Investigating the Molecular Basis for the Constitutive Activity of the Nuclear Hormone Receptor CAR

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der

Mathematisch-Naturwissenschaftlich-Technischen Fakultät

(mathematisch-naturwissenschaftlicher Bereich)

der Martin-Luther-Universität Halle-Wittenberg

von Herrn Björn Anselm Windshügel

geb. am 1. Januar 1976 in Bietigheim-Bissingen

Gutachter:

1. Prof. Dr. Wolfgang Sippl

2. PD Dr. Wolfgang Brandt

3. Prof. Dr. Antti Poso

Halle (Saale), den 19. Juli 2006

Acknowledgements

List of Original Publications

This doctoral dissertation is based on the following publications, referred to in the text as Roman numerals **I-III**.

I   Björn Windshügel, Johanna Jyrkkärinne, Antti Poso, Paavo Honkakoski and Wolfgang Sippl
Molecular dynamics simulations of the human CAR ligand-binding domain: deciphering the molecular basis for constitutive activity
J. Mol. Model., 11:69-79 (2005)

II   Johanna Jyrkkärinne, Björn Windshügel, Janne Mäkinen, Markku Ylisirniö, Mikael Peräkylä, Antti Poso, Wolfgang Sippl and Paavo Honkakoski
Amino acids important for ligand specificity of the human constitutive androstane receptor
J. Biol. Chem., 280:5960-5971 (2005)

III   Björn Windshügel, Johanna Jyrkkärinne,Jenni Vanamo, Antti Poso, Paavo Honkakoski and Wolfgang Sippl
Comparison of homology models and X-ray structures of the nuclear receptor CAR: Assessing the structural basis of constitutive activity
J. Mol. Graph. Model., in press

Contents

Chapter 1

Introduction

1.1 Nuclear Hormone Receptors

1.1.1 General Introduction

Biological systems are often characterised by a great complexity accomplished by a multitude of diverse interactions between its integral parts. This applies not only for the macroscopic level (*e.g.* organisation and concerted action of swarm forming animals) but also holds true for the smallest biological building blocks, the cells, whether of protozoan or metazoan nature. Intracellular communication is required for retention of the organisation and the physiological properties of the cell as well as its adaption to varying conditions. In multicellular organisms also the intercellular interactions play a significant role exemplified by diverse endocrine signals.

Referred to as signal transduction, intracellular communication processes often comprise successive biochemical reactions triggering changes in the gene expression profile, energy status or cell locomotion, respectively. Signals comprise small molecules such as steroid and thyroid hormones as well as cyclic nucleotides and phosphoinositide derivatives.

During the last decades nuclear receptors (NRs) have been emerged as key elements in the intracellular signal transduction of metazoans (Owen and Zelent, 2000). By responding to a large variety of hormonal and metabolic signals, NRs act as ligand-activated transcription factors, thus playing a crucial role in the regulation of gene expression. Moreover, NRs are targeted by other signalling cascades and integrate diverse signal transduction pathways involving them

1

in numerous physiological processes comprising development, differentiation, homeostasis and reproduction (Mangelsdorf et al., 1995).

Although the signal molecules such as steroid and thyroid hormones have been isolated in the early 20th century, the targets of those compounds remained unknown for several decades. In 1974, the correlation between hormone action and alterations in the gene expression status was reported (Ashburner et al., 1974). Later studies revealed the classic model of the NR signalling pathway described in detail in the next section (Yamamoto, 1985). The first NRs were cloned in 1985 and represent the starting point of the modern NR research (Hollenberg et al., 1985; Miesfeld et al., 1986; Green et al., 1986). Additional NRs were subsequently identified suggesting the existence of a large NR superfamily that has been evolved from one ancestral orphan receptor and is composed of altogether six sub-families (NR1-NR6) (Petkovich et al., 1987; Evans, 1988; Laudet, 1997). The numerous and often delusive denotations of NRs finally lead to a unified nomenclature system that relies on the homology to other NRs in the most conserved regions (Committee, 1999).

The number of NR genes between species differs significantly. As an example, 21 NR genes have been revealed in *Drosophila melanogaster* whereas in *Caenorhabitis elegans* more than 270 genes have been identified (Robinson-Rechavi et al., 2002). In humans altogether 48 NRs have been discovered so far. This number is close to that of known NR genes in mice (49) (Robinson-Rechavi and Laudet, 2003). Indeed, the number of functionally different NRs is by far larger due to alternative splicing processes (Zhou and Cidlowski, 2005).

1.1.2 Signaltransduction

The main steps of the protein biosynthesis comprise transcription and translation processes that are strictly regulated. Usually, transcription is prevented by the chromatin into which the DNA is assembled. Chromatin is the structural building block of a chromosome composed of nucleosomes (Kornberg, 1974). Each nucleosome is composed of a core constituted by histone proteins around which the DNA is wrapped. Besides providing the lowest level of DNA compaction, nucleosomes are also important for gene regulation. Depending on the acetylation state of histone proteins, the chromatin adopts a more condensed or

a more open form, that prevents or allows the access of the basal transcription machinery, thus repressing or initiating protein biosynthesis.

NRs regulate the gene expression by modulating the histone acetylation status of chromatin at their target gene, thus initiating or silencing the first step of the protein biosynthesis, the transcription process. NRs recognise and bind to specific binding sites in the promoter region of the gene referred to as response element (RE) (Chandler et al., 1983). Depending on the type of RE, NRs not only stimulate gene expression (positive RE), but also may have silencing effects via negative elements that are located in close vicinity of the transcription initiation site or even downstream of the TATA box (Belandia et al., 1998; Perez-Juste et al., 2000; Saatcioglu et al., 1993).

The canonical core recognition motif of REs consists of a central hexameric element having the consensus sequence 5'-AGGTCA-3' (Beato et al., 1995). Number and configuration of the core motif as well as the 5'-flanking region determines the specificity and affinity of the NR (Mader et al., 1993; Juge-Aubry et al., 1997). The length of the spacer region between the core motifs influences the NR specificity as well (Naar et al., 1991; Umesono et al., 1991).

Usually, NRs bind as homo- (Type I) or hetero-dimer (Type II) to their respective REs whose core motifs can be configured as direct repeats (DR), everted repeats (ER) or palindromes. Steroid hormone receptors (*e.g.* ER, AR, GR) almost exclusively recognise REs organised as palindromes whereas non-steroidal receptors (*e.g.* VDR, PPAR, RXR) recognise response elements of different configurations (Kishimoto et al., 2006).

Activation of gene expression requires co-activators and other protein factors to be recruited to the promoter-bound NR that serves as nucleation site for a large multi-protein complex containing histone modifying and chromatin re-modelling activities (Acevedo and Kraus, 2004). Usually, un-liganded NRs are complexed to co-repressors such as the silencing mediator of retinoid and thyroid receptors (SMRT) or the nuclear receptor co-repressor (NCoR) (Chen and Evans, 1995; Hörlein et al., 1995) both recruiting histone deacetylases and chromatin remodelling proteins thus rendering the promoter transcriptionally silent (Kraus and Wong, 2002).

Distinct groups of co-activators with different properties are necessary for NR-dependent transcription: Bridging co-activators act as connectors between NRs and proteins carrying histone modifying or chromatin remodelling activities

(Glass and Rosenfeld, 2000). These type of co-activators, formed by members of the steroid receptor co-activator (SRC) (Leo and Chen, 2000) family, bind to the NR via a the nuclear receptor interaction domain (NRID), an amphipathic helix containing a conserved LxxLL motif (L is leucine whereas x is any amino acid) (Heery et al., 1997; Darimont et al., 1998). Other types of co-activators are constituted by histone-modifying proteins (histone acetyltransfereases, histone methyltransferases and histone kinases) and chromatin-remodelling complexes both opening the promoter by unpacking the DNA from histones (Kraus and Wong, 2002; Dilworth and Chambon, 2001; Hsiao et al., 2002). Furthermore, mediator complexes facilitate binding of RNA polymerase II to the promoter and thus help to establish the pre-initiation complex consisting of general transcription factors (*e.g.* TFIID) and the RNA polymerase II (Rachez and Freedman, 2001).

In contrast to NR activation, antagonist binding can stabilise the NR-co-repressor interactions or even prevent the NR from adopting an active conformation as shown for raloxifene or tamoxifen in the estrogen receptor (Jackson et al., 1997; Brzozowski et al., 1997; Shiau et al., 1998).

1.1.3   Structural Organisation

Nuclear receptors share a conserved structual and functional organisation (Fig. 1.1). Altogether four distinct regions have been characterised which comprise the N-terminal region A/B, a conserved DNA binding domain (DBD, region C), a linker region D and a ligand binding domain (region E). Some nuclear receptors also contain a C-terminal extension (region F) of yet unknown function.

1.1.3.1   N-Terminal Domain

This region is also referred to as modulatory domain due to its promoter and cell context dependent activities (Tora et al., 1988; Berry et al., 1990; Vegeto et al., 1993). The N-terminal region displays most variability both in length and sequence among the NR domains (Krust et al., 1986; Segraves, 1991). For instance, the N-terminal domains of the vitamin D and mineralocorticoid receptor comprise 23 and 602 amino acids, respectively. For each NR multiple modulatory domains are generated by alternative splicing processes, different promoters as well as varying translational start sites thus resulting in various receptor
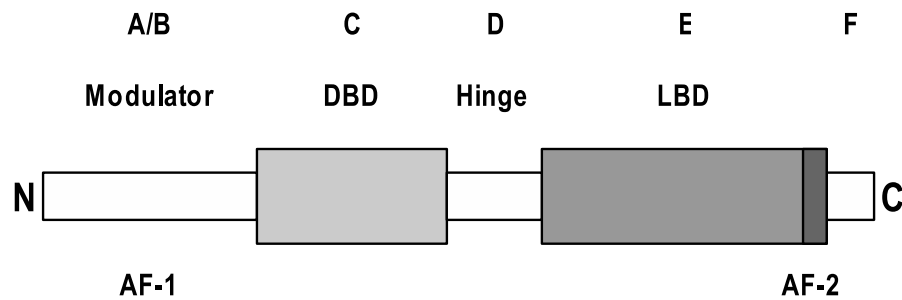
Figure 1.1: Schematic representation of the nuclear receptor architecture. Abbreviations: AF-1/AF-2, activation function-1/2; C, C terminus; DBD, DNA binding domain; LBD, ligand binding domain; N, N terminus.

isoforms with distinct biological function (Conneely et al., 1987; Kastner et al., 1990). The A/B region usually contains a ligand independent transcriptional activation function (AF-1) whose activity is regulated by its phosphorylation state (Tora et al., 1989; Shao and Lazar, 1999). Due to missing structural data the three-dimensional organisation of the A/B region is unknown and the available limited structure-function data is controversial (Wärnmark et al., 2003).

