4 DISCUSSION

A detailed knowledge of molecular activation mechanisms of the TSH receptor by investigation of structure-function relationships is of vital importance to understanding the molecular reasons of frequently occurring receptor dysfunctions (Schöneberg T 2002, Seifert R 2002).

Moreover, the elucidation of structurally localized patterns for inactivation and activation of the TSHR does not require an understanding of the complete activation mechanism, but is helpful in revealing new ideas and sites for pharmacological interventions and additionally opens the door for new therapeutic perspectives treating the TSHR mediated diseases of the thyroid directly at the TSHR.

The aim of this study was to provide conceptual advance in the understanding at the molecular level, of how the activation signal is transduced between the large extracellular ectodomain into the transmembrane domain after extracellular hormone binding. The use of a combination of site directed mutagenesis and molecular homology modeling crosses the border between physiological and structural aspects. These studies are focused on the regions potentially involved such as the leucine rich repeat domain, epitopes of the hinge region, the extracellular loops and the extracellular portions of the transmembrane helices. Additionally, insights into the signalling process have been revealed by docking of a small molecular weight ligand that activates the TSHR and the homologous LHCGR. In summary, for the first time different determinants and mechanisms in these regions that are important for the signal transduction and/or the regulation of different activity states were identified and described.

4.1 A tightly packed signalling interface between the extracellular- and the serpentine domain

4.1.1 Structural-functional features of the N-terminal cysteine-box 1 and the LRR hormone binding domain

We modeled the 3D structure of the LRRD of all three human GPHRs (TSHR, FSHR, LHCGR) based on the hNogo-receptor ectodomain (Kleinau G 2004). Compared to the previously suggested GPHR LRRD models (Moyle WR 1995, Jiang X 1995, Bhownick N 1996, Smits G 2002, 2003) this molecular model was characterized by several new features, like a fold stabilizing integrated N-terminal cysteine-box 1 and the enlarged radius of the LRRD (Figure 3.2). In 2005 the X-ray structure of the FSHR LRRD/FSH complex was published (Fan QR 2005). Our GPHR LRRD models provided in 2004 and based on the

hNogo-receptor template showed very high structural similarities to the crystal structure of the FSHR LRR domain and confirmed the general and detailed structural predictions for the GPHR LRRDs, which was also discussed by others (Vischer HF 2006, Fan QR 2006).

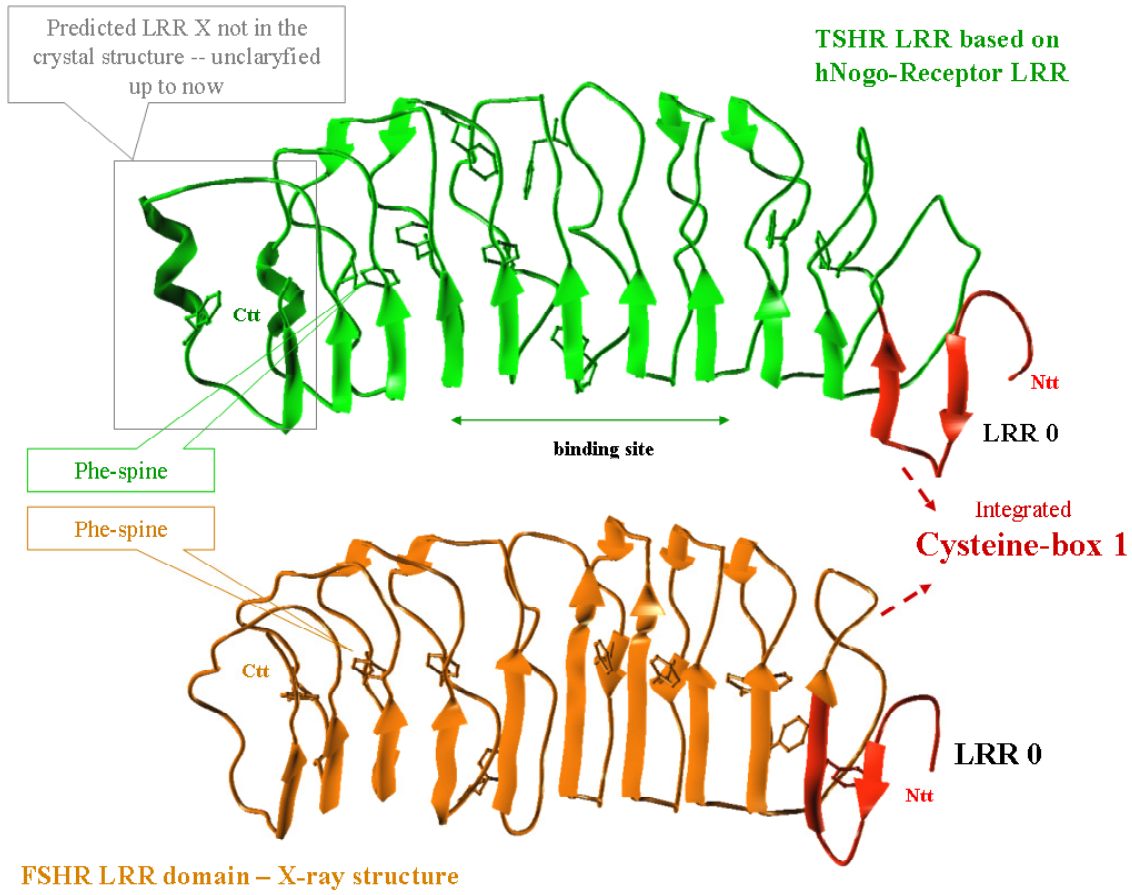


Figure 4.1: ***Comparison between the molecular model of the TSHR LRRD based on the hNogo-receptor structure and the FSHR LRRD crystal structure***
The X-ray structure of the FSHR LRR domain confirmed structural predictions for the GPHR LRR domains based on the structural template of the hNogo-receptor ectodomain including the N-terminal cysteine-box similar to the GPHR cysteine-box 1.

Our molecular homology model as well as the X-ray structure of the FSHR LRRD reveal the amino acids of cysteine-box 1 as an integral part of the LRRD stabilizing the N-terminal fold via disulfide bridges (Figure 4.1). The 4 cysteines interacting via two disulfide bridges and C-b1 formes an antiparallel and a parallel β-strand that are involved in the interaction with the hormone (Fan QR 2005 (b), 2007).

The structure-stabilizing function of the helices observed for LRRDs identified for many proteins (Enkhbayar P 2003) is partially adopted by aromatic interactions of phenylalanines ('Phe-spine') in the interior of the LRRD fold (Figure 4.1).

An eleventh repeat, comprising amino acids 258 to 278 at the C-terminal site of the LRRD, does not appear in the FSHR LRRD X-ray structure, but can be predicted based on the hNogo-receptor LRRD structure and the amino acid sequence similarity to the TSHR. This region connecting the LRRD and the C-b2, which most probably stabilizes the C-terminal LRRD fold follows the general architecture of this class of LRRDs (Kajava AV 2002).

4.1.2 The LRRD is oriented in spatial proximity to the serpentine domain and linked via cysteine-box 2

A tightly packed molecular model of selected extracellular components was built by extensive consideration of known functional data and 3D models of the LRRD, the C-b2 including S281 and a C-terminal epitope of C-b3 based on homologous structures (Results 3.1.1-3) (Kleinau G 2004).

The structural component C-b2 is attached back-to-back to the 11th β-strand of the LRRD, very likely via a short turn/loop containing S281 (Figure 3.4) (Nakabayashi K 2003, Jäschke H 2006 (a)). Irrespective of the question whether the homologous turn configuration suggested by sequence identity is assembled in the right orientation, a turn/loop structure very likely organizes how the LRR is spatially arranged (e.g. transversal or lateral) relative to C-b2/C-b3 and to the extracellular loops of the serpentine domain. Their mutual orientation obviously has a function as a potential pivot or hinge between them. This is consistent with data based on mutational activation at position S281 (Duprez L 1997, Nakabayashi K 2000), which indicate that minor conformational changes in the S281 turn/loop activate the receptor. This may support the idea of a mechanism in which a certain spatial adjustment between the LRRD and the following C-b2 (and remaining components of the ectodomain) via the S281 turn/loop is responsible for TSHR activation, including the first steps of signalling caused by hormone binding to the LRRD.

4.1.3 Epitopes Y279-K291 of cysteine-box 2 and P400-D410 of cysteine-box 3 are localized at the interface between the ecto- and serpentine domain

Based on new structural features for the hinge region, identified by sequence homology with a complex of the IL8 ligand and the IL8RA peptide, the hypothesis was generated that the structural components C-b2 and C-b3 interact and are located close to TMH1 (Results 3.1.2-3). One consequence of our model is a potential disulfide bridge between C283 and C408 or C284 and C408. Our findings are in agreement with suggestions of disulfide bridges between

C-b2 and C-b3 based on biochemical data (Ciullo I 2003, Chen CR 2003 (b)). Moreover, since the C-terminal end of the LRRD is connected to C-b2 via the S281 turn/loop and C-b3 is connected to the transmembrane domain TMH1, a tightly packed structure including the C-terminal end of the LRRD, the C-b2 and the C-b3 of the ectodomain was postulated (Figure 3.4).

