residues and M22, respectively, from the crystal structure of the M22-TSHR LRD complex. In the case of the K1-70-TSHR ECD model, after a coordinate superimposition of the structures of the TSH-TSHR ECD and K1-70-TSHR LRD complexes, the atomic coordinates of TSH from the model of the TSH-TSHR ECD complex were substituted for the atomic coordinates of the same TSH residues and M22, respectively, from the crystal structure of the K1-70-TSHR LRD complex.

The program MODELLER (Sali & Blundell 1993) within the Discovery Studio 3.5 suite of software was used for comparative modelling. ‘Check Structure’ and ‘Profiles-3D’ functionalities of Discovery Studio 3.5 were used for structure validation. Superimpositions of the coordinates of peptide chains and optimisation of the interfaces (using the CHARMm forcefield) were carried out by Discovery Studio 3.5.

Glycosylation study of glycoprotein hormones and their receptors

The glycoprotein hormone and glycoprotein hormone receptor amino acid sequences from different species, including mammals, birds, reptiles, amphibians and fishes, were obtained from the UniProt database of The European Bioinformatics Institute (EBI) (http://www.ebi.ac.uk/uniprot/). Only hormone chain sequences with greater than 60 residues and receptor ECD sequences with greater than 255 residues for the LH/CG and FSH receptors and greater than 305 residues for the TSH receptor were considered in this study. We studied 62 amino acid sequences of glycoprotein hormone α chains, 68 of FSH β chains, 85 of LH β chains, 9 of CG β chains, 40 of TSH β chains, 34 amino acid sequences of FSHR ECDs, 24 of CG/LH receptor (LHR) ECDs and 19 of TSHR ECDs. Amino acid sequence alignments of the different species of each receptor studied or the different species of each hormone chain studied were produced by the program ClustalW (Thompson et al. 1994). Sequence alignments (Supplementary Alignments 2, 3, 4, 5, 6, 9, 10 and 11) were not manually modified except in the case of a few clear alignment errors, and these did not involve glycosylation site candidates.

Glycosylation site candidates were identified as those that complied with the consensus sequence Asn-Xxx-Ser/Thr in which Xxx is any amino acid except proline (Kornfeld & Kornfeld 1985). An in-house program (GLYCOS, Núñez Miguel R, unpublished) was used for fast identification of glycosylation site candidates. The hTSH and hTSHR residues equivalent to the residues that are glycosylated in any of the glycoprotein hormone or glycoprotein hormone receptor sequences studied were mapped on the molecular surface of the TSH-TSHR ECD complex. The residues of the human hormone or receptor sequences that are aligned in the alignment of sequences from different species (eight family alignments; Supplementary Alignments 2, 3, 4, 5, 6, 9, 10 and 11) were considered to be the equivalent residues. Then, the TSH and TSHR equivalent residues identified from the eight family alignments were assessed in three alignments using only the human sequences (Supplementary Alignments 7, 8 and 12). Once all glycosylation site candidates were mapped, areas that are not glycosylated were identified on the molecular surfaces of the TSHR ECD and TSH. The identification of the non-glycosylated areas were made visually using the molecular graphics visualisation programs RasMol (http://sourceforge.net/projects/openrasmol/) and PyMOL (DeLano 2002).

The program PyMOL (DeLano 2002) was used for the visualisation of the glycosylation site candidates highlighted on the molecular surfaces of the hormones and receptors studied and for the generation of protein structure figures.

Results

Glycosylation study

Amino acid sequence alignments of GPHs and their respective receptors from different species were performed separately for each receptor and for each hormone chain (Supplementary Alignments 2, 3, 4, 5, 6, 9, 10 and 11).
Then, glycosylation site candidates were identified as described in ‘Methods’ section. An additional spatial condition was applied to the Asn residues that may be glycosylated as they have to be accessible to luminal oligosaccharyltransferase.

**Glycosylation of glycoprotein hormones**

A glycosylation study of the glycoprotein hormones was carried out for each of the five subunits: the α subunit and the four β subunits, FSH β, LH β, CGH β and TSH β. Each subunit is composed of 4 beta strands (β1–β4 for each of the α and β subunits) and 3 loops (αL1, αL2 and αL3 for the α subunit and βL1, βL2 and βL3 for the β subunit) (Fig. 1A).

A set of 62 amino acid sequences of the GPH α chain from different species was used to study glycosylation in the GPH α chain family (Table 1, Supplementary Alignment 2). Two glycosylation site candidates in the GPH α chain family are conserved in all 62 species studied except for Pike eel (Supplementary Table 1) in which Asn52 is substituted by Asp. The two conserved glycosylation sites of the human GPH α chain are glycosylated in the crystal structures of hFSH (Fox et al. 2001) and hGC (Lapthorn et al. 1994, Wu et al. 1994) at αAsn52 and αAsn78. The hFSH structure reveals that the sugar moiety at αAsn52 forms a hydrogen bond with βTyr58 and contributes to heterodimer stability. In hCG, the equivalent residue to βTyr58 is βPhe64, which is unable to form a hydrogen bond, but capable of hydrophobic interaction with the hydrophobic side of the sugar ring at αAsn52 (Fox et al. 2001). Lapthorn et al. (1994) reported contacts between hCG αAsn52 carbohydrate and hCG residues βTyr59, βVal62, βPhe64, βAla83 and βThr97.
A set of 68 amino acid sequences of FSH β from different species was used to study glycosylation in the FSH β chain family (Table 1, Supplementary Alignment 3). Glycosylation of the FSH β chain family is not as highly conserved as the GPH α chain family. For example, four of the seven glycosylation site candidates in the FSH β chain family were found only in four different fishes and were not present in other species (Supplementary Table 2). The four different glycosylation site candidates occur separately, one in each of the four different fish sequences. A different glycosylation site at FSH β position 24 is equivalent to glycosylation sites found in the other GPH β chains (TSH β position 23, LH β and CG β positions 30; Table 2, Supplementary Tables 2, 3, 4, 5; Supplementary Alignment 7), but some fishes do not have this glycosylation site (Supplementary Table 2). The crystal structure of hFSH in complex with the FSHR ECD (Jiang et al. 2012) and one of the two hormone–receptor complexes in the asymmetric unit of the crystal structure of the hFSH in complex with the FSHR LRD (Fan & Hendrickson 2005) show glycosylation of both FSH β residues Asn7 and Asn24. The other complex in the asymmetric unit in the Fan and Hendrickson (2005) structure, as well as the structures of hFSH solved by Fox et al. (2001) do not show glycosylation at FSH β Asn24.