1.1.3.2 DNA Binding Domain

Nuclear receptors recognise their specific target gene via the DNA binding domain (DBD). The DBD represents the most conserved domain in NRs that is composed of two zinc-finger motifs and a C-terminal extension (CTE) domain encoded by roughly 60-70 amino acids (Aranda and Pascual, 2001). Each zinc atom is coordinated in a tetrahedral arrangement by four highly conserved cysteine residues (Freedman et al., 1988). The three-dimensional structure of the DBD has been revealed by NMR and X-ray studies (Fig. 1.2) (Hard et al., 1990; Luisi et al., 1991; Schwabe et al., 1993).

Recognition of the target DNA is achieved by the so-called P-box located in the first zinc finger whereas the second one harbours the D-box which provides the dimerisation determinants (Umesono and Evans, 1989). Unlike the core DBD sequence, the CTE is not conserved among NRs and may adopt diverse structural motifs with different kinds of function comprising sequence recognition or NR dimerisation, respectively (Khorasanizadeh and Rastinejad, 2001; Claessens and Gewirth, 2004).
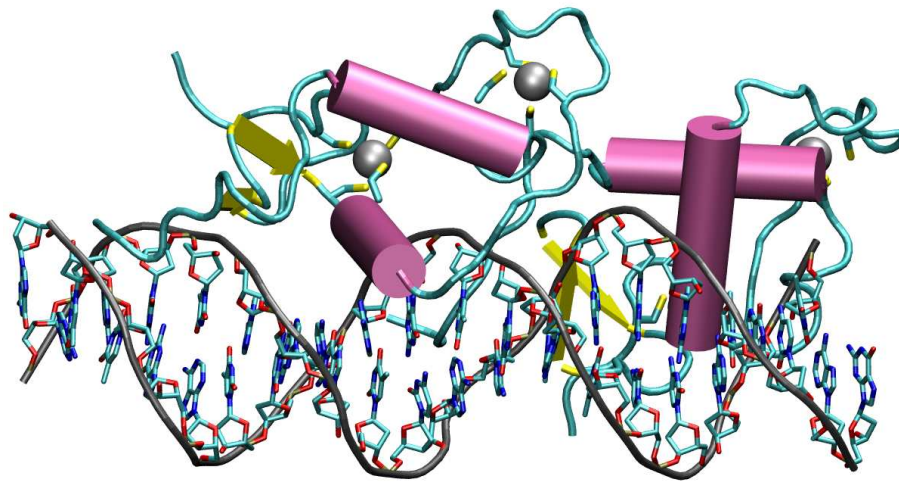
Figure 1.2: Two zinc finger motifs of the NR Rev-Erbα bound to their cognate response element. Helices are coloured in magenta, β-sheets in yellow. Zinc atoms complexed by cysteines are shown in grey.

1.1.3.3 Hinge Region

Similar to the modulatory N-terminal domain, the hinge region exhibits high variability in both, sequence and length. Located between the highly structured DBD and LBD, the hinge provides flexibility to the NR enabling accomodation to different heterodimerisation partners and different types of response elements. Additionally, region D contains a nuclear localisation signal.

1.1.3.4 Ligand Binding Domain

The LBD is a multifunctional domain that, besides the ligand binding site (Dobson et al., 1989), also carries the ligand-dependent transcription activation function 2 (AF-2) (Zenke et al., 1990; Danielian et al., 1992) as well as a dimerisation motif for RXR (Lees et al., 1990; Fawell et al., 1990) and a nuclear localisation signal (Picard and Yamamoto, 1987). Additionally, the LBD harbours binding sites for heat shock and co-regulatory proteins (Housley et al., 1990; Glass and Rosenfeld, 2000).

Despite a considerable variability in sequence, the LBDs of all nuclear receptors possess a canonical structure in which 12 to 14 α-helices, together with a 2- to 5-stranded β-sheet, are arranged in an antiparallel, three-layered helix sandwich
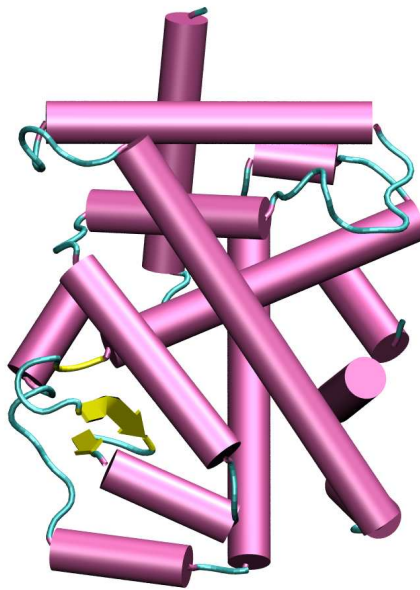
Figure 1.3: The ligand binding domain of the retinoid acid-related orphan receptor β (RORβ). Helices are coloured in magenta, β-strands are shown in yellow and loops are coloured cyan.

(Figure 1.3) (Wurtz et al., 1996; Bourguet et al., 1995; Wagner et al., 1995). Located between the outer sandwich layers, the ligand binding pocket (LBP) is constituted by amino acid of mainly hydrophobic character (Renaud et al., 1995). The size of the LBP can diverge considerably among different NRs ranging from 100 Å^3 (ERRα) up to 1300 Å^3 (PPARγ) (Nolte et al., 1998; Greschik et al., 2002; Kallen et al., 2004). Exceptions are NURR1 and DHR38 lacking any ligand binding cavity (Wang et al., 2003b; Baker et al., 2003).

1.2 The Subfamily NR1I

1.2.1 Vitamin D Receptor

The chicken vitamin D receptor was cloned almost 20 years ago using a monoclonal antibody raised against the purified protein and an intestinal cDNA expression library (McDonnell et al., 1987). Its human ortholog was first cloned in 1988 (Baker et al., 1988). The natural ligand of VDR is the biologically most active vitamin D metabolite, 1α,25-dihydroxyvitamin D_3 (1α,25-$(OH)_2$-D_3), a hormone involved in calcium homeostasis. In addition to its effect in calcium metabolism, 1α,25-$(OH)_2$-D_3 has also potent immunomodulatory effects. VDR is the only nuclear protein binding 1α,25-$(OH)_2$-D_3 with high affinity (K_d = 0.1 nM). VDR belongs to the classic endocrine receptor subgroup of the NR superfamily which also contains receptors for retinoid acid, thyroid hormone, estradiol and others (Carlberg, 1995).
Analogues of 1α,25-$(OH)_2$-D_3 are used in the treatment of hyperproliferative diseases such as psoriasis and different types of cancers (Hansen et al., 2001) as well as bone disorders such as osteoporosis (Brown, 2001). New studies have also adressed the VDR as a potential target for the treatment of multiple sclerosis, type I diabetes and inflammatory bowel diseases (Nagpal et al., 2005).

1.2.2 Pregnane X Receptor

First cloned in *Xenopus laevis* as a novel member of the NR superfamily, this receptor was initially termed xONR1 (xenopus orphan nuclear receptor 1) and later referred to as BXR (benzoate X receptor) (Smith et al., 1994; Blumberg et al., 1998a). A related mouse gene was identified in 1997 and the term PXR was introduced (Kliewer et al., 1998). Also known as SXR (steroid and xenobiotic receptor) and PAR (pregnane-activated receptor), the PXR was initially believed to respond to endogenous C21 steroids, so-called pregnanes (Blumberg et al., 1998b; Bertilsson et al., 1998). Later studies revealed PXR as a central xenobiotic receptor that responds to many clinically used drugs, among them calcium channel blockers, statins, antidiabetic drugs and HIV protease inhibitors (Handschin and Meyer, 2003).
 As a matter of fact, PXR is by far the most promiscuous NR wich can bind a diverse array of structurally distinct compounds that vary in molecular weight
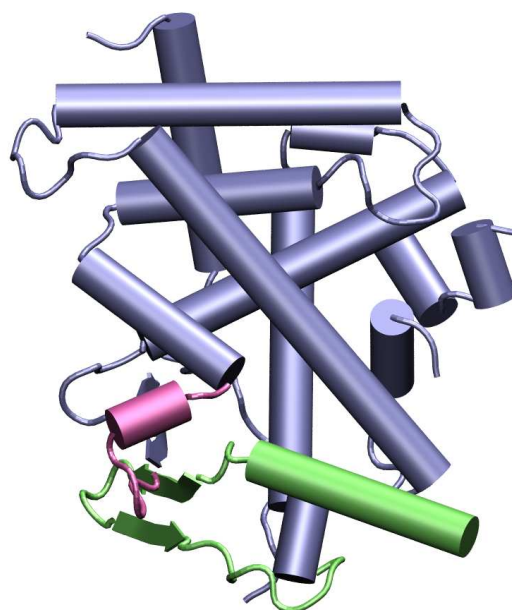
Figure 1.4: X-ray structure of the pregnane X receptor LBD (PDB code 1NRL). Deviations from the common NR topology are coloured in green (60 amino acid insert carrying two additional β-strands and a helix) and magenta (unfolded H6 and broken H7).

from 200 to 800 Da (Kliewer et al., 2002). In contrast to the closely related receptor VDR, the binding affinity of PXR ligands is significantly lower, usually in the micromolar or even millimolar range (Kliewer et al., 2002). The large promiscuity in ligand binding is enabled by a LBD insert consisting of approximately 60 amino acids that constitutes an integral part of the ligand binding pocket. Furthermore, helix unwinding and breaking enlarges the ligand binding crevice (Fig. 1.4). By possessing a flexible and conformable LBP, PXR is enabled to adapt to structurally diverse ligands as revealed by PXR crystals complexed with various ligands (Watkins et al., 2001; Watkins et al., 2003b; Watkins et al., 2003a; Chrencik et al., 2005).

PXR is expressed predominantly in the liver and intestine and regulates gene expression as a heterodimer together with the retinoid X receptor (RXR) (Kliewer et al., 1998; Lehmann et al., 1998). As a key regulator in drug metabolism, PXR target genes encompass all three phases of drug metabolism (Orans et al., 2005). The primary objective of PXR is represented by CYP3A4 that is known to metabolise the majority of drugs in use today. Moreover, PXR responds to

endogenous toxins such as lithocholic acid by regulating expression of bile acid metabolising enzymes and transporters, thus preventing hepatorenal toxicity (Staudinger et al., 2001; Xie et al., 2001).

