3 RESULTS

Based on the designed structural models, and using an iterative strategy of sequence-structure-function analysis (SSFA), hypotheses for potential signalling sensitive amino acids or epitopes were derived. This information was proofed by site-directed mutagenesis approaches to confirm or to refine structural-functional hypotheses. The predicted structure-function relationships were developed by the FMP Berlin. The experimental proof of functional-structural relationships by site directed mutagenesis was contributed by the cooperation partner of the University of Leipzig, III. Medical Department, Endocrinology.

3.1 Molecular analysis of receptor components

In the following chapters the results of bioinformatic procedures, molecular modeling and mutagenesis studies are described in detail. The chapters are sorted starting with the first N-terminal domain and progress towards the C-terminus in the following (structural) order:
1. cysteine-box 1 (C-b1), 2. leucine-rich-repeat domain (LRRD), 3. cysteine-box 2 (C-b2), 4. vicinity of S281 at C-b2 (plus ECL1), 5. cysteine-box 3 (C-b3), 6. extracellular loop 2 (ECL2), 7. extracellular loop 3 (ECL3), 8. docked complexes of low molecular weight (LMW) agonists within the serpentine domains of the TSHR and the LHCGR.
Regarding the aim to reveal insights into the extracellular signalling mechanism by complexed receptor components, extracellular structures, e.g. between LRRD/C-b2/C-b3, C-b2/C-b3, or C-b2/ECL1, and experimental proofs of derived functional hypotheses were described.
Features, possibilities and results of the designed Sequence-Structure-Function Analysis resource for GPHRs that represents a fundamental basis and result of this work are also given.

3.1.1 Cysteine-box 1 and the leucine-rich repeat hormone binding domain

3.1.1.1 Molecular Model

Studies were carried out to examine the optimal structural template for the cysteine-box 1 and the LRR domain and of the GPHRs (Kleinau G 2004). The LRRD from the hNogo-receptor ectodomain (pdb entry code 1OZN) (HE XL 2003) was identified in fourteen different LRRD structures (Enkhbayar P, 2003) and chosen as the best representative of the LRRD features of all three human GPHRs based on the highest amino acid sequence similarity (Figure 3.1).

N-terminal Cysteine-box 1

		Res No	
	hNogoR	26	P C PGA. C V C YNEPK
	hTSHR	23	G C SSPP C E C HQEED
	hFSHR	17	G C HHRI C H C SN...
	hLHCGRR	29	L C PEP. C N C VPD..

LRRD repeat No.

Sequence similarities			β x_1	β x_2	β L	β x_3	β L	β x_4	β x_5	nx<u>x</u>Lxx Lxxx<u>xx</u>	**a**	**F**	xxLxxxx
LRR0	hNogoR	39	V	T	T	S	C	P	Q	QG LQAVPVGIPAAS			
	hTSHR	37	F	R	V	T	C	K	D	IQRIPSLPPST....			
	hFSHR	28	R	V	F	L	C	Q	E	SK VTEIPSDLPRNA			
	hLHCGR	40	G	A	L	R	C	P	G	PT AGL.........			
LRRI	hNogoR	60	Q	R	I	F	L	H	G	NR ISHVPAA	**S**	**F**	RACRNL
	hTSHR	55	Q	T	L	K	L	I	E	TH LRTIPSH	**A**	**F**	SNLPNI
	hFSHR	49	I	E	L	R	F	V	L	TK LRVIQKG	**A**	**F**	SGFGDL
	hLHCGR	52	T	R	L	S	L	A	Y	LP VKVIPSQ	**A**	**F**	RGLNEV
LRRII	hNogoR	84	T	I	L	W	L	H	S	NV.LARIDAA	**A**	**F**	TGLALL
	hTSHR	79	S	R	I	Y	V	S	I	DVTLQQLESH	**S**	**F**	YNLSKV
	hFSHR	73	E	K	I	E	I	S	Q	NDVLEVIEAD	**V**	**F**	SNLPKL
	hLHCG	76	I	K	I	E	I	S	Q	IDSLERIEAN	**A**	**F**	DNLLNL
LRRIII	hNogoR	108	E	Q	L	D	L	S	D	NAQLRSVDPA	**T**	**F**	HGLGRL
	hTSHR	104	T	H	I	E	I	R	N	TRNLTYIDPD	**A**	**L**	KELPLL
	hFSHR	98	H	E	I	R	I	E	K	ANNLLYINPE	**A**	**F**	QNLPNL
	hLHCGR	101	S	E	I	L	I	Q	N	TKNLRYIEPG	**A**	**F**	INLPGL
LRRIV	hNogoR	133	H	T	L	H	L	D	R	CG LQELGPG..	**L**	**F**	RGLAAL
	hTSHR	129	K	F	L	G	I	F	N	TG LKMFPDLTK	**V**	.	YSTDIF
	hFSHR	123	Q	Y	L	L	I	S	N	TG IKHLPDVHK	**I**	.	HSLQK.
	hCG/LHR	126	K	Y	L	S	I	C	N	TG IRKFPDVTK	**V**	**F**	SSESN.
LRRV	hNogoR	157	Q	Y	L	Y	L	Q	D	NA.LQALPDD	**T**	**F**	RDLGNL.
	hTSHR	154	F	I	L	E	I	T	D	NPYMTSIPVN	**A**	**F**	QGLCNET
	hFSHR	147	V	L	L	D	I	Q	D	NINIHTIERN	**S**	**F**	VGLSFES
	hLHCGR	151	F	I	L	E	I	C	D	NLHITTIPGN	**A**	**F**	QGMNNES
LRRVI	hNogoR	181	T	H	L	F	L	H	G	NR ISSVPER	**A**	**F**	RGLHSL
	hTSHR	180	L	T	L	K	L	Y	N	NG FTSVQGY	**A**	**F**	NGT.KL
	hFSHR	173	V	I	L	W	L	N	K	NG IQEIHNC	**A**	**F**	NGT.QL
	hLHCGR	177	V	T	L	K	L	Y	G	NG FEEVQSH	**A**	**F**	NGT.TL
LRRVII	hNogoR	205	D	R	L	L	L	H	Q	NR.VAHVHPH	**A**	**F**	RDLGRL.
	hTSHR	203	D	A	V	Y	L	N	K	NKYLTVIDKD	**A**	**F**	GGVYSGP
	hFSHR	196	D	E	L	N	L	S	D	NNNLEELPND	**V**	**F**	HGA.SGP
	hLHCGR	200	T	S	L	E	L	K	E	NVHLEKMHNG	**A**	**F**	RGA.TGP
LRRVIII	hNogoR	229	M	T	L	Y	L	F	A	NN LSALPTE	**A**	**L**	APLRAL
	hTSHR	229	S	L	L	D	V	S	Q	TS VTALPSK	**G**	**L**	EHL...
	hFSHR	221	V	I	L	D	I	S	R	TR IHSLPSY	**G**	**L**	ENL...
	hLHCGR	225	K	T	L	D	I	S	S	TK LQALPSY	**G**	**L**	ESI...
LRRIX	hNogoR	253	Q	Y	L	R	L	N	D	NPWVCDCRARP	**L**	**W**	AWLQKFR
	hTSHR	250	K	E	L	I	A	R	N	TWTLKKLPLSL	**S**	**F**	LHL....
	hFSHR	242	K	K	L	R	A	R	S	TYNLKKLPTLE	K	L	VAL....
	hLHCGR	246	Q	R	L	I	A	T	S	SYSLKKLPSRE	**T**	**F**	VNL....
LRRX	hNogoR	280	G	S	S	S	E	V					
	hTSHR	273	T	R	A	D	L	S	Y	PSHCCAFKNQKKI			
	hFSHR	265	M	E	A	S	L	T	Y	PSHCCAFANWRRQ			
	hLHCGR	269	L	E	A	T	L	T	Y	PSHCCAFRNLPTK			

Figure 3.1: ***LRRD repeat-wise sequence alignments of N-terminal cysteine-box 1, LRRDs of the hGPHRs and from the hNogo-receptor ectodomain***

*The sequence from hNogo receptor ectodomain (pdb entry code 1OZN) best matches the LRRDs of the GPHRs according to sequence similarity, sequence length per repeat and number of repeats out of fourteen LRRD structures. A new structural feature is also the additional N- terminal repeat (named LRR0) including C41 provided by the cysteine-box 1 (background grey). Secondly, a further additional β-strand occurs at the C-terminal end (LRRX). Sequence names of crystal structures are underlined. Beside the number of each repeat (repeats 0-X) the patterns of LRR sequence similarity positions are given by x1x2Lx3Lx4x5nxxLx**a**F**x**x. The highly conserved residues, like leucine, are marked by upper case letters. The conserved phenylalanine residues forming a 'Phe spine' are marked by bold **F**. At lower case letter positions 'n' and 'a', sequences contain only less conserved N and A. Additional residues in the loop region occurring in the template and GPHR sequences are marked as x. The β-strands are indicated by β. The potential binding site comprises residue positions x1 to x5.*

Nine typical repeats match the sequence of the three receptors best and the N-terminal directly attached cysteine-box of the hNogo-receptor structure best fits to C-b1 of GPHR sequences (Figure 3.1).

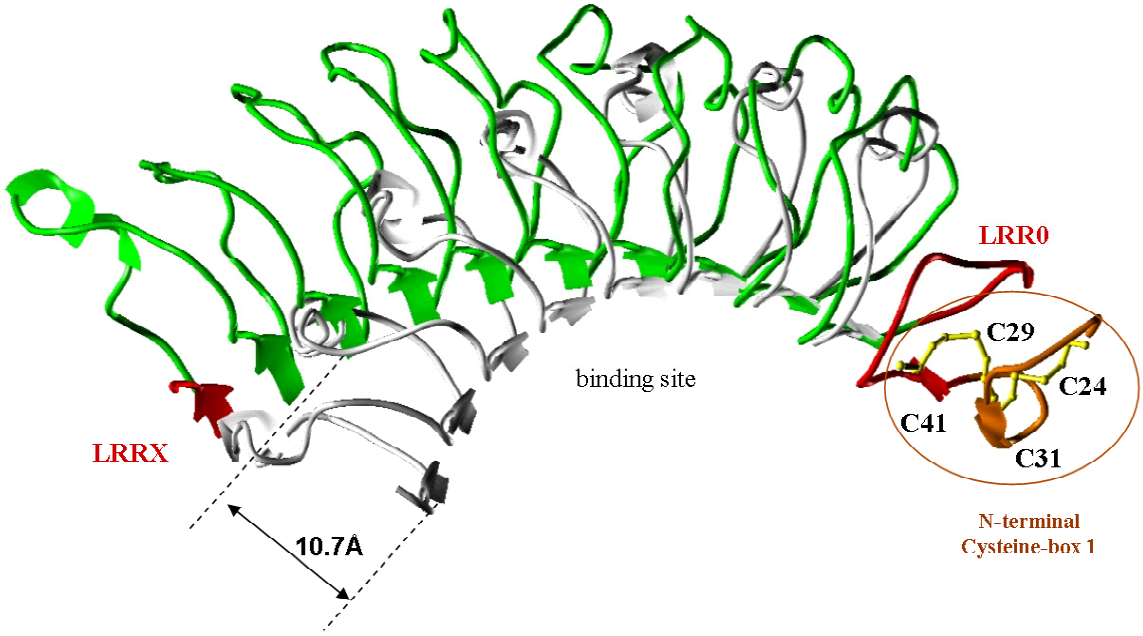


Figure 3.2: ***Comparison between structural models of the LRR hormone binding domain.***
Comparison of the previous model of the TSHR (pdb entry code 1XUM, grey) based on Ribonuclease Inhibitor template (pdb entry code 2BNH) and the newly generated homologous LRRD structure for the TSHR (red-green coloured) based on the X-ray structure of the hNogo-receptor ectodomain (pdb entry code 1OZN). The new LRRD-model offers a much larger radius and two additional parallel β-strands (red) for the inner convex arch of the hormone-binding region compared to the previous model. In our new TSHR LRRD model amino acids T273-S278 arrange the inner region of the additional 11th β-strand (LRRX), which is consistent with previously observed functional data. Strong structural importance of these residues in this 11th β-strand for expression has also been shown for the mutants T272A, L273A, T274A, Y275A (TSHR: D276, L277, S278, Y279) in the LHCGreceptor (Zeng H 2001). Chimeras between the TSHR and the LHCGR, comprising this region (Nagayama Y 1991), show that substitutions of TSHR residues L270HLTRADLS278 with corresponding amino acids of the LHCGR resulted in an impaired TSH binding and abolished cAMP response.

The previous LRRD templates such as the ribonuclease inhibitor (RI) (pdb entry code 2BNH) had significantly lower sequence similarity (~15%, PAM250 matrix) to the TSHR LRRD sequence, compared to the new hNogo-receptor template (34%).

The putative disulfide bridges between the cysteines C24, C29, C31 and C41 in the homologous TSHR model (Figure 3.2) are stabilizing an anti-parallel and parallel β-strand as integral parts of the LRRD, where the latter participates as an additional parallel strand (LRR0) to the convex binding face.

According to the LRRD pattern rule, an additional β-strand (named β-strand X) is adopted at the C-terminal side in all three GPHRs. Based on this match, nine complete LRRs (LRRI-IX) plus one additional β-strand (LRR0) at the N-terminus and plus one at the C-terminal side (LRRX) form 9+2 β-strands lining the inner convex surface of the hormone-binding region (Figure 3.2).

The new LRRD model for the GPHRs offers also a much larger radius for the inner convex arch of the hormone-binding region (Figure 3.2). This resulting radius is rather a shallow deflection, more like a “scythe blade”-shape than a “horseshoe"-shape, demonstrated by the superimposition of TSHR models using the previous and our new template (Figure 3.2).

The LRRD structure of the hNogo-receptor ectodomain contains a ‘Phe spine’ inside the LRRD for the stabilization of the fold, instead of the missing helices at the concave outer face. The importance of phenylalanines of such ‘Phe spine’ motiffor the overall fold, in this case for the GPHRs, is demonstrated by a homozygous mutation at the LHCGR, where a F194V mutation abolishes trafficking to the membrane causing male pseudohermaphroditism (Gromoll J 2002).

This molecular model of the LRRD with the additional repeat eleven attached the C-b2 directly to the LRRD based on the amino acid sequence of the TSHR. Therefore, further modeling and experimental studies on the central extracellular cysteine-boxes were aimed to build complexes between the LRRD and C-b2/3, which is described in the next chapter.

3.1.2 Structural model and functionalities of cysteine-box 2 and 3

3.1.2.1 Molecular models of C-b2 and C-b3 complexed with the LRRD

Systematic searches with fragmented portions of the TSHR ectodomain in the protein structure database for sequence similarities identified a new homologous structural template for C-b2 and C-b3 of the TSHRs ectodomain (Kleinau G 2004).

