UNIVERSITY OF LATVIA
Faculty of Biology

Dāvids Fridmanis

Doctoral Thesis

Melanocortin receptor family evolution with the focus on ACTHR functional specificity

Promotion to the degree of Doctor of Biology
Molecular Biology

Supervisor: Dr. biol. Jānis Kloviņš

Riga 2010

ABSTRACT

Melanocortin receptors (MCRs) are members of G - protein coupled receptor (GPCR) family. In vertebrates large proportion of GPCRs encoding genes are intronless or have a lower density of introns when compared with GPCRs of invertebrates. MCRs play an important role in a variety of essential functions such as the regulation of pigmentation, energy homeostasis, and steroid production. Among other MCRs adrenocorticotropic hormone (ACTH) receptor (MC2R) stands out with its highly specific membrane expression that is limited to adrenal cells and unlike the other members of MCR family, that recognize different melanocortin peptides, the ACTHR solely binds to ACTH. The aim of this research was to study the evolution of MCR family in context with evolutionary expansion of whole GPCR group and identify structural regions of ACTHR that determine its functional dissimilarities compared with, other MCRs. During the search of to ancestral similar MCRs in bony fishes: *Takifugu rubripes* and *Danio rerio* orthologs of all five MCR subtypes were identified, however there were no MCR identified in tunicates: *Coina intestinalis* and *Coina savignyi* nor in nematode: *Caenorhabditis elegans*. These findings combined with information of recently discovered MCRs in lamprey: *Lampetra fluviatilis* allow to conclude that MCRs appeared on an early stage of vertebrate evolution and this family evolved particularly rapidly soon after the split of lamprey and mammalian lineage. It is also concluded that specialisation of ACTHR took place during the same period of rapid evolution and to identify regions of this protein that determine its specificity in total 15 chimeric receptors were created by replacement of selected ACTHR domains with corresponding parts of MC4R. Chimeric receptors were tested for their membrane localization and ability to bind NDP-MSH and ACTH (1-24) and induce cAMP response. The results indicate that N-terminal part of ACTHR significantly decreases the membrane trafficking and third and fourth transmembrane region is main determinant of ACTHR intracellular trapping. In addition it was found that the fourth and fifth transmembrane domains in the ACTHR are involved in ACTH binding selectivity. During the research an analysis method to quantify cell membrane localization of recombinant receptors fused with enhanced green fluorescent protein from the results of confocal fluorescent microscopy was developed

KOPSAVILKUMS

Melanokortīna receptori (MCR) pieder pie ar G-proteīnu saistīto receptoru (GPCR) saimes. Mugurkaulniekos liela daļa šīs saimes receptorus kodējošo gēnu nesatur intronus vai arī tiem, salīdzinot ar bezmugurkaulniekiem, ir zemāks intronu blīvums. MCR ir svarīga loma tādu nozīmīgu funkciju kā, pigmentācija, enerģijas homeostāze asteroīdu produkcija regulācijā. Starp citiem MCR adrenokortikotropā hormona receptors (ACTHR) izceļas ar savu specifisko membrānas ekspresiju, kas tiek veikta tikai virsnieru šūnās, un atšķirībā no citiem MCR, kas atpazīst dažādus melanokortīna peptīdus, ACTHR saistās tikai ar adrenokortikotropo hormonu (ACTH). Šī darba mērķis bija pētīt MCR saimes evolūciju visas GPCR grupas evolucionārās ekspansijas kontekstā un identificēt strukturālos ACTHR reģionus, kas nosaka tā funkcionālās atšķirības. Darba gaitā ,meklējot priekšteču MCR līdzīgus receptorus kaulzivīs *Takifugu rubripes* un *Danio rerio*, tika atrasti ortologi visiem pieciem zīdītāju MCR, taču MCR netika atrasti nedz tunikātos *Ciona intestinalis* un *Ciona savignyi*, nedz nematodē *Caenorhabditis eleagns*. Apvienojot šos atklājumus ar nesen publicēto informāciju par MCR nēģī *Lampetra fluviatilis* bija iespējams secināt, ka MCR radās agri mugurkaulnieku attīstības sākumposmā un šī saime attīstījās īpaši strauji drīz pēc nēģu un zīdītāju priekšteču šķelšanās. Vēl tika secināts, ka ACTHR specializācija arī noritēja šajā pašā mugurkaulnieku evolūcijas posmā un, lai identificētu receptora specifiskumu nosakošos reģionus kopumā tika izveidoti 15 himeri receptori, kuros izvēlēti ACTHR reģioni tika aizvietoti ar homologiem MC4R reģioniem. Visiem himerajiem receptoriem tika veikti membrānas lokalizācijas, NDP-MSH un ACTH (1-24) saistīšanās un funkcionālās aktivitātes testi. Iegūtie rezultāti norādīja, ka ACTHR N-terminālā daļa ievērojami samazināja receptora transportu uz šūnas membrānu, un, ka trešā un ceturtā transmembrānas domēnu reģions nodrošina ACTHR iekššūnas aizturi. Papildus tika noteikts, ka ACTHR ceturtā un piektā transmembrānas domēnu reģions ir iesaistīs ACTH saistīšanas selektivitātes nodrošināšanā. Pētījumu gaitā arī tika izstrādāta analītiska metode ar uzlaboti zaļi fluorescējošo proteīnu sajūgtu rekombinantu receptoru lokalizācijas kvantifikācujai izmantojot konfokālo fluorescences mikroskopiju.

Doctoral Thesis were carried out at Latvian Biomedical Research and Study Centre during 2004-2010.

TABLE OF CONTENTS

Abstract 2
Kopsavilkums 2
Table of contents 3
Used abbrevations 4
Introduction 5
1 Summary of the existing scientific information in the research area 7
1.1 G-protein coupled receptors (GPCR) 7
1.1.1 Structure 7
1.1.2 Heterotrimeric G-protein 11
1.1.3 Signal transduction 14
1.1.4 Systematic and phylogeny of GPCRs 20
1.1.5 Intronless state of mammalian GPCR genes and intron evolution 24
1.2 Melanocortin system 27
1.2.1 Melanocortin (MC) peptides 27
1.2.2 Melanocortin receptors (MCRs) 29
1.2.3 Other elements of MC system 34
1.3 Chimerical proteins: prospects of application 37
1.4 Fluorescent proteins 38
2 Materials and methods 42
2.1 E. coli stain 42
2.2 Mammalian expression cell lines 42
2.3 DNA agrose gel electrophoresis 42
2.4 DNA fragment purification and extraction from agarose gel 42
2.5 Plasmid DNA extraction 42
2.6 Preparation of transformation competent *E. coli* cells 42
2.7 PCR directed mutagenesis 42
2.8 Restriction of DNA 43
2.9 Ligation of DNA fragments 43
2.10 Transformation of *E. coli* cells 43
2.11 Sequencing of DNA 43
2.12 Transfection and cultivation of mammalian expression cell lines 43
2.12.1 Creation of semi-stable HEK 293 EBNA cell lines for cAMP and binding assays 43
2.12.2 Transfection of BHK cells for confocal microscopy 44
2.13 cAMP response assay 44
2.13.1 Stimulation 44
2.13.2 Chromatography 44
2.14 Binding assay 45
2.14.1 Preparation of membranes 45
2.14.2 Binding and transfer to separation filters 45
2.15 Preparation of confocal microscopy samples 45
3 Results 47
3.1 The Melanocortin System in *Fugu*: Determination of POMC/AGRP/MCR Gene Repertoire and Synteny, As Well As Pharmacology and Anatomical Distribution of the MCRs 47
3.2 Formation of new genes explains lower intron density in mammalian Rhodopsin G protein-coupled receptors. 67
3.3 Identification of domains responsible for specific membrane transport and ligand specificity of the ACTH receptor (MC2R) 87
4 Discussion 101
Conclusions 111
Main thesis for defence 112
Acknowledgements 113
References 114

USED ABBREVATIONS

7TM .....7 transmembrane domain region
aa .....amino-acid
ACTH.....Adrenocorticotropic hormone
ACTHR .....Adrenocorticotropic hormone receptor
AF-WGA.....Alexa Fluor™ 633 labeled wheat germ agglutinin
AgRP .....Agouti related protein
ARG .....ancestral receptor group
ASIP .....Agouti or Agouti siganlling prrotein
C-terminus....carboxyl-terminus
CNS .....central nervous system
cAMP .....cyclic adenosine monophosphate
DNA .....deoxyribonucleic acid
EC50.....effective concentration (at which is produced 50% of of maximum response)
Ki.....inhibition constant
Kd.....dissociation constant
Bmax .....Amount of ligand required to saturate a population of receptor
EGFP .....enhanced green fluorescent protein
GFP.....green fluorescent protein
ER.....endoplasmatic reticulum
FCS.....foetal calf serum
Gα.....alpha subunit of heterotrimeric G-protein
Gβ.....beta subunit of heterotrimeric G-protein
Gγ .....gamma subunit of heterotrimeric G-protein
GDP.....guanidine diphosphate
GPCR .....G-protein coupled receptor
GTP .....guanidine triphosphate
HMM.....The Hidden Markov Model
kb.....Kilo-bases
MC.....melanocortin
MCR.....melanocortin receptor
MC#R.....melanocortin receptor type #
MC2R.....see ACTHR
MP .....Maximum Parsimony
MRAP.....melanocortin receptor accessory protein
MRAP-α.....melanocortin receptor accessory protein alpha isoform
MRAP-β.....melanocortin receptor accessory protein beta isoform
MRAP2.....melanocortin receptor accessory protein type 2
MSH .....Melanocyte-Stimulating Hormone
mRNA .....messenger ribonucleic acid
N-terminus....amino-terminus
NDP-MSH....[Nle4 D-Phe7] alpha-MSH
POMC.....proopiomelanocortin precursor
RNA .....ribonucleic acid
ROI.....regions of interest
RT-PCR.....reverse transcription polymerase chain reaction
SDS.....sodium dodecyl sulfate
TM.....transmembrane domain

INTRODUCTION

GPCRs are one of the largest protein families in the mammalian genome with over 800 genes in humans. GPCRs detect extracellular signals and transduce them into the cell through the activation of heterotrimeric G-proteins. This family includes receptors for many hormones, neurotransmitters, chemokines, calcium ions, and sensory molecules. All, currently known, GPCRs are sharing the same overall seven α-helical transmembrane (TM) domain core structure. The Rhodopsin family is the largest sub-family among GPCRs. There are no Rhodopsin GPCRs in plants and this group can only be found in animals. Rhodopsin GPCR family is one of the gene families in mammals with a large proportion of intronless genes. This is very uncommon, since the overall density of intronless genes in human is only 12.3%. Usually the intronless state is an indicator that a particular gene family is undergoing active changes and this peculiarity can be used to study the evolutionary history of introns. The number of GPCR genes has increased approximately six fold from nematodes to mammals.

The melanocortin system is involved in regulation of many different physiological functions, including the formation of pigmentation, adrenocortical steroidogenesis, energy homeostasis, exocrine secretion, sexual functions, inflammation and immunomodulation, thermoregulation and cardiovascular regulation. Term “melanocortin” is referred to peptides originated from proopiomelanocortin precursor (POMC), which exhibits melanotropic and adrenocorticotropic activity. MCs are produced mainly in frontal and intermediate lobe of pituitary. There are five types of melanocortin receptors (MCRs) that are all members of Rhodopsin GPCRs family and on the level of amino-acid sequence all MCRs share 39-61% identity. Despite this similarity all MCRs are displaying unique expression pattern and pharmacological profile, binding various melanocortins with distinctive affinities. ACTHR is the most specialized MCR, it is specifically binding only ACTHR and is predominantly expressed in adrenal cortex, where it mediates the effect of ACTH on glucocorticoid synthesis. Genetic defects in ACTHR gene are causing Familial glucocorticoid deficiency (FGD) - a rare autosomal recessive disorder.

The aim of this research was to study the evolution of MCR family and identify structural regions of ACTHR that determine its functional dissimilarities compared with, other MCRs. To reach this goal following tasks were established:

1. Identify MCRs in bony fishes *Takifugu rubripes* and *Danio rerio*;
2. Determine functional properties of *Takifugu rubripes* MCR which is most similar to mammalian ACTHR;
3. Search for "ancestral receptors groups" that would be present in vertebrates, tunicates and nematodes and identify the group which could be ancestral to MCRs;
4. Search for the explanations concerning existence of introns in *Takifugu rubripes* MCR genes and general intron absence in vertebrate GPCR genes;
5. Create chimeric receptors by replacement of selected ACTHR domains with corresponding parts of MC4R, determine their pharmacological properties upon binding and activation with NDP-MSH and ACTH and quantify their membrane trafficking;
6. Identify functional domains of ACTHR that determine its expression and ligand binding selectivity.

1 SUMMARY OF THE EXISTING SCIENTIFIC INFORMATION IN THE RESEARCH AREA

1.1 G-protein coupled receptors (GPCR)

GPCRs are one of the largest protein families in the mammalian genome with over 800 genes in humans (Fredriksson *et al.* 2003). GPCRs detect extracellular signals and transduce them into the cell using activation of GDP–GTP (guanidine diphosphate-guanidine triphosphate) exchange on heterotrimeric G-proteins. This family includes receptors for many hormones, neurotransmitters, chemokines, calcium ions, and sensory molecules. The Rhodopsin family is the largest sub family among GPCRs There are no Rhodopsin GPCR in plants (Fredriksson *et al.* 2005) and this group of GPCRs can be found only in animals. The number of GPCR gene has increased approximately six fold from nematodes to mammals (Fredriksson *et al.* 2005).

1.1.1 Structure

Despite the wide chemical and functional variety of the bound signal molecules, all, currently known, GPCRs are sharing the same overall seven α-helical transmembrane (TM) domain core structure (Fig.1). (Pierce *et al.* 2002) Such similarity in structure indicates that all GPCRs are most certainly evolutionary related and moreover it has been discovered that bacteriorhodopsin also has a 7TM, but instead of coupling to a G-protein, it functions as a light driven proton pump. However such structural similarity suggests that it might be related to ancestral GPCR.

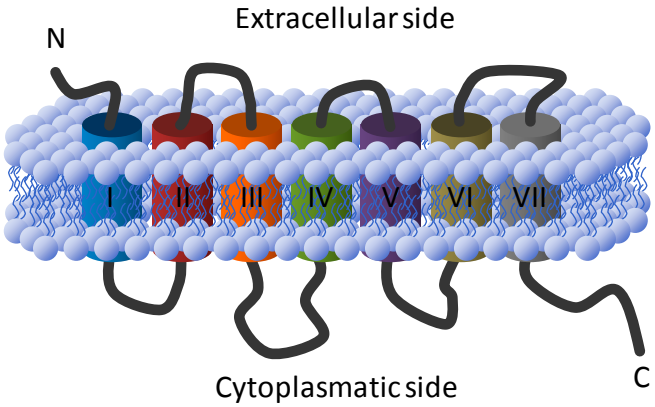


Figure 1. Schematical depiction of G-protein coupled receptor structure: seven cell membrane spanning α-helical transmembrane domains, N-termini is always located in extracellular side and C-terminus is always located in cytoplasmatic side.

First insight into the structural composition of GPCRs came in early 1980's. The breakthrough, discovery of complete amino-acid sequence of bovine rhodopsin, was achieved by two independent groups simultaneously (Ovchinnikov *et al.* 1982; Hargrave *et al.* 1983) and it was followed by molecular cloning of all four human opsins (Nathans *et al.* 1984; Nathans *et al.* 1986) and the hamster β-adrenergic receptor (Dixon *et al.* 1986). At this point it was noticed that secondary structure predictions of all these proteins shared similar structural pattern that was formed by seven hydrophobic membrane spanning α-helices, extracellular N-termini and intracellular C-termini. Discoveries of other around 70 similar receptors, which took place during the next decade, confirmed that such secondary structure is common to all members of GPCR super-family. (Probst *et al.* 1992). Nevertheless, despite identification of such, for that time, large number of new GPCRs, it took ten years of studies before the first insight into the three-dimensional tertiary structure was gained. It was accomplished at first by acquiring of 9Å projection map of bovine rhodopsin, which through electron crystallography of two-dimensional crystals (Schertler *et al.* 1993) and then by combining these findings with the sequence analysis, in order to allocate the individual TMs (Baldwin 1993). The rhodopsine was yet again used as model GPCR not for the historical reasons but because of its rigid structure caused by covalently bound ligand. In these studies it was acknowledged that TM domains of GPCRs are indeed organised in bundle forming a cavity in the centre, which could possibly act as ligand binding pocket for most of the GPCRs. For the next almost ten years this model along with the few later low resolution models served as the only sources of GPCR structure information and were extensively used in comparison modelling of other –similar receptors. (Costanzi *et al.* 2009) The next big breakthrough was achieved in 2000 when Palczewski and his colleagues acquired the first high resolution (2,8Å) X-ray crystal structure of bovine rhodopsin. Analysis of this model revealed exact geometry of each TM, geometries of large portion of intracellular and extracellular domains (loops, N- and C-termini), number of intramolecular interactions including disulfide bridge that connects extracellular side of TM3 with the second extracellular loop, as well as structures, in following years identified as rhodopsin specific (β-hairpin conformation for the second extracellular loop). (Palczewski *et al.* 2000) (Fig. 2)

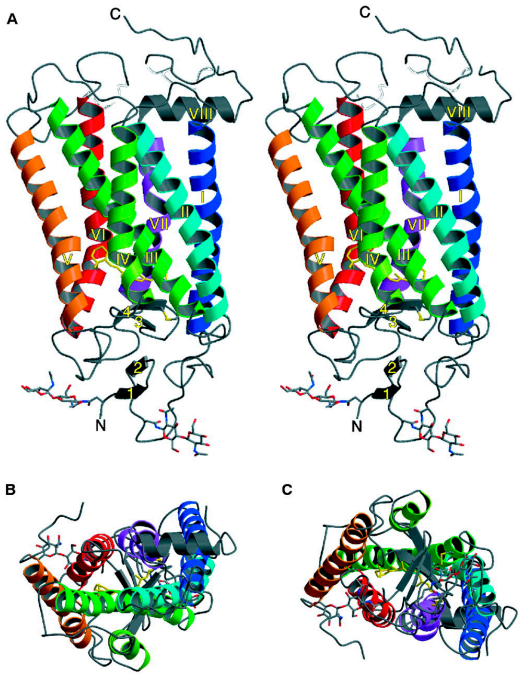


Figure 2. Ribbon drawings of rhodopsin. (**A**) Parallel to the plane of the membrane (stereo view). A view into the membrane plane is seen from the cytoplasmic (**B**) and intradiscal side (**C**) of the membrane. (Figure and figure legend are from: Palczewski *et al.* 2000)

After publication of the x-ray structure of rhodopsin, just like in a case with the low resolution model, was extensively used as template for the homology modelling of other GPCRs. Such models were mainly supported with the indirect experimental data acquired by usage of mutation techniques. However, two problems were undermining the credibility these results when used for modelling of GPCR-ligand complexes. First, these models could not be validated due to the lack of second-reference GPCR, and, second, since the structure acquired by Palczewski and his colleagues was the one of inactive rhodopsin, there were no information regarding the changes receptor is undergoing during activation process. (Costanzi *et al.* 2009)

In the following years the changes in structure of receptor during activation process have been studied very intensively using electron crystallography (Krebs *et al.* 2003; Ruprecht *et al.* 2004) and solid-state nuclear magnetic resonance (NMR) spectroscopy (Crocker *et al.* 2006). Results gained from these experiments suggested that the signal transduction promoting large scare conformational changes during activation of the receptor are occurring suddenly rather than gradually and that changes in conformation during the first phases of the activation process are concerning only several amino –acid residues located within the receptor binding pocket.

The most recent and ground breaking advances in GPCR structure discovery were achieved in the last few years. There are 10 new high-resolution structures solved: one of

thermo-stabilised bovine rhodopsin (Standfuss *et al.* 2007), two of bovine rhodopsin in an active-like state (Park *et al.* 2008; Scheerer *et al.* 2008), three of the human β_2 adrenergic receptor (Cherezov *et al.* 2007; Rasmussen *et al.* 2007; Hanson *et al.* 2008), one of the human adenosine A_{2a} receptor (Jaakola *et al.* 2008), one of the turkey β_1 adrenergic receptor (Warne *et al.* 2008) and two of the squid rhodopsin which unlike previous ones is a G_q coupled receptor (Murakami *et al.* 2008; Shimamura *et al.* 2008).

Such high rate of success, if compared to previous period of researches, can be mainly attributed by the usage of novel GPCR stabilisation and acquisition approaches. The stabilisation techniques (including thermal-stabilisation) comprised addition of receptor conformation stabilising ligands, lipids or antibodies at the stage of crystallisation, as well as modification of desirable GPCRs - proteolysis, mutatagenesis, deletion of flexible regions or insertion of highly structured ones was usually employed. Receptors modified in this way were over-expressed and purified from recombinant mammalian or insect cell expression systems. (Tate *et al.* 2009)

It was self-evident that such large amount of new discoveries led to comparison of different GPCR structures. During these studies it was acknowledged that overall helical framework and positions of helical irregularities of the analysed receptors indeed are well conserved, supporting the previously created models that were based on the first rhodopsin x-ray structure. However the extracellular surfaces of rhodopsin and hormone receptors were significantly different. Binding pocket of rhodopsin is closed/covered with the structure formed by N-terminus of the receptor, β-hairpin structure in the second extracellular loop and various elements of other extracellular loops, while binding pockets of hormone receptors are open to permit access of ligand. (Tate *et al.* 2009)

Regarding the activation of the GPCRs, it was established that in the inactive state two of Rhodopsin GPCRs (Class A GPCRs) characteristic motifs (E/D)RY, located in intracellular part of TM3, and NPxxY, located in intracellular part of TM7, are spatially separated by a "wall" of hydrophobic amino-acid residues located in intracellular part of TM2 and TM6. After the binding of ligand a series of intrahelical rearrangements are taking place. During these rearrangements "the wall" between (E/D)RY and NPxxY motifs is removed thus permitting the formation of structure crucial for G-protein binding. (Fig. 3) (Tate *et al.* 2009)

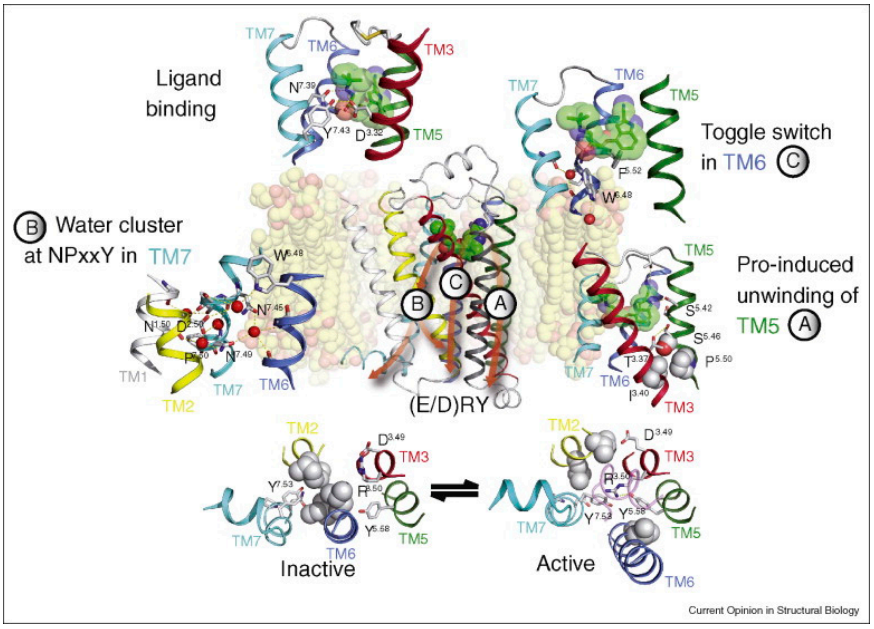


Figure 3. Four interhelical interaction networks important for receptor activation of Class A GPCRs, depicted in the structure of β1AR. Top panel: the rotamer toggle switch in TM6, formed by W6.48 and F6.52. Its rearrangement upon agonist binding has been related to the movement of TM6 (Rao *et al.* 1996). Left panel: a complex hydrogen bond network between TM1, TM2, TM6 and TM7 involving some of the most highly conserved residues in Class A GPCRs (N1.50, D2.50, W6.48, (N/S)7.45, N7.49, P7.50) and water molecules. The position of the water molecules has been derived from the structure of the β2AR. Right panel: a TM3/TM5 interaction between S5.46 and T3.37 can be directly disrupted by agonist binding, which could lead to the rearrangement of the Pro-induced unwinding of TM5 (stabilised by I3.40), and to a conformational change in the cytoplasmic side of TM5. A similar network is present in rhodopsin (Shi *et al.* 2002). These networks define three possible independent pathways of activation (A, B and C) that converge on the (E/D)RY motif. Bottom panel: detail of the (E/D)RY region in the inactive (β1AR structure, left) and active ($G\alpha_t$ peptide-bound opsin structure, right) conformations. In the inactive state (left), R3.50 is constrained in an energetically unfavourable conformation because of the ionic interaction with E/D3.49. The (E/D)RY motif in TM3 is isolated from the NPxxY motif in TM7 by a wall of bulky hydrophobic residues of TM2 and TM6 (spheres). Upon activation (right), TM6 moves away from the bundle and the NPxxY motif in TM7 rearranges, bringing Y7.53 close to R3.50. This rearrangement is probably related to the rearrangement of the water cluster (left panel), as water molecules seem responsible for stabilising the π-helix segment observed in the inactive state (Pardo *et al.* 2007). (Figure and figure legend are from: Tate *et al.* 2009)

1.1.2 Heterotrimeric G-protein

At the present moment there have been identified 21 α, 6 β and 12 γ (Tab. 1) mammalian heterotrimeric G-protein subunits. Most of those are encoded by separate genes while some are formed through alternative splicing. (Downes *et al.* 1999; Milligan *et al.* 2006) Since the Gα subunit is the one that most frequently determines the type of affected intracellular signalling pathway, it has been studied more intensively than other subunits. Basing on primary sequence similarity, all Gα were divided in four main classes: $G_s\alpha$, $G_{i/o}\alpha$, $G_{q/11}\alpha$ and $G_{12/13}\alpha$ (Tab. 1). (Simon *et al.* 1991) However the structural studies revealed that despite the class membership all Gα has the same conserved protein fold,

formed by GTPase domain and helical domain. Structure of the GTPase domain, main function of which is hydrolysis of GTP, is highly similar to structure found in other G-proteins including ones of the Ras superfamily. The helical domain on the other hand is unique to Gα. It is formed by bundle of six α-helices and covers nucleotide binding pocket as a lid, thus burying bound nucleotides in interior of the protein. The vast majority of Gα are post-translationally modified with the fatty acid at their N-termini, some are additionally myristoylated at their N-termini, and these modifications determine membrane localisation of particular Gα as well as some protein-protein interactions. (Oldham *et al.* 2008)

Gβ contains seven WD40 sequence repeats that form seven-bladed β-propeller structure, which is known to serve as a platform for the protein complex assembly or mediator of transient interplay among other proteins. N-terminus of Gβ is folded in α-helix that forms the coiled-coil with the N-terminus of Gγ. The C-terminus of Gγ is also bound to Gβ (to blades five and six) thus forming a complex that can only be separated in highly denaturing conditions. All Gγ are post-translationally modified by isoprenylation of their N-termini. (Oldham *et al.* 2008) It should also be mentioned that although most of the Gβ can form the complex with the most of Gγ, not all of ~60 possible combinations have been detected this far. (Clapham *et al.* 1997)

Table 1. The family of mammalian heterotrimeric G-protein subunits: function and regulation (Table and table legend are from Milligan *et al.* 2006)

Family	Subtype	Effectors	Expression	Disease relevance	Pharmacological modulation
$G_s\alpha$	$G_{s(S)}\alpha$ $G_{s(L)}\alpha$ $G_{s(XL)}\alpha$ $G_{olf}\alpha$	Adenylyl cyclases ↑ ($G_{s,s(XL),olf}\alpha$) Maxi K channel ↑ ($G_s\alpha$) Src tyrosine kinases (c-Src, Hck) ↑ ($G_s\alpha$) GTPase of tubulin ↑ ($G_s\alpha$)	$G_s\alpha$: ubiquitous $G_{olf}\alpha$: olfactory neurons, certain CNS ganglia; digestive and urogenital tract	$G_{s(XL)}\alpha$: brachydactyly, trauma-related bleeding tendency, neurological problems $G_s\alpha$: McCune–Albright syndrome, cholera, pseudohypoparathyroidism type Ia/b, testotoxicosis, adenomas of pituitary and thyroid	$G_s\alpha$: CTX $G_{olf}\alpha$: CTX
$G_{i/o}\alpha$	$G_{o1}\alpha$ $G_{o2}\alpha$ $G_{i1-i3}\alpha$ $G_z\alpha$ $G_{t1/2}\alpha$ $G_{gust}\alpha$	Adenylyl cyclase ↓ ($G_{i,o,z}\alpha$) Rap1GAPII-dependent ERK/MAPkinase activation ↑ ($G_i\alpha$) Ca^{2+} channels ↓ ($G_{i,o,z}\alpha$) K^+ channels ↑ ($G_{i,o,z}\alpha$) GTPase of tubulin ↑ ($G_i\alpha$) Src tyrosine kinases (c-Src, Hck) ↑ ($G_i\alpha$) Rap1GAP ↑ (G_za) GRIN1-mediated activation of Cdc42 ↑ ($G_{i,o,z}\alpha$) cGMP-PDE ↑ ($G_t\alpha$) $G_{gust}\alpha$: ?	$G_{o1-2}\alpha$: neurons, neuroendocrine cells, astroglia, heart $G_{i1-i3}\alpha$: neurons and many others $G_z\alpha$: platelets, neurons, adrenal chromaffin cells, neurosecretory cells $G_{t1}\alpha$: rod outer segments, taste buds $G_{t2}\alpha$: cone outer segments $G_{gust}\alpha$: sweet and/or bitter taste buds, chemoreceptor cells in the airways	$G_i\alpha$: whooping cough, adrenal and ovarian adenomas $G_t\alpha$: congenital cone dysfunction, night blindness	$G_{o(1/2)}\alpha$: PTX $G_{i1-i3}\alpha$: PTX $G_z\alpha$: ? $G_{t1/2}\alpha$: PTX, CTX $G_{gust}\alpha$: PTX
$G_{q/11}\alpha$	$G_q\alpha$ $G_{11}\alpha$ $G_{14}\alpha$ $G_{15}\alpha$ $G_{16}\alpha$	Phospholipase Cb isoforms ↑ p63-RhoGEF ↑ ($G_{q/11}\alpha$) Bruton's tyrosine kinase ↑ ($G_q\alpha$) K^+ channels ↑ ($G_q\alpha$)	$G_{q/11}\alpha$: ubiquitous $G_{15/16}\alpha$: hematopoietic cells	$G_{q/11}\alpha$: dermal hyperpigmentation and melanocytosis?	$G_{q/11}\alpha$: YM-254890 $G_{14}\alpha$: ? $G_{15}\alpha$: ? $G_{16}\alpha$: ?
$G_{12/13}\alpha$	$G_{12}\alpha$ $G_{13}\alpha$	Phospholipase D ↑ Phospholipase Ce ↑ NHE-1 ↑ iNOS ↑ E-cadherin-mediated cell adhesion: ↑ p115RhoGEF ↑ PDZ-RhoGEF ↑ Leukaemia-associated RhoGEF (LARG) ↑ Radixin ↑ Protein phosphatase 5 (PP5) ↑ AKAP110-mediated activation of PKA ↑ HSP90 ↑	Ubiquitous	Recent SNPs identified but no disease correlation yet	$G_{12}\alpha$: ? $G_{13}\alpha$: ?
$G\beta/\gamma$	β_{1-5} γ_{1-12}	PLCbs ↑ Adenylyl cyclase I ↓ Adenylyl cyclases II, IV, VII ↑ PI-3 kinases ↑ K^+ channels (GIRK1,2,4) ↑ Ca^{2+} (N-, P/Q-, R-type) channels ↓ P-Rex1 (guanine nucleotide exchange factor for the small GTPase Rac) ↑ c-Jun N-terminal kinase (JNK) ↑ Src kinases ↑ Tubulin GTPase activity ↑ G-protein-coupled receptor kinase recruitment to membrane ↑ Protein kinase D ↑ Bruton's tyrosine kinase ↑ p114-RhoGEF ↑	$\beta_1\gamma_1$: retinal rod cells $\beta_3\gamma_8$: retinal cone cells β_5: neurons and neuroendocrine organs β_5(L): retina Most cell types express multiple β and g subtypes	$G\beta_3$: atherosclerosis, hypertension, metabolic syndrome	$G\beta\gamma$: ?

1.1.3 Signal transduction

The mammalian GPCR-ligand system is considered very complex. This complexity at the greatest part is determined by a large number of GPCRs majority of which are capable to bind more than one ligand and even larger number of ligands that are usually capable to bind to several GPCRs. This picture is becoming even more complex if the differential binding affinities are taken into account. However signal transduction system from receptor to heterotrimeric G-protein if far from less complex if compared to GPCR-ligand system, although the number of GPCR activated Gα is significantly lower. There are several aspects of G-protein activation process that are increasing the complexity of signal transduction.

There are two models proposed to explain the process of activated GPCR and G-protein encounter process. The first, "collision coupling" model states that GPCR – G-protein coupling occurs after activation of receptor and prior activation, receptor and G-protein are freely diffunding in plasma membrane. (Tolkovsky *et al.* 1978) Second, recently developed, "precoupling" model on the other hand points that G-proteins can interact with the GPCR prior its activation and during ligand binding process G-protein is only being activated. (Gales *et al.* 2006) The latest of these two models could explain the rapid intracellular response that occurs after receptor activation, however it cannot explain the high level signal amplification, because according to scenario one receptor molecule activates only one precoupled G-protein molecule. Interestingly there are several studies that were carried out in the same GPCR – G-protein system, but results of which supported opposing models. (Hein *et al.* 2005) It is possible that both models are true, meaning, that some types of the receptors are predominantly precoupled while others are freely diffunding in plasma membrane and even more that some types of receptors are functioning under some kind of a mixed coupling model.

As mentioned in previous section after the binding of the ligand receptor is changing its conformation creating the G-protein binding pocket, formed by intracellular parts of TM3 and TM7. The last 11 amino-acid residues of C-terminus (Kisselev *et al.* 1998; Koenig *et al.* 2002) and amino-acid residues located in fourth α-helix and sixth β-sheet (Bae *et al.* 1997; Onrust *et al.* 1997; Bae *et al.* 1999) are regions of Gα that are interacting with this binding pocket during G-protein activation process. Even more, there are reports that receptor and Gβγ complex interaction, at least in some cases, is also crucial for receptor and Gα interaction stabilisation. (Taylor *et al.* 1994; Taylor *et al.* 1996) The

complexity of these interactions is determined by the distinctive ways of how the various amino acid residues in various positions of described regions are contributing to receptor and G-protein coupling. It has been shown that not only complete exchange of C-terminus of Gα (Conklin *et al.* 1993) but also a simple two amino acid residue dissimilarity at near previously described interaction regions (Bae *et al.* 1999) can alter the specificity of receptor and G-protein coupling. This observation is also true regarding the GPCRs. In this case, since the 3D structure of the GPCRs is not as rigid as the structure of G-proteins, even changes that occurred spatially far from surfaces of interaction can have influence on coupling. Even more, beyond specificity of Gα, there are studies which reveal the coupling preference of some Gβγ complexes to specific receptors. (Hou *et al.* 2000; McIntire *et al.* 2001) In fact all studies performed in this field, if taken together, reveal the existence of complex and tingled network of specificity defining interactions.

Since there are more than 800 different GPCRs and only 21 Gα in mammalian organisms, it is clear that each Gα must be able to couple to more than one receptor. However some receptors have been identified that are able to couple with more than one type of Gα. Even more, this dual coupling is ligand dependant, meaning that choice of Gα depends on specific ligand, receptor is bound to. (McLaughlin *et al.* 2005; Mukhopadhyay *et al.* 2005; Perez *et al.* 2005) Existence of such dual action mechanism raises the complexity level of GPCR – G-protein signal transduction even higher.

Coupling to activated receptor triggers the conformational changes in G-protein. As the result Gα subunit of G-protein opens the nucleotide binding pocket and exchanges its GDP for GTP. The exchange of nucleotide is rather passive process and mainly relies on excess of GTP concentration over GDP concentration in the cytoplasm of the cell. The exchange of nucleotide triggers further structural rearrangements as the result of which Gα adopts its active conformation and dissociates from the Gβγ complex and from the receptor. During the following phase of signal transduction active Gα and free Gβγ complex are acting as two separate proteins. While in its active conformation the Gα is able couple with its target protein, usually enzyme, like, adenyl-cyclase or protein kinase, but coupling to ion channels or other types of cell signal system proteins also occurs. This coupling activates the target protein, which in its turn alters the concentrations of intracellular secondary messenger molecules. The type of activated target protein depends on the type of Gα (Tab. 1) Since the Gα possesses GTPase activity, during the target protein activation the bound GTP is being slowly hydrolysed to GDP thus inactivating Gα.

After the inactivation Gα couples to Gβγ complex forming initial heterotrimeric G-protein. After disassociation from the Gα some types of Gβγ complex are also taking part in signal transduction. These are also coupling and activating their own target protein, that usually triggers concentration alterations of intracellular secondary messenger that is distinctive from one altered by its Gα. Since not all of the Gβγ complexes are capable to participate directly in signal transduction, the range of target proteins Gβγ can couple to is significantly lower (Tab. 1 Fig. 4) (Alberts *et al.* 2002).

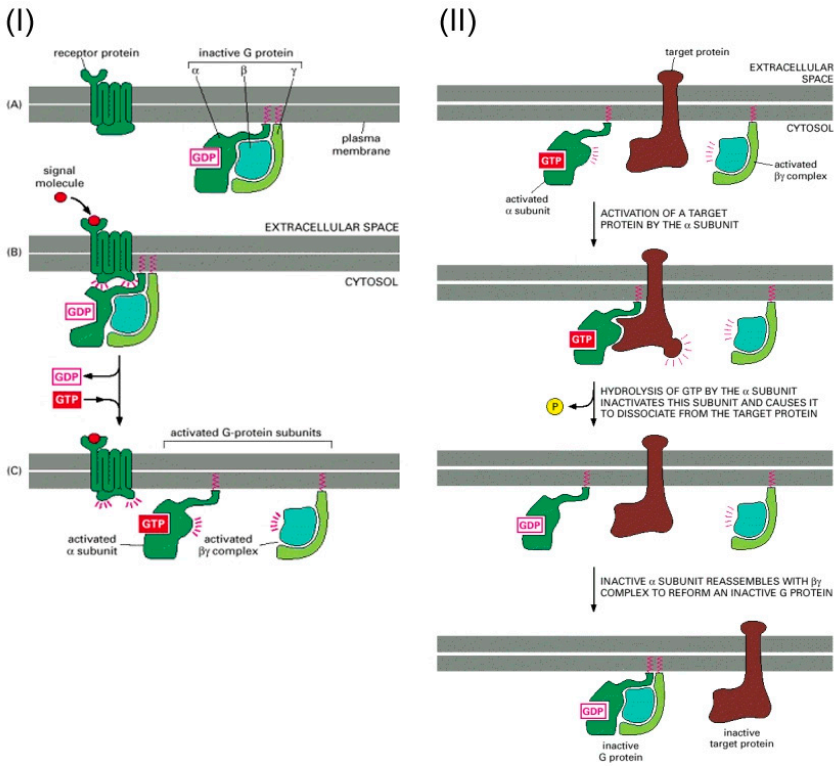


Figure 4. **Panel (I)** The disassembly of an activated G-protein into two signalling components. (A) In the unstimulated state, the receptor and the G protein are both inactive. Although they are shown here as separate entities in the plasma membrane, in some cases, at least, they are associated in a preformed complex. (B) Binding of an extracellular signal to the receptor changes the conformation of the receptor, which in turn alters the conformation of the G protein that is bound to the receptor. (C) The alteration of the α subunit of the G protein allows it to exchange its GDP for GTP. This causes the G protein to break up into two active components—an α subunit and a βγ complex, both of which can regulate the activity of target proteins in the plasma membrane. The receptor stays active while the external signal molecule is bound to it, and it can therefore catalyze the activation of many molecules of G protein. **Panel (II)** The switching off of the G-protein α subunit by the hydrolysis of its bound GTP. After a G-protein α subunit activates its target protein, it shuts itself off by hydrolyzing its bound GTP to GDP. This inactivates the α subunit, which dissociates from the target protein and reassociates with a βγ complex to re-form an inactive G protein. Binding to the target protein or to a membrane-bound RGS protein (not shown) usually stimulates the GTPase activity of the α subunit; this stimulation greatly speeds up the inactivation process shown here. (Figure and figure legend are from Alberts *et al.* 2002)

Described above is the classical understanding of G-protein signalling, which, to date, is still true but highly simplified. Exclusions and control mechanisms for almost every step of this process have been identified.

From functional point of view there are several types of ligands. Agonists are ligands that activate receptor; antagonists - ligands that are binding to receptor at the same binding pocket as agonists not activating the receptor and prohibiting agonists access to the binding pocket; inverse agonists that are inactivating constitutively active receptors (those usually have pharmacaphore fused to their N-terminus and do not have any agonists at all) positive allosteric modulators - ligands that are binding to receptor at the another-secondary binding pocket (different from the one occupied by agonists and antagonists), enhancing the agonist binding or receptor activation efficiency; and negative allosteric modulators, that are binding to receptor at the another-secondary binding pocket, diminishing the agonists binding or receptor activation efficiency. As one can understand from these descriptions already at ligand binding stage there exists an elaborate control system. It should be mentioned that most of the receptors have no natural antagonist or allosteric modulator, however these ligands types are very extensively used in pharmacy. (Zheng *et al.* ; Bond *et al.* 2006; Wang *et al.* 2009)

In previously described - classical understanding receptor is viewed as monomer, however the reality is quite different: large number of GPCRs is known to form homo and hetero oligomers, formation of which itself is a well known phenomenon. In the world of proteins, oligomerization has been recognized as advantageous from several aspects. Most obvious of these is the saving of space in the carrier of genetic material, for large proteins that are formed by number of identical subunits it is more convenient to store information of single domain in one - small gene, than store it in large gene encoding reputing multi-domain protein or multiple genes encoding the same subunit. Homodimerization of two identical GPCRs allows them to act in *trans* on one another, meaning that, due to involvement of larger number of molecular interactions, which are "knitting" various parts of proteins together, from conformational point of view dimers are more rigid than monomers and because of this rigidness they, upon ligand binding, are changing their conformation more reluctantly, however, conformation changes of one partner are subsequently forcing the second to follow more rapidly. If compared graphically the activation curve of dimer would be steeper than activation curve of monomer, thus dimers would exhibit better "molecular on-off switch" properties. Another advantage that homo- or heterooligomerization of GPCRs could provide is additional regulation leverages. The

simplest is based on the fact that interaction of two GPCRs can stabilize both receptors in conformation that is more/less suitable for ligand binding, hence increasing/decreasing the receptor affinity for the ligand. In this case it is possible to regulate the sensitivity of receptors by the means of receptor concentration changes. However, usually regulations that involve oligomerization are more complex. Formation of GPCR-GPCR hetrodimer has been shown to result in changes of ligand binding or G-protein coupling profile. For example Dopamine D1 and D2 receptors expressed individually are coupling to $G_s\alpha$ and $G_i\alpha$ respectively, while D1/D2 heterodimers are coupling to $G_q\alpha$. When referring to heteromerisation it should also be kept in mind that the term includes not only oligomerization of different GPCRs but also formation of GPCR and other non-GPCR protein complexes. It has been established that number of GPCRs, require the presence of specific accessory proteins for the functional activity. (Klemm *et al.* 1998; Rios *et al.* 2001; Millar *et al.* 2010)

Desensitization and internalisation (or sequestration) of GPCRs are processes that also affect GPCR signalling. Main purpose of these processes is to inactivate GPCR, making the cell unresponsive, after reception of the initial signal thus prohibiting the unnecessarily continuous activation of intracellular signalling cascades and saving the resources of the cell. These processes are also crucial on the scale of whole organism, in remotely localized cell response synchronisation it allows the cells that receive the signal later to catch up with the ones activated earlier. Effects of desensitization and internalisation are usually reversible, shortly after inactivation receptors are reactivated and cell is again ready “for action”. Desensitization is relatively fast process, it often takes place immediately after activation of G-protein and runs in the time scale from few milliseconds up to few seconds. GPCRs are desensitizated by phosphorylation of their intracellular loops and/or C-terminus and subsequent coupling to arrestin, a soluble cytoplasm protein, which prohibits receptor coupling to G-protein. Phosphorylation is carried out by G-protein coupled receptor kinases (GRKs) which are activated by various elements of intracellular signalling cascades initiated by the receptor itself. Internalisation on the contrary is rather slow process and runs in the time scale from few seconds up to few minutes. It starts with the recruitment of clathrin by GPCR bound arrestin, further newly formed complex is moved to region of cell membrane called clathrin coated pit where it is endocyted and transported to intracellular compartments of the cell. (Fig. 5)

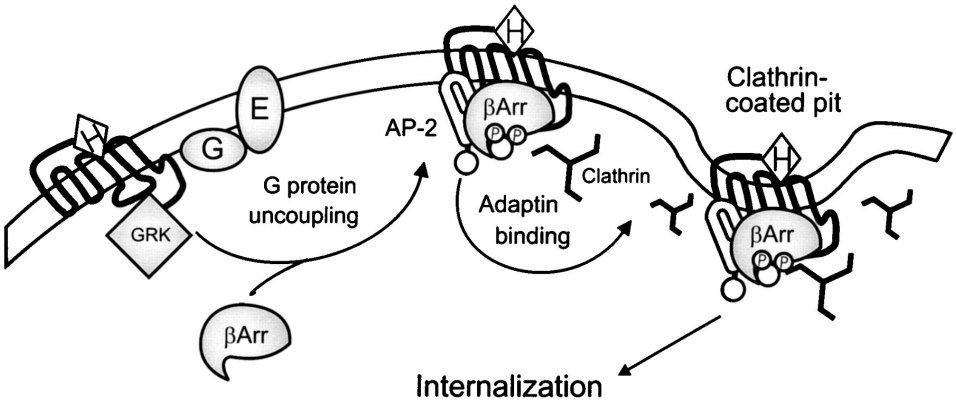


Figure 5. Molecular mechanisms involved in the GRK- and β-arrestin-dependent desensitization and internalization of GPCRs. GPCR activation leads to GRK-dependent phosphorylation of intracellular serine and threonine residues that facilitate the translocation and binding of β-arrestin proteins to the receptor. β-arrestins, via their association with the β2-adaptin subunit of the AP-2 heterotetrameric adaptor complex, target GPCRs to clathrin-coated pits. In addition to their association with β2-adaptins, β-arrestins also bind clathrin. The GPCR is subsequently internalized via clathrin-coated vesicles. AP-2, AP-2 heterotetrameric adaptor complex; βArr, β-arrestin; H, hormone; P, phosphate group. (Figure and figure legend are from Ferguson 2001)

After internalisation, there are two possible scenarios, receptor is ether dephosphorylated and recycled to membrane or ubiquitinated and assigned for degradation. The choice of scenario usually depends on the type of receptor. There are receptors that are actually signalling only once it their lifetime and there are receptors that are being constantly recycled. In fact there are receptors that are being constantly internalised even without any receptor activation. Such situation has been observed in cell types which require rapid return of the sensitivity to certain stimuli and constant receptor internalization is believed to provide supply of active "sensitive" receptors. (Fig. 6)

Although internalisation usually terminates the signalling of GPCRs, there are also some exclusions. In previously described scenario desensitization and internalisation took place after disassociation of receptor-Gα complex, thus only receptor was internalised, however there are situation when ligand-receptor-Gα is so highly stable that it does not disassociate and constantly activates target enzyme. It has been observed that in such cases whole membrane bound signalling machinery is internalised forming constantly signalling cell structures called signalosomes or signalling endosomes, which have been detected to remain active for as long as 30 minutes. In conclusion it was recently discovered that there are distinct GPCR signalling pathways, involving extracellular signalling-regulated kinases, that are activated through receptor coupling to arrestin, (Ferguson 2001; Wolfe *et al.* 2007; Xu *et al.* 2007; Kovacs *et al.* 2009; Rosenbaum *et al.* 2009; Jalink *et al.* 2010)

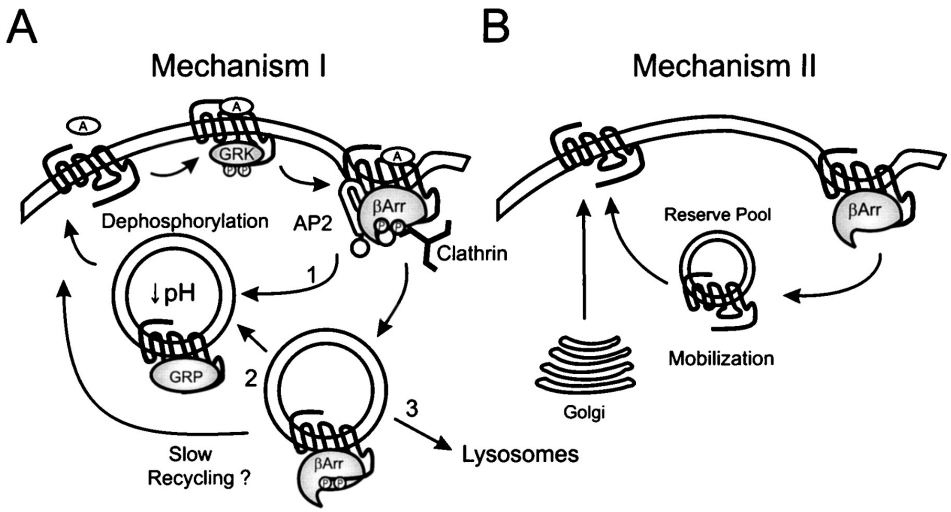


Figure 6. β-Arrestin-dependent regulation of GPCR trafficking and the mechanisms involved in GPCR resensitization. Panel A, mechanism I: GPCR resensitization is achieved by the clathrin-dependent internalization of GPCRs to endosomes, and the receptors are subsequently dephosphorylated and recycled back to the cell surface (1). Alternatively, GPCRs that internalize with β-arrestin either dissociate from β-arrestin in the intracellular compartment, and are both dephosphorylated and recycled (2) or are retained within large core vesicles and/or are targeted for degradation in lysosomes (3). Panel B, mechanism II: resensitization of GPCRs that are internalized with β-arrestin bound and do not recycle is mediated by the mobilization of a reserve pool of intracellular receptors and/or de novo receptor synthesis in the Golgi apparatus. A, agonist; AP-2, AP-2 heterotetrameric adaptor complex; βArr, β-arrestin; P, phosphate group. (Figure and figure legend are from Ferguson 2001)

1.1.4 Systematic and phylogeny of GPCRs

First attempts to classify receptors of GPCR superfamily were undertaken soon after discovery of Rhodopsin, however it was only in 1994 when the first universal classification system was created. (Attwood *et al.* 1994) At that moment there were ~400 identified GPCRs and it was established that although similar in their secondary structure alignment of amino-acid sequences of all known receptors presents no apparent over all similarity. For this reason Attwood and his colleague focussed upon identification of "fingerprints" that would be characteristic for specific groups of GPCRs. Such approach turned out to be very successful and as the result all GPCRs were divided six classes or clans: A, B, C, D, E, and F and a number of subclasses or subclans that were labelled with roman numbers. The main advantage of this system, besides effective division, was that it was designed to cover not only receptors found in mammalians but also GPCRs found in other forms of life, for example classes D, E and F that unite Fungal mating pheromone receptors, cAMP receptors and archaebacterial opsins respectively as well as subclass IV in class A that unites invertebrate opsin receptors. (Attwood *et al.* 1994) Although effective and in following years widely accepted "fingerprints" classification system had its opponents. Main arguments of opposition considered low amino-acid sequence

homology among different GPCRs and advisability of amino-acid sequence usage as main criteria of division. These researchers claimed that, due to low sequence homology, secondary structure similarity of GPCRs must be the result of molecular convergence and usage of bound ligands and location of binding pocket as main criteria of classification would much more informative (Bockaert *et al.* 1999; Qian *et al.* 2003).

Nevertheless the search of the new classification system that would be based upon sequence comparison and phylogenetical analysis were continued and in 2003 Fredriksson and his colleagues (Fredriksson *et al.* 2003) succeeded in creating GRAFS classification system for all GPCRs found in human. In this system all human GPCR Pool is divided in five main families: Glutamate (formed by 15 receptors), Rhodopsin (formed 241 nonolfactory receptors, and 460 olfactory receptors), Adhesion (formed by 24 receptors), Frizzled (formed by 24 receptors) and Secretin (formed by 15 receptors) receptor families, and the name designated to the classification system is formed by the first letters of receptor family names. Interestingly this study also confirmed the results gained from "fingerprints" classification because rhodopsin, secretin and glutamate receptor families are corresponding to classes A, B and C respectively, however adhesion and frizzled are not corresponding to any of the "fingerprints" classes. (Fig.7)

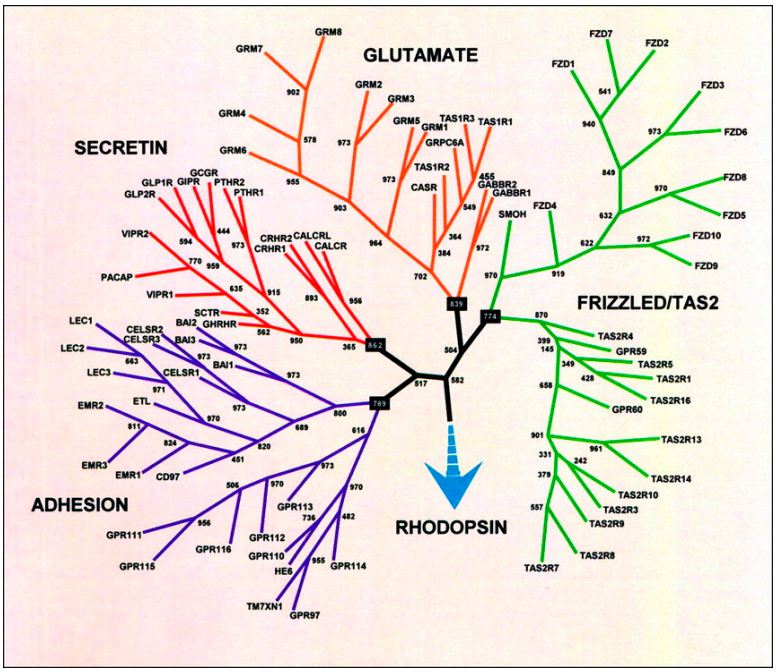


Figure 7. Phylogenetic relationship between the GPCRs (TMI–TMVII) in the human genome. The tree was calculated using the maximum parsimony method on 1000 replicas of the data set terminally truncated GPCR as described under Materials and Methods. The position of the rhodopsin family was established by including twenty random receptors from the rhodopsin family. These branches were removed from the final figure and replaced by an arrow toward the rhodopsin family analysis in Fig. 8. (Figure and figure legend are from Fredriksson *et al.* 2003)

Further, since the Rhodopsin receptor family was the largest, researchers decided to investigate it in detail and the results were very intriguing. Phylogenetical analysis revealed that receptors of this family are divided in four - distinct groups designated as: α, β, γ and δ. Even further three of these groups were formed by a number of clearly distinguishable receptor clusters. So α group was formed by prostaglandin, amine, opsins, melatonin and MECA (Melanocortin, Endothelial differentiation, Cannabinoid and Adenosin binding) receptor clusters, γ group was formed by SOG (Somatostatin, Opioid and Galanin), Melanin-concentrating hormone and Chemochine receptor clusters and δ group was formed by MAS-related (MAS1 oncogene receptor related), Glycoprotein, Purin and Olfactory receptor clusters. (Fig. 8)

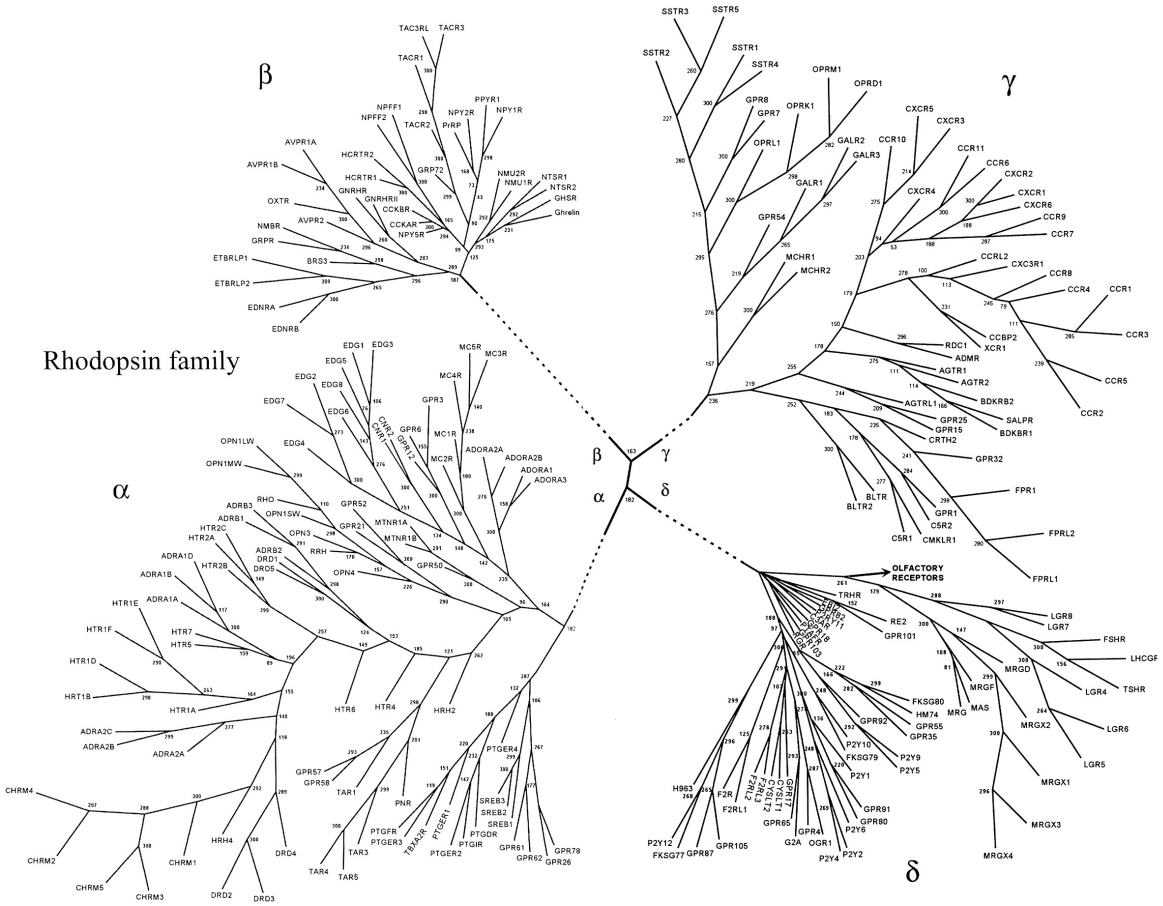


Figure 8. The phylogenetic relationship between GPCRs (TMI–TMVII) in the human rhodopsin family. The tree was calculated using the maximum parsimony method on 300 replicas. The position of the olfactory cluster was established by including 17 diverse random receptors from the olfactory cluster. These branches were removed from the final figure and replaced by an arrow toward the olfactory receptor cluster. (Figure and figure legend are from Fredriksson *et al.* 2003)

Interesting as it is from the classification point of view, this research also provided considerable evidence that all GPCRS are evolutionary related. However the question: why there is such low overall sequence similarity remained open. To resolve it Fredriksson and his colleagues generated a hidden Markov model (HMM) for every identified receptor family. Further they extracted TM domain parts of the model sequences and created separate alignments for each TM domain. Acquired results revealed that there are number of conserved motifs within various TM domains and although none of them is shared by all five families, each is connecting families in groups of two or more, which suggests that after split from common ancestor each family accumulated a set of family specific changes retaining only a small portion of ancestral sequence located in various regions, thus prohibiting similarity detection in large scale sequence comparisons. (Fredriksson *et al.* 2003)

Although very well built GRAFS classification system had one serious shortcoming: it was built on comparison of receptors found in only one organism. To compensate it authors in their further publication were focusing on receptors found in other fully sequenced organisms (Fredriksson *et al.* 2005). In this study they used previously generated HMMs and generated new ones for the groups of GPCRs that are not found in mammals, but are described in other classification systems. These HMMs were subsequently used to group all identified GPCRs of studied fully sequenced organisms into a families and the acquired results were intriguing. They revealed that all GPCR families arose prior the split of nematodes from chordate linage and that there are only few linage specific groups of receptors which led to conclusion that GRAFS classification is relevant in all analyzed bilateral species. Taken together they also highlighted the important role GPCR played during the evolution “higher species” (Fig. 9). (Fredriksson *et al.* 2005; Schioth *et al.* 2005)

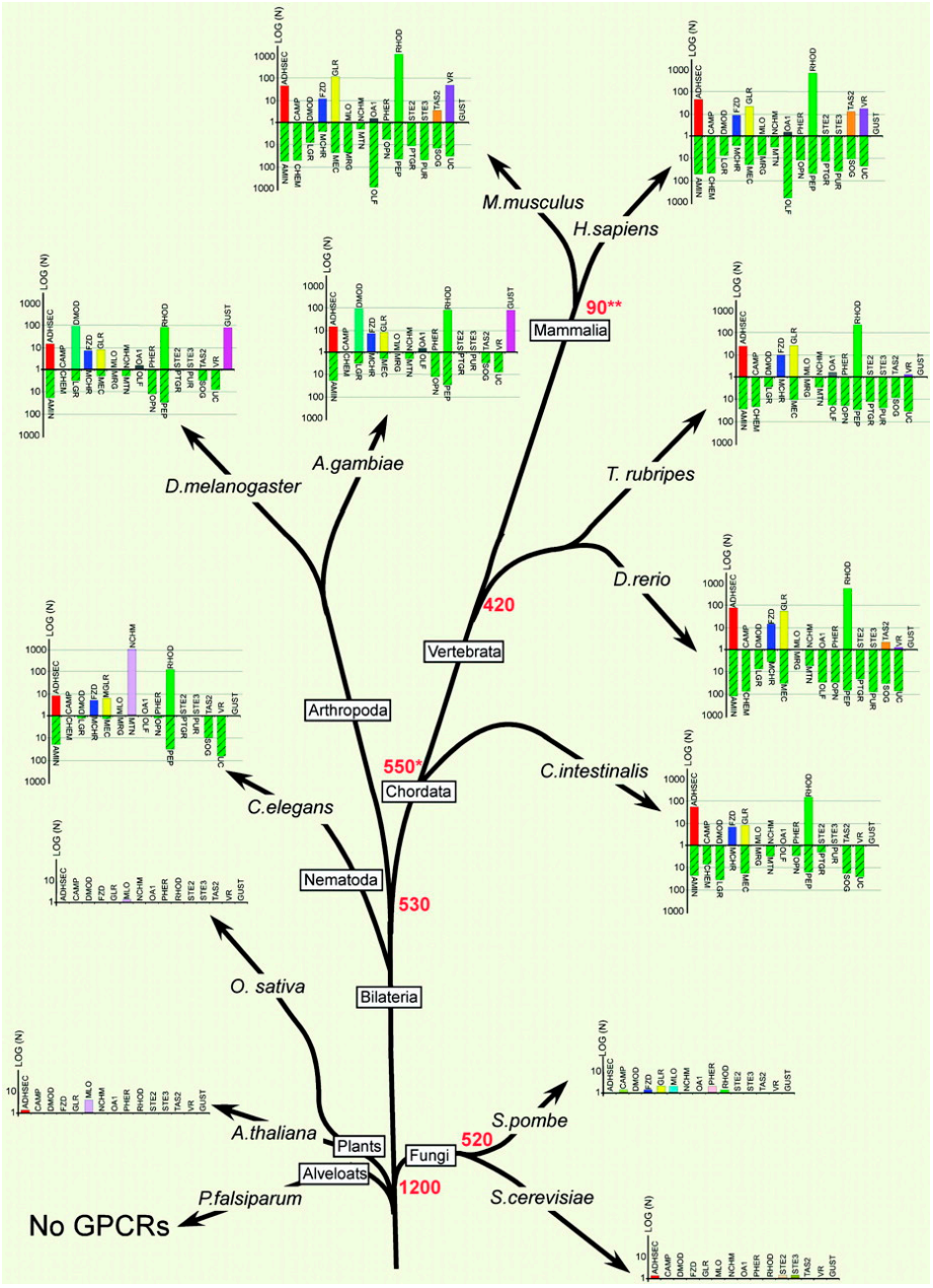


Figure 9. Evolutionary tree with the number of GPCRs in different species indicated in graphs. The graph displays the logarithm of the number of GPCRs at the y-axis, with the main class above the x-axis and the subdivision of Rhodopsin GPCRs below the x-axis (hatched). The numbers at the nodes indicate the time in millions of years since the split at that node occurred, taken from fossil data. (Figure and Figure legend are from Fredriksson *et al.* 2005)

1.1.5 Intronless state of mammalian GPCR genes and intron evolution

Rhodopsin GPCRs is one of the gene families in mammals with a large proportion of intronless genes (Brosius 1999; Gentles *et al.* 1999). This is very uncommon, since the overall density of intronless genes in human is only 12,3% (in other eukaryotic organisms varies from 2,7% in *Caenorhabditis elegan* to 97,7% in *Encephalitozoon cuniculi*). Usually the intronless state is an indicator that a particular gene family is undergoing active

changes and since the Rhodopsin GPCR family is very large this peculiarity, along with the facts that several of these receptors are recognized as ancient and that many groups have expanded at different stages of evolution, can be used to study the evolutionary history of introns. (Brosius 1999; Gentles *et al.* 1999; Sakharkar *et al.* 2004).