1.2.3 Constitutive Androstane Receptor

CAR was discovered in 1994 by screening a cDNA library with a degenerate oligonucleotide based on a conserved region of the DNA binding domain (Baes et al., 1994). Originally denoted as MB67, CAR was found to exhibit an intrinsic basal activity in cell based reporter assays which was in contrast to other NRs known so far. In 1997, a mouse homologue of MB67 was discovered and denoted CAR (Choi et al., 1997). Since the receptor was found to form a heterodimer with RXR in presence of retinoids and rexinoids the term CAR initially stood for "constitutive activator of retinoid response". Upon discovery of testosterone metabolites 5α-androstan-3α-ol and 5α-androst-16-en-3α-ol, both repressing the basal activity the meaning of CAR was changed into "constitutive androstane receptor" (Forman et al., 1998). The major expression sites of CAR are the liver and the intestine (Wei et al., 2000; Wei et al., 2002). Lower levels of CAR were found in heart and skeletal muscles as well as the kidney (Baes et al., 1994; Choi et al., 1997). Alternative spliced variants of CAR have been also detected in the adrenals, brain, spleen, prostate and testis that were found to be affected in their functional properties (*e.g.* loss of basal activity) (Lamba et al., 2004; Arnold et al., 2004).

1.2.3.1 Signal Transduction

As a nuclear receptor with intrinsic basal activity, the ligand-independent gene expression must be repressed in order to acquire responsiveness to activating compounds. Usually, classical nuclear receptors reside permanently at the nucleus, thus agonists or antagonists act directly through binding to the NR. By contrast, un-liganded CAR is retained in the cytoplasm of liver cells by the recently discovered cytoplasmic CAR retention protein (CCRP) (Kawamoto et al., 1999; Kobayashi et al., 2003), thus preventing an unregulated gene expression. Together with the molecular chaperone HSP90 (heat shock protein 90) and the CCRP, CAR constitutes a ternary complex that is associated with microtubules

at the cell membrane (Yoshinari et al., 2003; Kobayashi et al., 2003; Koike et al., 2005). Upon activation, CAR is translocated into the nucleus regulated by a leucine-rich region in the C-terminal region of the LBD (Zelko et al., 2001).

Currently, three distinct activation mechanisms are known:

- Direct ligand binding dissociates the ternary complex allowing the translocation of CAR into the nucleus.

- An indirect activation mechanism involves a poorly defined signalling cascade that requires the recruitment of protein phosphatase 2A to the ternary complex (Yoshinari et al., 2003). Very recently, the dephosphorylation of a specific serine has been revealed as an essential step for nuclear translocation of murine CAR (Hosseinpour et al., 2006). Most known CAR agonists act through this type of mechanism exemplified by the prototypic CAR inducer phenobarbital as well as lithocholic acid and steroids such as estrone and estradiol (Honkakoski et al., 1998b; Moore et al., 2000; Kawamoto et al., 2000).

- Finally, a ligand independent activation is mediated by the co-activator protein peroxisome proliferator-activated receptor γ co-activator-α (PGC-1) targeting CAR to nuclear speckles (Shiraki et al., 2003).

Once in the nucleus, CAR forms a heterodimer with the retinoid X receptor. Early studies demonstrated binding to response elements of DR5 organisation (two direct repeats with a spacer of 5 bp) (Baes et al., 1994; Choi et al., 1997). Later on, the CAR/RXR heterodimer was shown to bind also to DR4 motifs within the enhancer PBREM (phenobarbital responsive enhancer module) of CYP2B genes (Honkakoski et al., 1998b; Sueyoshi et al., 1999). Furthermore, a CAR binding site with ER6 configuration has been reported for the CYP3A4 gene (Sueyoshi et al., 1999).
Several co-activators have been revealed as capable to interact with nuclear CAR such as SRC-1 (Forman et al., 1998; Tzameli et al., 2000; Jyrkkärinne et al., 2003), SRC-3 (Kim et al., 1998), GRIP-1 (Min et al., 2002), PGC-1 (Shiraki et al., 2003) and TIF-2 (Frank et al., 2004). The processes finally leading to CAR-dependent gene expression have not been elucidated so far.

1.2.3.2 CAR Ligands

According to the provoked biological response, ligands can be categorised into agonists, antagonists or inverse agonists, respectively. Agonists are defined as compounds with activating effects on their target protein, whereas antagonists show inhibiting effects. The synonym inverse agonist arose upon discovery of constitutively active receptors, whose basal activity is abolished by those type of compounds.

Phenobarbital was the first known CAR agonist activating CYP2B10 gene expression in mice via the PBREM (Honkakoski et al., 1998b). Phenobarbital is the prototype of a group of structurally unrelated chemicals activating a subset of of P450 within the CYP2A, 2B, 2C and 3A subfamilies (Okey, 1990; Waxman and Azaroff, 1992; Honkakoski and Negishi, 1997). PB-type inducers comprise a structurally different family of CAR agonists including chlorpromazine, phenytoin, dichlorodiphenyltrichloroethane, the pesticide contaminant 1,4-bis[2-(3,5-dichloropyridyloxy)]benzene (TCPOBOP) as well as polychlorinated biphenyls (Honkakoski et al., 1998a; Wei et al., 2002; Wang et al., 2004a; Jackson et al., 2004; Tzameli et al., 2000).

Indeed, assignment of a ligand to a certain category is often problematic. Depending on the cell line used for assays, compounds can turn out to be agonists, antagonists or inverse agonists. The varying expression pattern of NR co-activators in different cell lines may lead to differential response of NR modulators (Smith et al., 1997; Liu et al., 2002). As an example, clotrimazole has been reported as agonist in HEK293 cells (Mäkinen et al., 2002; Jyrkkärinne et al., 2003; Honkakoski et al., 2004), whereas in a CV-1 cell line clotrimazole behaves as an inverse agonist and even no effect on CAR activity was observed in DLD cells (Moore et al., 2000; Toell et al., 2002). Additionally, the specific human CAR agonist 6-(4-chlorophenyl)imidazo[2,1-b][1,3]thiazole-5-carbaldehyde O-(3,4-dichlorobenzyl)oxime (CITCO) has been found to strongly activate CAR in CV-1 cells whereas in FLC-7 cells CITCO activity was weak (Maglich et al., 2002; Kobayashi et al., 2005). There are not only different biological responses in varying cell lines but also species specific differences exemplified by the pesticide contaminant TCPOBOP that strongly activates mouse CAR but does not show any impact on human CAR (Tzameli et al., 2000; Moore et al., 2000). Results for the antiemetic drug meclizine are similar. While an efficient agonist

in mouse CAR, meclizine suppresses CAR-mediated transactivation of human CAR (Huang et al., 2004b). In contrast, several different HMG-CoA reductase inhibitors, often clinically used in the treatment of hypercholesterolemia, were found to strongly enhance either mouse, rat or human CAR-mediated transcriptional activity in cell-based reporter gene assays (Kobayashi et al., 2003). Also Yin Zhi Huang, a decoction of Yin Chin (*Artemisia capellaris*) and three other herbs, widely used in Asia to prevent or treat neonatal jaundice, activates CAR in both, humans and mice (Huang et al., 2004a).

1.2.3.3 Regulation of Drug Metabolism

The human body is exposed to a huge variety of chemical compounds (also called xenobiotics) of which many enter the body via the lung, gastrointestinal system and skin. Once in the body, soluble toxins are detained penetrating the cell by the plasma membrane whose hydrophobic character is an efficient protector against those kind of substances. However, more lipophilic compounds easily cross the membrane and may accumulate in the cell to toxic levels. To protect oneself from such potentially harmful substances, the human body has evolved mechanisms encompassing the biotransformation of lipophilic to more soluble compounds and transport processes that readily eliminate the transformed xenobiotics from the cell and finally the body. The biotransformation of xenobiotics is subdivided into two distinct phases:

Phase I comprises oxidation, reduction or hydroxylation processes catalysed by members of the cytochrome P450 superfamily (CYPs), flavinmonooxygenases, alcohol dehydrogenase, hydrolases and also decarboxylases (Grant, 1991). Members of the CYP3A and CYP2B subfamily are the most relevant cytochromes for xenobiotic metabolism. CYP3A4 is the most abundant cytochrome in the human liver and small intestine metabolising approximately half of the prescription drugs in use today (Guengerich, 1999).

Phase II is characterised by coupling reactions catalysed by specific transferases that conjugate the compound with an endogenous hydrophilic substance (Bock et al., 1987). Compounds used for conjugation comprise glucuronic acid, sulfate, amino acids (in particular glycin) as well as S-adenosylmethionin (Caldwell, 1982).

Upon biotransformation, compounds are discharged from the cell via specific transporters into the blood or bile from where they are finally eliminated from the body via the kidney or gastrointestinal system, respectively (Kim, 2002). This phase is also often referred to as *Phase III*.

In absence of foreign compounds drug metabolising enzymes and transporters are usually expressed at low level in the cell (Handschin and Meyer, 2003) whereas in the presence of a potentially harmful compound gene expression is increased dramatically (Remmer, 1958; Conney et al., 1960). This adaptive increase in gene expression, termed xenobiotic induction, is reversible upon removal of the compound (Negishi and Honkakoski, 2000).

Regulation of gene expression is achieved by nuclear receptors functioning as xenosensors. Especially PXR and CAR have been emerged as major players in drug metabolism (Honkakoski et al., 2003). By recognising a wide array of structurally diverse compounds, both nuclear receptors subsequently activate the gene expression of cytochromes, transferases and transporter proteins (Willson and Kliewer, 2002).

CAR transcriptionally regulates not only the gene expression of main metabolising enzymes CYP3A4 and CY2B10 (Goodwin et al., 2002; Honkakoski et al., 1998b) but also other members of the 2B, 2C and 3A family. Table 1.1 summarises all known CYPs controlled in a CAR dependent manner. Most of those have been found in rats and mice. Besides CYPs, CAR also regulates the gene expression of phase II enzymes such as UDP-glucuronosyltransferases (UGTs), glutathione-S-transferases (GSTs) and sulfotransferases (SULTS) (Maglich et al., 2002; Ueda et al., 2002). Additionally, also genes coding for transporters of biotransformed xenobiotics like multidrug resistance protein 1 (MDR1) (Maglich et al., 2002; Ueda et al., 2002; Burk et al., 2005) and resistance-associated proteins 1 to 3 (MRP1-3) (Kast et al., 2002; Cherrington et al., 2002; Maher et al., 2005) are transcriptionally regulated by CAR (see also Table 1.1)

1.2.3.4   Role in Bilirubin Clearance

Besides foreign compounds also endogenous toxins such as bilirubin need to be cleared from the body. Bilirubin is the oxidative product of the protoporphyrin ring of the heme group found in proteins such as hemoglobin, myoglobin and the cytochromes P450. 250-400 mg bilirubin are produced in adults per day

