Chapter 1

Introduction

1.1 General Introduction

More than 100 years have passed since Emil Fischer compared in 1890 the interaction between a ligand molecule and a protein binding site with a key that fits a lock. Since then, tremendous advances in X-ray crystallography and computer technology have been achieved and nowadays allow the complex drug-receptor interaction to be visualised at an atomic scale. The possibility to see and understand how drugs are accommodated within a protein binding site and how they might exert their effects opens a bright future for rational drug design.

In this work, the focus has been set on molecular modelling studies of the human histamine H_3-receptor. This receptor is involved in the regulation of important physiological processes and could be a potential target for the treatment of numerous diseases. The hH_3R is involved in cognition, food intake, sleep and pain perception, thus the list of diseases that might be influenced via this receptor is long: Alzheimer, schizophrenia, attention-deficit hyperactivity disorder (ADHD), obesity,. . .

The H_3R receptor belongs to the family of G-protein coupled receptors, on which a short introduction is given below.

1.2 G-Protein Coupled Receptors

1.2.1 Pharmaceutical Relevance and Classification

The human genome codes for approximately 30000 proteins, of which, however, only 10% are currently classified as druggable targets. The term *druggable target* thereby refers to

proteins of pharmaceutical interest that can be addressed by small drug-like molecules that bind to a defined cavity within the protein, thereby exerting their effect. Excluded from this definition are molecules targeting for example protein-protein interfaces, as they do not provide any well defined binding pocket. The pool of human druggable targets can be subdivided into target families of which kinases, proteases, ion-channels and GPCRs are the most important (see figure 1.1). Sequence analysis has led to the assumption that the human genome could code for approximately 1000 GPCRs, [1] a large number of which are however still orphan receptors (i.e. receptors for which no ligand is known and no function has been determined so far). Currently, more than 50% of all drugs on the market exert their function targeting a GPCR, which stresses the outstanding pharmaceutical importance of this protein family. [2] The large number of GPCR protein sequences has led to a complex classification scheme consisting of 6 families (A–F), each containing several subfamilies. Human GPCR sequences can only be found within families A–C, while families D–F contain sequences of yeast, amoeba and archea. Family A, which is also termed rhodopsin family, is the largest group, comprising receptors for most amine neurotransmitters, many neuropeptides, purines, prostaglandines, cannabinoids, . . . Ligands targeting family A GPCRs either bind to the transmembrane helices (amines) or within the extracellular loop region (peptides). Family B, the secretin/glucagon family, contains receptors for peptide hormones, including secretin, glucagon, and calcitonin. The ligand binding domain is located within the *N*-terminal end of intermediate length. Family C, the metabotropic glutamate receptor family or calcium sensor family, is the smallest group comprising metabotropic glutamate receptors, $GABA_B$ receptors and calcium-sensing receptors. The ligands bind within the long extracellular tail. [1] All histamine receptors can be classified as biogen aminergic GPCRs, one of the several subfamilies within family A GPCRs. Other subfamilies in family A are the muscarinic acetylcholine, adrenergic, dopamine, serotonine, octopamine and trace amine-receptors.

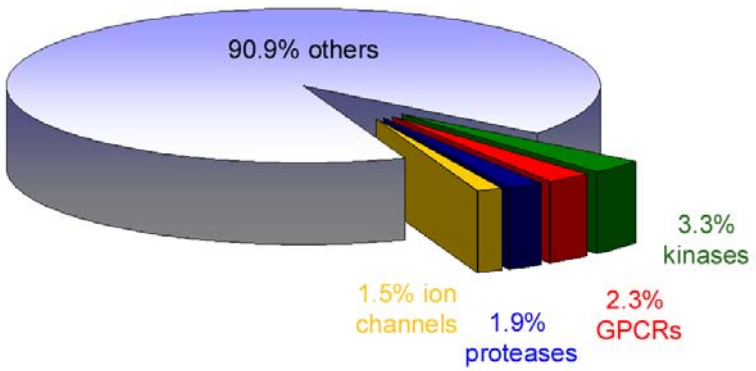

Figure 1.1: Important target families in the human genome.