Furthermore, the FSH β sequences of thirteen species of fish lack the third Cys (βCys20), which is bonded to the last Cys (βCys104) in all other FSH β sequences. In Gasterosteus aculeatus (three-spined stickleback fish), βCys20 has been substituted for Asn forming a glycosylation site candidate. In the crystal structure of the FSH-FSHR complex (Fan & Hendrickson 2005), βCys20 appears to have its side chain buried by βCys104. Accordingly, the Asn that substitutes βCys20 in Gasterosteus aculeatus is unlikely to be accessible to the oligosaccharyltransferase. However, the absence of this disulphide bond in Gasterosteus aculeatus is likely to make the C-terminus of FSH β flexible unburying the side chain of the Asn in Gasterosteus aculeatus and making it accessible to the oligosaccharyltransferase. It is not clear whether the βAsn20 of Gasterosteus aculeatus is or is not glycosylated.

Amino acid sequences (n=85) of the LH β chain from different species were used for studying glycosylation of the LH β chain family (Table 1, Supplementary Alignment 4). LH β chain glycosylation site candidates at positions 13 and 30 are equivalent to the two FSH β at positions 7 and 24, respectively (Table 2, Supplementary Alignment 7). It is interesting to note that primates, more precisely catarrhines (apes and Old World monkeys), have the two
glycosylation site candidates, whereas humans have lost the most conserved site (at position 13) of the family but have kept the catarrhini-specific site (at position 30). In the human sequence, Asn13 is conserved, but residue Thr15, needed for the complete glycosylation site consensus sequence, has been substituted for Ile. Takifugu rubripes (pufferfish) has also substituted this Thr for Met, whereas the other 83 species studied have the glycosylation site consensus sequence (Supplementary Table 3). A crystal structure of LH β is available but Weisshaar et al. (1991) describe the carbohydrate structures of hLH at Asn30 in the β chain, established by 1D and 2D 1H-NMR spectroscopy, composition analysis, methylation analysis and fast-atom-bombardment mass spectrometry (FAB-MS). This is consistent with the analysis based on amino acid sequence alignment.

A set of 9 amino acid sequences of CG β from primates were used to study glycosylation of the CG β chain family (Table 1, Supplementary Alignment 5). Glycosylation of CG β is different for New World monkeys compared to catarrhines (Supplementary Table 4). Further to the glycosylation site at position 30 that is present in all sequences, catarrhines have a glycosylation site at position 13, whereas the New World monkeys have a glycosylation site candidate at position 127. The conserved glycosylation site at position 30 is equivalent to glycosylation at positions 24 in FSH β and 30 in LH β (Table 2, Supplementary Alignment 7), whereas the glycosylation site at position 13 (conserved in catarrhines) is equivalent to glycosylation sites at positions 7 in FSH β and 13 in LH β (Table 2). The CG β chain is highly conserved with one glycosylation site present in all 40 sequences studied. This is at position 23 (Supplementary Table 5) and is equivalent to FSH β24, LH β30 and CG β30 (Table 2, Supplementary Alignment 7). No experimental structure of TSH β is available at present, but expression of wild-type hTSH and hTSH with Asn to Gln mutations showed that the in vitro activity of hTSH lacking the βAsn23 oligosaccharide expressed in CHO-K1 cells was increased 2–3-fold compared with wild type (Grossmann et al. 1995).

Most of the GPH β chains show a highly conserved glycosylation site located in strand β1 (corresponding to...
hTSH β position 6, hFSH β7, hLH β13 and hCG β7; Table 2 and Supplementary Alignment 7), but this is not present in TSH β chains of any of the species studied. Most of the TSH β sequences have a Glu or Asp amino acid instead of an Asn at this position except Monodelphis domestica (grey short-tailed opossum, South American marsupial) that has a Gly (position β6) instead of Asn while retaining the other glycosylation consensus amino acid Thr. In the CG β family, this site (β7) is glycosylated only in the catarrhines studied and not in the New World monkeys. In the case of the LH β family, this glycosylation site is conserved in all 85 studied sequences except human and one fish, and in the FSH β family, it is conserved in 63 of the 68 sequences studied.

Glycosylation of glycoprotein hormone receptors

The ECD of the glycoprotein hormone receptors has been divided into an N-terminal cap (NT-cap), 11 leucine-rich repeats (LRRs), a hairpin loop (H-loop), a 12th incomplete leucine-rich repeat and a C-terminal cap (CT-cap) (Fig. 1B). The NT-cap together with LRRs 1–10 forms the leucine-rich repeat domain (LRD), whereas LRRs 11 and 12, the H-loop and the CT-cap form the hinge region. TSHR amino acid sequences defining each TSHR ECD structural element are shown in Fig. 2.

A set of 34 amino acid sequences of FSHR ECDs from different species was used to study the glycosylation of the FSH receptor family (Table 1, Supplementary Alignment 9).

<table>
<thead>
<tr>
<th>TSHR</th>
<th>NT-cap</th>
<th>(23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSHR</td>
<td>1st repeat</td>
<td>(36)</td>
</tr>
<tr>
<td>TSHR</td>
<td>2nd repeat</td>
<td>(53)</td>
</tr>
<tr>
<td>TSHR</td>
<td>3rd repeat</td>
<td>(77)</td>
</tr>
<tr>
<td>TSHR</td>
<td>4th repeat</td>
<td>(102)</td>
</tr>
<tr>
<td>TSHR</td>
<td>5th repeat</td>
<td>(127)</td>
</tr>
<tr>
<td>TSHR</td>
<td>6th repeat</td>
<td>(152)</td>
</tr>
<tr>
<td>TSHR</td>
<td>7th repeat</td>
<td>(176)</td>
</tr>
<tr>
<td>TSHR</td>
<td>8th repeat</td>
<td>(201)</td>
</tr>
<tr>
<td>TSHR</td>
<td>9th repeat</td>
<td>(227)</td>
</tr>
<tr>
<td>TSHR</td>
<td>10th repeat</td>
<td>(248)</td>
</tr>
<tr>
<td>TSHR</td>
<td>11th repeat</td>
<td>(271)</td>
</tr>
<tr>
<td>TSHR</td>
<td>H-loop</td>
<td>(292)</td>
</tr>
<tr>
<td>TSHR</td>
<td>CT-cap</td>
<td>(403)</td>
</tr>
</tbody>
</table>