Another interesting fact that makes this group attractive for studies is, that although many of Rhodopsin GPCRs in mammals are intronless, most GPCRs from invertebrates contain at least one intron. Previous studies have proposed explanations to this unevenness, leading to two different conclusions. The first: number of papers suggests that the low number of introns in mammalian GPCRs is related to gene multiplication through retroposition, which means that new genes are intronless (Brosius 1999; Gentles *et al.* 1999). The second conclusion, which is quite controversial, suggests that there occurred a major loss of introns within the Rhodopsin GPCR family. It was argued that the majority of genes lost an "ancient" intron located in the conservative DRY motif of Rhodopsin GPCR (Bryson-Richardson *et al.* 2004).

After the discovery of introns in 1977 (Berget *et al.* 1977; Chow *et al.* 1977; Jeffreys *et al.* 1977), there has been continuous discussion about the functions, origin and time of origin of introns.

Regarding the functions of introns in the cell it was only the question of careful practical research and it was not long before the answer giving discoveries started to come forth. At the present moment there are at least six identified functions of introns in the cell: introns are one of the sources of non coding RNAs, a source of regulatory elements and act in alternative splicing and trans-splicing. Introns also function as enhancers of meiotic crossing over within coding sequences, in exon-shuffling and as signals for mRNA export from the nucleus and nonsense-mediated decay (Fedorova *et al.* 2003).

However the answers to questions concerning intron evolution are more difficult to obtain. In this field two main theories have been proposed. The "introns early theory" which states that introns as such must be very old formations that existed in common prokaryote and eukaryote ancestors, and that prokaryotes must have lost them during their evolution (Darnell 1978; Doolittle 1978; Gilbert 1978). As opposite is the "introns late theory" which proposes that introns must be a recent addition – selfish DNA, inserted through transposition and spread very quickly all over the genome of an ancient eukaryote before the split of the main eukaryotic lineages (Cavalier-Smith 1985). The latter theory was supported by Nicholas J. Dibb and Andrew J. Newman proposed the "proto-splice" site theory which states that conservative introns flanking MAG|R sequence (| - represents

the position of the intron, M represents A or C and R represents A or G) could serve as a signal for intron insertion (Dibb *et al.* 1989).

Ever since the beginning of formation of these two contradicting theories there has been a large amount of papers from different groups that took a side of either opinion. Some of these papers were clearly theoretical, mostly stating arguments against one or another opinion (Cavalier-Smith 1985; Fedorov *et al.* 2004), while the vast majority were built on practical investigations of modern introns (Lee *et al.* 1991; Giroux *et al.* 1994; Figueroa *et al.* 1995; de Souza *et al.* 1996; Cho *et al.* 1997; de Souza *et al.* 1997; Gilbert *et al.* 1997; Hankeln *et al.* 1997; Long *et al.* 1998; O'Neill *et al.* 1998; Tarrio *et al.* 1998; Venkatesh *et al.* 1999; Long *et al.* 2000; Fedorov *et al.* 2001; Roy *et al.* 2002; Fedorov *et al.* 2003; Roy *et al.* 2003; Tarrio *et al.* 2003; Sadusky *et al.* 2004; Roy *et al.* 2005c; Roy *et al.* 2005a; Roy *et al.* 2005b) Eventually, during the last years, there has been a growing number of publications from authors that support the "mixed theory of introns", which states that introns are extremely old formations, but some intron insertion is still occurring (de Souza *et al.* 1998; Roy *et al.* 1999; Roy *et al.* 2003).

Gene duplication is another process that goes hand in hand with intron gain and loss and also might explain the intronless state of mammalian Rhodopsin GPCR genes. It is considered to play a crucial role in development and the evolution of life, even in the RNA world (Brosius 2003) and taken together fits very well in the evolution of mammals. According to Hurles review (Hurles 2004) duplication of gene before its alternation is crucial for survival, because every gene has its function and most of the changes are leading to inactivation, which in a case of one gene may cause death of organism, but in a case of two copies one, not necessarily the new copy, might become open for changes and in time acquire new function. It is certain, that the fastest way for organism to change is through modification of regulatory pathways, a specially the ones modifying metabolism in general and ones involved in organism development, from fertilized ovule to adult organism. And since the receptors are integral part of these regulatory pathways, the fact of their amplified duplication and specific evolution raises no surprise. There is another crucial thing. A newly arisen gene, that has been duplicated in a single genome must become fixed in the population and preserved over time before it can be observed in evolutionary comparisons. The explanation of how receptor and other genes have overcome these difficulties lies in the way of formation of new species. The main requirement for formation of new species is geographical or other isolation of the group of individuals, which, in a case of a small group, automatically leads to high level of

inbreeding and greatly raises the chance of any duplication to become fixed and preserved, if it is harmless or improving. (Hurles 2004)

The gene duplication mechanisms through RNA intermediates is perhaps of most interest in this context since it is one of the oldest and originates from the time when modern life exchanged the carrier of its genetic material from RNA to DNA and it is also the one that involves intron loss. There are at least three other known single gene duplication mechanisms that are taking place in modern cells but they involve mechanisms connected with DNA rearrangement and whole gene (including introns) duplication. All of these mechanisms have been observed to happen with different frequency in different situations and they can appear at different stages of evolution (Danielson *et al.* 1999). A number of specific duplication events have been reported in different organisms and these can be lineage-specific and sometimes even species-specific (Lespinet *et al.* 2002).

1.2 Melanocortin system

The melanocortin system is involved in regulation of many different physiological functions, including the formation of pigmentation (De Wied *et al.* 1982; Spencer *et al.* 2008), adrenocortical steroidogenesis (Simpson *et al.* 1988; Allolio *et al.* 1997; Chida *et al.* 2007), energy homeostasis (Sina *et al.* 1999; Yeo *et al.* 2000; Lu 2001; Williams *et al.* 2001; Ellacott *et al.* 2004; Butler 2006; Ellacott *et al.* 2007; Muceniece *et al.* 2007), exocrine secretion (Chen *et al.* 1997; Thiboutot *et al.* 2000), sexual functions (Argiolas *et al.* 2000; Wessells *et al.* 2000; Pfaus *et al.* 2004), inflammation and immunomodulation (Haycock *et al.* 1999; Starowicz *et al.* 2003; Muceniece *et al.* 2005; Lasaga *et al.* 2008), thermoregulation (Fan *et al.* 2005; Fan *et al.* 2007) and cardiovascular regulation (Humphreys 2004; Rinne *et al.* 2008).

If looking from the functional point of view all elements of this system can be divided in three groups: signalling molecules (hormones) or melanocortins when referred to this system, melanocortin receptors that are located in the membrane of the cell and are interacting with signalling molecules thus transducting signal to interior of the cell, and the elements of this system that modify signalmolecule-receptor interaction.

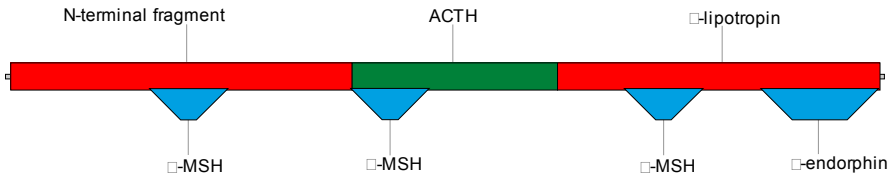
1.2.1 Melanocortin (MC) peptides

Term “melanocortin” is referred to peptides originated from proopiomelanocortin precursor (POMC), which exhibits melanotropic and adrenocorticotropic activity. MCs are produced mainly in frontal and intermediate lobe of pituitary, *Nucleus arcuatus* and

Nucleus trectus siolitarii, however lower levels of production have been found in almost all organs. They are formed from three different POMC regions and they all share the same MC conservative -His-Phe-Arg-Trp- (-H-F-R-W-) sequence that serves as pharmacophore for melanocortin receptors (MCRs). (Eberle 1988) γ-melanocyte stimulating hormone (MSH) is formed from N-terminal part of POMC, adrenocorticotropic hormone (ACTH) and α-MSH are partially overlapping peptides and are formed from the central region of POMC. β-MSH, β-lipotropic hormone (LPH) and γ-LPH are all derivates of C-terminal part of POMC (Smith *et al.* 1988) (Fig. 10)

Clear view on a wide variety of functions in regulating by melanocortins was acquired with description of the first two Caucasian children lacking products of POMC gene. Both patients displayed number of metabolism disorders as the result of hypocortisolaemia, both developed severe obesity caused by hyperphagia and both had a pale skin and red hair. (Krude *et al.* 1998) The designation of the last symptom, red hair, as POMC characteristic, however was placed under the question because in 2006 a child of the Turkish origin lacking POMC products, but retaining black hair, was identified. (Farooqi *et al.* 2006)

α-MSH and ACTH are the most abundantly studied melanocortins and as previously mentioned these are overlapping peptides thus they share a significant sequence identity. These two peptides are the most studied melanocortins since they are mediating most of melanocortin system effects. As the name of α-MSH suggests the first identified



ACTH	H_2N-SYSMEHFRWGKPVGKKRRPVKVYPNGAEDESAEAFPLEF-OH
α-MSH	AC-YSMEHFRWGKPV-NH_2
β-MSH	H_2N-AEKKDEGPYRMEHFRWGSPPKD-OH
γ-MSH	YVMGHFRWDRFG-OH
NDP-MSH	AC-YSXEHZRWGKPV-NH_2

Figure 10. Schematical representation of proopiomelanocortin precursor (upper part) and Sequence alignment of melanocortin peptides (lower part). All melanocortins, except for NDP-MSH, are formed from various regions of proopiomelanocortin precursor. All of them share conservative His-Phe-Arg-Trp (HFRW) motif. NDP-MSH is synthetic analogue of α-MSH, X represents norleucine and Z represents D-phenylalanine.

function of this peptide was stimulation of skin pigmentation formation. In a skin, α-MSH is produced in keratinocytes in response to increased UV-radiation, it acts as local hormone and after release it is binding to the receptors on the surface of melanocytes thus stimulating the production of eumelanin. (De Wied *et al.* 1982; Bohm *et al.* 2005; Spencer *et al.* 2008) In central nervous system α-MSH, β-MSH, γ-MSH and ACTH are produced in frontal and intermediate lobe of pituitary. Here the α-MSH, β-MSH and γ-MSH are also acting as local hormones binding to receptors on the surface of neurons further actions of which cause inhibition of the food intake and regulation of energy homeostasis (Yeo *et al.* 2000; Lu 2001; Williams *et al.* 2001; Harrold *et al.* 2003; Ellacott *et al.* 2004; Butler 2006; Ellacott *et al.* 2007) regulation of sexual functions (Argiolas *et al.* 2000), thermoregulation (Fan *et al.* 2005; Fan *et al.* 2007) and cardiovascular regulation (Humphreys 2004). ACTH on the contrary is being mainly released into the blood stream, and carried to adrenals where binding to its receptors is promoting the release of corticosteroids (Simpson *et al.* 1988; Allolio *et al.* 1997; Chida *et al.* 2007). It should be also acknowledged that there are reports of melanocortins, in various parts of organism, playing a crucial role in imunomodulation and regulation of inflammation, (Haycock *et al.* 1999; Starowicz *et al.* 2003; Lasaga *et al.* 2008) as well as regulation of exocrine secretion (Chen *et al.* 1997; Thiboutot *et al.* 2000) however the mechanisms involved in these functions at this point are unclear.

1.2.2 Melanocortin receptors (MCRs)

Using set of molecular cloning techniques at the beginning of the 1990s a new family of GPCRs was identified that were specifically binding melanocortin (melanotropic and adrecocorticotropic) peptides. Chronologically the first identified receptors were melanocyte stimulating hormone receptor (MSHR or MC1R) (Chhajlani *et al.* 1992; Mountjoy *et al.* 1992) and adrenocorticotropic hormone receptor (ACTHR or MC2R) (Mountjoy *et al.* 1992). Other three receptors, MC3R (Gantz *et al.* 1993a), MC4R (Gantz *et al.* 1993b) and MC5R (Gantz *et al.* 1994) were identified few years later and were named by the order of discovery, since at that time there were no information regarding their functions.

On the level of amino-acid sequence all MCRs share 39-61% identity. Interestingly MC3R, MC4R and MC5R seems to be more related to each other sharing, 55-61% sequence identity, than to MC1R and ACTHR, sharing only 43-46% sequence identity.

The least similar in this group of receptors are MC1R and ACTHR sharing only 39% sequence identity.

The MCRs are mainly acting through the $G_s\alpha$ subunit that is coupled to adenyl-cyclase (Chhajlani *et al.* 1992; Mountjoy *et al.* 1992; Gantz *et al.* 1993a; Gantz *et al.* 1993b; Gantz *et al.* 1994; Theodoropoulou *et al.* 2008), however there are also reports of MC3R acting through members of $G_{q/11}\alpha$ subunit family that are coupled to pathways involving phospholipase C-mediated hydrolysis of phosphoinositides and mobilization of intracellular Ca^{2+} (Konda *et al.* 1994).

Although all MCRs are predominantly acting through same type of Gα subunit, they all are displaying unique pharmacological profile, binding various melanocortins with distinctive affinities (Tab. 2). In a way MC3R and ACTHR are the most specialized ones. MC3R is the only of MCR that among other melanocortins effectively binds γ-MSH and ACTHR is specifically binding only ACTH. It has recently been discovered that, besides the melanocortin conservative -H-F-R-W- motif, the second pharmacophore -Lys-Lys-Arg-Arg- (-K-K-R-R-), located in the central part of ACTH, is also required for the activation of ACTHR. (Kovalitskaia *et al.* 2008)

As mentioned previously first identified MCRs were MC1R and ACTHR. The discovery of both receptors was accomplished in year 1992. Interestingly MC1R was discovered by two independent research groups that were led by Roger D. Cone and Jarl E.S. Wikberg (Chhajlani *et al.* 1992; Mountjoy *et al.* 1992), while ACTHR was discovered only by the Roger D. Cone's research group (Mountjoy *et al.* 1992). Both groups were using the same degenerate primer PCR technique. Primers of both groups were targeted for conservative GPCR TM7 region, and the only difference in strategy lied in choice of template DNA. Cone's research group used cDNA library acquired from Melanoma cells while Wikberg's group used genomic DNA of Melanoma cells.

At the present moment it is established that MC1R is predominantly expressed in melanocytes (Tab. 3) where it senses nanomolar changes of α-MSH concentrations and

Table 2. The melanocortin system (Gantz *et al.* 2003)

Receptor	Potency of Ligands	Antagonists	Primary Functions
MC1R	α-MSH=ACTH>β-MSH>_-MSH	Agouti	Pigmentation, inflammation
MC2R	ACTH	Agouti	Steroidogenesis
MC3R	α-MSH=β-MSH=γ-MSH=ACTH	Agouti, AGRP	Energy homeostasis
MC4R	α-MSH=ACTH>β-MSH>γ-MSH	Agouti, AGRP	Energy homeostasis, erectile function
MC5R	α-MSH>ACTH>β-MSH>γ-MSH		Sebaceous gland secretion

MC1R, melanocortin receptor-1; MSH, melanocyte-stimulating hormone; ACTH, adrenocorticotropic hormone; AGRP, agouti-related protein.

stimulates the formation of skin and hair pigmentation. Besides melanocytes its expression has also been detected in other types of tissues: keratinocytes, fibroblasts, endothelial cells, antigen-presenting cells and leukocytes, where it mediates the anti-inflammatory effects of melanocortines. (Gantz *et al.* 2003)

At the beginning the naming of the newly identified receptor as adrenocorticotropic was quite controversial, because there were no direct evidences of particular GPCR and ligand binding only indirect indications, like sequence similarity with the MCRs and expression in the adrenal cortex, a tissue type known to express real ACTHR. The inability to gain a conclusive proof was mainly because of difficulties to gain a sound heterologous receptor expression in endogenous MCRs expressing cell lines. Thus results acquired, at that time, could not accurately characterise pharmacological properties of expressed protein.(Noon *et al.* 2002)

The situation was changed in 1995 when Rong Qiu and his colleagues using the adrenal cell line Y1 created Y6 and OS3 cell lines lacking the endogenous MCRs expression. (Schimmer *et al.* 1995) However, the question regarding the reasons of such expression selectivity remained open. Further studies of this subject, performing comparison of ACTHR-GFP (Green fluorescent protein) fusion protein expression in CHO and Y6 cell lines showed that in all, but adrenal, cell lines this receptor is retained in endoplasmatic reticulum, which suggested that there exists some kind of ACTHR arrest signal, which is bypassed in adrenal cells. (Noon *et al.* 2002) After this study there were no significant advancements until 2005 when Matherell and his colleagues performing genetic analysis on a family with Familial glucocorticoid deficiency (FDG), identified Melanocortin receptor accessory protein (MRAP) - a small membrane protein which presence is necessary for ACTHR to successfully reach the cell plasmatic membrane. (Metherell *et al.* 2005)

ACTHR is mainly expressed in *zona reticularis* and *zona fasciulata* of adrenal cortex and in adipocytes (Tab. 3) (Boston *et al.* 1996; Beuschlein *et al.* 2001) and up to

Table 3. Expression and genomic localisation of all Melanocortin receptors

Receptor	Expression location	Genomic Localization
MC1R (MSH-R)	Melanoccytes	16q2
MC2R (ACTH-R)	Adrenal cortex, adipocytes	18p11.2
MC3R	Hypothalamus, limbic system, placenta, digestive tract	20q13.2
MC4R	Hypothalamus, limbic system, cerebrum, brain steam	18q22
MC5R	Muscles, liver, spleen, lungs, barin, adipocytes...	18p11

date it is the smallest identified GPCR. It consists only of 297 amino-acid residues with a total molecular weight of 33kDa (Mountjoy *et al.* 1992; Clark *et al.* 1996). In the adrenal cortex ACTHR mediates the effect of ACTH on glucocorticoid (Fig. 11)(Clark *et al.* 1996) and its own mRNA synthesis (Penhoat *et al.* 1989; Naville *et al.* 1999). In the mouse adipocytes it has been shown to play a role in lypolysis (Boston *et al.* 1996; Boston 1999), leptin (Norman *et al.* 2003) and interleukin 6 production (Jun *et al.* 2010), however the functions of ACTHR in human adipocytes has not been clarified yet.

Genetic defects in ACTHR gene are causing Familial glucocorticoid deficiency (FGD) - a rare autosomal recessive disorder, which is characterized by low or undetectable plasma cortisol levels and an excess of plasma ACTH. It becomes apparent in childhood as frequent hypoglycaemia and/or infective episodes accompanied by excessive skin pigmentation. For all (FGD) patients that had the alterations in ACTHR gene, at least one of gene alleles was coding partially functional receptor, while the other one was frequently coding completely inactive receptor (Elias *et al.* 1999; Penhoat *et al.* 2002)

Despite the already identified receptors, there were numerous evidences suggesting the existence of other high affinity MCRs in central nervous system (CNS) and other tissues besides the adrenal cortex and melanocytes. The authors of the time pointed that melanocortin peptides have a various effects upon cerebration, for example, memory (Sandman *et al.* 1969; Garrud *et al.* 1974), thermoregulation (Feng *et al.* 1987) etc. and

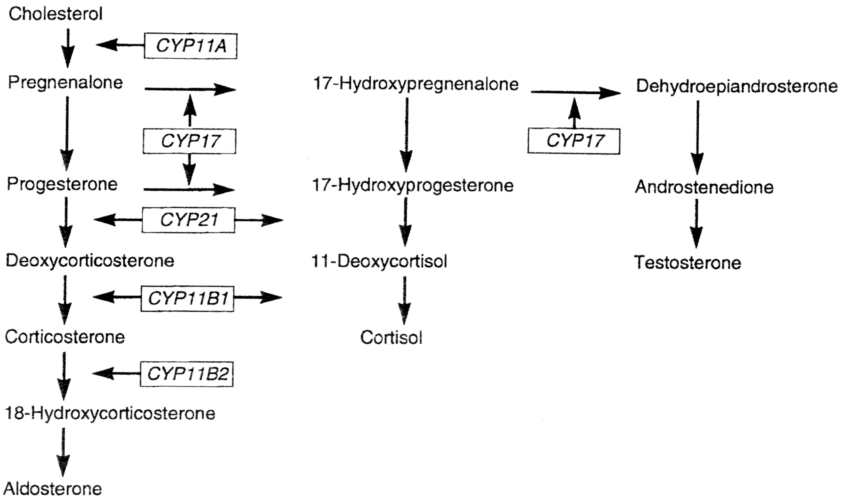


Figure 11. ACTH Regulates the expression of certain adrenal steroidogenic enzyme genes as depicted in the boxes which catalyze the steps in glucocorticoid, mineralcorticoid and androgen synthesis as indicated. (Figure and figure legend are from: Clark *et al.* 1996)

there were also reports of high affinity MSH and ACTH binding sites in Lacrimal glands, CNS and peripheral tissues. (Tatro *et al.* 1987; Hnatowich *et al.* 1989; Tatro 1990; Salomon *et al.* 1993)

Because of these evidences the searches for new MCRs were continued and as the result, using Northern Blot hybridisation method, the expression of a new - MC3R was detected in CNS and placenta. The cDNA of this receptor was acquired using reverse transcription from cells of stomach, duodenum and pancreas. (Gantz *et al.* 1993a) Further, using *in situ* hybridisation, it was established that mRNA of MC3R is transcribed in more than thirty nucleuses of rat CNS, mainly in hypothalamus and other structures of limbic system. (Roselli-Rehfuss *et al.* 1993) It is also notable that MC3R is only MCR which binds γ-MSH and receptor affinity for this ligand is close to affinity for other melanocortins. Further studies of this receptor determined that it is mainly involved in regulation of energy homeostasis. (Tab. 3) (Gantz *et al.* 2003)

MC4R was discovered the same year as MC3R by the same group of researchers using the same set methods. (Gantz *et al.* 1993b) Detailed neuroanatomical *in situ* hybridisation mapping revealed that mRNA transcription of this receptor is wider spread than transcription MC3R mRNA. mRNA of MC4R was detected in almost every part of the brain - in ~148 different cerebral nucleuses including cerebral cortex, thalamus, hypothalamus, limbic system, brain stem and medulla (Tab. 3) (Mountjoy *et al.* 1994). As for the functions of this receptor, it was already known prior to MCR discovery that melanocortin peptides α-MSH and ACTH can influence the food intake of laboratory animals. (Panskepp *et al.* 1976; Poggioli *et al.* 1986) Later experiments with the synthetic nonselective MCR agonists and antagonists confirmed these observations. Laboratory Rats treated with the Malanoton II (synthetic agonist) displayed dislike for the offered food (Thiele *et al.* 1998) which was followed by delay of food intake (Fan *et al.* 1997). The treatment with the SHU9119 (synthetic antagonist) on the other hand increased the day, night and fastening induced food intake. (Fan *et al.* 1997; Kask *et al.* 1998c) It was also established that simple limitation of the access to the food decreased the transcription level of POMC coding mRNA in *Nucleus arcuatus*. (Brady *et al.* 1990; Kim *et al.* 1996) More detailed studies involving MC4R selective antagosnist HS014 revealed that treatment of laboratory rats with this substance induced severe food intake during the daytime. (Kask *et al.* 1998a; Kask *et al.* 1998b; Kask *et al.* 1998d) Taken together all these facts suggest that endogenous ligands of MC4R are food intake inhibiting signals which are preventing animal overeating through stabilisation of amounts of daily food intake.

The last identified receptor of this family was MC5R. It was discovered a year later by the group that identified previous two receptors. (Gantz *et al.* 1994) This is the most abundantly expressed MCR, the expression has been identified in muscles, thymus, spleen, ovary, testis, adrenal cortex, lungs brain and adypocytes. (Tab. 3) There is very little known about its functions and only detected anomaly of MC5R knock-out mice was the dysfunction of skin exocrine, especially sebaceous, glands. (Chen *et al.* 1997)

1.2.3 Other elements of MC system

Besides MCs and their receptors there are six other elements in MC system: Agouti signalling peptide (ASiP) (or Agouti in mice), Agouti related protein (AgRP), Mahogany, Syndecan-3, Melanocortin receptor accessory protein-α (MRAP-α), MRAP-β and MRAP2.

Agouti and AgRP are paracrine signalling molecules. Currently they are the only identified endogenous GPCR antagonists, thus making MC system unique. Agouti is specifically binding to all MCRs except for MC5R while specific binding of AgRP is limited to MC3R and MC4R. Another specificity of these elements is that they are very large peptides, consisting of ~130 amino-acid residues. (Lu *et al.* 1994; Shutter *et al.* 1997)

As the names of both proteins indicate Agouti was discovered prior to AgRP. (Lu *et al.* 1994) The name *agouti* comes from zoology where it is used for description of specific fur colour (yellow band on otherwise black or brown background) of various mammalians. Besides yellow fur colour *agouti* mice were tended to develop obesity, insulin resistance increased somatic growth and tumorogenesis. After discovery of its coding gene, it was noticed that all these effects are caused by widespread and unregulated expression of Agouti protein while in genetically healthy mouse its expression is limited to skin where it inhibits the action of α-MSH and ACTH on MC1R. (Gantz *et al.* 2003) Unlike Agouti the human homologue - ASiP is expressed in a broader range of cell types that includes: adipocytes, testis, ovary, heart, foreskin, kidney and liver (Wilson *et al.* 1995), but since the agouti hair colour has never been observed in humans its participation in formation of pigmentation is implausible, even more the question regarding the real functions of this protein in humans remains open.

AgRP was discovered three years after Agouti. The coding sequence of this peptide was acquired thanks to the high level of sequence homology of both proteins. As mentioned above AgRP although similar to Agouti has a distinct pharmacological profile binding only to MCRs located in CNS. Besides differential pharmacology, the expression pattern of both proteins is also distinctive - AgRP is mainly expressed in *nucleus arcuatus*

that is located within the hypothalamus (Shutter *et al.* 1997). The expression of the shorter transcript (lacking the 5' noncoding exon) has also been detected in adrenal cortex, testis, lung, kidney and dorsal root ganglia (Beltramo *et al.* 2003). The physiological role of AgRP in hypothalamus is to bind to adjacently localised MC3R and MC4R and inhibit their food intake suppressing action and thus stimulate the appetite (Ilnytska *et al.* 2008). Studies concerning the function of AgRP expressed in adrenal cortex have shown that, besides ACTHR and MC5R, minor amounts of MC4R and possibly MC3R are also being expressed in adrenal cortex were they are inducing cortisol release and, since AgRP is antagonist of these receptors, it might participate in regulation of cortisol release. It should be also mentioned, another theory regarding the action of AgRP in adrenal cortex states that there might also be other than MCRs high affinity binding sites for AgRP (Doghman *et al.* 2007). However there is very little or no knowledge about functions of AgRP when expressed in other parts of organism.

Mahogany and Syndecan-3 are single pass transmembrane proteins that are believed to alter the activity of Agouti and AgRP respectively. Mahogany is mainly expressed in brain and skin, and it has been shown to be a low affinity receptor for Agouti but not AgRP. The idea that it might be part of MC system came from observation that *mahadony* mutation completely suppresses *agouti* caused fur colouring and health problems (obesity, insulin resistance, etc.) in mice. Knowing that Mahogany is only a low affinity receptor however, it is hard to explain such dramatic effect. Taken together with the discovery that Mahogany has its own, Agouti independent, effect on metabolic rate it leaves the question of its membership with the MC system open. Syndecan-3 is mainly expressed in the hypothalamus. Experiments involving Syndecan-3 coexpression with MC4R revealed the increasing efficiency of AgRP antagony. It has been therefore proposed that Syndecan-3 might act as AgRP coreceptor. This hypothesis however is also disputable since Syndecan-3 knock-out mice did not have any specific phenotype. (Gantz *et al.* 2003)

The last three members of MC system are MRAP-α, MRAP-β and recently discovered MRAP2. All three are coupling to MCRs and, in the case of ACTHR, are promoting membrane transport or, in the case of other MCRs, inhibiting their functional activity. (Chan *et al.* 2009) All three MRAP's are singe α-helical TM proteins. MRAP-α and MRAP-β are formed by alternative splicing from one gene. The gene is located in the locus 22.1 of the long arm of chromosome 21 (21q22.1), it is ~25Kb long and consists of six exons. 172 amino-acid isoform α is translated from mRNA formed by first five exons

of the gene while 102 amino-acid isoform β is formed from mRNA by third, fourth and sixth exons of the gene. Translation initiation codon is located in the third exon and TM of protein is encoded by fourth exon of the gene, thus both proteins have identical N-terminus and TM, but distinctive C-terminus. Analysis of expression patterns revealed that different types of tissues have differential preference for MRAP isoforms. Both subtypes are expressed in adrenals, testis, breast, ovary, adipocytes, skin and jejunum. MRAP-α alone is also expressed in thyroid, lymph nodes, ileum, liver, stomach and pituitary, while MRAP-β alone is expressed only in brain. (Metherell *et al.* 2005) Such diverse expression pattern of α and β isoforms might be explained by fact that ACTHR affinity for ACTH depends on coexpressed MRAP isoform (Roy *et al.* 2007). As mentioned previously MRAP2 was discovered quite recently, its gene is located in the locus 14.3 of the long arm of chromosome 6 (6q14.3), in total it consists of four exons and encodes 205 amino-acid protein. MRAP2 is expressed in brain and adrenals. (Chan *et al.* 2009) MRAP and MRAP2 share relatively low 39% sequence homology in their N-terminal and TM region. Cross species studies of these proteins revealed two primate specific aspects of MRAP genetics. MRAP-β and first two exons (containing 5'UTR) are found only in this linage of vertebrates. Interestingly when all three proteins are compared with their respective homologues from other organisms there was a higher sequence similarity among MRAP2s than among MRAPs, which could indicate that MRAP2 more resembles the ancestral MRAP. (Chan *et al.* 2009; Webb *et al.* 2010) Regarding the mechanism of MRAP action, in recent few years a large number of important findings have been made. First major discovery, revealed that, in the cell, MRAP exists as a very stable, SDS resistant, homodimer. This by itself isn't a unique case, because many membrane proteins, including GPCRs (Mandrika *et al.* 2010) are known to form dimers. The MRAP dimer however has the unique orientation of monomers - they are antiparallel, meaning that C-terminus of first monomer on the cytoplasmatic side of cell membrane while C-terminus of second on the extracellular side of cell membrane. (Sebag *et al.* 2007; Cooray *et al.* 2008) It was later reported that region formed by amino-acid residues 18 to 21 (-L-D-Y-L-),, located in N-terminal tail, is crucial for ACTHR ligand recognition. Substitution of these amino-acid residues with alanine resulted in protein that promoted only the trafficking of receptor to plasma membrane but not signalling. This observation was also supported by the fact that ACTHR complex with MRAP2, which lacks this region, but is able to couple to ACTHR and promote its transport to cell membrane, displayed no functional activity. (Sebag *et al.*

2009; Webb *et al.* 2009) These studies also revealed that region formed by amino-acid residues 31 to 37 (-L-K-A-N-K-H-S-), located in adjacently to N-terminal part of TM region, is important for MRAP to form antiparallel dimmers. Removal of this domain resulted in formation of protein which N-terminus is always located on the extracellular side of the cell membrane. At the present moment it is believed that during translation of MRAP molecule it is always inserted into the membrane with its N-terminus located inside the lumen of endoplasmatic reticulum, however the previously mentioned TM N-adjacent region is making this orientation somewhat unstable and part of the molecules are swapping to alternative one. It is possible that this swapping continues on until the antiparallel dimer is formed. (Sebag *et al.* 2009) The last, but not least, of these major discoveries revealed that TM domain of MRAP is responsible for specific coupling to ACTHR, although the nature of this interaction at the present moment is unclear. (Sebag *et al.* 2009; Webb *et al.* 2009) In the conclusion it should be mentioned that previously described wide expression of all types of MRAP could indicate that this protein is also participating in regulation of other functions, outside melanocortin system. (Hinkle *et al.* 2009; Webb *et al.* 2010)

1.3 Chimerical proteins: prospects of application

Directed modification of proteins, in order to improve or modify their characteristics is one of most important goals of biotechnology. Even more, this also is very useful tool to study protein biochemistry. Today *in vitro* DNA synthesis and recombinant DNA technology have made it possible to synthesize virtually any type of polypeptide, leaving only one question: what protein do we want to synthesize? There are two types of directed modification: formation of large random mutation libraries or modification through exchange of various regions of protein with other already existing ones. As stated by Nixon and his colleagues these regions might be individual amino-acid substitutions (point mutations) as well as secondary structure domains or whole subunits. (Nixon *et al.* 1998) Usually when it is needed for the resulting protein to be with the same functional properties, but with altered physical parameters, substitution of individual amino-acids or domains with homologue regions from other-similar proteins are applied. However when working with the chimerical proteins it should be kept in mind that introduction of large scale changes might disrupt the network of complex interactions, thus rendering acquired protein in functionally inactive state. Taking in to an account this information, introducing of point mutations might seem to be a safer way to go, however it

should be noted that exchange of defined structural domains to homologue structural domains can be successfully implemented without the loss of protein functional activity. (Beguin 1999)

Although complex in theory, creation of chimerical proteins have been successfully applied to create new chimerical restriction enzymes (Chandrasegaran *et al.* 1999). Regarding receptors this method has been used to identify functions of specific domains or to identify domains responsible for specific functions, for example to determine domains of VPAC receptor that are responsible for binding with calcium effector (Langer *et al.* 2002), or to identify which parts of D1B receptor are involved in signal transduction (Tumova *et al.* 2003). This approach has also been successfully implemented in studies of MCRs. In this way domains of MC3R and MC4R that determine distinctive coupling efficiency to G-protein were determined (Kim *et al.* 2002), domains of MC1R and MC3R that are involved in ligand binding (Schioth *et al.* 1998), domains of MC3R that determine receptors specific affinity for γ-MSH (Oosterom *et al.* 1999), as well as domains of MC4R that are involved of receptor binding with AgRP (Yang *et al.* 1999).

Also it should not be forgotten that fluorescent proteins are also often fused with the protein of interest. The purpose of such fusion is to measure the expression levels, localisation in cell, co-localisation, concentration or other parameters of studied protein.

1.4 Fluorescent proteins

At the present moment the most abundantly used fluorescent protein is green fluorescent protein (GFP), which was initially acquired from in northwest Pacific Ocean living jellyfish *Aequorea victoria*. *Aequorea victoria* GFP is 238 amino-acids long protein that, while absorbing blue light (maximal absorption at 395 and 475 nm), emits green light (maximal emission at 509 nm), it is very resistant to cell proteases, provides stable fluorescence, doesn't photobleach for at least 10 minutes and most importantly it doesn't require any special cell cofactors to retain full functionality. (Chalfie *et al.* 1994).

X-ray crystallography analysis of GFP that was carried out in 1996 revealed that its structure visually resembles 42Å high "barrel", with 24Å diameter. This barrel like structure is formed by 11 parallel, almost ideal cylinder forming, in helix twined β-sheets. Chromophore - *p*-hydroxybenzylidene-imidazolidone, which is formed by cyclization of Ser^{65} and Gly^{67} and oxidation of α-β bond of Tyr^{66}, is located within the α-helix at the centre of this "barrel". (Fig. 12 and 13) This structure of β-sheets enclosed chromophore is

determining the high resistance of GFP to proteases and other agents of environment, because most of proteases are unable to cleave β-sheets and their barrel-type placement serves as the barrier, preventing influence of various chemical agents on action of chromophore. (Ormo *et al.* 1996)

After its discovery number of research groups started to improve the properties of GFP. One of the first was attempt to reduce its molecular mass, however it turned out to be unsuccessful. It was established that GFP can be truncated only by one amino-acid from N-terminus and 10 to 15 amino-acids from C-terminus before losing its fluorescence ability. Further attempts of improvement included introduction of point mutations. Results acquired from these studies can be divided in two groups: mutations that influence the relative proportions of absorbance peaks and mutations that shift absorption and emission spectra.

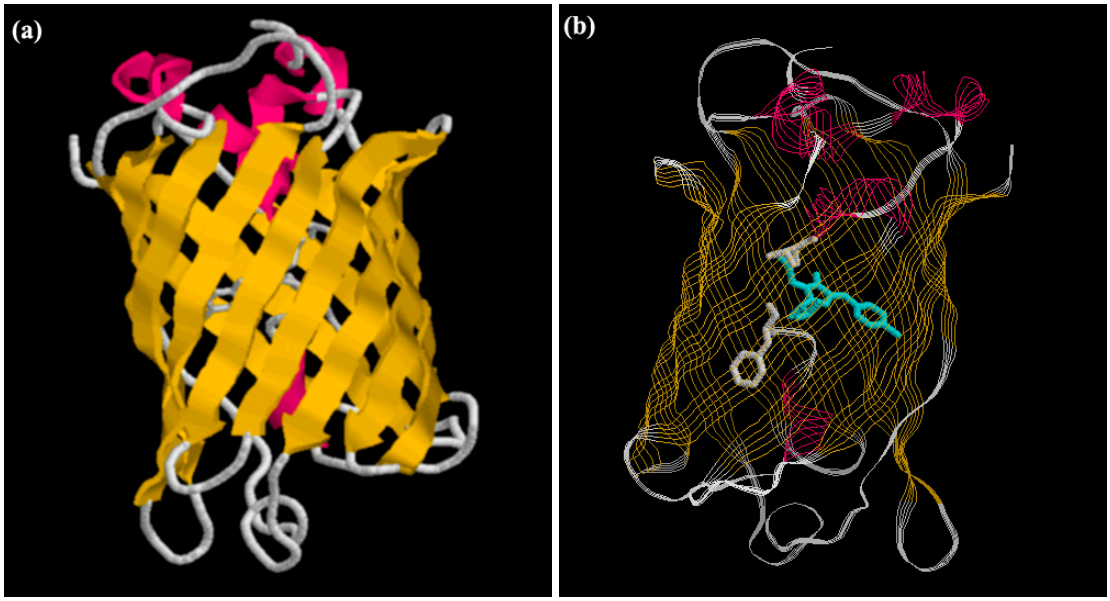


Figure 12. Schematically depicted 3D structure of GFP
(a) β-sheet structured regions are coloured in - , α-helix structured regions are coloured in - and unstructured regions are coloured in -
(b) chromophore - *p*-hydroxybenzylidene-imidazolidone is coloured in - and to chromophore proximal amino-acids: Ser^{65} and $Gly^{67}$are coloured in -
Images were created using OpenRasMol software (Bernstein 2000) and structure published by Ormo and colleagues (Ormo *et al.* 1996)

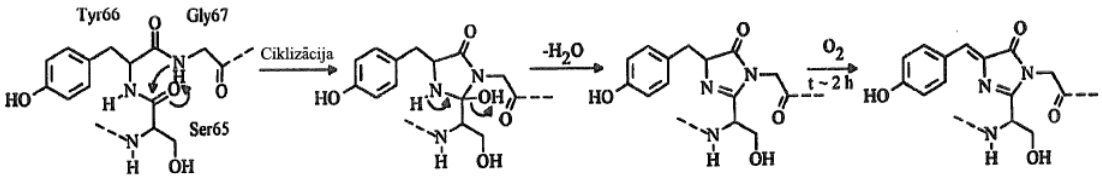


Figure 13. Formation of GFP chromophore - *p*-hydroxybenzylidene-imidazolidone. (Image from Cubitt *et al.* 1995)

The most striking examples of changes in relative proportions of absorbance peaks are mutations $Thr^{203}\rightarrow Ile$ and $Glu^{222}\rightarrow Gly$, that are simplifying absorption spectra to one peak 395 and 475 nm respectively, and $Ile^{167}\rightarrow Thr$ that inverts the absorption peak intensities of native GFP

Most of at the present moment identified absorption and emission spectra shifting mutations are located in chromophore proximal amino acids. (Tab. 4)

Exchange of the central tyrosine (Tyr^{66}) with other aromatic amino-acid (Trp, His or Phe) results in shift of absorption and emission spectra to shorter wavelength. Because of it's only slightly shifted absorption and greatly shifted emission, fluorescent protein with $Tyr^{66}\rightarrow His$ mutation has been designated the name of blue fluorescent protein (BFP). However Despite these intriguing results all Tyr mutants had one common flaw – significantly lower fluorescence level. (Cubitt *et al.* 1995)

It was also established that Ser^{65} exchange for Cys, Leu, Val or Ala resulted in loss of 395 nm 475 nm absorption peak and shift of 475 nm absorption peak to slightly longer wavelengths. Fluorescent protein with the greatest shift of absorption and emission peaks to longer wavelength formed with introduction of mutation $Ser^{65}\rightarrow Thr$. This mutant possessed three excellent properties: its emitted fluorescence was six times greater, its photobleaching was significantly slower and oxidation of its Tyr^{66} α-β bond, during formation of chromophore, was four times faster. (Cubitt *et al.* 1995)

Based on these observations several research groups tried to improve the properties of GFP and its colour derivatives. Some of the improvements included attempts to maximize the expression yield of properly folded fluorescent proteins while others included attempts to improve properties involving efficiency and stability of fluorescent proteins. First approach, for example, was used by Yang and his colleagues, they created synthetic EBFP (Enhanced blue fluorescent protein) gene sequence containing codons preferentially found in highly expressed human proteins, additionally they introduced four point mutations ($Phe^{64}\rightarrow Leu$, $Ser^{65}\rightarrow Thr$, $Tyr^{66}\rightarrow His$, and $Tyr^{145}\rightarrow Phe$) in order to

optimize the expression of protein. (Yang *et al.* 1998) Second approach was used by several groups. Cormack and his colleagues and Mena and his colleagues created a libraries of random point mutations in region adjacent to chromophore of GFP and BFP respectively. Afterwards members of both groups selected those mutants that presented improvement in stability and fluorescence intensity. As the result Cormack acquired $Phe^{64}\rightarrow$Leu and $Ser^{65}\rightarrow$Thr mutations containing EGFP (Enhanced green fluorescent protein), that had 30 fold greater fluorescence intensity than native GFP (Cormack *et al.* 1996) and Mena acquired $Phe^{64}\rightarrow$Leu, $Val^{150}\rightarrow$Ile, $Val^{150}\rightarrow$Ile and $Val^{224}\rightarrow$Arg mutations containing Azurite, that had 1.6 fold increased quantum yield and 40 fold reduced photobleaching (Mena *et al.* 2006).

Table 4. The effect ov mutations within the chromophore region of GFP (Table and legand are from Cubitt *et al.* 1995)

Ser^{65} mutations	Absorption max (nm)	Emission max (nm)	**Tyr^{66} mutations**	Absorption max (nm)	Emission max (nm)
Native GFP	395/475	508, 503	Natīvais GFP	395	508
$Ser^{65}\rightarrow$Ala	–/471	–/504	$Tyr^{66}\rightarrow$Phe	360	442
$Ser^{65}\rightarrow$Cys	–/479	–/507	$Tyr^{66}\rightarrow$His	382	448
$Ser^{65}\rightarrow$Leu	–/484	–/510	$Tyr^{66}\rightarrow$Trp	436	485
$Ser^{65}\rightarrow$Thr	–/488	–/511			

For Tyr^{66} mutants 475 nm absorption peak was very low and was considered as insignificant.

Suggested literature to gain basic understanding on the subject of the Doctoral thesis:
Fedorova *et al.* (2003), Dibb (1993), Logsdon *et al.* (1998), de Souza *et al.* (1998), Tuteja (2009), Jalink *et al.* (2010), Millar *et al.* (2010), Fredriksson *et al.* (2003), Schioth *et al.* (2005), Gantz *et al.* (2003), Cone (2006), Webb *et al.* (2010)

2 MATERIALS AND METHODS

2.1 E. coli stain

DH5alfa - F-, phi80dlacZDelta M15, Delta(lacZYA-argF), U169, deoR, recA1, hadR17(rk-, mk+),gal-, phoA, supE44, Lambda-, thi-1, gyrA96m relA1.

2.2 Mammalian expression cell lines

HEK 293 (EBNA) - Human Embryonic Kidney cell line (Invitrogen)
BHK – Baby Hamster Kidney cell line (Invitrogen)

2.3 DNA agrose gel electrophoresis

DNA agrose gel electrophoresis was carried out in TAE buffer (40 mM Tris-acetate pH 8,3, 20 mM Sodium acetate, 1 mM EDTA) 1,2% agarose (Fermentas, #R0491) gel with ethidium bromide to final concentration 1μg/ml added.

2.4 DNA fragment purification and extraction from agarose gel

For all DNA gel extraction and purification steps Silica Bead DNA Purification and Gel Extraction Kit (Fermentas #K0513) was used.

2.5 Plasmid DNA extraction

Bacterial cells acquired from single colony or frozen stock were inoculated in 6ml (miniprep) or 150 ml (midiprep) of 2xYT media (16 g/l Bacto tryptone, 10 g/l Bacto yeast extract and 5 g/l Sodium chloride) with ampicillin added to final concentration 1mg/ml and incubated for 16-24h at 37ºC with aeration (~200rpm). For miniprep DNA extraction GeneJet™ Plasmid Miniprep Kit (Fermentas #K0503) was used and for midiprep DNA extraction JETSTAR 2.0 Plasmid Midiprep Kit (Genomed #210050) was used

2.6 Preparation of transformation competent *E. coli* cells

Bacterial cells acquired from single colony or freezed stock of desired stain were inoculated in 6ml of LB media (10 g/l Bacto tryptone, 5 g/l Bacto yeast extract and 10 g/l Sodium chloride) with no antibiotics added and incubated for 16-24h at 37ºC with aeration (~200rpm). ~1ml from overnight culture was inoculated into 150ml of Psi broth (20 g/l Bacto tryptone, 5 g/l Bacto yeast extract, 5 g/l magnesium sulphate and pH 7,6 was adjusted using potassium hydroxide). The volume of overnight culture was scaled up or down as needed so the result was $0{,}03 \leq OD590 \leq 0{,}04$ and incubated at 37ºC with aeration (~200rpm) till OD_{590} = 0,3 - 0,4. Further culture was incubated on wet ice for 15min and transferred to two precooled 50ml centrifugation tubes afterwards Cells were pelleted for 5min at 3220×g. Supernatant was discarded and 20ml of TfbI (30mM potassium acetate, 100mM rubidium chloride, 10 mM calcium chloride, 50mM magnesium chloride, 15% v/v glyserol and pH 5,8 was adjusted using acetic acid) were added to each tube. Pelleted cells were resupended, incubated on wet ice for 15min and pelleted repeatedly for 5min at 3220×g. Supernatant was discarded and 2ml of TfbII (10mM MOPS, 75 mM calcium chloride, 10mM rubidium chloride, 15% v/v glyserol and pH 6,5 was adjusted using sodium hydroxide) were added to each tube. Pelleted cells were resupended, incubated on wet ice for 15min, quick-freezed in liquid nitrogen in 100μl aliquots and stored in at least -70ºC

2.7 PCR directed mutagenesis

For creation of mutant receptor genes Overlapping ends PCR directed mutagenesis approach was applied. During the first stage of the process following reaction mix was prepared in two microcentrifuge tubes: 5μl of 10× amplification buffer, 1μl of 10 mM dNTP Mix, 5μl of 10μM mutant DNA primer Fw or Rs, 5μl of 10μM standard DNA primer Rs or Fw, 1μl of DNA of interest, 1μl (~4u) of *Pfu* DNA Polymerase (Fermentas, #EP0509) and MiliQ water up to 50.μl. Afterwards the tubes were placed in thermal cycler and after 2min of predenaturation at 95ºC, 40 cycles of following conditions were applied: Denaturation, 15s at 95ºC, Annealing, 15s at tº that depended on each primer pair individually (usually ~55ºC), Synthesis, 1min 72ºC. After the last cycle additional step of Synthesis was applied: 7min 72ºC. Prior the next reaction both

products were purified from remainder of PCR primers using agarose gel electrophoresis and extracted from acquired slices of gel. During the second stage of the process ollowing reaction mix was prepared: 5µl of 10×amplification buffer, 1µl of 10 mM dNTP Mix, 5µl of 10µM standard DNA primer Fw, 5µl of 10µM standard DNA primer Rs, 2,5µl of both in previous stage amplified PCR products, 1µl (~4u) of *Pfu* DNA Polymerase (Fermentas, #EP0509) and MiliQ water up to 50µl. Afterwards the tube was placed in thermal cycler and after 2min of predenaturation at 95ºC, 40 cycles of following conditions were applied: Denaturation, 15s at 95ºC, Annealing, 15s at 55ºC, Synthesis and 1min 72ºC. After the last cycle additional step of Synthesis was applied: 7min 72ºC. Acquired PCR product was purified and its quality was tested by agarose gel electrophoresis.

2.8 Restriction of DNA

For restriction of DNA the following reaction mix was prepared: 2µl of 10×appropriate restriction buffer (restriction buffer Fermentas #BR5, restriction buffer G Fermentas, #BG5), 5-18µl of DNA (volume depended on type and concentration of DNA), 0,5µl of each of required restriction enzymes (*Hind*III Fermentas, #ER0501 or *Xho*I Fermentas, #ER0691 or *Bam*HI Fermentas, #ER0051), and MiliQ water up to 50µl. The reaction mix was incubated for 1h at 37°C.

2.9 Ligation of DNA fragments

For ligation of DNA fragments following reaction mix was prepared: 1µl of 10×ligation buffer, 1µl of 10mM rATP, 1µl of Vector DNA, 7µl of Fragment DNA, 1µl of T4 DNA ligase (Fermentas, #EL0011). The reaction mix was incubated overnight at 22°C.

2.10 Transformation of *E. coli* cells

Stock of transformation competent *E. coli* cells was removed from freezer (≤ -70°C) and thawed on wet ice. 10µl of ligation mix or 1µl of plasmid solution was added afterwards tube was incubated on wet ice for 30 min. The cells were heat shocked for 45 s at 42°C after which incubated on wet ice for 2 more min. Further the cells were transferred to 0,9ml of prewarmed LB (10 g/l Bacto tryptone, 5 g/l Bacto yeast extract and 10 g/l Sodium chloride) media with no antibiotics added, incubated for 1h at 37ºC with aeration (~200rpm) and pelleted for 5min at 800×g. Most of supernatant was carefully discarded leaving ~100 µl of media. Cell pellet was then resuspended in this remainder and spread on pre warmed media-LB agar (10g/l Bacto tryptone, 5g/l Bacto yeast extract and 10g/l Sodium chloride, 15g/l Bacto agar, ampicilin 1mg/ml) plate.

2.11 Sequencing of DNA

For sequencing of DNA following reaction mix was prepared: 2µl of 5×amplification buffer (400mM TrisCl pH 9.0, 10mM $MgCl_2$), 1µl of BigDye® Terminator v3.1 Mix (Applied Biosystems), 1µl of 5µM DNA sequencing primer, 1µl of matrix DNA and 5µl MiliQ water. Afterwards the tubes were placed in thermal cycler and after 1min of predenaturation at 96ºC, 35 cycles of following conditions were applied: Denaturation, 10s at 96ºC, Annealing, 10s at 52ºC and Synthesis, 4 min 60ºC. After end of reaction DNA from reaction mix was precipitated by addition of 1µl of 7,5M ammonium acetate and 30µl of 96% ethanol. Sample was vortexed, incubated at room temperature for 10 min and centrifuged for 15min at 20000×g. Supernantant was carefully discarded and 50µl of 70% ethanol were added. Then sample was repeatedly centrifuged for 5min at 20000×g, supernantant was carefully discarded and pellet was air-dried at 55ºC for 10 min.

2.12 Transfection and cultivation of mammalian expression cell lines

2.12.1 Creation of semi-stable HEK 293 EBNA cell lines for cAMP and binding assays

70µl of OptiMEM (Gibco/Invitrogen, 31985-047) media were transferred to sterile tube and 20µl of FuGENE (Boehringer Mannheim, Germany) transfection reagent were added directly to OptiMEM media and the mix was incubated in the room temperature for 5min. In another sterile tube 7µg of plasmid DNA and up to 30µl of OptiMEM media were mixed, then FuGENE containing mix from the first tube was transferred to second by direct pipeting into the solution. Combined mix was transferred by dripping into media (Dulbecco modified Eagle media (DMEM) (Gibco/Invitrogen, 31331-028), with 10% Fetal calf serum (FCS) (Gibco/Invitrogen, 10270-106), 250µg/ml geneticin (Gibco/Invitrogen, 10131-027), 100U/ml penicillin, 100

µg/ml streptomycin (Gibco/Invitrogen, 15140-122) and 25µg/ml Amphotericin B (Gibco/Invitrogen, 15290-026)) of mammalian cell culture containing 100mm petri plate and the cell culture was incubated for 24h in CO_2 incubator (37ºC, 5% CO_2). Ater this period the Hygromycin B (Invitrogen, 10687-010) was added to final concentration 200 µg/ml and the cell culture was cultivated for ~3 weeks, during which every 48 h culture media was replaced and Hygromycin B added.

2.12.2 Transfection of BHK cells for confocal microscopy

1 µg of plasmid DNA was diluted in 100 µl of serum-free DMEM (with, 100 U/ml penicillin, and 100 µg/ml streptomycin) and 2 µl of TurboFect™ (Fermentas, #R0531) were added to the diluted DNA. Acquired solution was incubated for 15-20 minutes at room temperature and evenly distributed on the bottom of well of a 24-well plate containing Ø12mm microscopy cover slip. 1 ml of ~105 BHK cells was gently layered top of the TurboFect™/ DNA mixture and cells were incubated at 37°C in a CO_2 incubator for ~24h prior microscopy analysis.

2.13 cAMP response assay

2.13.1 Stimulation

H^3 labelled adenosine (Amersham Pharmacia Biotech, TRK343) was added to cell culture containing petri plate to final concentration 5 µCi/ml (50µl of adenosine for 10ml of media) and cell culture was incubateded at 37°C in a CO_2 incubator for 2h. For the first dilution in a suitable tube ligand: NDP-MSH (Neurosystem France) or ACTH (Neurosystem France) was diluted with Basal buffer (137mM NaCl, 5mM KCl, 0,44mM KH2PO4, 4,2mM NaHCO3, 1,2mM MgCl2, 20mM TES, 1mM CaCl2, 10mM glikoze, pH7,4) to the final volume of 150µl and concentration that exceeded the required three times. 100µl of Basal buffer were added to another nine tubes and a certain amount of ligand mix, that depended necessary step of dilution, from the first tube was transferred to the next one, the mix in the second microcentrifuge tube was thoroughly mixed and the same amount as before was transferred to next microcentrifuge tube, and so on until all ligand dilutions were prepared. 50µl of all dilutions were transferred to 96-well plate in two replicates and 50µl of pure Basal buffer were added to each well. For the negative and positive control 100µl of pure basal buffer and 100µl of 15µM forskolin in basal buffer respectively were added to separate wells in two replicates. The plate then was stored at 37ºC until needed further. H^3 adenosine labelled cells containing petri plate was removed from CO_2 incubator, media discarded, cells rinsed down with 5ml of Basal buffer, transferred to centrifuge tube and pelleted for 5min at 130×g. Supernantant was carefully discarded and cell pellet gently resuspended in 1100µl of Basal buffer with 0,5 mM IBMX. The Cell suspension was incubated at 37°C for 10min and then 50 µl of it were added to every well of previously prepared 96-well plate. After 10min incubation at 37°C cells were pelleted for 2min at 1260×g and supernantant was discarded. Finally 200µl of 0,33M percloric acid were added to every well, the plate was covered with parafilm and freezed.