1.2.2 Molecular Structure and Signal Transduction

GPCRs are heptahelical transmembrane proteins located in the cytoplasma membrane and have the common function of transducing extracellular signals into intracellular responses. Binding domains for the extracellular ligands thereby either lie within the helix bundle or within the *N*-terminal region of the receptor. A wide variety of ligands was reported to bind to GPCRs including biogen amines (*e.g. histamine, dopamine, acetylcholine*), amino acids (*e.g. glutamate*), ions (*e.g.* Ca^{2+}), lipids (*e.g. prostaglandines, leukotrienes*), peptides and proteins (*e.g. angiotensin, bradykinin, endorphins*). Agonist binding provokes a conformational change of the receptor that results in an alteration of the intracellular interaction with the heterotrimeric G-protein (see figure 1.2). [3] From coupling with the trimeric G-protein that assures a subsequent signal transduction in all GPCRs, the name *G-protein coupled receptor* originated. Following receptor activation, the heterotrimeric G-protein (composed of an α-, β- and γ-subunit) itself undergoes a conformational change that leads to the exchange of GDP for GTP bound to the α-subunit. The G$\alpha\beta\gamma$ complex breaks down into an α-subunit and the $\beta\gamma$-complex. The α-subunit can then stimulate effector molecules such as cyclases, phosphodiesterases and phospholipases, while for the $\beta\gamma$-complex in some cases a direct interaction with ion channels has been reported. As a result, the production of several second messengers such as cAMP (cyclic 3',5'-adenosine monophosphate), diacylglycerol or inositol(1,4,5)-trisphosphate is modulated. Finally, these second messengers can either prompt a fast cellular response (e.g. a modulation of intracellular ion concentration or regulation of enzyme activity) or cause a long-term biological effect by influencing transcription factors thereby regulating gene expression. The process is terminated when the catalytic conversion of GTP to GDP occurs in the α-subunit. αGDP dissociates from the effector and reunites with the $\beta\gamma$ complex.

A sufficient signal amplification is guaranteed, as each agonist/receptor complex can activate several G-proteins, which in turn can produce a large number of second messengers. The specificity of the biological response depends mainly on the nature of the α-subunit, of which more than 20 subtypes are known. [1]

Recently, in some cases, also G-protein independent signalling by GPCRs has been observed, thus further broadening the molecular mechanisms by which these receptors transduce extracellular signals. An example is the coupling of the G-protein coupled receptor kinase (GRK) to the phosphorylated *C*-terminal ending of GPCRs, thus labelling this receptor for endocytosis. [1, 3]

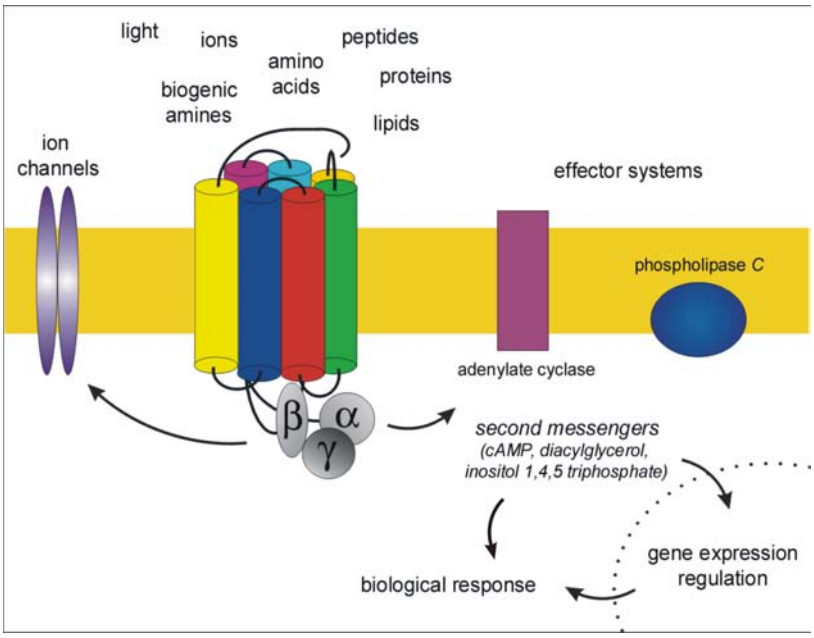

Figure 1.2: Diagram of GPCR function.