Phase	Target	Organism	Reference
I	CYP1A1	Mouse	(Maglich et al., 2002)
	CYP1A2	Mouse	(Maglich et al., 2002)
	CYP2A4	Mouse	(Maglich et al., 2002)
	CYP2B1	Rat	(Muangmoonchai et al., 2001)
	CYP2B2	Rat	(Xiong et al., 2002)
	CYP2B6	Human	(Sueyoshi et al., 1999)
	CYP2B10	Mouse	(Honkakoski et al., 1998b)
	CYP2C8	Human	(Ferguson et al., 2005)
	CYP2C9	Human	(Ferguson et al., 2002b)
			(Gerbal-Chaloin et al., 2002)
	CYP2C19	Human	(Ferguson et al., 2002a)
	CYP2C29	Mouse	(Jackson et al., 2004)
	CYP3A1	Rat	(Smirlis et al., 2001)
	CYP3A4	Human	(Goodwin et al., 2002)
	CYP3A5	Human	(Burk et al., 2004)
	CYP3A11		(Yamazaki et al., 2005)
	ALDH1	Mouse	(Maglich et al., 2002)
	FMN	Mouse	(Ueda et al., 2002)
II	GST	Mouse	(Maglich et al., 2002)
	SULT1A1	Mouse	(Maglich et al., 2002)
	SULT2A9	Mouse	(Saini et al., 2004)
	UGT1A1	Human	(Sugatani et al., 2001)
III	MDR1A	Mouse	(Maglich et al., 2002)
	MRP1	Mouse	(Maglich et al., 2002)
	MRP2	Mouse	(Kast et al., 2002)
	MRP3	Mouse	(Maglich et al., 2002)

Table 1.1: Overview of enzymes and transporters for each phase of the biotransformation and excreation process regulated in a CAR-dependent manner. Abbreviations: ALDH, aldehyde dehydrogenase; CYP, cytochrome P450; FMN, flavin monooxygenase; GST, glutathione-S-transferase; MDR, multidrug resistance protein; MRP, multidrug resistance-associated protein; UGT, UDP-glucuronosyltransferase; SULT, Sulfotransferase.

and the removal from the body is exerted by the liver, where bilirubin is glu-curonidated and secreted into the bile. A disturbed bilirubin clearance results in increased serum concentrations and finally in jaundice which is particularly common in neonates (Dennery et al., 2001) Due to its high hydrophobicity bilirubin accumulates in the central nervous system and sustained hyperbiliru-binemia may lead to neurotoxicity and encephalopathy (Roy-Chowhury et al., 1995).

The elimination of bilirubin is initiated by uptake across the sinusoidal hep-atocyte membrane via the organic anion transporting polypeptide 2 (OATP2) (König et al., 2000). In the cytosol bilirubin is associated with ligandin, a homo-or heterodimer of glutathione-S-transferase (GST) A1 and A2, transporting bilirubin to the endoplasmic reticulum (Mannervik, 1985). Reached the endo-plasmic reticulum, bilirubin is glucuronidated by the UDP-glucuronosyltrans-ferase (UGT1A1) (Tukey and Strassburg, 2000). Finally, hydrophilic bilirubin glucuronide is secreted across the canalicular membrane into the bile via the multi drug resistance-associated protein 2 (MRP2) (Kamisako et al., 1999).

The nuclear receptor CAR has been revealed as a key regulator of the bilirubin clearance pathway (Yamamoto et al., 2003). Expression of OATP2, GST A1/A2 and UGT1A1 as well as MRP2 is under transcriptional control of CAR (Wagner et al., 2005; Maglich et al., 2002; Xie et al., 2003; Sugatani et al., 2001; Kast et al., 2002). In case of elevated bilirubin levels CAR is activated by translocation into the nucleus. Similar to phenobarbital, bilirubin does not activate CAR directly but achieves translocation via an indirect mechanism (Huang et al., 2003). The role of CAR for bilirubin clearance is emphasised by the molecular basis of neonatal jaundice, where low CAR expression in newborns results in low UGT1A1 activity leading to raised bilirubin serum concentrations (Burchell et al., 1989; Huang et al., 2003).

1.2.3.5 CAR and Energy Metabolism

CAR not only operates as central component in the response to xenobiotic and endobiotic stress but also responds to nutritional stress by increasing the metabolism of thyroid hormones T_3 (triiodothyronine) and T_4 (thyroxine) which are the predominant regulators of the basal metabolic rate. Serum levels of T_3 and T_4 are directly correlated with energy expenditure and caloric loss.

Prolonged fasting periods lead to a significant drop of thyroid hormone levels by a mechanism poorly understood so far. Metabolism of thyroid hormones is accomplished by miscellaneous pathways (Visser, 1996). Both, T_3 and T_4 are inactivated either by deiodinases D1 and D3 deiodinating the inner ring or by phase II drug-metabolising enzymes UGT and SULT, respectively (Visser et al., 1993; Visser et al., 1998). Upon fasting, glucagon induces increase of cAMP levels that lead to gene expression of the NR co-activator PGC-1 via the cAMP response element-binding protein (CREB) (Yoon et al., 2001; Herzig et al., 2003). Subsequently, CAR is translocated into the nucleus by PGC-1 in a ligand independent manner where it transcriptionally activates UGTs and SULTs (Shiraki et al., 2003; Maglich et al., 2004; Qatanani et al., 2005).

1.2.3.6 Adverse Effects

Both CAR and PXR are activated by many different compounds resulting in gene expression of metabolising enzymes and transporters. As a result, biotransformation and transport processes dispose those undesirable compounds from the body. Beside this positive impact also adverse effects such as drug-drug interactions are well known as shown for the antidiabetic drug troglitazone that is metabolised into a toxic compound by CYP3A4 in a PXR dependent manner (Yamazaki et al., 1999). Another example is the cancer therapeutics taxol that is rapidly cleared from the body due to its activating effects on PXR (Kostrubsky et al., 1998). Also for CAR drug-drug interactions have been reported: The anticonvulsant drug phenytoin affects the pharmacokinetics of co-administered antineoplastics cyclophosphamide and ifosfamide by induction of CYP2B6 via the CAR (Ducharme et al., 1997; Williams et al., 1999).

CAR not only induces hepatic expression of detoxification enzymes but also modulates the liver size upon acute xenobiotic stress, augmenting the ability to clear an unwanted compound (Diwan et al., 1992; Whysner et al., 1996). By contrast, chronially elevated levels of CAR activators lead to hepatotoxicity as shown for the analgesic paracetamol and the formerly applied anesthetic agent carbon tetrachloride that may finally result in hepatocarcinogenesis as shown for the prototypic CAR activator phenobarbital and the pesticide contaminant TCPOBOP (Yamazaki et al., 2005; Zhang et al., 2002; Yamamoto et al., 2004; Huang et al., 2005).

1.2.3.7 Therapeutic Potential

The current market for nuclear receptor targeted drugs is estimated to be 10-15% of the 400 billion dollar global pharmaceutical market (Goodwin and Moore, 2004). Several compounds are clinically used successfully with the estrogen receptor as one of the most important targets. Selective estrogen receptor modulators such as raloxifene and tamoxifen are used in the treatment of osteoporosis and breast cancer, respectively (Deroo and Korach, 2006). The PPARγ is the target of thiazolidinediones, used as effective insulin-sensitising drugs in type II diabetes (Semple et al., 2006).

Analogues of the natural compound 1α,25-(OH)$_2$-D$_3$ that binds to the vitamin D receptor are used in treatment of hyperproliferative diseases such as psoriasis and different types of cancers (Hansen et al., 2001) as well as bone disorders such as osteoporosis (Brown, 2001). New studies have also adressed the VDR as potential target for the treatment of multiple sclerosis, type I diabetes and inflammatory bowel diseases (Nagpal et al., 2005).

Currently, agonists of oxysterol receptor LXR and bile acid receptor FXR are evaluated in clinical studies for their potential in the treatment of atherosclerosis and cholestasis, respectively (Joseph and Tontonoz, 2003; Claudel et al., 2003).

Phototherapy has been a widely used therapy of neonatal jaundice for which also the application of phenobarbital, clofibrate and D-penicillamine has been considered (Dennery, 2002). Phenobarbital prevents the accumulation of bilirubin by improving conjugation in a CAR-dependent fashion. However, treatment with phenobarbital will result in immediate side effects such as somnolence and stupor and may even have neurotoxic effects (Hansen and Tommarello, 1998). The fact that a compound of the traditional chinese medicine Yin Zhi Huang reduces bilirubin-levels by activating the constitutive androstane receptor makes CAR an interesting pharmacological target for the development of improved therapeutics for neonatal jaundice.

Moreover, CAR offers a potential target in therapy of obesity and cholestasis. During fasting CAR lowers thyroid hormone levels, thereby restricting caloric loss. As a lesser reduction of thyroid hormone levels would increase weight loss during caloric restriction, selective CAR antagonists would represent a target for the treatment of obesity (Maglich et al., 2004). Detoxification

of bile acids involves sulfation by SULTs transferring a sulfonyl group from 3'-phosphoadenosine-5'-phosphosulfate (PAPS) to the acceptor molecule. Gene expression of either PAPS synthetase as well as the corresponding SULT is regulated by CAR, thus an agonist could be applied in the therapy of bile acid disorders, such as cholestasis (Saini et al., 2004).

1.3   Aim of the Work

The human constitutive androstane receptor (CAR) is a key regulator in the gene expression of enzymes and transporters involved in metabolism of endogenous and foreign compounds. Unlike other NRs, CAR possesses an intrinsic basal activity *in vitro* that can be either repressed or enhanced by inverse agonists or agonists, respectively. Usually, NRs are activated in a ligand-dependent manner in which agonist binding rearranges the C-terminal helix H12 into an active position (Li et al., 2003). Available structural information of constitutively active NRs (*e.g.* ERRγ, PPAR) suggests that H12 is permanently in the active conformation, even in absence of any ligand. According to three-dimensional data, the constitutive activity of NRs is mainly achieved by specific side chain interactions between the LBD and H12 including vdW/hydrophobic contacts (*e.g.* in ERRγ) as well as salt bridges and hydrogen bonds (*e.g.* PPAR), respectively (Hong et al., 1999; Molnar et al., 2005). Due to the different types of interaction and the large variety of residues contributing, a structural mechanism for CAR constitutive activity cannot be simply derived from available X-ray data of other NRs.

Therefore, the main goal of this work was the generation of a homology model for the human CAR LBD in order to ascertain molecular determinants for the structural basis of the constitutive activity. The reliability of the model should be verified by experimental mutagenesis studies. Furthermore, the three-dimensional model should be used in order to study the interactions with known CAR agonists.

Several structurally diverse ligands are known to modulate the CAR activity in a yet unknown manner. In order to identify amino acids determining the ligand specificity, the ligand binding pocket of CAR should be investigated in detail. Potential amino acids critical for ligand binding should be selected based on structural data from the homology and the effect of point mutations should be studied experimentally. Additionally, an homology model of the inactivated CAR should be generated in order to propose a potential molecular mechanism of inverse agonist action.

Finally, the model should be used in the search for novel CAR agonists. Based on database searches and virtual screening approaches potential agonists should be selected and tested for activating effects on CAR in experimental assays.