A glycosylation site candidate present in one reptile (Podarcis sicula) and one fish (Gadus morhua) occurs in a segment of their amino acid sequences that corresponds to an insertion relative to the amino acid sequence of the hFSHR (Supplementary Table 6). The FSHR shows much less conservation in the glycosylation site pattern compared to the GPHs. Only one absolutely conserved glycosylation site, located in the 7th LRR, is observed among the FSHR family that corresponds to Asn191 in the hFSHR (Supplementary Table 6). The crystal structures of the human FSH-FSHR LRD (Fan & Hendrickson 2005) and the human FSH-FSHR ECD (Jiang et al. 2012) show sugar molecules at FSHR glycosylation site Asn191. One of the two glycosylation site candidates located in repeat 8 (position 199) is present in all mammals, birds, reptiles and amphibians, whereas it is absent in fish. A glycosylation site candidate, present only in humans, is located in the H-loop (position 318, not visible in the FSHR ECD crystal structure) (Supplementary Table 6).

A set of 24 amino acid sequences of the LH/CG receptor from different species was used to study glycosylation of the LH family (Table 1, Supplementary Alignment 10). There are two absolutely conserved glycosylation site candidates in the LH family located in the 7th repeat (position 195) and the H-loop (position 299). The former is equivalent to glycosylation site Asn191 in the hFSHR (Table 3, Supplementary Alignment 12). In addition to the two absolutely conserved sites, LHRs of mammals have another four glycosylation sites (Supplementary Table 7). No crystal structure of the LHR is available at present. Davis et al. (1997) engineered a series of rat LHR constructs in which all but one glycosylation site candidate had been disrupted by mutation. Each construct was deglycosylated using PNGaseF resulting in a reduction in mass on SDS-PAGE, indicating that all six glycosylation site candidates of rat LHR are glycosylated. Glycosylation of rat LHR at only Asn195 and Asn174 is sufficient for high-affinity hormone binding (Davis et al. 1997).

Amino acid sequences of the TSHR from 19 different species were used to study glycosylation in the TSHR family (Table 1, Supplementary Alignment 11). The TSHR family shows three absolutely conserved glycosylation site candidates located in the 3rd and 7th repeats and the N-terminus of the H-loop (positions 77, 198 and 302, respectively). Mammals and chicken have another glycosylation site in repeat 6 (position 177) (Supplementary Table 8). The crystal structure of the hTSHR (Sanders et al. 2007) shows that the glycosylation sites candidates Asn77, Asn99, Asn113, Asn177 and...
Asn198 are glycosylated. Neither TSHR crystal structures determined so far included Asn302 (Sanders et al. 2007, 2011); hence, glycosylation at this site could not be confirmed. However, studies with tryptic fragments of the TSHR expressed in CHO cells showed that Asn302 was likely to be glycosylated (Tanaka et al. 1998). The TSHR glycosylation site formed by the sequence Asn99-Xxx-Ser101 present in humans is conserved in all mammals except in the two primates studied, which have Ser-to-Asn mutation in the consensus sequence. In terms of the relationship between the human TSHR sequence and the sequences of the two primates, our study cannot discern whether in humans, mutation at TSHR position 101 occurred back to Ser or the common ancestor of human and the two primates had a Ser at position 101 that was subsequently substituted by Asn.

There is one absolutely conserved glycosylation site in all three glycoprotein hormone receptors located in repeat 7. In the case of hTSHR, this glycosylation site corresponds to Asn198. In addition, a glycosylation site candidate located in the hairpin loop (position 302), immediately after cysteine appears in all TSHR sequences studied, in the FSHR sequences of all mammals, 1 reptile, 1 amphibian and 3 fishes and in the LHR sequences of two fishes.

### Comparative model of the structure of the TSH-TSHR ECD complex

The comparative model of the structure of the TSH-TSHR ECD complex is composed of the crystal structure of hTSHR residues 24–255, from the K1-70-TSHR LRD complex (Sanders et al. 2011); a comparative model of hTSHR residues 256–411, based on the crystal structure of the FSHR ECD bound to FSH (Jiang et al. 2012); and a comparative model of hTSH, based on the crystal structure of FSH bound to the FSHR ECD (Jiang et al. 2012).

The model of the structure of the TSH-TSHR ECD complex (Fig. 1C) shows TSH bound to the concave surface of the receptor LRD in a similar way as FSH is