1.2.3 The Structure of Bovine Rhodopsin

A major breakthrough in the understanding of the GPCR receptor family was achieved in 2000 when the crystal structure of bovine rhodopsin was resolved [4] and for the first time detailed structural insights were gained (see figure 1.3). The crystallisation of other members of this protein family is hampered by usually low expression levels, which leads to problems in generating sufficient protein material for a crystallisation procedure as well as difficulties in the crystallisation process itself. [5]

Bovine rhodopsin represents a unique member among GPCRs as its intrinsic ligand *11-cis*-retinal is covalently bound to a lysine residue via a protonated Schiff-base linkage. Photon absorption results in retinal isomerisation to the *all-trans*-configuration triggering the activation of the protein. After photon capture, rhodopsin relaxes through a series of photoproducts until the active conformation, meta II, is formed which is capable of interacting with the G-protein. Hydrolysis of the Schiff-base results in the decay of meta II into the apoprotein opsin and *all-trans*-retinal. In vertebrates, *all-trans*-retinal is then regenerated in retinal pigment epithelial cells. [6]

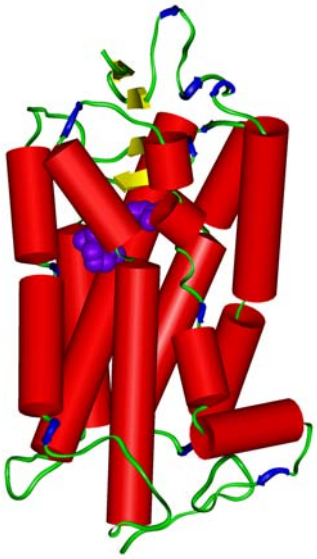

Figure 1.3: Crystal structure of bovine rhodopsin. Helical segments are depicted as red columns, beta-sheets as yellow arrows, turns as violet coil. The intrinsic ligand *11-cis*-retinal is displayed in its van der Waals representation.

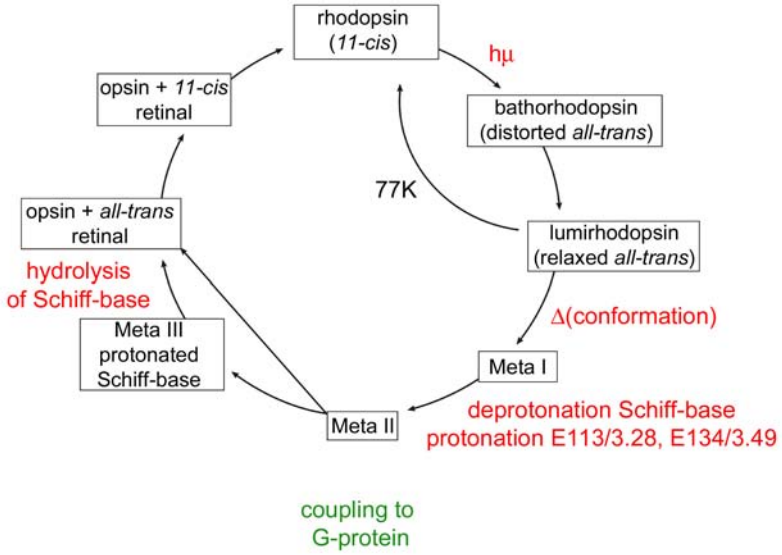

Figure 1.4: Intermediates in the photocycle: absorption of a photon results in the isomerisation of *11-cis*-retinal to a distorted *all-trans*-intermediate. The twisted double bond in bathorhodopsin relaxes to give lumirhodopsin. In lumirhodopsin the β-ionone moiety flips from its original position towards TM3 and TM4. Then, significant conformational changes take place and result in the meta I intermediate. Several de/protonation events produce meta II that represents the agonist-activated rhodopsin capable of interaction with the G-protein. Meta II finally decays to opsin and *all-trans*-retinal that is regenerated to *11-cis*-retinal outside the photoreceptor cells. The fact that lumirhodopsin can be reconverted to rhodopsin at 77 K where conformational changes are mainly impeded implies that no significant structural rearrangement takes place during the initial relaxation steps. [7]