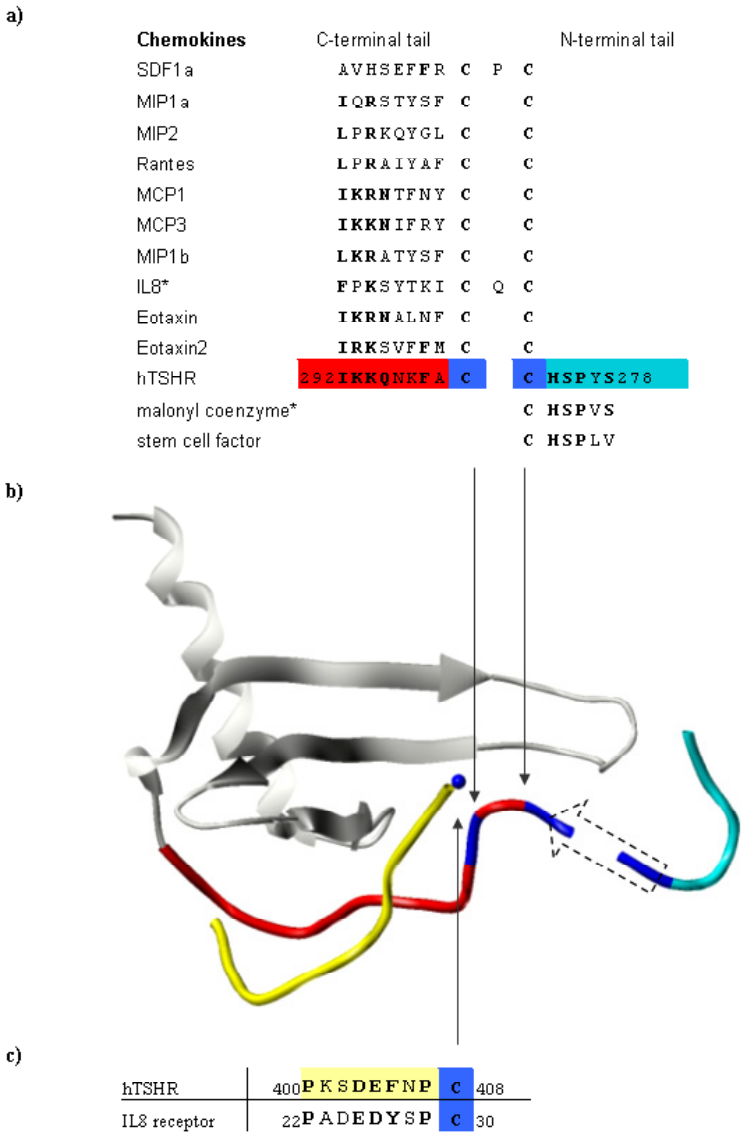


Figure 3.3: ***Homologous structural templates identified by sequence similarity***
a) *Partial sequence alignments for portions with sequence similarity of different chemokines (determined structures) and C-b2 region of the hTSHR upstream (red) of double cysteines C283 C284 (blue). Downstream of double cysteines high sequence similarity (cyan) to fragments with turn conformation from stem cell factor (pdb entry code 1EXZ) and from Malonyl Coenzyme (pdb entry code 1MLA) was identified. Residues with high similarities are marked in bold. Sequence alignment is shown in reversed order to be consistent with the anti-parallel orientation in the structure at Figures b, c. Arrows indicate the corresponding cysteine in the crystal structure. Asterisk marks the used templates in b.* ***b)*** *Conformational portions based on high sequence similarity adopted to the TSHR receptor are highlighted at the structure complex of IL8 monomer (gray) and its bound peptide (yellow) of the N-terminal tail of the IL8RA receptor fragment (pdb entry 1ILQ); red: region of sequence similarity to C-b2 region C284aFxxQKKI292 of TSHR; blue: IL8 cysteines homologous to the TSHR; yellow: IL8RA fragment of the N-terminal extracellular tail with sequence similarity to the TSHR P400xxDEFNPC408, cysteine (TSHR C408) is placed in ideal distance to cysteines of the other chain coming from C-b2 (TSHR C283, C284) to form a disulfide bridge; cyan: turn/loop conformation for a fragment of Malonyl Coenzyme as an additional template fragment for the TSHR consecutively residues P280SxC283.* ***c)*** *Partial sequence alignments of the IL8RA N-terminal peptide and the TSHR Cysteine-box-3 portion, homologous residues are marked in bold. Arrow is indicating the corresponding position in the X-ray structure.*

The sequence of the chemokine from the NMR structure of a complex between chemokine IL8/CXCL8 and chemokine receptor fragment IL8RA/CXCR1 (pdb entry code 1ILQ) (Figure 3.3b) (Skelton NJ 1999) showed high sequence similarity to the C-b2 portion (C283CAFKNQKKI292) of the TSHR (Figure 3.3).

The consensus sequence CCxFxxQKKI was derived from a sequence alignment to other chemokines (Fig 3.3a). Moreover, the bound peptide in the complex structure, representing a portion of the IL8RA N-terminal tail, has high sequence similarity to the TSH receptor sequence of C-b3 that was identified with the consensus sequence pattern P400xxDEFNPC408 (Figure 3.3c).

This structural template applied to the TSHR ectodomain model by substituting amino acid side chains from C-b2 (C283CAFKNQKKI292) and C-b3 (P400KSDEFNPC408) resulted in homologous model residues with complementary properties matching as interacting side chains like the aromatic interaction of F286 in C-b2 and F405 in C-b3 (Figure 3.4b). Subsequently, using this template, C408 of C-b3 is situated ideally close to C283 (or C284) of C-b2 to form a disulfide bridge. Our findings support the hypothesis of disulfide bridges between cysteines C283/C284 of C-b2 either to C398/C408 or to the reverse order C408/C398 of C-b3 (Figure 3.4b).

The TSHR sequence S278YPSHC283 directly following the last β-strand of LRRX and preceding C283 of C-b2 (including S281) matched best with sequences of structural fragments containing a turn/loop conformation such as SVPSHC from the Malonyl Coenzyme (pdb entry 1MLA) and VLPSHC from a stem cell factor (pdb entry 1EXZ) (Figure 3.3a). These findings confirm previous suggestions that S281 is an integral part of a turn or loop conformation (Nakabayashi K 2003) (Figure 3.3b). Such a turn built into our model links the LRRD and C-b2 and places the C-terminal portion of the LRR domain directly above the C-b2/C-b3 linkage (Figure 3.4). The function of this turn as a potential spatial pivot or hinge between LRRD and C-b2/C-b3 is consistent with the data for mutational activation at S281 (Jäschke H 2006 (a)).

The assembled model (Figure 3.4) provides support for a very compact structural arrangement of the C-b1, LRRD, S281 turn/loop, C-b2 and C-b3 proximal to the extracellular end of TMH1. Subsequently, the residues D403EFNPC408 of the C-b3 are located at particularly prominent interface positions of the ectodomain closest to the transmembrane domain. The hydrophilic residues D403, E404 and N406 are therefore be highly likely to participate in the intramolecular signal transduction from the ectodomain towards the serpentine domain and are thus suggested for site directed mutations (Fig 3.4b).

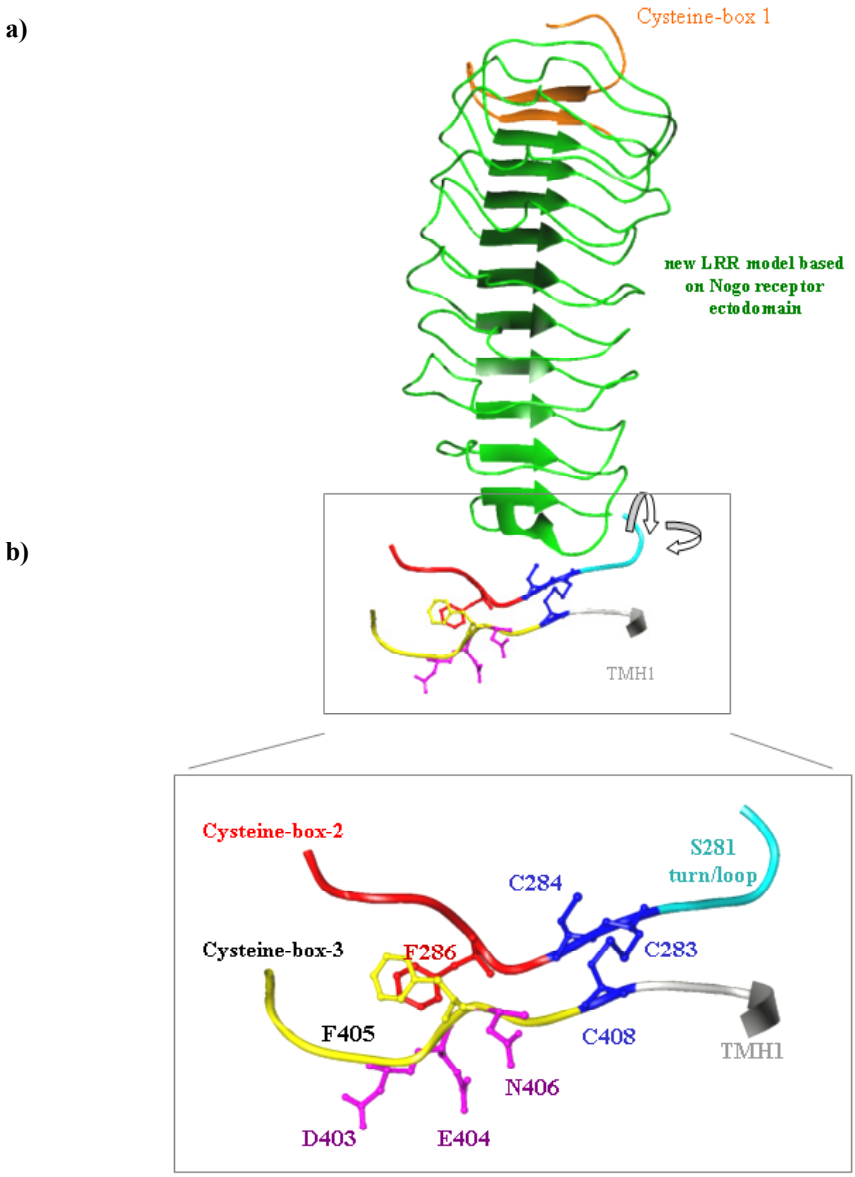


Figure 3.4: ***Homology model of the TSHR ectodomain and details of C-b2 (red) and C-b3 (yellow) interaction***
This assembled model demonstrates a tightly packed structural arrangement of Cysteine-box 1 (orange), LRRD (green), S281 turn (cyan), Cysteine-box 2 (red) and Cysteine-box 3 (yellow). The structural components are adopted from the homologous crystal structures of the a) hNogo-receptor ectodomain and b) the complex of IL8 and the N-terminal peptide of the IL8RA. For clarity remaining portion of C-b2 and C-b3 are not visualized. The backbone conformations of the C-b2 and C-b3 models adopted from IL8 and IL8RA complex structure. Aromatic interaction of F286 (C-b2) and F405 (C-b3); disulfide bridge between C408 (C-b3) and C283 (C-b2) (blue); S281-loop/turn conformation adopted from Malonyl Coenzyme (cyan); hydrophilic residues D403, E404, N406 at C-b3 are located at an interface position at the ectodomain.

Furthermore, based on the model of a tightly packed structural arrangement of C-b2 and C-b3 via disulfide bridges, and considering activation by tryptic clipping (Chen CR 2003), it was

hypothesized, that apart from S281, additional amino acids in C-b2 could also be sensitive for constitutive activation by mutations.

3.1.2.2 Functional characterization of amino acids

To provide support for the hypothesis of functional importance of the tightly packed C-b2 and C-b3 epitopes, the hydrophilic residues of C-b2: K287, N288, Q289, K290, K291, R293 and of C-b3: T399-D410 were tested by single alanine substitutions and side chain variations.

3.1.2.2.1 Mutations of residues in cysteine-box 2

All alanine mutations in C-b2 revealed a cell surface expression of more than 75% compared to wt (Table 3.1). The mutants also showed a basal cAMP accumulation comparable with the wt, except mutation K291A, which is characterized by a 3.3-fold increase of basal cAMP production. K291A also revealed increased constitutive activity compared to the wt as determined by linear regression analysis of constitutive activity as a function of TSHR expression determined by ^{125}I-bTSH binding.

Transfected construct	Cell surface expression	cAMP accumulation		
		basal	stimulated	constitutive activity
	% of wt TSHR	relative to wt basal		fold over wt basal
pSVL	12.8 ± 2.3	0.5 ± 0.1	0.7 ± 0.1	n.d.
TSHR wt	100	1	14.2 ± 1.4	1
C-b2				
K287A	80.3 ± 5.5**	1.4 ± 0.5	24.6 ± 1.2**	0.8 ± 0.1
N288A	82.4 ± 4.6***	1.1 ± 0.1	9.4 ± 1.7*	1.1 ± 0.5
Q289A	78.3 + 0.6**	1.0 + 0.2	8.6 + 0.8***	0.9 + 0.4
K290A	79.5 ± 2.3***	1.4 ± 0.3	20.2 ± 0.9**	0.8 ± 0.2
K291A	76.3 ± 1.3***	3.3 ± 0.9**	20.1 ± 1.6**	2.9 ± 0.4**
R293A	87.6 ± 6.1*	1.4 ± 0.2	15.1 ± 1.6	1.9 ± 0.9

Table 3.1: ***Functional characterization of alanine substitutions in C-b2***
*The mutated TSH receptors were cloned into the expression vector pSVL and transiently expressed in COS-7 cells. Characterization of the constructs was performed by determination of cell surface expression, cAMP accumulation and constitutive activity calculated by linear regression analysis of constitutive activity as a function of TSHR expression investigated by ^{125}I-bTSH binding) (see Experimental Procedures). Data are given as mean ± SEM of three independent experiments, each carried out in duplicate. The wt receptor and empty pSVL vector were used as controls. Statistical analysis was carried out by t test using GraphPad Prism 4.03 for Windows (*** p < 0.001 extremely significant; ** p 0.001 to 0.01 very significant; * p 0.01 to 0.05 significant; p > 0.05 not significant); n.d. not determined.*

In contrast to N288A and Q289A, R293A, which were characterized by a slightly decreased or wt like TSH induced cAMP response, substitutions of K287, K290 and K291 by alanine resulted in an increased TSH stimulated cAMP accumulation.

Since other amino acids in C-b2 (other than the known position S281) were hypothesised to be possible further determinants of the signalling process of the TSHR, the new CAM K291 at the C-terminal part of the C-b2 epitope was identified. This lysine is localized in a cluster of positively charged amino acids that was characterized as a tryptic clipping site (Chen CR 2003). Tryptic action leads also to receptor activation.

3.1.2.2.2 Mutations of residues in cysteine-box 3

Initially, the three predicted hydrophilic amino acids D403, E404 and N406 in C-b3 were experementially studied by site-directed mutagenesis. In order to investigate the functional properties of these residues in intramolecular signal transduction, the positions of the acidic residues D403 and E404 were replaced by a non-polar and a basic amino acid, alanine and lysine, respectively. N406 was substituted by alanine.

Transfected construct	Cell surface expression	^{125}J-bTSH binding	cAMP accumulation		IP accumulation	
	FACS *% of TSHR wt*	B_{max} *% of TSHR wt*	basal *fold over wt TSHR basal*	stimulated *fold over wt TSHR basal*	basal *fold over wt TSHR basal*	stimulated *fold over wt TSHR basal*
TSHR wt	100	100	1	14,4 ± 0.7	2.1 ± 0.1	38.9 ± 1.5
D403A	27.22 ± 0.90	35.5 ± 2.1	1.9 ± 0.1	6.8 ± 0.7	2.9 ± 0.1	2.9 ± 0.4
D403K	12.29 ± 0.75	12.2 ± 0.9	0.9 ± 0.1	3.2 ± 0.4	2.4 ± 0.3	2.5 ± 0.9
E404A	95.40 ± 4.60	89.0 ± 8.6	0.6 ± 0.2	16.7 ± 0.6	2.1 ± 0.2	47.0 ± 0.8
E404K	35.32 ± 0.96	49.7 ± 3.5	4.8 ± 1.0	13.4 ± 0.7	2.3 ± 0.2	10.5 ± 0.9
N406A	66.40 ± 3.60	70.7 ± 7.1	5.2 ± 1.3	12.8 ± 0.6	2.1 ± 0.2	19.6 ± 0.5

Table 3.2: ***Characterization of mutations at the TSHR epitope D403-N406***
The mutated TSH receptors were cloned in the expression vector pSVL and transiently expressed in COS-7 cells. Characterization of the constructs was performed by determination of TSH binding, cell surface expression, cAMP accumulation, IP accumulation. The wt receptor and empty pSVL vector were used as controls.

Three new mutations (D403A, E404K and N406A) in the ectodomain of the TSHR that cause constitutive cAMP activity were identified. Aspartate 403 and asparagine 406 are highly conserved in the GPHR family and the D403A and N406A mutations lead to similar levels of constitutive activity. In contrast, the E404A mutation does not affect basal cAMP accumulation, whereas the E404K mutation leads to an increased basal cAMP accumulation compared to that of the wt TSHR.