2.13.2 Chromatography

Prior chromatography Dowex columns (Poly-Prep chromatography columns (BIO-RAD, 731-1550) filled with AG 50W-X8 Resin (BIO-RAD, 142-1451)) were equilibrated by applying 2×10ml of H_2O and Alumina columns (Poly-Prep chromatography columns (BIO-RAD, 731-1550) filled with Aluminium oxide (Sigma, A-9003)) were equilibrated by applying 8ml of 0,1M imodazole (Sigma, I-0125) solution. Previously prepared reaction plate was thawed and centrifuged for 10min at 1260×g. Supernatant from each well was transferred to separate Dowex column and as the control 750µl of 0,33M percloric acid - 650cpm/ml ^{14}C-cAMP (Amersham Pharmacia Biotech, CFA442) solution were added to each column. Then columns were eluted with 2ml of H_2O and the flow through was collected in scintillation vials that were supplemented with 4ml of scintillation cocktail Optiphase HiSafe 3 (PerkinElmer, 1200-437). Dowex columns were placed upon Alumina columns and eluted once more with 10ml H_2O (flow through was not collected). After sample transfer from Dowex, Alumina columns were eluted with 4ml of 0,1M imodazole solution. and the flow through was collected in scintillation vials that were supplemented with 7ml of Optiphase HiSafe 3scintillation cocktail. In order to estimate efficiency of each column pair the control containing 750 µl of 0,33M percloric acid - 650 cpm/ml ^{14}C-cAMP solution and 2ml scintillation cocktail Optiphase HiSafe 3 was added to three separate scintillation vials. Finally for all samples in all scintillation vials ^{3}H and ^{14}C emission was measured.

2.14 Binding assay

2.14.1 Preparation of membranes

Cells in culture containing petri plate were washed down using media, transferred to centrifugation tube and pelleted for 4min at 340×g. Supernatant was discarded, cell pellet resuspended in 200µl pf BB buffer (25mM HEPES pH7,4, 2,5mM $MgCl_2$, 1mM $CaCl_2$) and transferred to microcentrifuge tube. Then cell suspension was homogenized for 40s using Jankie&Kunkel IKA-WERK Ultra-Turrax Homogenizer and 450µl of BB buffer were added to homogenizate. Resulting solution was centrifuged for 3min at 160×g and supernatant was carefully transferred to clean microcentrifuge tube and repeatedly centrifuged for 20min at 20000×g. Supernatant was discarded, pellet was resuspended in 400µl of BB buffer and the membrane suspension was homogenized once more for 10s.

2.14.2 Binding and transfer to separation filters

For the first dilution in a microcentrifuge tube the ligand was diluted with the BBB buffer to the final volume of 100µl and concentration that exceeded the required four times. Choice of variable ligand depended on the type of assay non labelled ligands (NDP-MSH or ACTH) were used for competition binding assays while radioactively labelled ([125I]-NDP-MSH (Euro Diagnostica)) was used for total binding assays. 100µl of BB buffer were added to another 11 microcentrifuge tubes and a certain amount of ligand mix, which depended on necessary step of dilution, from the first microcentrifuge tube was transferred to the next one, the mix in the second microcentrifuge tube was thoroughly mixed and the same amount as before was transferred to next microcentrifuge tube, and so on until all ligand dilutions were prepared. The following mix was prepared in two replicates for each dilution: 25µl of diluted ligand, 25µl of 0,6nM [125I] NDP-MSH, for competition binding assay or BBB buffer for total binding assay and 50µl of membrane suspension. Incubation in room temperature for 3 h was terminated by transfer of binding mix to 0,3% polietilenamine saturated GF/C Filtermat A filter (Wallac Oy, Turku, Finland) using TOMTEC Mach III cell harvester (Orange, CT, USA). After transfer filter was dried for 5 min at 60°C, and MeltiLex A scintillation sheet (Perkin Elmer) was melted-on. Measurement of β-radiation in each position was carried out using Wallac 1450 automatic scintillation counter (Perkin Elmer).

2.15 Preparation of confocal microscopy samples

24h after transfection cell growth media was removed and coverslip was rinsed with 1ml 1×PBS (137mM NaCl, 2,7mM KCl, 8,1mM Na_2HPO_4, 1,5mM KH_2PO_4, pH7,4) buffer solution. 100µl of AlexaFluor® 633 labeled wheat germ agglutinin (AF-WGA) (Molecular Probes-Invitrogen)/PBS solution were placed upon the coverslip and incubated for 2min at 37ºC. After removal of AF-WGA/PBS solution coverslip was rinsed with 1ml 1×PBS buffer solution. Then 1ml of 4% formaldehyde (Sigma)/PBS solution was added and incubated for 10min at 37ºC. After removal of Formaldehyde/PBS solution the coverslip was rinsed three times with 1ml of 1×PBS buffer solution. Further 6µl of 1,5µg/ml DAPI (Sigma-Aldrich, 32670)/PBS solution were placed at the center of microscopy slide and coverslip was placed upon this droplet with the cells facing down.

3 RESULTS

3.1 The Melanocortin System in *Fugu*: Determination of POMC/AGRP/MCR Gene Repertoire and Synteny, As Well As Pharmacology and Anatomical Distribution of the MCRs

Contributions:

- Cloning of the Fugu ACTHR coding sequence;
- Expression of Fugu and human ACTHRs in M3 melanoma cells;
- Carrying out of cAMP assays on Fugu and human ACTHRs expressed in M3 melanoma cells;
- Contribution in a form of discussions to development of hypothetical model which is explaining the formation of MCRs encoding genes.

The Melanocortin System in *Fugu*: Determination of POMC/AGRP/MCR Gene Repertoire and Synteny, As Well As Pharmacology and Anatomical Distribution of the MCRs

Janis Klovins,† Tatjana Haitina,* Davids Fridmanis,*† Zuzana Kilianova,‡ Ivo Kapa,*† Robert Fredriksson,* Nicole Gallo-Payet,‡ and Helgi B. Schiöth**

*Department of Neuroscience, Uppsala University, Uppsala, Sweden; †Biomedical Research and Study Centre, University of Latvia, Riga, Latvia; ‡Service of Endocrinology, Department of Medicine, University of Sherbrooke, Sherbrooke, Quebec, Canada

The G-protein–coupled melanocortin receptors (MCRs) play an important role in a variety of essential functions such as the regulation of pigmentation, energy homeostasis, and steroid production. We performed a comprehensive characterization of the MC system in *Fugu* (*Takifugu rubripes*). We show that *Fugu* has an AGRP gene with high degree of conservation in the C-terminal region in addition to a POMC gene lacking γ-MSH. The *Fugu* genome contains single copies of four MCRs, whereas the MC3R is missing. The MC2R and MC5R are found in tandem and remarkably contain one and two introns, respectively. We suggest that these introns were inserted through a reverse splicing mechanism into the DRY motif that is widely conserved through GPCRs. We were able to assemble large blocks around the MCRs in *Fugu*, showing remarkable synteny with human chromosomes 16 and 18. Detailed pharmacological characterization showed that ACTH had surprisingly high affinity for the *Fugu* MC1R and MC4R, whereas α-MSH had lower affinity. We also showed that the MC2R gene in *Fugu* codes for an ACTH receptor, which did not respond to α-MSH. All the *Fugu* receptors were able to couple functionally to cAMP production in line with the mammalian orthologs. The anatomical characterization shows that the MC2R is expressed in the brain in addition to the head-kidney, whereas the MC4R and MC5R are found in both brain regions and peripheral tissues. This is the first comprehensive genomic and functional characterization of a GPCR family within the *Fugu* genome. The study shows that some parts of the MC system are highly conserved through vertebrate evolution, such as regions in POMC coding for ACTH, α-MSH, and β-MSH, the C-terminal region of AGRP, key binding units within the MC1R, MC2R, MC4R, and MC5R, synteny blocks around the MCRs, pharmacological properties of the MC2R, whereas other parts in the system are either missing, such as the MC3R and γ-MSH, or different as compared to mammals, such as the affinity of ACTH and MSH peptides to MC1R and MC4R and the anatomical expression pattern of the MCRs.

Introduction

The Japanese pufferfish (*Takifugu rubripes*), or *Fugu*, has become subject of increasing interest for diversity of biological research during the past few years, as the *Fugu* was the second vertebrate species whose entire genome was sequenced (Aparicio et al. 2002). Although there are approximately 400 Myr since the linage leading to teleosts split from the lineage leading to mammals (Carroll 1988), it is clear, that human and fish have large similarities in their genome and proteome organization. *Fugu* can thus serve as a good model organism for better understanding of the function and formation of the human genome. The genome of *Takifugu rubripes* is only 365 Mb (one eighth of that of human) and about 33,000 genes (Genscan) have been predicted so far (which is close to that estimated for human). Although rough analysis of the *Fugu* genome data indicates that 75% of the human proteome have a *Fugu* counterpart, more precise analysis is needed and information on functional characteristics of the protein families in *Fugu* is very limited (Aparicio et al. 2002).

The G-protein–coupled receptors (GPCRs) are one of the largest protein families in vertebrate genomes and are the single most pursued group of proteins in drug discovery. There exist over 800 GPCR genes in the human genome and about 40% to 50% of modern drugs are targeted at these receptors. Despite this and the mounting genetic information, only few GPCR families from lower vertebrates have been pharmacologically characterized. The melanocortin receptors (MCRs) are GPCRs and belong to the rhodopsin group (Fredriksson et al. 2003). MCRs respond to the pro-opiomelanocortin (POMC) cleavage products melanocyte-stimulating hormones (α-MSH, β-MSH, and γ-MSH) and adrenocorticotropic hormone (ACTH), all of them possessing agonistic properties on MCRs. The MC system is unique in the sense that it also has two endogenous antagonists named agouti (ASIP) and agouti-related peptide (AGRP). There are five subtypes of MCRs in mammals and aves named MC1R to MC5R (for review see Schiöth [2001]) (Gantz and Fong 2003). In mammals, MC1R is expressed in melanocytes and has a main role in determination of skin and hair pigmentation by regulation of the dark eumelanin synthesis, as a response to α-MSH (Rana et al. 1999). Binding of the antagonist ASIP, however, inactivates the signalling pathway and leads to synthesis of yellow phaeomelanin (Lu et al. 1994). Expression of the MC1R is also detected in other cell types of skin and in a number of peripheral tissues and cells (Chhajlani 1996), including leukocytes, where it mediates the broad anti-inflammatory actions, prompting interest from the pharmaceutical industry to find agonist at the MC1R. The MC2R is expressed in adrenal cortex, where it mediates the effects of ACTH on steroid secretion. The mammalian MC2R differs pharmacologically from

Key words: GPCR, ACTH, AGRP, MSH, melanocortin receptor.

E-mail: helgis@bmc.uu.se.

Mol. Biol. Evol. 21(3):563–579. 2004
DOI: 10.1093/molbev/msh050
Advance Access publication December 23, 2003
Molecular Biology and Evolution vol. 21 no. 3

other MCRs in that it is activated only by ACTH and has no affinity for MSH peptides (Schiöth et al. 1996). Some expression of MC2R has been found in human adipose tissue, but its role in this tissue is not clear. The MC3R and MC4R are expressed in several brain regions, particularly in the hypothalamus. These receptors have gained great attention during past years because of their involvement in regulation of energy homeostasis. The MC4R is one of the best-characterized monogenic factors of obesity. A number of mutations in this receptor lead to obese phenotypes in humans (Vaisse et al. 1998; Farooqi et al. 2000; Miraglia Del Giudice et al. 2002). Although both receptors are involved in regulation of the energy balance, mice deficient in one of these receptors display different phenotypes (Huszar et al. 1997, Chen et al. 2000). They also differ in pharmacology, as the MC3R has unique preference for γ-MSH among different subtypes of MCRs. MC5R is expressed in a number of human peripheral tissues, including adrenal gland, adipocytes, leukocytes, and others (Chhajlani 1996). The functional properties of MC5R are, however, still not well understood, with the exception of its participation in exocrine function, regulating sebaceous gland secretion in mice (Chen et al. 1997).

The MSH peptides have been intensively studied in lower vertebrates, and it seems that the POMC gene has arisen early in chordate evolution. Two copies of this gene are found in several ray-finned fish species (Okuta et al. 1996; Danielson et al. 1999) and lamprey (Takahashi et al. 1995). It is more likely, however, that the duplications in ray-finned fishes are late evolutionary event and do not represent an ancestry organization of this gene (Danielson et al. 1999). All fish POMC genes contain α-MSH, β-MSH, and ACTH. These sequences, especially that of α-MSH, are very conserved among different vertebrate species. Notable is that the γ-MSH region in ray-finned fishes is quite variable and degenerate, either missing the γ-MSH core motif (Amemiya et al. 1997; Dores et al. 1997) or even the complete γ-MSH sequence (Kitahara et al. 1988; Lee et al. 1999). AGRP has been found in chicken (Takeuchi, Teshigawara, and Takahashi 2000), but the evolutionary origin of AGRP or ASIP is obscure.

Our group has recently cloned the MC4R from zebrafish, goldfish, and dogfish and the MC5R from zebrafish (Ringholm et al. 2002, Cerda-Reverter et al. 2003, Ringholm et al. 2003), indicating high conservation in structure and pharmacology of these two receptors. We also found that MC4R is involved in central regulation of food intake in goldfish (Cerda-Reverter et al. 2003). No functional information is however available about the other MCR subtypes in nonmammalian species. Comparative pharmacological analysis of the entire MCR repertoire has not been previously performed beyond the mammalian lineage.

In this paper we report the identification of the full repertoire of *Fugu* MCR genes and their ligands. We describe genomic structure, synteny with the human genome, and determine the pharmacological profile and anatomical distribution of the receptors, including the first characterization of MC1R and MC2R from any nonmammalian species and delineate the vertebrate evolution of the MCR system.

Materials and Methods

Identification of MCR Genes

Two closely related pufferfish, *Takifugu rubripes* and *Tetraodon nigroviridis*, as well as zebrafish, *Danio rerio*, MCR and AGRP genes were identified from the whole-genome shotgun databases found at http://www.ncbi.nlm.nih.gov and other genomic Web pages: http://fugu.hgmp.mrc.ac.uk; http://genome.jgi-psf.org containing *Fugu* sequences, http://www.genoscope.cns.fr containing *Tetraodon* sequences, and http://www.sanger.ac.uk/Projects/D_rerio/ for zebrafish sequences. TBlastN searches were carried out on the databases using a number of MCR, AGRP, and ASIP sequences, including these of human, mouse, chicken, and known receptor sequences from fish species. BioEdit software (Hall 1999) was used for local Blast searches. The following contigs were identified and retrieved from databases. *Takifugu rubripes* scaffold numbers (only scaffolds from assembly release 3 are shown): MC1R (scaffold_431), MC2R, and MC5R (scaffold_1144), MC4R (scaffold_662), AGRP (scaffold_3097), and AGRP-like motifs A1 and A2 (scaffold_26029 and scaffold_305, respectively). *Tetraodon nigroviridis* contig numbers (Assembly version 6 at http://www.genoscope.cns.fr): MC1R (1071_4), MC2R (36902_1), MC4R (4178_2), MC5R (6538_1), and AGRP (227_2). *Danio rerio* contig numbers (Assembly 06/Assembly Zv1 at http://www.sanger.ac.uk/Projects/D_rerio/): MC1R (z06s08391), MC2R and MC5bR (z06s005606), and MC3R (z06s010438). To identify the full-length genes, contigs were then analyzed with Genscan (http://genes.mit.edu/GENSCAN.html) or with SpliceView (http://l25.itba.mi.cnr.it/~webgene/wwwspliceview.html) followed by manual correction.

Comparative Synteny Analysis Between *Fugu* and Human

Sequence from MCR and AGRP containing *Fugu* scaffolds were used in repeated BlastN search against different releases of genome shotgun sequence data sets. Identified scaffolds were downloaded and assembled to extend the sequence around the genes of interest. Sequences were assembled manually, based on results of pairwise Blast. Four contigs were obtained containing the following scaffolds (scaffolds starting with M represent Assembly release 3; scaffolds starting with S represent Assembly release 2): MC1R (M000431, S000194), MC2R and MC5R (M001144, M003114, M000533, S003468, S001069, S002629), MC4 (M000683, M000622, S000593, S000417, S002052), and AGRP (M 003097, S000686). Assembled contigs were subjected to Genscan analysis. Obtained protein sequences were individually compared with human genome assembly at NCBI using BlastP. Sequence similarity was evaluated based on quality and length of the match, including "score" and "Expect" values. Human genome map positions were identified for selected sequences from matching reads at http://www.ncbi.nlm.nih.gov/mapview/. Identified human sequences were blasted back to *Fugu* genome databases using TBlastN to ensure the correctness of similarity and

search for putative duplicate genes. Additional human genes within and from the surrounding of the established synteny regions were used to search against *Fugu* databases using TBlastN to identify *Fugu* genes not predicted by Genscan. This search, however, did not produce any results on scaffolds included in assembled contigs.

Cloning and Sequencing

The MCR sequences were amplified with Pfu Turbo DNA polymerase (Stratagene) from genomic DNA or cDNA. *Fugu* genomic DNA and a whole animal for tissue analysis were provided by Dr. Greg Elgar from *Fugu* Genomics Group (UK). PCR products were purified from 1% agarose gel using Gel Extraction Kit (Qiagen), followed by incubation with Taq polymerase (Invitrogen) in a reaction volume of 20 μl, containing 250 μM dATP, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 2 mM $MgCl_2$ and were cloned into pCRII vector and transformed into TOP10 cells (TOPO TA-Cloning Kit [Invitrogen]). Where possible, the entire coding sequences, including flanking regions, were cloned into TOPO cloning vector for sequencing. Sequencing reactions were performed using the ABI PRIZM Big Dye Terminator cycle sequencing kit according to the manufacture's recommendations and analyzed on ABI PRIZM-310 or 3100 Automated Sequencers (Applied Biosystems). Sequences were compiled and aligned using the Seqman program from the DNAstar package (Lasergene). Sequences were compared against database assemblies using BlastN and BlastX. For expression, full-length coding sequences were amplified by means of PCR from receptor gene containing TOPO plasmids with Pfu polymerase using *Hind*III and *Xho*I restriction sites containing primers for N- terminus and C-terminus, respectively. Obtained fragments were then digested with both restriction enzymes and gel purified before ligation into modified pCEP expression vector (Lundell et al. 2001). All constructs were sequenced to ensure identity of sequence to the original.

Alignments and Phylogenetic Analysis

Alignment of predicted full-length amino acid sequences for identified genes with other known MCR and ligand sequences were made using ClustalW version 1.8 software (Thompson, Higgins, and Gibson 1994). The following sequences (with their accession codes) were retrieved from GenBank for this analysis: *Homo sapiens* (Hsa) MC1R (NM_002386), MC2R (NM_000529), MC3R (XM_009545), MC4R (NM_005912), MC5R (XM_008685), AGRP (NM_001138), ASIP (NM_001672), *Mus musculus* (Mmu) AGRP (NM_007427), ASIP (NM_015770), POMC (NM_000939.1) *Gallus gallus* (Gga) MC1R (D78272), MC2R (AB009605), MC3R (AB017137), MC4R (AB012211), MC5R (AB012868), AGRP (AB029443), POMC (AB019555), *Xenopus laevis* (Xla), POMC (M11346), *Oncorhynchus mykiss* (Omy), POMC B (X69809), *Danio rerio* (Dre) MC1R (NM_180970), MC2R (NM_180971), MC3R (NM_180972), MC4R (AY078989), MC5aR (AY078990) and MC5bR (AY078991), POMC (NM_181438), and *Takifugu rubripes* (Tru) POMC (AAL11984). The identified genes have the following accession numbers: *Takifugu rubripes* MC1R (AY227791), MC2R (AY227793), MC4R (AY227794), MC5R (AY227796), and AGRP (BK001439); *Tetraodon nigroviridis* (Tni) MC1R (AY332238), MC2R (AY332239), MC4R (AY332240), and MC5R (AY332241). Phylogenetic trees were constructed by MEGA version 2.2. (Kumar et al. 2001) using maximum-parsimony and distance neighbor-joining methods. Human cannabinoid 2 receptor (hCB2 [S36750]) was used to root the receptor tree. The outgroup sequence that was used for AGRP/ASIP phylogeny was made artificially from alignment consensus sequence randomizing all positions except the fully conserved ones. Bootstrapping was performed with 1,000 random replicates.

Cell Culture and Transfection

HEK 293-EBNA cells were transfected with 2 to 5 μg of the constructs using FuGENE Transfection Reagent (Roche) according to the manufacturer's instruction. The cells were grown in Dulbecco's modified Eagle's medium (DMEM)/ Nut Mix F-12 with 10% fetal calf serum containing 0.2 mM L-glutamine, 250 μg/ml G-418, 100 U/ml penicillin, 100 μg/ml streptomycin and 2.5 μg/ml amphotericin B (Gibco BRL). Semistable cell lines were obtained for every construct expressing *Fugu* MCR by selecting cells for growth on medium containing 100 μg/ml hygromycin B (Invitrogen). Hygromycin was first added to medium 24 h after transfection. The MC2R was expressed in M3 cells. M3 cells were purchased by ATCC and were cultured in Kaighn's modification of Ham's F-12 medium (F-12K) supplemented with 15% horse serum (ATCC), 2.5% fetal bovin serum, and 1% GlutaMAX (Invitrogen). The cells were kept in humidified atmosphere of 95% air and 5% CO_2 at 37°C. One day before transfection, 200 000 cells were plated in 35-mm culture dishes. Transfections were carried out with 1.5 μg DNA using Lipofectamine Reagent PLUS (Invitrogen) according to the manufacturer's instructions. Transiently transfected cells were used 72 h after transfection.

Radioligand Binding

HEK 293-EBNA cells expressing *Fugu* receptors were harvested from plate and resuspended in binding buffer composed of 25 mM HEPES (pH 7.4) containing 2.5 mM $CaCl_2$, 1 mM $MgCl_2$, and 0.2 % bacitracin. To obtain the membrane, preparation cells were homogenized using UltraTurrax. The cell suspension was centrifuged for 3 min at 1,300 g and membranes were collected from the supernatant by centrifugation for 15 min at 14,000 g. The pellet was resuspended in binding buffer. The binding was performed in a final volume of 100 μl for 3 h at room temperature. Saturation experiments were carried out with serial dilutions of [^{125}I] (Nle4,D-Phe7) α-MSH (NDP-MSH). Nonspecific binding was determined in presence of 1 μM unlabeled NDP-MSH. Competition experiments were performed with constant 0.6 nM concentration of [^{125}I]NDP-MSH and serial dilutions of competing unlabeled ligands: NDP-MSH, α-MSH, β-MSH, γ-MSH,

ACTH(1-24), MTII, and HS024 (Neosystem). The membranes were collected by filtration on glass fibre filters, Filtermat A (Wallac), using a TOMTEC Mach III cell harvester (Orange, Conn.). The filters were washed with 5 ml per well of 50 mM TrisHCl (pH 7.4) and dried. MeltiLex A scintillator sheets (Wallac) were melted on dried filters and radioactivity was counted with automatic Microbeta counter 1450 (Wallac). Binding assay was performed in duplicates from at least three independent experiments. Nontransfected cells did not show any specific binding with [^{125}I]NDP-MSH. The results were analysed with the Prism version 3.0 software package (Graphpad).

cAMP Assay in HEK 293 EBNA Cells

Cyclic adenosine 3′:5′–cyclic monophosphate production was determined on semistable HEK 293 EBNA cells expressing appropriate the MCR. A confluent layer of semistable cells was incubated for 3 h with 2.5 μCi of [^{3}H]ATP (specific activity 29 Ci/mmol [Amersham]) per ml of medium. Cells were collected, washed with PBS, and incubated for 10 min in PBS containing 0.5 mM isobutylmethylxanthine (IBMX) (Sigma). Stimulation reaction was performed for 10 min in a final volume of 150 μl of PBS supplemented with 1 mM IBMX containing approximately 2×10^5 cells and various concentrations of α-MSH peptide. After incubation, cells were centrifuged and 200 μl of 0.33 M perchloric acid (PCA) was added to pellets to lyse the cells. After centrifugation, 200 μl of lysate was added to Dowex 50W-X4 resin (Bio-Rad) column (Bio-Rad, poly-prep columns), previously washed with 2×10 ml H_2O. As an internal standard, 750 μl 0.33 M PCA containing 0.5 nCi/ml [^{14}C]cAMP (Amersham) was added to column. Columns were washed with 2 ml H_2O to remove ATP, which was collected in scintillation vials to estimate the amount of unconverted [^{3}H]ATP. Four milliliters of Ready Safe scintillation cocktail (PerkinElmer) was added to the vials before counting. Dowex columns were then placed over aluminia (Sigma) columns (prewashed with 8 ml 0.1 M imidazole), and the cAMP was transferred onto the aluminia column using 10 ml H_2O. cAMP was eluted from aluminia column using 4 ml 0.1 M imidazole and collected into scintillation vials to which 7 ml of scintillation fluid was added. ^{3}H and ^{14}C were counted on Tri-carb liquid scintillation beta counter. The amount of obtained [^{14}C]cAMP was expressed as a fraction of total [^{14}C]cAMP ([^{14}C]cAMP/total [^{14}C]cAMP) and was used to standardize [^{3}H]cAMP to column efficiency. Results were calculated as the percent of total [^{3}H]ATP (obtained as a sum of [^{3}H]ATP from first column and [^{3}H]cAMP from second column) to [^{3}H]cAMP and used to determine EC_{50} values by nonlinear regression using the Prism version 3.0 software. All experiments were made in duplicates and repeated three times.

cAMP Assay in M3 Cells

Intracellular cAMP production induced by MC2R expressed in M3 cells was measured by sequential chromatography on Dowex and alumina columns as described for HEK 293 EBNA cells with following modifications (Gallo-Payet and Payet 1989). Seventy-two hours after transfection, MC2R transiently transfected M3 cells plated on 35-mm culture dishes were incubated 1 h at 37°C with complete culture medium containing 2 μCi/ml [^{3}H]-adenine (NEN). The cells were then washed twice with Hank's buffer saline and equilibrated in the same buffer containing 1 mM isobutyl methylxanthine (IBMX) for 15 min at 37°C. ACTH was added to the incubation medium for further 15 min at 37°C. The reaction was ended by aspiration and adding 1 ml ice-cold 5% (w/v) trichloroacetic acid. Cells were harvested, and 100 μM ice-cold 5 mM ATP plus 5 mM cAMP solution was added to the mixture. Cellular membranes were pelleted at 5,000 g for 15 min at 4°C and the supernatants sequentially chromatographed on Dowex and alumina columns according to the method of Salomon, Londos, and Rodebell (1974), allowing the elution of [^{3}H]-ATP and [^{3}H]-cAMP, respectively. cAMP formation was calculated as percent conversion = [[^{3}H]-cAMP/ ([3H]-cAMP + [3H]-ATP)] × 100 and expressed as fold stimulation over basal cAMP level.

RT-PCR and Southern Analysis

The total RNA was isolated from number of peripheral tissues (head-kidney, eye, skin, muscle, intestine, and gut) and several brain regions (telencephalon, optic tectum, hypothalamus, cerebellum, and brain stem). Tissues were dissected from frozen animal and kept in RNAlater (Ambion) before RNA extraction. Tissues were homogenized and total RNA was isolated using Rneasy Mini Kit (Qiagen), including processing with DNA shredder and DNase treatment (Qiagen) as recommended by manufacture's protocol. Because some of the samples retained genomic DNA, total RNA was exposed for another treatment with 1U/μl RNase-free DnaseI (Roche) for 10 min, followed by heat inactivation of DNase for 5 min at 70°C. Absence of genomic DNA in RNA preparations was confirmed in PCR reactions with primers specific to β-actin gene using 10 to 100 ng of total RNA as a template and genomic DNA as a positive control. Messenger RNA was reverse transcribed using the First Strand cDNA Synthesis kit (Amersham). The produced cDNA was used as a template for PCR, with the specific primers for the receptor genes (see below). The quantity of cDNA was roughly estimated with PCR reactions on β-actin gene by ethidium bromide-stained agarose gel inspection of PCR product taken after cycles 20, 25, 30, 35, and 40 of PCR reaction. The conditions for the PCR on receptor genes were 1 min initial denaturation, 20 s at 95°C, 30 s at appropriate annealing temperature, 60 s at 72°C for 30 or 40 cycles, and finished by 5 min at 72°C, using Taq polymerase (Invitrogen). The following primers were used: 5′-TCT CCT TCT GCA TCC TGT CG-3′ and 5′-CAT CAA GGA GAA GCT GTG CT-3′ for β-actin gene (expected size 317 bp); 5′-TCA CGG GCC AAA GCA CCA GG-3′ and 5′-TCA TCG TCT ACC ACA CTG AT-3′ for MC1R gene (expected size 430 bp); 5′-TTA GCG CCA CCT CCA GTT TA-3′ and 5′-ACG TGG TGG ACT CGC TGC-3′ for MC2R gene (expected size

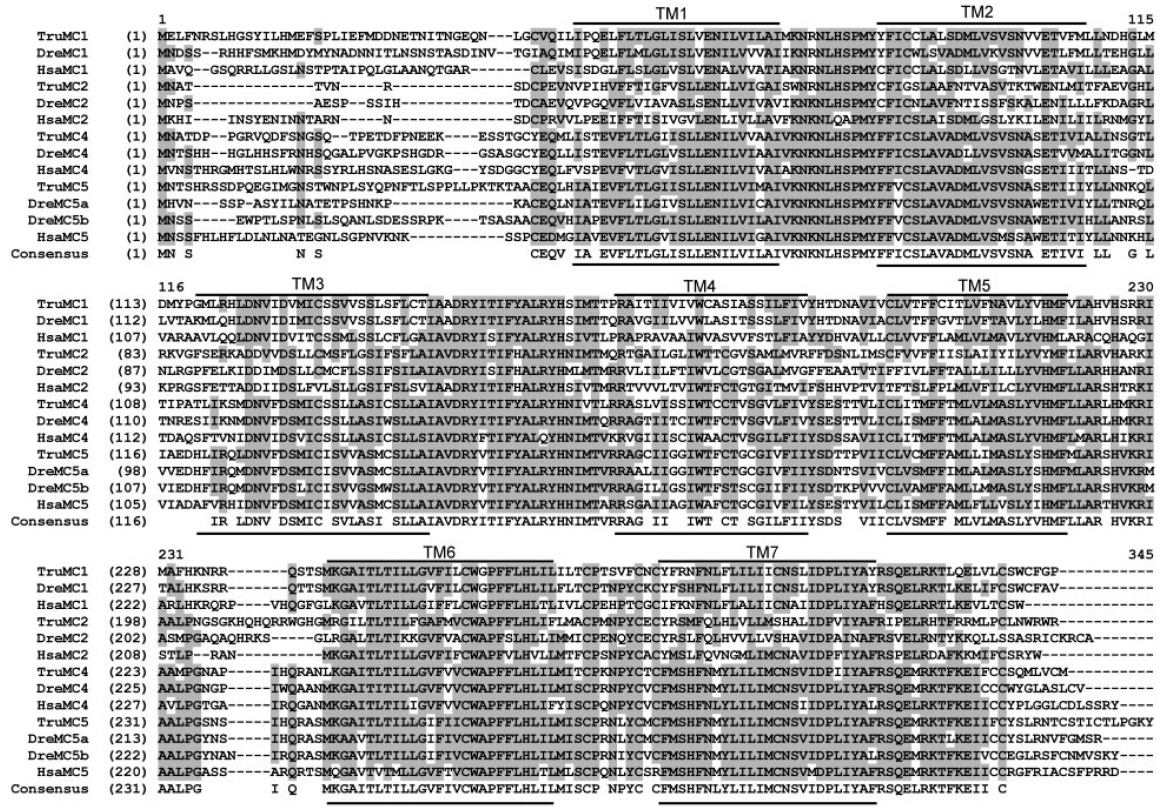


FIG. 1.—Amino acid sequence alignment of MCRs made using ClustalW version 1.8. Putative transmembrane (TM) regions are marked with lines. Gray boxes show conserved amino acid positions. Abbreviations used are human (Hsa), zebrafish (Dre), and *Fugu* (Tru). The accession numbers and Latin names are listed in *Materials and Methods*.

for cDNA 626 bp); 5′-TCA CAT GCA CAC CAG CAT TT-3′ and 5′-TGC TGG CGC GCC TGC ACA TG-3′ for MC4R gene (expected size 332 bp); 5′-TCA GTA CTT ACC TGG AAG AG-3′ and 5′-ACC TCC TTA ACA ACA AGC AGC-3′ for MC5R gene (expected size of cDNA 709 bp). The PCR products were analyzed on 1% agarose gel. DNA from the gel was transferred onto nylon filters overnight, using 0.4 M NaOH. The filters were hybridized with a random-primed ^{32}P-labeled, receptor-specific probe. Probes were generated by Megaprime kit (Amersham Biosciences) using sequence-verified PCR products amplified from plasmids containing *Fugu* MCR genes. Hybridization was carried out at 65°C in 25% formamide, 6 × SSC, 10% dextran sulfate, 5× Denhardt's solution, and 0.1% SDS overnight. The filters were washed three times in 0.2 × SSC, 0.1% SDS for 1 h at 65°C and exposed to autoradiography film (Amersham). As positive controls in the Southern blots, the PCR products obtained from genomic DNA was used. A number of MC2R and MC5R RT-PCR bands were cut out from agarose gel, purified, and sequenced to ensure their identity and correct splicing, because these RT-PCR products were of different size as compared with genomic DNA as a result of the presence of introns in these genes. The RT-PCR reactions with 30 and 40 cycles and respective Southern blots were performed at least two times each.

Results

Gene Identification and Sequence Analysis

Melanocortin Receptors

The shotgun sequence databases of two closely related pufferfishes, *Fugu* (*Takifugu rubripes)* and *Tetraodon* (*Tetraodon nigroviridis*), as well as zebrafish (*Danio rerio*) were screened using Blast search to identify orthologs of the MCR gene family. This search revealed MC1R, MC2R, MC4R, and MC5R gene sequences from *Takifugu rubripes* databases and corresponding sequences from *Tetraodon nigroviridis* database (see *Materials and Methods*). In addition to previously known MC4R, MC5aR, and MC5bR genes obtained by our group (Ringholm et al. 2002), MC1R, MC2R, and MC3R gene sequences were found in zebrafish databases. Apparently, *Fugu* lacks a gene for the MC3R. The new genes were designated according to their sequence similarity to known MCRs from human and other species and phylogeny (see accession numbers in *Materials and Methods*). Alignment of the protein sequences of the receptors with the previously cloned human, chicken, and zebrafish receptors is shown in figure 1. Maximum-parsimony analysis of this group of receptors is shown in figure 2, representing consensus tree, using human cannabinoid 2 (hCB2) receptor as an outgroup. Both neighbor-joining and maximum-parsimony analysis

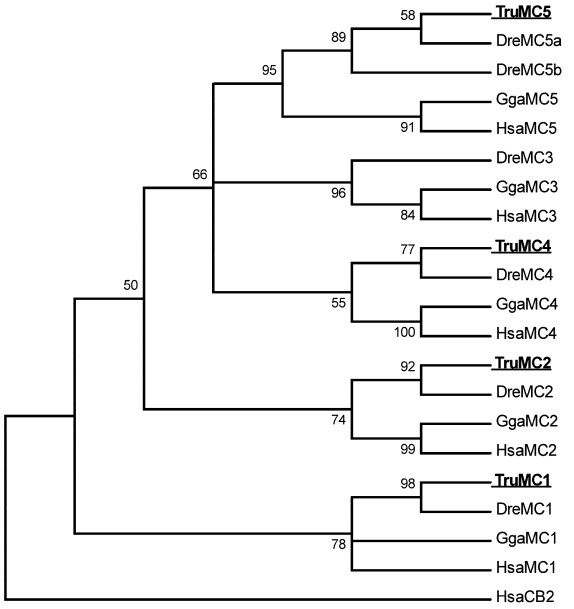


FIG. 2.—Phylogenetic analysis of MCRs using full-length amino acid sequences. The consensus tree was generated by maximum-parsimony analysis (MEGA version 2.0) The human cannabinoid 2 receptor (hCB2) sequence was used to root the tree. The numbers above nodes indicate percentage of bootstrap replicates. The abbreviations used are human (Hsa), chicken (Gga), zebrafish (Dre), and *Fugu* (Tru).

resulted in the same topology of the trees. It is interesting to note that the MC5R in *Fugu* is an ortholog of the MC5a in zebrafish. The sequences of identified *Fugu* MCR genes and flanking regions were confirmed by sequencing of genomic PCR products. Some minor differences were found in MC1R receptor gene sequence as compared with the released genome sequence data. When analyzing the gene structure of the *Fugu* MCR, we surprisingly found that *Fugu* contains one intron in MC2R (position 353 from CDS start) gene and three introns in MC5R gene (positions 156, 263, and 452). None of the previous available sequences of MCRs contains introns (with an exception of human MC1R, where extension of coding sequence has been detected in mRNA because of an alternative splicing [Tan et al. 1999]). *Tetraodon* MC2R and MC5R genes had the same genomic structure, indicating that introns arose before the split of *Fugu* and *Tetraodon* in the pufferfish lineage, which is roughly estimated to be 20 MYA (Crnogorac-Jurcevic et al. 1997). The coding sequences of the ortholog MCR genes are highly conserved between *Fugu* and *Tetraodon* (84% to 93% similarity). Moreover, there is remarkable similarity in both length and sequence of the introns in the MC2R gene (60% similarity). The introns of MC5R gene have, however, diverged considerably in length, but still retain sequence similarity. To test the correctness of predicted splice products for *Fugu* MC2R and MC5R genes, they were verified by sequencing of RT-PCR products obtained using total RNA from *Fugu* brain. We were able to obtain and sequence full-length gene cDNA sequence for MC5R, whereas in case of the MC2R, RT-PCR products lacked 200 bp from 5′ end. In both cases, the splicing occurred at the predicted positions in these brain cDNAs.

POMC

We also searched for the ligand peptides to obtain a better picture on evolution of the MC system. Figure 3*A* displays the alignment of POMC genes from several relevant species. It should be noted that the overall similarity for POMC gene sequence is quite low or about 40%, but the parts coding for specific active peptides have remarkable high homology. Comparing MSH peptides from human and *Fugu*, we can see that α-MSH differs only in one position (98% similarity), ACTH peptides have 72% similarity, being almost identical (differs in one position) for ACTH(1-24), and β-MSH has 81% similarity. As observed for other teleosts (Takahashi et al. 2001), the γ-MSH region is not present in Fugu POMC.

AGRP

We have also searched for AGRP/ASIP-related genes. Although we were not able to identify such genes in zebrafish, the *Fugu* search revealed three cystein-rich motifs with high similarity to the C-terminal part of AGRP/ASIP family. Figure 3*B* shows amino acid alignment of these sequences compared with peptides from other vertebrate species. We were able to identify and confirm the full-length structure of one such gene. Phylogenetic analysis placed this gene with the AGRP peptides, and the gene was accordingly named *Fugu* AGRP. For the other two motifs (A1 and A2), full-length genes were not identified because of incompleteness of sequence data in one case and unconformable Genscan predictions in the other case. The difficulties in analysing the N-terminal part are related to the lack of sequence homology for this part of the peptide. Figure 3*C* represents the results of phylogenetic analysis using maximum-parsimony method. We also found that *Fugu* AGRP gene structure resembles that of other vertebrates as seen in figure 3*B* and *D*, indicating conservative intron positions, even though the length of the introns is variable. The *Fugu*, however, has one additional 5′ intron in the AGRP sequence.

Comparative Synteny Analysis

To gain more information about organization and synteny of the genome regions in *Fugu* containing MCR genes, we assembled from 170 kb to 400 kb large contigs using sequence data from different releases of *Fugu* genome consortium. First we found that MC2R and MC5R are located in close proximity in both *Fugu* (separated by 2.6 kb) and zebrafish (6.1 kb) (see fig. 4*B*). In both cases, the MC2R and MC5R genes face each other in sense of direction of transcription. This localization and orientation of the two genes appears to be very conserved among vertebrates. Thus, the human MC2R and MC5R orthologs located on chromosome 18 (18p11.2) are separated by 57.8 kb, whereas the mouse has 67.7 kb "inserted" between these genes. The MC2R and MC5R loci have also been mapped on the same region of chicken chromosome 2 (Schiöth et al. 2003).

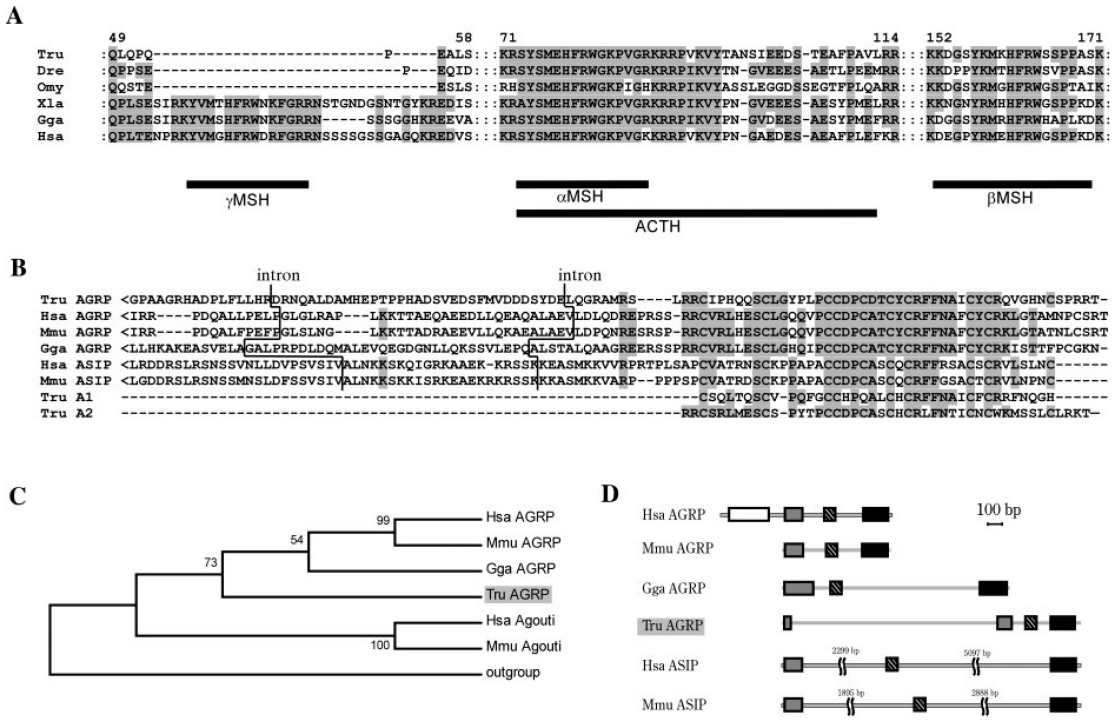


FIG. 3.—Alignment of the amino acid sequences of the pro-opiomelanocortin (POMC) propeptide (*A*) and the C-terminal part of the agouti-signalling peptide (ASIP)/agouti-related peptide (AGRP) (*B*) from different vertebrate species. The gray boxes show conserved amino acid positions. The sequences corresponding to the MSH peptides are marked with black boxes. The intron positions in the AGRP/ASIP alignment are shown by vertical lines. Two cystein-rich motifs were found in *Fugu* and named Tru A1 and Tru A2 and added to the AGRP alignment. The phylogenetic analysis of the C-terminal part of AGRP/ASIP is shown in (*C*). The tree was generated by maximum-parsimony analysis (MEGA 2.0). The numbers above the nodes indicate the percentage of the bootstrap replicates. (*D*) The genomic structure of the exon-introns of the full-length genes in this study. The abbreviations used are human (Hsa), mouse (Mmu), chicken (Gga), frog (Xla), rainbow trout (Omy), zebrafish (Dre), and *Fugu* (Tru).

To estimate the level of synteny shared between the human and *Fugu* genomes, we compared the content and order of genes from assembled *Fugu* contigs with available human genome maps at NCBI map search. Figure 4*A* shows the result of this analysis. Genes identified on *Fugu* contigs were named according to the orthologous human gene names. Where official names for human genes did not exist, we used interim names from UniGene cluster or gene model names from LocusLink. In total 20 genes were identified to share synteny with corresponding human genome regions. Thus nine genes from "MC1R contig" have their orthologs on two close loci of HSA16q with the same gene order. The segment of three genes including MC1R, however has been inverted either in humans or *Fugu*. "MC2R and MC5R contig" share eight genes with the HSA18p in locus containing both the MC2R and MC5R genes. The order and colinearity of these genes seems not to be disrupted between the two species. Finally three genes from "MC4R contig," including MC4R, have orthologs on HSA18q. This contig in contrast to the previous two has 10 genes mapping to other human chromosomes. These genes, with one exception, are organized in three clusters, which correspond to HSA5p, HSA 14q11.2, and HSA3q, sharing the order of genes with their human counterparts. Apparently there have been more translocations in this region, and genes were either inserted into this locus in *Fugu* lineage or moved from this region in human. Because it is believed that *Fugu* has undergone additional genome duplication early in teleosts lineage, we searched for copies of the identified genes elsewhere in *Fugu* genome. However, this search did not produce any positive results, with exception of genes from families with many copies and high sequence similarity (e.g., cadherin [CDH] gene family).

Pharmacological Properties of *Fugu* MCRs

For pharmacological testing of the *Fugu* MCR repertoire, the coding regions of the receptors were cloned into an expression vector containing cytomegalovirus (CMV) promoter. All receptors were expressed in HEK-293 EBNA cells except for the MC2R that was tested separately in M3 cells (see below). Semistable cell lines were used for radioligand binding and cAMP assays. The endogenous peptides we used were of the human origin. α-MSH is fully conserved between the species, but there are minor differences in the sequences of β-MSH and ACTH. These differences are not within the predicted core-binding region of the peptides, but it is not known if these may affect the pharmacology.

Radioligand Binding

We performed saturation analysis using high-affinity ligand [^{125}I] NDP-αMSH and competition-binding analysis using NDP-MSH, α-MSH, β-MSH, γ-MSH,

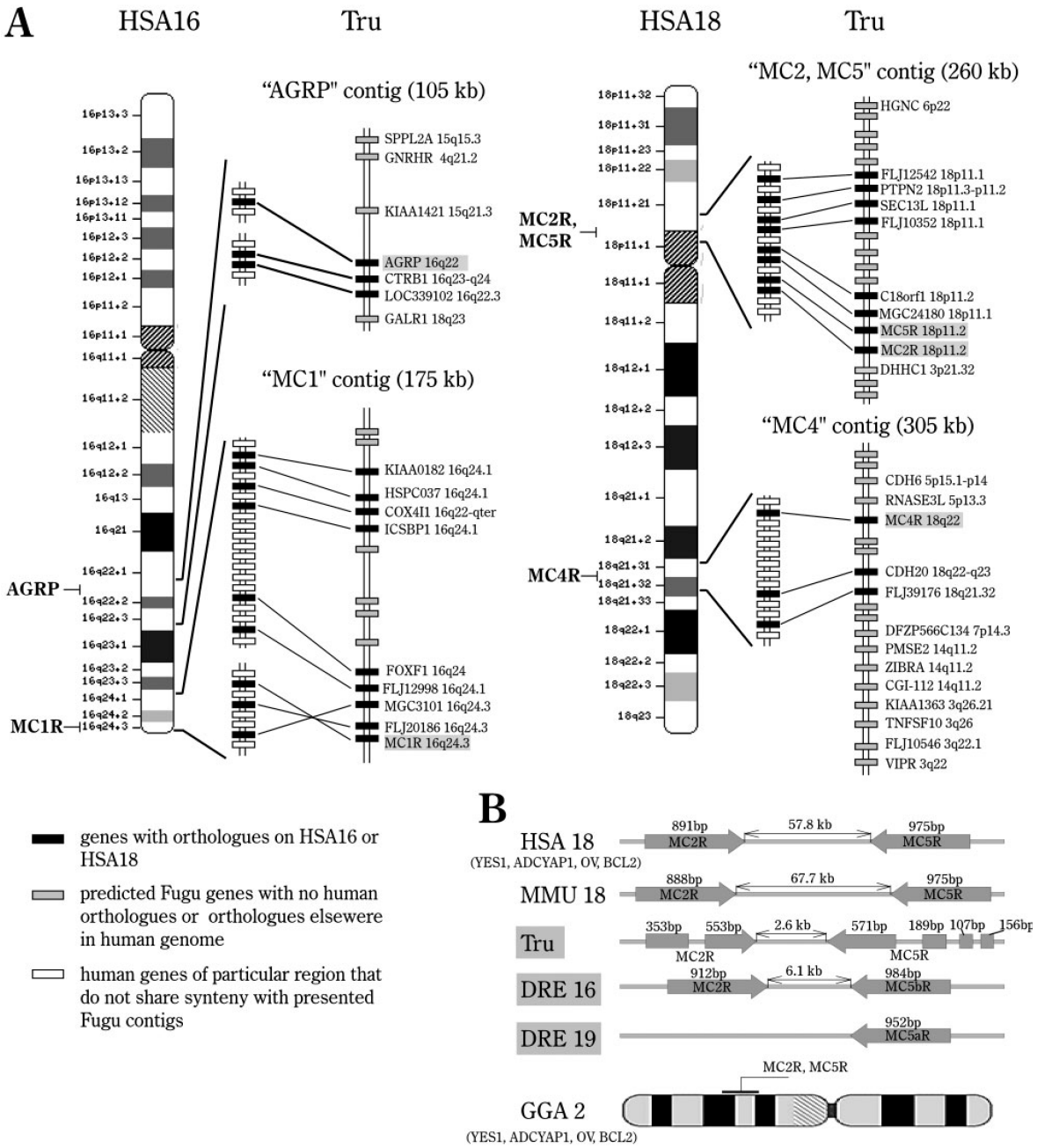


FIG. 4.—(*A*) The *Fugu*, HSA16, and HSA18 comparative gene maps. The genes were named according to the human orthologs. The syntenic genes are displayed as black boxes and linked by lines. The *Fugu* genes with orthologs elsewhere in the human genome are denoted by gray boxes. The gray boxes with no name indicate the genes whose human ortholog could not be identified. The human genes that do not share synteny with the *Fugu* contigs are indicated with white boxes. (*B*) The gene linkage for MC2R and MC5R in several vertebrate species. The distances between the stop codons of the MC2R and MC5R are obtained from the genome assembly data. The direction of the transcripts is shown with arrows. The chicken MC2R and MC5R genes were mapped on the same locus of chromosome 2 (Schiöth et al. 2003). Other genes that share synteny between the species are listed in parenthesis below the species name. The abbreviations used are human (HSA), zebrafish, (DRE), *Fugu (*Tru), mouse (MMU), and chicken (GGA).

ACTH(1–24), MTII, and HS024 as competitors. MTII and HS024 are synthetic ligands widely used for physiological studies. The results of binding experiments are shown in figure 5. Table 1 shows Kd and Ki values obtained from saturation and competition experiments, respectively, for the MSH-binding *Fugu* receptors. Table 1 also includes previously published results for the human MCRs for comparison, tested with the same methodological approach.

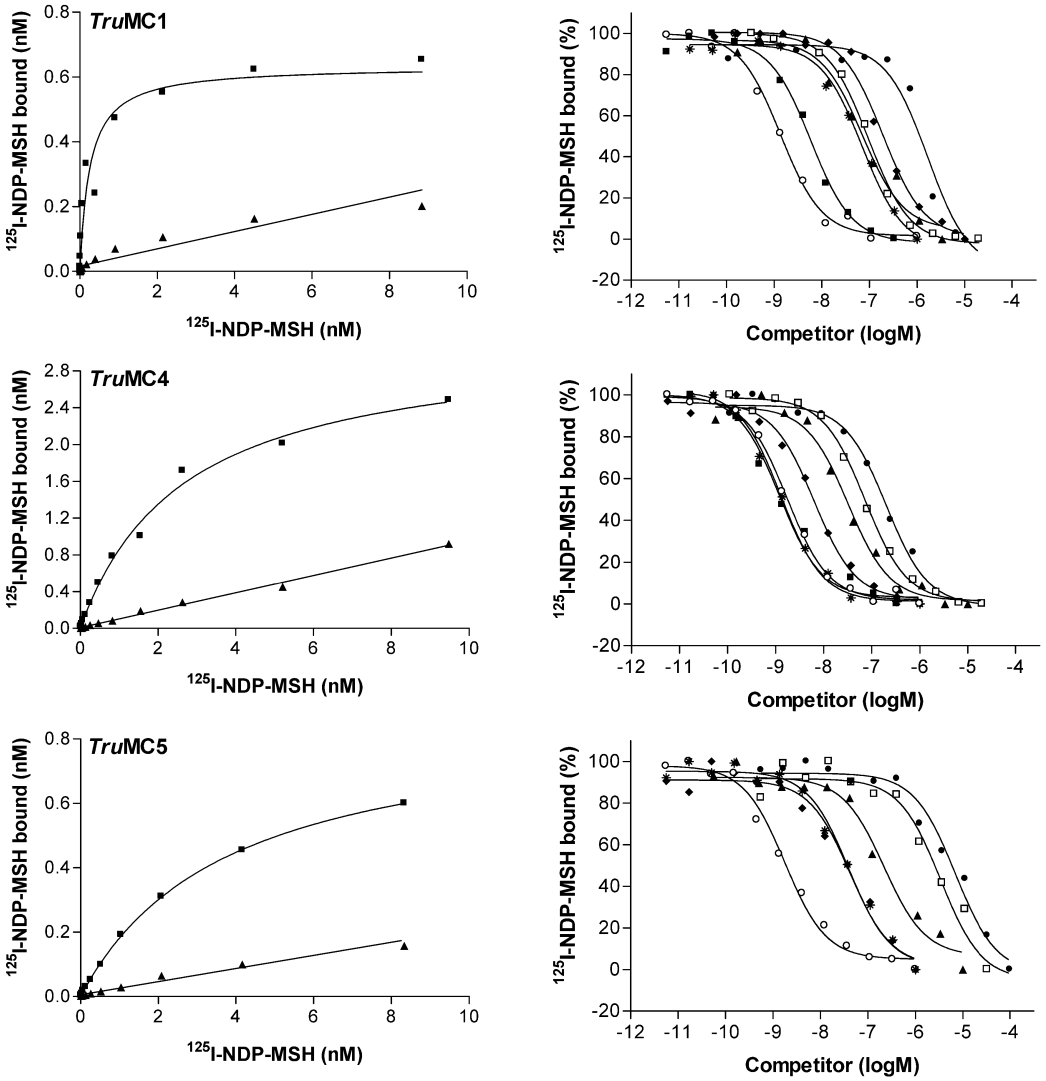


FIG. 5.—Saturation and competition curves for the *Fugu* MC1R, MC4R, and MC5R expressed in intact transfected HEK-293 cells. The saturation curves (left) were obtained with ^{125}I-labeled NDP-MSH, and the figure shows total binding (filled square) and binding in the presence of 2 μM cold NDP-MSH (filled triangle). The lines represent the computer-modeled best fit of the data, assuming that ligands bound to one site. The competition curves (right) for NDP-MSH (open circle), α-MSH (filled triangle), β-MSH (open square), γ-MSH (filled circle), ACTH (filled square), MTII (star), and HS024 (filled rhomb) were obtained by using a fixed concentration of ~0.6 nM ^{125}I-labeled NDP-MSH and varying concentrations of the non-labeled competing peptide.

It should be mentioned that these human MCR binding values have been tested repeatedly for over a decade with very consistent results. The results show that NDP-MSH binds to a single saturable site for all *Fugu* receptors, except for the MC2R that was tested separately in M3 cells (see below). The affinity for labeled NDP-MSH was high, indicating that this radioligand is appropriate for characterization of the *Fugu* MCRs, except the MC2R. The affinity of NDP-MSH was similar for both the labeled (saturation) and cold (competition) experiments and also similar between the *Fugu* and human orthologs. The endogenous ligands α-MSH, β-MSH, and γ-MSH had clearly lower affinity (64-fold, 7-fold, and 94-fold, respectively) for the *Fugu* MC1R than the human MC1R. The ACTH(1–24), which is generally considered to be equipotent to the full-length ACTH(1–39), had however only 4.7-fold higher affinity for the *Fugu* MC1R as compared with the human MC1R. The synthetic ligand MTII had about 10-fold lower affinity for the *Fugu* MC1R, whereas HS024 had similar affinity for the *Fugu* and human orthologs. In contrast to the *Fugu* MC1R, the *Fugu* MC4R had 23-fold, sevenfold, and 277-fold higher affinity for α-MSH, β-MSH, and γ-MSH, respectively, as compared with the human MC4R. MTII had higher affinity, whereas HS024 had lower affinity for the *Fugu* MC4R as compared with the human MC4R. In similar

Table 1
K_d and K_i Values (Mean ± SEM) Obtained from Saturation and Competition Curves, Respectively, for Melanocortin Peptide Analogs on MCRs Transfected HEK 293 and COS Cells

Ligand	*Tru* MC1	*Hsa* MC1[a]	*Hsa* MC3[e]	*Tru* MC4	*Hsa* MC4[e]	*Tru* MC5	*Hsa* MC5[b]
[^{125}I] NDP-MSH (K_d)	0.116 ± 0.003	0.109 ± 0.062	0.412 ± 0.121	1.52 ± 0.01	1.78 ± 0.36	1.47 ± 0.45	5.05 ± 1.00
NDP-MSH (K_i)	0.266 ± 0.097	0.046 ± 0.011	0.319 ± 0.064	1.24 ± 0.02	1.96 ± 0.39	1.65 ± 0.42	2.39 ± 0.10
α-MSH (K_i)	13.4 ± 2.4	0.210 ± 0.089	21.2 ± 5.3	22.8 ± 0.3	522 ± 122	120 ± 36	8240 ± 1670
β-MSH (K_i)	17.1 ± 0.8	2.53 ± 0.93	15.1 ± 3.4	52.7 ± 0.6	387 ± 208	1790 ± 470	14400 ± 1670
γ_1-MSH (K_i)	253 ± 14	2.68 ± 0.35[b]	7.45 ± 2.55	187 ± 45	51800 ± 12000	3440 ± 670	42600 ± 6600
ACTH (1–24) (K_i)	0.983 ± 0.194	0.209 ± 0.052[c]	32.8 ± 6.7	1.14 ± 0.15	755 ± 151	ND	2760 ± 780[c]
MTII (K_i)	7.62 ± 1.12	0.686 ± 0.109[d]	52.6 ± 7.9	1.00 ± 0.14	6.60 ± 0.82[f]	24.4 ± 7.4	46.1 ± 7.9[d]
HS024 (K_i)	22.4 ± 1.1	18.6 ± 3.3[d]	15.1 ± 3.0	5.36 ± 0.64	0.341 ± 0.089	27.4 ± 8.1	3.29 ± 1.15[d]

NOTE.—Values are nmol × L^{-1}. ND = not determined.
[a] Schiöth et al. 1997*c*.
[b] Schiöth et al. 1995.
[c] Schiöth et al. 1997*a*.
[d] Kask et al. 1998.
[e] Schiöth et al. 2002.
[f] Schiöth et al. 1997*b*.

fashion as the MC4R, the *Fugu* MC5R had higher affinity for the natural peptides (α-MSH [68-fold], β-MSH [8-fold], and γ-MSH [12-fold]) than the human MC5R. MTII had also higher affinity for the *Fugu* MC5R, whereas HS024 had lower affinity than the human MC4R.

Cyclic AMP Assay

We also tested cells expressing the *Fugu* receptors in a cAMP assay to determine the ability of these receptors to couple to G-proteins and induce accumulation of cAMP upon exposure to natural ligands. The results are shown in figure 6. All the MCRs, except the MC2R, reached maximal levels of response, when stimulated with α-MSH. As α-MSH had very low potency to the *Fugu* MC1R, we also tested ACTH(1–24) at this receptor. The *Fugu* MC1R showed clearly higher potency for ACTH(1–24) α-MSH than in agreement with the binding data (table 1). In contrast to the MC1R, the MC4R and MC5R responded with high potency to α-MSH.