1.2.4 Agonists, Antagonists and Inverse Agonists

Based on the biological response they provoke, ligands can be classified into agonists, partial agonists, antagonists, and inverse agonists.

Historically, a ligand that binds to a receptor and subsequently governs receptor activation has been referred to as an *agonist*; while a ligand that binds to a receptor without causing activation but impeding agonist binding, has been termed *antagonist*. Thus, although both molecules bind to the receptor (governed by their affinity), in the most simple cases only agonists will also possess an efficacy that results in receptor activation. *Partial agonists* are compounds that have a sub maximal tissue response even if they fully occupy the receptor.

The discovery of constitutive receptor activity has resulted in a reassessment of these terms. Constitutive activity describes the effect that receptors can exhibit an appreciable level of activity even in absence of any agonist. *Inverse agonists* are ligands that reduce this constitutive activation. *Neutral antagonists* restore the system towards the constitutive level of activity while agonists further activate the receptor. Figure 1.2 shows the classical two-state model that is nowadays however discussed controversially, as there is evidence for different coexisting receptor conformations.

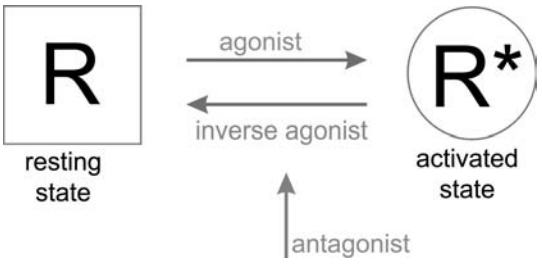

Figure 1.5: The two state model of receptor activity. The receptor exists in an equilibrium of activated and resting state. Antagonists preserve the level of constitutive activity, inverse agonists reduce activity and agonists enhance it.

1.3 Histamine

Histamine (2-[4-imidazolyl]ethyl amine) is a basic amine formed from L-histidine by histidine decarboxylase or an ubiquitous L-amino acid decarboxylase. High concentrations of histamine are found in the lungs, the skin and the gastrointestinal tract, where histamine functions as a mediator. In these tissues, histamine is predominantly stored in mast cells. During inflammatory or allergic reactions, histamine is liberated from mast cells by exocytosis. Alternatively, histamine is liberated upon destruction of these cells or by chemical substances (histamine liberators). When functioning as a mediator, histamine

exerts its effect predominantly via the H_1R, H_2R and the H_4R. Histamine additionally acts as a neurotransmitter in the central nervous system (CNS), where it interacts with the postsynaptically expressed H_1R and H_2R and praesynaptically expressed H_3R.

Two tautomeric forms of histamine exist, termed NπH and NτH (see figure 1.6). The ratio between the NτH and the NπH in an aqueous environment amounts to approximately 4:1. The pK_a of the α-amino group is 9.73 while the pK_a of the imidazole moiety is 5.91. [8] Histamine can thus exist as a dication whereby the protonated imidazole moiety is stabilised through delocalisation of the positive charge.

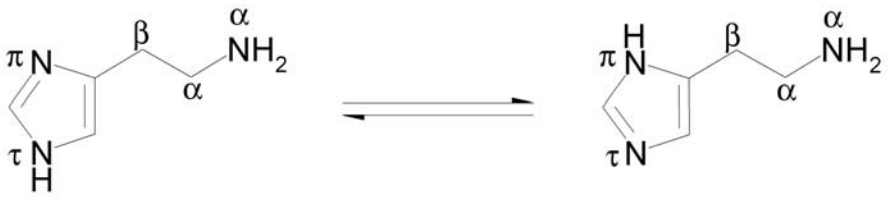

Figure 1.6: Tautomeric forms of histamine.