Chapter 2

Computational Methods

2.1 Homology Modelling

All data of the world wide published protein structures are accessed at the
Protein Data Bank (PDB) of the Rutgers university (New Jersey, USA) (Bern-
stein et al., 1977; Berman et al., 2000). Despite the considerable increase of this
database in recent years (36000 entries, April 2006) the number of known pro-
tein sequences (Swiss-Prot/TrEMBL databases: 3 million entries, April 2006)
by far exceeds that for solved 3D structures. To obtain structural information
of non-crystallised proteins theoretical approaches such as threading or homol-
ogy modelling procedures can be applied (Lengauer, 2003). Those methods
utilise the fact that proteins with a considerably high sequence identity share
a similar structure. This became evident by the X-ray studies on hemoglobin
as well as myoglobin and also holds true for the nuclear receptor superfamily
(Perutz et al., 1965; Wurtz et al., 1996). Despite a considerable diversity in their
sequences the general NR fold is very similar.

Threading methods are applied when no structural information of sequentially
related proteins is available (Bowie et al., 1991; Jones et al., 1992; Jones and
Thornton, 1993). A probable three-dimensional fold of the target sequence can
be suggested by comparison to a library containing representative protein folds
(Lo Conte et al., 2000) through which the sequence is threaded. By applying a
scoring function the fitness of the alignment between the target sequence and
each protein in the structure database is assessed.

Homology modelling procedures aim to reproduce the three-dimensional struc-

ture of the protein of interest based on available structural information from a protein of related sequence (template). The accuracy of comparative models correlates with the sequence identity between the target protein and its template (Chothia and Lesk, 1986). Low sequence identities (less than 25 %) most probably lead to models of poor quality (Sander and Schneider, 1991; Rost, 1999; Abagyan et al., 1994), whereas values of about 30 % are expected to allow generation of reasonable models having more than 85 % of the C_α atoms within 3.5 Å of the correct position (Marti-Renom et al., 2000). Reaching more than 50 % sequence identity, average comparative models may approach the accuracy of low resolution X-ray structures or NMR structures of medium resolution. Usually, the homology modelling procedure follows the basic protocol suggested by Greer (Greer, 1980; Greer, 1990):

- Identification of a template structure with reliable sequence identity

- Borrowing backbone coordinates for core secondary structures

- Construction of segments for which coordinates cannot be obtained from the template

- Assignment of side chain conformations

2.1.1 Template Selection

The accuracy of template selection is a critical step in the homology modelling process as the choice of a wrong or inappropriate template may result in an inadequate model. Therefore, the starting point of every modelling approach is the identification of proteins that qualify as feasible templates by both, a sequence related to that in the target as well as the availability of structural data. Potential templates can be identified by search methods such as FASTA, BLAST or PSI-BLAST (Pearson, 1990; Altschul et al., 1990; Altschul et al., 1997) scanning protein databases such as Swiss-Prot and TrEMBL (Boeckmann et al., 2003) for sequentially related proteins. The resulting proteins can be examined for existing structural data using the PDB The sequence identity is not an exclusive argument for the template selection. Also the quality of the template structure has to be taken into account. Beside the resolution of the X-ray crystal also

stereochemical parameters (see section 2.6 for details) as well as the complete-
ness of the structure has to be considered.

Once a suitable template protein has been selected alignment algorithms are
applied to determine regions of similar sequence between model and tem-
plate which are believed to correspond to the same three-dimensional structure
(Chothia and Lesk, 1986).

2.1.2 Sequence-Structure Alignment

Sequence alignment and secondary structure predictions provide information
about structurally conserved regions (SCRs) for which backbone coordinates
can be transferred from the template into the model. Consisting of secondary
structural elements such as helices and β-sheets, SCRs serve as scaffold for
the construction of structural variable regions (SVR) such as loops. Often cor-
responding loops in template and target differ in sequence and length, thus
preventing adoption of those coordinates from the template structure. There-
fore, an appropriate conformation for the backbone of those regions needs to
be created. Techniques generating loop conformations can be divided into loop
search methods and *de novo* generation procedures (Johnson et al., 1994).

In principle, loop search approaches scan a database containing protein or pep-
tide structures for segments of similar or identical sequence that fit properly
into the model's spatial environment (Jones and Thirup, 1986).

The *de novo* modelling approach is applied when the loop search method fails to
meet the geometrical criterion, for example due to steric clashes with conserved
segments. Based on randomly assigned values for dihedral angles a loop is
built between two conserved segments (Go and Scheraga, 1970; Shenkin et al.,
1987).

2.1.3 Assignment of Side Chains

The emerging structure of sequence-structure alignment is the backbone of the
target protein to which side chains have to be assigned. Several side chain
prediction algorithms have been developed in the last years of which most are
based on a rotamer library of discrete side chain conformations (Canutescu
et al., 2003). Early attempts used fixed rotamer libraries often in conjunction

with Monte Carlo simulations for selection of suitable conformations (Ponder and Richards, 1987; Holm and Sander, 1992). A frequently used algorithm is implemented in the programme SCWRL (**S**ide **C**hain placement **W**ith a **R**otamer **L**ibrary). Based on a backbone-dependent rotamer library the algorithm adds side chains to a protein backbone (Bower et al., 1997; Dunbrack Jr., 1999). The library provides lists of χ_1-χ_2-*chi*$_3$-χ_4 values and their relative probabilities for residues at given ϕ/ψ values and explores these conformations to minimise side-chain- backbone and side chain-side chain clashes. In order to incorporate side chain conformations of conserved amino acids directly from the template amino acids can be excluded from the assignment procedure.

2.2   Force Field Methods

Normally, structures emerging from the modelling process contain steric errors due to unusual bond lengths and angles resulting in unfavourable high energies. In order to relax and refine the structure, force field methods such as minimisation approaches and molecular dynamics (MD) simulations are applied. Beside force field methods also quantum mechanical and empirical approaches can be applied:

Quantum mechanical calculations comprise *ab initio* and semi-empirical methods in which electrons are explicitly represented. Both attempts base on approximations of the Schrödinger equation. Semi-empirical methods illustrate a simplified approach of *ab initio* approaches as the calculation of speed limiting integrals are replaced by approximations that are faster to compute. Due to the computational demanding calculations quantum mechanical approaches are restricted to systems of relatively small size.

Force field methods represent a simplified approach as the electron configuration of the system is neglected. The fundamental idea is the description of the molecular system by Newtonian mechanics for which the term molecular mechanics has been introduced. Thus, the energy of a system is described as a function of the nuclear coordinates. In contrast to quantum mechanical methods, empirical force fields can also handle large systems containing hundreds of thousands of atoms.

Several different empirical force fields are available targeting different classes of biomolecules (Mackerell Jr., 2004):

AMBER	**A**ssisted **M**odel **B**uilding and **E**nergy **R**efinement (Pearlman et al., 1995)
CHARMM	**C**hemistry at **HAR**vard **M**acromolecular **M**echanics (Brooks et al., 1983)
OPLS	**O**ptimised **P**otential for **L**iquid **S**imulations (Jorgensen and Tirado-Rives, 1988)
GROMACS	**GRO**ningen **M**achine for **C**hemical **S**imulations (Berendsen et al., 1995)

Typically, an empirical force field is composed of several energy terms describing bonded and non-bonded interactions within the system to be studied (Equation 2.1).

$$E = E_{bond} + E_{theta} + E_{phi} + E_{impr} + E_{elec} + E_{vdw} \tag{2.1}$$

Bonded interactions can be subdivided into terms describing bond stretching, angle bending and bond rotation (torsion angle). Bond stretching and angle bending are treated by a simple harmonic potential (Hooke's law, Equations 2.2 and 2.3), whereas torsion angle rotations are described by a sinusoidal or cosinusoidal term (Equation 2.4).

$$E_{bond} = \sum k_b (r - r_o)^2 \tag{2.2}$$

$$E_\theta = \sum k_\theta (\theta - \theta_0)^2 \tag{2.3}$$

$$E_\phi = \sum | k_\phi | - k_\phi cos(n\phi) \tag{2.4}$$

Non-bonded interactions comprise electrostatic as well as atom-atom repulsion and dispersion (van der Waals) interactions. The majority of biomolecular force fields treat the electrostatic interactions using Coulomb's law (Equation 2.5) whereas van der Waals (vdW) interactions are typically considered by a Lennard-Jones potential (Equation 2.6).

$$E_{elec} = \sum_{excl(i,j)} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{2.5}$$

$$E_{vdw} = \sum_{excl(i,j)} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) sw(r_{ij}^2 r_{on}^2 r_{off}^2) \tag{2.6}$$

Usually, chemical processes take place in a solvent that affects the behaviour of the system. To incorporate solvent effects into the calculations water molecules can be treated explicitly (Sagui and Darden, 1999). Another approach are continuum solvent models such as the generalised Born model (Still et al., 1990)

or the analytic continuum electrostatics (Schaefer and Karplus, 1996; Schaefer et al., 1998) that are used to represent the electrostatic contribution to the free energy of solvation (Constanciel and Contreras, 1984).

The number of non-bonded interactions scales with the square of the number of atoms thus it is reasonable keeping the number of atoms at a minimum thereby reducing computational costs. One approach is the united atom model in which hydrogen atoms are subsumed with the heavy atoms to which these are bonded. To consider hydrogen bond and salt bridge formation, polar hydrogens are usually treated explicitly.

The Lennard-Jones potential drops rapidly with distance. Therefore simplifications such as a cut-off can be applied by which all pairwise interactions further apart from the cut-off edge are ignored. Long range electrostatic interactions usually are likewise treated with a cut-off or by a more sophisticated approach such as Ewald summation or particle-mesh ewald method (Ewald, 1921; Darden et al., 1993; Essmann et al., 1995).

2.2.1 Energy Minimisation

The potential energy of a system corresponds to a function of the coordinates thus a system with N atoms matches an energy function with $3N$ dimensions. Minimum points of the multi-dimensional energy surface represent stable states of the system and therefore are of interest in molecular modelling. The conformation with the lowest potential energy is referred to as global minimum whereas all other minimum points are denoted as local minima. To identify geometries constituting a local or even the global minimum of the energy surface, minimisation algorithms can be applied.

Several approaches for the localisation of minimum points have been developed that can be categorised into simpler, non-derivative methods comprising energy based methods such as the simplex algorithm and the more efficient derivative approaches using the first (*e.g.* steepest descent, conjugate gradient) and second derivative of the energy function (*e.g.* Newton Raphson and variants thereof). The first derivative of the energy function provides information of the gradient and the new direction for the next minimisation step. Additionally, the second derivative uses the curvature of the energy function to identify the search direction.