There are general difficulties of expressing MC2R in common cell lines as observed by many researchers because of failed transport of MC2R to the membrane (Noon et al. 2002). Expression of the MC2R has only been achieved in some adrenocortical and melanoma cell lines (Schimmer et al. 1995; Penhoat et al. 2000). We therefore used M3 melanoma cell line to functionally express *Fugu* MC2R and tested it in cAMP assay. These cells possess some basal activity in response to ACTH because of endogenous MC1R. However, expression of human MC2R has proved to be successful, showing a response to ACTH at much lower doses and with higher maximum response than the endogenous MC1R response. Our results in figure 7 show the response of ACTH in M3 cells transfected with the human MC2R, which is in good agreement with previous data (unpublished data). Moreover, we also received reproducible ACTH dose-dependent stimulation of at least twofold higher cAMP production in M3 cells transiently transfected with the *Fugu* MC2R. The results from the *Fugu* MC2R, although it seems to have slightly lower affinity ($EC_{50} = 9.7 \times 10^{-9}$ M) are comparable with human MC2R ($EC_{50} = 2.0 \times 10^{-9}$ M). No activity was observed in the same cells stimulated with NDP-MSH.

Tissue Distribution of *Fugu* MCRs

The tissue distribution of the four *Fugu* MCRs was determined by RT-PCR. Total RNA from a number of tissues, including different brain regions, was isolated from an adult animal and cDNA was obtained. We tested the cDNA by PCR using β-actin primers and confirmed the integrity of the mRNA. We determined the approximate amounts of mRNA in preparations by estimating the kinetics of the PCR reaction. Results of these experiments on actin gene are shown in the figure in Supplementary Material online. Even though the expression levels of β-actin varied slightly among the different types of tissues, this experiment gave a rough estimation of quality and amounts of RNA and helped to interpret the results of receptor-specific RT-PCR in more quantitative terms. For example, because of elastic structure of the *Fugu* skin, we were not able to extract RNA of sufficient amount and quality. Even though β-actin RT-PCR product is observed after cycle 40 on cDNA of this tissue, the kinetics of PCR reaction indicates that it would not be possible to detect genes whose expression levels are lower than β-actin. On the other hand, we see that β-actin mRNA in brain is highly expressed compared with other tissues. This fact must be considered when analyzing the expression of the MCRs, because presence of transcripts in tissues with highly expressed mRNA may be caused by unspecific transcription events and can be detectable in RT-PCR after high number of cycles. We therefore used different cycle numbers for RT- PCR to better estimate the quantity and specificity of the expression.

The results of the RT-PCR for each of the receptor gene are shown in figure 8. Each pair of PCR primers was designed to be specific for one receptor subtype. Specificity of the PCR reactions was estimated by Southern blot with the *Fugu* MCR subtype-specific probes. No cross-hybridization was detected in our samples (data not shown). To get an impression of the expression levels for *Fugu* MCRs, we

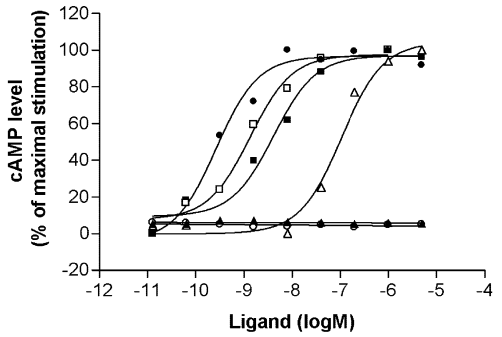


FIG. 6.—Generation of cAMP in response to α-MSH for the MC1R (open triangle), MC4R (filled square), and MC5R (open square) and to ACTH for MC1R (filled circle) in intact transfected HEK-293 cells. Untransfected HEK-293 cells showed no adenylate cyclase activity in response to α-MSH (filled triangle) or ACTH (open circle). The cAMP assay was performed in duplicates and repeated three times for each receptor.

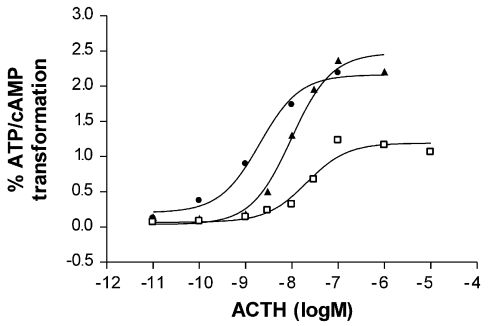


FIG. 7.—ACTH dose-response curve in native (open square), human MC2R (filled circle), and *Fugu* MC2R (filled triangle) transfected M3 cells. Because M3 cells endogenously express the MC1R at low levels, the basal level of cAMP production was estimated with every experiment. Accumulated [^{3}H]-cAMP concentrations were determined and expressed as a percentage of total [^{3}H]-ATP, calculated as described in *Materials and Methods*. Results are the means ± SE of triplicate determinations in three different experiments.

analyzed the PCR products taken from cycles 30 and 40 of the PCR. Average results from at least three independent experiments are displayed in table 2. The MC1R showed only weak signals and could only be seen in three brain parts on the Southern blot, indicating very low levels of expression. Expression of MC2R (seen with EtBr staining) was found in telencephalon and hypothalamus, but could be detected with hybridization assay in head-kidney (corresponding to the adrenal gland) at the cycle 40. RT-PCR of MC4R was detected in high levels (seen on EtBr gel after cycle 30) in all brain parts except the brain stem and slightly lower levels in head-kidney and gut. A signal for the MC5R was found in brain, head-kidney, and eye.

Discussion

There are two remarkable genomic features of the MCRs in *Fugu*. Firstly, there exist introns in two of the subtypes that are not found in any MCR in any other species (see below). Secondly, it seems almost certain that the *Fugu* does not have any MC3R subtype and that we have identified the full set of the MCR genes in *Fugu*, based on Blast search results of the nearly completed genomes, *Takifugu rubripes* with 95% genome coverage and *Tetraodon nigroviridis* with 83% coverage. It is unlikely (probability of 0.008) that MC3R gene is not found in these fishes because of an incomplete data set. Considering the other subtypes, the *Fugu*, *Tetraodon*, and zebrafish MCRs share high amino acid identity with the respective human orthologs (53% to 69% identity), which is higher than the identity of MCRs between the subtypes (38% to 60%). The phylogenetic analysis indicates clearly that newly identified receptors are orthologs of human MCR family (fig. 2). The TM2, TM3, TM6, and TM7 in the *Fugu* MCRs display the highest degree of conservation with MCRs from mammalian species, whereas they show almost no similarity in the N-terminal and C-terminal regions. This is in line with results that suggest involvement of TM2, TM3, TM6, and TM7 in the recognition and binding of MSH peptides (Schiöth 2001).

We showed that there exist functionally spliced introns in two of the MCR genes in *Fugu*. This is intriguing, considering that none of the MCRs cloned so far, including the MCRs in zebrafish, have introns in their coding sequences. Moreover, our group has cloned additional MCR genes from the teleost linage, but we have not found any introns in these genes (unpublished results). In general, positions and the number of introns are very conserved between different species of vertebrates. Nevertheless, several reports exists about recent gain or differences in intron composition in vertebrates (Logsdon, Stoltzfus, and Doolittle 1998; Stoltzfus et al. 1997; O'Neill et al. 1998) and fish species in particular (Venkatesh, Ning, and Brenner 1999; Figueroa et al. 1995; Sandford et al. 1997). It is intriguing that the MC2R intron is located at the same position as one of the MC5R introns. Moreover, introns in this very same amino acid sequence (DRY) can be found in many GPCRs. This is one of the most important motifs that distinguishes the large rhodopsin group of GPCRs and is believed to be crucial in keeping the receptors in an inactivated form in absence of a ligand. The question arises as to whether this is an ancient intron that has been lost in all other members of MCR family but remained intact in two *Fugu* genes, as suggested by Logan et al. (2003), or whether these introns have been inserted into the genome later in the evolution of *Fugu*. We tend to believe that the introns found in MCR genes in *Fugu* are the result of recent insertions. Firstly, MC2R and MC5R are the most divergent members in the MCR family and may therefore have split from a common ancestor earlier than other MCR genes, where no traces of introns have been found. It seems unlikely that two genes could retain introns in *Fugu*, whereas many other MCR genes from related teleost and more ancestral species are intronless (unpublished results). Secondly, the genetic code of the DRY motif can form a protosplice site C/A, A, G, R, which is believed to be target site for insertion of spliceosomal introns by a process called reverse splicing (Dibb and Newman 1989; Dibb 1991, 1993). Although this theory is still debated, recent gene findings in "lower" animals and plants confirm the preferable insertion of spliceosomal introns into protosplice sites (Di Maro et al.

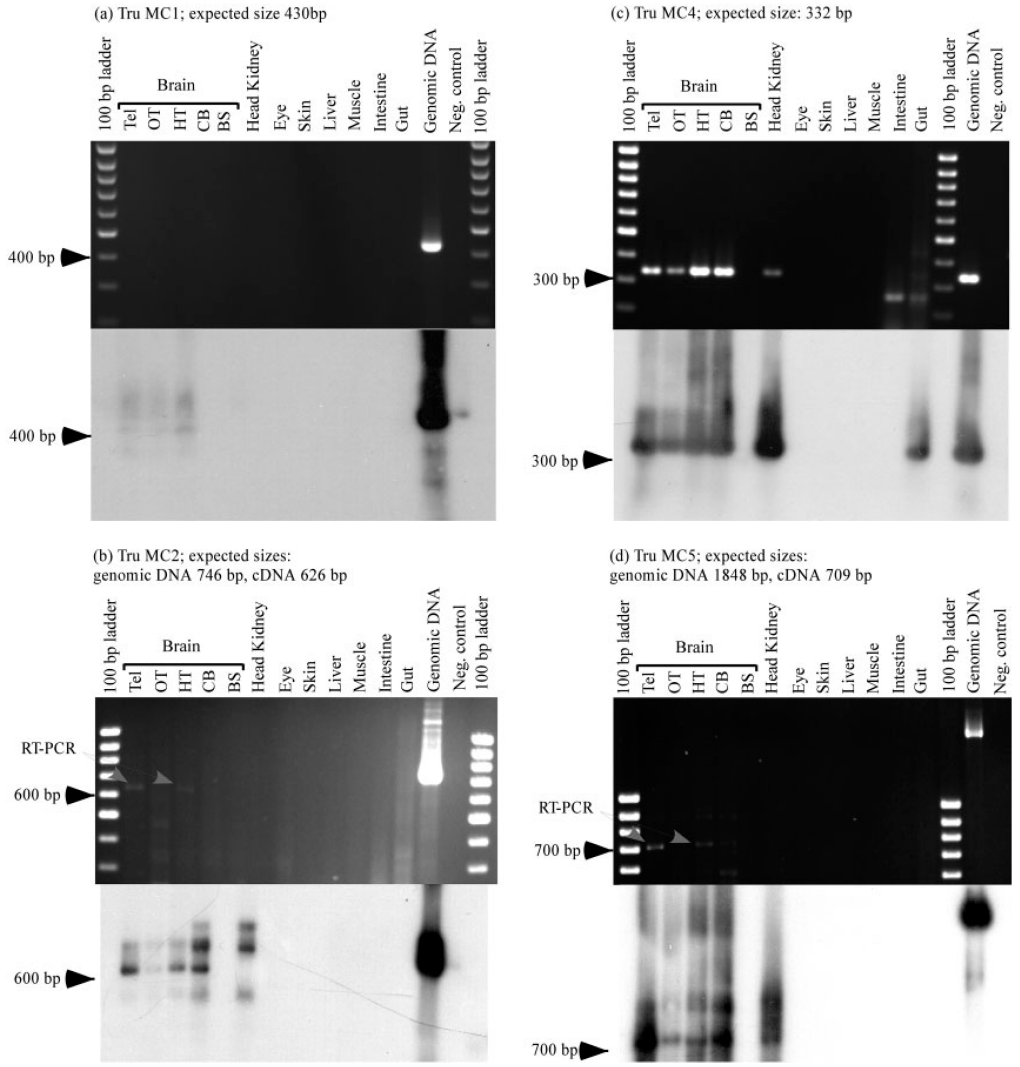


FIG. 8.—Expression of *Fugu* MC1R (*A*), MC2R (*B*), MC4R (*C*), and MC5R (*D*) mRNA as determined by RT-PCR. The tissues, controls, and expected sizes of the PCR products are denoted at the top of each figure. Ethidium bromide–stained agarose gels are presented on top and autoradiographs of Southern blots, hybridized with gene-specific probes are shown below. The names of the brain parts are abbreviated as follows: telencephalon (Tel), optic tectum (OT), hypothalamus (HT), cerebellum (CB), and brain stem (BS).

2002; Funke et al. 1999). We thus believe that many examples of intron position in DRY motif of GPCR can rather be explained by intron insertion that, in some cases, may have occurred late in the vertebrate evolution such as in the MCRs in *Fugu*, whereas others, such as the ones in galanine or neurokinin (NK) receptors, have occurred much earlier, giving rise to conserved introns in mammals and "lower" vertebrates (unpublished data). Moreover, the fact that rhodopsin GPCRs that have the DRY intron are not phylogenetically clustered, as far as we can determine, also indicates that the suggestion of common DRY intron in the GPCRs may be wrong. For example, the MC, NK, and galanine receptors all belong to different clusters of rhodopsin GPCRs or the α, β, and γ subgroups of the rhodopsin family, respectively (Fredriksson et al. 2003). It is possible that intron insertion process in pufferfish is more frequent taking advantage of the protosplice site compared with other species, and it is possible that it has played some role in large genomic rearrangements, resulting in the compactness of *Fugu* genome.

It seems evident that not only the four MCRs are conserved in *Fugu* but also their ligands show a high degree of conservation. It is well established that there exist a high level of similarity between the α-MSH peptide sequence in POMC among vertebrate species. As seen from figure 3, the MSH peptides contain amino acid sequence HFRW known to be the core sequence in ligand-receptor recognition. It has also been known that a number

Table 2
Anatomical Localization of the *Fugu* MCR mRNA Expression As Estimated by RT-PCR with Cycle Numbers 30 and 40

	Brain											
Fugu Gene Receptor	Tel	OT	HT	CB	BS	Head-Kidney	Eye	Skin	Liver	Muscle	Intestine	Gut
MC1 cycle 30												
MC1 cycle 40	*	*	*									
MC2 cycle 30												
MC2 cycle 40	+ *	*	+ *	*		*						
MC4 cycle 30	+ *	+ *	+ *	+ *								
MC4 cycle 40	+ *	+ *	+ *	+ *		+ *						*
MC5 cycle 30	*	*	*	*		*	*					
MC5 cycle 40	+ *	*	+ *	+ *		*	*					
β-actin	25	25	25	25	30	30	30	35	25	30	35	30

NOTE.—Plus signs indicate specific PCR fragments visible on ethidium bromide–stained gels. Specific PCR bands observed in Southern blots are marked with asterisks. The bottom row represents the results of kinetics of β-actin–specific PCR. The number of PCR cycles shows when the PCR product can be detected on an EtBr-stained gel.

of teleost fishes lost the region in the POMC gene coding for γ-MSH. This fact becomes particularly interesting in *Fugu*, which is missing MC3R, because MC3R in mammals has selectivity for γ-MSH. The role and importance of MC3R in teleost fish species remains, however, obscure, considering that the zebrafish has a MC3R.

Fairly recently, it was established that MCRs have not only POMC products as ligands but also AGRP and ASIP. AGRP has previously been shown to exist in chicken, and our data show that there exists an Agrp peptide in fish as well. It is thus conceivable that the dual ligand system consisting of both MCR agonist and antagonist arose early in vertebrate evolution. The Cys rich part in the C-terminal of the *Fugu* AGRP seems to be very well conserved, whereas the other part is much less conserved. This is in agreement with data showing that this part of the peptide is important for the interaction with MCRs (Tota et al. 1999). We also found partial genes that may account for additional genes that are likely to represent additional members of the AGRP/ASIP group of peptides. The low similarity in the N-terminal region and inconsistency of several gene prediction programs made it impossible with certainty to predict the full length of these genes, but blasts and phylogeny experiments (data not shown) indicate that these are more similar to AGRP than the agouti peptides.

The availability of the complete repertoire of MCR clones from a nonmammalian vertebrate species made it in particular interesting to investigate the functionality of these receptors, and we performed a thorough characterization of the pharmacological properties. One of the important characteristics of the mammalian MC1R is that it has very high affinity to α-MSH and similar or slightly lower affinity for ACTH. Surprisingly, we found that the *Fugu* MC1R had much lower affinity for α-MSH as compared with the human MC1R. In similar fashion, the *Fugu* had also lower affinity for the β-MSH and γ-MSH, whereas the potency order of these endogenous peptides was the same as for the mammalian MC1Rs. Remarkably, we also found that ACTH had more than 10-fold higher affinity than α-MSH to the *Fugu* MC1R. Our new data may suggest that the early vertebrate MCR had preference to ACTH peptides, whereas the sensitivity for the shorter POMC products such as α-MSH, β-MSH, and γ-MSH has appeared as the MCR subtypes gained more specified functions. This could also fit to the notion that the short MSH peptide sequences are copies of the original ACTH peptide sequence in the POMC gene. Previous studies have also indicated the mouse MC1R has much lower affinity for ACTH than α-MSH (Mountjoy 1994), suggesting that the mouse MC1R (that is perhaps evolving faster than the human one, considering that the mouse genome evolves about twice as fast as the human one) is losing its affinity for ACTH peptides. Taken together, this may support the hypothesis that the MC1R has evolved from being an ACTH-preferring receptor to being a specific α-MSH receptor in mammals.

One of the remarkable properties of the MCR family in mammalian species is that one of the MCRs does not bind MSH peptides at all. The MC2R is not able to bind MSH peptides, only recognizing ACTH peptides. It is not known how this unique property has evolved, and the MC2R in *Fugu* is the most "distant" receptor of this type that has been characterized. The pharmacological data show that the *Fugu* MC2R responds to ACTH, but it is not activated by α-MSH, showing similar characteristics as the mammalian ACTH receptors. This clearly indicates that the specific ACTH selectivity and perhaps functionality in mediating steroidoneogenesis arose early in evolution of vertebrates and that this property is likely to be crucial for its function in most vertebrates.

The MC4R is one of the most pursued GPCR for drug development, yet it is surprising how low affinity it has for the natural MSH peptides. The *Fugu* MC4R shows about 10-fold higher affinity for the α-MSH and β-MSH as compared with the human MC4R, whereas, in similar fashion as the MC1R, the affinity is much higher for ACTH. The high potency of ACTH at the MC1R, MC2R, and MC4R in *Fugu* provides further support to the possibility that it was indeed the ACTH that was the "original" ligand for the early MCRs, which apparently was created very early during vertebrate evolution. One of the most distinct pharmacological characteristics of the MC4R in mammals is a particularly low affinity for γ-MSH. The MC3R has preference to γ-MSH and, together with the fact that these two MCRs are the most abundantly

expressed in brain of the MCR repertoire in mammals, has made this property very valuable in physiological studies aiming to delineate the specific central effects of the MCRs. Remarkable, γ-MSH has several hundred–fold higher affinity for the *Fugu* MC4R than it has for the human MC4R. This is very interesting, as it is conceivable that the MC4R has evolved in such manner that it loses its binding ability to γ-MSH. The role of γ-MSH is still quite obscure, but these data clearly indicate the importance of γ-MSH to not interfere with the MC4R and suggest that "this property" has become specifically important in "higher" vertebrates. It is also possible that this inability to bind γ-MSH has been lost in the *Fugu* lineage, as it does not have the γ-MSH sequence in POMC. The inability of the human MC4R to bind γ-MSH has been linked to Tyr268 on the top of the TM6 (Oosterom et al. 2001). It is interesting to note that the *Fugu* MC4R is lacking the corresponding Tyr, as can be seen in figure 1. It is also interesting that we do not observe preference of the *Fugu* MC4R to β-MSH in a similar way as we have previously shown for both the rat and human MC4R (Schiöth et al. 2002). There is increasing evidence that β-MSH may have a specific role for the feeding response through MC4R in mammals (Harrold, Widdowson, and Williams 2003). The lack of specificity for the β-MSH at the *Fugu* MC4R may suggest that this is a property that has evolved at later stages of vertebrate evolution. The MC5R in *Fugu* has in general higher affinity for the natural MSH peptides as compared with the human ortholog. This is particularly evident for α-MSH, which has about 70-fold higher affinity for the *Fugu* receptor, indicating that the mammalian MC5Rs may have lost their affinity to the α- MSH, β-MSH, and γ-MSH, but so little is known about the functional role of this receptor that it is difficult to speculate about the importance of this. We have also shown that all the MCRs can functionally couple to the Gs-linked signalling pathway in response to α-MSH and/or ACTH in line with previous results on the mammalian MCRs.

Although the RT-PCR is not well suited for quantitative analysis of gene expression, this approach is very effective in detecting expression of genes sets in a wide range of tissues. The expression of the MC2R in the brain is intriguing. The MC2R was found in four of the five brain regions analyzed (see fig. 8 and table 2), and sequencing of the RT-PCR products confirmed the predicted splicing of this gene in the brain. After the first cloning of the MC2R and the subsequent demonstration that this was indeed an ACTH receptor in mammals, several groups preformed an extensive search for this receptor in the brain (Xia and Wikberg 1996). This was because of the physiological textbook perception that ACTH played a role in a central negative feedback (Motta, Mangili, and Martini 1965). The MC2R has not been found in brain of any mammal, and the central ACTH-binding sites that were found early on can probably be attributed to the MC3R (Schiöth et al. 1997*a*). It is possible that the presence of the MC2R in the *Fugu* brain may suggest that there exists a negative feedback loop involving this receptor that may have been later taken over by, for example, the MC3R in mammals. Expression of mammalian MC4Rs has only been found in brain tissue. The fact that we have detected MC4R transcripts in telencephalon, optic tectum, hypothalamus, cerebellum, head-kidney. and gut of *Fugu* is in agreement with previously found observations from zebrafish (Ringholm et al. 2002), where MC4R was in addition to brain found to be expressed in the eye, GI tract, and ovaries. In chicken, the MC4R is expressed in a wide variety of peripheral tissues, including the heart, adrenal glands, ovaries, testes, spleen, adipose tissue, and eye, as well as the brain (Takeuchi and Takahashi, 1998). It is therefore possible that restriction of MC4R expression to the brain is a feature of mammalian species and has developed because of needs for more specific CNS-driven regulation of metabolic and behavioral processes. It has recently been demonstrated that MC4R participates in central regulation of food intake in teleost fish (Cerda-Reverter et al. 2003), but the presence of this receptor in peripheral tissues may suggest that it may also have additional functions in fishes. The MC5R is expressed in a wide range of tissues not only in mammals but also in chicken (Takeuchi and Takahashi 1998) and zebrafish. Our results show that also the MC5R in *Fugu* has both central and peripheral distribution, as we were able to detect the MC5R transcripts in brain, head-kidney, and eye, even though the expression pattern seems to be less widely spread than in many other species. It seems thus evident that the MC5R took on both central and peripheral functions early in vertebrate evolution.

Gene duplications occur by both individual and/or block duplications. The large-scale duplications, including polyploidizations, are believed to be important for shaping vertebrate genomes (Ohno 1970). Two rounds of large-scale duplications are proposed to have occurred in early vertebrate ancestry, resulting in up to four copies of each gene in mammals, which originate from a common ancestral gene in cephalochordates. This is now known as the "2R hypothesis" or the "one to four model" (Lundin 1993; Holland et al. 1994). When analyzing the neighboring regions of MCR genes, we found remarkable synteny in gene content and order between *Fugu* and corresponding human genome regions on chromosomes 16 and 18. Although the *Fugu*, like most teleost fishes, most probably underwent a lineage-specific genome duplication (the third round) approximately 250 MYA (Postlethwait et al. 1998), we were not able to detect duplicated genes in the three *Fugu* regions studied. Nevertheless, the finding of such extent of synteny is remarkable, taking into consideration more than 400 Myr since split of the lineages leading to teleosts and mammals. The highly conserved linkage of MC2R and MC5R genes is found in all representatives of different vertebrate classes (see figure 4) (Schiöth et al. 2003). The functional relevance of this tandem remains a mystery. This conserved linkage suggests that these receptors arose from a common ancestor by a local gene duplication event. It is intriguing that the MC2R and MC5R subtypes are evolutionary the most distant among the MCRs as seen from the phylogenetic analysis (fig. 2). This indicates that a local duplication responsible for this linkage occurred very early, most probably before the two putative tetraploidizations, which, according to the 2R hypothesis, took place before the origin of gnathostomes (Holland et al. 1994;

Sidow 1996). The phylogenetic analysis clearly shows that MC4R and MC3R are more closely related to each other than to other MCRs. This suggest that if the MC2R and MC5R linkage have arisen from common ancestor through a local duplication, the MC3R might have arisen later from the common MC5R-related ancestor gene, perhaps through tetraploidization event. MC4R is, however, most likely the result of local duplication from MC5R that occurred after the last common tetraploidization in vertebrate lineage. Alternatively, MC1R and MC2R might have common ancestor gene and appeared through the genome duplication events. Although there is less similarity between these two receptors than in MC3/4/5R group, this scenario might be explained by different evolution rates, perhaps as MC2R has the most divergent pharmacological/physiological characteristics. An important question is when these duplications took place. According to our findings of MCR genes in lamprey (preliminary unpublished results), the appearance of MC1/2R and MC3/4/5R ancestor gene linkage can be dated before the appearance of gnathostomes (jawed vertebrates) but after the first tetraploidization. Taking into account the high similarity between MC3R, MC4R, and MC5R and presence of all these receptors in zebrafish, MC4R or MC3R might appeared through local duplication of MC5R soon after the last genome duplication after the split of gnathostomes. Figure 9 summarizes above-mentioned observations and provides a putative scheme for evolution of MCR genes through the duplication events. The order of appearance for MC3R and MC4R remains unclear. However, MC4R is located on the same chromosome in human, mouse, and chicken, thus indicating that MC4R might be a duplicate of a common ancestor to the MC5R branch. It is intriguing that the *Fugu* MC5R gene, which is linked with the MC2R gene, is, according to phylogenetic analysis, an ortholog of the zebrafish MC5aR gene. In zebrafish, however, it is the other one, or the MC5bR gene, that is linked with the MC2R gene. This would mean that different copies of MC2R gene have been lost in *Fugu* and zebrafish, providing further strong evidence for the ancient origin of the remarkable MC2R and MC5R gene linkage. It also indicates that two copies of MC2R gene existed in the teleost lineage before the split of *Fugu* and zebrafish. This is in agreement with our previous molecular clock investigation of the two zebrafish MC5Rs, dating them to the predicted teleost-specific genome duplication 250 MYA.

In summary, we have characterized the gene repertoire of the MC system in *Fugu*. We show that there exists an AGRP peptide gene in addition to a POMC gene. The receptor gene repertoire does not include an MC3R gene, but the other receptors are present in single copies. The MC2R and MC5R are found in tandem, and both of them have introns that we suggest may have been inserted in a protosplice site in the *Fugu* lineage through a reverse splicing mechanism. We show that there exist remarkable synteny of both HSA16 and HSA18 with large contigs within the *Fugu* genome that include all four *Fugu* MCRs. The pharmacological characterization is the first that has been performed on a complete set of MCRs from any nonmammalian species. The results show that ACTH has

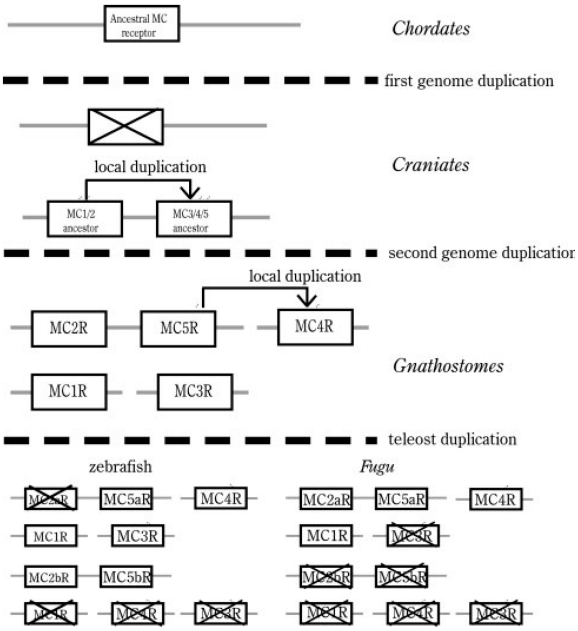


FIG. 9.—A hypothetical duplication scheme for the MCRs during vertebrate evolution according to the 2R theory. The dashed lines represent the putative events of total-genome duplication in the vertebrate lineage. The genes that were lost during evolution are marked X. The possible local duplication events are shown with arrows. Note that different copies of MC2R genes have been lost in *Fugu* and zebrafish.

remarkable high affinity for the MC1R and MC4R, whereas α-MSH had lower affinity. The MC2R in *Fugu* is an ACTH receptor that does not respond to α-MSH. The anatomical characterization indicates that the MC2R is expressed in the brain in addition to the head-kidney, whereas the MC4R and MC5R are found in both central and peripheral tissues. Overall, this is the first comprehensive genomic and functional characterization of a GPCR family within the *Fugu* genome. The studies provide insight into how genes that code for a system of peptides binding to GPCRs, that like many other such systems are involved in multiple and diverse functions in both central and peripheral tissues, may have gained and lost its functions through the evolution of vertebrates.

Acknowledgments

We thank Dr. Earl T. Larsson, Uppsala University, for expert assistance during dissection of *Fugu* fish. Dr. Janis Klovins was supported by the Wenner-Gren foundation and by a Marie Curie Fellowship of the European Community programme "Improving the Human Research potential and the Socio-Economic Knowledge Base" under contract number HPMF-CT-2002-01786. Dr. Nicole Gallo-Payet is a holder of a Canada Research Chair in Endocrinology of the Adrenal Gland. The studies were supported by the Swedish Research Council (VR, medicine), Canadian Institute for Health Research (MOP 10998), the Swedish Society for Medical Research (SSMF), Svenska Läkaresällskapet, Åke Wikberg Foun-

dation, The Novo Nordisk Foundation, Magnus Bergwall Foundation, and Melacure Therapeutics AB, Uppsala.

Literature Cited

Amemiya, Y., A. Takahashi, R. M. Dores, and H. Kawauchi. 1997. Sturgeon proopiomelanocortin has a remnant of gamma-melanotropin. Biochem. Biophys. Res. Commun. **230**: 452–456.

Aparicio, S., J. Chapman, E. Stupka et al. (38 co-authors). 2002. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. Science **297**:1301–1310.

Carroll, R. L. 1988. Vertebrate paelontology and evolution. W.H. Freeman and Co., New York.

Cerda-Reverter, J. M., A. Ringholm, H. B. Schiöth, and R. E. Peter. 2003. Molecular cloning, pharmacological characterization, and brain mapping of the melanocortin 4 receptor in the goldfish: involvement in the control of food intake. Endocrinology **144**:2336–2349.

Chen, A. S., D. J. Marsh, M. E. Trumbauer et al. (18 co-authors). 2000. Inactivation of the mouse melanocortin-3 receptor results in increased fat mass and reduced lean body mass. Nat. Genet. **26**:97–102.

Chen, W., M. A. Kelly, X. Opitz-Araya, R. E. Thomas, M. J. Low, and R. D. Cone. 1997. Exocrine gland dysfunction in MC5-R-deficient mice: evidence for coordinated regulation of exocrine gland function by melanocortin peptides. Cell **91**:789–798.

Chhajlani, V. 1996. Distribution of cDNA for melanocortin receptor subtypes in human tissues. Biochem. Mol. Biol. Int. **38**:73–80

Crnogorac-Jurcevic, T., J. R. Brown, H. Lehrach, and L. C. Schalkwyk. 1997. *Tetraodon fluviatilis*, a new puffer fish model for genome studies. Genomics **41**:177–184.

Danielson, P. B., J. Alrubaian, M. Muller, J. M. Redding, and R. M. Dores. 1999. Duplication of the POMC gene in the paddlefish (*Polyodon spathula*): analysis of gamma-MSH, ACTH, and beta-endorphin regions of ray-finned fish POMC. Gen. Comp. Endocrinol. **116**:164–177.

Dibb, N. J. 1991. Proto-splice site model of intron origin. J. Theor. Biol. **151**:405–416.

———. 1993. Why do genes have introns? FEBS Lett. **325**:135–139.

Dibb, N. J., and A. J. Newman. 1989. Evidence that introns arose at proto-splice sites. EMBO J. **8**:2015–2021.

Di Maro, A., E. Pizzo, M. V. Cubellis, and G. D'Alessio. 2002. An intron-less betagamma-crystallin-type gene from the sponge *Geodia cydonium*. Gene **299**:79–82.

Dores, R. M., T. R. Smith, D. A. Rubin, P. Danielson, L. E. Marra, and J. H. Youson. 1997. Deciphering posttranslational processing events in the pituitary of a neopterygian fish: cloning of a gar proopiomelanocortin cDNA. Gen. Comp. Endocrinol. **107**:401–413.

Farooqi, I. S., G. S. Yeo, J. M. Keogh, S. Aminian, S. A. Jebb, G. Butler, T. Cheetham, and S. O'Rahilly. 2000. Dominant and recessive inheritance of morbid obesity associated with melanocortin 4 receptor deficiency. J. Clin. Invest. **106**:271–279.

Figueroa, F., H. Ono, H. Tichy, C. O'Huigin, and J. Klein. 1995. Evidence for insertion of a new intron into an MHC gene of perch-like fish. Proc. R. Soc. Lond. B Biol. Sci. **259**:325–330.

Fredriksson, R., M. C. Lagerström, L-G. Lundin, and H. B. Schiöth. 2003. The G-protein coupled receptors in the human genome form five main families: phylogenetic analysis, paralogon groups and fingerprints. Mol. Pharmacol. **63**:1256–1272.

Funke, R. P., J. L. Kovar, J. M. Logsdon, Jr., J. C. Corrette-Bennett, D. R. Straus, and D.P. Weeks. 1999. Nucleus-encoded, plastid-targeted acetolactate synthase genes in two closely related chlorophytes, *Chlamydomonas reihardtii* and *Volvox carteri*: phylogenetic origins and recent insertion of introns. Mol. Gen. Genet. **262**:12–21.

Gallo-Payet, N., and M. D. Payet. 1989. Excitation-secretion coupling: involvement of potassium channels in ACTH-stimulated rat adrenocortical cells. J Endocrinol. **120**:409–421

Gantz, I., and T. M. Fong. 2003. The melanocortin system. Am. J. Physiol. Endocrinol. Metab. **284**:E468–E474.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. **41**:95–98.

Harrold, J. A., P. S. Widdowson, and G. Williams. 2003. Beta-MSH: a functional ligand that regulated energy homeostasis via hypothalamic MC4-R? Peptides **24**:397–405.

Holland, P. W., J. Garcia-Fernandez, N. A. Williams, and A. Sidow. 1994. Gene duplications and the origins of vertebrate development. Development (suppl.) **1994**:125–133.

Huszar, D., C. A. Lynch, V. Fairchild-Huntress et al. (14 co-authors). 1997. Targeted disruption of the melanocortin-4 receptor results in obesity in mice. Cell **88**:131–141.

Kask, A., F. Mutulis, R. Muceniece, R. Pahkla, I. Mutule, J. E. Wikberg, L. Rago, and H. B. Schiöth. 1998. Discovery of a novel superpotent and selective melanocortin-4 receptor antagonist (HS024): evaluation in vitro and in vivo. Endocrinology **139**:5006–5014.

Kitahara, N., T. Nishizawa, K. Iida, H. Okazaki, T. Andoh, and G. I. Soma. 1988. Absence of a gamma-melanocyte-stimulating hormone sequence in proopiomelanocortin mRNA of chum salmon *Oncorhynchus keta*. Comp. Biochem. Physiol. B **91**:365–370.

Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. Distributed by the authors, Arizona State University, Tempe.

Lee, J., P. Danielson, C. Sollars, J. Alrubaian, P. Balm, and R. M. Dores. 1999. Cloning of a neoteleost (*Oreochromis mossambicus*) pro-opiomelanocortin (POMC) cDNA reveals a deletion of the gamma-melanotropin region and most of the joining peptide region: implications for POMC processing. Peptides **20**:1391–1399.

Logan, D. W., R. J. Bryson-Richardson, K. E. Pagan, M. S. Taylor, P. D. Currie, and I. J. Jackson. 2003. The structure and evolution of the melanocortin and MCH receptors in fish and mammals. Genomics **81**:184–191.

Logsdon, J. M., Jr.,A. Stoltzfus, and W. F. Doolittle. 1998. Molecular evolution: recent cases of spliceosomal intron gain? Curr. Biol. **8**:R560–R563.

Lu, D., D. Willard, I. R. Patel et al. (11 co-authors) 1994. Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. Nature **371**:799–802.

Lundell, I., H. Eriksson, U. Marklund, and D. Larhammar. 2001. Cloning and characterization of the guinea pig neuropeptide Y receptor Y5. Peptides **22**:357–363.

Lundin, L. G. 1993. Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. Genomics **16**:1–19.

Miraglia Del Giudice, E., G. Cirillo, V. Nigro, N. Santoro, L. D'Urso, P. Raimondo, D. Cozzolino, D. Scafato, and L. Perrone. 2002. Low frequency of melanocortin-4 receptor (MC4R) mutations in a Mediterranean population with early-onset obesity. Int. J. Obes. Relat. Metab. Disord. **26**: 647–651.

Motta, M., G. Mangili, and L. Martini. 1965. A "short" feedback loop in the control of ACTH secretion. Endocrinology **77**:392–395.

Mountjoy, K. G. 1994. The human melanocyte stimulating hormone receptor has evolved to become "super-sensitive" to melanocortin peptides. Mol. Cell. Endocrinol. **102**:R7–R11

Noon, L. A., J. M. Franklin, P. J. King, N. J. Goulding, L. Hunyady, and A. J. Clark. 2002. Failed export of the adrenocorticotrophin receptor from the endoplasmic reticulum in non-adrenal cells: evidence in support of a requirement for a specific adrenal accessory factor. J. Endocrinol. **174**:17–25.

Ohno, S. 1970. Evolution by gene duplication. Springer-Verlag, Berlin.

Okuta, A., H. Ando, H. Ueda, and A. Urano. 1996. Two types of cDNAs encoding proopiomelanocortin of sockeye salmon, *Oncorhynchus nerka*. Zoolog. Sci. **13**:421–427.

O'Neill, R. J., F. E. Brennan, M. L. Delbridge, R. H. Crozier, and J. A. Graves. 1998. De novo insertion of an intron into the mammalian sex determining gene, SRY. Proc. Natl. Acad. Sci. USA **95**:1653–1657.

Oosterom, J., K. M. Garner, W. K. den Dekker, W. A. Nijenhuis, W. H. Gispen, J. P. Burbach, G. S. Barsh, and R. A. Adan. 2001. Common requirements for melanocortin-4 receptor selectivity of structurally unrelated melanocortin agonist and endogenous antagonist, Agouti protein. J. Biol. Chem. **276**:931–936.

Penhoat, A., D. Naville, H. El Mourabit, A. Buronfosse, P. Durand, and M. Begeot. 2000. Functional expression of the human ACTH receptor gene. Endocr. Res. **26**:549–557.

Postlethwait, J. H., Y. L. Yan, M. A. Gates et al. (29 co-authors). 1998. Vertebrate genome evolution and the zebrafish gene map. Nat. Genet. **18**:345–349.

Rana, B. K., D. Hewett-Emmett, L. Jin et al. (12 co-authors). 1999. High polymorphism at the human melanocortin 1 receptor locus. Genetics **151**:1547–1557.

Ringholm, A., R. Fredriksson, N. Poliakova, Y. L. Yan, J. H. Postlethwait, D. Larhammar, and H. B. Schiöth. 2002. One melanocortin 4 and two melanocortin 5 receptors from zebrafish show remarkable conservation in structure and pharmacology. J. Neurochem. **82**:6–18.

Ringholm, A., J. Klovins, R. Fredriksson, N. Poliakova, E. T. Larson, J. P. Kukkonen, D. Larhammar, and H. B. Schiöth. 2003. Presence of melanocortin (MC4) receptor in spiny dogfish suggests an ancient vertebrate origin of central melanocortin system. Eur. J. Biochem. **270**:213–221.

Salomon, Y., C. Londos, and M. Rodebell. 1974. A highly sensitive adenylate cyclase assay. Anal. Biochem. **58**:541–548.

Sandford, R., B. Sgotto, S. Aparicio et al. (12 co-authors). 1997. Comparative analysis of the polycystic kidney disease 1 (PKD1) gene reveals an integral membrane glycoprotein with multiple evolutionary conserved domains. Hum. Mol. Genet. **6**:1483–1489.

Schimmer, B. P., W. K. Kwan, J. Tsao, and R. Qiu. 1995. Adrenocorticotropin-resistant mutants of the Y1 adrenal cell line fail to express the adrenocorticotropin receptor. J. Cell. Physiol. **163**:164–171.

Schiöth, H. B. 2001. The physiological role of melanocortin receptors. Vitam. Horm. **63**:195–232.

Schiöth, H. B., A. A. Bouifrouri, R. Rudzish, R. Muceniece, H. Watanobe, J. E. Wikberg, and D. Larhammar. 2002. Pharmacological comparison of rat and human melanocortin 3 and 4 receptors in vitro. Regul. Pept. **106**:7–12.

Schiöth, H. B., V. Chhajlani, R. Muceniece, V. Klusa, and J. E. Wikberg. 1996. Major pharmacological distinction of the ACTH receptor from other melanocortin receptors. Life Sci. **59**:797–801.

Schiöth, H. B., R. Muceniece, M. Larsson, and J. E. Wikberg. 1997*a*. The melanocortin 1, 3, 4 or 5 receptors do not have a binding epitope for ACTH beyond the sequence of alpha-MSH. J. Endocrinol. **155**:73–78.

Schiöth, H. B., R. Muceniece, F. Mutulis, P. Prusis, G. Lindeberg, S. D. Sharma, V. J. Hruby, and J. E. Wikberg. 1997*b*. Selectivity of cyclic [D-Nal7] and [D-Phe7] substituted MSH analogues for the melanocortin receptor subtypes. Peptides **18(7)**:1009–1013.

Schiöth, H. B., R. Muceniece, J. E. Wikberg, and V. Chhajlani. 1995. Characterisation of melanocortin receptor subtypes by radioligand binding analysis. Eur. J. Pharmacol. **288**:311–317.

Schiöth, H. B., S. Petersson, R. Muceniece, M. Szardenings, and J. E. Wikberg. 1997*c*. Deletions of the N-terminal regions of the human melanocortin receptors. FEBS Lett. **410**:223–228.

Schiöth, H. B., T. Raudsepp, A. Ringholm, R. Fredriksson, S. Takeuchi, D. Larhammar, and B. P. Chowdhary. 2003. Remarkable synteny conservation of melanocortin receptors in chicken, human, and other vertebrates. Genomics **81**:504–509.

Sidow, A. 1996. Gen(om)e duplications in the evolution of early vertebrates. Curr. Opin. Genet. Dev. **6**:715–722.

Stoltzfus, A., J. M. Logsdon, Jr.,J. D. Palmer, and W. F. Doolittle. 1997. Intron "sliding" and the diversity of intron positions. Proc. Natl. Acad. Sci. USA **94**:10739–10744.

Takahashi, A., Y. Amemiya, M. Nozaki, S. A. Sower, and H. Kawauchi. 2001. Evolutionary significance of proopiomelanoocortin in agnathan and chondrichthyes. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. **129**:283–289.

Takahashi, A., Y. Amemiya, M. Sarashi, S. A. Sower, and H. Kawauchi. 1995. Melanotropin and corticotropin are encoded on two distinct genes in the lamprey, the earliest evolved extant vertebrate. Biochem. Biophys. Res. Commun. **213**: 490–498.

Takeuchi, S., and S. Takahashi. 1998. Melanocortin receptor genes in the chicken-tissue distributions. Gen. Comp. Endocrinol. **112**:220–231.

Takeuchi, S., K. Teshigawara, and S. Takahashi. 2000. Widespread expression of agouti-related protein (AGRP) in the chicken: a possible involvement of AGRP in regulating peripheral melanocortin systems in the chicken. Biochim. Biophys. Acta **1496**:261–269.

Tan, C. P., K. K. McKee, D. H. Weinberg, T. MacNeil, O. C. Palyha, S. D. Feighner, D. L. Hreniuk, L. H. Van Der Ploeg, D. J. MacNeil, and A. D. Howard. 1999. Molecular analysis of a new splice variant of the human melanocortin-1 receptor. FEBS Lett. **451**:137–141.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

Tota, M. R., T. S. Smith, C. Mao, T. MacNeil, R. T., Mosley, L. H. Van der Ploeg, and T. M. Fong. 1999. Molecular interaction of agouti protein and agouti-related protein with human melanocortin receptors. Biochemistry **38**:897–904.

Vaisse, C., K. Clement, B. Guy-Grand, and P. Froguel. 1998. A frameshift mutation in human MC4R is associated with a dominant form of obesity. Nat. Genet. **20**:113–114.

Venkatesh, B., Y. Ning, and S. Brenner. 1999. Late changes in spliceosomal introns define clades in vertebrate evolution. Proc. Natl. Acad. Sci. USA **96**:10267–10271.

Xia, Y., and J. E. Wikberg. 1996. Localization of ACTH receptor mRNA by in situ hybridization in mouse adrenal gland. Cell Tissue Res. **286**:63–68.

Claudia Kappen, Associate Editor

Accepted October 27, 2003

3.2 Formation of new genes explains lower intron density in mammalian Rhodopsin G protein-coupled receptors.

Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 43 (2007) 864–880
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Formation of new genes explains lower intron density in mammalian *Rhodopsin* G protein-coupled receptors

Davids Fridmanis [a,b], Robert Fredriksson [b], Ivo Kapa [a], Helgi B. Schiöth [b], Janis Klovins [a,b,*]

[a] *Biomedical Research and Study Centre, University of Latvia, Ratsupites 1, LV1067 Riga, Latvia*
[b] *Department of Neuroscience, Uppsala University, BMC, Box 593, SE751 24, Uppsala, Sweden*

Received 19 May 2006; revised 6 October 2006; accepted 2 November 2006
Available online 17 November 2006

Abstract

Mammalian G protein-coupled receptor (GPCR) genes are characterised by a large proportion of intronless genes or a lower density of introns when compared with GPCRs of invertebrates. It is unclear which mechanisms have influenced intron density in this protein family, which is one of the largest in the mammalian genomes. We used a combination of Hidden Markov Models (HMM) and BLAST searches to establish the comprehensive repertoire of *Rhodopsin* GPCRs from seven species and performed overall alignments and phylogenetic analysis using the maximum parsimony method for over 1400 receptors in 12 subgroups. We identified 14 different Ancestral Receptor Groups (ARGs) that have members in both vertebrate and invertebrate species. We found that there exists a remarkable difference in the intron density among ancestral and new *Rhodopsin* GPCRs. The intron density among ARGs members was more than 3.5-fold higher than that within non-ARG members and more than 2-fold higher when considering only the 7TM region. This suggests that the new GPCR genes have been predominantly formed intronless while the ancestral receptors likely accumulated introns during their evolution. Many of the intron positions found in mammalian ARG receptor sequences were found to be present in orthologue invertebrate receptors suggesting that these intron positions are ancient. This analysis also revealed that one intron position is much more frequent than any other position and it is common for a number of phylogenetically different *Rhodopsin* GPCR groups. This intron position lies within a functionally important, conserved, DRY motif which may form a proto-splice site that could contribute to positional intron insertion. Moreover, we have found that other receptor motifs, similar to DRY, also contain introns between the second and third nucleotide of the arginine codon which also forms a proto-splice site. Our analysis presents compelling evidence that there was not a major loss of introns in mammalian GPCRs and formation of new GPCRs among mammals explains why these have fewer introns compared to invertebrate GPCRs. We also discuss and speculate about the possible role of different RNA- and DNA-based mechanisms of intron insertion and loss.

Keywords: GPCR; Evolution; Introns; *Rhodopsin*

1. Introduction

Since the discovery of introns in 1977 by several independent research groups (Berget et al., 1977; Chow et al., 1977; Jeffreys and Flavell, 1977), there has been continuous discussion about the purpose of introns, how are they formed and when they arose, if they are still being formed or if they are mostly of ancient origin. There are at least six distinct functions of introns in the cell. Introns are one of the sources of non-coding RNAs, a source of regulatory elements and act in alternative splicing and trans-splicing. Introns also function as enhancers of meiotic crossing over within coding sequences, in exon-shuffling and as signals for mRNA export from the nucleus and nonsense-mediated decay (Fedorova and Fedorov, 2003).

Two main theories concerning intron evolution have been proposed. First, the "introns early theory" states that introns as such must be very old formations that existed in

* Corresponding author. Fax: +3717442407.
E-mail address: klovins@biomed.lu.lv (J. Klovins).

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.11.007

common prokaryote and eukaryote ancestors, and that prokaryotes must have lost them during their evolution (Darnell, 1978; Doolittle, 1978; see also review Fedorova and Fedorov, 2003; Gilbert, 1987). And second, the "introns late theory" proposes that introns must be a recent addition—selfish DNA, inserted through transposition and spread very quickly all over the genome of an ancient eukaryote before the split of the main eukaryotic lineages (Cavalier-Smith, 1985). The latter theory was supported by Nicholas J. Dibb and Andrew J. Newman who proposed the "proto-splice" site theory which states that conservative introns flanking MAG|R sequence (| represents the position of the intron, M represents A or C, and R represents A or G) could serve as a signal for intron insertion (see also reviews Cavalier-Smith, 1991; Dibb, 1993; Palmer and Logsdon, 1991). Ever since the formation of these two contradicting theories, there has been fierce debate on this matter (Cho and Doolittle, 1997; de Souza et al., 1996, 1997; Fedorov et al., 2001, 2003; Figueroa et al., 1995; Gilbert et al., 1997; Giroux et al., 1994; Hankeln et al., 1997; Lee et al., 1991; see also review Logsdon et al., 1998; Long et al., 1998; Long and Rosenberg, 2000; O'Neill et al., 1998; Roy et al., 2002, 2003; Roy and Gilbert, 2005a,b,c; Sadusky et al., 2004; Tarrio et al., 1998, 2003; Venkatesh et al., 1999). This eventually led to formation of a new "mixed theory of introns" (de Souza et al., 1998; Roy et al., 1999, 2003).

Another process that goes hand in hand with intron gain and loss is gene duplication. It is considered to play a crucial role in development and the evolution of life, even in the RNA world (Brosius, 2003). The gene duplication mechanisms through RNA intermediates is perhaps of most interest in this context since it is one of the oldest and originates from the time when modern life exchanged the carrier of its genetic material from RNA to DNA and it is also the one that involves intron loss. There are at least three other known single gene duplication mechanisms that are taking place in modern cells but they involve mechanisms connected with DNA rearrangement and whole gene (including introns) duplication. All of these mechanisms have been observed to happen with different frequency in different situations and they can appear at different stages of evolution (Danielson and Dores, 1999). A number of specific duplication events have been reported in different organisms and these can be lineage-specific and sometimes even species-specific (Lespinet et al., 2002).

Introns have been extensively studied from the perspective of their origin and functions, however there is still much to clarify from the perspective of possible mechanisms and dynamics of intron gain and loss. The families of genes of predominantly intronless members (like number of vertebrate G protein-coupled receptor (GPCRs) families) are good candidates for investigation of intron gain and loss since the intronless state is an indicator that a particular gene family is undergoing active changes. The density of intronless genes in different eukaryotic organisms varies from 2.7% in *Caenorhabditis elegan* to 97.7% in *Encephalitozoon cuniculi* (12.3% in human), however there are relatively few studies that address these issues (Brosius, 1999; Gentles and Karlin, 1999; Sakharkar et al., 2004).

One of the gene families in mammals with a large proportion of intronless genes is the *Rhodopsin* GPCRs (Brosius, 1999; Gentles and Karlin, 1999). GPCRs are one of the largest protein families in the mammalian genome with over 800 genes in humans (Fredriksson et al., 2003). GPCRs detect extracellular signals and transduce them into the cell using activation of GDP–GTP exchange on heterotrimeric G-proteins. This family includes receptors for many hormones, neurotransmitters, chemokines, calcium ions, and sensory molecules. Despite the great variety of ligands that bind to these proteins, they all share the same seven transmembrane (TM) structure (Pierce et al., 2002). The *Rhodopsin* family is the largest sub family among GPCRs and is useful to study the evolutionary history of introns because it contains several receptors that are recognized as ancient and also many groups that have expanded at different stages of evolution. Moreover, all the *Rhodopsin* GPCRs can be aligned in the 7TM region. There are no *Rhodopsin* GPCR in plants (Fredriksson and Schioth, 2005) and this group of GPCRs can be found only in animals. Bacteriorhodopsin has a 7TM but does not couple to a G protein. The number of GPCRs has increased approximately 6-fold from nematodes to mammals (Fredriksson and Schioth, 2005). Interestingly, although many of *Rhodopsin* GPCRs in mammals are intronless, most GPCRs from invertebrates contain at least one intron. Previous studies have proposed an explanation which has resulted in two very different conclusions. The first papers suggest that the low number of introns in mammalian GPCRs is related to gene multiplication through retroposition, which means that new genes are intronless (Brosius, 1999; Gentles and Karlin, 1999). The second conclusion, which is quite controversial, suggests that there occurred a major loss of introns within the *Rhodopsin* GPCR family. It was argued that the majority of genes lost an "ancient" intron located in the conservative DRY motif of *Rhodopsin* GPCR (Bryson-Richardson et al., 2004).

We also have recently established a detailed phylogenetic relationship of *Rhodopsin* GPCRs (Bjarnadottir et al., 2006; Fredriksson et al., 2003; Lagerstrom et al., 2006). Moreover, the amount of primary sequences from different genomes has become much more complete. This provides a good opportunity to study the introns of the *Rhodopsin* GPCR family in detail considering the different theories about intron loss and insertions. In this study, we present detailed intron analysis of the human *Rhodopsin* GPCR family by tracking back their ancestral origin and showing the differences in intron density in GPCR groups that are only found in higher vertebrates compared to groups-containing orthologous genes in lower animals.

2. Materials and methods

All protein sequences for human GPCRs from the *Rhodopsin* family, except olfactory receptor group, were

downloaded from http://www.ncbi.nlm.nih.gov/ according to ID numbers given in Fredriksson et al. (2003). The corresponding mRNA sequences were downloaded through the following reference links in Genbank files. All intron positions inside these mRNA sequences were identified by performing BLAT searches on the http://genome.ucsc.edu/ server (Kent, 2002), against the human genomic sequence assembly. We also identified 7TM domain region for all receptors using NCBI Conserved Domain Search v2.03 available on http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml. The results from these searches that exceeded 700 bp in length were considered as true, others were checked again with TMpred—Prediction of Transmembrane Regions and Orientation software (Hofmann and Stoffel, 1993) available on http://www.ch.embnet.org/software/TMPRED_form.html. After removing the long N- and C-terminals from some receptors, identified by RPS-BLAST searches at http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi, these protein sequences were aligned with ClustalW 1.81 (Thompson et al., 1994) using default settings. We constructed the Hidden Markov Model (HMM) from these alignments using the HMMbuild application with default settings and calibrated using HMMcalibrate application both from HMMER 2.2 package (Eddy, 1998).

Ciona savignyi genomic sequences-containing FASTA files were downloaded from http://mendel.stanford.edu/sidowlab/ciona.html, divided into smaller files of two Megabytes (2 MB) and used in gene prediction with Genscan (Burge and Karlin, 1997) using the human parameter file HumanIso.smat after which all complete predictions were gathered automatically using batch scripts into a set of predicted *C. savignyi* genes. These coding sequences were then transformed to protein sequences and used for further investigation. FASTA files containing the protein versions of the Genscan predicted gene-sets for other organisms were downloaded from ftp://ftp.ncbi.nlm.nih.gov/genomes/ for *Caenorhabditis elegans*, and ftp://ftp.ensembl.org/pub/ for *Mus musculus*, *Takifugu rubripes*, *Danio rerio*, and *Ciona intestinalis*. These predicted protein sets were searched against the HMM-model using HMMsearch from HMMER 2.2 with a cut off at $E = 10$. All hits with an *E*-value below 0.01 were considered correct and all hits with *E*-values between 0.01 and 10 were manually inspected using BLASTP searches against the NCBI GenPept dataset. For each predicted protein, the top five hits were manually inspected and a minimum of four of these had to be a GPCR for inclusion in the particular dataset, all other predicted proteins were excluded from further analysis.

The human *Rhodopsin* GPCRs were divided into 12 groups according to Fredriksson et al. (2003) (Table 1) and used to subdivide the ones found in the HMM-searches. A BLAST database was constructed from the human *Rhodopsin* GPCRs and the *Rhodopsin* GPCRs from the other species were searched against this database. A cut-off value of $E = 1e-9$ was used and the five top hits were manually inspected for each receptor in order to identify the most similar human GPCRs from the database. The requirement for being placed in a given group was to have at least four of the five best hits from the specific group. The receptors that did not match these criteria were excluded. After division into groups, receptor sequences were separated in different FASTA files, according to their organism and group identification and then the list was expanded with sequences from the corresponding group of human receptors (later called input FASTA files). This part of the work, from creation of HMM to subdivision in groups was performed according to Fredriksson and Schioth (2005). Then we aligned the predicted receptors from every organism and every group with the corresponding group of human

Table 1
Description of the human *Rhodopsin* GPCRs groups

Description	Group name	Number of receptors
Alpha		
Prostaglandin receptors and orphan receptors SREB 1–3, GPR26, GPR61, GPR62, and GPR78	PTGER	15
Bioamine GPCRs binding 5-hydroxytryptamine, dopamine, histamine, trace amines, adrenalin, and acetylcholine	AMIN	42
Opsin/putative opsin receptors and orphan receptors GPR21 and GPR52	OPN	9
Melatonin and orphan receptor GPR50	MTN	3
Receptors for phospholipids (EDG), melanocortin, cannabinoids, and somatostatin receptors together with three orphan GPCRs	MECA	22
Beta		
Receptors for NPY, tachykinins, neurotensin, orexin, neuromedin, NPFF, PrRP, GnRH, CCK, etc.	PEP	35
Gamma		
Somatostatin, opsin, and galanin receptors	SOG	15
Receptors for melanocyte concentrating hormone	MCH	2
Bradykinin receptors and receptors/putative receptors for chemokines	CHEM	42
Delta		
MRG and MAS receptors	MRG	8
Orphan LGR receptors and receptors for relaxin, FSH, TSH, and LH	LGR	8
Purin/putative purin receptors, formyl-peptide receptors, retinoic acid receptors, and orphan GPCRs	PUR	42

Classification is based on Fredriksson et al. (2003).

receptors using the above-mentioned FASTA files as input data. This was done with ClustalX 1.83 (Thompson et al., 1997) using the default settings. We then used these alignment files to construct Maximum Parsimony (MP) trees using the Mega 2.1 software (Kumar et al., 2001) with pairwise deletion and a bootstrap value of 100 while other settings were default. These MP trees were used to "correct" data in the input FASTA files, which means that we removed duplicate genes and replaced the predicted sequences with already identified gene sequences for previously cloned genes. The test to identify duplicate genes was performed by separate alignment (using ClustalX 1.83) of sequences that were located on the same branch and with bootstrap values exceeding 96. Criteria for identification as a duplicate gene and for shorter sequence exclusion from further analysis was 95% identity in the 7TM as recognized by NCBI Conserved Domain Search v2.03 available on http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml. The replacement of the predicted sequence with the corresponding cloned gene sequence was performed by identifying the cloned sequence in the NCBI GenPept database using BLAST using each predicted protein sequence as bait. Sequences were considered identical if 95% of the amino acid residues of the 7TM region were recognized as identical by the NCBI Conserved Domain Search v2.03. This "correction" of acquired data accounts for the difference between the numbers of receptors in analyzed organisms in this study and the previously published Fredriksson and Schioth (2005). After correcting the data in the input FASTA files, receptors from every organism and every group were again aligned with the corresponding group of human receptors and these alignments were again used for MP tree construction. This was performed using the same software parameters as before. Subsequently, the MP trees were used to determine ancestral receptor groups (ARG). ARG in this study is defined as a group of related GPCRs whose orthologue genes can be found in the closest common ancestor but this term does not reflect the actual ancestral state. The criteria for designation of a group as an ARG required that the human receptor was located in a separate branch with orthologue receptors from all vertebrates (*M. musculus*, *T. rubripes*, *D. rerio*) and together with receptors from at least two invertebrate species (*C. intestinalis, C. savignyi, C. elegans*).