Released histamine is rapidly inactivated by histamine *N*-methyltransferase (HNMT, EC 2.1.1.8) or diamine oxidase (DAO, EC 1.4.3.6). The HNMT plays the dominant role in histamine metabolism within the human airways and gut, and is the only enzyme responsible for termination of the neurotransmitter action. [9] The HNMT inactivates histamine by transferring a methyl group from S-adenosyl-L-methionine to the Nτ-atom of the imidazole ring, yielding methylhistamine and S-adenosyl-L-homocysteine. The inactive methylhistamine is excreted in the urine or can be further metabolised by DAO or MAO into Nτ-methyl-imidazole-acetaldehyde, which can in turn be further oxidised by aldehyde dehydrogenase into Nτ-acetylimidazole acetic acid (see figure 1.7). The pathway of histamine metabolism starting with DAO is only relevant in the periphery.

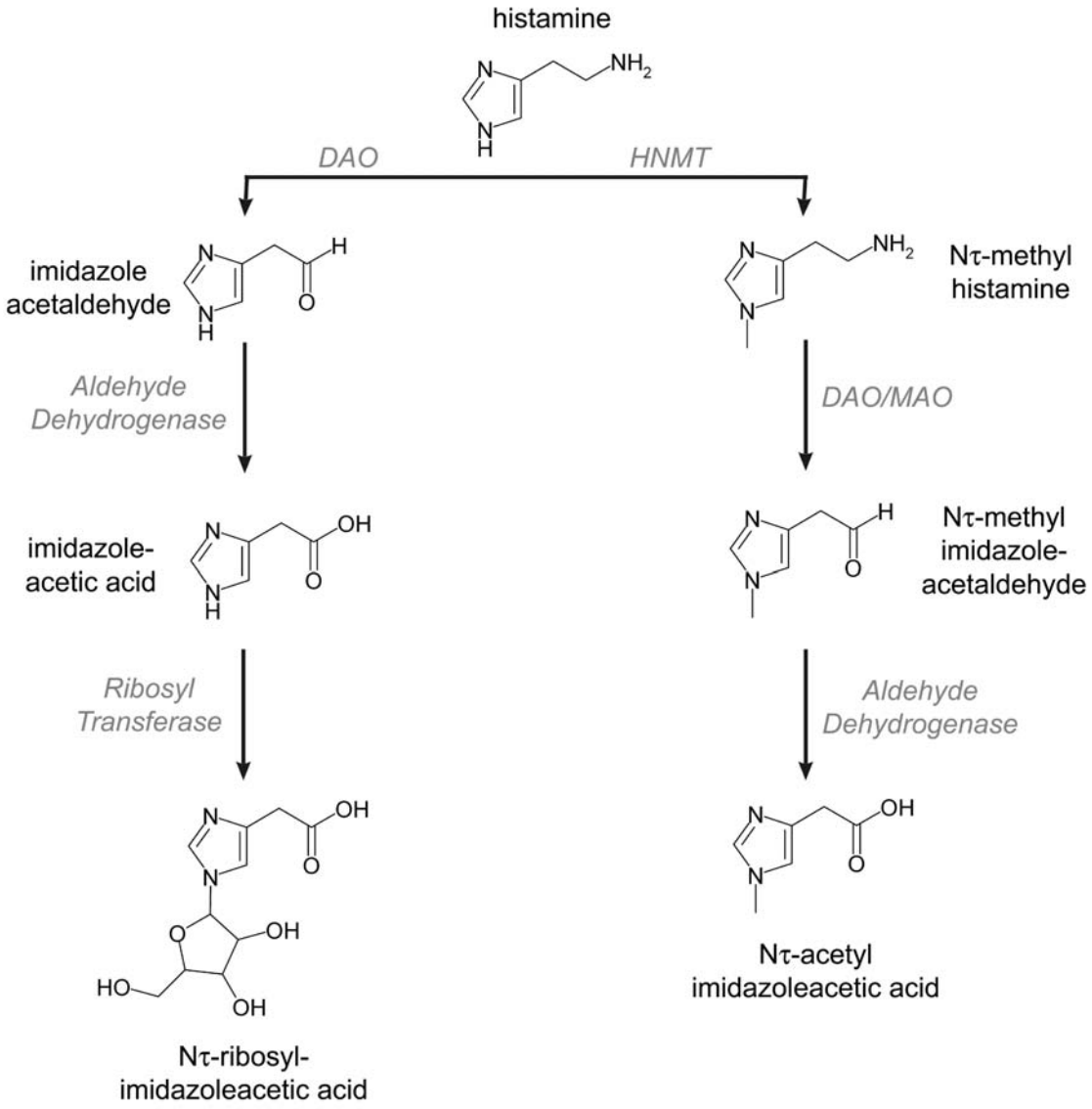

Figure 1.7: Metabolism of histamine.

1.4 The Histamine Receptor Family

The family of histamine receptors comprise to current knowledge four members, termed H_1-H_4. Compounds targeting the peripheral H_1-receptor are used in the therapy of allergic asthma and allergies; while compounds targeting the H_1R in the CNS can be used as sedatives or antiemetics. Antagonists addressing the H_2-receptor can down regulate gastric secretion and are widespread drugs for the treatment of peptic ulcer. To date no drugs targeting the H_3- or H_4-receptor are on the market, although some hH_3R ligands have by now reached clinic phase II. [10] It is expected that these agents have utility in the treatment of obesity, pain, cognitive disorders and allergic rhinitis. Finally, the most recently discovered H_4-receptor has been reported to play an important role in

inflammatory processes, therefore a broad spectrum of therapeutic application is expected also for this newest member of the histamine receptor family. Table 1.1 gives an overview on the main actions produced by histamine on the four receptors.

Table 1.1: Main actions of histamine on histamine receptors.