The epitopes Y279-K291 and P400-D410 of C-b2 and 3 of the ectodomain are structurally localized at the interface between the ecto- and serpentine domains and were therefore proposed as an important mediator for the intramolecular signal transduction from the ectodomain to the serpentine domain or as fundamental for the stabilization of the basal conformation. In the assembled models the two linked C-b2/C-b3 motifs are arranged across the serpentine domain, despite being constrained by the fact that they are connected to TMH1. This hypothesis lead to mutagenesis studies of several amino acids in these epitopes which have identified new key residues participating in intramolecular signal transduction at the interface to the serpentine domain. Using a combined modelling/mutagenesis approach, the functional importance of the N-terminal epitopes at extracellular C-b2 (Y279-K291) and within C-b3 (P400-D410) for the intramolecular signalling processes in the TSHR were demonstrated (Kleinau G 2004, Müller S & Kleinau G 2006).

4.1.4 Intramolecular switches for signalling activity in the TSHR ectodomain

4.1.4.1 Signalling sensitive amino acids in cysteine-box 3

Three amino acids in the ectodomain (D403, E404, N406) in proximity to TMH1, can act as switches for constitutive activity of the TSHR induced by point-mutations (Tables 3.2-5).

D403 – Mutation of D403 with a smaller neutral amino acid (alanine) causes a CAM very likely by loss of side chain interaction. In contrast a lysine mutant at position D403 with a more bulky and basic side chain strongly disturbs the receptor expression probably by a bulky/electrostatic repulsion, which obviously only has an effect on the overall fold but no effect on basal activity. Substitutions of D403 side chain to the bulky hydrophobic amino acid leucine, and even to the charge keeping glutamate but extending the side chain length, showed high constitutive activation, whereas substitutions with reduced side chain length such as asparagine and serine do not exhibit constitutive activity. Therefore, the hypothesis of a tight spatial environment by van der Waals interactions surrounding D403, which would explain that smaller amino acids are tolerated, and maintain the basal wt TSHR conformation at this position was derived (Figure 4.2).

E404 - At position E404 substitutions to amino acids of variable characteristics except to aspartate are CAMs. Thus for E404 a H-donator as interaction partner is assumed (Figure 4.2).

N406 - All substitutions at N406 except the mutant N406K are CAMs, implying that the interaction partner of N406 should provide a negative charge or at least a strong capacity of accepting H-bond(s) (Figure 4.2).

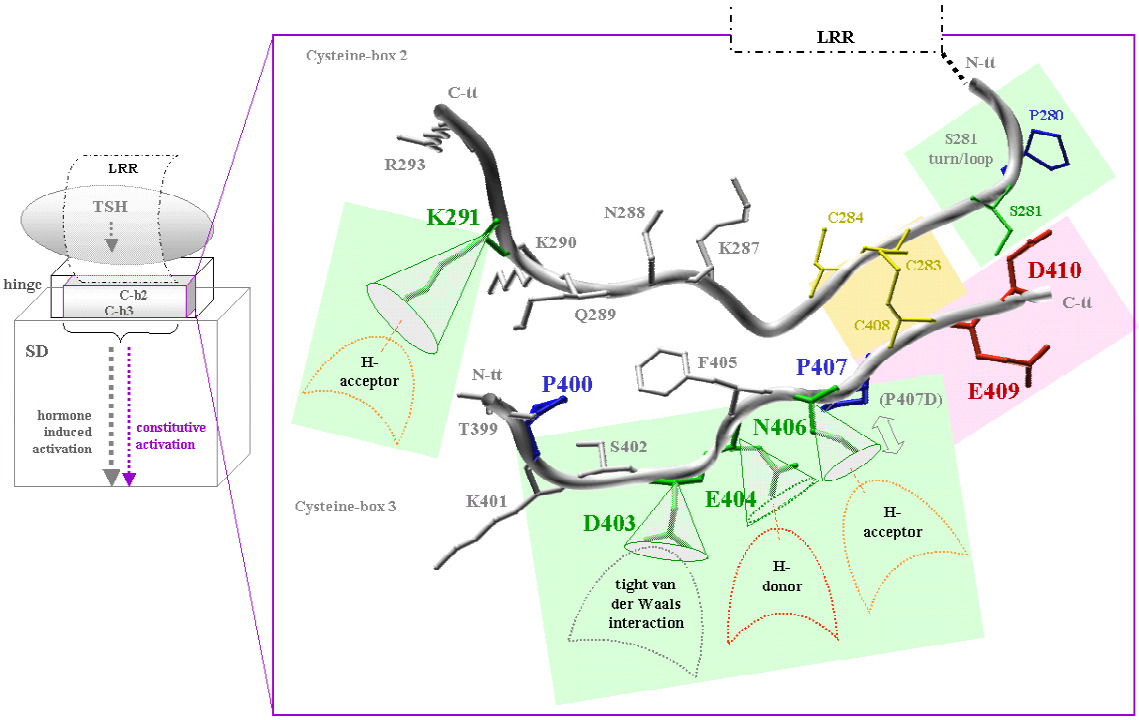


Figure 4.2: ***Backbone model for the 'C-b2/C-b3' - micro domain of the TSHR***
Predicted and confirmed key residues maintaining the basal wt conformation (green boxes) and transducing the activation signal as well (red box) are colored, which indicates their major functional phenotype as determined by mutagenesis (green: CAM, red: inactive, blue: prolines effects on backbone fold (partialy inactive mutations and constitutive activation), white: no effect on signalling, yellow: cysteines). Approximate orientations of essential side chains mutated in this study are shown as cones and the properties of their potential interacting counterparts are color-coded (orange dotted: H-bond acceptors; red dotted: H-bond donators; grey dotted: tight van der Waals interactions). Beside the disulfide bridge (C408/C283 or C284, yellow boxed) as a known conformational constraint, further effects on the backbone conformation were identified for prolines P400, and P407 (shown in blue). The amino acids P407, E409, and D410 are essential for the stabilization of the active conformation after hormone interaction. The identification of K291 at Cb-2 as a new CAM provides the second conformational constraint at Cb-2 for the basal state and refines the known tryptic cleavage site in this region.

Taken together, the results of this study show that constitutive activation at D403, E404 and N406 seems to be initiated by disruption of a constraining hydrogen bond network or a

destabilization of a tightly packed environment surrounding C-b3 that is necessary for the maintenance of the basal state of the wt TSHR (Figure 4.2).

Due to a tightly packed structural arrangement of C-b2 and C-b3 via disulfide bridges it was hypothesized that apart from S281 in C-b2 and the CAM region D403 E404 N406 in C-b3 additional amino acids could also be sensitive for constitutive activation by mutation.

Indeed, mutations P400A and P407D in C-b3 constitutively activate the TSHR (Tables 3.4-5). We conclud that the constitutive activation caused by the P400A mutant is based most probably on a structural shift of adjacent residues that are involved in the stabilization of the basal TSHR conformation. Furthermore it can be assumed, that the constitutive activation observed for mutant P407D is based on repulsion of a H-acceptor moiety (dotted arrow in Figure 4.2), which we already characterized as an interaction partner for N406. It is most likely localized in close spatial neighbourhood to P407.

Mutation E409A is characterized by a strongly decreased capability for cAMP mediated signalling. Furthermore, mutations E409K and D410K are characterized by a cell surface expression of 63% and 90% respectively of the wt TSHR but a complete loss of cAMP mediated signalling. Therefore, E409 and D410 are necessary to maintain the hormone induced active conformation of GPHRs. These amino acids are named ‘activator component’ (red boxed in Figure 4.2). Mutants at P407 with decreased activity but reasonable cell surface expression indicate the ability of P407 to stabilize the hormone induced active state very likely via a backbone orientation that is suitable for the activator component comprising E409 and D410.

4.1.4.2 Constitutively activating mutations in cysteine-box 2

S281 - The sequence of the C-b2 region around S281 is highly conserved among glycoprotein hormone receptors and constitutive activation of the TSHR by pathogenic mutations at S281 (N,T,I) (Duprez L 1997, Kopp P 1997) demonstrates a pivotal role of this residue for intramolecular signal transduction. The functional importance of this micro-domain is supported by the vicinity of S281 to the C-b2 epitope, in which mutations at positions P280, C283, and C284 also lead to constitutive activity (Ho SC 2001, Zheng H 2001) demonstrating that the C-b2 epitope 279 YPSHCC 284 can act as an intramolecular switch for receptor activation (Ho SC 2001).

Our group investigated all possible natural amino acid substitutions at position S281 and their effects on receptor activation (Table 3.6) (Jäschke H 2006 (a)) to reveal deeper insights into

the activation mechanism at this position and to confirm the relative location of S281 suggested by our structural model of C-b2 and C-b3 (chapter 4.1.3).
Surprisingly, not only S281 substitutions by small residues A and G were well tolerated at this position, but bulky aromatic amino acids W, F, Y, and H also showed high expression levels and only a slightly increased basal activity. The small side chain of alanine at position 281 does not yield constitutive activity indicating that it interferes neither with the inactive turn/loop conformation of the S281 region nor with the surrounding spatial vicinity.
Our 3-D model (Figure 3.5) suggests that the side chains of aromatic residues at position 281 sit at an angled orientation and additionally they lack a bulky branch at the Cβ-atom. Furthermore, the large ring systems are directed into a pocket, which accepts aromatic amino acids well. This data support the fact that the native side chain S281 facilitates the turn/loop conformation rather than that it directly interacts with a partner in the serpentine domain and that the S281 microdomain is embedded in an aromatic vicinity.
Taken together, our study provides evidence for (i) a tight conformational packing of the S281 turn/loop region, which tolerates only very small or aromatic side chains at position 281, and (ii) an aromatic vicinity of S281 region, which is necessary for receptor trafficking and signalling. (iii) Our refined TSHR model and mutagenesis data at the TMH2/ECL1 junction suggest spatial proximity of the micro-domain S281 to the junction of TMH2 and ECL1.