In order to verify, correct and extend the ARG assignment, we performed internal BLAST analysis for invertebrate organism receptors using the BLAST package 2.2.7 (Altschul et al., 1997). During this process, using the application formatdb, we created a number of BLAST databases from the *C. intestinalis*, *C. savignyi*, and *C. elegans* receptors-containing input FASTA files and then we carried out BLAST searches, leaving all settings default, with the application blastall, using the same files as input files. Individual BLAST results were carefully evaluated based on the e-value for each receptor from each BLAST database. In this case the group of receptors was designated as ARG if the following criteria were achieved: (1) presence of orthologue human receptors from the potential ARG in the top results of BLAST analysis within or immediately following the paralogue group of invertebrate receptors. (2) Presence of orthologue invertebrate receptor in the top results of BLAST analysis within or immediately following the paralogue group of human receptors. Each member of human and invertebrate groups of paralogue receptors has to follow these criteria to be joined in the ARG. We believe that these rather rigid criteria provide strong evidence that all the ARG that we identified do indeed have ancestral origin. It is also clear that there can be other ARGs in humans because there is a possibility that some of ARGs might have disappeared during evolution of both *Ciona* and *C. elegans*.

After identification of the ARGs, we calculated the total intron density in the protein coding sequence of human *Rhodopsin* GPCRs and the intron density within the 7TM domain region for both the ARG members and the non-ARG members. This was done by dividing the total length of the nucleotide sequence in kilobases with the total number of introns. Intron positions for invertebrate ARG members were identified manually by BLAT searches against the genomic sequence for the particular organism. For *C. intestinalis*, and *C. elegans*, this was done using the web interface on the http://genome.ucsc.edu/ server, while for *C. savignyi* this was performed with the application BLAT downloaded from http://genome.ucsc.edu/. Intron positions in ARGs were analyzed by aligning *H. sapiens*, *M. musculus*, *C. intestinalis, C. savignyi*, and *C. elegans* receptors from each ARG separately and marking the identified intron positions.

We also performed detailed intron position analysis of the *alpha*, *beta*, and *gamma* groups of the human *Rhodopsin* family GPCR receptors. The *Delta* group was excluded from this analysis because the vast majority of its members were intronless and the only group of receptors that contain introns are those with very large N-terminal tails. This was done by aligning all protein sequences (using ClustalX 1.83) and creating a graph in which every section of the *X*-axis represented one amino acid position and the *Y*-axis represents the number of introns within the receptors.

3. Results

We started our study with downloading protein and mRNA sequences for 240 human *Rhodopsin* GPCRs. 126 (52.5%) of these receptors were intronless in the protein coding region, 29 (7.5%) contained introns located outside the 7TM domain region, and 85 (35.5%) contained one or more introns within the 7TM domain region. Nomenclature of *Rhodopsin* GPCR subgroups is shown in Table 1.

3.1. Receptor extraction and division into groups

We identified 320 possible *Rhodopsin* GPCRs from the mouse (*Mus musculus*) Genscan predictions. After division into groups, 29 of them did not fit into any of the groups defined by Fredriksson et al. (2003) and these were

therefore labeled as unclassified and excluded from further investigation. The remaining 291 predictions were searched against the NCBI GenPept database and replaced with the correct protein sequences of previously identified receptors. This search revealed that only four of the predictions could be referred to as unidentified GPCRs; two of them were located within the purine (PUR) receptor group, and one in the chemokine (CHEM) and mas related gene (MRG) receptor groups.

In zebrafish (*Danio rerio*), we identified 513 possible *Rhodopsin* GPCRs which is 1.6-fold more than in mouse. Division into groups identified 74 predictions that did not fit in any of the groups and these were therefore excluded from further analysis. It is also notable that there was no prediction fitting the MRG group (Table 2). During prediction identification and replacement with correct receptor sequences, we noticed that in many cases one corrected sequence matched two predictions. One of them was usually the N-terminal part of the sequence and the other the C-terminal part. The reason for such great excess in the number of receptors could be the artifacts of the zebrafish Genscan predictions, possibly due to gaps in the genome assembly. This analysis also revealed that only 15% of the predictions were previously identified receptors.

The group distribution in fugu (*Takifugu rubripes*) was similar to that found in zebrafish, but with far less receptors. There were 218 possible *Rhodopsin* GPCRs, 35 of which were referred to as unclassified (excluded), only 20% of the predictions were previously identified as GPCRs and none that could be classified as a MRG receptor.

In the *Ciona intestinalis* GenScan predictions, we identified 152 possible *Rhodopsin* GPCRs; 27 of which were excluded from further analyses as unclassified during the group division. This analysis also showed that *C. intestinalis* probably has no PUR, MRG, and melanocyte concentration hormone (MCH) receptors (Table 2). It is also notable that only seven of the predictions matched previously identified proteins.

In *Ciona savignyi*, we identified only 87 possible *Rhodopsin* GPCRs. Division into groups revealed results similar to *C. intestinalis* with no PUR, MRG, and MCH receptors. 34 of the predictions did not fit into any of the groups and were therefore marked as unclassified and excluded from further analysis. One reason for the small number of predicted receptors in *C. savignyi* is that its genome sequence has not been fully completed yet. None of these predictions turned out to be previously identified proteins.

The analysis of *Caenorhabditis elegans* Genscan predictions gave us 112 possible receptors. We were unable to add 60 of them to any of the 12 defined groups and therefore they were also marked as unclassified and excluded from further analysis. Unlike the other species, only a few of the receptor groups from vertebrates were found in *C. elegans*. These were the biogenic amine and similar receptors (AMIN), the melanocortin, phospholipid binding, cannabinoid and adenosine (MECA), the opsins (OPN), the peptide binding (PEP), and long N-termini GPCR (LGR) receptor groups. Because *C. elegans* is one of most studied model organisms, it was not a surprise that almost all of the predictions, except three in AMIN group, were previously identified receptors. Results of these analyses are summarized in Table 2.

3.2. Ancestral receptor groups and intron position analysis

During our research, we identified 14 ARGs spread throughout the *Rhodopsin* family. All ARGs possess two properties. These groups are small and formed by less than five receptors, with two exceptions: *Prostaglandin* ARG from the *alpha* group and the *Glycoprotein* ARG from the *delta* group, both with eight members. Most of the receptors identified as ARG members, contained at least one intron within their coding sequence, with the exception of the receptors from *Muscarinic* ARG that all are intronless and receptors from *Adrenergic beta* ARG that are mostly intronless. There are seven ARGs within the *alpha* group and this group is thus the largest ARG holder among all main groups. These ARGs are: *Muscarinic* (CHRM), *Adrenergic beta* (ADRB), *Serotonin type two* (HTR2), *Opsin* (OPN), *Melatonin* (MTNR), *Adenosin binding*

Table 2
Number of predictions from each analysed organism identified as members of the *Rhodopsin* GPCR family and the number of previously identified receptors in each group for each organism (adapted from Fredriksson and Schioth (2005) with minor changes)

		Alpha					Beta	Gamma			Delta			Total classified	Unclassified
		AMIN	MECA	MTN	OPN	PTGER	PEP	CHEM	MCH	SOG	LGR	MRG	PUR		
Mus musculus	Predicted	51	22	3	7	13	38	54	1	14	8	30	50	291	29
	Identified	51	22	3	7	13	38	53	1	14	8	29	48	287	
Danio rerio	Predicted	108	33	5	33	15	57	66	4	32	9	0	77	439	74
	Identified	13	11	3	12	3	7	3	3	6	2	0	5	68	
Takifugu rubripes	Predicted	34	15	3	22	13	32	23	2	10	3	0	26	183	35
	Identified	17	6	0	6	0	1	1	2	1	0	0	1	35	
Ciona intestinalis	Predicted	22	17	4	5	2	16	6	0	18	35	0	0	125	27
	Identified	0	1	0	3	0	3	0	0	0	0	0	0	7	
Ciona savignyi	Predicted	12	2	3	1	1	2	2	0	9	21	0	0	53	34
	Identified	0	0	0	0	0	0	0	0	0	0	0	0	0	
Caenorhabditis elegans	Predicted	18	1	0	1	0	31	0	0	0	1	0	0	52	60
	Identified	15	1	0	1	0	31	0	0	0	1	0	0	49	

(ADORA), and *Prostaglandin* (PTGR). The numbers of receptors in these groups generally vary from three to five (as mentioned before, with the exception of the PTGR ARG). Another important fact is that only the *alpha* group contains intronless genes, when considering the ancestral receptor groups. The *alpha* ARGs holds 34% of the total of all *alpha* group members.

Muscarinic ancestral receptor group in humans is the largest within the *alpha* AMIN ARGs. It consists of five human CHRM receptors (CHRM1–5) and was defined as an ARG after identification of homologous groups of receptors in other selected organisms. As mentioned in Section 2, the homologous groups of receptors in other organisms were identified by aligning all human AMIN receptors with possible AMIN receptors from other organisms and subsequently using those alignments to create MP trees. The CHRM receptors were found on a separate distinct branch formed by the receptors from human and the specific organism, thus allowing us to define this group as ARG. Homologous groups of receptors in *M. musculus* consisted of five receptors, while there were nine in *D. rerio*, four in *T. rubripes*, four in *C. intestinalis*, three in *C. savignyi*, and one in *C. elegans* (Table 3). BLAST searches of each AMIN receptor against all human and other organism AMIN receptors confirmed that the selected receptors from other species were closely related to the human CHRM receptors. As previously mentioned, all human and mouse CHRM receptors are intronless within their coding sequence. However, in contrast to the mammals, all of the invertebrate CHRM receptors, except one from *C. intestinalis* and one from *C. savignyi*, contained several introns within their coding sequence. By analyzing the intron positions in the alignment of these receptor protein sequences (Fig. 1a), we identified a pattern of only two introns common for the two *C. intestinalis* and two *C. savignyi* receptors, while other introns are scattered all around in the receptor sequence. Interestingly, the one common intron position is found within the conserved DRY motif.

Adrenergic beta ancestral receptor group was the second *alpha* AMIN ARG that we identified. In humans, this group contains ADRB1, ADRB2, and ADRB3 receptors. From the MP trees, *M. musculus* has three receptors, *D. rerio* five, *T. rubripes* two, and *C. intestinalis* also two receptors. This was not sufficient to identify this group as an ARG, because it did not reach the important criteria of having similar receptors in at least two of the invertebrate organisms. But identification of one receptor from *C. savignyi* using BLAST analysis allowed us to define this group as ARG (Table 3 and Supplementary Figure 2). Similar to the CHRM ARG, the human ADRB ARG receptors also contain intronless receptors. Of the human receptors from this ARG, the ADRB3 is the only receptor with one intron, which separates only the last six C-terminal amino acids from the rest of the protein. Alignment analysis showed that there were not alignable introns to develop a reasonable comparison of the mammal and invertebrate intron positions. The intron positions in both *Ciona* are, however, very similar. Two positions are identical and the third position, within one of the *C. intestinalis* receptors, is close by (Fig. 1b).

Serotonin type two ancestral receptor groups was the third and last *alpha* AMIN ARG. It consists of three human *Serotonin type two* receptors (HTR2A-C). The group was defined after identification of homologous groups of receptors in *M. musculus* (three receptors), *D. rerio* (two receptors), *T. rubripes* (one receptor), *C. intestinalis* (one receptor), and *C. savignyi* (one receptor). The identification of homologous groups of receptors in other organisms was unambiguous since receptors formed separate, easily distinguishable branches. BLAST analysis verified the results and did not provide any additional information for this ARG (Table 3 and Supplementary Figure 2). All human serotonin type two receptors contain a number of introns inside their coding sequences, HTR2A and HTR2B have two introns and HTR2C three introns. The number of introns in each mouse receptor is identical to the human ortholog, while the number of introns in closely related receptors from both the *Cionas* is different, three in *C. intestinalis* and four in *C. savignyi*. The positions of the introns located after the second TM domain and in the middle of the fourth TM domain were exactly the same for all the HTR2 receptors (Fig. 1c).

Opsin ancestral receptor group. During the initial stages of our work, we identified a considerable number of possible *Opsins* in all of the organisms analysed. The OPN group of receptors is well studied because of its evolutionarily conserved function in light detection. The MP trees show a group of only four receptors that could be identified as members of an ARG. The receptors in the OPN ARG are: OPN1LW, OPN1MW, OPN1SW, and RHO in human and

Table 3
Number of predictions and/or previously identified receptors in each Ancestral Receptor Group for each organism

	Alpha							Beta					Gamma	Delta
	CHRM	ADRB	HTR2	OPNR	MTNR	ADORA	PTGR	AVPR	GNRHR	HCRTR	CCKR	TACR	GALR	LGR
Homo Sapiens	5	3	3	4	3	4	8	4	2	2	2	3	3	8
Mus musculus	5	3	3	3	3	4	8	4	1	2	2	3	2	8
Danio rerio	9	5	2	11	5	10	9	4	4	1	3	5	4	9
Takifugu rubripes	4	2	1	5	3	4	7	2	3	1	1	4	3	3
Ciona intestinalis	4	2	1	2	4	3	2	1	3	1	2	1	1	3
Ciona savignyi	3	1	1	1	3	2	1	1	0	1	1	1	1	2
Caenorhabditis elegans	2	0	0	1	0	1	0	2	1	0	2	2	0	1

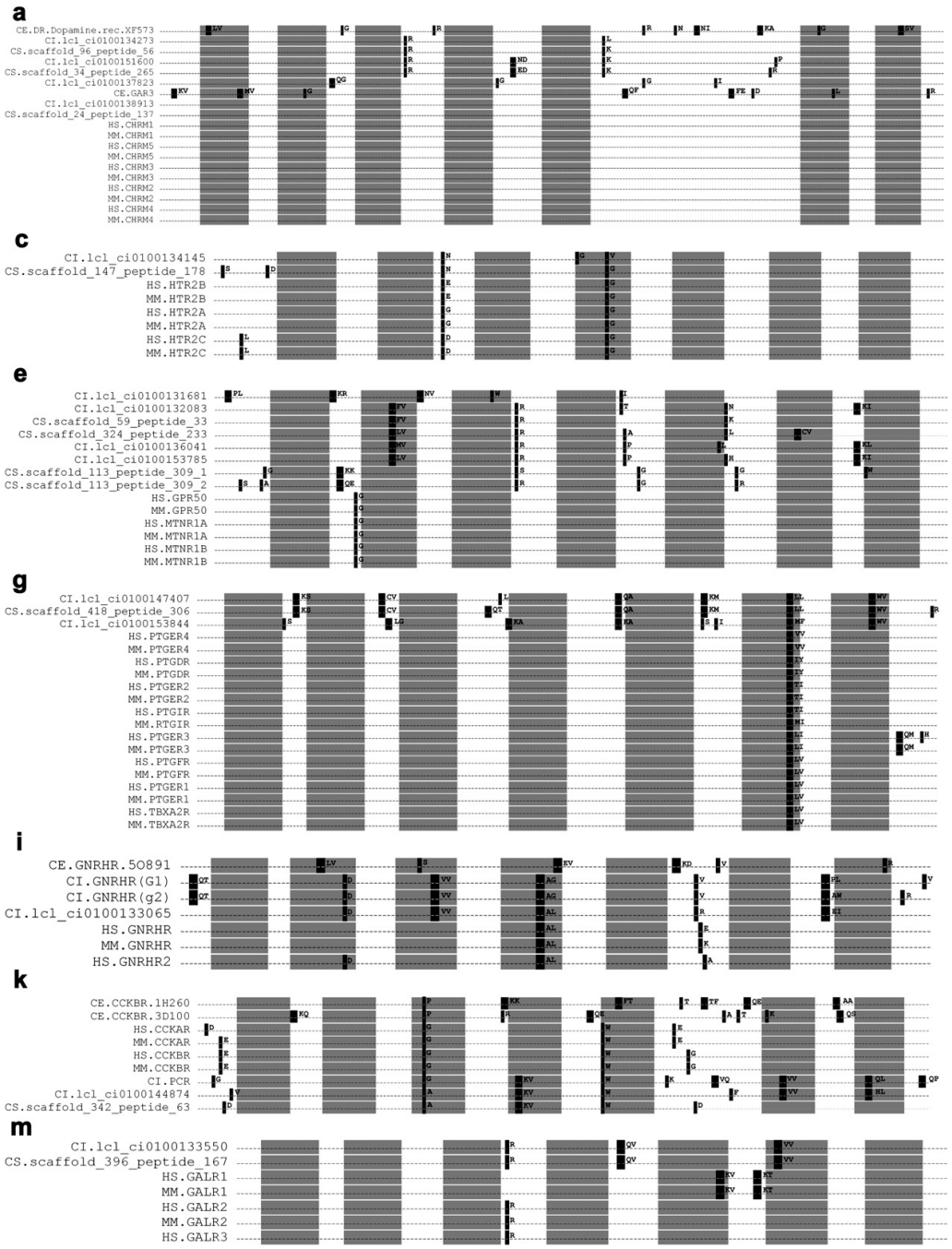


Fig. 1. Schematic presentation of protein sequence alignments for (a) *Muscarinic*, (b) *Adrenergic beta*, (c) *Serotonin type two*, (d) *Opsin*, (e) *Melatonin*, (f) *Adenosin binding*, (g) *Prostaglandin*, (h) *Arginine vasopressin*, (i) *Gonadotropin-releasing hormone*, (j) *Hypocretin*, (k) *Cholecystokinin*, (l) *Tachykinin*, (m) *Galanin*, and (n) *Glycoprotein* Ancestral Receptor Groups, which were used to identify common and possible ancient intron positions for GPCRs in mammalian and invertebrate organisms. Dashed lines represent the receptor amino acid sequence, shaded parts represent TM domains and black rectangles and the letters near them show amino acid residues, whose codons contained introns. If an intron is located between codons of two amino acids, both of them are marked. In cases when the dashed line stops after an intron-representing rectangle, the amino acid sequence of the predicted protein no longer resembled a receptor sequence.

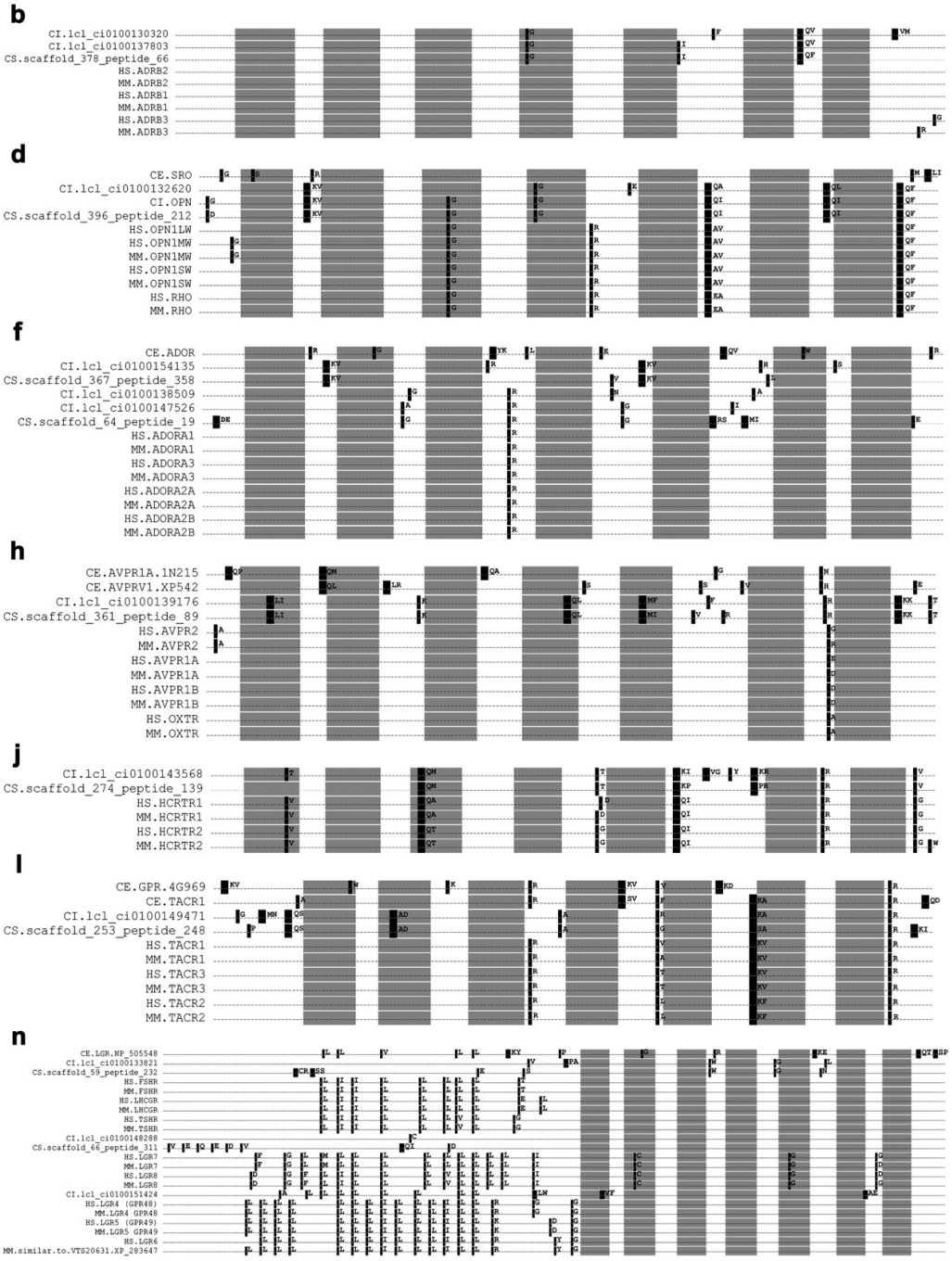


Fig. 1 (*continued*)

the number of previously identified receptors in other organisms were four receptors in *M. musculus*, eleven in *D. rerio*, five in *T. rubripes*, two in *C. intestinalis*, one in *C. savignyi*, and one in *C. elegans*. BLAST analysis revealed that three other *C. intestinalis* receptors; RRH, GPRLike, and lcl|ci0100132620, were more similar to human OPN3, OPN4, and RRH receptors than to GPR21 and GPR52 receptors which suggested that the OPN3, OPN4, and

RRH are also members of the OPN ARG. However, since we did not identify any orthologous or closely related receptors in *C. savignyi*, probably because of the incomplete genomic sequence, these receptors were excluded from the ARG. BLAST analysis also revealed that the *Opsin* in *C. elegans* is very distant from the OPN receptors in other organisms. The human members of the OPN ARG are the most intron rich of the *alpha* ARGs (Table 3 and Supplementary Figure 2). They all have four introns within the protein coding sequence with the exception of OPN1MW which has five introns. The same situation was observed in mouse: the receptors have the same number of introns at the same position. The *C. savignyi* and the *C. intestinalis* OPN receptors contain seven introns and three of them have the same position as the introns in human and mouse: middle of the third TM domain, after the fifth TM domain and after the seventh TM domain. The second receptor from *C. intestinalis* has a similar pattern of introns except it has one extra intron, before the fifth TM domain, and is missing two introns, one in the N-terminal tail and the one in the middle of third TM domain. The introns within the *C. elegans* receptor did not mach any of the introns in the other organisms (Fig. 1d).

Melatonin ancestral receptor group. After division into groups, we were able to select a number of MTN-like receptors for an ARG in all analyzed organisms, except for *C. elegans*. The MTN receptor group in humans is one of the smallest *Rhodopsin* GPCR groups with three receptors: MTNR1A, MTNR1B, and GPR50. MP analysis showed that receptors from invertebrates are relatively distant from their mammalian relatives. However, BLAST analyses showed that these receptors, although distant, very likely share a close origin with the mammalian MTN receptors, thus suggesting that all MTN receptor group should be considered ancestral. The number of identified MTN-like receptors in other organisms were as follows, three in *M. musculus*, *D. rerio*, *T. rubripes*, *C. savignyi*, and four in *C. intestinalis* (Table 3 and Supplementary Figure 2). All human and mouse MTNs contained only one intron inside the protein-coding sequence while predicted receptors from *C. intestinalis* and *C. savignyi* contained a number of introns in various positions. Five of these intron positions, taking in account that some shift has occurred, are common to most of the *Ciona* receptors in different combinations but none of them matches to the position of a single mammalian intron. It is worth to mention that one of the common *Ciona* introns is located within the conserved DRY motif analogue NRY sequence (Fig. 1e).

Adenosin binding ancestral receptor group. Looking at MP trees, we see a group of receptors forming a separate branch common to all organisms. In humans, this group consisted of four ADORA receptors: ADORA1, both ADORA2 (A and B), and ADORA3. There were four similar receptors in *M. musculus*, ten in *D. rerio*, four in *T. rubripes*, three in *C. intestinalis*, two in *C. savignyi*, and one in *C. elegans*. BLAST analysis confirmed that selected receptors in other organisms are closely related to human ADORA receptors (Table 3 and Supplementary Figure 2). Analyzing intron positions in *Ciona* receptors, whose number varied from four to eight, we noticed that, unlike in previously analyzed ARGs, intron positions have very little conservation and form only pairs of identical intron positions. Nevertheless, one of the intron positions in the two receptors from *C. intestinalis* and one from *C. savignyi* were identical to an intron position that was common for all mammalian receptors within this group. Again, one of the *C. intestinalis* receptors (lcl_ci0100154135) contains an intron within the conserved DRY-motif analogue ERY sequence (Fig. 1f).

Prostaglandin ancestral receptor group is one of the largest ARGs and has eight receptors in humans: PTGDR, PTGER1, PTGER2, PTGER3, PTGER4, PTGFR, PTGIR, and TBXA2R. The MP trees indicate that this group is an ARG. The number of homologous receptors in *M. musculus* was the same as in humans, while in other organisms it varies from one in *C. savignyi* to nine in *D. rerio* (seven in *T. rubripes* and two in *C. intestinalis*). The interesting thing about this ARG is that, unlike previous cases where invertebrate receptors, although related, were forming a separate branch, receptors of this group were integrated which indicated that they were very similar to the human and mouse PTGER4 receptor. BLAST analysis confirmed these observations (Table 3 and Supplementary Figure 2). Intron position analysis revealed that most of the human and mouse *Prostaglandin* receptors contained only one identical intron position located at the C-terminal part of the sixth TM domain. Exceptions were both human PTGER3 receptor with one and mouse PTGER3 receptor with two extra introns at the C-terminal tail. All three *Ciona* receptors contained four introns located in the same positions, three introns located at approximately the same position (apparently the result of some intron sliding), and two additional introns in two receptors located at unique positions. Comparing the intron position from invertebrate and mammalian receptors, we identified one intron position that is identical to all analyzed receptors (Fig. 1g).

Analyzing receptors from the *beta* group, we identified the five following ARG: *Arginine vasopressin* (AVPR), *Gonadotropin-releasing hormone* (GNRHR), *Hypocretin* (HCRTR), *Cholecystokinin* (CCKR), and *Tachykinin* (TACR). Among the other *Rhodopsin* GPCR groups, the *beta* group ARGs are characterised with a small number of members. Three of the *beta* ancestral groups; GNRHR, HCRTR, and CCKR, have only two members. In total, this ARG held 38% of the *beta Rhodopsin* GPCR.

Arginine vasopressin ancestral receptor group is the largest in the *beta* group. In humans, it is composed of four receptors: AVPR1A, AVPR1B, AVPR2, and OXTR. Homologous groups of receptors in mouse, *D. rerio*, *T. rubripes* and both species of *Ciona*, were identified during phylogenetic analysis, as usual they formed separate branches. Similar receptors in *C. elegans* were identified through BLAST analysis. In total, we identified four receptors in *M. musculus*, four in *D. rerio*, two in *T. rubripes*, one in *C. intestinalis*, one in *C. savigny*, and two in *C. elegans*

(Table 3 and Supplementary Figure 2). The situation with intron positions in *C. intestinalis* and *C. savigny* is the following: both receptors have approximately the same number of introns and their positions are identical. The exceptions are within the fifth intracellular loop, where positions of two introns from *C. savigny* receptors do not match intron positions from *C. intestinalis* receptor. The two *C. elegans* receptors have a number of introns and only two are located at the same position. The number and positions of the introns in human and mouse receptors were identical: two introns in the AVPR2 receptors and one in other receptors. Comparing intron positions in mammalian and invertebrate receptors, we found one common intron position for all receptors with the position differing in one amino acid towards N-terminal direction between *Ciona* and *C. elegans* and towards C-terminal direction between *Ciona* and mammalian receptors. This observation indicates that these intron positions might have originated from one intron position (Fig. 1h).

Gonadotropin-releasing hormone ancestral receptor group, as noted previously, consists only of two receptors: GNRHR and GANRHR2. The MP trees show homologous groups of receptors in *M. musculus* (one receptor), *D. rerio* (four receptors), *T. rubripes* (three receptors), and *C. intestinalis* (three receptors), but there were no receptors in *C. savignyi* and *C. elegans* that were located on the separate branch along with the human GNRHR receptors. One receptor in *C. elegans* was found with BLAST. The reason for our inability to find GNRHRs in *C. savignyi* might relate to the incompleteness of its genomic sequence (Table 3 and Supplementary Figure 2). Analysis of intron positions in human and mouse receptors revealed one common intron position located within the fourth TM domain, one semi-common intron position located within the fifth intracellular loop, and one intron position that was present only in the human GANRHR2 receptor. The second was considered ambiguous because the intron position in the GNRHR2 receptor is shifted by one amino acid in C-terminal direction. Intron positions in the *C. intestinalis* receptors formed a pattern of five identical intron positions and two extra intron positions at N- and C-terminal tails of GNHR(g1) and GNHR(g2) receptors, respectively, which was not present in the paralogue receptor sequences. The *C. elegans* receptor has six introns. Comparing the intron positions between species, we identified one identical intron position (within the fourth TM domain), one semi-common position (within the fifth intracellular loop) where the intron position in the human GANRHR2 receptor and *C. intestinalis* has been shifted one amino acid to C- and N-terminal direction, respectively, and one that is present only in human GNRHR2 and *C. intestinalis* receptors (Fig. 1i).

Human *Hypocretin ancestral receptor group*, like GNRHR ARG, is very small and consists only of two receptors: HCRTR1 and HCRTR2. Orthologue groups of receptors in *M. musculus* (two receptors), *D. rerio* (one receptor), *T. rubripes* (one receptor), and *C. intestinalis* (one receptor) were identified through phylogenetic analysis. The receptor from *C. savignyi* was identified via BLAST analysis (Table 3 and Supplementary Figure 2). The number of introns within the coding sequences of *C. intestinalis* and *C. savignyi* receptors is nine and six, respectively. Positions of all six found in *C. savignyi* matched the ones found in *C. intestinalis*. However the C-terminal amino acid sequence from the *C. savignyi* receptor was very distinct from the sequences of the same region in other receptors, while the protein sequence from the first intron position to the end is well-conserved, indicating the possible inaccuracy in the gene prediction. Human and mouse HCRTR receptors have six introns, with the exception of the mouse HCRTR2 receptor that contains one extra intron in the C-terminal tail. The third intron in the human HCRTR1 receptor has been shifted by one amino acid in the C-terminal direction. Notably, all six identical intron positions found in human and mouse are also present in both the *Ciona* receptors (Fig. 1j).

Cholecystokinin ancestral receptor group, like the two previous groups in humans, contains two receptors: CCKAR and CCKBR. All receptors similar to *Cholecystokinin* receptors from other organisms were identified by phylogenetic analysis. BLAST analysis verified these results. We found two homologous receptors in *M. musculus*, three in *D. rerio*, one in *T. rubripes*, two in *C. intestinalis*, one in *C. savignyi*, and two in *C. elegans* (Table 3 and Supplementary Figure 2). All human and mouse receptors in this group have four introns. The intron positions found within the third and fifth TM domains were located at identical positions, whereas positions of other introns found within the N-terminal tail and within the fifth-intracellular loop showed apparent position sliding. Receptors obtained from both *Ciona* species contained introns, that varied in number from five to nine. Five of these introns were located at identical positions, and one exhibited obvious signs of intron sliding. It should be noted that the *C. savignyi* receptor has a highly diverged sequence after the last intron compared to other orthologue receptors, probably due to inaccuracy in the gene prediction. Introns in the *C. elegans* receptor have only one identical position. Comparing the intron positions between species, we identified one position matched in all CCKR receptors and one that was present in all receptors except *C. elegans* (Fig. 1k).

Tachykinin ancestral receptor group in humans consists of three receptors: TACR1, TACR2, and TACR3. This group of receptors was defined as an ARG after identification of orthologue receptors in *M. musculus* (three receptors), *D. rerio* (five receptors), *T. rubripes* (four receptors), *C. intestinalis* (one receptor), *C. savignyi* (one receptor), and *C. elegans* (two receptors). All receptors were identified by search through MP trees and BLAST analysis provided only verification of data (Table 3 and Supplementary Figure 2). Analyzing intron positions in both *C. elegans* receptors, we found that although the number of introns they have is similar (seven and eight), their positions are different. In total, only four intron

positions for these receptors were identical. The intron positions in the *Ciona* receptors were, in most cases, identical but only differed in the N- and C-terminal tails. Human and mouse TAC receptors have four introns, whose positions matched and were present also in invertebrate organisms. It is worth noticing that one of the positions, located within a conserved DRY motif, is present in all human, mouse and *C. elegans* receptors but not in *Ciona*. On the other hand, an intron position located at the end of fifth-intracellular loop was found in all receptors except *C. elegans* GPR receptor (Fig. 1l).

Galanin ancestral receptor group. Within the *Rhodopsin* family *gamma* group, we identified only one ARG. It consisted of three *Galanin* (GALR) receptors: GALR1, GALR2, and GALR3 (5% of *gamma* group members). Orthologue receptors in *M. musculus* (two receptors), *D. rerio* (four receptors), *T. rubripes* (three receptors), and *C. savignyi* (one receptor) were identified by MP analysis. Only one receptor from *C. intestinalis* was identified using BLAST analysis (Table 3 and Supplementary Figure 2). As we can see in Fig. 1m, introns in GALR ARG receptors form three groups with different intron positions. The first group consists of receptors found in *C. savignyi* and *C. intestinalis*, they have three introns located within the conserved DRY motif, the fourth extracellular loop and the sixth TM domain. The second group consists of human and mouse GAL1 receptors with two introns located within the fifth TM domain and the fifth intracellular loop. The third group is formed by human and mouse GAL2 and human GAL3 receptors and these receptors have only one intron located within the conserved DRY motif (Fig. 1m).

Glycoprotein ancestral receptor group. The *delta* group had, like the *gamma* group, only one relatively large ARG. *Glycoprotein* (LGR) ARG in humans consists of eight members: LGR4, LGR5, LGR6, LGR7, LGR8, FSHR, TSHR, and LHCGR. Our analysis revealed a group of eight homologous receptors in *M. musculus*, nine in *D. rerio*, three in *T. rubripes*, three in *C. intestinalis*, two in *C. savignyi*, and one receptor in *C. elegans*. All receptors, except one from *C. savignyi*, were identified already at the stage of phylogenetic analysis. The remaining receptor was found through BLAST (Table 3 and Supplementary Figure 2). This is the intron-richest ARG, the number of introns in human and mouse receptors varied from nine, in FSHR and TSHR through ten in LHCGR and up to seventeen in LGR4, LGR5, LGR7, LGR8, and mouse LGR6 receptors. The human LGR6 receptor has 16 introns. The number of introns in the invertebrate receptors is unique for every receptor and varies from one to twelve. One other interesting thing about LGR receptors is that they have extremely long N-terminal tails where most of the human and mouse receptors have introns. We were able to identify only one receptor from *C. intestinalis* that had a similar pattern of intron positions, seven of them matched with intron positions in human receptors (Fig. 1n). The results obtained from the analyses mentioned above are summarized in Table 3, Fig. 1, and Supplementary Figure 2. It should be mentioned here that there could be more undetected ARGs in the analysed pool of genes since we base our analysis on a relatively small number of ancient species (*Ciona* and *C. elegans*) in which some receptors from ARGs might have disappeared only in these lineages.

3.3. Intron density analysis

After we identified the ARGs, we calculated total intron density within the entire coding sequence and within the 7TM region for all human *Rhodopsin* GPCRs, *Rhodopsin* GPCR ARG receptors, and non-ARG *Rhodopsin* GPCRs. Calculations showed that the intron density in ARG receptors is 3.5-fold higher than in non-ARG receptors when calculated for the entire coding sequence and 2.3-fold higher when calculated for the 7TM region. Results from this analysis are summarized in supplementary Table 1.

4. Discussion

It is well established that vertebrate GPCR genes are characterised by a large proportion of intronless genes and a lower density of introns compared to invertebrate GPCRs (Brosius, 1999; Bryson-Richardson et al., 2004; Gentles and Karlin, 1999). None of the previous studies have, however, provided detailed analysis of the origin of these introns, but instead have mainly looked at the numbers within different species. It is important to differentiate between the different phylogenetic branches because, although the GPCR family consists of structurally similar proteins, this family is still very heterogeneous with respect to evolutionary origin. The recent development of a new evolutionary-based classification system for the GPCR family (Fredriksson et al., 2003) and the increased availability of full genome sequences from a number of organisms has made it possible to perform more detailed studies on this topic. In this study, a combination of HMM and BLAST searches provided a fairly comprehensive repertoire of the GPCRs, given the fact that we used large and complete seeding sets of GPCRs (Fredriksson and Schioth, 2005). The choice of genomes used in our study was defined for two reasons. First, we intended to identify only those ARGs that humans had possibly retained during evolution from its ancestors at the base of the *Deuterostomes* lineage (this included the *C. elegans* lineage which, according to traditional phylogeny of the species of animals, separated before the split of *Deuterostomes* and *Protostomes*, and explains exclusions of unclassified receptors from the analysis). This also explains exclusion of insect species from our analysis. And, second, most of the genomes that we studied are almost completely sequenced which has made it possible to track each phylogenetic branch and include virtually all receptors and introns within the *Rhodopsin* GPCRs. This work has identified numerous new sequences as GPCRs and we present alignment and phylogenetic analyses of 1402 receptors from seven species, to identify the

receptors that are likely to have ancestral origin (belonging to an ARG) and which introns are found in conserved positions.

Interestingly, we found that there exists a remarkable difference in the intron density among ancestral and new *Rhodopsin* GPCRs. The intron density within ARG members is more than 3.5-fold higher than within non-ARG members and more than 2-fold higher if we look only at the 7TM region. This difference is even more notable since some other receptor groups with introns (NPFF and OPR for example) are very likely to be ancestral, but were not included in the ARG since they did not pass all the criteria described in Materials and Methods. This suggests that the new GPCR genes have predominantly been formed intronless while the ancestral receptors are likely to have accumulated introns during the evolution. It seems apparent that most of the groups of receptors that have undergone very recent expansion, such as the MRG (found only in mammals), PUR (found only in vertebrates), CHEM (found several receptors in *Ciona*, but distant from vertebrate receptors), and olfactory (found only in vertebrates) receptors are mostly intronless.

We suggest that the high number of new, intronless, GPCR genes can be explained by gene duplication through an RNA intermediate, since this is the only mechanism of gene duplication that results in formation of intronless genes from intron-containing genes (Brosius, 1999; Gentles and Karlin, 1999). It is also notable that we find many intronless GPCR genes located very close to each other in the genome, for example CCR1, CCR3, CCR5, CCR2, CXCR5, and CCR9 from the CHEM group, MRG receptors in mouse (Zylka et al., 2003), the trace amine receptors in zebrafish (Gloriam et al., 2005), and the olfactory receptors (Newman and Trask, 2003) and these are therefore probably duplicated through local, DNA-based, duplication mechanisms. It is thus possible that mRNA-based duplication has played a very important role in the formation of new receptor genes while DNA-based duplications may have multiplied these intronless genes through local and perhaps, in some cases, block- or genome-wide duplications (Fig. 2).

Besides intron density analysis, we also compared intron positions within receptors of ARG from the mammalian and invertebrate species to identify those mammalian introns that seemed to be ancient. Three important findings were identified as result of this analysis. First, we noticed that only a few of *Ciona* and none of *C. elegans* receptors have less than three introns while a large proportion of mammalian receptors contain less than three introns and a number are intronless genes. Detailed calculations confirmed that there is a high difference in intron density between invertebrate and mammalian receptors. We found that the density was 4.49 Intron/Kb and 1.55 Intron/Kb within the 7TM region of the invertebrate (5.54 Intron/Kb for *C. elegans*, 4.70 Intron/Kb for *C. intestinalis*, and 3.59 Intron/Kb for *C. savignyi*) and mammalian (1.58 Intron/Kb for human and 1.52 Intron/Kb for *M. musculus*) ARG receptors, respectively. Second, most of the intron positions found in mammalian ARG receptor sequences were also present in orthologue invertebrate receptors (mostly in *Ciona*), indicating that these intron positions are indeed ancient. And, third, most of the receptor genes had highly similar if not identical intron position patterns within a group of paralogue receptors in each species, either in humans or *Ciona*. This phenomenon was observed in every paralogous group consisting of more than one receptor in both humans and *Ciona*.

The differences in intron density between orthologue mammalian and *Ciona* receptors could be explained by extensive intron loss in mammalian genes or intensive intron gain in C*iona* genes after the split of these two lineages. However, that does not explain the highly similar intron position pattern within groups of paralogue receptors since the intron loss and gain process is supposed to be at random. One of the explanations could be that they result from recent DNA-based gene duplications. To verify this possibility, we collected data about the chromosomal localization for all ARG members and identified those located in close proximity to each other on the same chromosome (Table 4). In doing this, we found six of them: CHRM1/CHRM4, OPN1LW/OPN1MW, ADORA1/

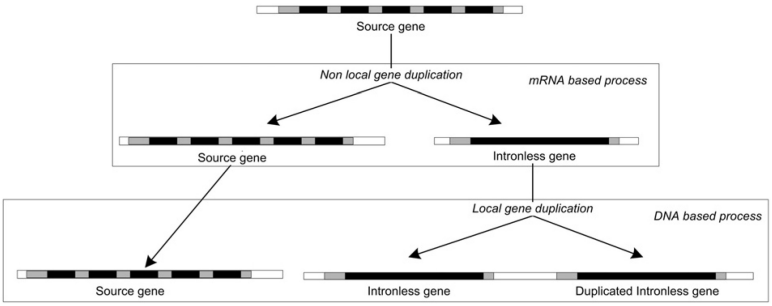


Fig. 2. Schematic presentation of the hypothetical chain of events that could have resulted in the extensive formation of intronless GPCR genes. The first event is the formation of a new intronless gene through an mRNA intermediate. Next, the newly formed gene is duplicated by other local (shown in figure) and genome-wide duplication mechanisms. White rectangles represent genomic DNA, grey rectangles represent non-coding parts of the gene (5′ and 3′ untranslatable regions and introns) and black rectangles represent protein coding sequences.

Table 4
Chromosomal location of human Ancestral Receptor Group members

Receptor	Location	Receptor	Location	Receptor	Location	Receptor	Location	Receptor	Location
CHRM1	***11q13***	OPN1LW	***Xq28***	PTGDR	14q22.1	GNRHR	4q21.2	GALR1	18q23
CHRM2	7q31–q35	OPN1MW	***Xq28***	PTGER1	***19p13.1***	GNRHR2	1q12	GALR2	17q25.3
CHRM3	1q41–q44	OPN1SW	7q31.3–q32	PTGER2	14q22	HCRTR1	1p33	GALR3	22q13.1
CHRM4	***11p12–p11.2***	RHO	3q21–q24	PTGER3	1p31.2	HCRTR2	6p11-q11	LGR4	11p14–p13
CHRM5	15q26	MTNR1A	4q35.1	PTGER4	5p13.1	CCKAR	4p15.1–p15.2	LGR5	12q22–q23
ADRB1	10q24–q26	MTNR1B	11q21–q22	PTGFR	1p31.1	CCKBR	11p15.4	LGR6	1q32.1
ADRB2	5q31–q32	GPR50	Xq28	PTGIR	***19q13.3***	AVPR1A	12q14–q15	LGR7	4q32.1
ADRB3	8p12–p11.2	ADORA1	***1q32.1***	TBXA2R	***19p13.3***	AVPR1B	1q32	LGR8	13q13.1
HTR2A	13q14–q21	ADORA2A	22q11.23	TACR1	2p12	AVPR2	Xq28	FSHR	***2p21–p16***
HTR2B	2q36.3–q37.1	ADORA2B	17p12–p11.2	TACR2	10q11–q21	OXTR	3p25	TSHR	14q31
HTR2C	Xq24	ADORA3	***1p21–p13***	TACR3	4q25			LHCGR	***2p21***

Receptors from one closely located Ancestral Receptor Group are marked in bold italic.

ADORA3, PTGER1/PTGIR/TBXA2R, and FSHR/LHCGR (Source gene/descendant gene or contrariwise), which indeed were probably formed recently through local DNA-based gene duplication mechanisms. Although a number of genes may have lost synteny because of chromosome rearrangements, we still find a large number of paralogue receptors with similar intron pattern randomly distributed in the genome, suggesting that other gene duplication mechanisms are likely involved in this process.

Some of the possible mechanisms explaining similar intron positions among paralogue genes that occur by gene duplication and subsequent conversion process are illustrated in Fig. 3. We chose duplication through an RNA intermediate because of its demonstrativeness. Alternatively, duplication may occur through any DNA-based mechanisms with subsequent gain of additional intron. It is clear that one of the duplicated genes can undergo the evolutionary processes of gaining new, losing old and changing positions (intron sliding) of existing introns, which would be a rather slow process. However, if duplication of this gene occurs, especially through RNA-based mechanisms, the new gene, although with identical coding sequence, has a different gene structure. It consequently increases the likelihood of replacing the intronless region in the new gene with an intron-containing region from the source gene, or in replacing the intron-containing region in the source gene with an intronless region from the new gene by a homologous sequence exchange mechanism (Roy, 2004) such as gene conversion, a process that homogenizes paralogous sequences retarding their divergence, and consequently obscuring their antiquity (Hurles, 2004). This process continues until both genes acquire similar intron–exon structure or until their coding sequence, through mutations, becomes too distinct for homologous recombination to take place. This implies that the more intronless genes are formed, the higher the possibility that the source gene will also lose its introns, like for the CHRM and ADRB receptors which all are intronless. And, on the contrary, if the new intronless gene appears in a family of highly similar genes with high intron density, it will most likely acquire the family characteristic introns.

We also performed detailed intron position analysis within the *alpha*, *beta* and *gamma* group of the human *Rhodopsin* family (Fig. 4). This analysis shows one common intron position that was specific for receptors from a number of different *Rhodopsin* GPCR groups. This intron position lies within a conserved DRY motif (usually found in the end of the third TM domain or at the beginning of the third-intracellular loop). All of the other positions, in which intron number was equal or greater than four, turned out to be group-specific which suggests that the introns in these positions were typical for one or two groups of closely related receptors and also present in one or two receptors from other groups. Careful tracking of receptors-containing introns within the conserved DRY motif showed that they are spread all around the *Rhodopsin* family (Supplementary Figure 2). There were five such receptors within the AMIN, ten within the PEP, two within the GAL, one within the MCH, and one within the PUR groups. Our results show that this intron position is by far the most common among GPCRs. The presence of this intron position in receptors in five (CHRM, MTNR, ADRB, TACR, and GALR) *C. intestinalis*, *C. savignyi*, and *C. elegans* ARGs indicates that this intron position could be rather ancient, as suggested by Bryson-Richardson and colleagues (Bryson-Richardson et al., 2004).

However, taking into account the fact that at least two of these introns, found in *T. rubripes* melanocortin type 2 and 5 receptors, are probably inserted recently (Klovins et al., 2004; Schioth et al., 2005) and taking notice that it is also found in the RE2 receptors which is the only DRY-intron-containing receptor in the distant *delta* group, it is very likely that some of these introns might have been recently inserted. This suggests that there exists a mechanism promoting the insertion of introns in this particular position. It could be speculated that introns from highly conserved DRY-motif-containing receptors might be transferred to other distinct receptors through previously mentioned sequence exchange mechanisms. Looking at the 3D structures and the functional importance of the DRY and surrounding regions, it is well established that this is one of the key regions for signal transduction mechanisms of GPCRs in general and it is believed that the DRY motif plays a key role in preventing

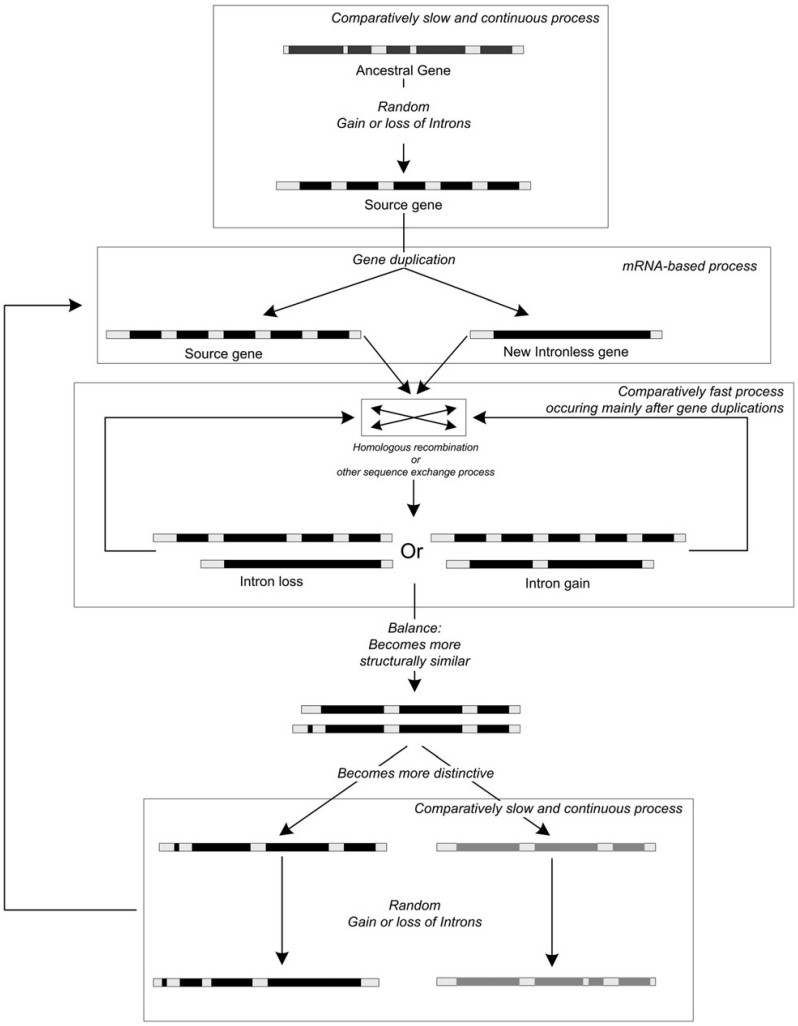


Fig. 3. Schematic presentation of the hypothetical chain of events that could have affected the intron density of GPCRs. The first box illustrates the process of gene evolution (coding sequence mutation and sliding, gain and loss of introns). This step is followed by gene duplication (in this case we have used mRNA-based for illustrative purposes) and homologous recombination-based exchange of intronless/intron-containing regions. This results in intron loss in the source and gain in the new gene. This process halts after the genes have become too distinct for exchange. White rectangles represent non-coding parts of the gene (5′ and 3′ untranslatable regions and introns) and black and grey rectangles represent protein-coding sequences.

constitutive activity of GPCRs (Costa and Cotecchia, 2005; Haitina et al., 2004; Lagerstrom et al., 2003). The DRY motif and its surrounding regions are therefore very well conserved in GPCRs and variation in coding DNA sequence is limited. The similarity of these regions may increase the likelihood of sequence transfer between two different receptors. Another possible mechanism is based on the observation that nucleotide sequence coding DRY motif may form a proto-splice site (MAG|R where M represents A or C, R represents A or G, and vertical line | represents the position of the intron), which is believed to be target site for insertion of spliceosomal introns. This theory was introduced by Dibb and Newman in 1989 (Dibb and Newman, 1989). Interestingly, we have found that other introns from GPCRs are often located between the second and third nucleotide of the arginine codon within motifs, similar to DRY (Supplementary Figure 2). To verify if corresponding DNA sequences of these motifs actually resemble a proto-splice site, we extracted them along with six flanking nucleotides from both sides and produced two alignments for DRY-motif (located between TMIII and TMIV) and other DRY-like motifs within the receptor sequences, respectively (Table 5). The consensus sequences are in both cases identical to the proto-splice sequence, further supporting the suggestion that the proto-splice

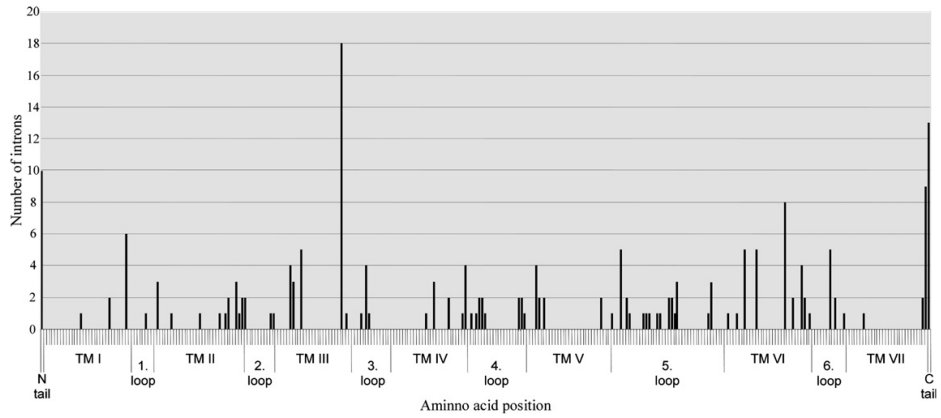


Fig. 4. Graph representing the total number of introns in every amino acid position of *Rhodopsin* family GPCRs, alpha, beta, and gamma group receptors. Minor divisions of the *X*-axis represent amino acid position and major divisions represent the region of receptor. The *Y*-axis indicates the number of introns.

Table 5
List of all *Rhodopsin* GPCRs that contained introns within the conserved DRY motif or within DRY related motif, the motifs coding sequence along with six flanking nucleotides nd consensus sequences

Receptor	DNA Sequence	Motif sequence	Receptor	DNA Sequence	Motif sequence
ADORA1	atccctctccg|gtacaagatg	LRY	NPFF1	gctgtggaaag|gttccgctgc	ERF
ADORA2A	atcccgctccg|gtacaatggc	LRY	NPFF2	gctgtagatag|gttccagtgt	DRF
ADORA2B	gtcccgctcag|gtataaaagt	LRY	NMBR	agcgccgacag|gtacagagcc	DRY
ADORA3	cttaccgtcag|atacaagagg	VRY	GRPR	tcggcagacag|atacaaagcc	DRY
DRD2	agcatcgacag|gtacacagct	DRY	BRS3	agcgctgacag|atacaaggca	DRY
DRD3	agcatagacag|gtacactgca	DRY	EDNRA	agtgttgacag|gtacagagca	DRY
DRD4	agcgtggacag|gttcgtggcc	DRF	EDNRB	agtattgacag|atatcgagct	DRY
OPN1LW	ggttggagcag|gtactggccc	SRY	OPRD1	ggcatcgtccg|gtacactaag	VRY
OPN1MW	ggttggagcag|gtactggccc	SRY	OPRK1	gtgatcatccg|atacacaaag	IRY
OPN1SW	ggctggagccg|gttcatccct	SRF	OPRM1	gtgattgtcag|atacaccaag	VRY
RHO	ggctggtccag|gtacatcccc	SRY	GALR2	tccctggacag|gtatctggcc	DRY
RRH	aaaaatgatag|atcttttgtg	DRS	GALR3	tccgtggacag|gtacctggcc	DRY
HTR4	tctctggatag|gtattacgcc	DRY	MCHR2	agtgtggacag|gtactttgcc	DRY
HTR7	agcattgacag|gtaccttggg	DRY	RGR	aaggggacag|aaacttcacc	DRN
TACR1	gcctttgatag|gtacatggcc	DRY	RE2	gccatcgaccg|ctactatgct	DRY
TACR2	gctgccgacag|gtacatggcc	DRY	TACR1v1	ctcaatgacag|gttccgtctg	DRF
TACR3	gcggtggacag|gtatatggct	DRY	NPFF2v1	cgccgggagag|gttcatcatg	ERF
Consensus 1:	G--GTGGACAG|GTACA-GGC-		Consensus 2:	GGC-GGG-CAG|GTACA--A-G	
Proto splice sitie:	MAG|R		Proto splice sitie:	MAG|R	

Consensus 1 is for all motifs, Consensus 2 is only for DRY related motifs. M represents A or C, R represents A or G. The DRY related motifs are shaded.

mechanisms are likely to have played an important role for insertion of new introns, at least in *Rhodopsin* GPCR family.

In summary, through this extensive study we identified groups of invertebrate receptors, which were highly similar to certain groups of human receptors, which subsequently were labeled as ARGs. We found that intron density among the human ARG members was considerably lower than in corresponding invertebrate receptors, and that most of the introns found among the human ARG members, including the one located in a conservative DRY motif, were present in the invertebrate receptors. It is unlikely that there has been major loss of introns among vertebrate GPCRs while formation of new GPCRs among vertebrates explains why these have fewer introns compared with invertebrate GPCRs. We speculate over the mechanisms that may explain these phenomena and the sequence of possible events. Moreover, we discuss the mechanisms that could explain the recent cases of intron insertion in the conservative DRY motif.

Acknowledgments

We thank Dr. Chris Pickering, Uppsala University for correcting the language of the Manuscript. Dr. Janis Klovins was supported by the European Union Marie Curie

Fellowship programme and by OE och Edla Johanssons Vetenskapliga Stiftelse and Davids Fridmanis was supported by FEBS Collaborative Experimental Scholarship for Central and Eastern Europe while working in Department of Neuroscience, Uppsala University and by The European Social Fund while working in Biomedical Research and Study Centre, University of Latvia and Dr. Robert Fredriksson by the Swedish Brain Research Foundation. The studies were supported by the Swedish Research Council (VR, medicine), The Novo Nordisk Foundation and the Magnus Bergwall Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2006.11.007.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Berget, S.M., Moore, C., Sharp, P.A., 1977. Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. Proc. Natl. Acad. Sci. USA 74, 3171–3175.

Bjarnadottir, T.K., Gloriam, D.E., Hellstrand, S.H., Kristiansson, H., Fredriksson, R., Schioth, H.B., 2006. Comprehensive repertoire and phylogenetic analysis of the G protein-coupled receptors in human and mouse. Genomics 88, 263–273.

Brosius, J., 1999. Many G-protein-coupled receptors are encoded by retrogenes. Trends Genet. 15, 304–305.

Brosius, J., 2003. Gene duplication and other evolutionary strategies: from the RNA world to the future. J. Struct. Funct. Genomics 3, 1–17.

Bryson-Richardson, R.J., Logan, D.W., Currie, P.D., Jackson, I.J., 2004. Large-scale analysis of gene structure in rhodopsin-like GPCRs: evidence for widespread loss of an ancient intron. Gene 338, 15–23.

Burge, C., Karlin, S., 1997. Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. 268, 78–94.

Cavalier-Smith, T., 1985. Selfish DNA and the origin of introns. Nature 315, 283–284.

Cavalier-Smith, T., 1991. Intron phylogeny: a new hypothesis. Trends Genet. 7, 145–148.

Cho, G., Doolittle, R.F., 1997. Intron distribution in ancient paralogs supports random insertion and not random loss. J. Mol. Evol. 44, 573–584.

Chow, L.T., Gelinas, R.E., Broker, T.R., Roberts, R.J., 1977. An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. Cell 12, 1–8.

Costa, T., Cotecchia, S., 2005. Historical review: Negative efficacy and the constitutive activity of G-protein-coupled receptors. Trends Pharmacol. Sci. 26, 618–624.

Danielson, P.B., Dores, R.M., 1999. Molecular evolution of the opioid/orphanin gene family. Gen. Comp. Endocrinol. 113, 169–186.

Darnell Jr., J.E., 1978. Implications of RNA–RNA splicing in evolution of eukaryotic cells. Science 202, 1257–1260.

de Souza, S.J., Long, M., Klein, R.J., Roy, S., Lin, S., Gilbert, W., 1998. Toward a resolution of the introns early/late debate: only phase zero introns are correlated with the structure of ancient proteins. Proc. Natl. Acad. Sci. USA 95, 5094–5099.

de Souza, S.J., Long, M., Schoenbach, L., Roy, S.W., Gilbert, W., 1996. Intron positions correlate with module boundaries in ancient proteins. Proc. Natl. Acad. Sci. USA 93, 14632–14636.

de Souza, S.J., Long, M., Schoenbach, L., Roy, S.W., Gilbert, W., 1997. The correlation between introns and the three-dimensional structure of proteins. Gene 205, 141–144.

Dibb, N.J., 1993. Why do genes have introns? FEBS Lett. 325, 135–139.

Dibb, N.J., Newman, A.J., 1989. Evidence that introns arose at proto-splice sites. EMBO J. 8, 2015–2021.