In general, minimisation techniques aim to locate the geometry containing the lowest potential energy by "walking" down an valley of energy surface. Ideally, the algorithm is capable to reach the bottom representing a minimum conformation. As the algorithms applied always search for minimum points on the energy surface, hills cannot be crossed thus only the valley containing the starting structure can be sampled.

2.2.2 Molecular Dynamics Simulations

Understanding the function of biological macromolecules requires knowledge of structure and dynamics. X-ray crystallography has gained insight into the assembly of protein and MD simulations have emerged as a powerful tool for studying protein dynamics (Karplus and Kuriyan, 2005). Beside other simulation methods, MD simulation techniques have been applied successfully for several tasks comprising conformational changes in proteins, enzyme catalysis and even protein folding (Karplus and McCammon, 2002).

Classical MD simulations explore the accessible conformational space on the energy landscape of a given molecule by addressing numerical solutions of Newton's equation of motion (F = m*a) on a molecular system. Integration at successive time steps yields a trajectory allowing to study time dependent properties of the system. The Verlet algorithm and modifications thereof are applied for integrating the equations of motion (Leach, 2001). From the first MD simulation of a small protein almost 30 years ago (McCammon et al., 1977) substantial progress in simulation algorithms (Mackerell Jr., 2004) and computer performance has been achieved that not only allowed the simulation of always larger systems (hundreds of thousands of atoms) but also longer duration increased from few picoseconds (McCammon et al., 1977) to microseconds (Duan and Kollman, 1998). Additionally, the description of the simulation environment changed from calculations in vacuum to more realistic models including explicit water molecules, counter ions and a more convenient treatment of the system boundaries and long range electrostatic interactions (Hansson et al., 2002). Nevertheless, MD simulations are approximative approaches exemplified by the usage of classical mechanics. Therefore, the quality of the data strongly depends on the system setup comprising the reliability of the model, the accuracy of the force field as well as the simulation software applied and last but not least, the users competence (van Gunsteren and Mark, 1998).

2.3 Molecular Interaction Fields

Beside covalent bonds also non-covalent interactions play an important role in biological processes comprising ligand and substrate binding as well as protein folding (Höltje et al., 2003). In order to study these interactions, molecular interaction fields (MIFs) can be calculated using several different programmes. The programme package GRID (Molecular Discovery Ltd., Pinner, UK) determines and visualises energetically favourable sites of interactions on molecules (Goodford, 1985). This can be used in the drug discovery process where GRID has been already applied successfully (von Itzstein et al., 1993). Moreover, GRID can be used as descriptors input for statistical programs such as CoMFA or GOLPE (Cramer et al., 1988; Pastor et al., 1997) and, additionally, GRID fields also serve as input for a docking program (GLUE).

In principle, the program places a grid around the molecule of interest. The energies of interactions between a user-defined chemical group (the "probe") and the molecule to be studied (the "target") are calculated at each vertex of the grid. Restricted to non-covalent interactions, the energy of interaction is described by vdW, electrostatic and hydrogen bond contributions (Equation 2.7). Different kinds of single and multi atom probes are available allowing to study sites of favourable interactions with water, ions or hydrophobic groups. The resulting energies can be displayed as three-dimensional contour surfaces, together with the structure of the target molecule.

$$E_{GRID} = \sum E_{LJ} + \sum E_{el} + \sum E_{HB} \qquad (2.7)$$

2.4 Molecular Docking

Docking procedures aim to find and evaluate possible binding modes of ligands in their receptor binding site in order to discover new potential hits or leads (Höltje et al., 2003). A database of compounds is screened against the target of interest, usually a protein. In the drug discovery process molecular docking approaches have been emerged as an integral part in structure-based drug design, lead optimisation and virtual high-throughput screening (HTS) (Klebe, 2000; Vangrevelinghe et al., 2003).

Numerous docking algorithms are available that can be categorised into genetic algorithms (Willett, 1995), approaches incrementally constructing the ligand in

the binding pocket (Leach and Kuntz, 1990; Rarey et al., 1996) and Monte Carlo methods, often used in conjunction with simulated annealing procedures (Goodsell and Olsen, 1990). Of those algorithms several have been implemented in popular docking programs such as DOCK (Kuntz et al., 1982), FlexX (Rarey et al., 1996), GOLD (Jones et al., 1997), AutoDock (Goodsell and Olsen, 1990) or Glide (Halgren et al., 2004).

Most docking programs consider only the flexibility of the ligand whereas the receptor is kept rigid. Apart from early attempts to include side chain flexibility into docking approaches (Leach, 1994), only few programs provide side chain or protein flexibility, respectively (FlexE and Slide) (Claussen et al., 2001; Zavodszky et al., 2002). A third will be available with the forthcoming new version of AutoDock.

2.4.1 Scoring

Docking algorithms generate a large number of potential binding modes of ligands at their binding pocket from which plausible solutions have to be selected. Scoring functions attempt to approximate the free energy of binding for a ligand-receptor complex thus selecting and ranking probable binding modes. Usually, scoring algorithms are an integral part of the docking procedure as the binding of ligands in their binding crevice is optimised according to the score. As the calculation of free energy of binding is computationally demanding and hence inefficient for docking approaches, faster and more approximative algorithms have been developed (Ajay and Murcko, 1995). Indeed, the majority of estimated scores acquired by scoring functions does not correlate well with experimentally determined ligand binding affinities (Wang et al., 2003a; Wang et al., 2004b). Although there are geometric decoys and false positives hits that may significantly decrease the hit rate (Graves et al., 2004), good scoring functions are able to identify the experimentally determined binding mode of a ligand-receptor complex in up to 75 % of a given test set (Wang et al., 2003a). However, developed to be applied for all kind of protein-ligand complexes, the significance of docking scores for a single protein is limited. A more reliable approach is the combination of two or more scoring functions (consensus scoring) which has been shown to improve the results (Charifson et al., 1999; Wang et al., 2003a). Typical representatives of consensus scoring functions are

CScore and X-Score (Clark et al., 2002; Wang et al., 2002). As X-Score comprises three related scoring functions that differ from each other only in the calculation of the hydrophobic effect term, CScore combines up to five different scoring algorithms including G-Score (Jones et al., 1997), D-Score (Kuntz et al., 1982), F-Score (Rarey et al., 1996), PMF-Score (Muegge and Martin, 1999) and ChemScore (Eldridge et al., 1997).

2.5   Virtual Screening

High throughput screening (HTS) is a well-established technology used in the drug discovery process by which hundreds of thousands of compounds can be tested within a short time. However, costs for biological testing and preclinical studies have reached 14 % of research and development (R&D) expenditure (Handen, 2002). In recent years computer based virtual screening (VS) methods have emerged as supporting techniques for rationalising drug discovery research by reducing the number of compounds to be tested by HTS to a limited number of candidates (Bajorath, 2002). VS concepts show a large diversity: Beside docking, also other 3D search techniques such as pharmacophore and QSAR models are applied in VS (Sheridan et al., 1989; Hopfinger et al., 1999; Wolber and Langer, 2005).

Besides the successful application in the discovery of several novel inhibitors for different types of proteins and even RNA (Perola et al., 2000; Doman et al., 2002; Filikov et al., 2000), VS techniques have been also applied effectively for NRs (Schapira et al., 2003a) resulting in novel agonists and antagonists for RAR and TR, respectively (Schapira et al., 2000; Schapira et al., 2003b).

Filtering methods are applied in order to exclude compounds with unwanted characteristics resulting in enrichment of libraries with preferred molecules. Computational approaches allow to recognise and remove substances carrying reactive or toxic moieties using dictionaries of undesired functional groups. Also the aqueous solubility can be predicted by various methods (Jorgensen and Duffy, 2002). A standard procedure is the application of Lipinski's "rule of five" as a measure for the "drug-likeness" (Lipinski et al., 2001). According to this rule, a compound is estimated to be poorly absorbed and unlikely permeates across cell membranes if two or more of the following criterions apply: A molecular weight larger than 500, a calculated logP value higher than 5 and the number of hydrogen bond donors and acceptors exceed 5 or 10, respectively. Other approaches identify compounds that can cross the blood-brain barrier and other ADMET parameters such as oral absorption and bioavailability (van der Waterbeemd et al., 1998; van der Waterbeemd and Jones, 2003).

2.6   Homology Model Evaluation

To ensure that the structure resulting from the modelling process is accurate homology models need to be analysed for their quality and reliability. Not only structural models can contain incorrect folded regions of smaller or larger extent but also X-ray structures were found to have major errors (Bränden and Jones, 1990).

Generally, methods evaluating the structural quality can be categorised into three different groups encompassing stereochemical accuracy, packing quality and folding reliability (Höltje et al., 2003). Stereochemical parameters comprise bond lengths, bond angles and torsion angles that are expected to differ only marginally from values revealed by X-ray crystallography. Also the planarity of peptide bonds and aromatic side chains need to be approximate to ideal values. The Ramachandran plot allows to inspect the distribution of the main chain torsion angles ϕ and ψ (Ramachandran et al., 1963). Usually, ϕ/ψ values are restricted to certain areas of the plot. Residues adopting unfavourable values are located in other regions and thus allow easy visual detection. Determination of stereochemical parameters can be accomplished by evaluation programmes such as PROCHECK and WHATCHECK (Laskowski et al., 1993; Hooft et al., 1996). PROCHECK comprises a suite of programmes that assess standard stereochemical parameters according to Morris et al. (*e.g.* ϕ/ψ distribution) and deviation from ideal bond lengths and angles according to Engh and Huber (Morris et al., 1992; Engh and Huber, 1991). Beside the distribution of ϕ and ψ torsion angles for which a Ramachandran plot is provided, also the correctness of side chain torsion angles χ_1 and χ_2 is analysed. Additionally, hydrogen-bond energies are calculated and secondary structure assignments are performed based on the method of Kabsch and Sander (Morris et al., 1992; Kabsch and Sander, 1983).

Not only the stereochemical parameters but also the overall three-dimen-sional fold of the model must meet various criteria. In order to determine whether a protein fold is correct or contains misfolded regions, several algorithms have been developed: The Profiles-3D method measures the compatibility of an amino acid sequence with a three-dimensional protein structure in a three-step process (Bowie et al., 1991; Lüthy et al., 1992). Information from the three-dimensional structure of the protein of interest is reduced to a one-dimensional

environment string containing information about the secondary structure as well as the position in the protein and the environment of each amino acid. Using a scoring matrix a 3D-profile is generated that is subsequently aligned with the corresponding amino acid sequence. The resulting profile provides information about the folding reliability for each amino acid. Additionally, the overall alignment score of the entire model is compared to that obtained for a correctly folded protein of similar size. The programme ProSa (protein structure analysis) represents another approach of quality assessment (Sippl, 1993). Adapted from knowledge based mean fields, the energy distribution in protein folds is analysed. Similar to the Profiles-3D approach resulting energies are plotted for each amino acid enabling recognition of misfolded regions.