K291 - Chen et al., 2003 (Chen CR 2003 (a)) have shown that a cluster of positively charged hydrophilic amino acids in C-b2 forms the site for tryptic clipping, which also results in partial receptor activation. Our structural model suggests that the region K287-K291 of C-b2 is located in spatial proximity to the CAM region D403 E404 N406 of C-b3 (Figure 4.3). Based on the combination of the data obtained from tryptic cleavage and our own results, conformational anchor point(s) within the C-b2 region can be assumed. This hypothesis was confirmed by identification of K291A as a CAM in C-b2 (Table 3.1) (Müller S & Kleinau G 2006).

4.1.5 Modulation of signalling activity by mutations of amino acids in cysteine-box 2 and cysteine-box 3

Our data (4.1.1-4) provide i) properties of potential interaction partners for the amino acids D403, E404, N406 recently identified as CAM positions; ii) evidence that adjacent prolines (P400, P407) are important for stabilizing the basal wt conformation of the TSHR. These two prolines are also key molecular players for the receptor activation process at C-b2 and C-b3

induced by TSH and provide a backbone orientation that is required for a iii) activator component, which very likely comprises E409 and D410 at the extreme C-terminal hinge region. Furthermore the structural-functional characterization of the disulfide linked C-b2/C-b3 epitopes as an interface between the ECD and SD was completed by identification of a iv) new anchor point at C-b2 (K291) and v) a detailed description of the vicinity around signalling sensitive amino acid S281 at C-b2. Based on our molecular model driven characterization of mutants an intramolecular signalling region at the junction of the N-terminal ECD with the SD can be outlined now.

The epitope P280 and S281 on the N-terminal side of C-b2, for which a mechanism of receptor activation by constitutively active mutants via the disturbance of structural constraints has been suggested (Nakabayashi K 2003, Jäschke H 2006 (a)), was named the anchor I fragment (green boxed in Figures 4.2 and 4.3). On the other hand five amino acids P400, D403, E404, N406, P407 at C-b3 can activate the TSHR through mutations causing release of constraints responsible for stabilization of the basal wt TSHR conformation. This corresponds to a region named the anchor II fragment (green boxed, Figures 4.2 and 4.3).

Mutations of amino acids P407, E409, and D410 (C-b3) have a strong effect on hormone induced signalling by partially inactivating the TSHR towards hormone induced receptor activation. P407 stabilizes most likely the hormone induced active state, very likely via a backbone orientation that is suitable for the activator component comprising E409 and D410.

Furthermore, the alanine mutant of K291 as a new CAM in C-b2 was identified by our group. The shared effect of receptor activation by tryptic cleavage at a cluster between K287- R293 or single mutations of the positively charged amino acid K291 at C-b2 leads in both cases to a release of the constrained portion at C-b2. Here, a conformational anchor for C-b2 (K291) was identified, which is named the anchor III fragment (green boxed in Figure 4.3).

The CAM positions described in C-b2/C-b3 give an image of conformationally anchored fragments (Figures 4.2, 4.3). Moreover, the side chain mutagenesis studies narrow down the characteristics of the potential interaction partners for this region by characterizing their properties.

Based on our findings, the following scenario for a molecular activation mechanism mediated by C-b2 and C-b3 is suggested.

The three anchor fragments and one activator component are tightly packed and flank the disulfide bridge, which holds C-b2 and C-b3 tightly together. Obviously, only the epitope K287- K290 and R293 is accessible, since mutations have no effects on signalling, which is consistent with a reported tryptic cleavage region (Chen CR 2003) (Figures 4.2, 4.3).

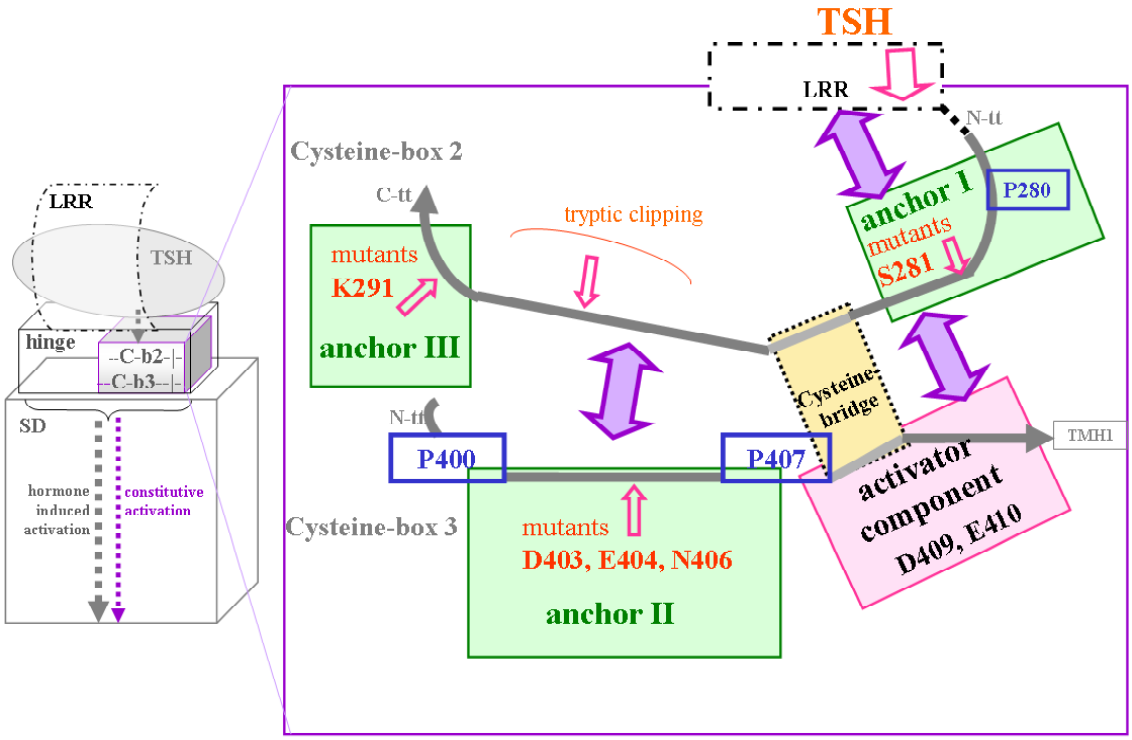


Figure 4.3: ***Suggested mechanism of an internal signal transmitter fragment***
Cartoon of three anchored fragments (green boxed) and of a component directly participating in the hormone induced activation process (red boxed). Regardless of which types of conformational interferences act on the constrained wt C-b2/C-b3 conformation, such as release of one of the anchors by side chain modifications, tryptic cleavage or hormone stimulation (red open arrows), in all cases there follows a conformational displacement of the sensitive C-b2/C-b3 portions against one other (violet double arrows), thus affecting the activator component. The cysteine bridge plays a pivotal role as a fulcrum, while the three highly conserved prolines (blue) assist this process by their defined backbone conformation.

All three anchor fragments stabilize the basal TSHR conformation. It is very likely, that the three highly conserved prolines (blue at Figures 4.2, 4.3) assist this process through their defined backbone conformation. Anchor fragments II and III are firmly locked by H-bond interactions. Anchor fragment I is held by tight 'knob and hole' van der Waals interactions with its environment (Jäschke H 2006). Since anchor fragment I immediately follows the LRR domain harbouring the hormone binding site, it is conceivable that a conformational change of the LRR upon hormone binding could be transferred to anchor fragment I. Such a conformational change could subsequently displace the identified sensitive C-b2/C-b3 portions relative to each other (violet double arrows in Figure 4.3). Regardless of the type of conformational interferences on the tightly constrained C-b2/C-b3 wt conformation, such as side chain modifications, tryptic cleavage or hormone stimulation (red open arrows Figure 4.3), they always release one of the anchor fragments. In all cases this would be followed by a conformational displacement of the sensitive C-b2/C-b3 portions against one other, which in

turn could result in receptor activation. In this process, the cysteine bridge plays a pivotal role as a fulcrum.

In summary, most likely these epitopes which have been characterized and the signalling determinants which were identified, are part of an intramolecular signal transmitter that is located at the interface between the ECD and SD. This intramolecular signal transmitter is fundamentaly involved in the signal transduction process and the regulation of the activity states of the TSHR.

4.1.6 Section summary

To identify determinants at the TSHR ectodomain that may be involved in signal transduction, this issue was first addressed by searching for homologous structural features. Based on high sequence similarity to the determined structures of the hNogo-receptor ectodomain and the intermolecular complex of the Interleukin-8 ligand (IL8) and the N-terminal peptide of the IL8 receptor (IL8RA) the hypothesis was developed that portions of the intramolecular components C-b2 and C-b3 of the TSHR ectodomain are interacting in tight spatial cooperation and are localized at the interface between the ecto- and serpentine domains. Indeed, point mutations within the D403EFN406 motif at C-b3 resulted in increased basal cAMP levels suggesting that this motif may be important for the transduction of the signal from the ectodomain to the transmembrane domain. Based on our structure-function studies a model is presented for properties of potential interaction partners for this region. Moreover, it is shown that P400 and P407 adjacent to this epitope are also important for stabilizing the partially active, basal conformation of the wild type TSHR. Additionally, the mutation K291A in the extracellular C-b2 was identified as a new CAM, which releases the basal conformation of the wild type receptor like the known tryptic cleavage in its close vicinity.