Doolittle, W.F., 1978. Genes in pieces: Were they ever together? Nature 272, 581–582.

Eddy, S.R., 1998. Profile hidden Markov models. Bioinformatics 14, 755–763.

Fedorov, A., Cao, X., Saxonov, S., de Souza, S.J., Roy, S.W., Gilbert, W., 2001. Intron distribution difference for 276 ancient and 131 modern genes suggests the existence of ancient introns. Proc. Natl. Acad. Sci. USA 98, 13177–13182.

Fedorov, A., Roy, S., Fedorova, L., Gilbert, W., 2003. Mystery of intron gain. Genome. Res. 13, 2236–2241.

Fedorova, L., Fedorov, A., 2003. Introns in gene evolution. Genetica 118, 123–131.

Figueroa, F., Ono, H., Tichy, H., O'Huigin, C., Klein, J., 1995. Evidence for insertion of a new intron into an Mhc gene of perch-like fish. Proc. Biol. Sci. 259, 325–330.

Fredriksson, R., Lagerstrom, M.C., Lundin, L.G., Schioth, H.B., 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol. Pharmacol. 63, 1256–1272.

Fredriksson, R., Schioth, H.B., 2005. The repertoire of G-protein-coupled receptors in fully sequenced genomes. Mol. Pharmacol. 67, 1414–1425.

Gentles, A.J., Karlin, S., 1999. Why are human G-protein-coupled receptors predominantly intronless? Trends Genet. 15, 47–49.

Gilbert, W., 1987. The exon theory of genes. Cold Spring Harb. Symp. Quant. Biol. 52, 901–905.

Gilbert, W., de Souza, S.J., Long, M., 1997. Origin of genes. Proc. Natl. Acad. Sci. USA 94, 7698–7703.

Giroux, M.J., Clancy, M., Baier, J., Ingham, L., McCarty, D., Hannah, L.C., 1994. De novo synthesis of an intron by the maize transposable element dissociation. Proc. Natl. Acad. Sci. USA 91, 12150–12154.

Gloriam, D.E., Bjarnadottir, T.K., Yan, Y.L., Postlethwait, J.H., Schioth, H.B., Fredriksson, R., 2005. The repertoire of trace amine G-protein-coupled receptors: large expansion in zebrafish. Mol. Phylogenet. Evol. 35, 470–482.

Haitina, T., Klovins, J., Andersson, J., Fredriksson, R., Lagerstrom, M.C., Larhammar, D., Larson, E.T., Schioth, H.B., 2004. Cloning, tissue distribution, pharmacology and three-dimensional modelling of melanocortin receptors 4 and 5 in rainbow trout suggest close evolutionary relationship of these subtypes. Biochem. J. 380, 475–486.

Hankeln, T., Friedl, H., Ebersberger, I., Martin, J., Schmidt, E.R., 1997. A variable intron distribution in globin genes of Chironomus: evidence for recent intron gain. Gene 205, 151–160.

Hofmann, K., Stoffel, W., 1993. TMbase—A database of membrane spanning proteins segments. Biol. Chem. Hoppe-Seyler 374, 166.

Hurles, M., 2004. Gene duplication: the genomic trade in spare parts. PLoS Biol. 2, E206.

Jeffreys, A.J., Flavell, R.A., 1977. The rabbit beta-globin gene contains a large insert in the coding sequence. Cell 12, 1097–1108.

Kent, W.J., 2002. BLAT—the BLAST-like alignment tool. Genome Res. 12, 656–664.

Klovins, J., Haitina, T., Fridmanis, D., Kilianova, Z., Kapa, I., Fredriksson, R., Gallo-Payet, N., Schioth, H.B., 2004. The melanocortin system in Fugu: determination of POMC/AGRP/MCR gene repertoire and synteny, as well as pharmacology and anatomical distribution of the MCRs. Mol. Biol. Evol. 21, 563–579.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17, 1244–1245.

Lagerstrom, M.C., Hellstrom, A.R., Gloriam, D.E., Larsson, T.P., Schioth, H.B., Fredriksson, R., 2006. The G protein-coupled receptor subset of the chicken genome. PLoS Comput. Biol. 2, e54.

Lagerstrom, M.C., Klovins, J., Fredriksson, R., Fridmanis, D., Haitina, T., Ling, M.K., Berglund, M.M., Schioth, H.B., 2003. High affinity

agonistic metal ion binding sites within the melanocortin 4 receptor illustrate conformational change of transmembrane region 3. J. Biol. Chem. 278, 51521–51526.

Lee, V.D., Stapleton, M., Huang, B., 1991. Genomic structure of *Chlamydomonas caltractin*. Evidence for intron insertion suggests a probable genealogy for the EF-hand superfamily of proteins. J. Mol. Biol. 221, 175–191.

Lespinet, O., Wolf, Y.I., Koonin, E.V., Aravind, L., 2002. The role of lineage-specific gene family expansion in the evolution of eukaryotes. Genome Res. 12, 1048–1059.

Logsdon Jr., J.M., Stoltzfus, A., Doolittle, W.F., 1998. Molecular evolution: recent cases of spliceosomal intron gain? Curr. Biol. 8, R560–R563.

Long, M., de Souza, S.J., Rosenberg, C., Gilbert, W., 1998. Relationship between "proto-splice sites" and intron phases: evidence from dicodon analysis. Proc. Natl. Acad. Sci. USA 95, 219–223.

Long, M., Rosenberg, C., 2000. Testing the "proto-splice sites" model of intron origin: evidence from analysis of intron phase correlations. Mol. Biol. Evol. 17, 1789–1796.

Newman, T., Trask, B.J., 2003. Complex evolution of 7E olfactory receptor genes in segmental duplications. Genome Res. 13, 781–793.

O'Neill, R.J., Brennan, F.E., Delbridge, M.L., Crozier, R.H., Graves, J.A., 1998. De novo insertion of an intron into the mammalian sex determining gene, SRY. Proc. Natl. Acad. Sci. USA 95, 1653–1657.

Palmer, J.D., Logsdon Jr., J.M., 1991. The recent origins of introns. Curr. Opin. Genet Dev. 1, 470–477.

Pierce, K.L., Premont, R.T., Lefkowitz, R.J., 2002. Seven-transmembrane receptors. Nat. Rev. Mol. Cell Biol. 3, 639–650.

Roy, S.W., 2004. The origin of recent introns: transposons? Genome. Biol. 5, 251.

Roy, S.W., Fedorov, A., Gilbert, W., 2002. The signal of ancient introns is obscured by intron density and homolog number. Proc. Natl. Acad. Sci. USA 99, 15513–15517.

Roy, S.W., Fedorov, A., Gilbert, W., 2003. Large-scale comparison of intron positions in mammalian genes shows intron loss but no gain. Proc. Natl. Acad. Sci. USA 100, 7158–7162.

Roy, S.W., Gilbert, W., 2005a. Complex early genes. Proc. Natl. Acad. Sci. USA 102, 1986–1991.

Roy, S.W., Gilbert, W., 2005b. The pattern of intron loss. Proc. Natl. Acad. Sci. USA 102, 713–718.

Roy, S.W., Gilbert, W., 2005c. Rates of intron loss and gain: implications for early eukaryotic evolution. Proc. Natl. Acad. Sci. USA 102, 5773–5778.

Roy, S.W., Nosaka, M., de Souza, S.J., Gilbert, W., 1999. Centripetal modules and ancient introns. Gene 238, 85–91.

Sadusky, T., Newman, A.J., Dibb, N.J., 2004. Exon junction sequences as cryptic splice sites: implications for intron origin. Curr. Biol. 14, 505–509.

Sakharkar, M.K., Chow, V.T., Chaturvedi, I., Mathura, V.S., Shapshak, P., Kangueane, P., 2004. A report on single exon genes (SEG) in eukaryotes. Front Biosci. 9, 3262–3267.

Schioth, H.B., Haitina, T., Fridmanis, D., Klovins, J., 2005. Unusual genomic structure: melanocortin receptors in fugu. Ann. NY Acad. Sci. 1040, 460–463.

Tarrio, R., Rodriguez-Trelles, F., Ayala, F.J., 1998. New *Drosophila* introns originate by duplication. Proc. Natl. Acad. Sci. USA 95, 1658–1662.

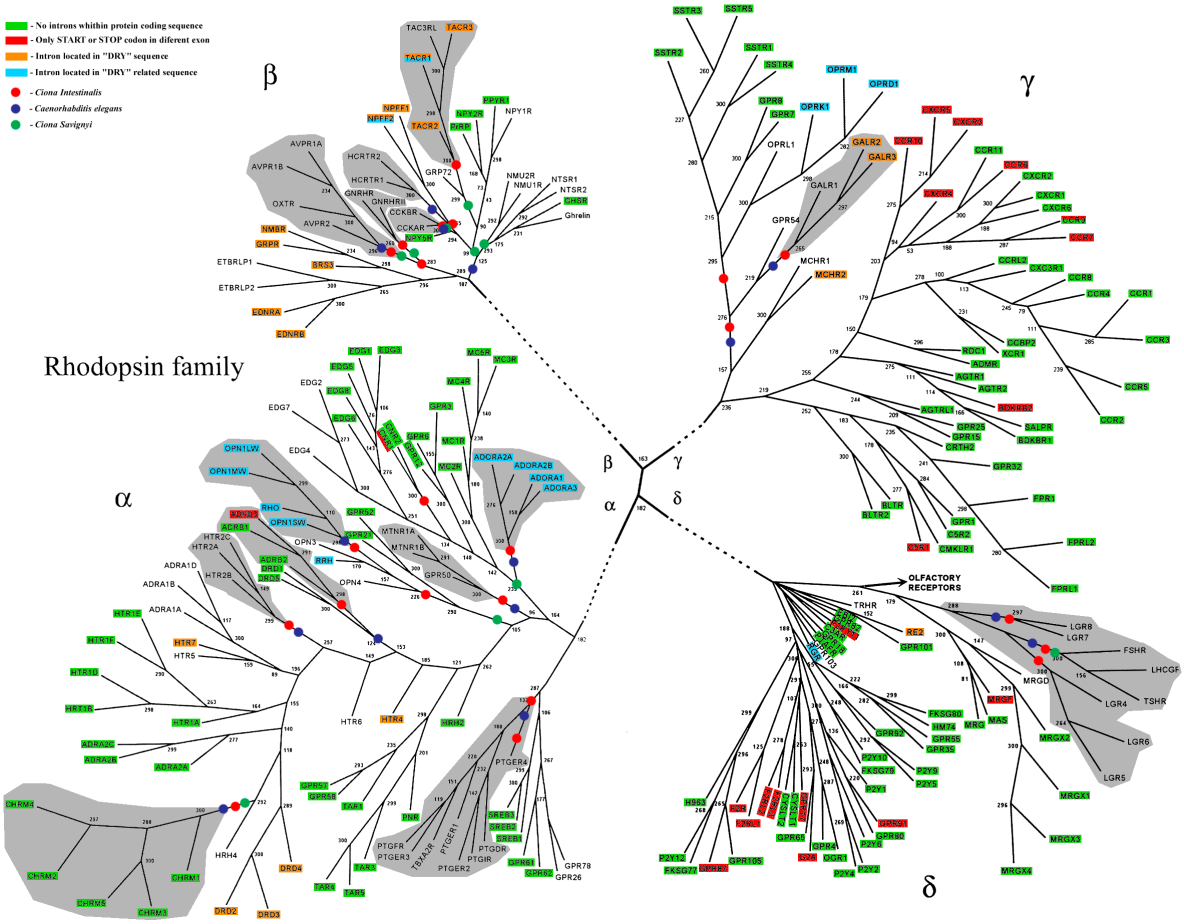
Tarrio, R., Rodriguez-Trelles, F., Ayala, F.J., 2003. A new *Drosophila* spliceosomal intron position is common in plants. Proc. Natl. Acad. Sci. USA 100, 6580–6583.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Venkatesh, B., Ning, Y., Brenner, S., 1999. Late changes in spliceosomal introns define clades in vertebrate evolution. Proc. Natl. Acad. Sci. USA 96, 10267–10271.

Zylka, M.J., Dong, X., Southwell, A.L., Anderson, D.J., 2003. Atypical expansion in mice of the sensory neuron-specific Mrg G protein-coupled receptor family. Proc. Natl. Acad. Sci. USA 100, 10043–10048.



Supplementary Figure 1. The phylogenetic relationship between GPCRs (TMI–TMVII) of the human Rhodopsin family, obtained from Fredriksson et al. (2003). The receptors labeled with: green was intronless within protein coding sequence, red contained only one intron within protein coding sequence, which separated only start or stop codon along with small number of surrounding amino acid codons, orange contained intron within conservative DRY motif and blue contained intron within motif to similar conservative DRY motif. Shaded fields mark Ancestral Receptor Groups (ARG). Coloured circles represent the results acquired from BLAST searches and displays the similarities of *Ciona intestinalis* (red), *Ciona savignyi* (blue), and *Caenorhabditis elegans* (green) receptors to groups of human receptors.

Supplementary Table 1 Total intron densities within whole (CDS) and only seven transmembrane region (TM7) coding sequence in human *Rhodopsin* GPCRs, members of Ancestral Receptor Group and non-Ancestral Receptor Group members

		In all receptors	In ARG members	In Non ARG members	Fold difference
ALL CDS	Total length of CDS	288130	75717	212413	
	No. of Introns	363	202	161	
	Introns/Kb	1.259847985	2.6678289	0.7579574	3.51
Only TM7 Regions	Total length of CDS	193466	45478	147988	
	No. of Introns	184	76	108	
	Introns/Kb	0.951071506	1.6711377	0.7297889	2.29

3.3 Identification of domains responsible for specific membrane transport and ligand specificity of the ACTH receptor (MC2R)

Molecular and Cellular Endocrinology 321 (2010) 175–183

Contents lists available at ScienceDirect

Molecular and Cellular Endocrinology

journal homepage: www.elsevier.com/locate/mce

Identification of domains responsible for specific membrane transport and ligand specificity of the ACTH receptor (MC2R)

Davids Fridmanis[a,b], Ramona Petrovska[a], Ineta Kalnina[a,b], Maija Slaidina[a,b], Raitis Peculis[a], Helgi B. Schiöth[b], Janis Klovins[a,b,*]

[a] Latvian Biomedical Research and Study Centre, Latvia, Ratsupites 1, LV-1067 Riga, Latvia
[b] Department of Neuroscience, Uppsala University, BMC, Box 593, SE751 24, Uppsala, Sweden

ARTICLE INFO

Article history:
Received 7 October 2009
Received in revised form 23 February 2010
Accepted 23 February 2010

Keywords:
GPCR
Rhodopsin
Melanocortin
ACTHR
Chimeric receptors
MRAP
GFP
Membrane transport

ABSTRACT

The adrenocorticotropic hormone (ACTH) receptor has highly specific membrane expression that is limited to adrenal cells; in other cell types the polypeptide fails to be transported to the cell surface. Unlike other evolutionarily related members of the melanocortin receptor family (MC1R–MC5R) that recognize different melanocortin peptides, ACTHR (MC2R) binds only ACTH. We used a mutagenesis approach involving systematic construction of chimeric ACTHR/MC4R receptors to identify the domains determining the selectivity of ACTHR membrane transport and ACTH binding. In total 15 chimeric receptors were created by replacement of selected domains of human ACTHR with the corresponding regions of human MC4R. We developed an analytical method to accurately quantify cell-membrane localization of recombinant receptors fused with enhanced green fluorescent protein by confocal fluorescence microscopy. The chimeric receptors were also tested for their ability to bind ACTH (1–24) and the melanocyte-stimulating hormone (MSH) analog, Nle4, DPhe7-α-MSH, and to induce a cAMP response. Our results indicate that substitution of the MC4R N-terminal segment with the homologous segment of ACTHR significantly decreased membrane transport. We also identified another signal localized in the third and fourth transmembrane regions as the main determinant of ACTHR intracellular retention. In addition, we found that the fourth and fifth transmembrane domains of the ACTHR are involved in ACTH binding selectivity. We discuss the mechanisms involved in bypassing these arrest signals via an interaction with melanocortin 2 receptor accessory protein (MRAP) and the possible mechanisms that determine the high ligand-binding specificity of ACTHR.

1. Introduction

All five melanocortin receptors are activated by binding of different melanocortins: α, β, and γ melanocyte-stimulating hormones (MSH) and adrenocorticotropic hormone (ACTH), that are all peptides derived from the large pro-opiomelanocortin precursor (Eberle, 1988) and share a conserved tetrapeptide sequence, His-Phe-Arg-Trp, that serves as a pharmacophore (Abdel-Malek, 2001). The melanocortins (MCs) regulate many different physiological functions including the formation of pigmentation (De Wied and Jolles, 1982; Spencer and Schallreuter, 2008), adrenocortical steroidogenesis (Simpson and Waterman, 1988; Allolio and Reincke, 1997; Chida et al., 2007), energy homeostasis (Sina et al., 1999; Yeo et al., 2000; Lu, 2001; Williams et al., 2001; Ellacott and Cone, 2004; Butler, 2006; Ellacott et al., 2007), exocrine secretion (Chen et al., 1997; Thiboutot et al., 2000), sexual function (Argiolas et al., 2000; Wessells et al., 2000; Pfaus et al., 2004), inflammation and immunomodulation (Haycock et al., 1999; Starowicz and Przewlocka, 2003; Lasaga et al., 2008), thermoregulation (Fan et al., 2005, 2007) and cardiovascular regulation (Humphreys, 2004; Rinne et al., 2008). The ACTH receptor (ACTHR) belongs to the MC receptor (MCR) family that comprises five members, MC1R–MC5R, where MC2R is the ACTH receptor. The MCRs all belong to the rhodopsin G-protein-coupled receptor (GPCR) family, α-group, MECA cluster (Fredriksson et al., 2003). MCRs act predominantly through $G_{\alpha s}$ subunits coupled to adenyl cyclase (Chhajlani and Wikberg, 1992; Mountjoy et al., 1992; Gantz et al., 1993a,b, 1994; Theodoropoulou et al., 2008), however there are also reports of MC3R acting through members of $G_{\alpha q/11}$ subunit family coupled to pathways involving phospholipase C-mediated hydrolysis of phosphoinositides and the mobilization of intracellular Ca^{2+} (Konda et al., 1994).

The MC3R, MC4R and MC5R share the highest amino-acid sequence identity among the MCRs (55–61%) while MC1R and ACTHR are less similar with only 39% sequence similarity. The

* Corresponding author at: Latvian Biomedical Research and Study Centre, Ratsupites 1, LV-1067 Riga, Latvia. Tel.: +371 67808003; fax: +371 67442407.
E-mail address: klovins@biomed.lu.lv (J. Klovins).

0303-7207/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.mce.2010.02.032

MC4R and ACTHR share 47% overall amino-acid sequence identity and 50% within the TM regions. Despite high similarity in receptor sequences and intracellular pathways activated, all five MCRs display different profiles of tissue-specific expression and distinct pharmacological profiles. ACTHR stands out from the other MCRs with very high expression and ligand-binding specificity. Unlike the other MCRs, it solely binds to ACTH and is predominantly expressed in the adrenal cortex (Mountjoy et al., 1992), although low expression levels have been also detected in adipose tissue (Boston, 1999). It is also one of the smallest GPCRs identified, comprising only 297 amino-acid residues with a total molecular weight of 33 kDa (Clark and Cammas, 1996). In the adrenal cortex ACTHR mediates the effects of ACTH on glucocorticoid synthesis (Clark and Cammas, 1996) and as well as on its own mRNA synthesis (Penhoat et al., 1989; Naville et al., 1999). Genetic defects in the gene encoding ACTHR cause familial glucocorticoid deficiency (FGD), a rare autosomal recessive disorder characterized by low or undetectable plasma cortisol levels and an excess of plasma ACTH. It becomes apparent in childhood as frequent hypoglycemia and/or infective episodes accompanied by excessive skin pigmentation (Elias et al., 1999; Penhoat et al., 2002).

Since the cloning of the receptor by Mountjoy et al. (1992) studies of the ACTHR have been hampered by the fact that heterologous functional expression of this receptor could only be achieved in cell lines that already express endogenous MCRs, resulting in high background levels of ACTH binding sites (Noon et al., 2002). The first breakthrough was achieved when Schimmer et al. (1995) obtained Y6 and OS3 cell lines (from adrenal cell line Y1) lacking endogenously expressed MCRs. Nevertheless the reason why ACTHR could only be expressed in cell lines that normally express MC receptors remained unanswered until 2005 when Metherell et al. (2005) performed genetic analysis on a family with FGD and identified the MCR accessory protein (MRAP) – a small membrane protein that is required for ACTHR to reach the plasma membrane. Since then it has been extensively studied, resulting in discovery of MRAP dual-topology dimerization (Sebag and Hinkle, 2007, 2008; Cooray et al., 2008), the identification of MRAP domains involved in ACTHR transport (Sebag and Hinkle, 2008; Webb et al., 2008), and the ACTHR amino-acid residues involved in the interaction with MRAP (Chung et al., 2008).

We have now performed a comprehensive investigation of the domains responsible for specific membrane transport and ligand specificity of the ACTHR using of series of chimeric receptors fused to fluorescent markers. During the process we developed an analytical method that allows quantification of cell-membrane co-localization of recombinant receptors based on statistical analysis of the florescence ratio of enhanced green fluorescence protein (EGFP) and Alexa Fluor 594-conjugated wheat germ agglutinin (AF-WGA) in the plasma membrane.

2. Materials and methods

2.1. Cloning vector

All recombinant constructs were assembled using vector pCEP4-GFP-C modified from vector pCEP4 (Invitrogen). Modification included insertion of the gene encoding enhanced green fluorescent protein (EGFP) at the 3′ end of the multiple cloning site. Insertion was performed by amplification of the EGFP gene from plasmid pEGFP-N1 (Invitrogen) plasmid using *Pfu* polymerase (Fermentas) and the following primers: 5′-GGG CTC GAG ATG GTG AGC AAG GGC GAG G-3′ and 5′-CCG GAT CCT TAC TTG TAC AGC TCG TCC-3′ that contain *Xho*I and *Bam*HI restriction sites (underlined); amplified fragments and the pCEP4 vector were cleaved with both restriction enzymes (Fermentas) and ligated using T4 DNA ligase (Fermentas).

2.2. Generation of chimeric ACTHR/MC4R clones

Human ACTHR and human MC4R coding sequences were inserted in pCEP4-GFP-C vector by amplification of both receptor genes from plasmids based on the pcDNA3.1 vector using *Pfu* polymerase (Fermentas) with following forward (F) and reverse (R) primers: MC2-F (5′-ATT GGG AAG CTT ATG AAG CAC ATT ATC AAC TCG-3′) and MC2R (5′-GTA T***CT CGA GTC CAG AGC CGG TTC CAG ATC C***CC AGT ACC TGC TGC AGA A-3′) for ACTHR and primers MC4-F (5′-ATT GGG AAG CTT ATG GTG AAC TCC ACC CAC CGT-3′) and MC4-R (5′-AAG CTC GAG ATA TCT GCT AGA CAA GTC A-3′) for MC4R. Primers included *Hind*III or *Xho*I restriction sites (underlined). The MC2R primer contains a 9 amino-acid linker sequence (**bold**–*italic*): GSGTGSGLE. Amplified fragments were cleaved with *Hind*III and *Xho*I restriction enzymes and ligated into pCEP4-GFP-C vector using T4 DNA ligase (Fermentas). All chimeric receptors were created using site-directed mutagenesis by overlap extension (Ho et al., 1989). Sequences of all acquired chimeric receptors were verified by sequencing of plasmids using BigDye v3.1 and the ABI Prism® 3130xl Genetic Analyzer sequencer system (Applied Biosystems) according to recommendations of the supplier. The sequences of all primers are provided in Supplementary Table 1.

2.3. Expression of receptor clones

For binding and cAMP studies, HEK293 EBNA cells were grown to 50–70% confluence and transfected with 10 μg of construct using FuGene™ transfection reagent (Roche Applied Science) according to manufacturer's instructions. Cells were grown in Dulbecco's modified Eagle medium: nutrient mixture F-12 (DMEM/F12) supplemented with 10% fetal calf serum (FCS) and 250 μg/ml Geneticin® (Invitrogen) in a humidified atmosphere of 95% air and 5% CO_2 (v/v). Total expression levels for each DNA construct were estimated by measuring direct EGFP fluorescence using a Wallac 1420 (VICTOR 3 V) Multilabel counter (PerkinElmer), the measurements were performed on three separate cell preparations. Semistable cell lines expressing the target receptor were obtained by selecting for growth in the presence of 200 μg/ml hygromycin B (Invitrogen), first added 24 h after transfection. Cells were grown continuously in the presence of hygromycin B and harvested for further experiments. For microscopy experiments BHK cells were grown on cover slips in DMEM/F12 supplemented with 5% FCS. Plasmids were transiently transfected using TurboFect™ transfection reagent (Fermentas) and the reverse transfection technique. Cover slips with attached cells were used for confocal microscopy after ∼24 h of incubation.

2.4. Binding studies

HEK293 EBNA cells expressing the recombinant receptors were harvested in 1× phosphate buffered saline (PBS) containing 1.47 mM KH_2PO_4, 4.29 mM Na_2HPO_4, 137 mM NaCl, 2.68 mM KCl, pH 7.4 and homogenized using an Ultra-Turrax homogenizer tool (Jankie & Kunkel IKA-WERK). The homogenized suspension was centrifuged for 3 min at 1300 rpm and supernatant was recentrifuged for 15 min at 15,000 rpm to pellet and concentrate the cell membranes. The pellet was resuspended in binding buffer (BB) containing 50 mM Tris–HCl pH 7.4, 2.5 mM $MgCl_2$, 1 mM $CaCl_2$, and 150 mM NaCl. Binding was performed in a final volume of 100 μl for 3 h at room temperature using 10–20 μg of membrane preparation per well. Saturation studies were performed with different concentrations of ^{125}I-labeled Nle4, DPhe7-α-MSH (NDP-MSH) while competition studies including serial dilutions of unlabeled NDP-MSH or ACTH (1–24) peptides (Neosystems, France) in the incubation mixture together with 0.6 nM ^{125}I-labeled NDP-MSH. Non-specific binding was defined as the amount of radioactivity remaining bound to cell membranes after incubation in the presence of 1000 nM unlabeled NDP-MSH or ACTH (1–24) depending on the type of experiment. Incubation was terminated by filtration through GF/C filters, Filtermat A (Wallac Oy, Turku, Finland), that had been pre-soaked in 0.3% polyethyleneimine (Sigma) using a TOMTEC Mach III cell harvester (Orange, CT). Filters were washed with 5.0 ml of 50 mM Tris–HCl pH 7.4 per well at 4 °C and dried at 60 °C. Dried filters were then treated with MeltiLex A (PerkinElmer Life Sciences) melt-on scintillation sheets and were counted using a Wallac 1450 (Wizard automatic Microbeta counter). Results were analyzed using Prism 3.0 GraphPad (San Diego, CA). All binding assays were performed in duplicate and were repeated at least three times.

2.5. Cyclic adenosine monophosphate (cAMP) response studies

Prior to stimulation, HEK293 cells expressing the receptor of interest were incubated for 3 h in media as specified above with addition of 2.5 μCi/ml of [^{3}H]ATP (specific activity 29 Ci/mmol; Amersham Biosciences). After incubation, cells were harvested with 5 ml of Na-Elliot buffer (137 mM NaCl, 5 mM KCl, 0.44 mM KH_2PO_4, 4.2 mM $NaHCO_3$, 1.2 mM $MgCl_2$, 20 mM HEPES, 1 mM $CaCl_2$, and 10 mM glucose, pH adjusted to 7.4), centrifuged and resuspended in Na-Elliot buffer containing 0.5 mM isobutyl methylxanthine (IBMX, Sigma). Afterwards the cells were incubated at 37 °C for 10 min. Stimulation was performed by incubating cells (∼2×10^5 cells) with different concentrations of α-MSH and ACTH (1–24) (diluted in Na-Elliot buffer containing 0.5 mM IBMX) for 20 min at 37 °C in a final volume of 150 μl. After incubation, cells were precipitated by centrifugation and lysed by the addition of 200 μl of 0.33 M perchloric acid, followed by freezing and thawing. After centrifugation the supernatant (200 μl) was transferred to prewashed (with 2× 10 ml H_2O) Dowex 50 W-X4 resin columns (Bio-Rad). To standardize the efficiency of the [^{3}H]cAMP columns an internal standard consisting of 750 μl of 0.33 M perchloric acid containing 0.5 nCi/ml [^{14}C]cAMP (Amersham Biosciences) was also added to each column. Columns were washed with 2 ml of H_2O to elute ATP and the flowthrough was collected in scintillation vials to estimate the amount of unconverted [^{3}H]ATP. 4 ml

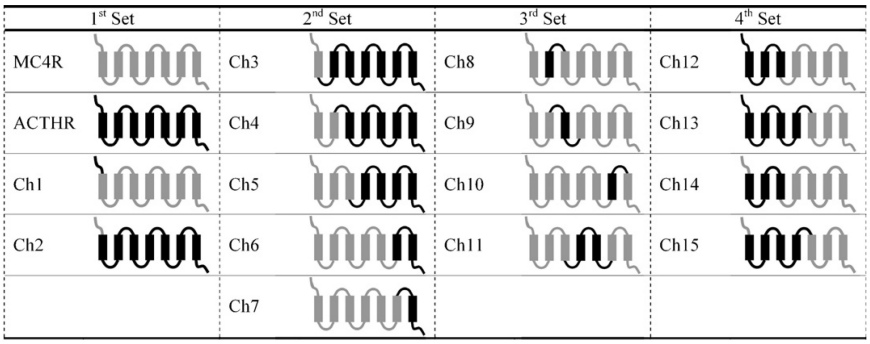


Fig. 1. Schematic presentation of the chimeric receptors. Regions from the MC4R are shaded grey, regions from the ACTHR are shaded black.

of Optiphase HiSafe 3 scintillation cocktail (Perkin Elmer) were added to each vial before counting. Dowex columns were placed over alumina (Sigma) columns (pre-washed with 8 ml of 0.1 M imidazole) and cAMP was eluted using 10 ml of H_2O. cAMP was then eluted from the alumina columns using 4 ml of 0.1 M imidazole and collected in scintillation vials containing 7 ml of Optiphase HiSafe 3 scintillation cocktail (PerkinElmer). ^3H and ^{14}C were counted using a Tri-carb liquid scintillation beta counter. Results were calculated as the percent of total [^{3}H]ATP (obtained as a sum of [^{3}H]ATP from first column and [^{3}H]cAMP from second column) to [^{3}H]cAMP, standardized by calculation of each column efficiency (the amount of obtained [^{14}C]cAMP was expressed as a fraction of total [^{14}C]cAMP) and used to determine 50% effective concentration (EC50) values by non-linear regression using PRISM 3.0 software. We also used LACNE^{TM} cAMP 384 Kit (PerkinElmer) to compare the influence of EGFP, and linker-fused EGFP, on cAMP response potency and ACTHR activity. All experiments were performed in duplicate and repeated three times.

2.6. Confocal microscopy studies

For these experiments BHK cells were used because they attach better to glass coverslips than HEK293 cells. To label the cells membrane, cover slips with attached BHK cells expressing different receptors were rinsed in 1× PBS and incubated for 3 min in Alexa Fluor® 633-labeled wheat germ agglutinin (AF-WGA) (Molecular Probes – Invitrogen) in 1× PBS solution at a final concentration of 10 μg/ml. After incubation, cover slips were rinsed twice with 1× PBS, fixed for 10 min with 4% formaldehyde (Sigma) in 1× PBS and rinsed twice with 1× PBS. The prepared samples were placed upon microscopy slides and inspected using a confocal fluorescent microscope (Leica DM 6000B).

At least two independent transfections were used for each receptor construct. Three cells were selecting for analysis according to the following criteria. First, a clear EGFP fluorescence emission signal was detected; second, the plasma membrane of the cell analyzed was fully labelled by AF-WGA with no or little Golgi complex labeling. To avoid bias due to the choice of the cell, inspection was performed separately for images acquired from EGFP fluorescence microscopy and AF-WGA fluorescence microscopy. From each analyzed cell, 18 evenly distributed cell cross-sectional images were captured along the z-axis (cell depth axis). These images were used for membrane expression/quantification measurements. The efficiency of EGFP-tagged receptor transport to the cell membrane was estimated by calculation of the EGFP/AF-WGA fluorescence intensity ratio at multiple points on the plasma membrane selected using randomly drawn linear regions of interest (ROI) across the cell images. To avoid bias due to the choice of ROIs, suitable ROIs were first selected on the AF-WGA image and the data for the same ROI were then acquired from the corresponding EGFP image (sample overlays of both images with ROIs are shown in Fig. 2b and d). Selected ROIs were considered as suitable for further analysis only if they crossed the cell membrane at two points and if AF-WGA fluorescence intensity peaks in the intensity graph (a sample graph of a ROI is shown in Fig. 2a and c) at these points were clearly distinguishable from background and noise peaks. 20 random points from each cell were selected, giving at least 120 points for analysis of each receptor construct per cell. All measurements were carried out using the Leica Confocal Software (LCS-Lite v2.61) quantification-profiling tool.

2.7. Statistical analysis of confocal microscopy data

After data collection, all six datasets acquired from six cells expressing the same receptor were tested for uniformity using the Kruskal–Wallis test (significance level alpha = 0.05). This step was performed for two reasons. First, because a relatively small number of cells were analyzed it was important to exclude results acquired from single cells that significantly differed from majority. Second, many transfection and microscopy experiments for different receptors were performed simultaneously, and this increased the probability of sample mix-up that would generate two data sets with significant differences in membrane transport. If the Kruskal–Wallis test showed that medians of acquired data varied significantly then Dunn's multiple comparison test (Dunn, 1964) (significance level alpha = 0.05) was performed to determine which of the data sets was different. Datasets that presented significant differences with more than two other datasets were replaced with values acquired from repeated independent transfection experiments and confocal microscopy analysis. Further, data acquired from different receptors were analyzed by comparing median values and interquartile ranges. The significance of differences in median values was estimated using the Kruskal–Wallis test (significance level alpha = 0.01). Because medians varied significantly the Kruskal–Wallis test was followed by Dunn's multiple comparison test (significance level alpha = 0.01). Modules of differences in rank sums acquired from this test were used to create a matrix table that was then used to cluster receptors by their differences in membrane export using a Euclidean distance clustering method. Clustering employed MultiExperiment Viewer v4.3 software (Saeed et al., 2003).

3. Results

We investigated the membrane transport and ligand-binding specificity of the human ACTH receptor using a series of constructs in which different domains of ACTHR and MC4R were exchanged. To allow fluorescence imaging of cells expressing these constructs by confocal microscopy all constructs created in this study contained EGFP fused to the C-termini of the receptors. Because GFP fusion to the C-terminus of ACTHR can impair receptor function (Roy et al., 2007) we inserted a serine- and glycine-rich linker sequence between the ACTHR and EGFP sequences as recommended by Shimozono and Miyawaki (2008). Fusion of EGFP to the C-terminus of ACTHR decreased the potency of the cAMP response after ACTH stimulation for the constructs both with and without the linker (EC50 = 3.05 nM and 1.11 nM, respectively) compared to wild-type ACTHR (EC50 = 0.03 nM) in cells transiently co-expressing MRAP. The construct containing the linker however showed increased efficacy (11.87 nM cAMP) compared to the EGFP only construct (7.514 nM cAMP) but was lower than wild-type ACTHR (28.27 nM cAMP).

The domain-exchange chimeric constructs assembled in this study, designated Ch1–15, are presented in Fig. 1. Alignments of all amino-acid sequences including wild type (wt) receptors and chimeric receptors are presented in Supplementary Fig. 1. All DNA constructs directed the expression of functional EGFP fusion proteins as analyzed by direct measurement of EGFP fluorescence signals (Supplementary Fig. 2). We conclude that the different chimeric constructs all encode functional fusion polypeptides that are reliably expressed following transfection. Very little between cell variation was observed in median values of EGFP/AF-WGA fluorescence intensities of cells expressing the same construct, as measured at multiple different points on the plasma membrane of each cell analyzed, and even from different transient transfection experiments (Supplementary Fig. 3).

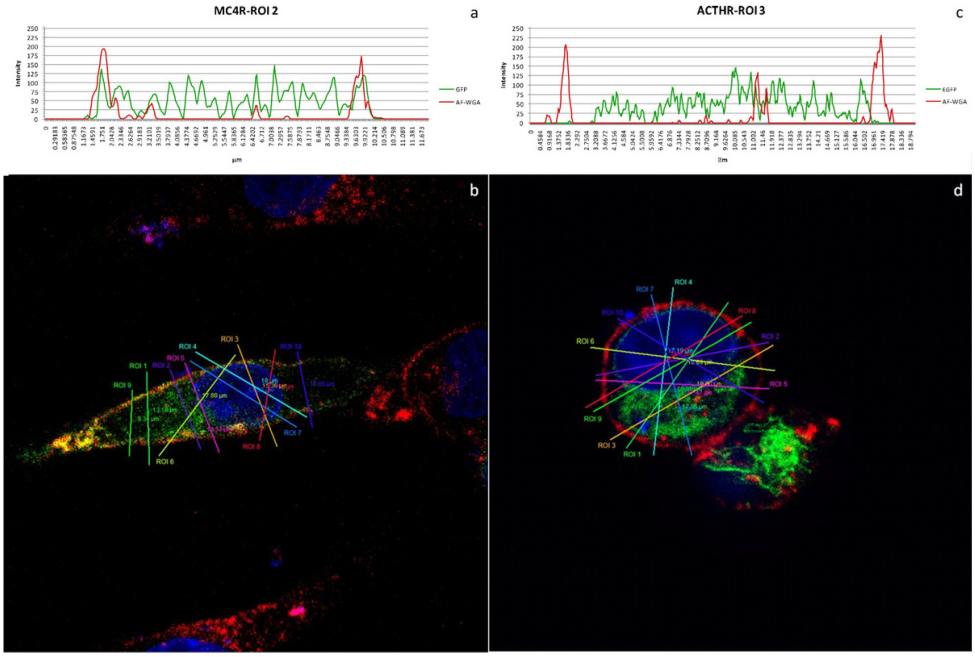


Fig. 2. Example of cell-surface fluorescence quantification of BHK cells expressing receptor-EGFP fusion proteins using confocal microscopy. (a) and (c) Plots of fluorescence intensity for MC4R and ACTHR respectively across selected regions of interest (ROI), the green line represents fluorescence level of receptor-EGFP fusion proteins, the red line represents the fluorescence level of AF-WGA (non-specific cell-surface glycoproteins). (b) and (d) Confocal fluorescence microscopy image of cells expressing MC4R-EGFP and ACTHR-EGFP fusion proteins, respectively (in green); the cell membrane is stained with AF-WGA (red); the colored lines indicated the ROIs analyzed.

3.1. *Role of the N-terminus in melanocortin receptor transport to the plasma membrane*

To investigate if the N-terminal segment of MC4R determines proper transport of the polypeptide to the plasma membrane, and the same region of ACTHR leads to transport failure, we created two ACTHR/MC4R chimeric constructs. In Ch1 the extracellular N-terminal segment of MC4R was replaced by the corresponding region from ACTHR, whereas Ch2 comprises ACTHR in which the N-terminal segment has been replaced by that of MC4R (Fig. 1). BHK cells transiently expressing these receptors were stained with AF-WGA that non-specifically labels cell-membrane glycoproteins and were subjected to confocal fluorescence microscopy.

This revealed that only cells expressing intact MC4R displayed EGFP localization overlapping with AF-WGA staining around the perimeter of the cell, in addition to dispersed intracellular fluores-

Table 1
Effects of NDP-MSH and ACTH (1–24) on [^{125}I]NDP-MSH binding and cAMP production by HEK293 EBNA cells transfected with chimeric receptors (means ± SEM).

Receptors	B_{max} ± SEM [^{125}I]NDP-MSH [CPM] × 10^{-3}	K_d ± SEM [^{125}I]NDP-MSH [nM]	K_i ± SEM NDP-MSH [nM]	K_i ± SEM ACTH(1–24) [nM]	EC50 ± SEM NDP-MSH [nM]	EC50 ± SEM ACTH(1–24) [nM]
MC4R[a]	31.7 ± 4.2	1.3 ± 0.9	5.4 ± 2.1	7.8 ± 2.6	3.5 ± 2.1	9.2 ± 2.59
ACTHR	ND	ND	ND	ND	ND	ND
Ch1[b]	13.7 ± 2.7	0.5 ± 0.2	6.4 ± 4.8	ND	24.5 ± 10.7	350.5 ± 209.5
Ch2	ND	ND	ND	ND	ND	ND
Ch3[b]	ND	ND	ND	ND	ND	ND
Ch4	ND	ND	ND	ND	ND	ND
Ch5[a]	ND	ND	ND	ND	ND	ND
Ch6[a]	11.2 ± 4.5	0.5 ± 0.2	1.2 ± 0.0	86.7 ± 24.9	ND	ND
Ch7[b]	13.5 ± 4.2	0.3 ± 0.0	3.8 ± 0.8	13.9 ± 3.1	132.5 ± 97.6	573.6 ± 526.5
Ch8[b]	ND	ND	ND	ND	ND	ND
Ch9[b]	ND	ND	ND	ND	ND	ND
Ch10[a]	5.1 ± 0.5	0.4 ± 0.1	16.8 ± 10.5	99.0 ± 31.0	479.5 ± 300.5	1144.0 ± 956.0
Ch11[a]	23.1 ± 2.3	1.2 ± 0.2	2.7 ± 0.6	19.7 ± 7.7	ND	ND
Ch12[b]	ND	ND	ND	ND	ND	ND
Ch13	ND	ND	ND	ND	ND	ND
Ch14[a]	7.8 ± 2.2	1.3 ± 1.0	5.3 ± 1.8	3.8 ± 1.0	ND	ND
Ch15[b]	3.6 ± 0.8	0.2 ± 0.1	1.5 ± 0.7	2.0 ± 1.4	6.1 ± 1.3	36.7 ± 3.8

B_{max} and K_d values were obtained from saturation curves using [^{125}I]NDP-MSH; K_i values were obtained from 0.6 nM [^{125}I]NDP-MSH competition curves using NDP-MSH and ACTH (1–24) as competitors; EC50 values were obtained from cAMP response curves using NDP-MSH and ACTH (1–24) as stimulators.
ND, not determined; SEM, standard error of the mean.
[a] Receptors with high membrane export.
[b] Receptors with low membrane export.

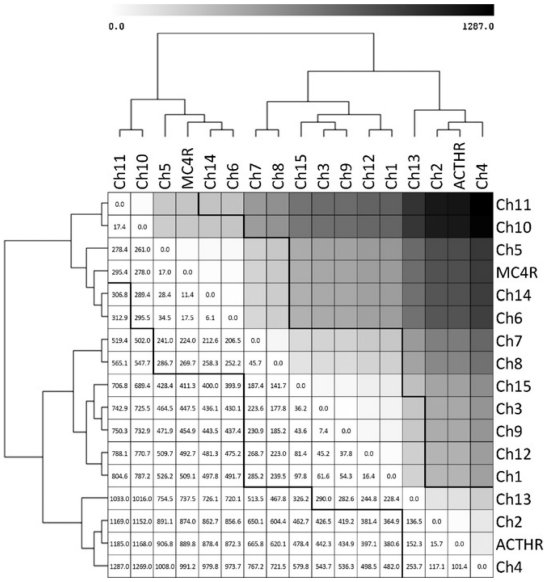


Fig. 3. Matrix table and receptor clustering tree for cell-surface presentation of the different receptors studied. Both were created using Dunn's multiple comparison test. The lower left part of the matrix table contains modules of differences in rank sums that were used to cluster the receptors according to differences in the efficiency of membrane export using a Euclidean distance-clustering method. The upper right part of matrix table provides a visual depiction of the matrix table; smallest differences in rank sums are shown in white, boxes in black correspond to the largest differences, and color transitions between white and black represent intermediate stages. The bold lines at upper right and lower left corner of the matrix table draw the borders between cells with significant difference in rank sums ($p < 0.01$) and cells with no significant difference in rank sums ($p > 0.01$).

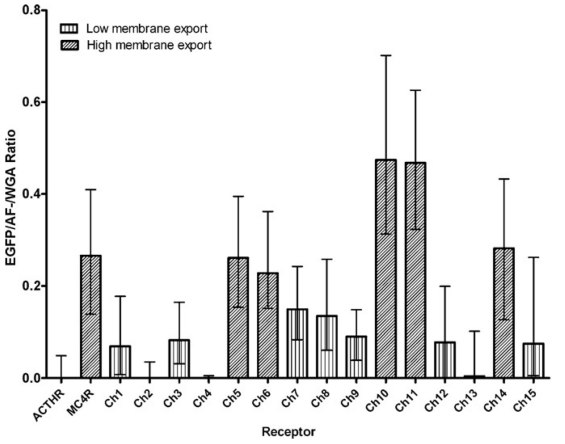


Fig. 4. Graph representing medians with interquartile ranges of EGFP/AF-WGA fluorescence intensity ratios at the plasma membrane for all tested receptors. Bars with vertical and angled cross-hatching represent receptors with low and high levels of membrane export, respectively.

cence (Fig. 2). By contrast, cells expressing intact ACTHR, which has been shown to be retained in endoplasmatic reticulum (ER) (Noon et al., 2002), and both chimeric constructs Ch1 and Ch2 showed a different pattern of intracellular EGFP distribution that was predominantly in the perinuclear area. Quantification of the EGFP/AF-WGA fluorescence intensity ratio in the plasma membrane confirmed the absence of ACTHR and Ch2 from the plasma membrane. However, this revealed that Ch1 is also trafficked to the cell membrane but in significantly ($p < 0.01$) lower amounts than MC4R (Figs. 3 and 4).

To evaluate the possibility that Ch1 and Ch2 are exported to the plasma membrane at levels not clearly detectable by fluorescence imaging, binding and cAMP assays were also performed on HEK293 EBNA cells semi-stably expressing these constructs. No detectable binding or cAMP accumulation was observed either with NDP-MSH or ACTH (1–24) on Ch2 receptor, while Ch1 showed a NDP-MSH binding affinity similar to that of MC4R. cAMP response EC50 values on stimulation with NDP-MSH and ACTH (1–24) were 7- and 38-fold lower, respectively (Table 1, Fig. 5).

3.2. Internal receptor regions responsible for the intracellular arrest of ACTHR receptor transport

To evaluate the roles of other ACTHR regions in transport arrest we created two further sets of chimeric receptors in which different transmembrane domains (TMs) of ACTHR were replaced with corresponding regions of MC4R and vice versa. This second set comprised five chimeric receptors, Ch3–Ch7, in which the number of ACTHR TM domains substituted by the corresponding MC4R domains was increased serially from TM1 to TM6 (Fig. 1). In a third set of four chimeric receptors, Ch8–Ch11, TM2, TM3, TM6 or TM4 + TM5 of MC4R were replaced with corresponding domains from ACTHR (Fig. 1).

Quantification of levels of EGFP in the plasma membrane of transiently transfected BHK cells revealed that Ch5, Ch6, Ch10 and Ch11 are effectively trafficked to the cell membrane. Ch3, Ch7, Ch8 and Ch9 are also present but in amounts significantly ($p < 0.01$) lower than those of the first group, whereas Ch4 was completely intracellulary retained, as estimated by a EGFP/AF-WGA ratio that was not significantly ($p > 0.05$) different from that of ACTHR (Figs. 3 and 4). As revealed by visual inspection pattern of fluorescence in Ch4 confocal microscopy images was similar to that in ACTHR confocal microscopy images.

Binding assays revealed that four out of nine chimeric receptors, Ch6, Ch7, Ch10 and Ch11, specifically bound both the tested ligands. Binding affinities were generally close to those of MC4R, with the exception of Ch6 and Ch10 binding to ACTH (1–24) that bound this ligand with >10-fold lower affinity. Furthermore, Ch7 and Ch10 also triggered a cAMP response but this was significantly lower than that produced by MC4R (Table 1 and Fig. 5). Together these results indicate that the TM regions 3 and 4 of MC4R are permissive for transport while the same regions of ACTHR prevent transport to the plasma membrane and the polypeptide is intracellulary retained.

To test the role of TM3 and TM4 in transport of ACTHR we created a further set of four chimeric receptors. Two, Ch12 and Ch14, contained the N-terminus of ACTHR up to TM3 fused to the C-terminus of MC4R including TM4 (Fig. 1). In the remaining two chimeric receptors, Ch13 and Ch15, the N-terminus of ACTHR up to TM4 was fused to the C-terminus of MC4R extending from TM4 (Fig. 1). Because of concerns that the extracellular N-terminal segment of ACTHR might also contribute to intracellular arrest, in Ch12 and Ch13 this segment was that of ACTHR whereas in Ch14 and Ch15 this was replaced by the extracellular N-terminal segment of MC4R (Fig. 1).

Confocal microscopy on cells expressing these constructs revealed that the combined TM3 and TM4 domains of ACTHR decrease transport to the cell membrane (Ch13 and Ch15) compared to the chimeras where only TM3 was replaced (Ch12 and Ch14). In addition, the extracellular N-terminal segment of ACTHR (Ch14 and Ch15) significantly reduced transport to the cell membrane compared to the reverse chimeras (Ch12 and Ch13) containing the corresponding MC4R N-terminal extracellular segment. Indeed Ch13, containing both TM3 and TM4 and N-terminal seg-

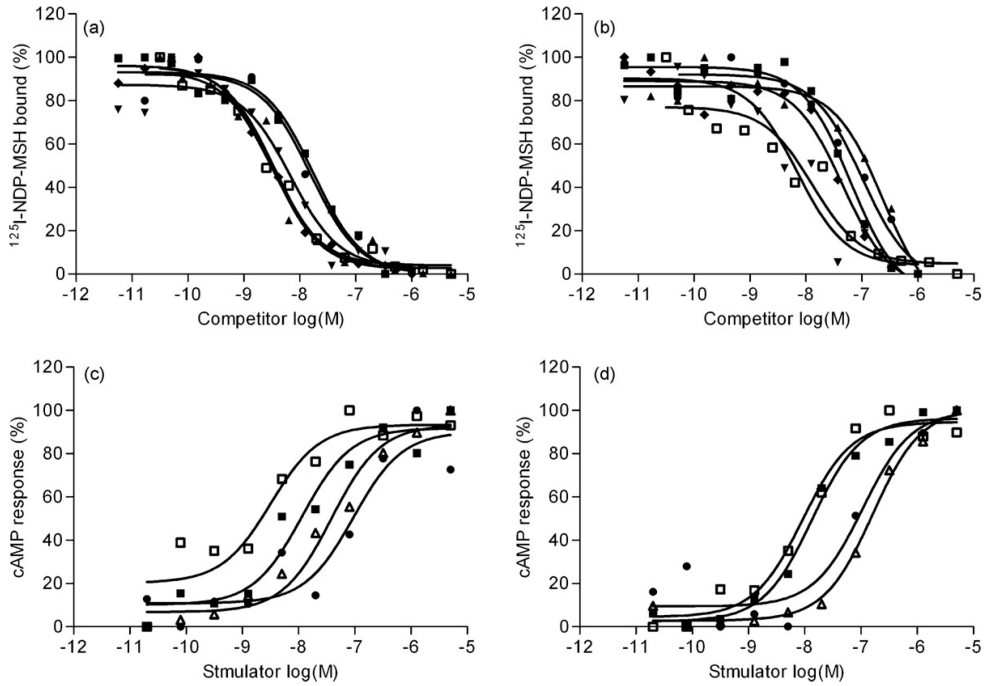


Fig. 5. Receptor ligand-binding analysis. (a) and (b) Competition binding curves; (c) and (d) cAMP response curves. (a) and (c) Binding curves from experiments with NDP-MSH; (b) and (d) binding curves from experiments with ACTH (1–24). Curves acquired from the analysis of different chimeric receptors are labeled as follows: (▲) Ch3; (■) Ch4; (●) Ch7; (♦) Ch9; (△) Ch16; (▼) Ch19 and (□) Ch20.

ment of ACTHR, was completely intracellulary retained (pattern of fluorescence in confocal microscopy images was similar to that in ACTHR confocal microscopy images), whereas Ch15, containing TM3 and TM4 from ACTHR and the N-terminal segment from MC4R was present in the cell membrane, although in amounts significantly lower than MC4R ($p < 0.01$). Ch12, containing TM3 and the N-terminal segment from ACTHR was found in the membrane in only small amounts. By contrast, Ch14, that contains TM1–TM3 from ACTHR and the N-terminal segment from MC4R, was effectively trafficked to the cell membrane in amounts not significantly ($p > 0.05$) different from MC4R (Figs. 3 and 4).

In ligand-binding and cAMP assays, Ch14 and Ch15 were both found to be active in the binding of both ligands tested, and with binding affinities close to those of MC4R. However, only the Ch15 cAMP response to NDP-MSH was comparable to that of MC4R, and the response to ACTH (1–24) was 4-fold lower (Table 1 and Fig. 5)

To accurately compare the levels of membrane transport of the different receptors we used EGFP/AF-WGA ratio quantification analyses followed by receptor clustering using modules of their rank differences acquired from Dunn's multiple comparison test (table in Fig. 3). This analysis confirmed that all the receptors tested can be subdivided into three groups based on location in the clustering tree (Fig. 3). The first group of receptors are located on the same major branch as MC4R and have a significantly higher ($p = 0.01$) EGFP/AF-WGA florescence ratio at the cell membrane than the other receptors (indicated by a bold line in the Fig. 3 table). These are thus considered to have 'high membrane export' levels. The second group are located on the same major branch as ACTHR and have a significantly lower ($p = 0.01$) EGFP/AF-WGA florescence ratio at the cell membrane compared to the other receptors. These receptors are thus considered to have 'no membrane export'. A third group of receptors are located on a separate branch from either MC4R or ACTHR and have a significantly lower ($p = 0.01$) EGFP/AF-WGA fluorescence ratio at the cell membrane compared to the first group but a significantly higher ($p = 0.01$) cell-membrane EGFP/AF-WGA florescence ratio compared to the second group. These receptors are thus considered to have 'low membrane export' levels (Fig. 3).

4. Discussion

The ACTHR has remarkably high ligand-binding specificity for ACTH and ignores all other melanocortins, including α-MSH that shares the same N-terminal 13 amino-acid sequence as ACTH. The molecular mechanisms underlying this endocrinologically important specificity have been difficult to study because of problems encountered in attempts to express ACTHR; to date no systematic mutagenesis experiments have been published on this receptor.

In the present study we constructed a series of chimeric ACTHR/MC4R receptors to identify the regions of ACTHR responsible for cell-surface expression and binding specificity. This technique has previously proved to be a powerful tool for the study of receptor structure and function (Frielle et al., 1988; Kobilka et al., 1988; Schioth et al., 1998b). MC4R was selected from among the other melanocortin receptors for the construction of chimeras because it is one of the best-characterized members of this receptor family. Chimeric receptors were fused at their C-termini with a reporter gene (EGFP) to allow visualization of their subcellular localization. To analyze cell-surface expression of chimeric receptors we developed an analytical method to accurately quantify cell-membrane co-localization based on statistical analysis of EGFP and AF-WGA fluorescence ratios in cell cross-sections. Because there was only low between cell variation in the fluorescence ratio we believe that this method is a powerful tool for the quantitative comparison of different levels of membrane expression that has been a major obstacle in previous mutagenesis studies on GPCRs.

It was previously shown that the extracellular loops of melanocortin receptors do not participate in ligand binding (Schioth et al., 1998a,b). Accordingly all the chimeric constructs assembled retained the original parental transmembrane (TM) domain sequences, and switch points were placed in the loop regions that share the lowest extent of amino-acid sequence homology. The first set of chimeric proteins comprised different exchanges between the N-terminal regions of MC4R and ACTHR. Analysis of cell-surface expression indicated that the N-terminal part of the receptors determines whether the polypeptide is trafficked to the cell surface (MC4R) or is intracellulary retained (ACTHR). However, simple replacement of the ACTHR N-terminal region by that from MC4R was insufficient for efficient cell-surface expression. This indicated a presence of another intracellular arrest signal within the ACTHR polypeptide. Previous studies established that melanocortin receptor regions TM1, TM2, TM3, TM6 and TM7 are predominantly involved in the binding of melanocortin peptides (Schioth et al., 1996, 1998a,b; Lagerstrom et al., 2003), and exchanges within these regions could therefore lead to incorrect receptor folding (partial or complete) and subsequent receptor degradation before it reaches the plasma membrane. This in turn could influence the total receptor expression thus increasing the complexity of how to interpretation of these results, particularly from quantification of receptor binding sites that were obtained using the total membrane preparation. It should therefore be mentioned that results acquired from the binding studies cannot be directly compared to microscopy results and we used these results mainly to characterize the receptor–ligand interaction. Analysis of further series of chimeric receptors demonstrated that domain replacements in the TM1–TM3 region (Ch3, Ch4, Ch8 and Ch9 in Fig. 1) not only caused loss of ligand binding, but also significantly reduced membrane transport (Table 1, Figs. 3 and 4). Furthermore, analysis of ACTHR chimeras containing increasing numbers of MC4R TM domain replacements (starting at the N-terminus) revealed that effective membrane transport only took place after replacement of the TM1–TM3 domains (Table 1, Figs. 3 and 4). These results indicate that the intracellular arrest signal within ACTHR is predominantly located within TM3, although it is also possible that the arrest signal is formed through interaction of several receptor regions. This hypothesis was confirmed after analysis of fourth set of chimeric receptors. Because partial intracellular retention was observed for only those receptors that contained either the N-terminal segment or both TM3 and TM4 from the ACTHR, we conclude that TM4 is also involved in generating the arrest signal. The only receptors that were completely intracellulary retained were those that contained both the N-terminal segment and also TM3 and TM4 from the ACTHR. Moreover, the gain of plasma-membrane expression is clearly detectable in identical constructs where ACTHR N-terminal segment is replaced by that of MC4R (Ch13 versus Ch15 and Ch12 versus Ch14). Similarly the replacement of N-terminal segment together with TM1–TM3 in Ch5 rescues the ACTHR from intracellular trapping compared to Ch2. Our results also indicate that intact TM1–TM3 region either from MC4R or ACTHR is an important prerequisite for high-level expression in the plasma membrane and changes in these regions may lead to misfoding.

Examination of the different MC receptor sequences (including those of different species) revealed that one of the primary unique features of the ACTHR is its short N-terminal region compared to the other subtypes. Analysis of the first set of chimeric receptors indicates that the N-terminal segment of ACTHR in Ch1 significantly decreased the amount of receptor at the plasma membrane, demonstrating that it could play an important role in the ACTHR intracellular retention. Membrane transport versus intracellular retention is frequently associated with the presence or absence of effective glycosylation signals (usually multiple) within this region (Petrecca et al., 1999; Boer et al., 2000) that are often required for appropriate plasma-membrane transport of cell-surface glycoproteins. We therefore performed a prediction search (NetNGlyc 1.0 Server software; available online at http://www.cbs.dtu.dk/services/NetNGlyc/) (Gupta and Brunak, 2002; Gupta et al., in preparation) to investigate potential glycosylation sites within the N-terminal segments of melanocortin receptors from five different species: *Homo sapiens*, *Mus musculus*, *Galus galus*, *Danio rerio* and *Takifugu rubripes*. The results are presented in Supplementary Table 2. This analysis indicates that human ACTHR (as well as ACTHR orthologs in general) contains only two sites for N-linked glycosylation, both with low glycosylation potential. Because effective membrane transport of other proteins often requires two or more sites of efficient glycosylation (Petrecca et al., 1999), the presence of only two weak potential glycosylation sites in the ACTHR receptor would suggest that this receptor is constrained in its ability to be transported to the cell membrane. By contrast, our analysis predicts that the human MC4R contains three glycosylation sites, all with significantly higher glycosylation potential. This could in part explain the efficient transport of MC4R to the cell surface. In addition, the majority (but not all) of MC4R polypeptides were found to contain a further potential glycosylation site within the first extracellular loop; this could possibly increase the efficiency of cell-surface transport.

In most cell types ACTHR does not appear to be transported to the cell membrane. However, in cells of the adrenal cortex the receptor is efficiently transported to the cell surface. The explanation for this is not fully understood. It is possible that intracellular arrest is bypassed in adrenal cells by additional post-translational modification or conformation changes, or the arrest signals could be masked with the help of another protein. Recent studies of the melanocortin receptor accessory protein (MRAP) have shown that changes take place following oligomerization of ACTHR with two antiparallel MRAP molecules (Sebag and Hinkle, 2007, 2008; Cooray et al., 2008; Webb et al., 2008). Indeed, such oligomerization could solve the ACTHR 'weak' glycosylation problem because MRAP itself has two sites for efficient N-linked glycosylation, thus increasing the overall glycosylation potential of the ACTHR–MRAP complex (Supplementary Table 2).