After the experimentally confirmed predictions about the influence of the three extracellular C-b3 residues, D403, E404 and N406, their involvement in the receptor activation were

characterized in more detail. The wild type amino acids were replaced with various side chain characteristics by site-directed mutagenesis (Table 3.3) (Müller S & Kleinau G 2006).

D403 - All substitutions at position D403 showed a cell surface expression of 50–80% compared with the wt TSHR, except D403L with 15%. For D403L a receptor expression of 50% was measured by means of the permeabilized cell assay, suggesting an increased receptor accumulation within the cells (data not shown). D403L and D403E revealed a strong increase in basal cAMP accumulation (4 – 7 fold) relative to the wt TSHR. D403S and D403N, with a cell surface expression of about 80% of the wt, exhibited a cAMP signal comparable to that of the wt TSHR. Basal activity calculated by linear regression analysis of constitutive activity as a function of TSHR expression was analyzed to evaluate the effect of receptor density on basal cAMP accumulation independently from their cell surface expression.

Transfected construct	Cell surface expression	cAMP accumulation		
		basal	stimulated	constitutive activity
	% of wt TSHR	relative to wt basal		fold over wt basal
pSVL	11.1 ± 1.5	0.4 ± 0.1	0.7 ± 0.1	n.d.
TSHR wt	100	1	13.8 ± 1.7	1
D403E	57.7 ± 4.7****	4.1 ± 1.2****	14.2 ± 4.3	4.3 ± 1.1*
D403L	15.3 ± 1.8****	6.7 ± 1.3****	9.5 ± 2.1*	21.3 ± 0.4****
D403N	79.2 ± 5.8****	1.4 ± 0.3	17.5 ± 2.8	1.3 ± 0.1
D403S	81.6 ± 1.0****	1.3 ± 0.5	15.5 ± 3.4	1.1 ± 0.0
E404D	84.5 ± 4.5*	1.2 ± 0.1	8.0 ± 0.6*	1.3 ± 0.0
E404L	76.7 ± 6.8**	3.8 ± 1.0****	16.1 ± 4.2	2.0 ± 0.2*
E404N	55.6 ± 4.5****	4.3 ± 0.4****	18.0 ± 1.5	5.6 ± 0.0****
E404S	75.7 ± 6.4***	1.7 ± 0.2****	22.2 ± 5.0	1.8 ± 0.0
N406D	30.2 ± 0.4****	6.8 ± 0.5****	11.9 ± 1.2*	16.2 ± 0.1****
N406K	32.0 ± 2.3****	1.4 ± 0.1	8.9 ± 1.0***	1.5 ± 0.0
N406L	34.3 ± 1.4****	3.0 ± 0.9***	10.8 ± 1.2***	2.5 ± 0.0***
N406Q	70.5 ± 4.6****	4.1 ± 0.3****	16.7 ± 1.7	2.0 ± 0.0*
N406S	73.1 ± 7.9**	3.1 ± 0.5****	16.7 ± 0.8	1.6 ± 0.1*

Table 3.3: ***Functional characterization of the CAM positions in C-b3 of the TSHR by side chain variation***

*For the determination of cell surface expression, constitutive activity calculated by linear regression analysis of constitutive activity as a function of TSHR expression investigated by ^{125}I-bTSH binding as well as cAMP accumulation (see Experimental Procedures), mutated TSH receptors were transiently expressed in COS-7 cells. Data are given as mean ± SEM of three independent experiments, each carried out in duplicate. The wt receptor and empty pSVL vector were used as controls. Statistical analysis was carried out by t test using GraphPad Prism 4.03 for Windows (*** p < 0.001 extremely significant; ** p 0.001 to 0.01 very significant; * p 0.01 to 0.05 significant; p > 0.05 not significant); n.d. not determined*

E404 - All substitutions at E404, except E404D, showed an increased basal cAMP signal (1.7 – 4.5 fold relative to wt TSHR), whereas E404N revealed the highest constitutive activity.

N406 - Substitutions to asparagine, lysine and leucine revealed only a cell surface expression of about 30% of the wt TSHR. All mutants except N406K showed an increased basal cAMP accumulation (3.0 – 6.8 fold relative to the wt).
Based on these results of different mutation phenotypes, interaction patterns or vicinity properties for these 3 signalling sensitive amino acids in C-b3 were developed (chapter *4.1.5*).

To test whether adjacent residues of the predicted CAM region, D403 E404 D406, are involved in the signalling process, their flanking amino acids in C-b3 were investigated by alanine substitution (Table 3.4).
Most of the alanine substitutions (T399A - D410A) gave rise to a cell surface expression 65-100% of the wt TSHR. Mutants P400A and F405A however displayed only 11-16% cell surface expression compared to wt TSHR. P407A and E409A were characterized by a cell surface expression of over 60% compared with wt and showed decreased cAMP accumulation (Table 3.4), indicating the importance of P407 and E409 for hormone induced signalling. The basal cAMP response of P400A is significantly increased (5-fold compared to wt) and showed the highest constitutive activity (Table 3.4).

Transfected construct	Cell surface expression	cAMP accumulation		
		basal	stimulated	constitutive activity
	% of wt TSHR	relative to wt basal		fold over wt basal
pSVL	12.8 ± 2.3	0.5 ± 0.1	0.7 ± 0.1	n.d.
TSHR wt	100	1	14.2 ± 1.4	1
C-b3				
T399A	80.4 ± 7.1**	1.5 ± 0.3	11.6 ± 1.2	1.2 ± 0.3
P400A	11.7 ± 1.6****	5.0 ± 2.0**	7.0 ± 2.5****	220.1 ± 2.2****
K401A	95.4 ± 8.7	1.5 ± 0.2	10.4 ± 0.2	1.7 ± 0.1
S402A	100.5 ± 7.6	1.6 ± 0.3	11.9 ± 0.8	1.2 ± 0.2
F405A	15.8 ± 1.0****	1.0 ± 0.0	1.8 ± 0.2****	n.d.
P407A	64.7 ± 2.2****	1.0 ± 0.1	4.2 ± 0.9****	0.9 ± 0.3
E409A	83.0 ± 4.6****	0.8 ± 0.1	5.4 ± 0.6****	0.7 ± 0.0
D410A	72.6 ± 4.4****	1.0 ± 0.0	10.1 ± 1.8*	0.3 ± 0.0*

Table 3.4: ***Functional characterization of alanine substitutions at C-b3***
*The mutated TSH receptors were cloned into the expression vector pSVL and transiently expressed in COS-7 cells. Characterization of the constructs was performed by determination of cell surface expression, cAMP accumulation and constitutive activity calculated by linear regression analysis of constitutive activity as a function of TSHR expression investigated by ^{125}I-bTSH binding. Data are given as mean ± SEM of three independent experiments, each carried out in duplicate. The wt receptor and empty pSVL vector were used as controls. Statistical analysis was carried out by t test using GraphPad Prism 4.03 for Windows (*** p < 0.001 extremely significant; ** p 0.001 to 0.01 very significant; * p 0.01 to 0.05 significant; p > 0.05 not significant); n.d. not determined.*

Furthermore, the signalling sensitive prolines in C-b3 (P400 and P407) were characterized in more detail by side chain variations to deliniate insights in the detailed molecular function. In

contrast to the alanine mutation, the alterations of P400 to D, K, and L were not constitutively active (Table 3.5). Mutants P407A, K, and L displayed basal cAMP accumulation comparable with wt TSHR and a strongly decreased cAMP response after treatment with TSH (Table 3.5). Interestingly, mutant P407D activated the cAMP cascade constitutively.

Transfected construct	Cell surface expression	cAMP accumulation		
		basal	stimulated	constitutive activity
	% of wt TSHR	relative to wt basal		fold over wt basal
pSVL	11.1 + 1.5	0.4 + 0.1	0.7 + 0.1	n.d.
TSHR wt	100	1	13.8 ± 1.7	1
P400D	9.9 ± 0.8****	0.6 ± 0.1*	0.8 ± 0.1****	n.d.
P400K	16.7 ± 2.8****	0.3 ± 0.1****	1.4 ± 0.2****	n.d.
P400L	11.9 ± 1.7****	0.6 ± 0.1*	3.3 ± 0.4****	n.d.
P407D	84.2 ± 4.5**	2.4 ± 0.4**	4.4 ± 1.1****	2.5 ± 0.1***
P407K	30.2 ± 3.6****	1.2 ± 0.1	7.9 ± 1.4**	1.5 ± 0.2
P407L	55.9 ± 5.8****	1.7 ± 0.0****	5.6 ± 0.3****	1.7 ± 0.0

Table 3.5: ***Functional characterization of side chain variations at prolines 400 and 407***
*COS-7 cells were transiently transfected with various mutant TSHRs. The wt receptor and empty pSVL vector were used as controls. Functional characterization was performed by determination of cell surface expression, cAMP accumulation and constitutive activity calculated by linear regression analysis of constitutive activity as a function of TSHR expression investigated by ^{125}I-bTSH binding). Data are given as mean ± SEM of three independent experiments, each carried out in duplicate. Statistical analysis was carried out by t test using GraphPad Prism 4.03 for Windows (*** p < 0.001 extremely significant; ** p 0.001 to 0.01 very significant; * p 0.01 to 0.05 significant; p > 0.05 not significant); n.d. not determined.*

These mutation phenotypes reveal detailed insights into the functional importance of the two prolines for the regulation of receptor activity states and activation that are discussed in section *4.1.5*.

A structural template has not yet been identified for the linker region between the N-terminal extracellular cysteine-boxes 2 and 3. Therefore this N-terminal receptor component could not be investigated in terms of structure function relationships in this study.
The extracellularly located loops connecting the transmembranal helices are thought to represent the junction between the ectodomain and transmembrane domain since they play a fundamental role in the signalling process of the TSHR and/or in the regulation of the activity states that is displayed by known pathogenic CAMs (Parma J 1995, Duprez L 1997, Fuhrer D 1997, Tonacchera M 1998, Claus M 2005). The extracellular loops are probably involved as signalling-transmitters between the N-terminal components and the TMHs via interactions, or directly contacting the hormone (Ryu KS 1996, Sohn J 2002). Therefore, in the following

sections, the structural-functional description of the ECLs by molecular modeling and mutagenesis will be described.

3.1.3 Extracellular loop 1 and its interaction with cysteine-box 2

3.1.3.1 Molecular model

The ECL1 was investigated in relation to amino acid S281 (CAM position) at C-b2, because interactions between these two receptor components were hypothesized based on our molecular modeling studies.

In detail, cysteine-box 2 (C-b2) (hTSHR: 282-301) and the linked cysteine-box 3 (C-b3) (hTSHR: 390-410) are positioned in close proximity (chapter *3.1.2*). Amino acids of the C-b3 are preceeding the N-terminal border of TMH1 of glycoprotein hormone receptors. It is obvious that the two highly conserved cysteines are close to TMH1 (C398 and C408) and known that they are linked by disulphide bridges to the highly conserved cysteines of C-b2 (C283, C284) (Nagayama Y 1991, Tanaka K 1998). Subsequently, C-b2 and C-b3 are close to each other and must be structurally close to TMH1. Serine 281 at the TSHR is directly beside cysteines 283 and 284 of C-b2 and therefore, it should also be spatially close to TMH1.

Due to 100% sequence homology between 278 SYPSHC 283 and the malonyl coenzyme structure (PDB: 1MLA), the S281 region (hTSHR: 279-281) is expected to form a loop/turn conformation. This joined model of C-b2 and 3, the leucine rich repeat motif, and the serpentine domain favours the orientation of C-b2 and the connected S281 region towards TMH1 (Figure 3.3) (Kleinau G 2004). Based on the structure of rhodopsin, the TMH2 is close to TMH1 and, therefore, S281 is very highly likely to be situated between TMH1 and the TMH2/ECL1 junction.

In the hTSHR serpentine model, the TMH2 α-helix is extended by one turn towards the junction with ECL1. Subsequently, apart from D474 and H478, Y481 is also facing the interior side of the extended helix TMH2 that points to the interior side of the receptor (Figure 3.5). Moreover, the ring systems of Y481 and H478 of the TMH2/ECL1 junction are able to interact with the aromatic system of Y279 from the S281 region in the joined model. Models of S281 mutants indicate that the small alanine interferes neither with the loop/turn conformation nor with the environment.

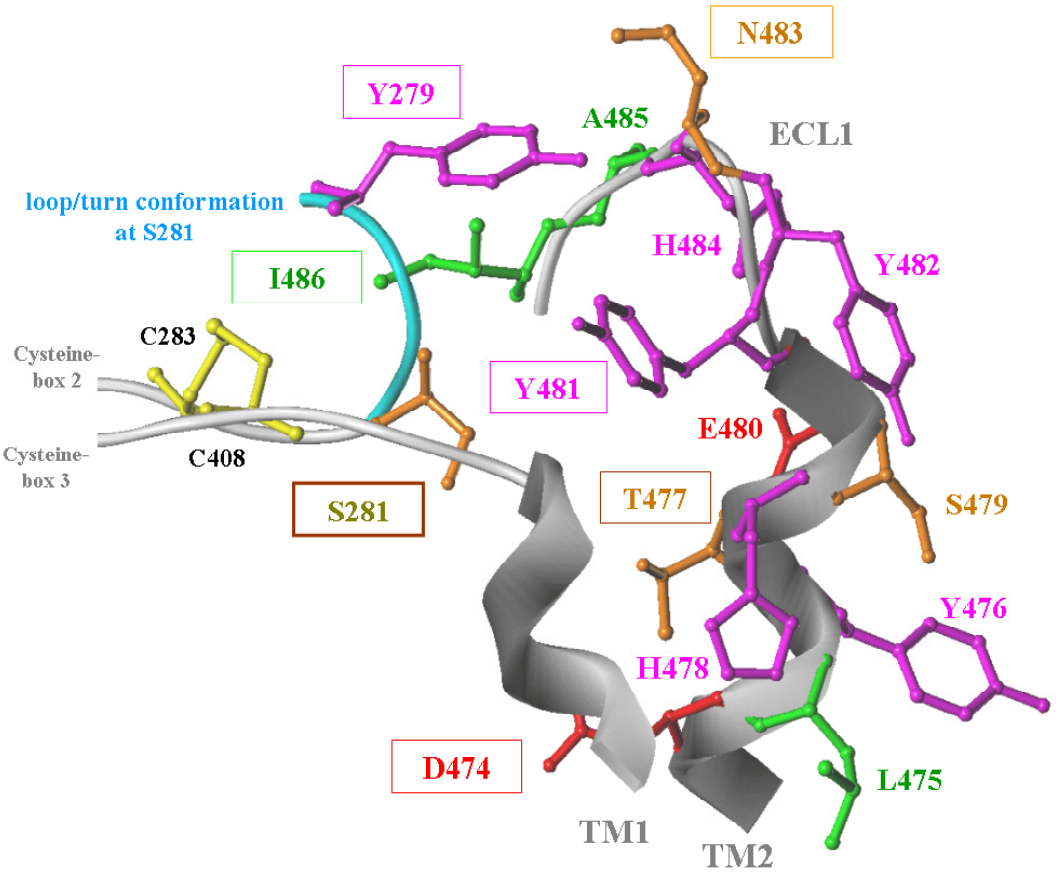


Figure 3.5: ***Three-dimensional model of the hTSHR - TMH2/ECL1 transition, and epitopes of C-b2 and C-b3***
The TMH2/ECL1 junction with amino acids of functional importance (boxed annotations) and a potential orientation of the S281 region are shown in detail. C-b3 organizes the disulfide-linked C-b2 and the S281 loop/turn region (hTSHR: 279-281) closely to the serpentine domain and favours an orientation of S281 towards ECL1. The residues D474, T477, Y481 and H484 of the TMH2 are oriented towards the interior side of the receptor. Aromatic residues Y279 and Y481 are spatially closest to S281. Amino acid coloring: aromatic side chains - magenta, hydrophobic side chains - green, cysteines -yellow, negatively charged side chains - red.