Taken together, an activation scenario at the C-b2/C-b3 unit can be provided. Three tightly packed anchor fragments (anchors I-III) most likely constrain the basal conformation. A disulfide bridge holds the C-b2/C-b3 portions in close proximity. Independent of the type of conformational interference - side chain modifications, tryptic cleavage or hormone stimulation - that acts on the constrained C-b2/C-b3 wild type conformation, it will always release one of the anchor fragments. Subsequently, this results in a conformational displacement of the C-b2/C-b3 portions relative to each other, inducing receptor activation.

4.2 A fundamental role of the extracellular loop 2 in the signal transduction process

The extracellular loop 2 is located in the centre at the junction between the ectodomain and the transmembrane domain of the TSHR. For the first time our group systematically characterized the amino acids within the ECL2 of the TSHR by site-directed mutagenesis to provide deeper insights into the functional and structural role of ECL2 in the processes of TSHR activation and intramolecular signal transduction (Kleinau G 2006 (a)). Furthermore, this study was aimed at revealing detailed insights into the activation mechanism of pathogenic constitutive activating mutations which have been identified at amino acid I568 in ECL2 (Claus M 2005).

4.2.1 Lysine 565 in the ECL2 is a key player in intramolecular signalling processes of the TSHR

The alanine scan of ECL2 revealed that the K565A mutation in ECL2 is basally inactive and shows an impaired TSH-induced response (Table 3.8).

We suggest two scenarios, which might explain the low activity of K565A: I. The cascade of constituents involved in the signal transmission process starting from the ligand-occupied ECD to intracellular effectors is interrupted by mutation of K565 via the breaking of hydrogen bonds or of electrostatic salt bridge interactions. II. K565 is important for the formation of the active receptor state conformation by binding to a new interaction partner after hormone induced signal initiation. The mutant is unable to stabilize the active receptor structurally and/or functionally.

The recently published mutants K660D (TMH7), E409K and D410K (extreme C-terminus of the ECD) (Claus M 2006) exhibit similar phenotypes. Our suggested inactivation mechanisms for mutant K565A at ECL2 can also be applied to these mutants and a functional interplay in the wt receptor during the hormone induced activation process between these two positively and two negatively charged amino acids at the extracellular side of the TSHR is conceivable.

4.2.2 Mutations with both decreased basal Gαs- and decreased hormone-induced Gαq activity

We identified mutants in ECL2 that are characterized by impaired basal Gαs mediated signalling activity (Table 3.8). In addition to our findings for ECL2, several other mutants with decreased basal activity but with cell surface expression comparable to the wt receptor have been characterized at ICL2, at the junction between TMH5/6 and ICL3 and in the

interior of the transmembrane domain (Neumann S 2005, 2006, Biebermann H 1998, Urizar E 2005 (a)). It is noteworthy that most of the mutations exhibiting decreased basal cAMP accumulation also exhibit decreased ligand-mediated Gαq activity. However, hormone-stimulated cAMP accumulation is not affected to the same degree as IP accumulation. Based on these observations, the hypothesis is derived that the activity state of basal Gαs signalling is correlated to the hormone induced Gαq conformation rather than to hormone induced Gαs mediated signalling mechanisms of the TSHR. This allows the postulation that the conformation of the basal Gαs mediated activity is a prerequisite for hormone induced IP accumulation at the TSHR.

4.2.3 The interface between the ECL2 and the TMH6

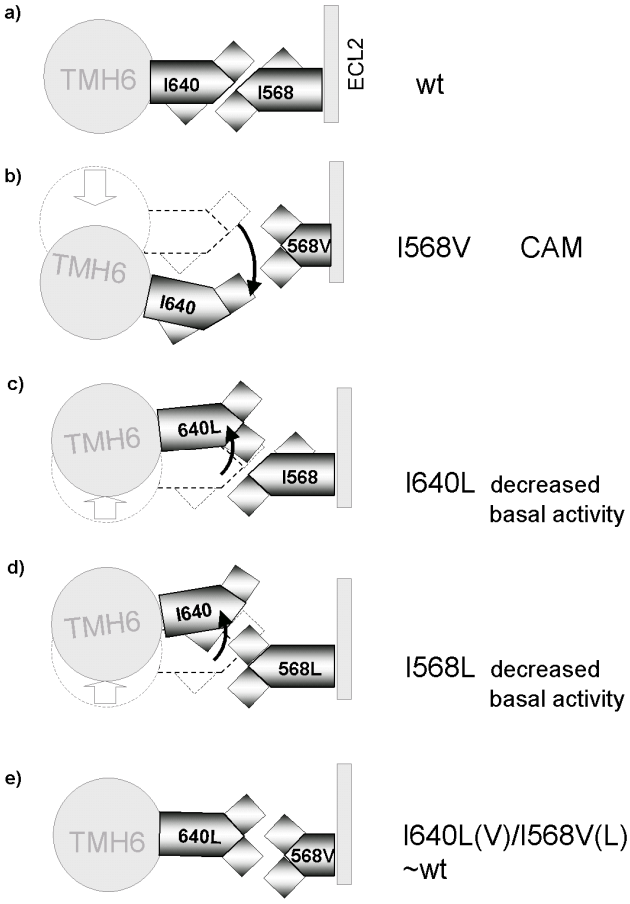
The TSHR model based on the X-ray structure of inactive bovine rhodopsin (Palczewski K 2000) suggests that the ECL2 is plugged nearly horizontally into the transmembrane domain on the extracellular side. Similar orientations of ECL2 have also been confirmed for other 7TMRs by mutational analysis (Liu S 2003, Costanzi S 2004, Wuller S 2004). Mutations of the hydrophobic amino acid I568 in ECL2 are able to activate the TSHR constitutively. The mutations to threonine and valine are known to be pathogenic (Parma J 1995, Claus M 2005). In our model I568 is pointing downwards, directly into a cleft between the transmembrane helices. To identify possible hydrophobic residues in the TMHs that interact with I568 as counterparts, systematically introduced side chain alterations at potential interaction sites of I568 with TMH1 (L417), TMH2 (I470), TMH6 (I640) or TMH7 (V664) were tested (Figure 3.7).

The mutant I640V showed constitutive activation. I640 is located in TMH6, close to the TSHR P639. This corresponds to the highly conserved proline P6.50 (Ballesteros & Weinstein number (Ballesteros JA 1995)) in GPCR family one, and is responsible for causing a kink at TMH6. The side chain of I640 points towards ECL2 (Figure 3.8) and the hydrophobic side chain interaction between I640 (TMH6) and I568 (ECL2) is essential to adjust TMH6 towards ECL2 in the basal active conformation (Figure 4.4a). Reduction of the side chain length at the interface between TMH6 and ECL2 by I640V at TMH6 or by I568V as the counterpart at ECL2 leads to a release of this interaction resulting in constitutive receptor activation, very likely by movement of TMH6 (Figure 4.4b). Our model can explain our observation of a decreased basal activity for the I640L mutant as well. The slight side chain alteration from isoleucine to leucine at position 640 by moving the branching further towards the tip of the side chain (from the beta-C atom to the gamma-C atom) causes a

different interlocked interaction with I568 in ECL2. The additional gamma methyl group constrains TMH6 in a direction opposite to the release and activation state, and thus it is moved to a more inactive position preventing the modification of the basal active conformation (Figure 4.4c). Although in the double mutant I640L/I568V the side chain of 568 is reduced, the additional methyl group of leucine enables an interaction between the two branched side chains, which is reflected by its wt behaviour (Figure 4.4e).

To validate our suggested scenario the reciprocal single mutants I568L and I640V were designed and the corresponding double mutant I640V/I568L. In the case of an intertwining side chain contact between I640 and I568, the introduction of the differently branched side chain of leucine (branch already on the tip of side chain see Figure 4.4) by mutant I640L on TMH6 should cause an interlocking of the side-chains (Figure 4.4c).

Figure 4.4*: Molecular model based scheme of wild type and mutant interactions between ECL2 (position 568) and TMH6 (position 640)*



a) In the wt receptor the hydrophobic side chain interaction between I640 (TMH6) and I568 (ECL2) is essential to position TMH6 towards ECL2 into the basal active conformation, **b),** Reduction of side chains by mutants I568V at ECL2 or by I640V as the counterpart in TMH6 leads to a release of the intertwined interaction and to a constitutive receptor activation by movement of TMH6, **c)** The slight side chain alteration from isoleucine to leucine at position 640 by an additional branching at the tip of the side chain (at gamma-C instead of at beta C-atom) causes a different interlocked interaction with I568 at ECL2. The additional gamma methyl group constrains TMH6 in an opposite direction compared to the state of release and activation, and thus is moved to a more inactive position preventing the adjustment of the basal active conformation**, d)** the reciprocal single mutant I568L in ECL2 shows exactly the same phenotype as I640L. **e)** Although in the double mutant I640L/I568V the side chain of 568 is reduced, the additional methyl group of leucine enables an intertwined interaction between the two branched side chains, which is reflected by its wt behavior. The same phenotype could be observed for the reciprocal double mutant I568L/I640V also pointing towards an intertwined interaction mechanism.