We also report that the combination of two intracellular retention signals in the ACTHR N-terminal segment and in TM3–TM4 is required for complete block of cell-surface transport. Receptors with only one regulatory element, such as Ch1 and Ch12 (containing only the ACTHR N-terminal segment) or Ch3 and Ch15 (containing only TM3–TM4), were transported to some degree to the cell membrane leading to functionally active receptor. However, ACTHR and Ch13, that contain both elements, were completely intracellulary retained. The presence of a 'dual protection' system to prevent cell-surface expression of ACTHR – the N-terminal segment and TM3–TM4 signals, could help to avoid potential leakage of ACTHR to the plasma membrane. Such leakage, even at low level, could radically change the way a cell responds to external stimuli.

In addition, conformation changes and interactions between TM domains could add a further level of complexity to the regulation of membrane transport. For example, this is seen when we compare plasma-membrane expression of chimeric receptors Ch2, Ch3 and Ch4. Receptor Ch3, comprising the N-terminal segment and TM1 domain from MC4R, is transported at only low levels to the membrane, whereas Ch2 (that contains only the N-terminal segment of MC4R) and Ch4 (that also contains the TM1–TM2 domains of MC4R) are both completely intracellulary retained.

The serious rearrangements in sequence of different chimeras used in this study can cause structural changes that would lead to immediate degradation due to loss of membrane protein characteristics. We used replacements that leave intact the original TM domain sequences in order to minimize the risk of misfold-

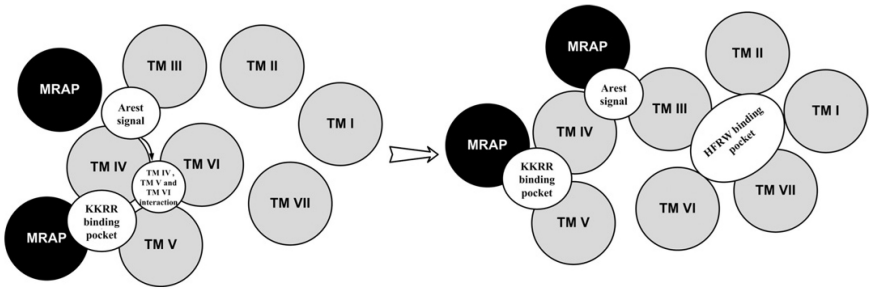


Fig. 6. Schematic presentation of the hypothetical ACTHR and MRAP complex before and after binding of the ACTH K-K-R-R pharmacophore. α-Helical transmembrane (TM) domains are depicted as grey circles, potential functional regions of the complex are marked with white circles. The black arrow illustrates a possible active conformational change taking place after binding of the ACTH K-K-R-R pharmacophore.

ing. Three chimeras (Ch2, Ch4 and Ch13) did however show total absence of co-localization with plasma-membrane marker and this might be an indication of such degradation. The possibility cannot be excluded that misfolding induced degradation may influence the membrane trafficking of other chimeras as well as ACTHR itself. However, no obvious influence of expression levels on binding- or plasma-membrane quantification data could be observed from our data. The chimeras Ch2, Ch4, Ch13 and ACTHR had total expression levels that are similar or even higher than those of MC4R. From the other hand the chimeras displaying the lowest total expression levels (Ch5, Ch7, Ch10 and Ch11) show medium to high plasma-membrane concentration and relevant binding and cAMP response affinities (Table 1 and Supplementary Fig. 2).

Although this work has focused primarily on the identification of regions responsible for the specificity of cell-surface expression of ACTHR, our data are also indicative of the regions that determine ligand-binding specificity. In Ch5, where TM4–TM7 of ACTHR are replaced by their MC4R counterparts, although the chimeric receptor was effectively transported to cell plasma membrane it failed to bind any of the receptor ligands tested. By contrast, Ch6, in which only TM6 and TM7 were substituted, and Ch11, with replacement of only TM4 and TM5, were both effectively transported to cell plasmatic membrane where they specifically bound both NDP-MSH and ACTH (1–24). This would suggest that the key determinants of ligand-binding reside within these TM domains. Nevertheless, ligand-induced cAMP production was impaired in these receptors. It is possible that the TM4 and TM5 regions (or some parts of them) interact with the TM6 and TM7 regions (or some parts of them) during the formation of the ligand-binding pocket, and heterologous replacement of these regions disrupts the conformation required binding of the melanocortin consensus pharmacophore His-Phe-Arg-Trp (H-F-R-W).

Notably, an interaction between ACTHR and MRAP is not only necessary for effective membrane export but is also required for binding of the ligand ACTH (Sebag and Hinkle, 2008). It is plausible to speculate that one of the MRAP antiparallel molecules binds to ACTHR at the TM3 and TM4 regions, thus masking the ER arrest signal, while the second one binds in the vicinity of TM4 and TM5 to constitute the binding pocket for the second ACTH pharmacophore Lys-Lys-Arg-Arg (K-K-R-R) (Kovalitskaia Yu et al., 2008). This scenario raises the possibility that binding of the ACTH-specific K-K-R-R pharmacophore leads to release of the TM4–TM5 binding interaction with TM6–TM7, allowing the formation of the pocket that interacts with the melanocortin consensus pharmacophore H-F-R-W, or that interaction initiates a conformation change that actively creates the H-F-R-W binding pocket (Fig. 6). Impaired cAMP signaling of some of the chimeric receptors could be explained by disrupted G-protein coupling to the modified receptors.

In summary, we report the identification of two independent regions responsible for the specificity of ACTHR cell-surface expression and propose that a dual protection system might be necessary to avoid receptor leakage to the cell membrane. Our results also indicate the specific receptor regions likely to determine ligand binding and cast light on potential mechanisms underlying the unique specificity of ACTHR for binding of its natural ligand, ACTH, and not of closely related melanocortins.

Acknowledgements

These studies were supported by the Latvian Council of Science Grant 09.1280.01, Latvian State Research Program 8.14. Dr. Janis Klovins was supported by the Novo European Union Marie Curie Fellowship Programme and by OE och Edla Johanssons Vetenskapliga Stiftelse. Davids Fridmanis was supported by a FEBS Collaborative Experimental Scholarship for Central and Eastern Europe and the Novo Nordisk Foundation while working in Department of Neuroscience at Uppsala University, and by the European Social Fund while working in the Biomedical Research and Study Centre at the University of Latvia. Dr. Helgi Schiöth was supported by the Swedish Research Council. We thank Dace Pjanova and Ruta Bruvere for help with confocal microscopy experiments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.mce.2010.02.032.

References

Abdel-Malek, Z.A., 2001. Melanocortin receptors: their functions and regulation by physiological agonists and antagonists. Cell. Mol. Life Sci. 58 (3), 434–441.
Allolio, B., Reincke, M., 1997. Adrenocorticotropin receptor and adrenal disorders. Horm. Res. 47 (4–6), 273–278.
Argiolas, A., Melis, M.R., Murgia, S., Schioth, H.B., 2000. ACTH- and alpha-MSH-induced grooming, stretching, yawning and penile erection in male rats: site of action in the brain and role of melanocortin receptors. Brain Res. Bull. 51 (5), 425–431.
Boer, U., Neuschafer-Rube, F., Moller, U., Puschel, G.P., 2000. Requirement of N-glycosylation of the prostaglandin E2 receptor EP3beta for correct sorting to the plasma membrane but not for correct folding. Biochem. J. 350 Pt 3, 839–847.
Boston, B.A., 1999. The role of melanocortins in adipocyte function. Ann. N.Y. Acad. Sci. 885, 75–84.
Butler, A.A., 2006. The melanocortin system and energy balance. Peptides 27 (2), 281–290.
Chen, W., Kelly, M.A., Opitz-Araya, X., Thomas, R.E., Low, M.J., Cone, R.D., 1997. Exocrine gland dysfunction in MC5-R-deficient mice: evidence for coordinated regulation of exocrine gland function by melanocortin peptides. Cell 91 (6), 789–798.
Chhajlani, V., Wikberg, J.E., 1992. Molecular cloning and expression of the human melanocyte stimulating hormone receptor cDNA. FEBS Lett. 309 (3), 417–420.
Chida, D, Nakagawa, S., Nagai, S., Sagara, H., Katsumata, H., Imaki, T., Suzuki, H., Mitani, F., Ogishima, T., Shimizu, C., Kotaki, H., Kakuta, S., Sudo, K., Koike, T., Kubo, M., Iwakura, Y., 2007. Melanocortin 2 receptor is required for adrenal gland development, steroidogenesis, and neonatal gluconeogenesis. Proc. Natl. Acad. Sci. U.S.A. 104 (46), 18205–18210.

Chung, T.T., Webb, T.R., Chan, L.F., Cooray, S.N., Metherell, L.A., King, P.J., Chapple, J.P., Clark, A.J., 2008. The majority of ACTH receptor (MC2R) mutations found in Familial Glucocorticoid Deficiency type 1 lead to defective trafficking of the receptor to the cell surface. J. Clin. Endocrinol. Metab..
Clark, A.J, Cammas, F.M., 1996. The ACTH receptor. Baillieres Clin. Endocrinol. Metab. 10 (1), 29–47.
Cooray, S.N., Almiro Do Vale, I., Leung, K.Y., Webb, T.R., Chapple, J.P., Egertova, M., Cheetham, M.E., Elphick, M.R., Clark, A.J., 2008. The melanocortin 2 receptor accessory protein exists as a homodimer and is essential for the function of the melanocortin 2 receptor in the mouse y1 cell line. Endocrinology 149 (4), 1935–1941.
De Wied, D., Jolles, J., 1982. Neuropeptides derived from pro-opiocortin: behavioral, physiological, and neurochemical effects. Physiol. Rev. 62 (3), 976–1059.
Dunn, O.J, 1964. Multiple comparisons using rank sums. Technometrics 6, 241–252.
Eberle, A.N., 1988. Chem. Physiol. Mech. Action.
Elias, L.L., Huebner, A., Pullinger, G.D., Mirtella, A., Clark, A.J., 1999. Functional characterization of naturally occurring mutations of the human adrenocorticotropin receptor: poor correlation of phenotype and genotype. J. Clin. Endocrinol. Metab. 84 (8), 2766–2770.
Ellacott, K.L., Cone, R.D., 2004. The central melanocortin system and the integration of short- and long-term regulators of energy homeostasis. Recent Prog. Horm. Res. 59, 395–408.
Ellacott, K.L., Murphy, J.G., Marks, D.L., Cone, R.D., 2007. Obesity-induced inflammation in white adipose tissue is attenuated by loss of melanocortin-3 receptor signaling. Endocrinology 148 (12), 6186–6194.
Fan, W., Morrison, S.F., Cao, W.H., Yu, P., 2007. Thermogenesis activated by central melanocortin signaling is dependent on neurons in the rostral raphe pallidus (rRPa) area. Brain Res. 1179, 61–69.
Fan, W., Voss-Andreae, A., Cao, W.H., Morrison, S.F., 2005. Regulation of thermogenesis by the central melanocortin system. Peptides 26 (10), 1800–1813.
Fredriksson, R, Lagerstrom, M.C., Lundin, L.G., Schioth, H.B., 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol. Pharmacol. 63 (6), 1256–1272.
Frielle, T., Daniel, K.W., Caron, M.G., Lefkowitz, R.J., 1988. Structural basis of beta-adrenergic receptor subtype specificity studied with chimeric beta 1/beta 2-adrenergic receptors. Proc. Natl. Acad. Sci. U.S.A. 85 (24), 9494–9498.
Gantz, I, Konda, Y., Tashiro, T., Shimoto, Y., Miwa, H., Munzert, G., Watson, S.J., DelValle, J., Yamada, T., 1993a. Molecular cloning of a novel melanocortin receptor. J. Biol. Chem. 268 (11), 8246–8250.
Gantz, I., Miwa, H., Konda, Y., Shimoto, Y., Tashiro, T., Watson, S.J., DelValle, J., Yamada, T., 1993b. Molecular cloning, expression, and gene localization of a fourth melanocortin receptor. J. Biol. Chem. 268 (20), 15174–15179.
Gantz, I., Shimoto, Y., Konda, Y., Miwa, H., Dickinson, C.J., Yamada, T., 1994. Molecular cloning, expression, and characterization of a fifth melanocortin receptor. Biochem. Biophys. Res. Commun. 200 (3), 1214–1220.
Gupta, R, Brunak, S., 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. Pac. Symp. Biocomput., 310–322.
Gupta, R., Jung, E., Brunak, S., in preparation. Prediction of N-glycosylation sites in human proteins.
Haycock, J.W, Wagner, M., Morandini, R., Ghanem, G., Rennie, I.G., MacNeil, S., 1999. Alpha-MSH immunomodulation acts via rel/NF-kappa B in cutaneous and ocular melanocytes and in melanoma cells. Ann. N.Y. Acad. Sci. 885, 396–399.
Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., Pease, L.R., 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77 (1), 51–59.
Humphreys, M.H, 2004. Gamma-MSH, sodium metabolism, and salt-sensitive hypertension. Am. J. Physiol. Regul. Integr. Comp. Physiol. 286 (3), R417–R430.
Kobilka, B.K., Kobilka, T.S., Daniel, K., Regan, J.W., Caron, M.G., Lefkowitz, R.J., 1988. Chimeric alpha 2-, beta 2-adrenergic receptors: delineation of domains involved in effector coupling and ligand binding specificity. Science 240 (4857), 1310–1316.
Konda, Y, Gantz, I., DelValle, J., Shimoto, Y., Miwa, H., Yamada, T., 1994. Interaction of dual intracellular signaling pathways activated by the melanocortin-3 receptor. J. Biol. Chem. 269 (18), 13162–13166.
Kovalitskaia Yu, A., Kolobov, A.A., Kampe-Nemm, E.A., Iurovskii, V.V., Sadovnikov, V.B., Lipkin, V.M., Navolotskaia, E.V., 2008. Synthetic peptide KKRR corresponding to the human ACTH fragment 15–18 is an antagonist of the ACTH receptor. Bioorg. Khim. 34 (1), 29–35.
Lagerstrom, M.C, Klovins, J., Fredriksson, R., Fridmanis, D., Haitina, T., Ling, M.K., Berglund, M.M., Schioth, H.B., 2003. High affinity agonistic metal ion binding sites within the melanocortin 4 receptor illustrate conformational change of transmembrane region 3. J. Biol. Chem. 278 (51), 51521–51526.
Lasaga, M., Debeljuk, L., Durand, D., Scimonelli, T.N., Caruso, C., 2008. Role of alpha-melanocyte stimulating hormone and melanocortin 4 receptor in brain inflammation. Peptides 29 (10), 1825–1835.
Lu, X.Y, 2001. Role of central melanocortin signaling in eating disorders. Psychopharmacol. Bull. 35 (4), 45–65.
Metherell, L.A., Chapple, J.P., Cooray, S., David, A., Becker, C., Ruschendorf, F., Naville, D., Begeot, M., Khoo, B., Nurnberg, P., Huebner, A., Cheetham, M.E., Clark, A.J., 2005. Mutations in MRAP, encoding a new interacting partner of the ACTH receptor, cause familial glucocorticoid deficiency type 2. Nat. Genet. 37 (2), 166–170.
Mountjoy, K.G, Robbins, L.S., Mortrud, M.T., Cone, R.D., 1992. The cloning of a family of genes that encode the melanocortin receptors. Science 257 (5074), 1248–1251.
Naville, D., Penhoat, A., Durand, P., Begeot, M., 1999. Three steroidogenic factor-1 binding elements are required for constitutive and cAMP-regulated expression of the human adrenocorticotropin receptor gene. Biochem. Biophys. Res. Commun. 255 (1), 28–33.
Noon, L.A, Franklin, J.M., King, P.J., Goulding, N.J., Hunyady, L., Clark, A.J., 2002. Failed export of the adrenocorticotrophin receptor from the endoplasmic reticulum in non-adrenal cells: evidence in support of a requirement for a specific adrenal accessory factor. J. Endocrinol. 174 (1), 17–25.
Penhoat, A., Jaillard, C., Saez, J.M., 1989. Corticotropin positively regulates its own receptors and cAMP response in cultured bovine adrenal cells. Proc. Natl. Acad. Sci. U.S.A. 86 (13), 4978–4981.
Penhoat, A, Naville, D., El Mourabit, H., Buronfosse, A., Berberoglu, M., Ocal, G., Tsigos, C., Durand, P., Begeot, M., 2002. Functional relationships between three novel homozygous mutations in the ACTH receptor gene and familial glucocorticoid deficiency. J. Mol. Med. 80 (7), 406–411.
Petrecca, K., Atanasiu, R., Akhavan, A., Shrier, A., 1999. N-linked glycosylation sites determine HERG channel surface membrane expression. J. Physiol. 515 (Pt 1), 41–48.
Pfaus, J.G, Shadiack, A., Van Soest, T., Tse, M., Molinoff, P., 2004. Selective facilitation of sexual solicitation in the female rat by a melanocortin receptor agonist. Proc. Natl. Acad. Sci. U.S.A. 101 (27), 10201–10204.
Rinne, P., Harjunpaa, J., Scheinin, M., Savontaus, E., 2008. Blood pressure regulation and cardiac autonomic control in mice overexpressing alpha- and gamma-melanocyte stimulating hormone. Peptides 29 (11), 1943–1952.
Roy, S, Rached, M., Gallo-Payet, N., 2007. Differential regulation of the human adrenocorticotropin receptor [melanocortin-2 receptor (MC2R)] by human MC2R accessory protein isoforms alpha and beta in isogenic human embryonic kidney 293 cells. Mol. Endocrinol. 21 (7), 1656–1669.
Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., Quackenbush, J., 2003. TM4: a free, open-source system for microarray data management and analysis. Biotechniques 34 (2), 374–378.
Schimmer, B.P, Kwan, W.K., Tsao, J., Qiu, R., 1995. Adrenocorticotropin-resistant mutants of the Y1 adrenal cell line fail to express the adrenocorticotropin receptor. J. Cell. Physiol. 163 (1), 164–171.
Schioth, H.B., Fredriksson, A., Carlsson, C., Yook, P., Muceniece, R., Wikberg, J.E., 1998a. Evidence indicating that the extracellular loops of the mouse MC5 receptor do not participate in ligand binding. Mol. Cell. Endocrinol. 139 (1–2), 109–115.
Schioth, H.B., Muceniece, R., Szardenings, M., Prusis, P., Wikberg, J.E., 1996. Evidence indicating that the TM4, EL2, and TM5 of the melanocortin 3 receptor do not participate in ligand binding. Biochem. Biophys. Res. Commun. 229 (3), 687–692.
Schioth, H.B., Yook, P., Muceniece, R., Wikberg, J.E., Szardenings, M., 1998b. Chimeric melanocortin MC1 and MC3 receptors: identification of domains participating in binding of melanocyte-stimulating hormone peptides. Mol. Pharmacol. 54 (1), 154–161.
Sebag, J.A., Hinkle, P.M., 2007. Melanocortin-2 receptor accessory protein MRAP forms antiparallel homodimers. Proc. Natl. Acad. Sci. U.S.A. 104 (51), 20244–20249.
Sebag, J.A, Hinkle, P.M., 2008. Regions of melanocortin2 receptor accessory protein necessary for dual topology and MC2 receptor trafficking and signaling. J. Biol. Chem..
Shimozono, S., Miyawaki, A., 2008. Engineering FRET constructs using CFP and YFP. Methods Cell Biol. 85, 381–393.
Simpson, E.R., Waterman, M.R., 1988. Regulation of the synthesis of steroidogenic enzymes in adrenal cortical cells by ACTH. Annu. Rev. Physiol. 50, 427– 440.
Sina, M., Hinney, A., Ziegler, A., Neupert, T., Mayer, H., Siegfried, W., Blum, W.F., Remschmidt, H., Hebebrand, J., 1999. Phenotypes in three pedigrees with autosomal dominant obesity caused by haploinsufficiency mutations in the melanocortin-4 receptor gene. Am. J. Hum. Genet. 65 (6), 1501–1507.
Spencer, J.D, Schallreuter, K.U., 2008. Regulation of pigmentation in human epidermal melanocytes by functional high affinity {beta}-MSH/MC4-R signalling. Endocrinology.
Starowicz, K., Przewlocka, B., 2003. The role of melanocortins and their receptors in inflammatory processes, nerve regeneration and nociception. Life Sci. 73 (7), 823–847.
Theodoropoulou, M.C, Bagos, P.G., Spyropoulos, I.C., Hamodrakas, S.J., 2008. gpDB: a database of GPCRs, G-proteins, effectors and their interactions. Bioinformatics 24 (12), 1471–1472.
Thiboutot, D., Sivarajah, A., Gilliland, K., Cong, Z., Clawson, G., 2000. The melanocortin 5 receptor is expressed in human sebaceous glands and rat preputial cells. J. Invest. Dermatol. 115 (4), 614–619.
Webb, T.R., Chan, L., Cooray, S.N., Cheetham, M.E., Chapple, J.P., Clark, A.J., 2008. Distinct MRAP domains are required for MC2R interaction and promotion of receptor trafficking. Endocrinology.
Wessells, H., Gralnek, D., Dorr, R., Hruby, V.J., Hadley, M.E., Levine, N., 2000. Effect of an alpha-melanocyte stimulating hormone analog on penile erection and sexual desire in men with organic erectile dysfunction. Urology 56 (4), 641–646.
Williams, G, Bing, C., Cai, X.J., Harrold, J.A., King, P.J., Liu, X.H., 2001. The hypothalamus and the control of energy homeostasis: different circuits, different purposes. Physiol. Behav. 74 (4–5), 683–701.
Yeo, G.S., Farooqi, I.S., Challis, B.G., Jackson, R.S., O'Rahilly, S., 2000. The role of melanocortin signalling in the control of body weight: evidence from human and murine genetic models. QJM 93 (1), 7–14.

Supplementary Table 1. Primers and source receptors used in the design of chimeric receptors and sequences of standard sequencing primers.

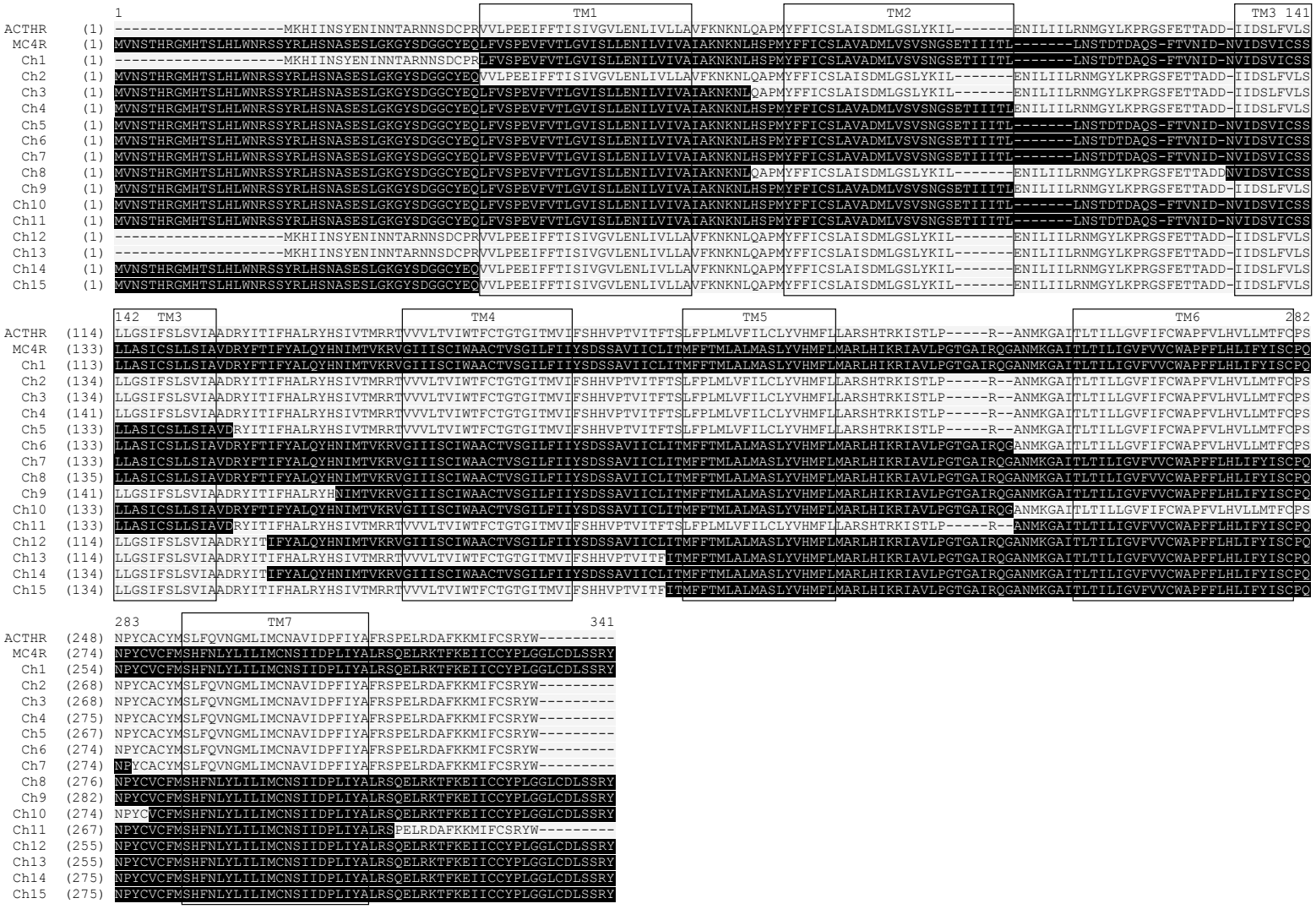
Construct	Reverse mutant primer (5'-3')	Forward mutant primer (5'-3')	5'-end Source Receptor	3'end Source Receptor
Ch1	AGACAAAAAGACGAGGACAGTCGGAATTAT	CTGTCCTCGTCTTTTTGTCTCTCCTGAGGT	MC2R	MC4R
Ch2	GCAAAACCACTTGCTCGTAGCACCCTCCAT	CTACGAGCAAGTGGTTTTGCCGGAGGAGAT	MC4R	MC2R
Ch3	GCCTGGAGATTCTTGTTCTTGGCTAT	ATAGCCAAGAACAAGAATCTCCAGG	MC4R	MC2R
Ch4	GATATTTTCCAGGGTGATGATAATGGTTTC	TCATCACCCTGGAAAATATCCTGATCATA	MC4R	MC2R
Ch5	GATGGTGATGTAGCGGTCCACTGCAATTGAAAG	CTTTCAATTGCAGTGGACCGCTACATCACCATC	MC4R	MC2R
Ch6	GGATGGTCAGTGTGATCGCTCCCTTC	GAAGGGAGCGATCACACTGACCATCC	MC4R	MC2R
Ch7	CAGGCGCAGTATGGATTCTGAGGACAA	CCTCAGAATCCATACTGCGCCTGCTAC	MC4R	MC2R
Ch8	CAATGACATTGTCATCGGCTGTGGTTTCAA	AGCCGATGACAATGTCATTGACTCGGTGAT	Ch3	MC4R
Ch9	TCATAATGTTGTGGTACCGCAGTGCGTGGA	GCGGTACCACAACATTATGACAGTTAAGCG	Ch4	MC4R
Ch10	TGAAGCACACGCAGTAGGGGTTACTTGGGC	CCCCTACTGCGTGTGCTTCATGTCTCACTT	Ch6	MC4R
Ch11	AAGGTAATCGCTCCCTTCATGTTGGCTCTGGGGAG	CTCCCCAGAGCCAACATGAAGGGAGCGATTACCTT	Ch5	MC4R
Ch12	AGAAGATAGTGATGTAGCGGTCCGCAGC	CCGCTACATCACTATCTTCTATGCTCTCCA	MC2R	MC4R
Ch13	ACATGGTGATGAAGGTGATCACTGTGGGC	GATCACCTTCATCACCATGTTCTTCACCATG	MC2R	MC4R
Ch14	GCAAAACCACTTGCTCGTAGCACCCTCCAT	CTACGAGCAAGTGGTTTTGCCGGAGGAGAT	MC4R	Ch12
Ch15	GCAAAACCACTTGCTCGTAGCACCCTCCAT	CTACGAGCAAGTGGTTTTGCCGGAGGAGAT	MC4R	Ch13

Sequencing Primer	Primer Sequence
CEP-F	GACAGAGAAGACTCTTGCGTTTCTGAT
CEP-R	CGGCCTCTAGAGTCGACTCATTTGTCA
GFP-seq-R	GCCGTCCAGCTCGACCAGGATGGGCACC

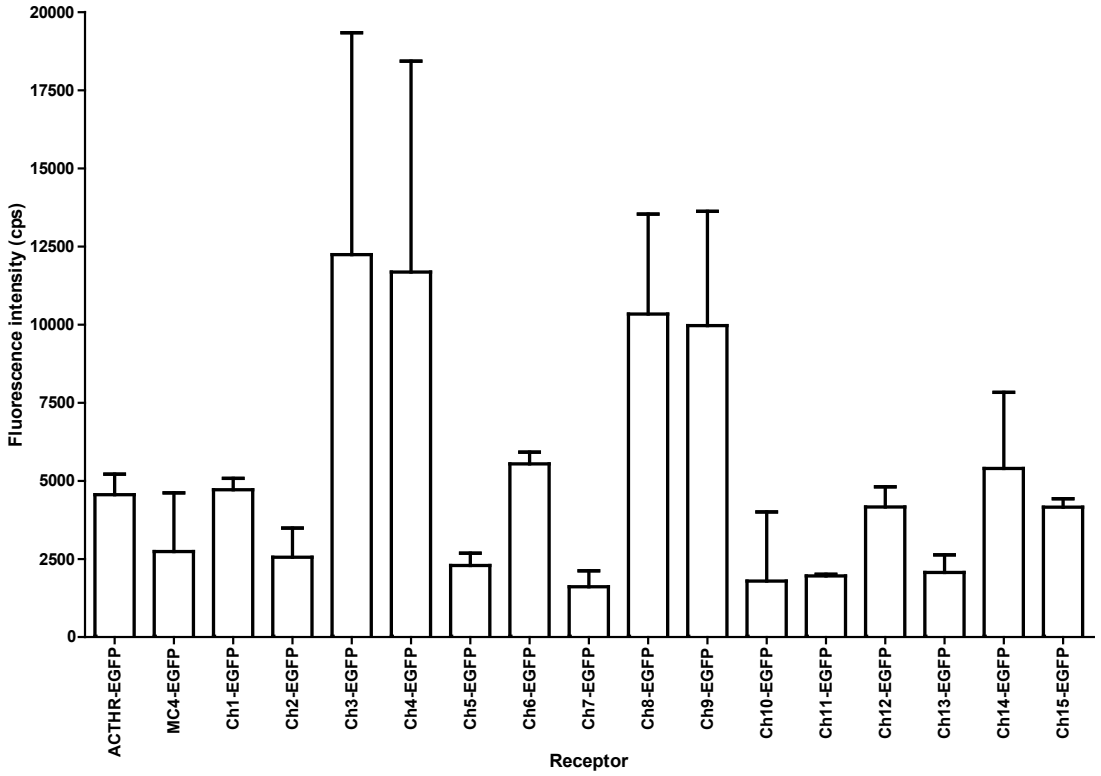
Supplementary Table 2. Glycolysation potential of Asn-Xaa-Ser/Thr N-glycosylation sequons (where Xaa is any amino acid except proline) of different melanocortin receptors from five different organisms. Analysis was performed using software at http://www.cbs.dtu.dk/services/NetNGlyc/output.php)

Receptor	Potential			
	1st site	2nd site	3rd site	4th site
MC1R-Danio_rerio	0.6917	0.6856	0.7544	0.7142
MC1R-Gallus_gallus	0.5553	0.6091	0.4963	
MC1R-Homo_sapiens	0.3064	0.6703		
MC1R-Mus_musculus	0.7467	0.6548		
MC1R-Takifugu_rubripes	0.7336	0.6467		
ACTHR-Danio_rerio				
ACTHR-Gallus_gallus	0.7973	0.6886		
ACTHR-Homo_sapiens	0.6352	0.525		
ACTHR-Mus_musculus	0.6659	0.4468		
ACTHR-Takifugu_rubripes	0.6843	0.6108		
MC3R-Danio_rerio	0.6036	0.7341	0.7267	
MC3R-Gallus_gallus	0.7129	0.8022	0.7106	
MC3R-Homo_sapiens	0.5997	0.669	0.6779	
MC3R-Mus_musculus	0.6129	0.7494	0.7078	
MC4R-Danio_rerio	0.6502	0.5473		
MC4R-Gallus_gallus	0.7666	0.5564		
MC4R-Homo_sapiens	0.7128	0.7492	0.6507	
MC4R-Mus_musculus	0.6817	0.7343	0.6569	
MC4R-Takifugu_rubripes	0.7518	0.6586		
MC5aR-Danio_rerio	0.1754	0.7163		
MC5bR-Danio_rerio	0.6808	0.7637	0.6173	
MC5R-Gallus_gallus	0.7249	0.679	0.7232	
MC5R-Homo_sapiens	0.5768	0.737	0.6586	0.5978
MC5R-Mus_musculus	0.7493	0.7493	0.74	0.4516
MC5R-Takifugu_rubripes	0.5917	0.685	0.6446	
MRAP-Homo_sapiens	0.7290	0.5636		
MRAP-ACTHR-complex	0.7290	0.5636	0.6352	0.525

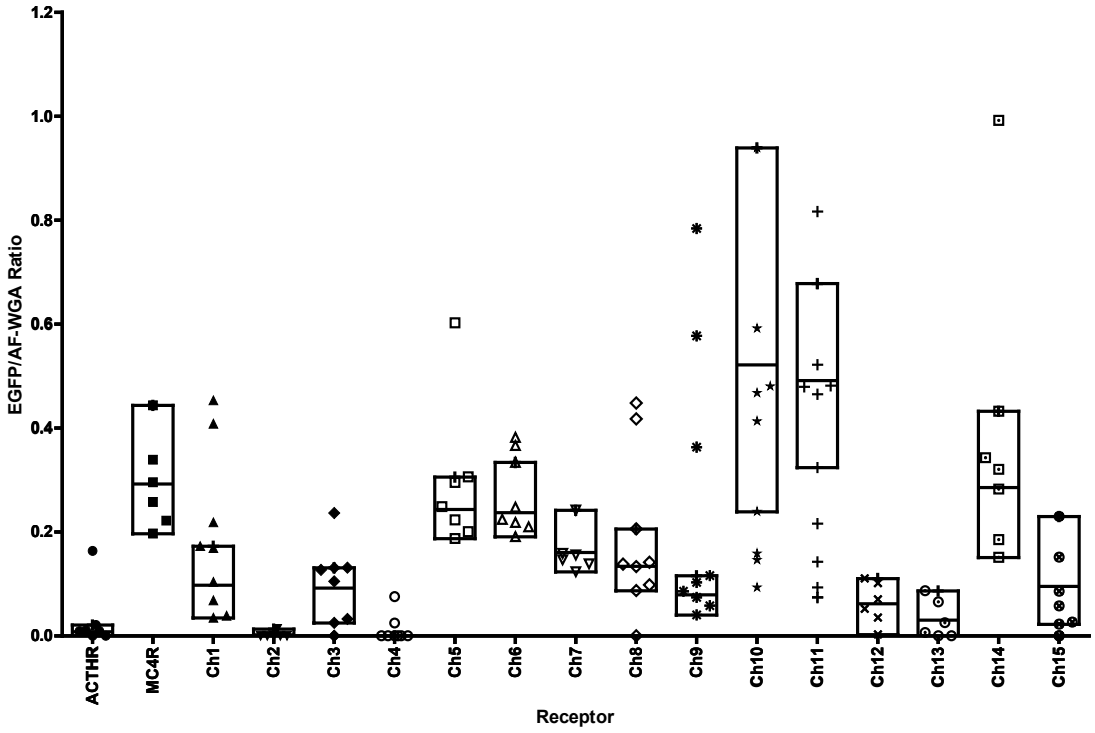
Supplementary Figure 1. Sequence alignment of all chimeras produced in this study. Regions of ACTHR are unshaded; MC4R regions are shaded in black. The approximate extents of the transmembrane domains are boxed.



Supplementary Figure 2. Receptor-EGFP fusion-protein expression levels acquired by direct EGFP fluorescence measurements.



Supplementary Figure 3. Median values of EGFP/AF-WGA fluorescence intensity ratios from selected points on the plasma membrane of each cell. The median values for each cell that met the criteria described in Materials and Methods and that were used for analysis of membrane-expression levels are enclosed in rectangles. Horizontal lines crossing each rectangle represent the means of the median values employed.



4 DISCUSSION

The aim of the research included studies of the MCR family evolution and the first task was to identify MCRs in bony fishes *Takifugu rubripes* and *Danio rerio*. Orthologs of all five mammalian melanocortin receptor subtypes were identified in *Danio rerio* while only four were found in *Takifugu rubripes* (there were no orthologs for MC3R). After identification MCRs coding sequences of *Takifugu rubripes* were cloned, expressed in mammalian cell lines and tested for their pharmacological properties. At that time it was already established that mammalian ACTHR is binding solely ACTH and can only be expressed in adrenal cell lines or cell lines endogenously expressing other MCRs. It was not a surprise therefore that ACTHR of *Fugu* displayed similar properties, interesting however was its expression pattern, ACTHR like other MCRs of *Fugu* were predominantly expressed in CNS and only low level expression was detected in the head of the kidney (adrenal cells of the fishes are found intermingling with kidney tissue). These findings are varying significantly from recent results acquired during characterisation of MCRs in european river lamprey (*Lampetra fluviatilis*) (Haitina *et al.* 2007). This jawless fish is believed to be very similar to early vertebrates and has two MCRs: MCAR and MCBR none of which is expressed in CNS, only in skin (MCAR, MCBR) and liver (MCBR) and at lower levels in heart and muscle (MCBR). The best explanation here would be that MCRs acquired their new regulatory functions soon after the final formation of this GPCR family, which occurred after the split of vertebrate and jawless fish lineages. This explanation is also supported by the phylogenetical analysis of MCAR and MCBR. Amino-acid sequence of MCAR turned out to be more similar to MC1R while amino-acid sequence of MCBR to MC4R, which in mammalians is predominantly expressed in CNS. Further, data from lamprey MCRs study are fitting very well in our proposed MCR family formation model (results section, first chapter, Fig. 9), although, there are no information confirming location of these genes in one locus similar to vertebrate ACTHR/MC5R locus. Additionally in regard to MC4Rs being more similar to MCBR than closely related MC5R and MC3R, some minor modifications should be implemented. In the improved version MC4R has been swapped with the MC3R (Fig. 14) and now is considered to be formed during the first duplication of ancestral MCBR gene, while MC3R now is considered to be "the youngest" member of the family. Besides sequence similarity this conclusion is also supported, although unconvincingly, by the study described in second chapter of results

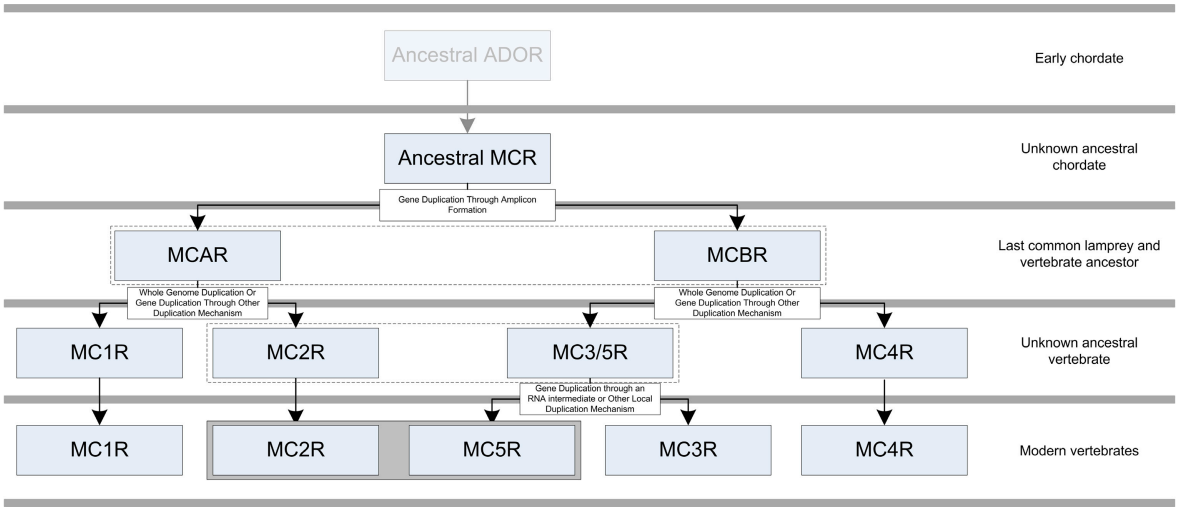


Figure 14. Schematical presentation of hypothetical chain of events that could have occurred during the formation of MCR family. MCR genes are depicted as shaded boxes. The duplication events are depicted as unshaded boxes located upon the arrow-line. The first stage was the formation of first melanocortin receptor from Adenosine binding receptor (ADOR). The second stage of events included the duplication of ancestral MCR. Designation of duplicated receptor genes as MCAR and MCBR is adopted from European river lamprey (*Lampetra fluviatilis*) which has only two MCRs, one similar to MC1R and ACTHR and other to MC3R, MC4R and MC5R, thus it is plausible that they are highly similar to ancestral MC1/2R and MC3/4/5R. Third stage of events included the duplication of MCAR gene to MC1R and MC2R genes and duplication MCBR gene to MC3/5R and MC4R genes. MC4R is believed to form during this duplication event because of its higher similarity to lamprey MCBR. Fourth stage of events included duplication MC3/5R gene to MC3R and MC5R genes. Enclosed in shaded frame are ACTHR and MC5R genes which are detected to be localized in one locus and have an opposite direction of transcription in all analyzed vertebrates. Since such situation is characteristic for local duplication via amplicon formation and ACTHR and MC5R are two least similar MCRs it is highly plausible that MC2/5R locus must be a direct descendant (or its copy) of original MCR locus which was formed during the first stage of events. Enclosed in dashed line frame are genes in ancestral MC2/5R locuses. Genomic organization of lamprey MCR genes is yet uncovered and might confirm the existence of MCA/BR locus. Existence of ancestral MC2/3/5R locus might never be confirmed since there are no "living early *gnathostomes*". The whole genome duplication is widely believed to take place upon transition from *craniates* to *gnathostomes* which in our proposed chain of events correspond to second stage.

section where MC3Rs whenever clustered were always located on the separate branch along with MC5Rs indicating slightly higher sequence similarity and thus more recent common ancestry. (Fig. 15)

Surprising, however, was the structure of *Fugu* ACTHR and MC5R genes. Both receptor genes contained introns which proved a unique property of *Fugu* MCR coding genes. Even more, one of the MC5R introns and ACTHR intron were located at the same position of conservative "DRY" motif coding region. This discovery was even more intriguing, because there were observations of vertebrate Rhodopsin GPCRs genes being predominantly intronless. Second chapter of the results section describes more detailed investigation of this phenomenon where we came to conclusion that genes encoding the

Rhodopsin GPCRs are predominantly intronless due to the rapid formation of new genes during the early stages of vertebrate evolution, suggesting that introns of *Fugu* ACTHR and MC5R encoding genes must be of a recent origin. During this study we came across a number of Rhodopsin GPCRs, located in various clusters of phylogenetic tree developed by Fredriksson et al. (2003), with intron at the same position of DRY motif and what's more interesting, there were receptors (also located in various clusters of phylogenetic tree) with intron in region that is encoding DRY related motif (motif with similar or identical amino-acid sequence, but located elsewhere, not at conservative position). The most obvious explanation of this DRY intron phenomenon as suggested by Bryson-Richardson et al. (2004) would be that intron at this position is of ancient origin and already existed in ancestral GPCR. However our findings of relatively recently inserted *Fugu* ACTHR and MC5R introns and existence of DRY related introns suggested that at least some of these introns have inserted recently and it is possible that this insertion occurred through proto-splice site mechanism.

During the study described in second chapter of the results section we also acquired some indications regarding the origin of MCR family. Here while attempting to divide the GPCR of analysed organisms according to GRAFS classification system we noticed that not all of clusters were represented in "lower animals". Detailed investigation revealed that, in fact, there are only 13 isolated receptor groups which were present in all analysed vertebrates and in at least two of three analysed invertebrate. These groups of receptors were subsequently designated as Ancient Receptor Groups (ARG). The search for such evolutionary ancient receptors, through the analysis of maximum parsimony phylogenetical comparison trees in the MECA group, which is formed by MC, Endothelial differentiation (EDG), Cannabinoid receptors (CN) and Adenosine binding (ADO) receptors, identified that ADOR is found in all analyzed organisms (Fig. 15). Interestingly, the sequence similarity analysis reveals that human MCRs are more similar to human ADORs (HMM generated model identity 29%) than to other receptors families of MECA cluster (HMM generated model identity 27% with EDGR, 26% with GPR3/6/12 and 24% with CNR). It is thus highly probable that ancestral MCR of early chordates most probably have evolved from receptor of ADOR family that belongs to ARG (Fig. 14). This speculation, if presumed to be true, shows how actually flexible is GPCR system in general, because this transition also involved important change of the type of specifically bound ligand from nucleotide to peptide.

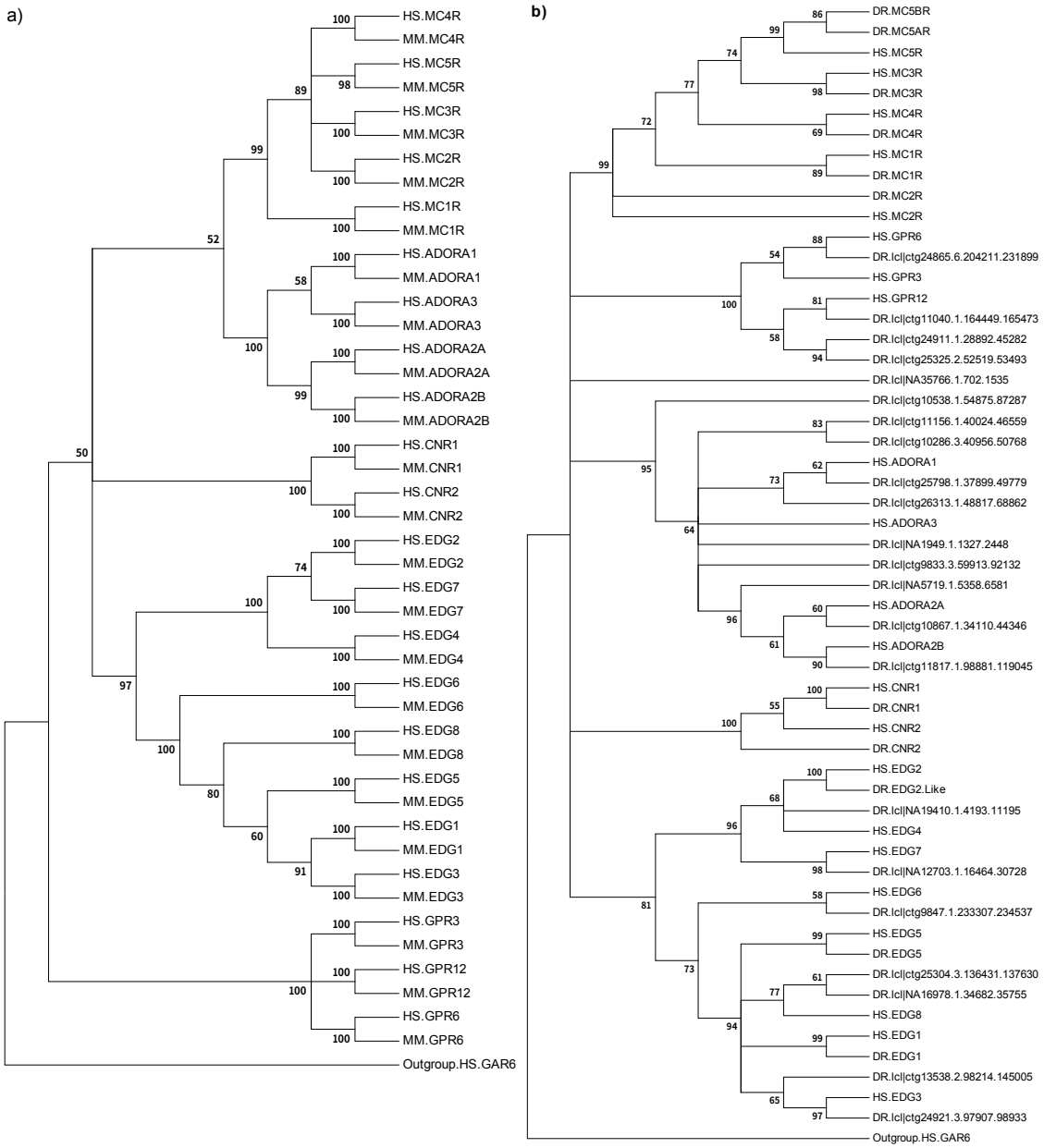
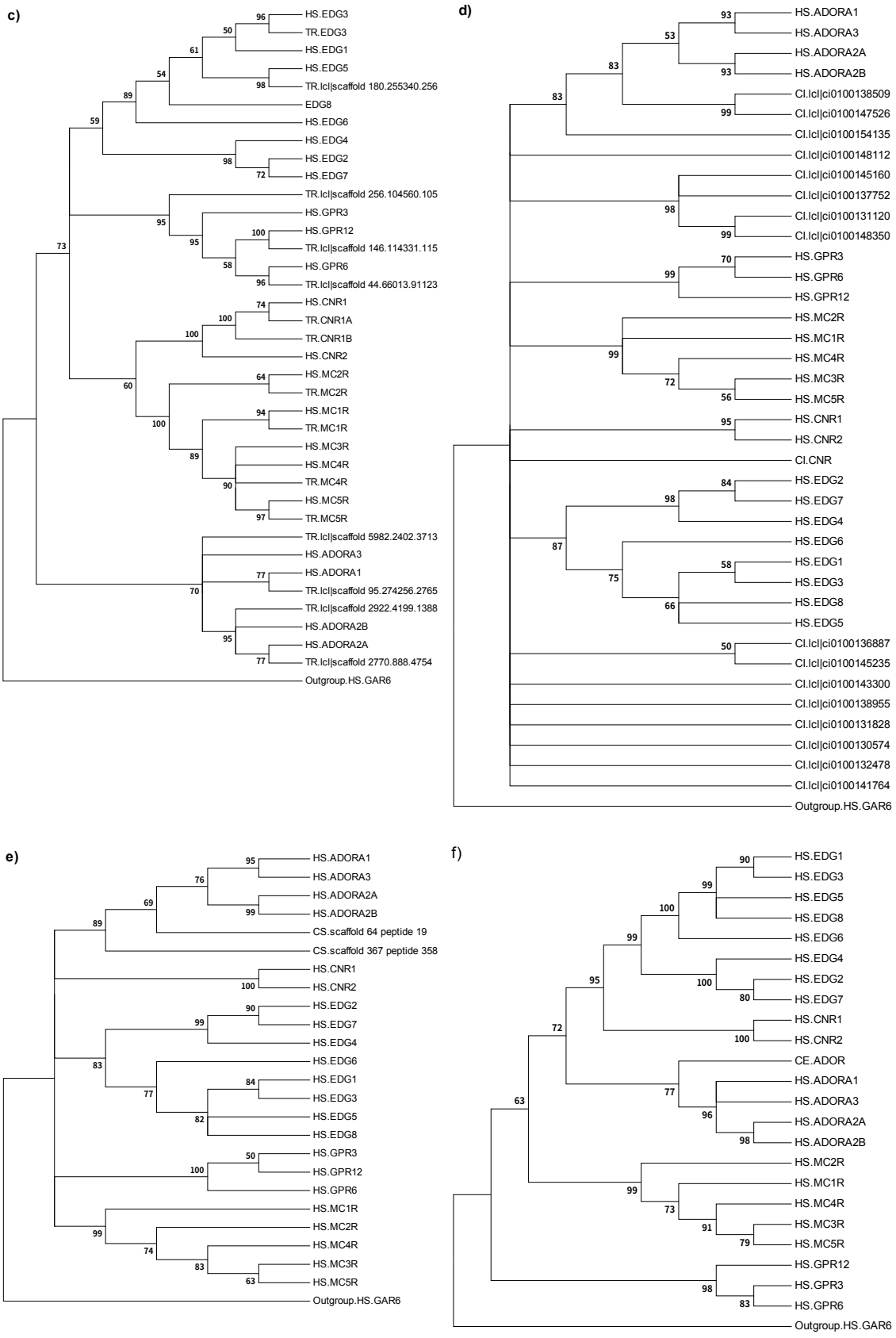


Figure 15. Maximum parsimony phylogenetical comparison trees of GRAFS classification system MECA group. In these trees human - *Homo sapiens* (HS) MECA group receptors are phylogenetically co analyzed with the MECA group receptors of a) house mouse - *Mus musculus* (MM), b) zebrafish - *Danio rerio* (DR), c) fugu -*Takifugu rubripes* (TR), d) transparent sea squirt - *Ciona intestinalis* (CI), e) pacific transparent sea squirt - *Ciona savignyi* (CS) and f) transparent nematode *Caenohabditis elegans* (CE). This analysis was used to identify receptor families in human that have orthologs in other analyzed vertebrate and invertebrate organisms. As clearly visible orthologs of ADOR receptor family can be identified in all analyzed organisms, indicating this receptor family being ancestral to all other MECA families. In these trees it is also visible that MC3R, whenever clustered, is always located on one branch with MC5R indicating higher sequence similarity


c)
HS.EDG3
TR.EDG3
HS.EDG1
HS.EDG5
TR.lcl|scaffold 180.255340.256
EDG8
HS.EDG6
HS.EDG4
HS.EDG2
HS.EDG7
TR.lcl|scaffold 256.104560.105
HS.GPR3
HS.GPR12
TR.lcl|scaffold 146.114331.115
HS.GPR6
TR.lcl|scaffold 44.66013.91123
HS.CNR1
TR.CNR1A
TR.CNR1B
HS.CNR2
HS.MC2R
TR.MC2R
HS.MC1R
TR.MC1R
HS.MC3R
HS.MC4R
TR.MC4R
HS.MC5R
TR.MC5R
TR.lcl|scaffold 5982.2402.3713
HS.ADORA3
HS.ADORA1
TR.lcl|scaffold 95.274256.2765
TR.lcl|scaffold 2922.4199.1388
HS.ADORA2B
HS.ADORA2A
TR.lcl|scaffold 2770.888.4754
Outgroup.HS.GAR6
d)
HS.ADORA1
HS.ADORA3
HS.ADORA2A
HS.ADORA2B
Cl.lcl|ci0100138509
Cl.lcl|ci0100147526
Cl.lcl|ci0100154135
Cl.lcl|ci0100148112
Cl.lcl|ci0100145160
Cl.lcl|ci0100137752
Cl.lcl|ci0100131120
Cl.lcl|ci0100148350
HS.GPR3
HS.GPR6
HS.GPR12
HS.MC2R
HS.MC1R
HS.MC4R
HS.MC3R
HS.MC5R
HS.CNR1
HS.CNR2
Cl.CNR
HS.EDG2
HS.EDG7
HS.EDG4
HS.EDG6
HS.EDG1
HS.EDG3
HS.EDG8
HS.EDG5
Cl.lcl|ci0100136887
Cl.lcl|ci0100145235
Cl.lcl|ci0100143300
Cl.lcl|ci0100138955
Cl.lcl|ci0100131828
Cl.lcl|ci0100130574
Cl.lcl|ci0100132478
Cl.lcl|ci0100141764
Outgroup.HS.GAR6
e)
HS.ADORA1
HS.ADORA3
HS.ADORA2A
HS.ADORA2B
CS.scaffold 64 peptide 19
CS.scaffold 367 peptide 358
HS.CNR1
HS.CNR2
HS.EDG2
HS.EDG7
HS.EDG4
HS.EDG6
HS.EDG1
HS.EDG3
HS.EDG5
HS.EDG8
HS.GPR3
HS.GPR12
HS.GPR6
HS.MC1R
HS.MC2R
HS.MC4R
HS.MC3R
HS.MC5R
Outgroup.HS.GAR6
f)
HS.EDG1
HS.EDG3
HS.EDG5
HS.EDG8
HS.EDG6
HS.EDG4
HS.EDG2
HS.EDG7
HS.CNR1
HS.CNR2
CE.ADOR
HS.ADORA1
HS.ADORA3
HS.ADORA2A
HS.ADORA2B
HS.MC2R
HS.MC1R
HS.MC4R
HS.MC3R
HS.MC5R
HS.GPR12
HS.GPR3
HS.GPR6
Outgroup.HS.GAR6

As mentioned previously experiments with fugu ACTHR revealed that it exhibits the same properties regarding the expression and ligand binging selectivity, which indicated the occurrence of this specialisation on an early stage of vertebrate evolution. When thinking of this in a broader picture it seems obvious to presume the presence of MRAP, which actually ensures correct transportation and ligand recognition of mammalian ACTHR, in bony fishes. Recent finding of gene encoding MRAP2 in zebrafish (Chan *et al.* 2009) confirms that this protein was an essential part of melanocortin system already prior the split of mammalian and teleost lineages. However this finding is accentuating the existence of two important questions: When MRAP was introduced in melanocortin system and how this high expression and ligand recognition specificity of the ACTHR is achieved? In regard to the first question some speculations can be acquired while summarising the information from first chapter of results section, previously mentioned lamprey study and one of recently published studies, authors of which declare MRAP and MRAP2 being bidirectional regulators of melanocortin receptor trafficking, which couple not only to ACTHR, but also to other members of MCR family, thus down-regulating their membrane expression (Chan *et al.* 2009). This later finding if combined with the fact of lamprey receptors not being ACTH specific, thus not requiring MRAP, suggests that MRAP, most probably, joined MC system only after the formation of first four MC receptors and after integration of MC system in regulation mechanisms of CNS. As for its initial functions those, most probably, were the regulation of MCR membrane transport. Further, from previously described scenario it seems to be obvious that specialisation of ACTHR, must have occurred only after the introduction of MRAP in to the MC system and my speculation would be that ACTHR expression specificity was the first one to occur since it is more close to the proposed initial function of MRAP. (Fig. 16)

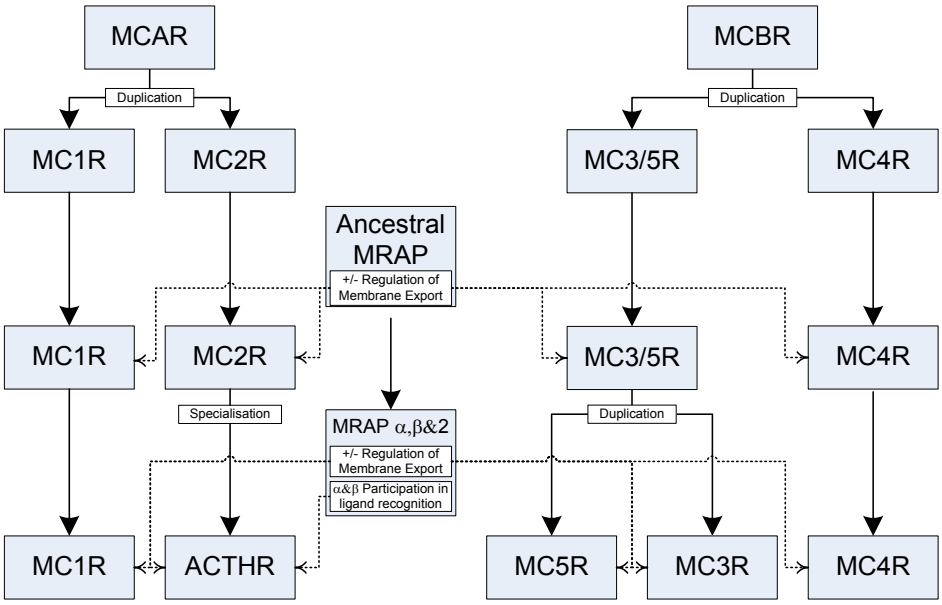


Figure 16. Schematical presentation of hypothetical chain of events that could have occurred during the introduction of MRAP to MC system. Proteins of MC system are depicted as shaded boxes, their specific functions as unshaded boxes inside box representing specific protein and dashed arrow-lines are pointing to targets of specific function. Time line flows for each protein are depicted as solid arrow-lines and specific events occurring on them are depicted as unshaded boxes located upon the arrow-line. Since none of european river lamprey MCRs are expressed in CNS and none display specificity to ACTH it is highly plausible that MRAP was introduced to MC system after the split of mammalian and lamprey lineages and possibly after the formation of first four. It has been shown that, besides coupling to ACTHR, MRAP also couples to other MCRs acting as bidirectional regulator, thus initial function of MRAP must have been regulation of MCRs membrane trafficking. Specialization of MC2R (transition to ACTHR) must have occurred after the introduction of MRAP and expression specificity must have been the first one to occur.