The assumption is, that polar side chains at position 281 and especially those hydrophobic side chains with additional branched bulky extensions at the Cβ atom (like V, I, T) probably clash with residues of the ECL1. Only those residues with an angled side chain and with a shallow aromatic ring system, like tyrosine (e.g. S281Y), are spatially tolerated and can also interact with the ring system of Y481. The remaining part of ECL1 is located in close proximity to the S281 region (Jäschke H 2006 (a)).

3.1.3.2 Functional characterization of amino acids

3.1.3.2.1 Mutations of S281 in cysteine-box 2

The data of mutation studies at S281 (Jäschke H 2006 (a)) are summarized in tables 3.6 and 3.7.

Basal and bTSH-stimulated cAMP accumulation - Specific constitutive activity was determined to evaluate all S281 mutants independently from their cell surface expression. Fifteen mutants at amino acid position 281 (S281L, D, I, V, M, E, N, T, Q, P, G, W, C, H, and F) revealed constitutive activity with varying degrees from 3.6 to 86.4-fold over wt hTSHR basal (Table 3.6, Figure 3.6). Replacement of S281 with the weakly hydrophobic and small A or the slightly hydrophilic Y with an aromatic side chain resulted in a basal cAMP activity comparable to that of the wt hTSHR. S281F, H, W mutants were characterized by only a slight constitutive activity, suggesting that aromatic side chains are tolerated well in the vicinity of S281 (Table 3.6, Figure 3.6).

Transfected construct	**Cell surface expression**	**cAMP accumulation**			**IP accumulation**	
	FACS *% of TSHR wt*	basal *fold over TSHR wt basal*	stimulated	slope *fold over TSHR wt basal*	basal *IPs (%(IPs+PI)*	stimulated
TSHR wt	100	1.0	13.3 ± 1.1	1.0	3.5 ± 0.4	38.1 ± 5.9
pSVL	7.1 ± 2.0	0.2 ± 0.1	0.8 ± 0.2	-	3.4 ± 0.1	3.6 ± 0.4
S281A	98.1 ± 2.8	0.8 ± 0.2	13.5 ± 1.0	0.9 ± 0.1	3.0 ± 0.7	36.7 ± 0.0
S281G	78.4 ± 4.5 **	2.5 ± 0.2 ***	12.7 ± 0.9	4.3 ± 0.6 ***	3.1 ± 0.2	22.3 ± 0.5 **
S281Y	83.1 ± 6.7 *	1.3 ± 0.1	14.4 ± 0.9	1.1 ± 0.1	2.5 ± 0.1	41.5 ± 0.5
S281H	95.3 ± 4.4	3.5 ± 0.6 **	12.8 ± 0.9	3.8 ± 1.1 ***	2.7 ± 0.3	42.4 ± 0.0
S281W	73.1 ± 2.3 **	3.8 ± 0.9 *	13.0 ± 0.3	4.2 ± 0.1 ***	3.1 ± 0.3	36.2 ± 0.1
S281F	83.0 ± 0.5 **	2.3 ± 0.3 **	10.0 ± 0.8	3.6 ± 0.2 **	2.8 ± 0.4	43.6 ± 0.7
S281T	75.8 ± 8.3 *	4.4 ± 0.5 ***	12.1 ± 0.9	7.0 ± 1.3 ***	2.6 ± 0.0	27.0 ± 1.8 *
S281V	44.5 ± 4.5 ***	7.6 ± 0.9 ***	9.6 ± 0.4	14.1 ± 0.3 ***	9.3 ± 2.3 **	41.2 ± 1.7
S281I	22.9 ± 2.4 ***	7.0 ± 0.7 ***	8.2 ± 0.1 *	25.1 ± 7.9 ***	13.1 ± 1.8 ***	23.3 ± 0.4 **
S281D	13.8 ± 3.4 ***	6.4 ± 1.5 **	8.8 ± 0.3 *	72.4 ± 22.8 ***	2.7 ± 0.5	2.9 ± 0.4 ***
S281E	41.9 ± 3.7 ***	7.7 ± 1.5 **	10.6 ± 0.4	9.9 ± 1.1 ***	3.4 ± 0.0	10.8 ± 0.8 ***
S281R	9.3 ± 1.0 ***	0.9 ± 0.1	2.9 ± 0.3 ***	-	2.9 ± 0.3	2.6 ± 0.4 ***
S281K	9.3 ± 0.7 ***	0.8 ± 0.1	3.3 ± 0.4 ***	-	2.9 ± 0.2	2.9 ± 0.4 ***
S281Q	86.0 ± 6.9	7.4 ± 0.9 ***	12.2 ± 1.3	6.9 ± 0.4 ***	3.5 ± 0.8	39.2 ± 2.5
S281N	55.7 ± 6.8 **	5.0 ± 1.6 **	10.5 ± 0.4	7.9 ± 0.1 ***	2.2 ± 0.3	27.4 ± 1.5 *
S281L	19.9 ± 1.0 ***	5.4 ± 0.2 ***	8.3 ± 0.0 *	86.4 ± 20.6 ***	2.8 ± 0.4	4.4 ± 1.2 ***
S281M	38.6 ± 4.0 ***	7.0 ± 1.7 **	13.6 ± 0.5	11.8 ± 1.1 ***	6.5 ± 1.8 *	34.7 ± 0.6
S281C	64.9 ± 3.2 ***	3.5 ± 0.5 **	14.4 ± 1.3	3.9 ± 0.1 ***	2.9 ± 0.2	29.7 ± 0.0 *

Table 3.6: ***Functional characterization of mutations at amino acid position 281 in the extracellular domain.***
COS-7 cells were transfected with wt hTSHR or various mutated hTSHRs. The pSVL vector was used as a control. The hTSHR is characterized by an elevated basal cAMP (40-50 pmol/ml) level compared to the pSVL vector alone (2-10 pmol/ml) (48). Therefore, cAMP accumulation is expressed relative to wt hTSHR basal level. Stimulated levels of cAMP and IP accumulation were determined after treatment of cells with 100 mIU bTSH/ml. Expression of wt and mutant hTSHRs was quantified on a FACS flow cytometer. Data are given as mean ± SEM of three independent experiments, each carried out in duplicate. (p = 0.01 to 0.05; ** p = 0.001 to 0.01; *** p < 0.001)*

In contrast, amino acids with positively charged side chains (S281K, R) led to loss of basal and bTSH-induced cAMP production, possibly due to strongly impaired cell surface expression (Table 3.6). Interestingly, despite very low cell surface expression, S281D was constitutively active (72.4-fold over wt hTSHR basal).

Effects of the mutants on phosphoinositide hydrolysis – All mutants, except S281I, V, showed no increase in basal IP accumulation compared to the wt hTSHR (Table 3.6). It is noteworthy that mutant S281I showed the highest basal IP production despite a cell surface expression of only 22.9% in comparison to the wt hTSHR. However, maximum bTSH-induced IP accumulation was reduced. As a likely consequence of their strongly reduced cell surface expression, mutants S281K, R, D and L revealed no activation of Gαq-mediated signalling in the presence of bTSH.

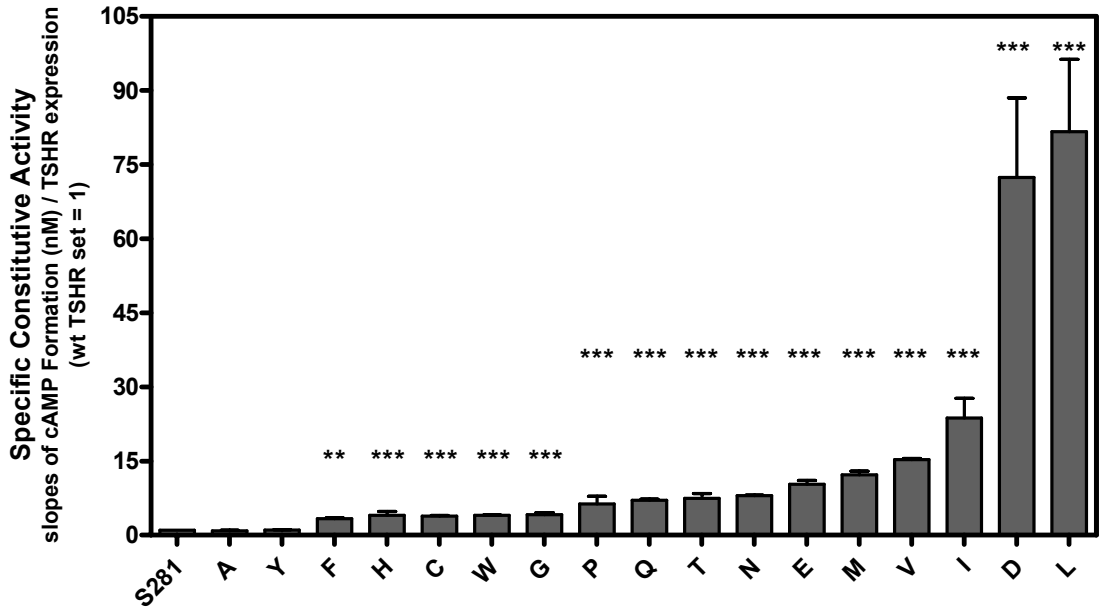


Figure 3.6: ***Substitution of S281 of the hTSHR to all amino acid residues led to various degrees of constitutive activity***
Specific constitutive activity is expressed as basal cAMP formation as a function of receptor expression determined by ^{125}I-bTSH binding. Slopes were calculated by linear regression analysis. Mutants S281R and S281K were excluded from the calculation because of their strongly impaired cell surface expression, which results in loss of ^{125}I-bTSH binding. All data are presented as mean ± SD of two independent experiments, each performed in duplicate.

3.1.3.2.2 Functional assessment of the hTSHR mutants at Y279 (C-b2), Y476, H478, Y481, Y482 and H484 (ECL1)

Basal and bTSH-stimulated cAMP accumulation - None of these mutants showed constitutive cAMP accumulation (Table 3.7). Four mutations, Y476A, H478A, Y482A and H484A, were

characterized by their cell surface expression and Gαs-mediated adenylate cyclase pathway activationafter stimulation with bTSH compared to the wild type (Table 3.7). Cells transfected with Y279A and Y481A displayed strongly impaired cell surface expression and the cAMP response was abolished. In contrast, replacement of Y279 with the large aromatic amino acid W had little effect on both cell surface expression (83.5%) and cAMP production (comparable to wt hTSHR). Similar characteristics were determined for Y481 when substituted by the nonpolar aromatic residue F (Table 3.7).

Transfected construct	Cell surface expression	cAMP accumulation		IP accumulation	
	FACS *% of TSHR wt*	basal *fold over TSHR wt basal*	stimulated	basal *IPs (%(IPs+PI)*	stimulated
TSHR wt	100	1	15.6 ± 1.5	2.2 ± 0.3	45.5 ± 1.1
pSVL	13.8 + 0.1	0.1 + 0.0	0.2 + 0.0	1.6 + 0.2	1.8 + 0.1
Y279A	20.2 ± 1.4 ***	0.3 ± 0.0 ***	4.8 ± 0.1 ***	2.5 ± 0.1	3.3 ± 0.4 ***
Y279N	14.5 ± 0.6 ***	0.8 ± 0.1	5.1 ± 0.6 ***	2.4 ± 0.2	2.4 ± 0.1 ***
Y279W	83.5 ± 3.6	0.8 ± 0.1	13.6 ± 4.1	2.4 ± 0.1	26.1 ± 2.7 **
Y476A	94.0 ± 8.4	1.1 ± 0.1	13.7 ± 2.1	2.5 ± 0.3	48.5 ± 4.4
H478A	103.1 ± 9.7	0.7 ± 0.1 *	15.4 ± 0.1	1.8 ± 0.0	49.8 ± 0.3
Y481A	22.3 ± 1.5 ***	0.6 ± 0.1 *	1.6 ± 0.1 ***	2.3 ± 0.0	2.4 ± 0.0 ***
Y481F	79.3 ± 5.4	1.0 ± 0.1	11.0 ± 1.2	2.7 ± 0.3	31.7 ± 0.6 ***
Y481N	17.2 ± 0.3 ***	0.8 ± 0.1	1.8 ± 0.2 ***	2.5 ± 0.2	2.5 ± 0.1 ***
Y482A	86.3 ± 2.0	0.6 ± 0.1 *	11.6 ± 1.6	2.9 ± 0.3	8.6 ± 0.6 ***
H484A	85.4 ± 2.6	0.5 ± 0.1 **	15.2 ± 1.6	2.4 ± 0.1	17.6 ± 0.6 ***

Table 3.7: ***Functional analysis of substitutions of aromatic residues within the TMH2/ECL1 transition.***
COS-7 cells were transfected with wt hTSHR or various mutant hTSHRs. cAMP accumulation is expressed relative to the wt hTSHR basal level. Stimulated levels of cAMP and IP accumulation were determined after treatment of cells with 100 mIU bTSH/ml. Expression of wt and mutant hTSHRs was quantified on a FACS flow cytometer. Data are given as mean ± SEM of three independent experiments, each carried out in duplicate. (p = 0.01 to 0.05; ** p = 0.001 to 0.01; *** p < 0.001).*

Effects of the mutants on phosphoinositide hydrolysis - As shown for cAMP production, no remarkable differences for mutants Y476A and H478A regarding phosphoinositide hydrolysis were determined when compared to the wt hTSHR (Table 3.7). Cells transfected with Y279A, N and Y481A, N did not show bTSH induced IP accumulation (Table 3.7). Similar to the cAMP production, substitutions to F at position Y481 and to W at position Y279, were well tolerated and activated Gαq like the wt hTSHR. Y482A and H484A, located in the C-terminal region of ECL1, showed wt properties with respect to cAMP accumulation and cell surface expression.

In conclusion, these mutations at ECL1 and S281 at C-b2 highlight the functional importance of the aromatic vicintiy of S281 comprised by amino acids of the ECL1.

Also for the second extracellular loop of the TSHR, pathogenic activating mutations are reported (I568T,V) (Parma J 1995, Claus M 2005), but neither molecular models nor detailed experimental studies were available for sequence structure function analysis. Therefore a project was launched to investigate the general and detailed role of the ECL2 in the TSHR.

3.1.4 Extracellular loop 2

3.1.4.1 Molecular Model

The structural model of the TSHR based on rhodopsin suggests that the ECL2 is plugged, nearly horizontally, into the transmembrane domain on the extracellular side.

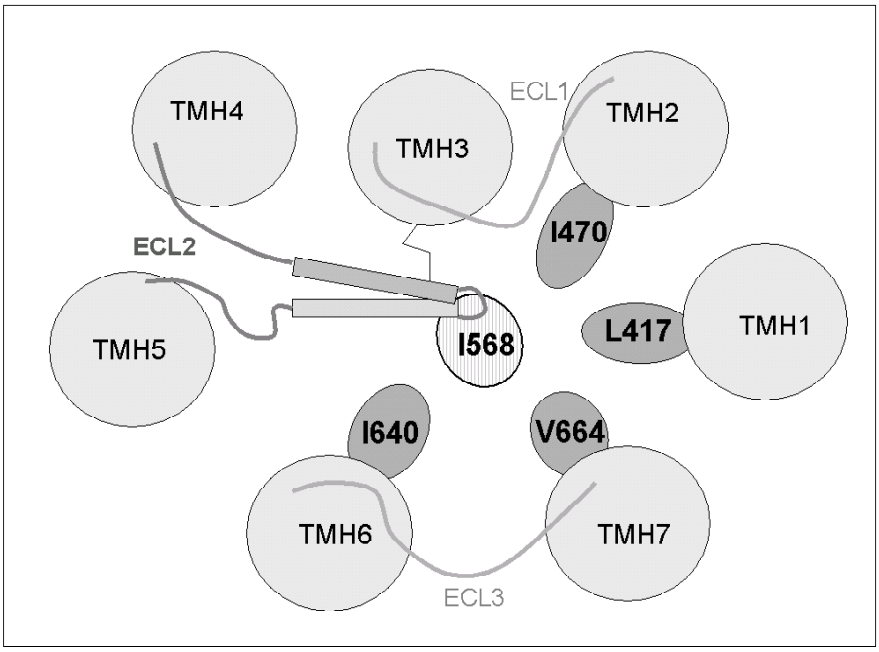


Figure 3.7*: Schematic representation of the localization of I568 at ECL2 and potential interaction partners in the transmembrane region*
The structural model of the TSHR based on rhodopsin suggests a localization of I568 at the tip of ECL2 embedded in a hydrophobic environment formed by the residues L417 (TMH1), I470 (TMH2), I640 (TMH6), and V664 (TMH7).