receptor	main action
hH_1R	• contraction of most smooth muscles in ileum, bronchi, uterus • increased vascular permeability • drop of blood pressure by dilatation of blood vessels • itching if injected to skin by stimulation of nerve endings • CNS: regulation of wakefulness
hH_2R	• stimulation of gastric secretion • cardiac stimulation (increase of rate and output of heart action)
hH_3R	• modulation of histamine release in histaminergic neurons • modulation of histamine synthesis in histaminergic neurons • modulation of the release of acetylcholine, dopamine, noradrenaline, serotonine, GABA, glutamate
hH_4R	• chemotaxis of eosinophils and mast cells to histamine • release of IL-16 from $CD8^+$ T cells

The first human histamine receptor cloned was the hH_2R, in 1991, [11] followed by the hH_1R in 1994, [12] the hH_3R in 1999 [13] and finally the hH_4R in 2000 [14–18]. Sequence analysis of the histaminergic receptor family revealed a low conservation within the family (20%) whereby the hH_3R and hH_4R share with approximately 40% the highest homology. Also the effector systems that are stimulated upon receptor activation are different within the receptor family: the hH_1R activates phospholipase C, the hH_2R stimulates the adenylyl cyclase, while the first cloned hH_3R subtype and the hH_4R inhibit cAMP production. As can be expected from the low degree of conservation, also the pharmacology of histamine receptor agonists and inverse agonists is clearly differentiable.

1.4.1 The Human H_3-Receptor

The histamine H_3-receptor was discovered in 1983 by Arrang and coworkers [19] and has been the focus of intense research over more than 20 years since then. Recently, several review articles have been published on the histamine receptor, [10] H_3R isoforms, [20, 21] on H_3R inverse agonists [22, 23] and agonists [24] that summarise the current knowledge on this receptor. Here, only a brief introduction on topics concerning this work can be given. Table 1.2 gives a short survey on interesting characteristics of the hH_3R.