<table>
<thead>
<tr>
<th>Pos*</th>
<th>Location</th>
<th>Number of sequences showing the site</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>FSHR</td>
<td>LHR</td>
</tr>
<tr>
<td>30</td>
<td>NT-cap</td>
<td>0/34</td>
<td>1/24</td>
</tr>
<tr>
<td>33</td>
<td>1st repeat</td>
<td>8/34</td>
<td>0/24</td>
</tr>
<tr>
<td>43</td>
<td>1st repeat</td>
<td>0/34</td>
<td>4/24</td>
</tr>
<tr>
<td>45</td>
<td>1st repeat</td>
<td>1/34</td>
<td>0/24</td>
</tr>
<tr>
<td>53</td>
<td>2nd repeat</td>
<td>12/34</td>
<td>0/19</td>
</tr>
<tr>
<td>58</td>
<td>2nd repeat</td>
<td>1/34</td>
<td>0/24</td>
</tr>
<tr>
<td>Insert.</td>
<td>Insertion after human</td>
<td>2/34</td>
<td>0/24</td>
</tr>
<tr>
<td>77</td>
<td>3rd repeat</td>
<td>0/34</td>
<td>5/24</td>
</tr>
<tr>
<td>99</td>
<td>3rd repeat</td>
<td>6/34</td>
<td>0/24</td>
</tr>
<tr>
<td>102</td>
<td>4th repeat</td>
<td>0/34</td>
<td>0/24</td>
</tr>
<tr>
<td>113</td>
<td>4th repeat, centre</td>
<td>0/34</td>
<td>0/24</td>
</tr>
<tr>
<td>124</td>
<td>4th repeat</td>
<td>1/34</td>
<td>0/24</td>
</tr>
<tr>
<td>143</td>
<td>5th repeat</td>
<td>2/34</td>
<td>0/24</td>
</tr>
<tr>
<td>177</td>
<td>6th repeat</td>
<td>1/34</td>
<td>11/24</td>
</tr>
<tr>
<td>198</td>
<td>7th repeat</td>
<td>34/34</td>
<td>24/24</td>
</tr>
<tr>
<td>206</td>
<td>8th repeat</td>
<td>22/34</td>
<td>0/24</td>
</tr>
<tr>
<td>218</td>
<td>8th repeat</td>
<td>2/34</td>
<td>0/24</td>
</tr>
<tr>
<td>237</td>
<td>9th repeat</td>
<td>1/34</td>
<td>0/24</td>
</tr>
<tr>
<td>276</td>
<td>11th repeat</td>
<td>13/34</td>
<td>1/24</td>
</tr>
<tr>
<td>288</td>
<td>11th repeat</td>
<td>2/34</td>
<td>0/24</td>
</tr>
<tr>
<td>292</td>
<td>H-loop</td>
<td>9/34</td>
<td>0/24</td>
</tr>
<tr>
<td>295</td>
<td>H-loop</td>
<td>0/34</td>
<td>9/24</td>
</tr>
<tr>
<td>302</td>
<td>H-loop</td>
<td>20/34</td>
<td>24/24</td>
</tr>
<tr>
<td>312</td>
<td>H-loop</td>
<td>0/34</td>
<td>2/24</td>
</tr>
<tr>
<td>316</td>
<td>H-loop</td>
<td>0/34</td>
<td>11/24</td>
</tr>
<tr>
<td>319</td>
<td>H-loop</td>
<td>1/34</td>
<td>0/24</td>
</tr>
<tr>
<td>326</td>
<td>H-loop</td>
<td>1/34</td>
<td>0/24</td>
</tr>
<tr>
<td>392</td>
<td>H-loop</td>
<td>7/34</td>
<td>1/24</td>
</tr>
</tbody>
</table>

Table 3 Glycosylation site candidates in the glycoprotein hormone receptor family.

*Position numbers correspond to the human TSHR amino acid sequence, Supplementary Alignment 12. bGlycosylation status cannot be assessed because the residue is located in an insertion with respect the human sequence.
bound to the FSHR LRD in the crystal structure (Fan & Hendrickson 2005). The structure of the TSHR ECD in the model displays a slightly curved tube shape formed by an NT-cap, 12 LRRs, a H-loop and a CT-cap (Fig. 2). The NT-cap includes two disulfide bonds connecting cysteines C24 with C29 and C31 with C41, which is a different arrangement than that in the FSHR ECD (Sanders et al. 2011, Jiang et al. 2012). The CT-cap includes two disulfide bonds connecting cysteines C283 with C398 and C284 with C408. An additional disulfide bond is present in the H-loop between cysteines C301 and C390. The TSHR H-loop is longer (103 residues) than that of the FSHR (60 residues) and the sulphated Tyr385 in the TSHR (Costagliola et al. 2002) is equivalent to FSHR Phe333, whereas the sulphated-Tyr in the FSHR is Tyr335 (Jiang et al. 2012).

In the model of the TSHR ECD, the structures of repeats 11 and 12 adopted highly similar structures to those of the FSHR ECD (Jiang et al. 2012). In contrast, the segment at the convex surface of repeat 10 of the TSHR shows a distribution of amino acid residues (Leu263-Leu272) different than that of the FSHR (Figs 2 and 3). An alignment as shown in Fig. 3A would leave two leucines exposed to the solvent and one serine buried in the hydrophobic core. It is accepted in the protein structure prediction field that errors in the structure of model proteins can be reflected, among other reasons, by buried polar/charged groups or high solvent-exposure of hydrophobic residues (Novotn’y et al. 1988, Nabuurs et al. 2006, Pugalenthi et al. 2006). Accordingly, the alignment was modified (Fig. 3B) so that only one phenylalanine is exposed to the solvent and further remodelling of this segment reduced Phe269 exposure to the solvent.

The structure of the TSH-TSHR ECD complex does not predict the interactions between TSH and the TSHR leucine-rich repeats 11 and 12 (residues 271–291 and 395–402, respectively), but it does predict interactions of residues from the H-loop of the receptor with hormone residues, in addition to the interaction of the sulphated Tyr385. In particular, the model predicts the interactions between residues from loop βL2 of the hormone and the TSHR H-loop residues Tyr387, Thr388 and Ile389. The interaction with Thr388 is polar, whereas the interactions with the other two TSHR residues are hydrophobic. The FSH-FSHR ECD crystal structure shows the interaction between FSH and the FSHR sulphated Tyr335 and in addition interactions are observed between residues from loop βL2 of the hormone and the FSHR H-loop residues Asp334 and Leu337. The interaction with Asp334 is electrostatic, whereas the interaction with Leu337 is hydrophobic.

Comparative model of the structure of the M22-TSHR ECD complex

The comparative model of the structure of the M22-TSHR ECD complex consists of the crystal structure of hTSHR residues 24–29, from the K1-70-TSHR LRD complex (Sanders et al. 2011); the crystal structure of hTSHR residues 30–255, from the M22-TSHR LRD complex (Sanders et al. 2007); a comparative model of hTSHR residues 256–411, based on the crystal structure of the FSHR ECD bound to FSH (Jiang et al. 2012); and the crystal structure of M22 from the M22-TSHR LRD complex (Sanders et al. 2007).

The model of the M22-TSHR ECD complex (Fig. 1D) shows M22 bound to the whole concave surface of the LRD of the TSHR and the hinge region. M22 is predicted to interact with all twelve LRRs and contacts the H-loop and the CT-cap of the hinge region of the TSHR ECD.

Comparative model of the structure of the K1-70-TSHR ECD complex

The comparative model of the structure of the K1-70-TSHR ECD complex is composed of: the crystal structure of hTSHR residues 24–255, from the K1-70-TSHR LRD complex (Sanders et al. 2011); a comparative model of hTSHR residues 256–411, based on the crystal structure of the FSHR ECD bound to FSH (Jiang et al. 2012); and the crystal structure of K1-70 from the K1-70-TSHR LRD complex (Sanders et al. 2011).