Pathogenic activating mutations, known at position isoleucine 568 (V, T) (Parma J 1995, Claus M 2005) were the initial points of this ECL2 study, since their molecular mechanism is not understood. Our initial TSHR SD model indicated that I568 is located at the tip of ECL2 and is directly flanked by the conserved C569, which is disulfide bridged to C494 at TMH3

and embedded between hydrophobic residues of TMH1, 2, 6, and 7. The side chain of I568 points downwards in a hydrophobic cleft between the TMHs 1, 2, 6, and 7, and towards potential interaction partners L417 (TMH1), I470 (TMH2), I640 (TMH6), and V664 (TMH7) (Figure 3.7).

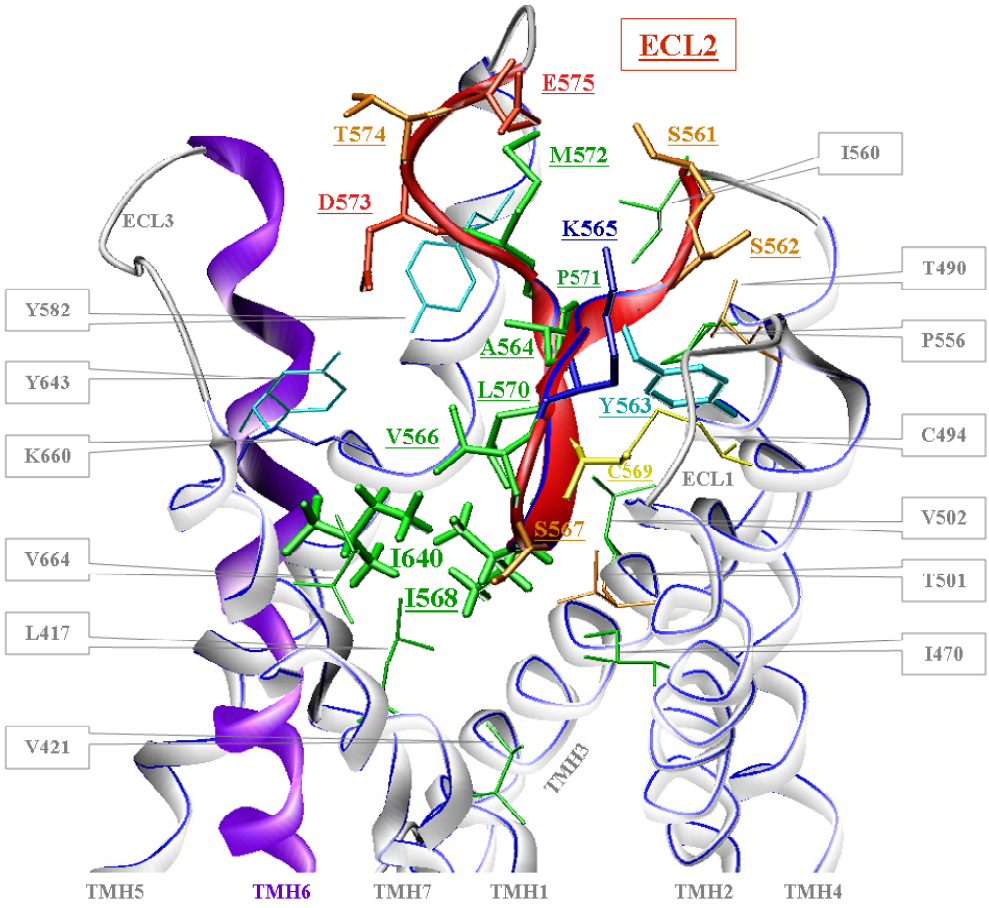


Figure 3.8: ***Molecular model image of ECL2 in the wild type TSHR***
The homology model of the TSHR is based on the X-ray structure of inactive bovine rhodopsin . Considering the partial sequence similarity between the ECL2 of the TSHR and bovine rhodopsin, the complementary environment comprised by the surrounding residues and the conserved disulfide bridge of ECL2-TMH3, a similar embedding of ECL2 into the TM bundle is feasible as in rhodopsin.
Considering for TSHR the ECL2 partial sequence similarity, the complementary of the environmental residues and the conserved disulfide bridge of ECL2-TMH3, a similar embedding of ECL2 into the TM bundle is feasible as in rhodopsin.
Since the ECL2 sequence of the TSHR is four residues shorter than the rhodopsin sequence (see figure 2.4a) the tip of ECL2 is slightly shortened, enabling interaction between I568 (green bold) of ECL2 (red) and I640 (green bold) of TMH6 (violet), which is consistent with our mutation data. All side chains of ECL2 (from S561 to E575) and residues of the TM bundle contacting ECL2 residues are visualized. Side chains of ECL2 amino acids are depicted in bold and labels are colored and underlined corresponding to surrounding amino acids. Color code of amino acids: green - hydrophobic, cyan - aromatic, yellow - cysteines, orange - hydrophilic serine and threonine, red - negatively charged, blue - positively charged.

In order 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 (see chapter *3.1.4.2*). Based on single and double mutations leading to restored wild type basal signalling a direct interaction between ECL2 and TMH6 was identified (see chapter *3.1.4.2*).
In conclusion, based on an initial molecular model of the TSHR, it was suggested and confirmed experimentally hydrophobic interaction partners between ECL2 and TMH6. Based on these findings a refined 3D model of the TSHR was presented, were the amino acid I568 at ECL2 interact directly with I640 at TMH6 (Figure 3.8). This interaction (and therefore the ECL2) is of high importance for the regulation of different activity states via justification of TMH6 (Kleinau G 2006 (a)).

3.1.4.2 Functional characterization of amino acids in the ECL2

All ECL2 residues (Figure 3.9) were substituted by alanine (Table 3.8) and hydrophobic amino acids were also substituted with phenylalanine (Table 3.9). Amino acids C569 and P571 of ECL2 were not considered in this mutagenesis approach since previous studies at both amino acids for the TSHR and the LHCGR (Kosugi S 1992, Ryu KS 1998 (a)) demonstrated that they are essential for correct folding of the receptors. For the FSHR, the threonine mutant of P519 (TSHR: P571) is known as an inactivating pathogenic mutant caused by intracellular trapping of the receptor (Meduri G 2003). Amino acid C569 is disulfide linked to the conserved C494 of TMH3 (Ballesteros & Weinstein number C3.25).

	TMH4								ECL2							TMH5	
hTSHR	561	**S**	**S**	**Y**	**A**	**K**	**V**	**S**	**I**		**C**	**L**	**P**	**M**	**D**	**T**	**E** 575
hLHCGR	506	**S**	**N**	**Y**	**M**	**K**	**V**	**S**	**I**		**C**	**F**	**P**	**M**	**D**	**V**	**E** 520
hFSHR	510	**S**	**N**	**Y**	**M**	**K**	**V**	**S**	**I**		**C**	**L**	**P**	**M**	**D**	**V**	**E** 524
bOPSD	176	**S**	**R**	**Y**	**I**	**P**	**E**	**G**	**M**	**Q**	**C**	**S**	**C**	**G**		**I**	**D**	**Y**	**Y** 192

Figure 3.9: ***Alignment of the ECL2 of human GPHRs and bovine Rhodopsin***
Underlined: different amino acids between the human TSHR, LHCGR and FSHR; Gray background: mutated amino acids of the TSHR

Cell surface expression - FACS analyses revealed that the mutants show a cell surface expression in the range of 50 to 110% of the wt TSHR with exception of mutants Y563A and I568A (Y563A 8.4%, I568A 18.7% of wt) (Table 3.8). The inactivity of the Y563A mutant is

strongly related to its low number of receptors at the cell surface. Therefore, this mutant was not considered in further functional description.

Transfected construct	Cell surface expression	cAMP accumulation		
		basal	stimulated	constitutive activity
	% of wt TSHR	relative to wt basal		fold over wt basal
pSVL	11.1 ± 1.5	0.4 ± 0.1	0.7 ± 0.1	n.d.
TSHR wt	100	1	13.8 ± 1.7	1
D403E	57.7 ± 4.7****	4.1 ± 1.2****	14.2 ± 4.3	4.3 ± 1.1*
D403L	15.3 ± 1.8****	6.7 ± 1.3****	9.5 ± 2.1*	21.3 ± 0.4****
D403N	79.2 ± 5.8****	1.4 ± 0.3	17.5 ± 2.8	1.3 ± 0.1
D403S	81.6 ± 1.0****	1.3 ± 0.5	15.5 ± 3.4	1.1 ± 0.0
E404D	84.5 ± 4.5*	1.2 ± 0.1	8.0 ± 0.6*	1.3 ± 0.0
E404L	76.7 ± 6.8**	3.8 ± 1.0****	16.1 ± 4.2	2.0 ± 0.2*
E404N	55.6 ± 4.5****	4.3 ± 0.4****	18.0 ± 1.5	5.6 ± 0.0****
E404S	75.7 ± 6.4**	1.7 ± 0.2****	22.2 ± 5.0	1.8 ± 0.0
N406D	30.2 ± 0.4****	6.8 ± 0.5****	11.9 ± 1.2*	16.2 ± 0.1****
N406K	32.0 ± 2.3****	1.4 ± 0.1	8.9 ± 1.0**	1.5 ± 0.0
N406L	34.3 ± 1.4****	3.0 ± 0.9**	10.8 ± 1.2**	2.5 ± 0.0**
N406Q	70.5 ± 4.6****	4.1 ± 0.3****	16.7 ± 1.7	2.0 ± 0.0*
N406S	73.1 ± 7.9**	3.1 ± 0.5****	16.7 ± 0.8	1.6 ± 0.1*

Table 3.8: ***Functional characterization of alanine mutations at ECL2***

COS-7 cells were transfected with the wt TSH receptor or described mutant TSH receptors. Because of the basal activity of the wt TSHR, cAMP levels are expressed as relative to wt TSHR basal (set at 1). Increase in cAMP and IP levels was determined after stimulation with 100 mU/ml bovine TSH. The TSH receptor cell surface expression was quantified on a FACS flow cytometer. Data are given as mean ± SD of two independent experiments, each carried out in duplicate. The pSVL vector was used as a control.

cAMP accumulation - Alanine and phenylalanine mutants of I568 were characterized by increased basal Gαs mediated cAMP signalling compared to wt (Table 3.8, Table 3.9).

In contrast, mutants S561A, S562A, K565A, S567A, and M572A displayed decreased basal cAMP accumulation. For K565A, no basal cAMP signalling was observed.

TSH mediated signalling is comparable to wild type or slightly decreased for all mutants (not under 50% compared to maximum of wt) except mutant K565A with strongly impaired signalling activity.

IP accumulation - Basal inositol phosphate levels of all mutants were comparable to that of the wt TSHR (Table 3.8). Ligand stimulated IP accumulation is markedly reduced by alanine mutants of S561, S562, L565, L570, M572 and by V566F (Table 3.9).

Transfected construct	Cell surface expression	cAMP accumulation		IP accumulation	
	FACS	basal	stimulated	basal	stimulated
	(% of wt)	*(relative to wt basal)*			
wt	100	1	15.1 ± 0.1	1.8 ± 0.4	42.3 ± 4.5
A564F	80.9 ± 4.8	0.6 ± 0.1	8.7 ± 0.9	3.0 ± 0.3	19.3 ± 1.3
V566F	55.7 ± 1.7	1.2 ± 0.1	5.6 ± 0.5	2.7 ± 0.1	2.9 ± 0.5
I568F	101.4 ± 7.4	6.3 ± 0.2	15.4 ± 0.5	2.9 ± 0.2	54.1 ± 2.5
L570F	68.9 ± 3.4	1.7 ± 0.3	10.7 ± 2.0	2.9 ± 0.1	18.3 ± 0.6
M572F	92.9 ± 3.5	1.2 ± 0.3	8.7 ± 0.3	2.4 ± 0.3	29.3 ± 2.8
pSVL		0.4 ± 0.1	0.6 ± 0.01	2.0 ± 0.6	2.0 ± 0.6

Table 3.9: ***Functional characterization of phenylalanine mutations at hydrophobic residues in ECL2***
The hydrophobic amino acids within the ECL2 of the TSHR were characterized by phenylalanine mutants. COS-7 cells were transfected with the wt TSH receptor or described mutant TSH receptors. Because of the basal activity of the wt TSHR, cAMP levels are expressed as relative to wt TSHR basal (set at 1). Increase in cAMP and IP levels was determined after stimulation with 100 mU/ml bovine TSH. The TSH receptor cell surface expression was quantified on a FACS flow cytometer. Data are given as mean ± SD of two independent experiments, each carried out in triplicate. The pSVL vector was used as a control.

In summary, several amino acids (K565, I568, M572) in ECL2 are signalling determinants characterized by activation or inactivation of TSHR caused by mutations. This underlines a fundamental role of the ECL2 in the signalling mechanism of the TSHR.

By these mutagenesis studies also we demonstrated, that in addition to the known pathogenic CAMs I568T,V (Parma J 1995, Claus M 2005), mutations I568A,F (Table 3.8, Table 3.9) are CAMs. Therefore, a tight hydrophobic interaction for I568 with its counterpart(s) is hypothesized to constrain the basal wt conformation since slight side chain alterations at I568 lead to a release of these constraints. Moreover, it is assumed that similar modifications at the counterpart side chains (shorter, longer or bulkier compared to the wt TSHR) might also result in constitutive activation of the TSHR.

3.1.4.3 Identification and characterization of amino acids that are involved in the constitutive activation caused by pathogenic mutation I568V

The molecular model of the TSHR suggests that ECL2 is plugged nearly horizontally into the transmembrane domain on the extracellular side (chapter *3.1.4.1*). Isoleucine 568 is located at the tip of ECL2 and is directly flanked by the conserved C569, which is disulfide bridged to C494 at TMH3 and embedded between hydrophobic residues of TMH1, 2, 6, and 7. The side chain of I568 points downwards in a hydrophobic cleft between the TMHs 1, 2, 6, and 7, and

towards potential interaction partners: L417 (TMH1), I470 (TMH2), I640 (TMH6), and V664 (TMH7) (Figure 3.8).

It was assumed, that also a slight side chain alteration (like I->V at the position 568), should also lead to a constitutive activation at the potential interacting residue. Thus, these four potential hydrophobic interaction partners for I568 were tested by constructing valine or alanine mutations (Table 3.10).

Indeed, the I640V mutant in TMH6 with reduced side chain length was a CAM (basal activity 240%, wt TSHR set at 100%). I640 is located at TMH6 close to P639 of the TSHR, which 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.

Location	Transfected construct	Cell surface expression	cAMP accumulation		IP accumulation	
		FACS	basal	stimulated	basal	stimulated
		(% of wt)	*(relative to wt basal)*			
	wt	100	1.0 ± 0.1	7.0 ± 0.9	2.6 ± 0.2	43.1 ± 3.1
TMH1	L417V	98.1 ± 7.4	0.8 ± 0.1	5.6 ± 0.7	3.1 ± 0.2	8.5 ± 0.3
TMH2	I470V	98.8 ± 7.5	0.9 ± 0.1	7.6 ± 1.6	3.1 ± 0.3	32.9 ± 2.0
TMH6	I640V	95.9 ± 8.6	2.4 ± 0.3	10.1 ± 2.1	2.9 ± 0.2	45.5 ± 1.6
TMH7	V664A	75.5 ± 8.5	1.1 ± 0.3	6.4 ± 1.1	2.7 ± 0.7	16.0 ± 0.7
	pSVL		0.6 + 0.1	0.6 + 0.1	2.0 + 0.0	2.1 + 0.0

Table 3.10: ***Functional characterization of single mutants at TMH1, 2, 6 and 7***
Potential hydrophobic interaction partners for I568 were constructed by substituting valine or alanine for the four potential interaction partners in TMH1, 2, 6, and 7. COS-7 cells were transfected with wt or mutant TSH receptors. Increase in cAMP and IP levels was determined after stimulation with 100 mU/ml bovine TSH. Data are given as mean ± SD of two independent experiments, each carried out in duplicate. The pSVL vector was used as a control.