In that case a similar effect should also be produced by mutant I568L in the counterpart ECL2, resulting in a decreased basal activity (Figure 4.4d).

Subsequently, the reciprocal double mutant I568L/I640V (CAM I640V) should restore the wild type basal cAMP activity (Figure 4.4e). Indeed, whereas I568V and all other known mutants at position 568 are CAMs (I568T,A,F) the reciprocal mutant I568L at ECL2 is basally inactive and thus identical to mutant I640L at TMH6.

Moreover, the combination of CAM I640V with the decreased basal active mutant I568L in the double mutant I640V/I568L confirmed our prediction by restoring the basal cAMP activity of the wild type protein. This reverse approach is consistent with our model and supports the suggested scenario presented in figure 4.4c and e.

4.2.4 TMH6 may glide along the ECL2 according to different receptor activity states

An established common mechanism for family A 7TMRs is an activation-associated see-saw movement of TMH6 around the conserved proline (P6.50, TSHR P639) as a pivot opening the intracellular access for G-proteins (Schwartz TW 2006). Our data and models support the following activation scenario for ECL2/TMH6 upon hormone stimulated activation: Since K565 at the top of ECL2 is identified as involved in the formation of a hormone-induced activated conformation at least a side chain movement but also a slight backbone displacement is likely at that particular ECL2 site (Figure 4.5a). Moreover, since ECL2 is plugged into the tranmembrane domain and is intercepted at three anchor points (N-terminal TMH4, C-terminal TMH5, and disulfide bridge at Cys569) consisting of a fold-essential proline (Pro571), a slight lever-like shift during hormone-induced activation is conceivable at least for a portion of ECL2. Such a slight shift induced in ECL2 would be sufficient to release the suggested interaction between TMH6 (I640) and ECL2 (I568) in the basal conformation, thus allowing a further gliding of TMH6 along ECL2 and the subsequent see-saw movement of TMH6 for complete activation (Figure 4.5a).

Our mutants reflect a dynamic interface, where TMH6 glides along ECL2 according to three different receptor states. The basal activity is constrained by side chain interactions (Figure 4.5b). The decreased basal activity is provoked by the differently branched, bulkier side chain of I640L in TMH6 (Figure 4.5c). The side chain interlocking of mutant I640L with I568, which most likely causes a slight movement of TMH6 in the opposite direction compared to the activation, diminishes G-protein interaction. The I640F,M mutations cannot form this interlocking side chain interaction, because of their differences in side chain branching and

flexibility compared to the leucine mutant. The aromatic ring system is angled at the beta carbon of the phenyl side chain at position 640 and cannot interlock with I568V.

Methionine is characterized by an unbranched side chain with more flexibility in its side chain orientation. In the activated state the constraint is released by mutants with smaller side chains at either of the counterparts or by hormone induced ECL2 shift, allowing movement of TMH6, which opens the intracellular access for G-proteins (Figure 4.5a).

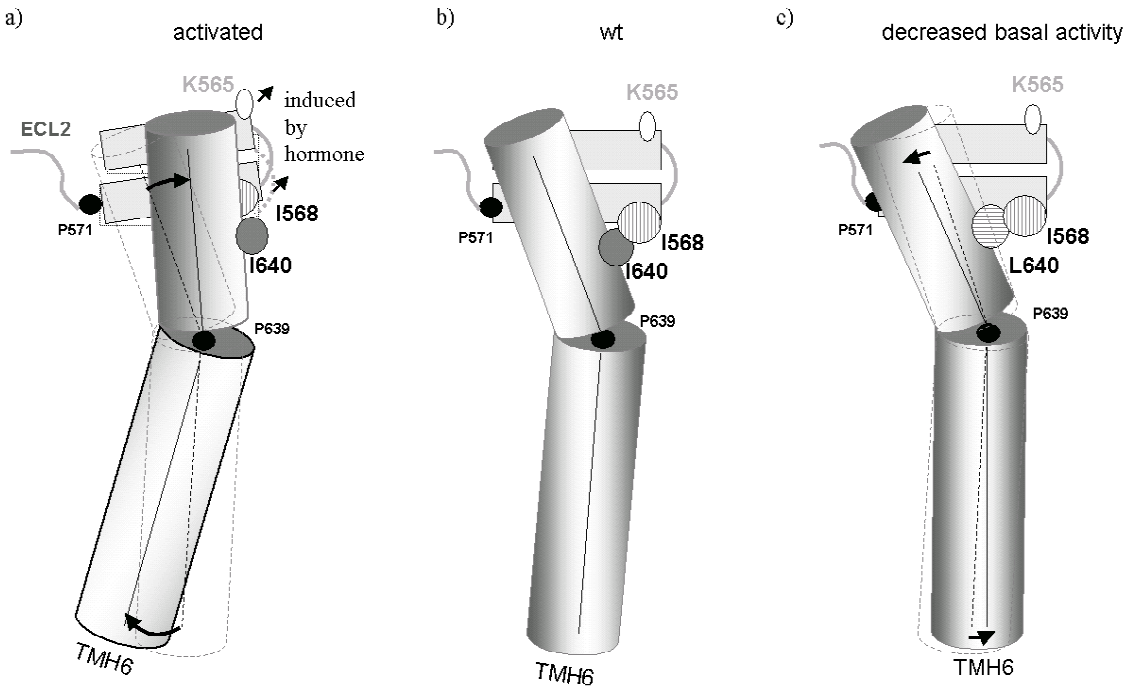


Figure 4.5*: The ECL2 regulates the TMH6 adjustment and influences directly the different activity states of the TSHR*

It has been established that a common mechanism of 7TMRs in family I is an activation-associated sew-saw movement of TMH6 around the conserved proline 6.50 (TSHR P639) as a pivot. Our mutants reflect an interface where TMH6 glides along ECL2 in a row of three different receptor states. **a)** In the activated state the rather constraint basal state is released either by mutants with reduced side chains or by a hormone-induced ECL2 shift. Since identified residue K565 in ECL2 is involved in forming a hormone-induced activated conformation a slight backbone displacement is also feasible at ECL2 (small arrows). Such an induced slight lever-like shift around the essential proline 571 of ECL2 would be sufficient to release the stabilized interaction between TMH6 and ECL2, and allow the see-saw movement of TMH6 for complete activation. **b)** The hydrophobic side chain interaction of I640 (TMH6) and I568 (ECL2) is constraining the basal (partially active) conformation of wild type TSHR. The TMH6 in this partially active state is slightly tilted to the vertical axis. **c)** The decreased basal activity is provoked by a differently branched bulkier side chain (I640L) at TMH6 by further interlocking with I568, which most likely causes a slight movement of TMH6 in the opposite direction to the activation conformation and diminishes G-protein interaction (see Figure 4.4c,d).

Taken together amino acids in ECL2 and TMH6 of the TSHR were identified, which are required for signalling and an activation mechanism is suggested at the interface between

TMH6 and ECL2, in which TMH6 glides along ECL2 according to the different receptor activation states.

4.2.5 Section Summary

A number of alanine mutations in extracellular loop two (ECL2) of the TSHR were found to increase or decrease basal activity compared to the wild type receptor. K565A was identified as a mutant with decreased basal activity and strongly impaired hormone induced signalling activity. To gain insights into how ECL2 mutants affect basal activity, we focused on the constitutively activating pathogenic mutant I568V in ECL2, which exhibits elevated basal activity. Because our molecular model suggests I568 is embedded in an environment of hydrophobic residues provided by the transmembrane helix bundle, we tested mutants in this region to identify potential interaction partner(s) for I568. Indeed, double-mutant I568V/I640L (ECL2/TMH6) suppresses the increased basal activity exhibited by I568V alone. This suggest a spatial and functional relationship between ECL2 and TMH6 in which side chain a interaction between I568 and I640 constrains the receptor in a conformation with low basal activity. While the single mutant I640L exhibits basal activity lower than the wild type, its differently branched and bulkier side chain complements the reduced side chain bulk in I568V, restoring wild type basal activity to the double mutant. The reciprocal double mutation I640V/I568L confirms this scenario.

These and other mutant phenotypes reported here support a dynamic interface between TMH6 and ECL2. Disruption of this critical interface for signalling by introduction of mutations in the TSHR can either increase or decrease basal activity.

4.3 The extracellular loop 3 is of high functional importance

The importance of ECL3 for ligand binding and signal transduction has been well studied for the FSHR and the LHCGR. *In vitro* mutagenesis for these two receptors revealed that ECL3 is important for receptor activation, but that single amino acids contribute differently (Ryu KS 1996, 1998 (b), Fernandez LM 1996 (a,b), Gilchrist RL 1996, Sohn J 2002). Therefore, the functions of theTSHR ECL3 were studied in more detail. This idea was supported by the occurrence of two different pathogenic constitutively activating mutations in the ECL3 of the TSHR: N650Y and V656F (Tonacchera M 1996, Fuhrer D 1997).

The highly conserved position K660 at the ECL3/TMH7 junction has been described previously as being important for the signalling of the LHCGR (K583) and the FSHR (K590)

(Sohn J 2002, Gilchrist RL 1996). For this reason and based on the functional characteristics of the TSHR K660A mutant, it was assumed that this amino acid might also be a key position for Gαs or Gαq activation in the TSHR.