In the model (Fig. 1E), K1-70 binds to the concave surface of the ECD of the TSHR (repeats 1–8), and it does not interact with the C-terminal repeats of the LRD neither with the hinge region as the K1-70 surface is too distant from the hinge region in the complex to interact.

Mapping GPH glycosylation site candidates on the structure of TSH

The GPH β chain N-linked glycosylation site candidates mapped on the hTSH β subunit correspond to positions 1, 6, 19, 23, 56, 91, 100, 106, 110, 112 and 116 (Supplementary Alignment 7). Glycoprotein hormone α chain N-linked glycosylation site candidates were mapped at hTSH α subunit positions: 52 and 78 (Supplementary Alignment 8).
Two areas free from glycosylation are observed on the hormone surface (Fig. 4). Area A1 corresponds to the loop βL3 and parts of strands β1 and β2 of the β subunit (Fig. 1A). Area A2 corresponds to the loop αL1 and parts of strands β2 and β3 of the α subunit (Fig. 1A). Mutations at the βL3 loop, within the area A1, that introduce positive charge (I58R, E63R, and L69R) increase hTSH binding to the TSHR and hTSH bioactivity by up to 150- and 60-fold, respectively (Grossmann et al. 1998). In addition, Jiang et al. (2012, 2014) identified a potential exo-site for additional FSH/FSHR interactions based on contacts observed in the trimeric arrangement of the FSH-FSHR ECD complex in the asymmetric unit. These additional FSH-FSHR interactions include FSHβ residues Tyr58, Thr60, Asp71, Leu73 and Thr75, within area A1. Area A2, which involves the αL1 loop, contains residues that have been proposed to be responsible for the higher binding affinity of bovine (b) and porcine (p) TSH compared to hTSH for binding to the hTSHR (Szkudlinski et al. 1996). These residues are TSHα, GlN13, Glu14, Pro16 and Gln20, which are lysines in bTSH and pTSH. Szkudlinski et al. (1996) generated TSH superactive analogues by mutating hTSHα residues Q13, E14, P16 and/or Q20 to lysines.

Mapping GPHR glycosylation site candidates on the structure of the TSHR ECD

The N-linked glycosylation site candidates present in at least one species of any of the three glycoprotein

Figure 1

β Glycosylation analysis of GPHs

α Szkudlinski

Supplementary Alignment 7

et al.

Fig. 1

β α

Glycosylation analysis of GPHs

β α

Glycosylation analysis of GPHs

β α

Jiang et al. (2012, 2014) identified a potential exo-site for additional FSH/FSHR interactions based on contacts observed in the trimeric arrangement of the FSH-FSHR ECD complex in the asymmetric unit. These additional FSH-FSHR interactions include FSHβ residues Tyr58, Thr60, Asp71, Leu73 and Thr75, within area A1. Area A2, which involves the αL1 loop, contains residues that have been proposed to be responsible for the higher binding affinity of bovine (b) and porcine (p) TSH compared to hTSH for binding to the hTSHR (Szkudlinski et al. 1996). These residues are TSHα, GlN13, Glu14, Pro16 and Gln20, which are lysines in bTSH and pTSH. Szkudlinski et al. (1996) generated TSH superactive analogues by mutating hTSHα residues Q13, E14, P16 and/or Q20 to lysines.

Mapping GPHR glycosylation site candidates on the structure of the TSHR ECD

The N-linked glycosylation site candidates present in at least one species of any of the three glycoprotein

Two areas free from glycosylation are observed on the hormone surface (Fig. 4). Area A1 corresponds to the loop βL3 and parts of strands β1 and β2 of the β subunit (Fig. 1A). Area A2 corresponds to the loop αL1 and parts of strands β2 and β3 of the α subunit (Fig. 1A). Mutations at the βL3 loop, within the area A1, that introduce positive charge (I58R, E63R, and L69R) increase hTSH binding to the TSHR and hTSH bioactivity by up to 150- and 60-fold, respectively (Grossmann et al. 1998). In addition, Jiang et al. (2012, 2014) identified a potential exo-site for additional FSH/FSHR interactions based on contacts observed in the trimeric arrangement of the FSH-FSHR ECD complex in the asymmetric unit. These additional FSH-FSHR interactions include FSHβ residues Tyr58, Thr60, Asp71, Leu73 and Thr75, within area A1. Area A2, which involves the αL1 loop, contains residues that have been proposed to be responsible for the higher binding affinity of bovine (b) and porcine (p) TSH compared to hTSH for binding to the hTSHR (Szkudlinski et al. 1996). These residues are TSHα, GlN13, Glu14, Pro16 and Gln20, which are lysines in bTSH and pTSH. Szkudlinski et al. (1996) generated TSH superactive analogues by mutating hTSHα residues Q13, E14, P16 and/or Q20 to lysines.

Mapping GPHR glycosylation site candidates on the structure of the TSHR ECD

The N-linked glycosylation site candidates present in at least one species of any of the three glycoprotein
hormone receptor families mapped on the human TSH receptor correspond to positions: 30, 33, 43, 45, 53, 77, 99, 102, 113, 124, 143, 177, 198, 218, 237, 276, 288, 292, 295, 302, 312, 316, 319, 326 and 392 (Supplementary Alignment 12).

Figure 5 shows all glycosylation site candidates of the three glycoprotein hormone receptors, that appear in at least one species, mapped on the structure of the human TSHR ECD and their relative position with respect to TSH from the comparative model of the structure of the TSH-TSHR ECD complex. Candidate sites 316, 319 and 326 are not shown as these residues have not been modelled. All hormone residues that are involved in interactions with the receptor are at least 4.0 Å distant from TSHR residues that are candidates for glycosylation. The closest residue is TSHβ Glu98, which is exactly 4.0 Å from TSHR Lys58 and is a glycosylation site candidate in the FSHR sequence (position 36) of one fish, Gadus morhua (Atlantic cod). However, the sequence of Gadus morhua FSHR shows uncertainty in its alignment with the hFSHR in its N-terminal 60 residues. It is highly probable that this glycosylation site candidate may correspond to a different residue in the human sequence, and therefore, this glycosylation site in Gadus morhua FSHR was not considered a candidate in this study.