The answer to the second question, however, cannot be acquired through merely theoretical means and since there were no publications of studies addressing it we decided to investigate this issue ourselves. Our approach, as described in third chapter of results section, was to create a number of chimerical ACTH/MC4 receptors and thus identify receptor domains responsible for ligand recognition and expression specificity. The first serious problem was the lack of rapid and effective quantification system for membrane proteins, because most of already developed methods involved either direct approach of protein extraction followed quantification (Scheurer *et al.* 2005) or, when using fluorescent labels, quantification of total cell emission (Vitko *et al.* 2007) and both of these approaches are time consuming. To solve this problem we developed new – quantification of cell-membrane co-localisation method. Our approach involved acquisition of confocal microscopy images of receptor-EGFP fusion-protein expressing cells and measurement of EGFP emission at one point of cell membrane. Since such single measurement is not sufficient to estimate the average emission of whole membrane and is even less sufficient

to estimate the average emission of protein expressing cell population, we measured emission intensity at 20 points of cell membrane for six cells from at least two transfection experiments (analyzing three or less cells per one experiment), thus in total acquiring 120 measurements for each expressed receptor. Further results from various receptors were compared, searching for significant differences in emission levels. The method itself turned out to be rapid, effective and reliable for analysis of a large number of microscopy experiments in relatively short period of time, however since it involves the estimation of emission differences the acquired quantification is relative and thus it requires the usage of references. In our case, the unchanged MC4R and ACTHR EGFP fusion-proteins were serving as such maximum and minimum references respectively. The analysis of chimerical ACTH/MC4 receptors revealed that, according to relative amount of receptors in membrane of the cell, they all can be subdivided in three groups. First group united receptors with fluorescence emission level at cell membrane similar to MC4R, second group included receptors that resembled ACTHR like fluorescence emission levels and third group of receptors with fluorescence emission levels lower than MC4R, but higher than ACTHR. Strict division of chimeric receptors in three groups was extraordinary and suggested the presence of more than one membrane trafficking regulation element within the structure of ACTHR, because intermediate states between maximum and minimum are usually achieved through existence of several regulation mechanisms. Detailed analysis revealed that indeed there are at least two elements; one located within N-terminal tail of the receptor and other within region formed by TM3 and TM4. Since extracellular part of the membrane proteins, which in the case of ACTHR includes N-terminus, is almost always being glycosylated and usually at multiple sites, the presence of strong glycosylation signals is commonly associated with membrane export promoting. In-silico analysis of N-termini of ACTHR from various species predicted only two glycosylation sites and both with low glycosylation potential, thus we speculated that such low glycosylation potential must be significantly decreasing the membrane transport of ACTHR molecules.

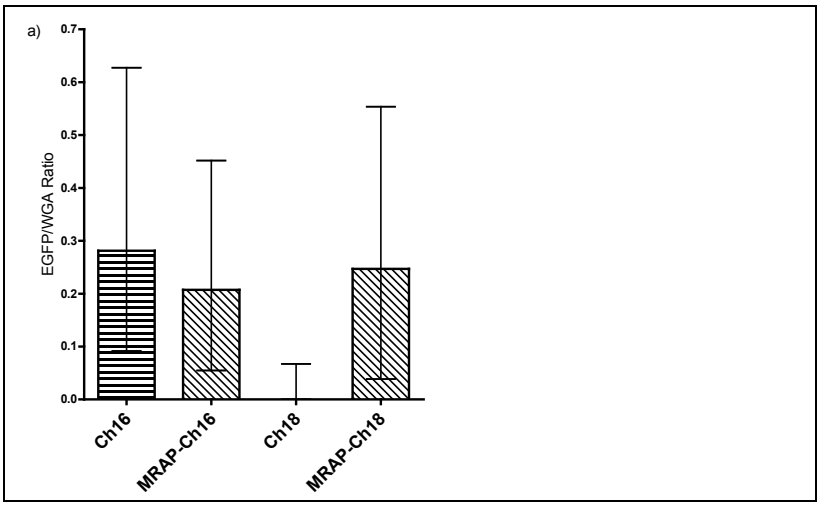


Figure 17. a) Graph representing medians with interquartile ranges of EGFP/AF-WGA fluorescence intensity ratio at cell plasma membrane for tested receptors. Bars with angled lines correspond to the receptors co-expressed with MRAP-β. b) Schematic presentation of the chimeric receptors. Regions from the MC4R are shaded grey, regions from the ACTHR are shaded black.

Because of a sequence based signalling nature the identification of N-termini as region containing regulation elements was rather simple, from logical point of view however, it was a lot harder to identify the location of second membrane transport regulation element. We encountered the difficulties to identify this region because it was a tertiary structure based signal which is formed only when both TM3 and TM4 of ACTHR are present. Our speculation regarding the nature of this element is that TM3 and TM4 are forming intracellular retention signal that is masked or disrupted after receptor coupling to MRAP. Our recent-unpublished results are also supporting these findings and as it is clearly seen in Fig 17, there was no fluorescence emission detected in the plasmatic membrane of BHK cells expressing Ch18-EGFP fusion-protein, which has both TM3 and TM4 of ACTHR, while, on the contrary, the emission was detected in the membrane of cells expressing Ch16-EGFP fusion-protein which has only TM3 of ACTHR. In order to verify that the lack of membrane trafficking of Ch18 was not caused by unspecific chimeric receptor misfolding we also co-expressed both receptors with MRAP-β and as one can see this promoted the membrane trafficking of Ch18-EGFP, leaving unchanged the Ch16-EGFP emission level. The existence of such dual protection system might be necessary to avoid “receptor leakage” to the cell membrane, since alone none of the signals could ensure complete receptor retention.

Besides the membrane trafficking analysed in third chapter of results section we also tested the functional properties of created chimerical receptors. Acquired results indicated that region formed by TM4 and TM5 of ACTHR is responsible for receptor

ligand recognition selectivity. Combining this knowledge with the discoveries of MRAP anti-parallel dimerisation and its involvement in ligand recognition we developed a model of possible interactions within ACTHR/MRAP complex and proposed the possible scenario of events which is ensuring the selectivity of ligand recognition. Since it was discovered that MRAP forms the anti-parallel dimers prior coupling to ACTHR and that region located in the N-terminal tail of MRAP is participating in ligand recognition, we proposed that the MRAP monomer whose N-terminus is located on extracellular side of membrane is interacting with TM4-TM5 region of ACTHR, thus forming ACTH -K-K-R-R- pharmacophore binding pocket, while monomer whose C-terminus is located on extracellular side of membrane is interacting with TM3-TM4 region of ACTHR, thus masking or "neutralising in another way" the arrest signal which is formed by this region.

CONCLUSIONS

1. Six genes encoding melanocortin receptors were identified in zebrafish - *Danio rerio*. Sequence similarity analysis revealed that there is one ortholog for mammalian MC1R, ACTHR, MC3R and MC4R and two orthologs for mammalian MC5R.
2. Four genes encoding melanocortin receptors were identified in fugu - *Takifugu rubripes*. Sequence similarity analysis revealed that those are orthologs for mammalian MC1R, ACTHR, MC4R and MC5R. There were no orthologs of mammalian MC3R.
3. Expression of *Takifugu rubripes* ACTHR is limited to adrenal cells lines or cell lines expressing endogenous MCRs and, like mammalian orthologue, it binds solely to ACTH.
4. There are 14 ancestral receptor groups (ARGs) spared throughout the Rhodopsin GPCR family. These groups are usually small – formed by less than five receptors and most of the receptors in these groups contain at least one intron within protein coding sequence.
5. Adenosine binding receptor ARG, which is located in the same MECA cluster as MCRs shares higher sequence similarity with MCRs than other MECA families, thus it is highly plausible that MCRs family originated from receptor of Adenosine binding receptor family
6. Genes encoding the Rhodopsin GPCRs are predominantly intronless due to the rapid formation of new genes during the early stages of vertebrate evolution and since there are no introns in MCRs of any other species the introns of Fugu ACTHR and MC5R encoding genes must be of a recent origin.
7. Substitution of MC4R N-terminal part with homologue part from ACTHR significantly decreased the membrane trafficking of MC4R and signal localized in third and fourth transmembrane region is the main determinant of intracellular trapping of ACTHR.

MAIN THESIS FOR DEFENCE

1. Melanocortin receptors have acquired their respective functions in an early stage of craniate evolution, since there are no MC receptors in “lower” Chordates, but all five types of MC receptors are already found in bony fishes
2. The Majority of vertebrate GPCRs are intronless due to rapid formation of new genes during the course of vertebrate evolution.
3. N-terminus and TM3-TM5 are regions of ACTHR that are determining receptor intracellular retention and ligand recognition specificity.
4. Developed quantification of cell-membrane co-localization is effective membrane export level comparison method.

ACKNOWLEDGEMENTS

The research was supported by FEBS Collaborative Experimental Scholarship for Central and Eastern Europe and Novo Nordisk Foundation while working in Department of Neuroscience, Uppsala University and by The European Social Fund while working in Biomedical Research and Study Centre, University of Latvia.

Special thanks to Dace Pjanova and Ruta Bruvere for help with confocal microscopy experiments, Ramona Petrovska for help with cultivation of mammalian cell cultures, Dr. Helgi B. Schiöth and his research group for kind reception while working in their laboratory, Jānis Kloviņš for help and patience while supervising my research activities, Liene Ņikitina-Zaķe for consultations regarding medical terminology, closest colleagues for random and not so random aid in my research and my family for support and understanding.

REFERENCES

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, D. J. (2002). Signaling through G-Protein-Linked Cell-Surface Receptors. in *Molecular biology of the cell***:** 852-870.

Allolio, B. and Reincke, M. (1997). "Adrenocorticotropin receptor and adrenal disorders." *Horm Res* **47**(4-6): 273-278.

Argiolas, A., Melis, M. R., Murgia, S. and Schioth, H. B. (2000). "ACTH- and alpha-MSH-induced grooming, stretching, yawning and penile erection in male rats: site of action in the brain and role of melanocortin receptors." *Brain Res Bull* **51**(5): 425-431.

Attwood, T. K. and Findlay, J. B. (1994). "Fingerprinting G-protein-coupled receptors." *Protein Eng* **7**(2): 195-203.

Bae, H., Anderson, K., Flood, L. A., Skiba, N. P., Hamm, H. E. and Graber, S. G. (1997). "Molecular determinants of selectivity in 5-hydroxytryptamine1B receptor-G protein interactions." *J Biol Chem* **272**(51): 32071-32077.

Bae, H., Cabrera-Vera, T. M., Depree, K. M., Graber, S. G. and Hamm, H. E. (1999). "Two amino acids within the alpha4 helix of Galphai1 mediate coupling with 5-hydroxytryptamine1B receptors." *J Biol Chem* **274**(21): 14963-14971.

Baldwin, J. M. (1993). "The probable arrangement of the helices in G protein-coupled receptors." *EMBO J* **12**(4): 1693-1703.

Beguin, P. (1999). "Hybrid enzymes." *Curr Opin Biotechnol* **10**(4): 336-340.

Beltramo, M., Campanella, M., Tarozzo, G., Fredduzzi, S., Corradini, L., Forlani, A., Bertorelli, R. and Reggiani, A. (2003). "Gene expression profiling of melanocortin system in neuropathic rats supports a role in nociception." *Brain Res Mol Brain Res* **118**(1-2): 111-118.

Berget, S. M., Moore, C. and Sharp, P. A. (1977). "Spliced segments at the 5' terminus of adenovirus 2 late mRNA." *Proc Natl Acad Sci U S A* **74**(8): 3171-3175.

Bernstein, H. J. (2000). "Recent changes to RasMol, recombining the variants." *Trends Biochem Sci* **25**(9): 453-455.

Beuschlein, F., Fassnacht, M., Klink, A., Allolio, B. and Reincke, M. (2001). "ACTH-receptor expression, regulation and role in adrenocortial tumor formation." *Eur J Endocrinol* **144**(3): 199-206.

Bockaert, J. and Pin, J. P. (1999). "Molecular tinkering of G protein-coupled receptors: an evolutionary success." *EMBO J* **18**(7): 1723-1729.

Bohm, M., Wolff, I., Scholzen, T. E., Robinson, S. J., Healy, E., Luger, T. A., Schwarz, T. and Schwarz, A. (2005). "alpha-Melanocyte-stimulating hormone protects from ultraviolet radiation-induced apoptosis and DNA damage." *J Biol Chem* **280**(7): 5795-5802.

Bond, R. A. and Ijzerman, A. P. (2006). "Recent developments in constitutive receptor activity and inverse agonism, and their potential for GPCR drug discovery." *Trends Pharmacol Sci* **27**(2): 92-96.

Boston, B. A. (1999). "The role of melanocortins in adipocyte function." *Ann N Y Acad Sci* **885**: 75-84.

Boston, B. A. and Cone, R. D. (1996). "Characterization of melanocortin receptor subtype expression in murine adipose tissues and in the 3T3-L1 cell line." *Endocrinology* **137**(5): 2043-2050.

Brady, L. S., Smith, M. A., Gold, P. W. and Herkenham, M. (1990). "Altered expression of hypothalamic neuropeptide mRNAs in food-restricted and food-deprived rats." *Neuroendocrinology* **52**(5): 441-447.

Brosius, J. (1999). "Many G-protein-coupled receptors are encoded by retrogenes." *Trends Genet* **15**(8): 304-305.
Brosius, J. (2003). "Gene duplication and other evolutionary strategies: from the RNA world to the future." *J Struct Funct Genomics* **3**(1-4): 1-17.
Bryson-Richardson, R. J., Logan, D. W., Currie, P. D. and Jackson, I. J. (2004). "Large-scale analysis of gene structure in rhodopsin-like GPCRs: evidence for widespread loss of an ancient intron." *Gene* **338**(1): 15-23.
Butler, A. A. (2006). "The melanocortin system and energy balance." *Peptides* **27**(2): 281-290.
Cavalier-Smith, T. (1985). "Selfish DNA and the origin of introns." *Nature* **315**(6017): 283-284.
Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. and Prasher, D. C. (1994). "Green fluorescent protein as a marker for gene expression." *Science* **263**(5148): 802-805.
Chan, L. F., Webb, T. R., Chung, T. T., Meimaridou, E., Cooray, S. N., Guasti, L., Chapple, J. P., Egertova, M., Elphick, M. R., Cheetham, M. E., Metherell, L. A. and Clark, A. J. (2009). "MRAP and MRAP2 are bidirectional regulators of the melanocortin receptor family." *Proc Natl Acad Sci U S A* **106**(15): 6146-6151.
Chandrasegaran, S. and Smith, J. (1999). "Chimeric restriction enzymes: what is next?" *Biol Chem* **380**(7-8): 841-848.
Chen, W., Kelly, M. A., Opitz-Araya, X., Thomas, R. E., Low, M. J. and Cone, R. D. (1997). "Exocrine gland dysfunction in MC5-R-deficient mice: evidence for coordinated regulation of exocrine gland function by melanocortin peptides." *Cell* **91**(6): 789-798.
Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G., Thian, F. S., Kobilka, T. S., Choi, H. J., Kuhn, P., Weis, W. I., Kobilka, B. K. and Stevens, R. C. (2007). "High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor." *Science* **318**(5854): 1258-1265.
Chhajlani, V. and Wikberg, J. E. (1992). "Molecular cloning and expression of the human melanocyte stimulating hormone receptor cDNA." *FEBS Lett* **309**(3): 417-420.
Chida, D., Nakagawa, S., Nagai, S., Sagara, H., Katsumata, H., Imaki, T., Suzuki, H., Mitani, F., Ogishima, T., Shimizu, C., Kotaki, H., Kakuta, S., Sudo, K., Koike, T., Kubo, M. and Iwakura, Y. (2007). "Melanocortin 2 receptor is required for adrenal gland development, steroidogenesis, and neonatal gluconeogenesis." *Proc Natl Acad Sci U S A* **104**(46): 18205-18210.
Cho, G. and Doolittle, R. F. (1997). "Intron distribution in ancient paralogs supports random insertion and not random loss." *J Mol Evol* **44**(6): 573-584.
Chow, L. T., Gelinas, R. E., Broker, T. R. and Roberts, R. J. (1977). "An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA." *Cell* **12**(1): 1-8.
Clapham, D. E. and Neer, E. J. (1997). "G protein beta gamma subunits." *Annu Rev Pharmacol Toxicol* **37**: 167-203.
Clark, A. J. and Cammas, F. M. (1996). "The ACTH receptor." *Baillieres Clin Endocrinol Metab* **10**(1): 29-47.
Cone, R. D. (2006). "Studies on the physiological functions of the melanocortin system." *Endocr Rev* **27**(7): 736-749.
Conklin, B. R., Farfel, Z., Lustig, K. D., Julius, D. and Bourne, H. R. (1993). "Substitution of three amino acids switches receptor specificity of Gq alpha to that of Gi alpha." *Nature* **363**(6426): 274-276.
Cooray, S. N., Almiro Do Vale, I., Leung, K. Y., Webb, T. R., Chapple, J. P., Egertova, M., Cheetham, M. E., Elphick, M. R. and Clark, A. J. (2008). "The melanocortin 2

receptor accessory protein exists as a homodimer and is essential for the function of the melanocortin 2 receptor in the mouse y1 cell line." *Endocrinology* **149**(4): 1935-1941.
Cormack, B. P., Valdivia, R. H. and Falkow, S. (1996). "FACS-optimized mutants of the green fluorescent protein (GFP)." *Gene* **173**(1 Spec No): 33-38.
Costanzi, S., Siegel, J., Tikhonova, I. G. and Jacobson, K. A. (2009). "Rhodopsin and the others: a historical perspective on structural studies of G protein-coupled receptors." *Curr Pharm Des* **15**(35): 3994-4002.
Crocker, E., Eilers, M., Ahuja, S., Hornak, V., Hirshfeld, A., Sheves, M. and Smith, S. O. (2006). "Location of Trp265 in metarhodopsin II: implications for the activation mechanism of the visual receptor rhodopsin." *J Mol Biol* **357**(1): 163-172.
Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A. and Tsien, R. Y. (1995). "Understanding, improving and using green fluorescent proteins." *Trends Biochem Sci* **20**(11): 448-455.
Danielson, P. B. and Dores, R. M. (1999). "Molecular evolution of the opioid/orphanin gene family." *Gen Comp Endocrinol* **113**(2): 169-186.
Darnell, J. E., Jr. (1978). "Implications of RNA-RNA splicing in evolution of eukaryotic cells." *Science* **202**(4374): 1257-1260.
de Souza, S. J., Long, M., Klein, R. J., Roy, S., Lin, S. and Gilbert, W. (1998). "Toward a resolution of the introns early/late debate: only phase zero introns are correlated with the structure of ancient proteins." *Proc Natl Acad Sci U S A* **95**(9): 5094-5099.
de Souza, S. J., Long, M., Schoenbach, L., Roy, S. W. and Gilbert, W. (1996). "Intron positions correlate with module boundaries in ancient proteins." *Proc Natl Acad Sci U S A* **93**(25): 14632-14636.
de Souza, S. J., Long, M., Schoenbach, L., Roy, S. W. and Gilbert, W. (1997). "The correlation between introns and the three-dimensional structure of proteins." *Gene* **205**(1-2): 141-144.
De Wied, D. and Jolles, J. (1982). "Neuropeptides derived from pro-opiocortin: behavioral, physiological, and neurochemical effects." *Physiol Rev* **62**(3): 976-1059.
Dibb, N. J. (1993). "Why do genes have introns?" *FEBS Lett* **325**(1-2): 135-139.
Dibb, N. J. and Newman, A. J. (1989). "Evidence that introns arose at proto-splice sites." *Embo J* **8**(7): 2015-2021.
Dixon, R. A., Kobilka, B. K., Strader, D. J., Benovic, J. L., Dohlman, H. G., Frielle, T., Bolanowski, M. A., Bennett, C. D., Rands, E., Diehl, R. E., Mumford, R. A., Slater, E. E., Sigal, I. S., Caron, M. G., Lefkowitz, R. J. and Strader, C. D. (1986). "Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin." *Nature* **321**(6065): 75-79.
Doghman, M., Soltani, Y., Rebuffet, V., Naville, D. and Begeot, M. (2007). "Role of Agouti-related protein in adrenal steroidogenesis." *Mol Cell Endocrinol* **265-266**: 108-112.
Doolittle, W. F. (1978). "Genes in pieces: Were they ever together?" *Nature* **272**(5654): 581-582.
Downes, G. B. and Gautam, N. (1999). "The G protein subunit gene families." *Genomics* **62**(3): 544-552.
Eberle, A. N. (1988). The Melanotropins. Chemistry, Physiology and Mechanisms of Action. Basel, S. Karger Publishers.
Elias, L. L., Huebner, A., Pullinger, G. D., Mirtella, A. and Clark, A. J. (1999). "Functional characterization of naturally occurring mutations of the human adrenocorticotropin receptor: poor correlation of phenotype and genotype." *J Clin Endocrinol Metab* **84**(8): 2766-2770.

Ellacott, K. L. and Cone, R. D. (2004). "The central melanocortin system and the integration of short- and long-term regulators of energy homeostasis." *Recent Prog Horm Res* **59**: 395-408.
Ellacott, K. L., Murphy, J. G., Marks, D. L. and Cone, R. D. (2007). "Obesity-induced inflammation in white adipose tissue is attenuated by loss of melanocortin-3 receptor signaling." *Endocrinology* **148**(12): 6186-6194.
Fan, W., Boston, B. A., Kesterson, R. A., Hruby, V. J. and Cone, R. D. (1997). "Role of melanocortinergic neurons in feeding and the agouti obesity syndrome." *Nature* **385**(6612): 165-168.
Fan, W., Morrison, S. F., Cao, W. H. and Yu, P. (2007). "Thermogenesis activated by central melanocortin signaling is dependent on neurons in the rostral raphe pallidus (rRPa) area." *Brain Res* **1179**: 61-69.
Fan, W., Voss-Andreae, A., Cao, W. H. and Morrison, S. F. (2005). "Regulation of thermogenesis by the central melanocortin system." *Peptides* **26**(10): 1800-1813.
Farooqi, I. S., Drop, S., Clements, A., Keogh, J. M., Biernacka, J., Lowenbein, S., Challis, B. G. and O'Rahilly, S. (2006). "Heterozygosity for a POMC-null mutation and increased obesity risk in humans." *Diabetes* **55**(9): 2549-2553.
Fedorov, A., Cao, X., Saxonov, S., de Souza, S. J., Roy, S. W. and Gilbert, W. (2001). "Intron distribution difference for 276 ancient and 131 modern genes suggests the existence of ancient introns." *Proc Natl Acad Sci U S A* **98**(23): 13177-13182.
Fedorov, A. and Fedorova, L. (2004). "Introns: mighty elements from the RNA world." *J Mol Evol* **59**(5): 718-721.
Fedorov, A., Roy, S., Fedorova, L. and Gilbert, W. (2003). "Mystery of intron gain." *Genome Res* **13**(10): 2236-2241.
Fedorova, L. and Fedorov, A. (2003). "Introns in gene evolution." *Genetica* **118**(2-3): 123-131.
Feng, J. D., Dao, T. and Lipton, J. M. (1987). "Effects of preoptic microinjections of alpha-MSH on fever and normal temperature control in rabbits." *Brain Res Bull* **18**(4): 473-477.
Ferguson, S. S. (2001). "Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling." *Pharmacol Rev* **53**(1): 1-24.
Figueroa, F., Ono, H., Tichy, H., O'Huigin, C. and Klein, J. (1995). "Evidence for insertion of a new intron into an Mhc gene of perch-like fish." *Proc Biol Sci* **259**(1356): 325-330.
Fredriksson, R., Lagerstrom, M. C., Lundin, L. G. and Schioth, H. B. (2003). "The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints." *Mol Pharmacol* **63**(6): 1256-1272.
Fredriksson, R. and Schioth, H. B. (2005). "The repertoire of G-protein-coupled receptors in fully sequenced genomes." *Mol Pharmacol* **67**(5): 1414-1425.
Gales, C., Van Durm, J. J., Schaak, S., Pontier, S., Percherancier, Y., Audet, M., Paris, H. and Bouvier, M. (2006). "Probing the activation-promoted structural rearrangements in preassembled receptor-G protein complexes." *Nat Struct Mol Biol* **13**(9): 778-786.
Gantz, I. and Fong, T. M. (2003). "The melanocortin system." *Am J Physiol Endocrinol Metab* **284**(3): E468-474.
Gantz, I., Konda, Y., Tashiro, T., Shimoto, Y., Miwa, H., Munzert, G., Watson, S. J., DelValle, J. and Yamada, T. (1993a). "Molecular cloning of a novel melanocortin receptor." *J Biol Chem* **268**(11): 8246-8250.

Gantz, I., Miwa, H., Konda, Y., Shimoto, Y., Tashiro, T., Watson, S. J., DelValle, J. and Yamada, T. (1993b). "Molecular cloning, expression, and gene localization of a fourth melanocortin receptor." *J Biol Chem* **268**(20): 15174-15179.
Gantz, I., Shimoto, Y., Konda, Y., Miwa, H., Dickinson, C. J. and Yamada, T. (1994). "Molecular cloning, expression, and characterization of a fifth melanocortin receptor." *Biochem Biophys Res Commun* **200**(3): 1214-1220.
Garrud, P., Gray, J. A. and de Wied, D. (1974). "Pituitary-adrenal hormones and extinction of rewarded behaviour in the rat." *Physiol Behav* **12**(1): 109-119.
Gentles, A. J. and Karlin, S. (1999). "Why are human G-protein-coupled receptors predominantly intronless?" *Trends Genet* **15**(2): 47-49.
Gilbert, W. (1978). "Why genes in pieces?" *Nature* **271**(5645): 501.
Gilbert, W., de Souza, S. J. and Long, M. (1997). "Origin of genes." *Proc Natl Acad Sci U S A* **94**(15): 7698-7703.
Giroux, M. J., Clancy, M., Baier, J., Ingham, L., McCarty, D. and Hannah, L. C. (1994). "De novo synthesis of an intron by the maize transposable element Dissociation." *Proc Natl Acad Sci U S A* **91**(25): 12150-12154.
Haitina, T., Klovins, J., Takahashi, A., Lowgren, M., Ringholm, A., Enberg, J., Kawauchi, H., Larson, E. T., Fredriksson, R. and Schioth, H. B. (2007). "Functional characterization of two melanocortin (MC) receptors in lamprey showing orthology to the MC1 and MC4 receptor subtypes." *BMC Evol Biol* **7**: 101.
Hankeln, T., Friedl, H., Ebersberger, I., Martin, J. and Schmidt, E. R. (1997). "A variable intron distribution in globin genes of Chironomus: evidence for recent intron gain." *Gene* **205**(1-2): 151-160.
Hanson, M. A., Cherezov, V., Griffith, M. T., Roth, C. B., Jaakola, V. P., Chien, E. Y., Velasquez, J., Kuhn, P. and Stevens, R. C. (2008). "A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor." *Structure* **16**(6): 897-905.
Hargrave, P. A., McDowell, J. H., Curtis, D. R., Wang, J. K., Juszczak, E., Fong, S. L., Rao, J. K. and Argos, P. (1983). "The structure of bovine rhodopsin." *Biophys Struct Mech* **9**(4): 235-244.
Harrold, J. A., Widdowson, P. S. and Williams, G. (2003). "beta-MSH: a functional ligand that regulated energy homeostasis via hypothalamic MC4-R?" *Peptides* **24**(3): 397-405.
Haycock, J. W., Wagner, M., Morandini, R., Ghanem, G., Rennie, I. G. and MacNeil, S. (1999). "alpha-MSH immunomodulation acts via rel/NF-kappa B in cutaneous and ocular melanocytes and in melanoma cells." *Ann N Y Acad Sci* **885**: 396-399.
Hein, P., Frank, M., Hoffmann, C., Lohse, M. J. and Bunemann, M. (2005). "Dynamics of receptor/G protein coupling in living cells." *EMBO J* **24**(23): 4106-4114.
Hinkle, P. M. and Sebag, J. A. (2009). "Structure and function of the melanocortin2 receptor accessory protein (MRAP)." *Mol Cell Endocrinol* **300**(1-2): 25-31.
Hnatowich, M. R., Queen, G., Stein, D. and LaBella, F. S. (1989). "ACTH receptors in nervous tissue. High affinity binding-sequestration of [125I]Phe2,Nle4]ACTH 1-24 in homogenates and slices from rat brain." *Can J Physiol Pharmacol* **67**(6): 568-576.
Hou, Y., Azpiazu, I., Smrcka, A. and Gautam, N. (2000). "Selective role of G protein gamma subunits in receptor interaction." *J Biol Chem* **275**(50): 38961-38964.
Humphreys, M. H. (2004). "Gamma-MSH, sodium metabolism, and salt-sensitive hypertension." *Am J Physiol Regul Integr Comp Physiol* **286**(3): R417-430.
Hurles, M. (2004). "Gene duplication: the genomic trade in spare parts." *PLoS Biol* **2**(7): E206.

Ilnytska, O. and Argyropoulos, G. (2008). "The role of the Agouti-Related Protein in energy balance regulation." *Cell Mol Life Sci* **65**(17): 2721-2731.
Jaakola, V. P., Griffith, M. T., Hanson, M. A., Cherezov, V., Chien, E. Y., Lane, J. R., Ijzerman, A. P. and Stevens, R. C. (2008). "The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist." *Science* **322**(5905): 1211-1217.
Jalink, K. and Moolenaar, W. H. (2010). "G protein-coupled receptors: the inside story." *Bioessays* **32**(1): 13-16.
Jeffreys, A. J. and Flavell, R. A. (1977). "The rabbit beta-globin gene contains a large large insert in the coding sequence." *Cell* **12**(4): 1097-1108.
Jun, D. J., Na, K. Y., Kim, W., Kwak, D., Kwon, E. J., Yoon, J. H., Yea, K., Lee, H., Kim, J., Suh, P. G., Ryu, S. H. and Kim, K. T. (2010). "Melanocortins induce interleukin 6 gene expression and secretion through melanocortin receptors 2 and 5 in 3T3-L1 adipocytes." *J Mol Endocrinol* **44**(4): 225-236.
Kask, A., Mutulis, F., Muceniece, R., Pahkla, R., Mutule, I., Wikberg, J. E., Rago, L. and Schioth, H. B. (1998a). "Discovery of a novel superpotent and selective melanocortin-4 receptor antagonist (HS024): evaluation in vitro and in vivo." *Endocrinology* **139**(12): 5006-5014.
Kask, A., Rago, L., Korrovits, P., Wikberg, J. E. and Schioth, H. B. (1998b). "Evidence that orexigenic effects of melanocortin 4 receptor antagonist HS014 are mediated by neuropeptide Y." *Biochem Biophys Res Commun* **248**(2): 245-249.
Kask, A., Rago, L., Mutulis, F., Pahkla, R., Wikberg, J. E. and Schioth, H. B. (1998c). "Selective antagonist for the melanocortin 4 receptor (HS014) increases food intake in free-feeding rats." *Biochem Biophys Res Commun* **245**(1): 90-93.
Kask, A., Rago, L., Wikberg, J. E. and Schioth, H. B. (1998d). "Evidence for involvement of the melanocortin MC4 receptor in the effects of leptin on food intake and body weight." *Eur J Pharmacol* **360**(1): 15-19.
Kim, C. S., Lee, S. H., Kim, R. Y., Kim, B. J., Li, S. Z., Lee, I. H., Lee, E. J., Lim, S. K., Bae, Y. S., Lee, W. and Baik, J. H. (2002). "Identification of domains directing specificity of coupling to G-proteins for the melanocortin MC3 and MC4 receptors." *J Biol Chem* **277**(35): 31310-31317.
Kim, E. M., Welch, C. C., Grace, M. K., Billington, C. J. and Levine, A. S. (1996). "Chronic food restriction and acute food deprivation decrease mRNA levels of opioid peptides in arcuate nucleus." *Am J Physiol* **270**(5 Pt 2): R1019-1024.
Kisselev, O. G., Kao, J., Ponder, J. W., Fann, Y. C., Gautam, N. and Marshall, G. R. (1998). "Light-activated rhodopsin induces structural binding motif in G protein alpha subunit." *Proc Natl Acad Sci U S A* **95**(8): 4270-4275.
Klemm, J. D., Schreiber, S. L. and Crabtree, G. R. (1998). "Dimerization as a regulatory mechanism in signal transduction." *Annu Rev Immunol* **16**: 569-592.
Koenig, B. W., Kontaxis, G., Mitchell, D. C., Louis, J. M., Litman, B. J. and Bax, A. (2002). "Structure and orientation of a G protein fragment in the receptor bound state from residual dipolar couplings." *J Mol Biol* **322**(2): 441-461.
Konda, Y., Gantz, I., DelValle, J., Shimoto, Y., Miwa, H. and Yamada, T. (1994). "Interaction of dual intracellular signaling pathways activated by the melanocortin-3 receptor." *J Biol Chem* **269**(18): 13162-13166.
Kovacs, J. J., Hara, M. R., Davenport, C. L., Kim, J. and Lefkowitz, R. J. (2009). "Arrestin development: emerging roles for beta-arrestins in developmental signaling pathways." *Dev Cell* **17**(4): 443-458.
Kovalitskaia, Y. A., Kolobov, A. A., Kampe-Nemm, E. A., Iurovskii, V. V., Sadovnikov, V. B., Lipkin, V. M. and Navolotskaia, E. V. (2008). "[Synthetic peptide KKRR

corresponding to the human ACTH fragment 15-18 is an antagonist of the ACTH receptor]." *Bioorg Khim* **34**(1): 29-35.
Krebs, A., Edwards, P. C., Villa, C., Li, J. and Schertler, G. F. (2003). "The three-dimensional structure of bovine rhodopsin determined by electron cryomicroscopy." *J Biol Chem* **278**(50): 50217-50225.
Krude, H., Biebermann, H., Luck, W., Horn, R., Brabant, G. and Gruters, A. (1998). "Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans." *Nat Genet* **19**(2): 155-157.
Langer, I., Vertongen, P., Perret, J., Waelbroeck, M. and Robberecht, P. (2002). "A small sequence in the third intracellular loop of the VPAC(1) receptor is responsible for its efficient coupling to the calcium effector." *Biochem Soc Trans* **30**(4): 447-450.
Lasaga, M., Debeljuk, L., Durand, D., Scimonelli, T. N. and Caruso, C. (2008). "Role of alpha-melanocyte stimulating hormone and melanocortin 4 receptor in brain inflammation." *Peptides* **29**(10): 1825-1835.
Lee, V. D., Stapleton, M. and Huang, B. (1991). "Genomic structure of Chlamydomonas caltractin. Evidence for intron insertion suggests a probable genealogy for the EF-hand superfamily of proteins." *J Mol Biol* **221**(1): 175-191.
Lespinet, O., Wolf, Y. I., Koonin, E. V. and Aravind, L. (2002). "The role of lineage-specific gene family expansion in the evolution of eukaryotes." *Genome Res* **12**(7): 1048-1059.
Logsdon, J. M., Jr., Stoltzfus, A. and Doolittle, W. F. (1998). "Molecular evolution: recent cases of spliceosomal intron gain?" *Curr Biol* **8**(16): R560-563.
Long, M., de Souza, S. J., Rosenberg, C. and Gilbert, W. (1998). "Relationship between "proto-splice sites" and intron phases: evidence from dicodon analysis." *Proc Natl Acad Sci U S A* **95**(1): 219-223.
Long, M. and Rosenberg, C. (2000). "Testing the "proto-splice sites" model of intron origin: evidence from analysis of intron phase correlations." *Mol Biol Evol* **17**(12): 1789-1796.
Lu, D., Willard, D., Patel, I. R., Kadwell, S., Overton, L., Kost, T., Luther, M., Chen, W., Woychik, R. P., Wilkison, W. O. and et al. (1994). "Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor." *Nature* **371**(6500): 799-802.
Lu, X. Y. (2001). "Role of central melanocortin signaling in eating disorders." *Psychopharmacol Bull* **35**(4): 45-65.
Mandrika, I., Petrovska, R. and Klovins, J. (2010). "Evidence for constitutive dimerization of niacin receptor subtypes." *Biochem Biophys Res Commun*.
McIntire, W. E., MacCleery, G. and Garrison, J. C. (2001). "The G protein beta subunit is a determinant in the coupling of Gs to the beta 1-adrenergic and A2a adenosine receptors." *J Biol Chem* **276**(19): 15801-15809.
McLaughlin, J. N., Shen, L., Holinstat, M., Brooks, J. D., Dibenedetto, E. and Hamm, H. E. (2005). "Functional selectivity of G protein signaling by agonist peptides and thrombin for the protease-activated receptor-1." *J Biol Chem* **280**(26): 25048-25059.
Mena, M. A., Treynor, T. P., Mayo, S. L. and Daugherty, P. S. (2006). "Blue fluorescent proteins with enhanced brightness and photostability from a structurally targeted library." *Nat Biotechnol* **24**(12): 1569-1571.
Metherell, L. A., Chapple, J. P., Cooray, S., David, A., Becker, C., Ruschendorf, F., Naville, D., Begeot, M., Khoo, B., Nurnberg, P., Huebner, A., Cheetham, M. E. and Clark, A. J. (2005). "Mutations in MRAP, encoding a new interacting partner of the ACTH receptor, cause familial glucocorticoid deficiency type 2." *Nat Genet* **37**(2): 166-170.

Millar, R. P. and Newton, C. L. (2010). "The year in G protein-coupled receptor research." *Mol Endocrinol* **24**(1): 261-274.
Milligan, G. and Kostenis, E. (2006). "Heterotrimeric G-proteins: a short history." *Br J Pharmacol* **147 Suppl 1**: S46-55.
Mountjoy, K. G., Mortrud, M. T., Low, M. J., Simerly, R. B. and Cone, R. D. (1994). "Localization of the melanocortin-4 receptor (MC4-R) in neuroendocrine and autonomic control circuits in the brain." *Mol Endocrinol* **8**(10): 1298-1308.
Mountjoy, K. G., Robbins, L. S., Mortrud, M. T. and Cone, R. D. (1992). "The cloning of a family of genes that encode the melanocortin receptors." *Science* **257**(5074): 1248-1251.
Muceniece, R., Zvejniece, L., Kirjanova, O., Liepinsh, E., Krigere, L., Vilskersts, R., Baumane, L., Gordjusina, V., Kalvinsh, I., Wikberg, J. E. and Dambrova, M. (2005). "Beta-MSH inhibits brain inflammation via MC(3)/(4) receptors and impaired NF-kappaB signaling." *J Neuroimmunol* **169**(1-2): 13-19.
Muceniece, R., Zvejniece, L., Vilskersts, R., Liepinsh, E., Baumane, L., Kalvinsh, I., Wikberg, J. E. and Dambrova, M. (2007). "Functional evaluation of THIQ, a melanocortin 4 receptor agonist, in models of food intake and inflammation." *Basic Clin Pharmacol Toxicol* **101**(6): 416-420.
Mukhopadhyay, S. and Howlett, A. C. (2005). "Chemically distinct ligands promote differential CB1 cannabinoid receptor-Gi protein interactions." *Mol Pharmacol* **67**(6): 2016-2024.
Murakami, M. and Kouyama, T. (2008). "Crystal structure of squid rhodopsin." *Nature* **453**(7193): 363-367.
Nathans, J. and Hogness, D. S. (1984). "Isolation and nucleotide sequence of the gene encoding human rhodopsin." *Proc Natl Acad Sci U S A* **81**(15): 4851-4855.
Nathans, J., Thomas, D. and Hogness, D. S. (1986). "Molecular genetics of human color vision: the genes encoding blue, green, and red pigments." *Science* **232**(4747): 193-202.
Naville, D., Penhoat, A., Durand, P. and Begeot, M. (1999). "Three steroidogenic factor-1 binding elements are required for constitutive and cAMP-regulated expression of the human adrenocorticotropin receptor gene." *Biochem Biophys Res Commun* **255**(1): 28-33.
Nixon, A. E., Ostermeier, M. and Benkovic, S. J. (1998). "Hybrid enzymes: manipulating enzyme design." *Trends Biotechnol* **16**(6): 258-264.
Noon, L. A., Franklin, J. M., King, P. J., Goulding, N. J., Hunyady, L. and Clark, A. J. (2002). "Failed export of the adrenocorticotrophin receptor from the endoplasmic reticulum in non-adrenal cells: evidence in support of a requirement for a specific adrenal accessory factor." *J Endocrinol* **174**(1): 17-25.
Norman, D., Isidori, A. M., Frajese, V., Caprio, M., Chew, S. L., Grossman, A. B., Clark, A. J., Michael Besser, G. and Fabbri, A. (2003). "ACTH and alpha-MSH inhibit leptin expression and secretion in 3T3-L1 adipocytes: model for a central-peripheral melanocortin-leptin pathway." *Mol Cell Endocrinol* **200**(1-2): 99-109.
O'Neill, R. J., Brennan, F. E., Delbridge, M. L., Crozier, R. H. and Graves, J. A. (1998). "De novo insertion of an intron into the mammalian sex determining gene, SRY." *Proc Natl Acad Sci U S A* **95**(4): 1653-1657.
Oldham, W. M. and Hamm, H. E. (2008). "Heterotrimeric G protein activation by G-protein-coupled receptors." *Nat Rev Mol Cell Biol* **9**(1): 60-71.
Onrust, R., Herzmark, P., Chi, P., Garcia, P. D., Lichtarge, O., Kingsley, C. and Bourne, H. R. (1997). "Receptor and betagamma binding sites in the alpha subunit of the retinal G protein transducin." *Science* **275**(5298): 381-384.

Oosterom, J., Nijenhuis, W. A., Schaaper, W. M., Slootstra, J., Meloen, R. H., Gispen, W. H., Burbach, J. P. and Adan, R. A. (1999). "Conformation of the core sequence in melanocortin peptides directs selectivity for the melanocortin MC3 and MC4 receptors." *J Biol Chem* **274**(24): 16853-16860.

Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y. and Remington, S. J. (1996). "Crystal structure of the Aequorea victoria green fluorescent protein." *Science* **273**(5280): 1392-1395.

Ovchinnikov, Y., Abdulaev, N., Feigina, M., Artamonov, I., Zolotarev, A., Kostina, M., Bogachuk, A., Miroshnikov, A., Martinov, V. and Koudelin, A. (1982). "The Complete Amino-Acid-Sequence of Visual Rhodopsin." *Bioorganicheskaya Khimiya* **8**: 1011-1014.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M. and Miyano, M. (2000). "Crystal structure of rhodopsin: A G protein-coupled receptor." *Science* **289**(5480): 739-745.

Panskepp, J., Reilly, P., Bishop, P., Meeker, R. B., Vilberg, T. R. and Kastin, A. J. (1976). "Effects of alpha-MSH on motivation, vigilance and brain respiration." *Pharmacol Biochem Behav* **5**(Suppl 1): 59-64.

Pardo, L., Deupi, X., Dolker, N., Lopez-Rodriguez, M. L. and Campillo, M. (2007). "The role of internal water molecules in the structure and function of the rhodopsin family of G protein-coupled receptors." *Chembiochem* **8**(1): 19-24.

Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H. W. and Ernst, O. P. (2008). "Crystal structure of the ligand-free G-protein-coupled receptor opsin." *Nature* **454**(7201): 183-187.

Penhoat, A., Jaillard, C. and Saez, J. M. (1989). "Corticotropin positively regulates its own receptors and cAMP response in cultured bovine adrenal cells." *Proc Natl Acad Sci U S A* **86**(13): 4978-4981.

Penhoat, A., Naville, D., El Mourabit, H., Buronfosse, A., Berberoglu, M., Ocal, G., Tsigos, C., Durand, P. and Begeot, M. (2002). "Functional relationships between three novel homozygous mutations in the ACTH receptor gene and familial glucocorticoid deficiency." *J Mol Med* **80**(7): 406-411.

Perez, D. M. and Karnik, S. S. (2005). "Multiple signaling states of G-protein-coupled receptors." *Pharmacol Rev* **57**(2): 147-161.

Pfaus, J. G., Shadiack, A., Van Soest, T., Tse, M. and Molinoff, P. (2004). "Selective facilitation of sexual solicitation in the female rat by a melanocortin receptor agonist." *Proc Natl Acad Sci U S A* **101**(27): 10201-10204.

Pierce, K. L., Premont, R. T. and Lefkowitz, R. J. (2002). "Seven-transmembrane receptors." *Nat Rev Mol Cell Biol* **3**(9): 639-650.

Poggioli, R., Vergoni, A. V. and Bertolini, A. (1986). "ACTH-(1-24) and alpha-MSH antagonize feeding behavior stimulated by kappa opiate agonists." *Peptides* **7**(5): 843-848.

Probst, W. C., Snyder, L. A., Schuster, D. I., Brosius, J. and Sealfon, S. C. (1992). "Sequence alignment of the G-protein coupled receptor superfamily." *DNA Cell Biol* **11**(1): 1-20.

Qian, B., Soyer, O. S., Neubig, R. R. and Goldstein, R. A. (2003). "Depicting a protein's two faces: GPCR classification by phylogenetic tree-based HMMs." *FEBS Lett* **554**(1-2): 95-99.

Rao, V. R. and Oprian, D. D. (1996). "Activating mutations of rhodopsin and other G protein-coupled receptors." *Annu Rev Biophys Biomol Struct* **25**: 287-314.

Rasmussen, S. G., Choi, H. J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R., Sanishvili, R., Fischetti, R. F., Schertler, G. F., Weis, W. I. and Kobilka, B. K. (2007). "Crystal structure of the human beta2 adrenergic G-protein-coupled receptor." *Nature* **450**(7168): 383-387.

Rinne, P., Harjunpaa, J., Scheinin, M. and Savontaus, E. (2008). "Blood pressure regulation and cardiac autonomic control in mice overexpressing alpha- and gamma-melanocyte stimulating hormone." *Peptides* **29**(11): 1943-1952.

Rios, C. D., Jordan, B. A., Gomes, I. and Devi, L. A. (2001). "G-protein-coupled receptor dimerization: modulation of receptor function." *Pharmacol Ther* **92**(2-3): 71-87.

Roselli-Rehfuss, L., Mountjoy, K. G., Robbins, L. S., Mortrud, M. T., Low, M. J., Tatro, J. B., Entwistle, M. L., Simerly, R. B. and Cone, R. D. (1993). "Identification of a receptor for gamma melanotropin and other proopiomelanocortin peptides in the hypothalamus and limbic system." *Proc Natl Acad Sci U S A* **90**(19): 8856-8860.

Rosenbaum, D. M., Rasmussen, S. G. and Kobilka, B. K. (2009). "The structure and function of G-protein-coupled receptors." *Nature* **459**(7245): 356-363.

Roy, S., Rached, M. and Gallo-Payet, N. (2007). "Differential regulation of the human adrenocorticotropin receptor [melanocortin-2 receptor (MC2R)] by human MC2R accessory protein isoforms alpha and beta in isogenic human embryonic kidney 293 cells." *Mol Endocrinol* **21**(7): 1656-1669.

Roy, S. W., Fedorov, A. and Gilbert, W. (2002). "The signal of ancient introns is obscured by intron density and homolog number." *Proc Natl Acad Sci U S A* **99**(24): 15513-15517.

Roy, S. W., Fedorov, A. and Gilbert, W. (2003). "Large-scale comparison of intron positions in mammalian genes shows intron loss but no gain." *Proc Natl Acad Sci U S A* **100**(12): 7158-7162.

Roy, S. W. and Gilbert, W. (2005a). "Complex early genes." *Proc Natl Acad Sci U S A* **102**(6): 1986-1991.

Roy, S. W. and Gilbert, W. (2005b). "The pattern of intron loss." *Proc Natl Acad Sci U S A* **102**(3): 713-718.

Roy, S. W. and Gilbert, W. (2005c). "Rates of intron loss and gain: implications for early eukaryotic evolution." *Proc Natl Acad Sci U S A* **102**(16): 5773-5778.

Roy, S. W., Nosaka, M., de Souza, S. J. and Gilbert, W. (1999). "Centripetal modules and ancient introns." *Gene* **238**(1): 85-91.

Ruprecht, J. J., Mielke, T., Vogel, R., Villa, C. and Schertler, G. F. (2004). "Electron crystallography reveals the structure of metarhodopsin I." *EMBO J* **23**(18): 3609-3620.

Sadusky, T., Newman, A. J. and Dibb, N. J. (2004). "Exon junction sequences as cryptic splice sites: implications for intron origin." *Curr Biol* **14**(6): 505-509.

Sakharkar, M. K., Chow, V. T., Chaturvedi, I., Mathura, V. S., Shapshak, P. and Kangueane, P. (2004). "A report on single exon genes (SEG) in eukaryotes." *Front Biosci* **9**: 3262-3267.

Salomon, Y., Zohar, M., Dejordy, J. O., Eshel, Y., Shafir, I., Leiba, H., Garty, N. B., Schmidt-Sole, J., Azrad, A., Shai, E. and et al. (1993). "Signaling mechanisms controlled by melanocortins in melanoma, lacrimal, and brain astroglial cells." *Ann N Y Acad Sci* **680**: 364-380.

Sandman, C. A., Kastin, A. J. and Schally, A. V. (1969). "Melanocyte-stimulating hormone and learned appetitive behavior." *Experientia* **25**(9): 1001-1002.

Scheerer, P., Park, J. H., Hildebrand, P. W., Kim, Y. J., Krauss, N., Choe, H. W., Hofmann, K. P. and Ernst, O. P. (2008). "Crystal structure of opsin in its G-protein-interacting conformation." *Nature* **455**(7212): 497-502.

Schertler, G. F., Villa, C. and Henderson, R. (1993). "Projection structure of rhodopsin." *Nature* **362**(6422): 770-772.
Scheurer, S. B., Rybak, J. N., Roesli, C., Brunisholz, R. A., Potthast, F., Schlapbach, R., Neri, D. and Elia, G. (2005). "Identification and relative quantification of membrane proteins by surface biotinylation and two-dimensional peptide mapping." *Proteomics* **5**(11): 2718-2728.
Schimmer, B. P., Kwan, W. K., Tsao, J. and Qiu, R. (1995). "Adrenocorticotropin-resistant mutants of the Y1 adrenal cell line fail to express the adrenocorticotropin receptor." *J Cell Physiol* **163**(1): 164-171.
Schioth, H. B. and Fredriksson, R. (2005). "The GRAFS classification system of G-protein coupled receptors in comparative perspective." *Gen Comp Endocrinol* **142**(1-2): 94-101.
Schioth, H. B., Yook, P., Muceniece, R., Wikberg, J. E. and Szardenings, M. (1998). "Chimeric melanocortin MC1 and MC3 receptors: identification of domains participating in binding of melanocyte-stimulating hormone peptides." *Mol Pharmacol* **54**(1): 154-161.
Sebag, J. A. and Hinkle, P. M. (2007). "Melanocortin-2 receptor accessory protein MRAP forms antiparallel homodimers." *Proc Natl Acad Sci U S A* **104**(51): 20244-20249.
Sebag, J. A. and Hinkle, P. M. (2009). "Regions of melanocortin2 receptor accessory protein necessary for dual topology and MC2 receptor trafficking and signaling." *J Biol Chem*.
Shi, L., Liapakis, G., Xu, R., Guarnieri, F., Ballesteros, J. A. and Javitch, J. A. (2002). "Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch." *J Biol Chem* **277**(43): 40989-40996.
Shimamura, T., Hiraki, K., Takahashi, N., Hori, T., Ago, H., Masuda, K., Takio, K., Ishiguro, M. and Miyano, M. (2008). "Crystal structure of squid rhodopsin with intracellularly extended cytoplasmic region." *J Biol Chem* **283**(26): 17753-17756.
Shutter, J. R., Graham, M., Kinsey, A. C., Scully, S., Luthy, R. and Stark, K. L. (1997). "Hypothalamic expression of ART, a novel gene related to agouti, is up-regulated in obese and diabetic mutant mice." *Genes Dev* **11**(5): 593-602.
Simon, M. I., Strathmann, M. P. and Gautam, N. (1991). "Diversity of G proteins in signal transduction." *Science* **252**(5007): 802-808.
Simpson, E. R. and Waterman, M. R. (1988). "Regulation of the synthesis of steroidogenic enzymes in adrenal cortical cells by ACTH." *Annu Rev Physiol* **50**: 427-440.
Sina, M., Hinney, A., Ziegler, A., Neupert, T., Mayer, H., Siegfried, W., Blum, W. F., Remschmidt, H. and Hebebrand, J. (1999). "Phenotypes in three pedigrees with autosomal dominant obesity caused by haploinsufficiency mutations in the melanocortin-4 receptor gene." *Am J Hum Genet* **65**(6): 1501-1507.
Smith, A. I. and Funder, J. W. (1988). "Proopiomelanocortin processing in the pituitary, central nervous system, and peripheral tissues." *Endocr Rev* **9**(1): 159-179.
Spencer, J. D. and Schallreuter, K. U. (2008). "Regulation of pigmentation in human epidermal melanocytes by functional high affinity {beta}-MSH/MC4-R signalling." *Endocrinology*.
Standfuss, J., Xie, G., Edwards, P. C., Burghammer, M., Oprian, D. D. and Schertler, G. F. (2007). "Crystal structure of a thermally stable rhodopsin mutant." *J Mol Biol* **372**(5): 1179-1188.
Starowicz, K. and Przewlocka, B. (2003). "The role of melanocortins and their receptors in inflammatory processes, nerve regeneration and nociception." *Life Sci* **73**(7): 823-847.

Tarrio, R., Rodriguez-Trelles, F. and Ayala, F. J. (1998). "New Drosophila introns originate by duplication." *Proc Natl Acad Sci U S A* **95**(4): 1658-1662.
Tarrio, R., Rodriguez-Trelles, F. and Ayala, F. J. (2003). "A new Drosophila spliceosomal intron position is common in plants." *Proc Natl Acad Sci U S A* **100**(11): 6580-6583.
Tate, C. G. and Schertler, G. F. (2009). "Engineering G protein-coupled receptors to facilitate their structure determination." *Curr Opin Struct Biol* **19**(4): 386-395.
Tatro, J. B. (1990). "Melanotropin receptors in the brain are differentially distributed and recognize both corticotropin and alpha-melanocyte stimulating hormone." *Brain Res* **536**(1-2): 124-132.
Tatro, J. B. and Reichlin, S. (1987). "Specific receptors for alpha-melanocyte-stimulating hormone are widely distributed in tissues of rodents." *Endocrinology* **121**(5): 1900-1907.
Taylor, J. M., Jacob-Mosier, G. G., Lawton, R. G., Remmers, A. E. and Neubig, R. R. (1994). "Binding of an alpha 2 adrenergic receptor third intracellular loop peptide to G beta and the amino terminus of G alpha." *J Biol Chem* **269**(44): 27618-27624.
Taylor, J. M., Jacob-Mosier, G. G., Lawton, R. G., VanDort, M. and Neubig, R. R. (1996). "Receptor and membrane interaction sites on Gbeta. A receptor-derived peptide binds to the carboxyl terminus." *J Biol Chem* **271**(7): 3336-3339.
Theodoropoulou, M. C., Bagos, P. G., Spyropoulos, I. C. and Hamodrakas, S. J. (2008). "gpDB: a database of GPCRs, G-proteins, effectors and their interactions." *Bioinformatics* **24**(12): 1471-1472.
Thiboutot, D., Sivarajah, A., Gilliland, K., Cong, Z. and Clawson, G. (2000). "The melanocortin 5 receptor is expressed in human sebaceous glands and rat preputial cells." *J Invest Dermatol* **115**(4): 614-619.
Thiele, T. E., van Dijk, G., Yagaloff, K. A., Fisher, S. L., Schwartz, M., Burn, P. and Seeley, R. J. (1998). "Central infusion of melanocortin agonist MTII in rats: assessment of c-Fos expression and taste aversion." *Am J Physiol* **274**(1 Pt 2): R248-254.
Tolkovsky, A. M. and Levitzki, A. (1978). "Mode of coupling between the beta-adrenergic receptor and adenylate cyclase in turkey erythrocytes." *Biochemistry* **17**(18): 3795.
Tumova, K., Iwasiow, R. M. and Tiberi, M. (2003). "Insight into the mechanism of dopamine D1-like receptor activation. Evidence for a molecular interplay between the third extracellular loop and the cytoplasmic tail." *J Biol Chem* **278**(10): 8146-8153.
Tuteja, N. (2009). "Signaling through G protein coupled receptors." *Plant Signal Behav* **4**(10): 942-947.
Venkatesh, B., Ning, Y. and Brenner, S. (1999). "Late changes in spliceosomal introns define clades in vertebrate evolution." *Proc Natl Acad Sci U S A* **96**(18): 10267-10271.
Vitko, I., Bidaud, I., Arias, J. M., Mezghrani, A., Lory, P. and Perez-Reyes, E. (2007). "The I-II loop controls plasma membrane expression and gating of Ca(v)3.2 T-type Ca2+ channels: a paradigm for childhood absence epilepsy mutations." *J Neurosci* **27**(2): 322-330.
Wang, L., Martin, B., Brenneman, R., Luttrell, L. M. and Maudsley, S. (2009). "Allosteric modulators of g protein-coupled receptors: future therapeutics for complex physiological disorders." *J Pharmacol Exp Ther* **331**(2): 340-348.
Warne, T., Serrano-Vega, M. J., Baker, J. G., Moukhametzianov, R., Edwards, P. C., Henderson, R., Leslie, A. G., Tate, C. G. and Schertler, G. F. (2008). "Structure of a beta1-adrenergic G-protein-coupled receptor." *Nature* **454**(7203): 486-491.

Webb, T. R., Chan, L., Cooray, S. N., Cheetham, M. E., Chapple, J. P. and Clark, A. J. (2009). "Distinct MRAP domains are required for MC2R interaction and promotion of receptor trafficking." *Endocrinology*.

Webb, T. R. and Clark, A. J. (2010). "Minireview: the melanocortin 2 receptor accessory proteins." *Mol Endocrinol* **24**(3): 475-484.

Wessells, H., Gralnek, D., Dorr, R., Hruby, V. J., Hadley, M. E. and Levine, N. (2000). "Effect of an alpha-melanocyte stimulating hormone analog on penile erection and sexual desire in men with organic erectile dysfunction." *Urology* **56**(4): 641-646.

Williams, G., Bing, C., Cai, X. J., Harrold, J. A., King, P. J. and Liu, X. H. (2001). "The hypothalamus and the control of energy homeostasis: different circuits, different purposes." *Physiol Behav* **74**(4-5): 683-701.

Wilson, B. D., Ollmann, M. M., Kang, L., Stoffel, M., Bell, G. I. and Barsh, G. S. (1995). "Structure and function of ASP, the human homolog of the mouse agouti gene." *Hum Mol Genet* **4**(2): 223-230.

Wolfe, B. L. and Trejo, J. (2007). "Clathrin-dependent mechanisms of G protein-coupled receptor endocytosis." *Traffic* **8**(5): 462-470.

Xu, Z. Q., Zhang, X. and Scott, L. (2007). "Regulation of G protein-coupled receptor trafficking." *Acta Physiol (Oxf)* **190**(1): 39-45.

Yang, T. T., Sinai, P., Green, G., Kitts, P. A., Chen, Y. T., Lybarger, L., Chervenak, R., Patterson, G. H., Piston, D. W. and Kain, S. R. (1998). "Improved fluorescence and dual color detection with enhanced blue and green variants of the green fluorescent protein." *J Biol Chem* **273**(14): 8212-8216.

Yang, Y. K., Dickinson, C. J., Zeng, Q., Li, J. Y., Thompson, D. A. and Gantz, I. (1999). "Contribution of melanocortin receptor exoloops to Agouti-related protein binding." *J Biol Chem* **274**(20): 14100-14106.

Yeo, G. S., Farooqi, I. S., Challis, B. G., Jackson, R. S. and O'Rahilly, S. (2000). "The role of melanocortin signalling in the control of body weight: evidence from human and murine genetic models." *QJM* **93**(1): 7-14.

Zheng, H., Loh, H. H. and Law, P. Y. "Agonist-selective signaling of G protein-coupled receptor: mechanisms and implications." *IUBMB Life* **62**(2): 112-119.