4.3.1 A hydrophobic cluster in the center of the third extracellular loop is important for TSH receptor signalling

The functional characterization of the TSHR ECL3 alanine substitutions revealed no significant effects on cell surface expression and TSH binding affinity. In contrast, mutations at the LHCGR positions K573, P575, I577 and N581 (TSHR: N650, P652, I654 and N658) resulted in strongly decreased receptor numbers, while expression of the other mutants was comparable to the wt LHCGR (Ryu KS 1996). Alanine mutations in the FSHR ECL3 also caused a decrease of receptor molecules on the cell surface (Ryu KS 1998 (b)). Mutation of the hydrophobic cluster in the FSHR ECL3 caused three-fold increased FSH binding affinities (Sohn J 2002), whereas TSHR mutants of the homologous hydrophobic residues showed no influence on TSH binding (Table 3.12). In summary, these data suggest that individual amino acids seem to play different roles in these three receptors regarding correct receptor folding.

Introduction of alanine for residues P, L, I, T, V in the central ECL3 had the strongest effect on cAMP production (Figure 3.12). However, contrary to the FSHR, the TSHR mutants did not totally abolish cAMP formation but displayed about 50% decreased cAMP activity compared to the wild type TSHR. In contrast to our findings, alanine substitutions in the LHCGR ECL3 caused only moderate effects on hCG-induced cAMP formation for most residues. Mutations P575A and V579A in the LHCGR (TSHR: P652A and V656A) were found to decrease stimulated cAMP production (Ryu KS 1996, 1998 (b)). Therefore, for TSHR and FSHR but not for LHCGR, a hydrophobic cluster in the center of ECL3 is most likely important for ligand-induced activation of the Gαs-mediated cAMP-signalling pathway. This suggests that the LHCGR ECL3 may act in a different manner in the process leading to Gαs activation. However, despite different functional characteristics of mutants in the central portion of ECL3, P652 and V656 (which are adjacent to the described hydrophobic cluster) as well as K660 (at the ECL3/TMH7 junction) are important for cAMP signalling in all three GPHRs. Substitutions of residues in the hydrophobic cluster of ECL3 also caused the strongest reduction of Gαq-mediated IP production (Table 3.12). Despite high sequence homology within ECL3 (Figure 3.12), comparison of results from alanine scanning mutagenesis of these closely related GPHRs reveals significant functional differences.

Individual residues seem to play different roles in these receptors regarding receptor folding, cAMP- and IP-signalling.

In addition, our results underline that the highly conserved K660 at the junction between ECL3 and TMH7 is essential in a similar manner for both signalling pathways in all three GPHRs. Based on relaxed side chain conformations of the TSHR serpentine model (Figure 3.11) all possible orientations of the K660 side chain using molecular dynamics simulations were considered and determined potential interaction partners that might be able to form salt bridges or hydrogen bonds with K660. Experimental verification of interactions of charged amino acids by double mutants excludes an interaction of K660 with suggested acidic residues. Therefore, an orientation of K660 towards the backbone of ECL2 forming a hydrogen bond to a backbone carboxyl-oxygen of the ECL2 is favored (Figure 3.11).

4.3.2 Section summary

Previous reports on the FSHR and LHCGR, suggest that the ECL3 could be a key domain for ligand binding and intramolecular receptor signalling. In contrast to ECLs 1 and 2 of the GPHRs, the ECL3 displays high sequence homology among the GPHRs, particularly in the central portion of the loop. Therefore, our group opted to identify amino acids with functional importance within ECL3 of the TSHR.

Single alanine substitutions of all residues in ECL3 were generated. Functional characterization revealed the importance of five amino acids in the central portion of ECL3 and K660 at the ECL3/TMH7 junction for TSHR signalling. Decrease of G_S activation and loss of G_q activation by substitutions of K660 demonstrates a role for this position for TSHR conformation and signal transduction. Our findings provide evidence that a hydrophobic cluster, comprising residues 652 – 656 of ECL3, strongly influences intramolecular signal transduction and G-protein activation of the TSHR.

4.4 TSHR activation by a low molecular weight ligand

Orally active low molecular weight (LMW) antagonists may have therapeutic potential for TSHR-mediated hyperthyroidism and LMW agonists might replace injected rhTSH in the diagnosis of thyroid cancer. Up to now, no low molecular weight ligand has been found to be a modulator of TSHR activity.

In 2002 van Straten and co-workers published the first LMW agonist (org41841) for the LHCGR (van Straten NC 2002). The thienopyrimidine org41841 was identified as a high

potency agonist (EC50: 20 nM) using the cyclic-AMP-response-element (CRE)-luciferase-reporter assay. Interestingly, org41841 did not compete with 125I-labeled LH for binding, indicating distinct binding sites for the two ligands. It was suggested that org41841 did not bind to the extracellular leucine rich repeat domain like the glycoprotein hormones but rather within the transmembrane domain of the receptor (van Straten NC 2002). It is known that small ligands bind to the transmembrane cores of rhodopsin and other GPCRs (Palczewski K 2000, Sun Y 2003, Shim JY 2003). However, direct evidence for the binding site for LMW ligands in GPHRs has not been presented.

Using a combination of molecular modeling and evaluation of the binding of org 41841 within several TSHR/LHCGR chimera mutants, it was recently reported that org41841 was found to bind within the seven-transmembrane domain of the TSHR and LHCGR (Jäschke H 2006). Specifically, this compound was shown to bind within a localized pocket between transmembrane helices 3, 4, 5, 6 and 7 and extracellular loop 2. Low molecular weight ligands of the LHCGR and TSHR are remarkable in that they are not likely to compete with the large native hormones for binding at the extracellular N-terminal domain.

The docking studies in this work represent an extension and refinement of the previous work on TSHR (Jäschke H 2006 (b)) to reveal information for ligand and receptor model optimization.

4.4.1 Modes of binding of a small agonistic molecule

Analysis of the docking models of this class of ligands at both LHCGR and TSHR provided two distinct docking orientations (Figure 3.14). The two orientations have nearly equivalent docking scores yet they represent two vastly divergent ligand positions.

The docking-models represent an extension of our previous work on the binding pocket for small molecules on the TSHR (Jäschke H 2006) and constitute the first reported docking models of this class of ligands to the LHCGR. In part, the models confirm our suggestion that the LHCGR contains a larger binding pocket than the TSHR. This has numerous consequences with regards to the binding mode of org41841 within the two homologous receptors. From an early stage in our modeling experiments it was apparent that org41841 could bind in two distinct orientations to both receptors with very similar docking scores. Common to both is the pivotal role of E3.37 as a hydrogen bond acceptor. Further, these two orientations were statistically more significant than all other docking orientations examined. They can be briefly explained by using the methoxy-phenyl moiety as a point of reference.

In orientation A, the methoxy-phenyl ring is seen in a position proximal to transmembrane helixes 4 and 5 and bordering the extracellular loop region (Figure 3.14). In orientation B, the methoxy-phenyl ring is seen in a position proximal to transmembrane helix 3 and more deeply buried within the transmembrane core.

These docking studies between a first LMW agonist and the TSHR (and LHCGR) give detailed information regarding the binding mode and contact points between the receptors and compound org41841. This information is a prerequisite for the rational modification of this compound scaffold with respect to the LMW agonist optimization and development of antagonistic or inverse agonistic properties.

4.4.3 Section summary

Many cognate low molecular weight (LMW) agonists bind to 7TMRs within their transmembrane helices. The thienopyrimidine org41841 was identified previously as an agonist for the LHCGR and suggested to bind within its TMHs because it did not compete for LH binding to the LHCGR ectodomain. Because of its high homology with the LHCGR, it was predicted by us that the TSHR might also be activated by org41841. It was shown that org41841 is a partial agonist for the TSHR but with lower potency than for the LHCGR. Analysis of 3-dimensional molecular models of the TSHR and LHCGR predicted a binding pocket for org41841 in common clefts between TMHs 3, 4, 5, 6 and 7 and extracellular loop 2 in both receptors. Evidence for this binding pocket was obtained in signalling studies with chimeric receptors that exhibited improved responses to org41841. Furthermore, a key receptor-ligand interaction between the highly conserved negatively charged E3.37 and the amino group of org41841 predicted by docking of the ligand into the 3-dimensional TSHR model was experimentally confirmed. The docking process reveals 2 favourable binding modes of org 41841 that reveal detailed information concerning properties of the binding pocket and interactions between the TSHR and the LMW ligand. Moreover, these studies provide the foundation for further virtual screening to identify potential new LMW scaffolds that bind to that binding pocket at the transmembrane helix bundle of the TSHR.

4.5 A Sequence-Structure-Function-Analysis resource

4.5.1 SSFA concept and modules

Experimentally identified inactivating and constitutive activating mutations of GPHRs provide clues about the involvement of particular residues in forming conformations which define different receptor activity states (Cotecchia S 2003, Kleinau G 2004, Urizar E 2005,

Puett D 2005, Zhang M 2005). Receptor phenotypes and their biological characterizations represent information about differences between the activity states (Ro to R*) (Kristiansen K 2004, Vauquelin G 2005). To improve the almost isolated consideration of point mutations in structure-function studies for the deduction of molecular activation mechanisms of GPHR, a retrieval system for a systematic and diversified analysis of GPHR mutation data is presented.