Four areas not showing glycosylation sites are identified on the surface of the hTSHR (Fig. 5). Area B1 includes the hTSHR segments R255-K261, S278-H282 and E394-I411 corresponding to TSHR segments of the convex surface of repeats 10 and 11 and the CT-cap. Area B2 corresponds to the C-terminal half of the H-loop. Area B3 includes TSHR segments of the convex surface of leucine-rich repeats 2, 3 and 4, whereas B4 includes segments of the convex surface of repeats 5, 6, 7 and 8.

Figure 5
Glycoprotein hormone receptor glycosylation site candidate residues (in red) mapped on a molecular surface representation of the structure of the human TSHR ECD. TSHα is shown in green, TSHβ in wheat and the TSHR in yellow. Candidates are present in at least one species of any of the three glycoprotein hormone receptors. Top panels are different points of view of the TSHR ECD related by rotations of ~90° along a horizontal axis. Bottom panels show the TSHR ECD in the same orientation as in the top panels but bound to TSH. N- and C-termini of the TSHR (N, C), TSHα (Nα, Cα), and TSHβ (Nβ, Cβ) are marked. Four receptor areas (B1–B4) free of glycosylation site candidates, that may be involved in additional interactions, are marked by blue ellipses. B1 includes TSHR segments of the convex surface of repeats 10 and 11 and the CT-cap. B2 corresponds to the C-terminal half of the H-loop. B3 includes TSHR segments of the convex surface of leucine-rich repeats 2, 3 and 4, whereas B4 includes segments of the convex surface of repeats 5, 6, 7 and 8. Copyright holder RSR Ltd.
involved in receptor activation. For example, TSHR S281T and S281N mutations cause increased TSHR cyclic AMP constitutive activity (Duprez et al. 1997) and the TSHR S281N mutation was reported to be associated with clinical hyperthyroidism (Gruters et al. 1998). In addition, TSHR P400A, D403E, D403L, E404K, E404L, E404N, N406A, N406D, N406L, N406Q, N406S and P407D (all contained within area B1) have been identified as constitutively activating mutants (Kleinau et al. 2004, Mueller et al. 2006), whereas TSHR D403N and D410N appear to be inactivating pathogenic mutations associated with a rare form of subclinical hypothyroidism (de Roux et al. 1996, Camilot et al. 2005). Furthermore, TSHR mutations E409K and D410K (also within B1) lead to reduction in cyclic AMP signalling (Claus et al. 2005). The FSHR ECD mutant P348K (equivalent to TSHR position 400 within B1) is inactive in functional studies, does not bind FSH and was found to be associated with primary amenorrhoea (Allen et al. 2003).

Area B2 (Fig. 5) corresponds to the C-terminal half of the H-loop. The C-terminal half of the H-loop (area B2) is free of glycosylation, whereas the N-terminal half of the H-loop shows several glycosylation site candidates. The H-loop is a good candidate for making interactions with the hormone, in particular with TSHα residues Q13, E14, P16 and Q20 found to be responsible for differences in binding affinity of hTSH, bTSH and pTSH for the TSHR (Szkudlinski et al. 1996). In addition, within B2, there are the sulphated Tyr385 that interacts with TSH (Costagliola et al. 2002, Jiang et al. 2012) and Asp382 that has been identified as making contact with bTSH (Mueller et al. 2008). In particular, TSHR substitutions D382A, D382N and D382K were reported to decrease TSH binding and have been proposed to influence TSHR signalling (Mueller et al. 2008).

Deletion of the rat TSHR segment Tyr387-Asp395 (equivalent to Tyr385-Asp395 of the hTSHR contained in B2) led to a decrease in high-affinity TSH binding
(Kosugi et al. 1991a) and the single substitution TSHR T388A (in B2 area) markedly decreased TSH binding to the receptor (Kosugi et al. 1991b). The hTSHR residues Thr388-Met396 (overlapping B2 area) were found to be important for TSH binding to the receptor and receptor cyclic AMP activity (Nagayama & Rapoport 1992). Furthermore, the TSHR germline mutation N372T (in B2) is responsible for clinical hyperthyroidism (Goza et al. 2008). In addition, TSHR residues between Lys371 and His384 (in B2) have been proposed to be involved in extracellular domain inverse agonistic activity (Mizutori et al. 2008).

Areas B3 and B4 (Fig. 5) are located in the convex surface of the LRD of the receptor and could be involved in dimerisation/oligomerisation. Dimerisation/oligomerisation of GPHRs has been reported (Chazenbalk et al. 1996, Graves et al. 1996, Latif et al. 2001, 2010, 2014, Fan & Hendrickson 2005, Urizar et al. 2005, Jiang et al. 2012, 2014, Jonas et al. 2015). As mentioned previously, Jiang et al. (2012, 2014) identified a potential exo-site for additional FSH/FSHR interactions that involve the FSHR residues Arg59, Glu84, Glu87 and Tyr110. These residues correspond to TSHR Arg65, Gln90, Glu93 and Tyr116, which are located within area B3. A study of extracellular domain multimer formation in which the extracellular domain of the TSHR linked to a glycoprophosidyl anchor was expressed, indicated that TSHR mutation Y116S (in B3) abrogated the observed multimer formation (Latif et al. 2010). FSHR-inactivating mutations I160T and A189V (equivalent to TSHR positions 167 and 196, respectively, within area B4) are associated with hypergonadotrophic ovarian dysgenesis and secondary amenorrhea, respectively (Aittomäki et al. 1995, Beau et al. 1998). In vitro analysis showed that in both cases, FSH binding and FSH-induced cyclic AMP production were decreased as a result of the mutations. In addition, LHR mutant F194V (equivalent to TSHR 196, within area B4) did not respond to hCG in vitro (Gromoll et al. 2002). These observations indicate that area B4 could have a role in receptor dimerisation/trimerisation linked to modulation of constitutive and/or stimulated activation of the GPHRs (Urizar et al. 2005, Jiang et al. 2014).

Discussion

The structure of the FSHR ECD in complex with FSH (Jiang et al. 2012) suggests that the second step in GPHR activation, after hormone binding to the receptor LRD could be a movement of the receptor hairpin loop (H-loop) towards the hormone, driven by an electrostatic interaction between the sulphated tyrosine present in the receptor H-loop and a positively charged pocket at the interface of the α- and β-subunits of the hormone.