Chapter 3

Generation of CAR Homology Models

3.1 Homology Modelling (I, II)

PXR shares the highest degree of sequence identity with CAR among all NRs. However, structural deviations from the common NR topology impede its usage as a single template. Therefore, a combined approach was applied in which VDR was used as second template for modelling regions of CAR where PXR differs from the canonical NR fold (Figure 3.1). VDR and PXR share a sequence identity of 36 % and 49 % with CAR, respectively. For either VDR or PXR several X-ray structures with reliable resolution are available (Rochel et al., 2000; Tocchini-Valentini et al., 2001; Watkins et al., 2001; Watkins et al., 2003b; Watkins et al., 2003a). Because of the significantly higher sequence identity with CAR, PXR qualified as main template for the modelling procedure.

The structural deviations of PXR from the canonical nuclear receptor fold are represented by an insertion domain between helix 1 and 3 consisting of an additional helical element and two β-strands (Watkins et al., 2001). Additionally, the conformation of H6 and H7 deviates from the common NR topology: H6 is completely unfolded whereas H7 is broken and kinked at about 90 degrees (see Fig. 1.4). The sequence of VDR also codes for a H1-H3 insert with comparable length. Predicted to be poorly structured and composed of a high percentage of negatively charged amino acids, the H1-H3 region affects receptor stability and thus interferes with crystallisation processes. In order to circumvent those

drawbacks, the H1-H3 region of the VDR has been manipulated by excision of a segment spanning 51 amino acids, finally permitting crystallisation (Rochel et al., 2000).
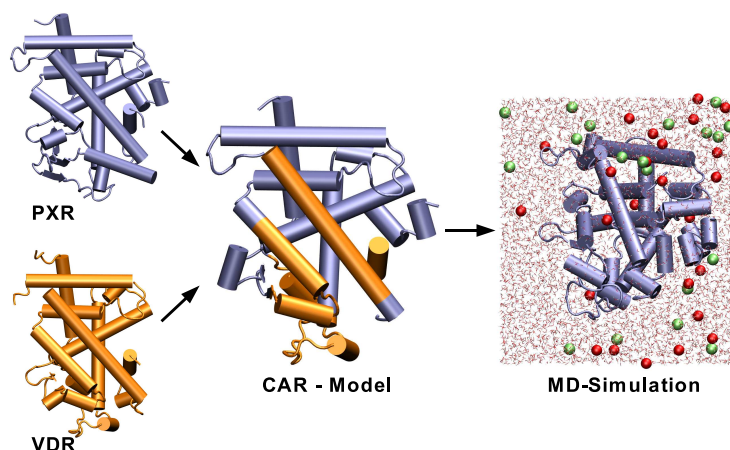
Figure 3.1: Generation of the CAR homology model. X-ray structures of PXR (blue) and VDR (orange) were used as templates in a combined approach. MD simulations were carried out in order to verify the model and to assess side chain flexibility of amino acids constituting the LBP.

Protein sequences for the target and template proteins were obtained from the Swiss-Prot/TrEMBL database (Boeckmann et al., 2003). The secondary structure of the CAR LBD was predicted by applying several different algorithms (PHD, PROF, PSIPRED, JPRED and SSPRO). The programmes were evaluated on their prediction accuracy using the sequence of human VDR and PXR for which X-ray crystal structures are available. The multiple sequence alignment of CAR with the related VDR and PXR receptors was carried out using CLUSTALW version 1.82 (Thompson et al., 1994).

Only one PXR X-ray structure exhibits a reliable resolution (PDB code 1NRL, resolution 2.0 Å) from which chain B was selected (Watkins et al., 2003a). For VDR several X-ray structures solved with high resolution are available that differ only marginally from each other regarding overall structural organisation and side chain conformations. Finally, the VDR X-ray structure complexed with the natural ligand vitamin D was selected (PDB code 1DB1, resolution 1.8 Å) as second template (Rochel et al., 2000).

3.1.1 CAR Model (I)

Coordinates for most amino acids were borrowed from the PXR structure (see Figure 3.1). To obtain the common NR fold, coordinates for H6 and H7 were taken from the VDR X-ray structure. Additionally, coordinates for helices H10/11 and the C-terminal H12 were adopted from the VDR template. Compared to PXR, VDR shares a significantly higher sequence identity with CAR within helix H10/11. Smaller amino acid side chains on H12 of CAR are believed to afford closer attachment of H12 to the LBD in CAR than observed in PXR (Dussault et al., 2002). This was also found in the VDR structure, thus coordinates for H12 were adopted from 1DB1. For the H1-H3 region (29 amino acids) the application of loop search approaches or *de novo* construction methods failed to determine a reliable conformation. Thus the protein backbone of the H1-H3 region was completely adopted from the VDR receptor. This was facilitated by the identical number of amino acids in the engineered segment of VDR and the corresponding region in CAR. The gap in the alignment between S337 and A338 was filled by carrying out a loop search procedure for amino acids A338 and M339 in INSIGHT II 2000 (Accelrys Inc., San Diego, USA).
Side chain conformations were assigned by the programme SCWRL version 2.95 (Bower et al., 1997; Dunbrack Jr., 1999). To ensure the accuracy of side chain prediction the programme was validated on the crystal structures of VDR and PXR. Special attention was paid on the reproduction of side chain conformations forming the LBP. For both receptors most side chains were assigned correctly, thus SCWRL qualified as applicable to the CAR model. In order to enhance the quality of side chain assignment, conformations of conserved amino acids were directly adopted from the template structures.

3.1.2 CAR/SRC-1 Model (I, II)

A second CAR model was generated including the NRID of the SRC-1 coactivator. Only few NR X-ray structures in complex with a coactivator peptide have been solved including also the PXR that has been crystallised with a SRC-1 peptide (Gampe Jr et al., 2000; Xu et al., 1999; Watkins et al., 2003a). To model CAR/SRC-1, coordinates for the coactivator (amino acids 682-696) were completely adopted from the crystal structure of PXR/SRC-1 (PDB code 1NRL, chain D).

3.1.3 CAR/NCoR Model (II)

From the available NR X-ray structures complexed with an antagonist, only the peroxisome proliferator-activated receptor α (PPARα, PDB code 1KKQ) has been co-crystallised with a corepressor NRID of SMRT (silencing mediator of retinoid and thyroid receptors) (Xu et al., 2002).

The beforehand generated CAR model and the X-ray structure 1KKQ were used as templates. Both structures were superimposed by fitting residues of helices H3, H4 and H10. Most amino acids for the CAR/NCoR model could be adopted from the CAR model whereas the PPARα crystal structure was used to model the C-terminal H11-H12 region (amino acids 452 to 463). Additionally, coordinates for the co-crystallised corepressor NRID of SMRT were completely adopted from 1KKQ into the model. In order to obtain the NCoR NRID, amino acids were changed where appropriate. Side chain conformations were assigned using SCWRL 2.95.

3.2 Model Refinement (I, II, III)

All CAR models were refined by minimisation procedures and MD simulation approaches. To mimic an aqueous environment, the models were placed in a solvent box. Water molecules were represented using the SPC model (Berendsen et al., 1981). Na$^+$ and Cl$^-$ ions were added to ensure the overall neutrality of the system and to simulate a physiological NaCl solution. The energy minimisation procedure was performed using the Steepest Descents algorithm within the GROMOS96 force field (Scott et al., 1999). 2000 steps of Steepest Descents minimisation were performed with a convergence criteria of 100 kJ mol^{-1} nm^{-1} The resulting minimised structure was used as input for subsequent MD simulations utilising the GROMOS96 force field. Periodic boundary conditions were applied to simulate a bulk fluid. Long-range electrostatic interactions were considered by applying the Particle-mesh Ewald method (Darden et al., 1993; Essmann et al., 1995). For the calculation of vdW interactions a cut-off of 0.9 nm was introduced. The system was kept at constant temperature of 310 K using a Berendsen thermostat with a coupling time of 0.1 ps (Berendsen et al., 1984). Also constant pressure was maintained by coupling to an external bath having a reference value of 10^5 Pa, with a coupling time of 1.0 ps and an isother-

mal compressibility of 4.5×10^{-10} Pa. The time step for the simulations was set to 1 fs. Bonds involving hydrogen atoms were constrained to their equilibrium length using the Lincs algorithm (Hess et al., 1997).

Depending on the model (CAR, CAR/SRC-1, CAR/NCoR) different kinds of equilibration procedures were applied. For CAR without a coactivator NRID four 100 ps equilibration runs with decreasing constraints on backbone atoms ($1000{-}100$ kJ mol^{-1}) were employed. The CAR/SRC-1 model was equilibrated using a MD simulation of 250 ps and having constraints of 1000 kJ mol^{-1} on backbone atoms which were excluded for the artificially folded H1-H3 region. Free MD simulations for CAR and CAR/SRC-1 were carried out for 2.0 and 2.25 ns, respectively.

The resulting trajectories were clustered using NMRCLUST by superimposition all frames on backbone atoms with subsequently grouping into clusters of similar conformation for amino acids facing H12 (Kelley et al., 1996). For each cluster a representative frame was selected and investigated further by assessing the stereochemical quality using PROCHECK, ProSa and Profiles-3D.

3.3 Molecular Docking (I, II, III)

The ligand docking of agonists and antagonists into the homology models CAR/SRC-1 and CAR/NCoR was performed using GOLD version 2.1 (Cambridge Crystallographic Data Centre, Cambridge, UK). GoldScore was chosen as scoring function. For each ligand three independent docking runs were performed each with a maximum allowed number of 10 poses. The ligands were automatically docked within a sphere of 20 Å radius from the side chain atom CD1 of amino acid L206.

All docking results were analysed in detail by visual inspection and grouped into clusters of similar conformations according to their RMSD value. The solution with the highest scoring value was selected for further examinations.

The binding mode of docked ligands was further examined by carrying out MD simulations for the CAR-ligand complexes using the simulation setup described for the CAR/SRC-1 model. GROMOS96 topologies for the ligands were generated manually.

Chapter 4

The Mechanism of Constitutive Activity (I)

4.1   Introduction

The metabolism of xenobiotics such as environmental pollutants, pesticides or drugs involves sequential steps of oxidation, mainly by cytochrome P450s (CYPs), and conjugation by various transferases into hydrophilic, water soluble derivatives that are easily excreted (Guengerich, 1989; Tukey and Strassburg, 2000; Borst and Elferink, 2002). An exposure to xenobiotics also causes an adaptive increase in the expression of metabolic enzymes, termed induction, resulting in a faster inactivation and elimination but sometimes increased toxic reactions or unwanted drug-drug interactions as shown for CYP3A4 (Okey, 1990; Willson and Kliewer, 2002). The constitutive androstane receptor (CAR) (Baes et al., 1994) and the pregnane X receptor (PXR) (Kliewer et al., 1998) are the main regulators of gene expression of metabolising enzymes in the liver and the intestine. Both CAR and PXR belong to the superfamily of nuclear hormone receptors. In humans, this family of ligand-activated transcription factors comprises 48 members that are involved in regulation of homeostasis, development, reproduction and metabolism (Laudet and Gronemeyer, 2001). All receptors share a common topology and are constituted of different functional domains. Typically, a nuclear receptor comprises a N-terminal domain of high variability, a conserved DNA binding domain (DBD) and a ligand binding domain (LBD) that is attached to the DBD via a linker region. Some receptors

show an additional C-terminal region of yet unknown function.