Interestingly, further modifications of I640 to methionine (elongated but flexible side chain), leucine (different branching), and phenylalanine (angled side chain) are not CAMs (Table 3.11). The I640L single mutant showed a decreased cAMP basal activity (basal activity ~40%, wt TSHR set at 100%).

Subsequently, it was opted to test whether this phenotype affects mutant I568V. Indeed, combining the CAM I568V in ECL2 and I640L in TMH6 in the double mutant I568V/I640L (Table 3.11) resulted in restoration of the wild type level of basal cAMP signalling. The cAMP accumulation of the double mutant I568V/I640L after hormone stimulation was comparable to I568V.

Location	Transfected construct	Cell surface expression	cAMP accumulation		IP accumulation	
		FACS	basal	stimulated	basal	stimulated
		(% of wt)	*(relative to wt basal)*			
	wt	100	1.0 ± 0.0	15.7 ± 0.8	2.9 ± 0.0	41.9 ± 0.9
TMH6	I640F	74.2 ± 7.6	0.9 ± 0.1	8.7 ± 0.2	3.0 ± 0.2	7.7 ± 0.1
TMI6	I640M	59.6 ± 2.6	1.2 ± 0.2	9.4 ± 1.6	4.2 ± 0.6	18.2 ± 0.1
TMH6	I640L	92.5 ± 4.7	0.4 ± 0.1	7.8 ± 1.1	3.0 ± 0.1	12.7 ± 0.1
TMH6	I640V	92.1 ± 2.0	2.9 ± 0.6	10.7 ± 1.5	n.d.	n.d.
ECL2	I568V	90.8 ± 2.1	2.9 ± 0.5	12.3 ± 0.6	2.7 ± 0.0	44.9 ± 0.0
ECL2	I568L	101.0 ± 1.2	0.4 ± 0.04	10.7 ± 1.6	n.d.	n.d.
ECL2/TMH6	I568V/I640L	76.1 ± 4.8	1.1 ± 0.0	11.2 ± 1.4	3.0 ± 0.1	18.7 ± 0.1
ECL2/TMH6	I568L/I640V	98.5 ± 5.0	1.0 ± 0.1	9.8 ± 1.6	n.d	n.d.
	pSVL		0.3 + 0.1	0.3 + 0.1	3.5 + 0.1	3.8 + 0.1

Table 3.11: ***Functional characterization of single and double-mutants at ECL2 and TMH6***
Increase in cAMP and IP levels was determined after stimulation with 100 mU/ml bovine TSH. Data are given as mean ± SD of two independent experiments, each carried out in duplicate. The pSVL vector was used as a control. n.d. = not determined

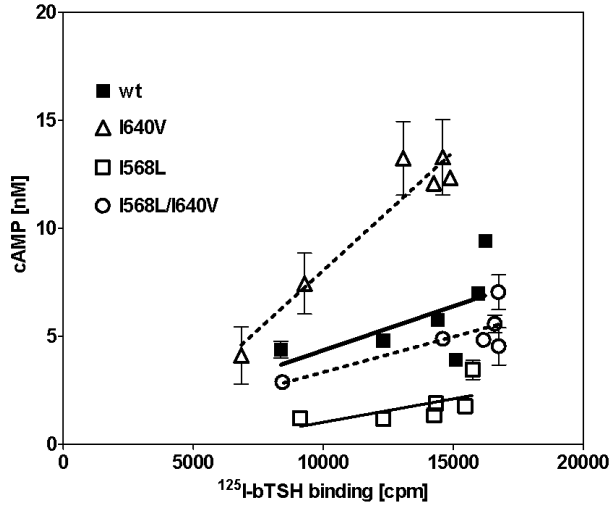
In conclusion, the double mutant between the CAM I568V at ECL2, that is characterized by a shorter side chain compared to wt and the more branched side chain at TMH6 (I640L), restored the basal activity of the wt TSHR. Furthermore, it was predictable that single mutant I568L and the reciprocal double mutation I568L/I640V must have similar phenotypes as shown for I568L (decreased basal activity) and I568I/I640L (restored wt).

We tested the single mutants I568L and the double mutant I640V/I568L. Indeed, as predicted, although I568V and all other known mutants at position 568 are CAMs (I568T,A,F) (Parma J 1995) (Table 3.8, Table 3.9), the mutation I568L is characterized by a basally decreased activity. This phenotype corresponds exactly to the phenotype of I640L. The combination of CAM I640V and mutant I568L with decreased basal cAMP activity in reciprocal double mutant I640V/I568L restored also the basal cAMP activity to that of the wild type. Linear regression analysis of the TSHR mutants I568V and I640V confirmed increased basal cAMP activity (Figure 3.10).

The expression level for mutants I568L, I640V and I568L/I640V is close to 100% and the signalling capacity induced by TSH is about 70% of wild type.

In summary, based on an initial molecular model of the TSHR, suggested hydrophobic interaction partners between ECL2 (I568) and TMH6 (I640) by single and double mutations were experimentally confirmed (Kleinau G 2006 (a)).

A)



B)

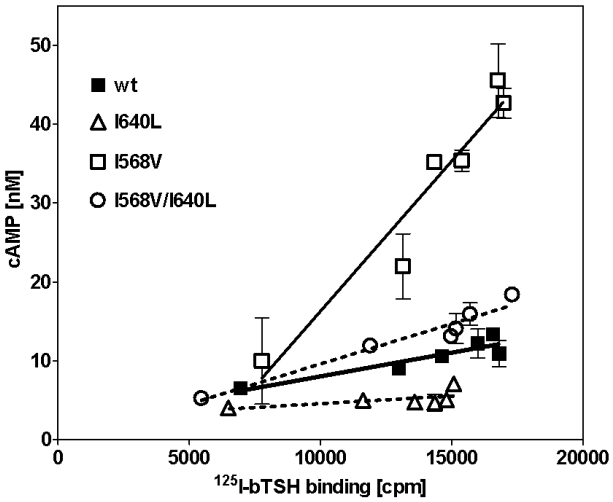


Figure 3.10: ***Linear regression analysis of basal activity as a function of TSHR expression (slopes) for wild type, double mutants and corresponding single mutants***
COS-7 cells were transiently transfected with increasing amounts of plasmids encoding the TSHR wt and mutants. Constitutive activity of mutants was determined by linear regression analysis of basal activity as a function of TSHR expression, investigated by 125I-bTSH binding, using the linear regression function of GraphPad Prism 4.03. The empty pSVL vector was used as control. A representative experiment of three linear regression analyses, each performed in duplicate, is shown.

3.1.5 Extracellular loop 3

3.1.5.1 Molecular Model

The highly conserved residue K660 at the ECL3/TMH7 junction has been described previously to be important for the signalling of the LHCGR (K583) and the FSHR (K590) (Sohn J 2002, Gilchrist GR 1996). To reveal the function and a potential interaction partner in the TSHR, we used the serpentine domain model of the TSHR to choose 3 negatively charged

amino acids for experimental verification. The ECL3 modeling procedure is described in section *2.1.2.5.* By molecular modeling (Claus M 2005), three negatively charged potential interaction partners of the positively charged K660 was determined: E409 (Figure 3.11a) and D410 (Figure 3.11a) in the N-terminus of TMH1, and D573 (Figure 3.11b) located in ECL2.

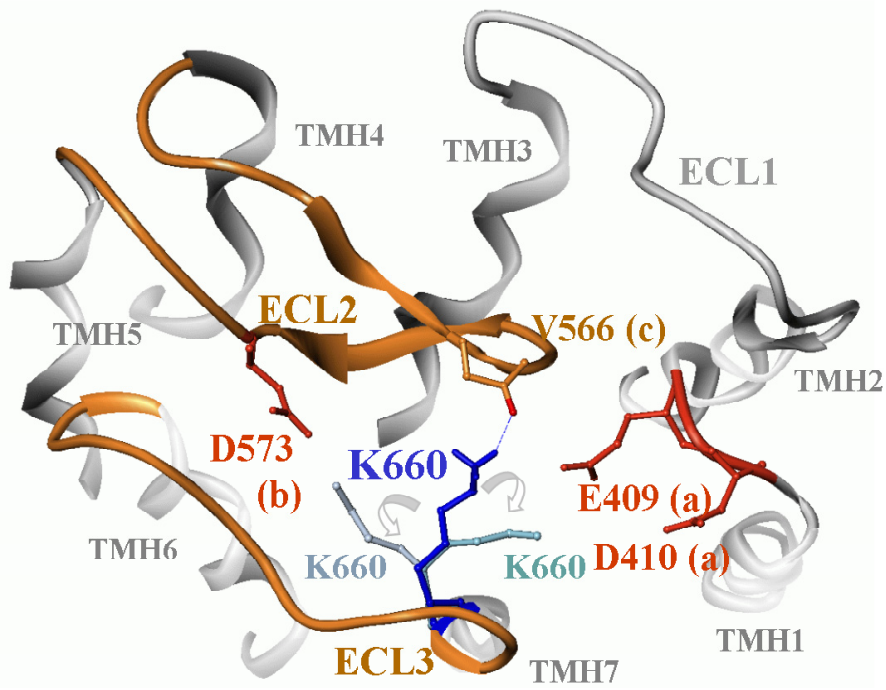


Figure 3.11: ***Backbone visualization of the wt TSHR with proposed orientations and interaction partners of K660 (ECL3)***
Homology model of the TSHR serpentine domain (based on the X-ray structure of rhodopsin) with detailed top view of the extracellular loop portions. Theoretical conformational orientations of the K660 side chain interacting with proposed amino acid partners are shown. Besides the three possible electrostatic interactions between the positively charged (blue) K660 (ECL3) and the negatively charged residues (red orange) E409 or D410 (N-terminus of TMH1) (a) or D573 (ECL2) (b), also the interaction with the backbone of V566 in ECL2 (c) is possible. Experimental verification of interactions between charged amino acids by double mutants excludes an interaction of K660 with one of the acidic residues. Therefore, an orientation of K660 towards the backbone of ECL2 forming a hydrogen bond (blue dashed line) to a backbone carbonyl-oxygen of V566 as a requirement for TSHR signalling-function is favored (c).

However, switching the charges in the respective positions by amino acid substitutions and construction of corresponding double mutants did not rescue the loss of function observed by single substitutions of K660.

Therefore an orientation of K660 towards the backbone of ECL2 forming a hydrogen bond to a backbone carboxyl-oxygen as a requirement for TSHR signalling is favorable (Figure 3.11c). However, since contacts of the K660 side chain to the carbonyl backbone cannot be

studied by standard mutagenesis approaches, this hypothetical interaction remains to be verified.

3.1.5.2 Functional characterization of amino acids in the ECL3

The functional characterization of the TSHR ECL3 alanine substitutions (Claus M 2005) revealed no significant effects on cell surface expression and TSH binding affinity (Table 3.12). In contrast, mutations at the LHCGR positions K573, P575, I577 and N581 (TSHR: N650, P652, I654 and N658) resulted in strongly decreased receptor numbers, whereas expression of the other mutants was comparable to the wt LHCGR (Ryu KS 1996).

GPHR	Helix6	ECL3								Helix7					
		650								660					
TSHR	A	I	L	N	K	**P**	**L**	**I**	**T**	**V**	S	N	S	**K**	I	L
FSHR	A	S	L	K	V	**P**	**L**	**I**	**T**	**V**	S	K	A	**K**	I	L
CG/LHR	A	A	F	K	V	**P**	L	I	T	**V**	T	N	S	**K**	V	L

Figure 3.12*: Influence on cAMP signalling activity of alanine mutations in the ECL3*
ECL3 and junctions to TMH6 and TMH7 of the TSHR, FSHR and LHCGR are shown. Bold: residues which significantly affect cAMP signalling after substitution for alanine; Underlined: residues that, after substitution for alanine, totally abolished cAMP activity. Mutation of L, I and T of the LHCGR does not significantly affect cAMP signalling.

Moreover, alanine mutations in the FSHR ECL3 also caused a decrease of receptor molecules on the cell surface (Ryu KS 1998 (b)). Furthermore, 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.

For the TSHR ECL3, the constitutively activating *in vivo* mutations, N650Y and V656F, are known (Tonacchera M 1995, Fuhrer D 1997). However, mutants N650A and V656A display basal cAMP activity comparable to the wt TSHR (Table 3.12), suggesting that only amino acid enlargement or increased side chain bulkiness, rather than size reduction causes constitutive activation. Substitutions of N650, K651, S657, N658, and S659 near the TSHR ECL3/TMH junctions affected TSH stimulated cAMP accumulation most weakly. In contrast,

substitution of alanine for residues P, L, I, T, V in the central ECL3 had the strongest effect on cAMP production (Table 3.12).

	^{125}I-bTSH binding		Cell surface expression	cAMP accumulation			IP accumulation	
Transfected construct	K_d	B_{max}	FACS	basal	stimulated	EC_{50}	basal	stimulated
	mU/ml bTSH	% of wt TSHR	% of wt TSHR	relative to wt basal		mU/ml bTSH	IPs (% (IP/IPs +PI))	
TSHR wt	5.1 ± 0.0	100	100	1	12.5 ± 1.7	0.4 ± 0.1	2.2 ± 0.1	44.9 ± 3.5
N650A	4.0 ± 1.1	98.6 ± 8.6	82.2 ± 2.5	0.9 ± 0.0	11.8 ± 1.5	0.2 ± 0.03	1.5 ± 0.1	26.9 ± 1.1 **
K651E	4.5 ± 0.4	96.7 ± 1.1	94.0 ± 0.1	0.7 ± 0.1	16.0 ± 1.1	0.2 ± 0.04	2.7 ± 0.1	33.9 ± 0.5 **
P652A	3.2 ± 0.5	70.5 ± 0.3	66.3 ± 3.4	0.9 ± 0.2	8.2 ± 0.8 *	0.2 ± 0.01	1.7 ± 0.1	18.2 ± 0.0 **
L653A	3.9 ± 0.7	80.7 ± 5.0	84.6 ± 3.2	0.8 ± 0.1	7.8 ± 0.7 **	0.6 ± 0.1	2.1 ± 0.0	8.0 ± 0.8 **
I654A	2.6 ± 0.2	70.4 ± 4.6	80.3 ± 2.5	0.8 ± 0.2	6.3 ± 0.2 **	1.3 ± 0.1**	2.4 ± 0.3	2.4 ± 0.1 **
T655A	4.5 ± 0.6	89.8 ± 7.7	91.8 ± 4.2	0.7 ± 0.3	6.4 ± 0.2 **	0.4 ± 0.02	2.4 ± 0.1	5.3 ± 0.8 **
V656A	3.7 ± 0.3	84.1 ± 6.7	85.6 ± 7.2	1.5 ± 0.3	4.7 ± 0.1 **	1.3 ± 0.1**	2.7 ± 0.5	2.6 ± 0.2 **
S657A	4.3 ± 0.2	91.7 ± 5.3	103.1 ± 3.2	0.8 ± 0.1	9.8 ± 1.1	0.6 ± 0.1	2.1 ± 0.3	19.7 ± 0.2 **
N658A	2.7 ± 0.1	70.7 ± 2.1	92.5 ± 2.4	0.8 ± 0.1	13.3 ± 0.6	0.3 ± 0.02	2.3 ± 0.1	21.4 ± 0.3 **
S659A	6.9 ± 0.5	108.3 ± 3.1	107.2 ± 4.8	1.4 ± 0.2	14.8 ± 0.6	0.5 ± 0.01	2.4 ± 0.3	42.8 ± 2.4
K660A	3.0 ± 0.4	69.0 ± 5.2	80.0 ± 3.6	0.8 ± 0.2	7.1 ± 1.0 **	0.8 ± 0.2	2.5 ± 0.1	2.9 ± 0.3 **
pSVL			4.9 ± 0.4	0.6 ± 0.0	0.5 ± 0.1		2.5 ± 0.4	2.6 ± 0.4

Table 3.12: ***Functional characterization of single alanine mutants in the ECL3***
*COS-7 cells were transfected with the wt TSHR or various TSHR mutants. The constitutive activity of TSHR is evident from elevated basal cAMP levels of cells expressing the wt TSHR in comparison to those expressing the pSVL vector alone. cAMP production is expressed relative to the wt TSHR basal level. Stimulated response of cAMP and IP accumulation was determined after treatment with 100 mU bTSH/ml. B_{max} and K_D values were determined by homologous competitive binding experiments. The TSHR cell surface expression was quantified on a FACS flow cytometer. The pSVL vector was used as a control. Data are given as mean ± SEM of three independent experiments, each carried out in duplicate. * $p<0.05$; ** $p<0.01$.*

However, contrary to the FSHR properties (Ryu KS 1998 (b), Sohn J 2002), TSHR mutants did not totally abolish cAMP formation but displayed about 50% decreased cAMP activity compared to the wt 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 the TSHR and the FSHR but not for the LHCGR, a hydrophobic cluster in the center of ECL3 is most probably important for ligand-induced activation of the Gαs-mediated signalling pathway. This suggests that the ECL3 of the LHCGR may act in a different manner in the process leading to Gαs activation (Figure 3.11). 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.