Models of the structure of the TSH-TSHR ECD complex obtained by us (presented here) and by others (Krause et al. 2012, Davies et al. 2014, Schaarschmidt et al. 2014) show that it is likely that activation of the TSHR by TSH proceeds via a similar mechanism. Human TSHα residues Gln13, Glu14, Pro16 and Gln20 are located in the TSH loop αL1 (area A2, Fig. 1A). Mutation of all four residues to lysines results in the increase of binding affinity of hTSH to the hTSHR (Szkudlinski et al. 1996, Szkudlinski 2015). Jiang et al. (2012) proposed, based on the crystal structure of the FSH-FSHR ECD complex, that the TSHα mutations Q13K, E14K, P16K and Q20K are concentrated at the top side of the sulphated-Tyr binding pocket of the hormone, generating additional positive charges for a stronger interaction with the sulphated-Tyr of the receptor. Our model of TSH-TSHR ECD also predicts an increase of the positive charge of the pocket between TSHα and β chain by those hTSHα mutations that would increase the attraction of the negatively charged TSH sulphated Tyr385 in the H-loop (in area B2) towards the hormone. In contrast, Krause et al. (2012) proposed, based on their model of the bTSH-hTSHR ECD, that the side chains of positively charged bTSHα residues (Lys13, Lys14, Lys16 and Lys20) interact with the negatively charged side chains of TSHR residues Glu303 and Asp382 located at the N- and C-terminus of the TSHR hinge region, respectively. TSHR residues Glu303 and Asp382 are located in area B2 which has been identified in our study for possible interactions. Although it is likely that the H-loop that includes area B2, would make contacts with the positive residues from the α chain of bovine TSH, these interactions cannot be predicted by our model based on human TSH as the α chain of hTSH has no positive residues at positions 13, 14, 16 and 20.

Conformational changes of the ligand upon receptor binding, as the one observed during FSHR are not unusual (Watanabe et al. 2002, Taylor et al. 2014). For example, changes in conformation of insulin on receptor binding have been recently reported based on crystallographic, NMR and comparative modelling analysis (Menting et al. 2014). In more detail, the C-terminal β-turn (residues B20–B23) and β-strand (B24–B27) of insulin undergoes hinge-like rotation away from its α-helical core to allow its conserved aromatic motif (Pheβ24, Pheβ25, Tyrβ26) to bind to the primary hormone-binding site (α-subunit domains L1 and αCT) of the insulin receptor (Menting et al. 2014).
Study of glycosylation patterns in a protein family or superfamily is an elegant method to identify new active or binding sites within the family. The functionally important active or binding sites should be free of glycosylation in all members of the family/superfamily assuming that the function is conserved within the family/superfamily. A different but related approach based on glycan wedge scanning was successfully used to investigate how the dimeric ECD of the GABA<sub>B</sub> receptor controls receptor activity (Rondard et al. 2008).

Our study has shown that FSHR amino acid sequences from all mammals, birds, reptiles and an amphibian had a potential glycosylation site at Asn199-XXX-Ser201, whereas this was absent at the equivalent positions in fish. Deglycosylation studies of rat FSHR indicated that Asn191 and Asn293 are glycosylated but Asn199 is not (Davis et al. 1995, Wheatley & Hawtin 1999). Asn199 is located in the concave surface of the 8th LRR and although Asn199 is not contacting the hormone in the crystal structure of the human FSH-FSHR ECD complex, a bulky carbohydrate at this position would most likely reduce FSH binding to the receptor. Further, because both Asn199 and Ser201, forming the consensus sequence Asn-XXX-Ser/Thr, are exposed, another spatial restraint may prevent Asn199 glycosylation. Consequently it is unlikely that Asn 199 is actually glycosylated in the FSHR molecule.

No glycosylation was observed in either LRR 12 or the CT-cap of the ECD in any receptor of any species studied. LRR 12 has only a short amino acid sequence and its structure is facing the hormone in the TSH-TSHR ECD model; however, at a distance too long for making contact. LRR 12 is the last repeat of the ECD and may contact the extracellular loops of the TMD. Glycosylation of any of the residues of the 12th LRR would introduce steric hindrance with the TMD due to the bulky structures of sugar molecules. The approximately ten-residue long CT-cap segment is also likely to interact with the extracellular loops of the TMD, as described above, and the presence of glycans would be disruptive for any such interactions. The TSHR sequences of three fishes show a glycosylation site candidate equivalent to hTSHR Asp276 located in the concave surface of repeat 11 of the TSHR ECD (Fig. 4). This region of the TSHR LRD does not interact with the hormone in our model and the closest hormone residue to TSHR Asp276 is TSHα Ser92 at 11.6 Å.

In our model of the TSHR ECD, the Asn302 glycosylation site is present in the H-loop, and it is close in space to the TSHR LRD and the TSHR TMD favouring glycan contacts with the ligands or the TMD.

Alternatively, glycosylated Asp302 may be prevented from making contacts with the ligands or the TMD as the H-loop is unstructured and flexible. Overall our model cannot discern whether the glycans located in the H-loop of the GPHRs (equivalent to hTSHR Asp302) may or may not make contacts with ligands and/or the TMD.

The crystal structures of the M22-TSHR LRD complex (Sanders et al. 2007) and the K1-70-TSHR LRD complex (Sanders et al. 2011) show that the sugar molecules present in the hTSHR LRD do not interact with the antibodies. From the glycosylation-site candidates of GPHRs found in this study, only the one at TSHR position 276 is likely to contact the stimulating antibodies but not the blocking antibodies (Fig. 6) in the comparative models of the M22-TSHR ECD and K1-70-TSHR ECD complexes. However, the TSHR residue at position 276 is only glycosylated in three fishes (Supplementary Table 8), in the LHR, the equivalent residue is glycosylated in one fish (Supplementary Table 7), and in the FSHR, the equivalent residue is glycosylated in horse, kangaroo, birds and 9 fishes (Supplementary Table 6). Autoantibodies to the hTSHR are known to stimulate the guinea pig thyroid (Adams & Purves 1956, Davies et al. 1978), the mouse thyroid (McKenzie 1958), the rat thyroid (Kosugi et al. 1991a,b, 1993, Furmaniak et al. 2013) and, probably, TSHR from other species sharing high homologous amino acid sequences. However, TSHR autoantibodies are unlikely to bind to the TSHR of species with lower amino acid homology. Furthermore, autoantibodies to the hTSHR are unlikely to bind to the TSHR of those fishes that show a glycosylation site candidate at position 276 due to steric hindrance with the sugar molecule, if the candidate site is actually glycosylated.