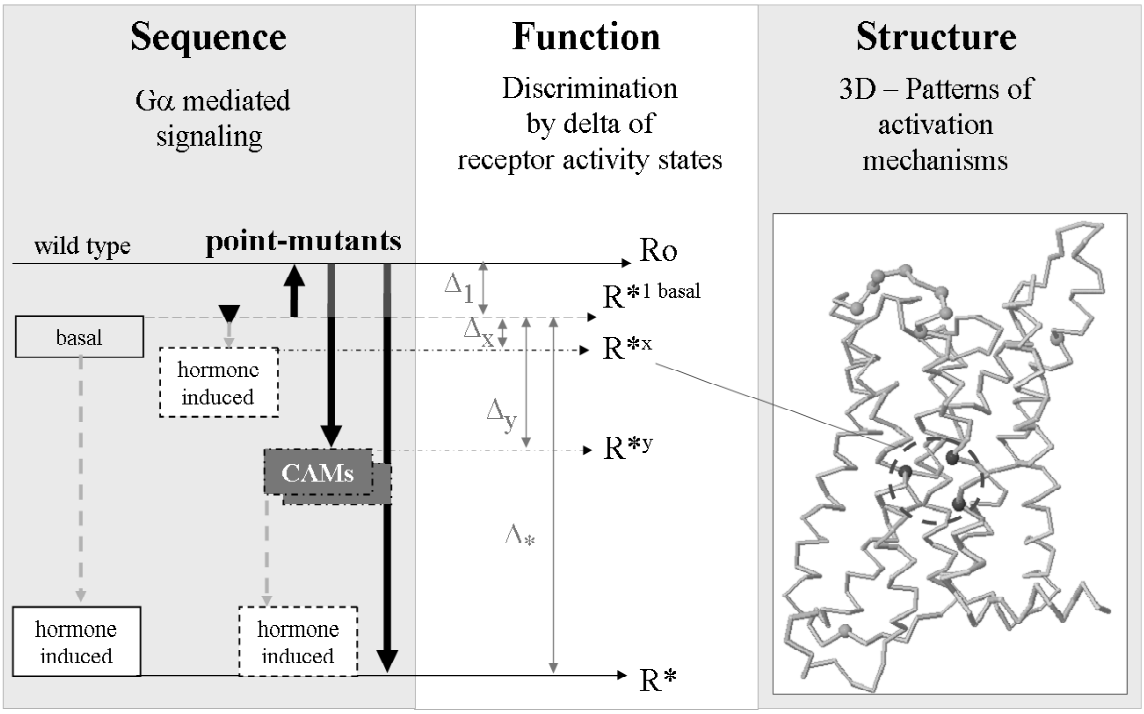


Figure 4.6: ***Basic concepts of Sequence-Structure-Function Analysis for GPHRs***
Starting point are the sequences of wild type and mutation phenotypes of GPHR. Different phenotypes that are generated by mutations (black bold arrows) like constitutively active mutants (CAMs) and hormones (grey dashed arrows) represent levels of different activation states (Ro: abolished basal activity; $R^{*1basal}$: basal activity, $R^{*x...y}$: intermediately active states, R: complete induced activation) of the receptor (function of the protein). Search and filtering tools of SSFA enable a semi-quantitative user-driven functional classification of mutant phenotypes that discriminates according to the difference Δ (gray arrows) between the respective activation states. Assignment and mapping of spatial locations for similar and different functionalities to 3D structural models (structure) leads to determinants that are responsible for forming conformations of corresponding activity states and finally to 3D-patterns of activation mechanisms. Example for the TSHR serpentine domain model: mutants inactivating the Gαs a n d Gαq mediated signalling simultaneously are clustered at the transmembrane core at TMH6 and TMH7 (dark balls, dashed circle) and indicate that wild type residues at these positions are significantly involved in the formation of a general active conformation enabling activation of Gαs and Gαq. On the other hand amino acids selectively responsible only for the Gαq signalling cascade are localized (grey balls) mainly in the extracellular (ECL3) and intracellular loops (ICL2).*

The basic concept of our sequence-structure-function analysis is outlined in figure 1 and the modules are shown in figure 2. A semi-quantitative classification of mutant phenotypes according to the different activity states they represent, and mapping of classified phenotypes

to the 3D-structures of the receptors, reveal interrelations between determinants of the corresponding activity state. More specifically, hormones and mutations give rise to different phenotypes that represent levels of various activation states, e.g. constitutive or basal activity (Figure 4.6).

Several tools allow the assignment and mapping of spatial locations for similar and different functionalities as 2D- (table) and 3D-(Model) outputs. This yields determinants that are responsible for forming conformations of corresponding activity states and finally, 3D-patterns of activation mechanisms.

4.5.2 SSFA Tools

For a sequence-structure-function analysis the mutational data set is combined with different tools to enable focused searches and generation of different output formats. Regardless of the type of sequence numbering, a multiple sequence alignment with a unified numbering system allows easy localisation of any residue number. This not only enables a helpful overview of which residues occur at corresponding positions in homologous GPHRs but also whether mutational data is available. The queries and outputs are designed with regard to: **a.** the GPHR subtypes TSHR (human), LHCGR (human and rat), FSHR (human and rat) used mostly for mutagenesis studies; **b.** the three different established sequence-numbering systems (starting with the first amino acid of the sequence (Num1); the Ballesteros-Weinstein numbering-system; numbering by the GPCRDB); **c.** structural epitopes or domains, **d.** specified amino acid or mutant properties, **e.** type of mutation (change of and to specific residue properties), **f.** and functional characteristics of mutations that allow a combinatorial search for various parameters.

Our compiled set of GPHR mutant phenotypic data (Figure 4.7) and their percentage values enables the use of the SSFA search and filtering tools to compare and discriminate the phenotypes according to the difference (Δ) between the respective activation states (Figure 4.6). This allows a semi-quantitative, user-driven functional classification of mutant phenotypes.

Data-Analysis - The ‘Search’ functions allows combinations of queries for specific assays used for the characterization of receptor phenotypes. This enables a precise definition of queries under inclusion and/or exclusion of user driven data-ranges of normalized standard assay values (Figure 4.7).

We implemented an analyzing strategy, which filters distant values. Therefore, our semi-quantitative tool is designed to classify the functional data into two rough classes.

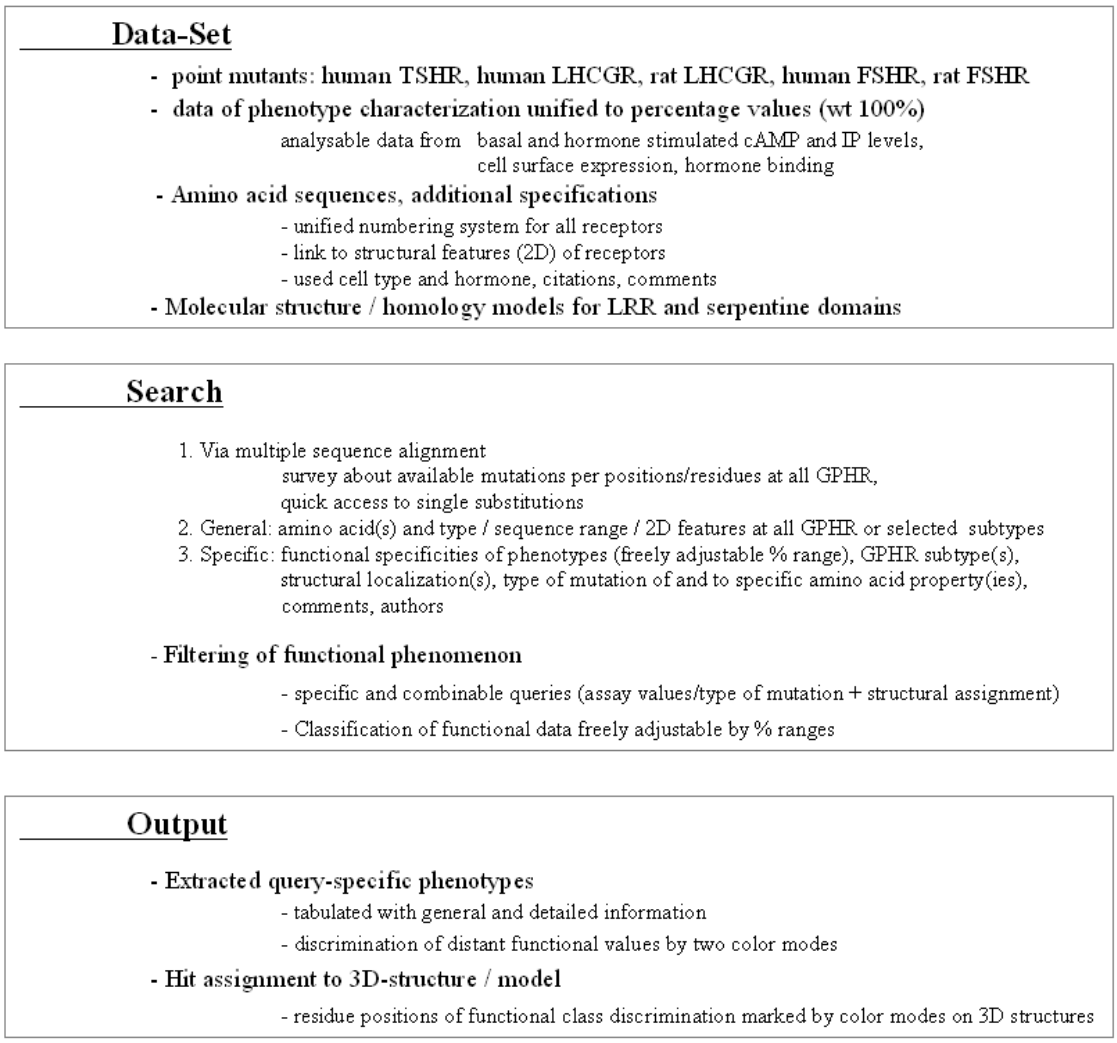


Figure 4.7: ***Scheme of modules and service tools of the SSFA resource***
Data set: specific and general phenotype data of GPHR mutations are compiled as unified percentage values allowing comparable analysis of basal and hormone induced cAMP and IP level data and surface expression and hormone binding throughout all mutants (~850 at 10/2006). Apart from the functional data, various side chain properties of wild type and mutant, location in structural regions and features, such as cell type and hormone type used, citations and comments, different numbering systems etc. are also enclosed as searchable attributes for each mutation. Search: The classification of functional data according freely adjustable percent ranges is the essential part of an advanced search to filter common, different or even contrary phenotype data. Output: Further filtering and discrimination of distant functional values by two color modes. Assignment to structural location given as 2D- (table) and 3D-(Model) output.