Despite high sequence homology within ECL3 (Figure 3.11), 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. Our results underline, however, that the highly conserved K660 at the junction between ECL3 and TMH7 is essential in a similar manner for both signalling pathways in all 3 GPHRs. Moreover, our experimental data provide evidence for a hydrophobic cluster, comprising residues 652 - 656 in the central part of ECL3, which is important for intramolecular signal transduction and G-protein activation by the TSHR.

Furthermore our studies finally led to a refinement of the initial TSHR model including components of the serpentine domain (e.g. ECLs 1-3). This model was used for docking studies between the TSHR (LHCGR) and an agonistic LMW (Moore S 2006) that binds in an allosteric binding pocket in the transmembrane domain (Jäschke H 2006 (b)).

Using a combination of molecular modeling and evaluation of the binding of org41841 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 (b)). Specifically, org41841 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 the TSHR are remarkable in that they are not likely to compete with the large native hormones for binding at the extracellular N-terminal domain.

3.2 Modes of binding of a small agonistic molecule

3.2.1 Molecular models of the TSHR and the LHCGR in complex with a LMW agonist

In this study, ICM software (Abagyan R 1994) was used to dock the compound, org41841 (Figure 3.13), within the previously described binding pockets of the TSHR and the LHCGR and the binding modes were analysed in detail.

The similarities and differences of the amino acid residues between the receptor subtypes were studied to elucidate those that form and cover the binding pocket for the two receptors. To facilitate the comparative relationship between selected residues of each receptor, a GPCR residue indexing system was utilized (i.e. E3.37 for both E506 at the TSHR and E451 at the LHCGR) (Ballesteros JA 1995, Costanzi S 2004).

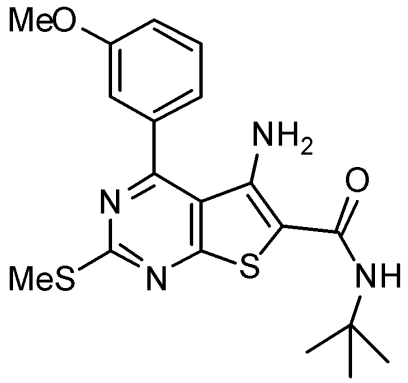


Figure 3.13: ***Org41841 – a LMW agonist for the LHCGR***
The thienopyrimidine org41841 was identified as a high potency agonist (EC_{50}: 20 nM) for the LHCGR. It is demonstrated that org41841 acts as a partial agonist at TSHR with a lower potency (EC_{50}: 7400 nM) than at the LHCGR.

Several conserved amino acids, such as the key glutamate (E3.37), as well as non-conserved amino acids, such as the F → T (position 5.42), Y → F (position 6.53) and L → F (position 570/515), were deemed relevant for the binding of 3 at both the LHCGR and the TSHR. The complete comparison of the amino acid residues that form the putative binding pockets is shown in table 3.13.

The resulting models of these docking studies represent an extension of our previous work on the TSHR (Jäschke H 2006 (b)) and constitute the first reported docking models of this class of ligands to LHCGR. In part, the models confirm our supposition that the LHCGR contains a larger binding pocket than the TSHR.

Binding mode A (Figure 3.14) - In the first binding mode, the 3-methoxyphenyl group of the compound is located at the same depth from the extracellular plane of the membrane as the amino acids 5.42 (TMH5), 4.60 (TMH4) and 3.42 (TMH3). Nine of the amino acids (Table 3.13, underlined) that directly surround the 3-methoxyphenyl ring system are different in the two receptors.

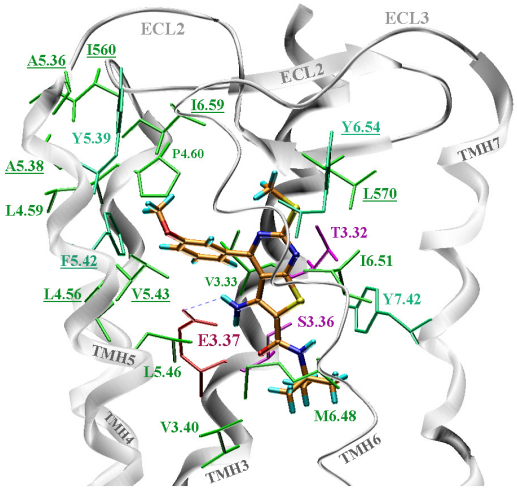
In the TSHR, the 3-methoxyphenyl is in an offset stacked aromatic-aromatic interaction with F5.85 (CG/LHR T5.85). In the LHCGR, the aromatic 3-methoxyphenyl group is embedded in a hydrophobic cavity between TMH4 and TMH5 and the plane of the phenyl ring is perpendicular to the plane of the cell membrane. Consequently, in LHCGR, the long axis of the thieno[2,3-d]pyrimidine moiety is not oriented perpendicularly to the plane of the membrane, as in the case of the TSHR.

Region	Pos.	TSHR	LHCGR
TMH3	3.33	Val 502	Val 447
	3.36	Ser 505	Ser 450
	3.37	Glu 506	Glu 451
TMH4	4.56	Leu 552	Ile 497
	4.57	Ala 553	Ala 498
	4.59	Leu 555	Leu 500
	4.60	Pro 556	Pro 501
ECL2		Ile 560	Val 505
		Ser 561	Ser 506
		Leu 570	Phe 515
		Pro 571	Pro 516
		Met 572	Met 517
		Asp 573	Asp 518
TMH5	5.34	Pro 577	Thr 522
	5.36	Ala 579	Ser 524
	5.37	Leu 580	Gln 525
	5.38	Ala 581	Val 526
	5.39	Tyr 582	Tyr 527
	5.42	Phe 585	Thr 530
	5.43	Val 586	Ile 531
	5.46	Leu 589	Leu 534
	5.47	Asn 590	Asn 535
TMH6	6.48	Met 637	Met 582
	6.51	Ile 640	Ile 585
	6.52	Ser 641	Ser 586
	6.53	Tyr 643	Phe 588
	6.54	Ala 644	Ala 589
	6.59	Ile 648	Ala 593
TMH7	7.39	Val 664	Val 609
	7.42	Tyr 667	Tyr 612

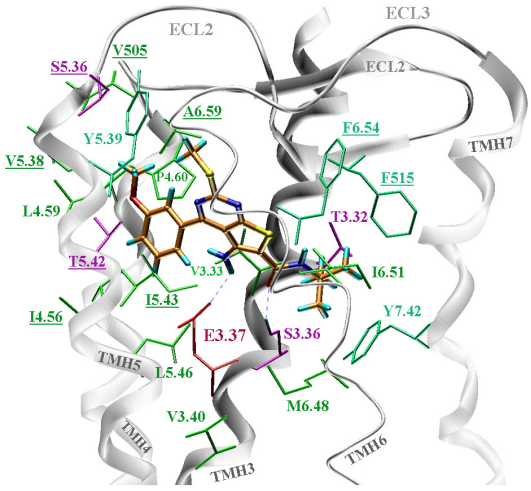
Table 3.13: ***Comparison of the binding pockets of Organon 41841 in the TSHR and in the LHCGR***
The binding pockets of the TSHR and the LHCGR are composed of 32 amino acids, 11 of which are divergent (underlined).Amino acid numbering is given by the sequence (including signal peptid) and by the Ballesteros-Weinstein numbering.

As another consequence of the differences in the amino acids that surround the phenyl moiety, the centroids of the docked compounds are shifted nearly 3Å further towards TMH4 and TMH5 in the LHCGR than in the TSHR.

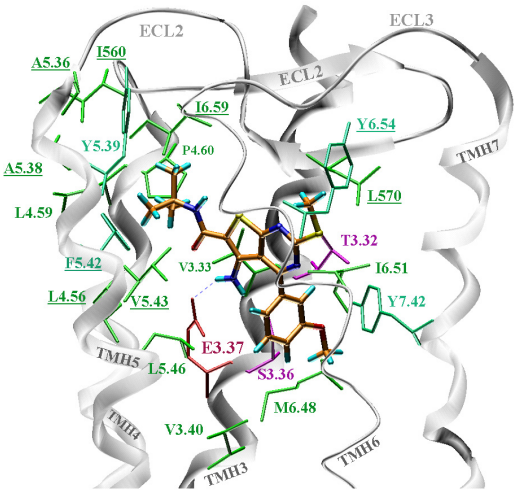
TSHR – 3 docking complex: Orientation A



LHCGR – 3 docking complex: orientation A



TSHR – 3 docking complex: Orientation B



LHCGR – 3 docking complex: orientation B

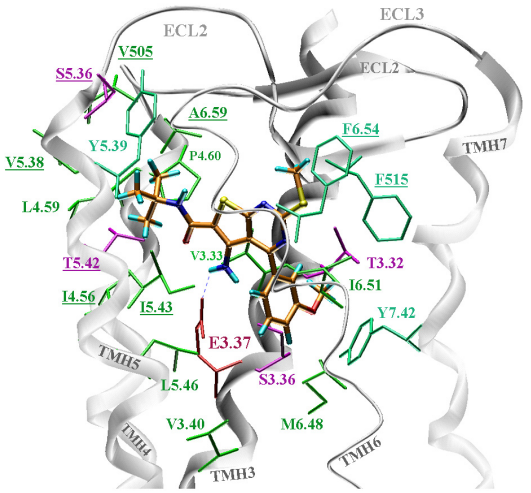


Figure 3.14: ***Docking poses of Org 41841 (3) within homology models of TSHR and LHCGR optimized to orientations A and B***
The binding-pocket is located within the extracellular half of the transmembrane helical bundle between TMH 3, 4, 5, 6, 7 and ECL 2 (backbone-ribbons in white). Amino acids that are not conserved between the two receptors are underlined. Coloring: green = hydrophobic side chains; green/blue = aromatic side chains; magenta = hydrophilic side chains; red = negatively charged side chains. Compound 3 is shown in orange with colored atoms.

In the LHCGR the carboxamido moiety is accommodated between amino acids T3.32, S3.36, F515 (ECL2), I6.51, F6.53 and Y7.42. Conversely, in the case of the TSHR it is oriented vertically and points toward M6.48. An H-bond between the carbonyl oxygen of the

carboxamido moiety of the three docked ligands and the amino acid S3.36 (TMH3) is present only in the case of LHCGR.

Binding mode B (Figure 3.14) - in the case of the second binding mode, the carboxamido moiety of the compound is located at the same depth from the extracellular plane of the membrane as the amino acids 5.42 (TMH5), 4.60 (TMH4) and 3.42 (TMH3), while the 3-methoxyphenyl moiety points toward M6.48. Due to the described differences in the binding pockets, the carboxamido moiety of the compounds is closer to TMH4 and TMH5 in LHCGR than in TSHR. The relative distances of the docked compounds from TMH4 and TMH5 in the TSHR and LHCGR (compound - distance at the TSHR, distance at the LHCGR) are 9.2Å and 7.7Å, respectively.

The predicted interaction best supported by experimental data is the H-bond observed between the amine functionality of org41841 and E3.37 (Jäschke H 2006 (a)). This potential interface was examined by converting the amine of org41841 to a dimethylamine moiety in compound 20 thus eliminating the H-bond donation capacity of the small molecule (Figure 3.15).

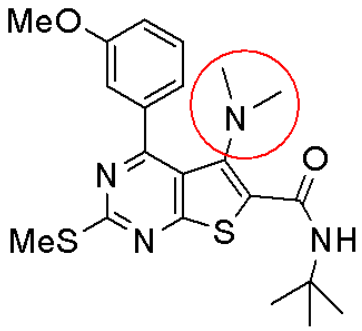


Figure 3.15: Dimethylamine moiety at compound 20
Initial docking experiments suggested a potential hydrogen bond between the amine functionality of org41841 and E3.37 in transmembrane helix 3 of both TSHR and LHCGR. To fully examine this it was chosen to eliminate this potential interaction using the small molecule as a point of manipulation. The dimethylation accomplished the exclusion of H-bond donation capability.

We observed that each of these presented changes at the TSHR or the LMW markedly reduce agonistic activity at the TSHR and, in terms of the activity of compound 20, LHCGR as well (Figure 3.16). This provides compelling experimental evidence supporting the existence of a critical H-bond between the small molecule and E3.37 and our proposed molecular docking arrangements of org41841within both the LHCGR and the TSHR.

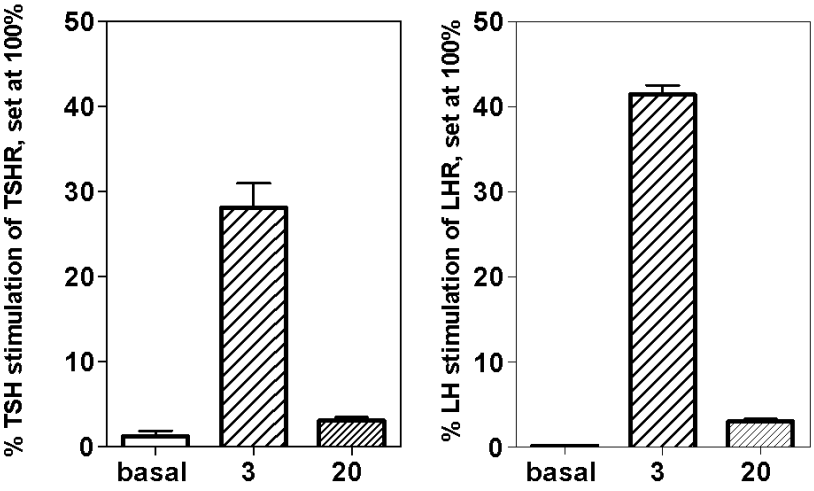


Figure 3.16: ***Analysis of org41841 (3) and dimethyamine org4184 (20) at both the TSHR and LHCGR***
Comparison of activation of the TSHR and the LHCGR by compounds 3 and 20 relative to basal activities of both receptors. Intracellular cAMP production was determined in response to 100 µM of each compound and is expressed as % of maximum response of TSHR/LHCGR to TSH (100 mU/ml)/LH (1000 ng/ml). The data are presented as mean ± SEM of two independent experiments, each performed in duplicate.