Our study confirms that the areas on the surfaces of GPHs that interact with their cognate receptors (Fig. 4) are free from glycosylation in each of the four hormones and for all species. In addition, the concave faces of LRR one to ten of the receptor LRDs are free from glycosylation sites, in each of the three receptors and for all species (Fig. 5). This is consistent with these repeats showing the strongest interactions with the hormones. Accordingly, interactions of glycans of one component of the hormone-receptor complex with the other component of the complex are unlikely or weak. This study does not predict stabilisation or destabilisation of the binding of glycoprotein hormones to their cognate receptors by their glycans. By mapping glycosylation patterns of GPHs and GPHRs on the comparative models of the structures of the TSH-TSHR...
ECD, M22-TSHR ECD and K1-70-TSHR ECD complexes, we have been able to identify two areas on GPHs free from glycosylation, that are not involved in currently known interactions. These areas are good candidates for making additional contacts with the GPHRs. In the case of the GPHR ECDs, four areas not showing glycosylation sites, that may make additional contacts, are identified on the surface of the receptors. Two areas are located in the convex surface, one in the long H-loop and one corresponds to the C-terminus of the extracellular domain. These areas are candidates for additional (currently unknown) contacts between the receptors and their cognate hormones.

 Supplementary data

This is linked to the online version of the paper at http://dx.doi.org/10.1530/JME-16-0169.

 Declaration of interest

All authors are employees of RSR Limited, a developer and manufacturer of diagnostic kits, including kits for measuring thyroid autoantibodies.

 Funding

The work was funded by RSR Limited.

 References

Adams DD & Purves HD 1956 Abnormal responses in the assay of thyrotropin. Proceedings of the University of Otago Medical School 34 11–12.


Bousfield GR, Butnev VY, Butnev VV, Hiramasa Y, Harvey DJ & May JV 2014 Hypo-glycosylated human follicle-stimulating hormone (hFSH(21/18)) is much more active in vitro than fully-glycosylated hFSH (hFSH(24)). Molecular and Cellular Endocrinology 382 989–997. (doi:10.1016/j.mce.2013.11.008)


Davies TF, Rees Smith B & Hall R 1978 Binding of thyroid stimulators to guinea pig testis and thyroid. Endocrinology 103 6–10. (doi:10.1210/endo-103-1-6)


Davis DP, Liu X & Segaloff DL 1995 Identification of the sites of N-linked glycosylation on the follicle-stimulating hormone (FSH) receptor and assessment of their role in FSH receptor function. Molecular Endocrinology 9 159–170. (doi:10.1210/me.9.2.159)

Davis DP, Rozell TG, Liu X & Segaloff DL 1997 The six N-linked carbohydrates of the lutropin/choriogonadotropin receptor are not absolutely required for correct folding, cell surface expression, hormone binding, or signal transduction. Molecular Endocrinology 11 550–562. (doi:10.1210/mend.11.5.9927)


Fox KM, Dias JA & vanRoey P 2001 Three-dimensional structure of human follicle-stimulating hormone. Molecular Endocrinology 15 378–389. (doi:10.1210/endo.15.3.0603)


mutation of the TSH receptor: coexpression in a hyperthyroid family member with a second activating somatic mutation. Thyroid 18 499–508. (doi:10.1089/thy.2007.0335)


Grossmann M, Weintraub BD & Szukdalski MW 1997 Novel insights into the molecular mechanisms of human thyrotropin action: structural, and therapeutic implications for the glycoprotein hormone family. Endocrine Reviews 18 476–501. (doi:10.1210/edrv.18.4.0305)


Further characterization of thyroid-stimulating antibodies from Graves’ patients. Biochemical and Biophysical Research Communications 180 1118–1124. (doi:10.1016/S0006-291X(05)81182-9)

Kosugi S, Ban T & Kohn LD 1993 Identification of thyroid-stimulating antibody-specific interaction sites in the N-terminal region of the thyrotropin receptor. Molecular Endocrinology 7 114–130. (doi:10.1210/mend.7.1.8095322)

Krause G, Kreuchwig A & Kleinau G 2012 Extended and structurally supported insights into extracellular hormone binding, signal transduction and organization of the thyrotropin receptor. PLoS ONE 7 e59290. (doi:10.1371/journal.pone.0059290)


Mizutori Y, Chen CR, McClachlan SM & Rapoport B 2008 The thyrotropin receptor hinge region is not simply a scaffold for the leucine-rich domain but contributes to ligand binding and signal transduction. Molecular Endocrinology 22 1171–1182. (doi:10.1210/me.2007-0407)


Mueller S, Kleinau G, Jaeschke H, Paschke R & Krause G 2008 Extended hormone binding site of the human thyroid stimulating hormone receptor: distinctive acid residues in the hinge region are involved in bovine thyroid stimulating hormone binding and receptor activation. Journal of Biological Chemistry 283 18048–18055. (doi:10.1074/jbc.M800449200)


Glycosylation analysis of GPHs and receptors

Research

R NÚÑEZ MIGUEL and others

58:1

51

Differences in interactions of thyroid stimulating and blocking autoantibodies with the TSH receptor. Journal of Molecular Endocrinology 49 137–151. (doi:10.1530/jme-12-0040)


Segaloff DL & Ascoli M 1993 The lutropin/choriogonadotropin receptor... 4 years later. Endocrine Reviews 14 324–347. (doi:10.1210/er.14.3.324)


Watanabe K, Chirgadze DY, Lietha D, de Jonge H, Blundell TL & Gherardi E 2002 A new crystal form of the NK1 splice variant of HGF/SF demonstrates extensive hinge movement and suggests that the NK1 dimer originates by domain swapping. Journal of Molecular Biology 319 283–288. (doi:10.1016/S0022-2836(02)00199-7)


Received in final form 3 November 2016
Accepted 13 November 2016
Accepted Preprint published online 14 November 2016