A freely adjustable coloring-system allows easy discrimination of similar and different functional effects of mutants and their corresponding properties by color-coding tabulated results. Colorable balls at C-α positions also enable the visualization of results from our approach on one 3D-structure per receptor as well as the loci for combinations of point

mutations with similar and different or even opposed phenotypes (Figure 4.8). The color(s) can be chosen according the discrimination of phenotype classes. This enables the identification, not just of clusters of loci generating common phenotypes, but also of spatially close wild type residues that could possibly be involved in interactions whose modification or disruption by mutants might therefore cause this phenotype.

4.5.3 Signalling specificities and interrelated determinants of GPHRs revealed by SSFAnalysis

4.5.3.1 Amino acids of high importance for the conformation of hormone induced receptor activation

SSFA queries for mutations that only effect hormone induced Gαq and Gαs mediated signalling can reveal positions that are explicit determinants. These mutants either prevent the formation of the active state upon hormone binding or might stabilize the inactive state and block further activation.

Such queries showed 12 TSHR mutations with less than 50% hormone induced Gαq mediated signalling compared to the wild type. Three of these mutations (TMH6-D633R (Neumann S 2001), TMH7- N670A (44) and N674A (Neumann S 2001)) simultaneously show inactivation for Gαs. Therefore nine mutations selectively affect Gαq mediated signalling (Figure 4.8). These nine amino acids are localized in the extracellular loops ECL1 (Jäschke H 2006), ECL3 (Claus M 2005), the junction of ECL3/TMH7 (Claus M 2005), and the ICL2 (Neumann S 2005).

Together with positions of mutants selectively initiating constitutive Gαq activity (see section below) these mutants can be considered as determinants responsible for forming a selective Gαq conformation. LHCGR and FSHR mutants with common phenotypes could not be retrieved. Thus currently only Gαq mediated signalling selective determinants could be identified. The amino acids of high importance for both Gαq and Gαs mediated signalling are located in the center of the transmembrane domain between TMH6 and TMH7. These positions are key-players both for the Gαq and Gαs mediated signalling process.

The compilation of unified percentage values for the mutational effects allowed exact discrimination between common loci/positions that are important for maintaining basal activity and for generating constitutive activity (as exemplarily shown in figure 4.8) for the first time. Moreover, the simultaneous consideration of different assay data make it possible

to distinguish between loci/positions that are generally important for all Gα subtype signalling pathways, and further to discriminate between positions which are important in determining conformations either for the Gαq or the Gαs subtype signalling pathways alone (Figure 4.8). Our results show that the conformation for Gαq selective signalling is not only dependent on the expected positions at the intracellular loops, but also on a defined spatial cluster of positions at the extracellular loops. These results have new implications for the understanding the molecular cascade of G- protein subtype dependent signal transduction through the transmembrane domain.

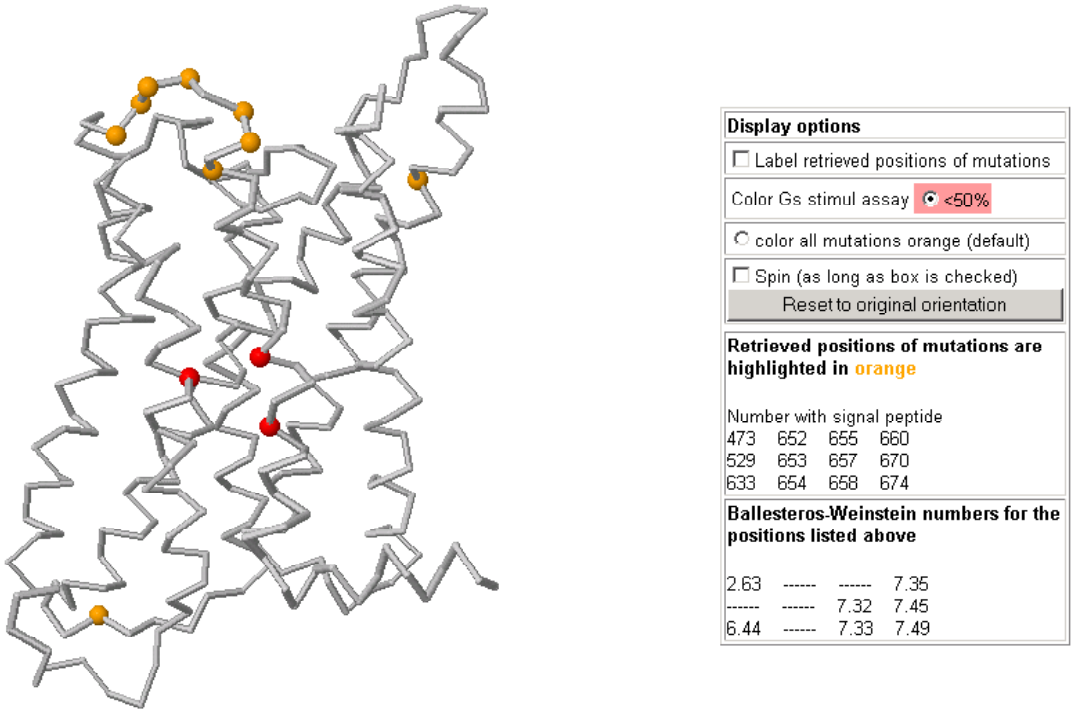


Figure 4.8: ***Amino acids of high importance for the hormone induced receptor conformation***
SSFA query revealing mutations that exclusively effect hormone induced Gαq and Gαs mediated signalling. The natural amino acids at these positions are explicit signalling determinants and mutations either prevent the formation of the active state upon hormone binding or might stabilize the inactive state and block further activation. Such a query reveals 12 TSHR mutations with less than 50% hormone induced mediated signalling activity compared to the wild type. Further query specifications (filtering) are: cell surface expression and maximum binding capability >50%; hormone independent basal activity between 50-140% compared to the wild type. The borderlines used are extracted by analysis of inactivating pathogenic mutations (iPMs) of GPHRs. For most of the characterized iPMs the receptor functionalities are less than 50% compared to the wild type. Three mutations (TMH6 D633R; TMH7 N670A, N674A) are simultaneously important for both Gαq and Gαs mediated signalling. In conclusion, these 3 amino acids are fundamental for both signalling pathways and 9 mutations exclusively impair the Gαq mediated signalling capability.

4.5.4 Section Summary

Comparisons between wild type and mutated glycoprotein hormone receptors TSHR, FSHR

and LHCGR are made so as to identify determinants involved in the molecular activation mechanism. The basic aims of this SSFA resource are i) the discrimination of receptor phenotypes according the differences between activity states which they represent, ii) the assignment of classified phenotypes to 3D-structural positions to reveal iii) functional-structural hotspots and interrelations between determinants that are responsible for corresponding activity states. Since it is hard to survey the vast amount of pathogenic and site-directed mutations available for GPHRs and to improve an almost isolated consideration of individual point mutants, a free and accessible system for systematic and diversified sequence-structure-function analysis (SSFA) were developed (http://www.fmp-berlin.de/ssfa). In order to combine all mutagenesis data into one set, the functional data were converted into unified scaled values. This at least enables their comparison in a rough classification manner. Complementary to known databases our data set and bioinformatics tools make it possible to link functional and biochemical specificities with spatial features to reveal concealed structure-function relationships by a semi-quantitative analysis. New interrelations of determinants important for selective G-protein mediated activation of GPHRs are resumed.

4.6 General importance of presented studies

Generation of homology modelling driven hypotheses, verification by site directed mutagenesis and mapping of motifs stepwise along different portions of a protein, proved to be a suitable strategy to elucidate sequence-structure-function related information about membrane proteins, especially for cases where there is a lack of structural data. By providing new insights into mechanisms that are involved in the activation (or inactivation) processes of the TSHR, these studies contribute significantly to filling the gap in understanding how the signal is transduced after extracellular hormone binding. Molecular understanding of pathogenic mutations such as constitutive activation of TSHR (e.g. at ECL2) not only helps to decipher the molecular activation mechanism of the receptor, it also provides the foundation for future studies on directed pharmacological intervention. However, the study provides the allosteric binding mode of a small molecule interacting with the transmembrane region at the TSHR for the first time. Finally, to rationalize the sequence-structure-function analysis (SSFA) using functional data from about 900 mutations, a freely accessible web-based information resource was developed for GPHRs.