3.3 A semi-quantitative data set for Sequence-Structure-Function Analysis at GPHRs

3.3.1 Data Set

As mentioned in the introduction (*1.3.2.*3), there are mutation databases for GPCRs that collect information concerning several types of mutations or functional studies (point-substitutions, deletions, mutagenesis studies, *in vivo* mutations etc.) in a non-quantitative manner (Kristiansen K 1996, Horn F 1998, University of Leipzig, www.uni-leipzig.de/innere/TSH). Therefore, a dataset of available GPHR phenotypes (manually) was designed and created that enables a semi-quantitative analysis of experimental data combined with structural information. In combination with bioinformatic tools (focused search functions, free adjustable output formats), this information source was a prerequisite to summarize and to evaluate known and new functional data. Finally; this tool is helpful for the evaluation of new hypotheses concerning intramolecular interactions (visualization of phenotypes at 3D structures), and it helps to identify signalling cascades or important areas for specific receptor functions (Kleinau G 2006 (b)).

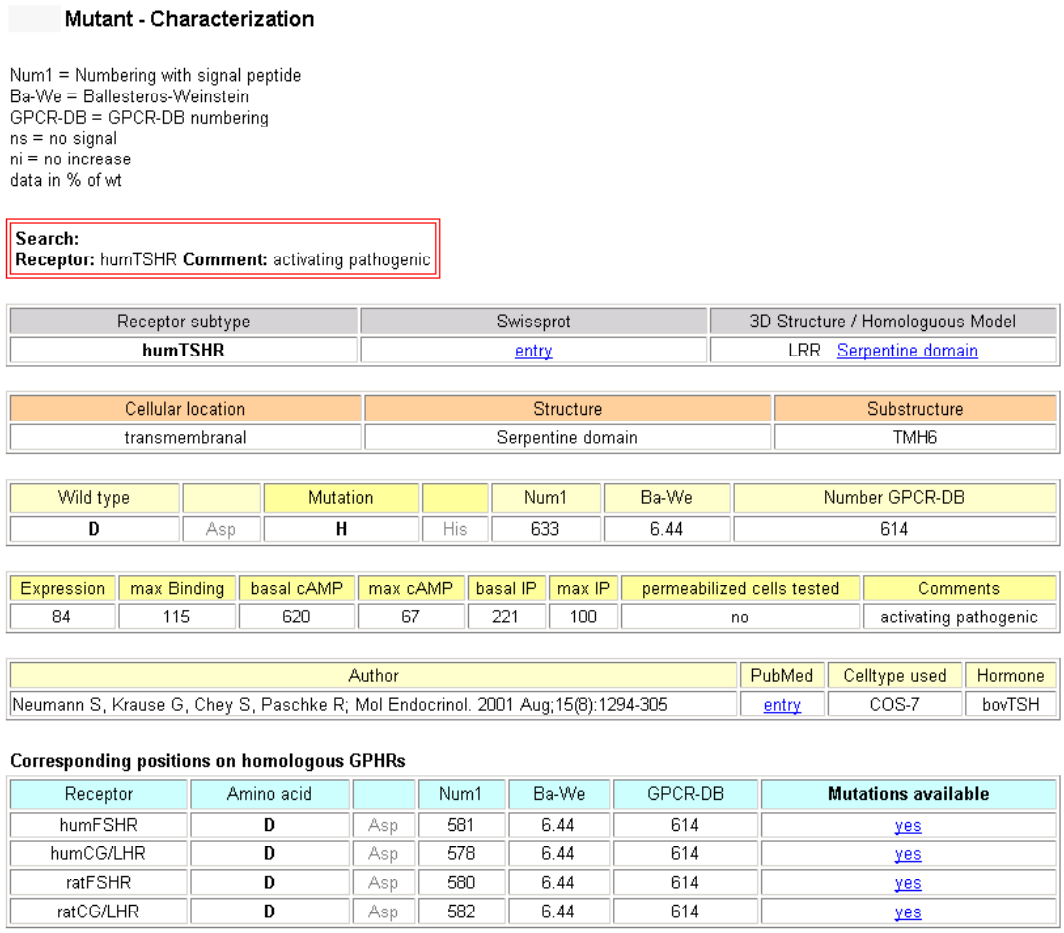
Utilizing the influence of single side chain substitutions on structure-functional effects, point mutational data are extracted from the literature as described in section *2.1.3.1*. About 900 mutants containing specific and basic information for the human TSHR, human LHCGR, human FSHR, rat LHCG, and for the rat FSHR were included in the data set at time.

Phenotypes of receptor mutants represent a certain activation state of the receptor and are therefore helpful for the analysis of activation mechanisms (chapter 1.3, Figure 1.11).
The data from the following standard assays for wild type and mutant phenotypes are included: a. cell surface expression level, b. hormone binding capability (maximum), c. basal Gαs mediated activity (cAMP accumulation), d. basal Gαq mediated activity (IP accumulation), e. Gαs mediated activity after hormone treatment (maximum), and f. Gαq mediated activity after hormone treatment (maximum).
The original data of each publication and assay are scaled to unified percentage values, which are calibrated to the corresponding wild type values of 100% and rounded up or down to the closest decimal place. Results presented as diagrams in publications are annotated simply with absolute values, therefore, the data extracted from such diagrams is not highly accurate. These mutations are marked in the comment field. However, in contrast to known mutation databases with GPCRs (Kristiansen K 1996, Horn F 1998) or GPHRs exclusively (van Durme J 2006, www.uni-leipzig.de/innere/TSH), a functional data set from mutation studies compiled for a semi-quantitative sequence-structure-function analysis is provided.
In addition to the functional data, also experimental conditions such as the cell system used (e.g. COS, HEK) or the type of hormone used for each experiment to enable comparison of data from the same or different experimental conditions are included. Additional features are the comparison of corresponding amino acids between all three GPHR sub-classes, the numbering using the three most popular numbering systems and the citation of the original study. Keywords in a 'comment' field are searchable via the 'Advanced Search', e.g. 'pathogenic mutations' or mutations causing promiscuous hormone binding.

3.3.2 SSFA Tools

For a sequence-structure-function analysis, the mutation data set is combined with different tools for focused searches and generating different output formats. Regardless of type of sequence numbering, a multiple sequence alignment with a unified numbering system allows an easy localisation of any residue number. This not only enables a helpful overview of which residues occur at corresponding positions in homologous GPHR but also whether mutational data is available (Figure 3.17). 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 (Ballesteros JA 1995); numbering by the GPCRDB); **c.** structural epitopes

or domains, **d.** specified amino acid or mutation properties, **e.** type of mutation (change of and to specific residue properties), **f.** functional characteristics of mutations that allow a combinatorial search for various parameters and **g.** different comments, e.g. *pathogenic mutations*



Mutant - Characterization

Num1 = Numbering with signal peptide
Ba-We = Ballesteros-Weinstein
GPCR-DB = GPCR-DB numbering
ns = no signal
ni = no increase
data in % of wt

Search:
Receptor: humTSHR **Comment:** activating pathogenic

Receptor subtype	Swissprot	3D Structure / Homologuous Model
humTSHR	entry	LRR Serpentine domain

Cellular location	Structure	Substructure
transmembranal	Serpentine domain	TMH6

Wild type		Mutation		Num1	Ba-We	Number GPCR-DB
D	Asp	**H**	His	633	6.44	614

Expression	max Binding	basal cAMP	max cAMP	basal IP	max IP	permeabilized cells tested	Comments
84	115	620	67	221	100	no	activating pathogenic

Author	PubMed	Celltype used	Hormone
Neumann S, Krause G, Chey S, Paschke R; Mol Endocrinol. 2001 Aug;15(8):1294-305	entry	COS-7	bovTSH

Corresponding positions on homologous GPHRs

Receptor	Amino acid		Num1	Ba-We	GPCR-DB	**Mutations available**
humFSHR	**D**	Asp	581	6.44	614	yes
humCG/LHR	**D**	Asp	578	6.44	614	yes
ratFSHR	**D**	Asp	580	6.44	614	yes
ratCG/LHR	**D**	Asp	582	6.44	614	yes

Figure 3.17: ***Detailed output***
Initially, the detailed output of queries initiated by the 'Advanced Search' and 'Basic Search' sections give an overview output of general information, like the availability and numbering of mutants. From this point, the user can choose the 'Details' option to obtain several specific pieces of information for the extracted mutant: query options overview, localization of mutated amino acid, the three different types of numbering, assay values that are available, comments, authors, and corresponding positions at homologous GPHRs. Here the visualization on a 3D model for each retrieved mutation is provided (if the mutation occurs in the SD or LRR domain).

Data-Analysis - The 'Advanced Search' functions allows combinations of queries for specific assays used for 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. An analyzing strategy that filters distant values is also implemented. Therefore, our semi-quantitative tool is designed to classify the functional data into two rough classes. A freely adjustable coloring-system allows easy discrimination of similar and different

functional effects of mutations and their corresponding properties by color coding tabulated results (Figure 3.18).

Compressed results

Search:
Receptor: humTSHR **Location:** all locations **Gs basal:** 1-9999 **Gs stimul:** 1-9999 **Expression:** 1-9999 **ni:** yes

Num1 = number with signal peptide
ns = no signal
ni = no increase
data in % of wt

Expression <50%

basal cAMP >149%

max cAMP <50%

Receptor subtype	Swissprot	3D Structure / Homologuous Model
humTSHR	entry	LRR Serpentine domain

64 hits were found (positions are mutated).

Cellular location				Structure		Sub-structure			
extracellular				hinge		Cysteine-box-3			
Wild type	Mutation	Num1	PubMed/Medline	Expression	max Binding	basal cAMP	max cAMP	basal IP	max IP
D	**A**	403	entry	27	35	47	36	138	7
D	**K**	403	entry	12	12	90	22	114	6
E	**A**	404	entry	95	89	60	116	100	121
E	**K**	404	entry	35	50	480	93	110	27
N	**A**	406	entry	35	50	520	89	100	50
E	**K**	409	entry	97	90	20	23	155	8
D	**K**	410	entry	55	63	20	25	155	8

Figure 3.18: ***Compressed, two class colored and tabulated results***
The tabulated and compressed output ('Data Analyzer') of the 'Advanced Search' section compares the callibrated percentage values for selected assays and sort's values into two classes. In this example a search for all mutations in the TSHR where the cAMP accumulation is characterized is presented. Cell surface expression and hormone stimulated cAMP accumulation levels below 50% (compared to the wild type) are colored red and basal cAMP levels greater than 150% are green (table not fully represented). Therefore, this example reveals mutants, which show deceased signalling activity with expression levels comparable to the wild type and also mutants with a decreased cell surface expression level and decreased signalling activity. The constitutive activation by the two CAMs (green) does not influence the hormone induced cAMP accumulation. This tool might enable quick visualization of functionalities assisted by user driven borderlines. If the classes are allocated to all searched assays in one go, the retrieved positions of mutants can be highlighted in color on 3D models.

With the use of colorable balls at C-α positions, results generated using our approach as well as the loci for combinations of point mutations with similar, different or even opposed phenotypes can also be visualized on one 3D-structure per receptor.

Colorable balls at C-α positions enable the visualization of results from our approach also on one 3D-structure per receptor as well as the loci for combinations of point mutations with similar and different or even opposed phenotypes. 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 spatially close wild type residues that might be feasible for interaction and whose modification or disruption by mutation might therefore cause this phenotype (Figure 3.19).

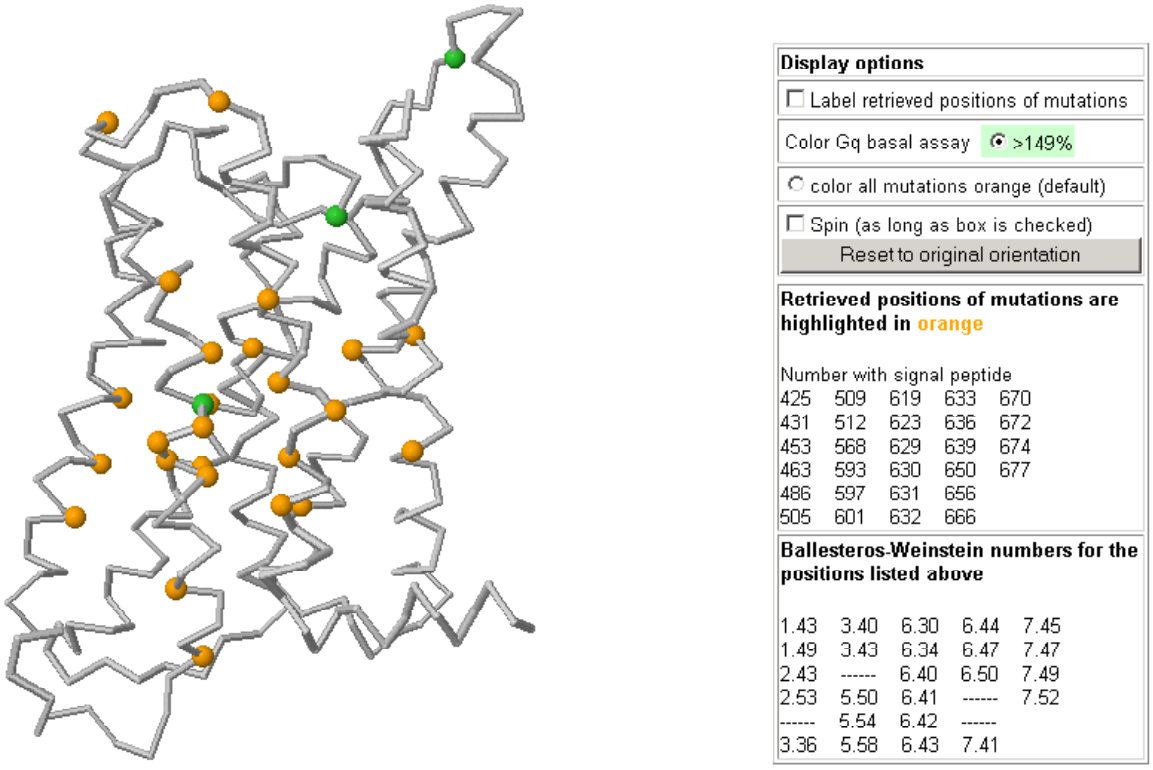


Figure 3.19: ***Constitutively activating mutations***
This picture represents the localization of CAMs (activity above 150%) in the serpentine domain of human TSHR. Given that the lowest basal Gαs signalling activity was shown to be between ~150-170% for active pathogenic mutants of the TSHR and the FSHR. This criteria of >150% of cAMP accumulation for extraction is used.The group of mutants (green) that are characterized by a simultaneous constitutive activation of Gαq (IP accumulation) and Gαs (cAMP accumulation) mediated signalling are especially interesting.

3.3.3 Pathogenic mutations

In a multiple alignment and tabulated overview, pathogenic mutations at the homologous GPHRs are summarized. The terms '*gain-* or *-loss of function*' mutations for *in vitro* results are strictly avoided, since these terms are related to the pathological *in vivo* findings. Moreover, TSHR and FSHR pathological mutations called '*gain-of-function*' mutants can be caused by different sources or can have different origins (e.g. constitutive activation or promiscuous binding of hormones). Therefore, for clear separation between *in vitro* and *in vivo* study-data the terms 'activating pathogenic' and 'inactivating pathogenic' as searchable terms for *in vivo* identified mutations are used.

In a particular multiple sequence alignment of GPHRs, the positions for *activating* and *inactivating pathogenic mutations* are marked in green and red and linked to PubMed/Medline citations or to a tabulated overview (if more than one pathogenic mutation is known). This tool simplifies the comparison of pathogenic mutations and helps to identify sequence positions and loci of common and different functionality for signalling among the GPHRs.

3.3.4 Constitutively activating mutations

In a separate alignment, mutations are marked that are characterized by an increased basal cAMP activity (more than 150% compared to wild type). These mutations are extracted from the data set automatically.

We used the criteria of >150% of cAMP accumulation for extraction given the circumstance, that for *activating pathogenic mutations* of the TSHR and the FSHR, the lowest experimentally shown basal Gαs signalling activity was ~150-170%.