For a number of nuclear receptors the three-dimensional structure of the LBD has been resolved by X-ray crystallography. The structures revealed a general architecture for nuclear hormone receptors and how endogenous ligands and xenobiotics interact with the binding site. The common fold of the LBDs comprises a three-layered anti-parallel helix sandwich composed of 12 helices and a three-stranded beta-sheet (Moras and Gronemeyer, 1998). The ligand binding pocket is located between the outer layers of the helix sandwich and is mainly formed by hydrophobic amino acid residues. A prerequisite for the stimulated gene transcription is the formation of receptor dimers that subsequently bind to specific DNA sequences and the binding of a co-regulator (McKenna et al., 1999). For example, PXR and CAR form a heterodimer with the retinoid X receptor (RXR) enabling recognition of several DNA response elements (Honkakoski et al., 1998b). Co-Activator binding to the receptor dimer-DNA-complex prompts the recruitment of other factors. The resulting multicomplex is then responsible for the local decondensation of chromatin. Finally, gene expression is initiated via direct and indirect interaction with the basal transcription machinery (Ueda et al., 2002).

The LBD harbours a short helix (H12, also termed AF-2) within its C-terminus that is responsible for the activation of transcription. Agonists and antagonists are able to modulate the gene activation via induction of a conformational change of AF-2 which in turn dictates complex formation with either co-activators or co-repressors. In case of an agonist-bound receptor, AF-2 covers the ligand binding pocket like a lid and simultaneously takes part in the formation, together with helix H3 and H4, of a hydrophobic groove. This groove represents the binding site for co-activators (Nolte et al., 1998). When the receptor binds an antagonist, this groove undergoes a reorientation and excludes AF-2, thus enabling binding of co-repressors.

Unlike classical nuclear receptors, the CAR possesses strong transcriptional activity in the absence of any added ligand (Baes et al., 1994). The basal activity of the CAR can be enhanced by agonist binding. However, currently only a few agonists for the human CAR have been described such as CITCO (6-(4-chlorophenyl)imidazo[2,1-b][1,3]thiazole-5-carbaldehyde O-(3,4-dichloro- benzyl)oxime) (Maglich et al., 2003) and tri-(*p*-methylphenyl)-phosphate (TMPP)

(Honkakoski et al., 2004). Another CAR ligand, clotrimazole, has given contrasting results (Maglich et al., 2003; Toell et al., 2002). In our own studies clotrimazole acts as a strong activator (Mäkinen et al., 2002; Jyrkkärinne et al., 2003).

In this work homology models of the human CAR ligand binding domain alone and in complex with a co-activator were generated since a X-ray crystal structure has not been determined so far. The aim of the modelling study was to elucidate the mechanism of the constitutive activity as well as to analyse the binding mode of known agonists and their effect on the CAR LBD. Our models are based on the crystal structures of the related nuclear receptors PXR and VDR. Despite the common nuclear receptor fold, CAR, PXR and VDR also exhibit differences in their architecture, which have to be considered during model building (Watkins et al., 2001; Rochel et al., 2000). The interactions between CAR, co-activators and ligands were studied applying molecular dynamic simulations (MD). Based on these simulations, the essential features responsible for the constitutive activity of CAR were identified and compared with experimental mutagenesis data.

4.2 Results

4.2.1 Homology Modelling

In order to apply the approach of homology modeling, a suitable template structure must be found for the sequence-structure alignment. In case of the CAR, sequences of the closely related VDR and PXR receptors of which crystal structures are available, show about 40 to 50 % sequence identity within the LBD, respectively. To test the reliability of a homology modelling procedure using only a single template structure, we generated models of the PXR and VDR LBD, respectively and compared them with their X-ray structures. Thus, the VDR homology model is based on the PXR template structure, and vice versa. The sequence identity between PXR and VDR in the LBD is 37 %. In contrast to the common topology of nuclear receptors, both nuclear receptors contain an additional domain inserted between helix H1 and H3. In case of the PXR, this domain consists of a helical segment and two beta-strands. The insertion domain in the X-ray structure of the VDR contains two additional helices and an artificial loop segment. A further deviation from the common topology in PXR occurs in the region of helices H6 and H7. The superimposition of the minimised and equilibrated VDR and PXR models and their corresponding crystal structure revealed large RMSD values, indicating that the strategy based on only one template structure does not yield a reliable receptor model. Therefore, we used a combined approach for the generation of the CAR homology model.

In order to account for the structural deviations of PXR, the VDR structure was used as scaffold for the modelling of the segment between H1 and H3 and helices H6 and H7. Additionally, the proximal helix H10/11 and the AF-2 domain were generated based upon the VDR template. Figure 4.2*A* shows the complete sequence alignment as well as the parts taken from each template. In order to analyse the influence of co-activator binding a second model with the complexed co-activator peptide (named CAR/SRC-1) was built following the procedure described above. The coordinates for the co-activator were taken from the PXR X-ray structure (Watkins et al., 2003a) that had been crystallised with a 15 amino acid peptide from the steroid receptor co-activator 1 (SRC-1) carrying the binding motif for nuclear receptors. The MD simulations of

```
CAR    106        LSKEQEELIRTLLGAHTRHMGTMFEQFVQFRPP----------------
PXR    142       GLTEEQRMMIRELMDAQMKTFDTTFSHFKNFRLPGV|REEAAKWSQVRKDLC
VDR    120    LRPKLSEEQQRIIAILLDAHHKTYDPTYSDFCQFRPP----------------


CAR    139    -------------AHLFIHHQPLPTLAPVLPLVTHFADINTFMVLQVIKFTKDLP
PXR    208    SLKVSLQLRGEDGSVWNYKPPADSGGKEIFSLLPHMADMSTYMFKGIISFAKVIS
VDR    209    -------------VRVNDGGGSVTLELSQLSMLPHLADLVSYSIQKVIGFAKMIP


CAR    181    VFRSLPIEDQISLLKGAAVEICHIVLNTTFCLQTQNFLCG--PLRYTIEDGARVG
PXR    263    YFRDLPIEDQISLLKGAAFELCQLRFNTVFNAETGTWECG--RLSYCLEDTAG-G
VDR    250    GFRDLTSEDQIVLLKSSAIEVIMLRSNESFTMDDMSWTCGNQDYKYRVSDVTKAG


CAR    234    FQVEFLELLFHFHGTLRKLQLQEPEYVLLAAMALFSPDRPGVTQRDEIDQLQEEM
PXR    315    FQQLLLLEPMLKFHYMLKKLQLHEEEYVLMQAISLFSPDRPGVLQHRVVDQLQEQF
VDR    305    HSLELIEPLIKFQVGLKKLNLHEEEHVLLMAICIVSPDRPGVQDAALIEAIQDRL


CAR    289    ALTLQSYIKGQQRRPRDRFLYAKLLGLLAELRSINEAYGYQIQHI---QGLS-AM
PXR    370    AITLKSYIECNRPQPAHRFLFLKIMAMLTELRSINAQHTQRLLRI---QDIHPFA
VDR    360    SNTLQTYIRCRHPPP|--LLYAKMIQKLADLRSLNEEHSKQYRCLSFQPECSMKL


CAR    340    MPLLQEICS
PXR    422    TPLMQELFGI
VDR    412    TPLVLEVFG
```

Figure 4.1: Sequence alignment of the two template structures PXR and VDR and the target sequence of CAR. Vertical lines indicate missing segments. Residues within the CAR sequence are coloured depending on the origin of the structural information used (PXR: red, VDR: green).

both homology models revealed that the overall fold of the nuclear receptors remained stable. The RMSD did not exceed a value of 0.2 nm within the backbone region. This might be due to the compact architecture of the three-layered helix sandwich that allows only limited motions of the individual domains.

The Ramachandran plots for CAR and CAR/SRC-1 models assessed 87.4 % and 86.4 % of the phi-psi torsion angles as being within the favourable region, respectively. The Profiles-3D scores for CAR and CAR/SRC-1 models, 98 and 116, are rather or very close to that expected for a high quality model of corresponding size (110 and 117, respectively). To account for the artificial fold taken from VDR and the ensuing additional helix, these segments were not constrained during the MD equilibration. Consistent with secondary structure predictions, this led to an unfolding of the helical segment resulting in a long disordered loop showing significant changes in RMSD (Fig. 4.3). In addition, increased
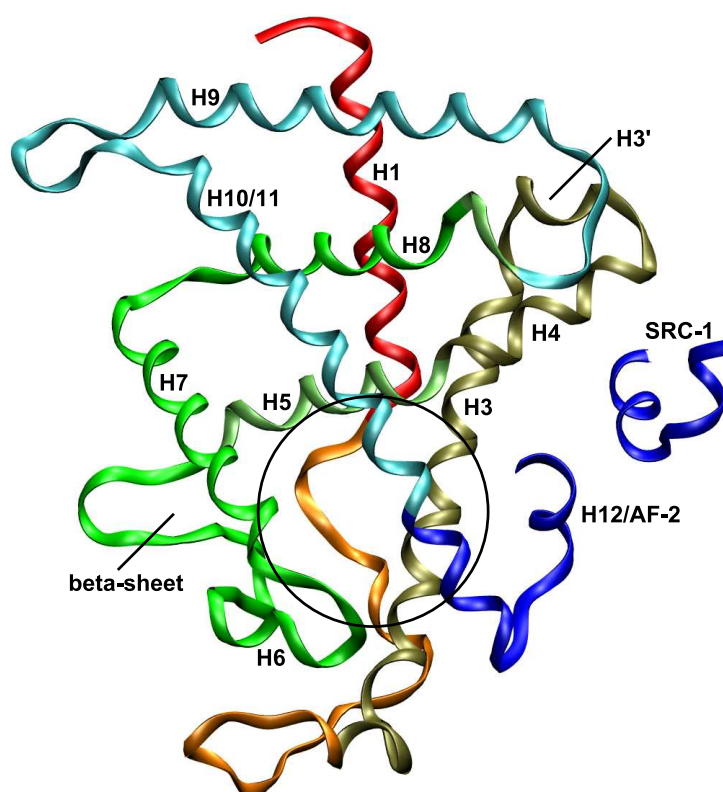
Figure 4.2: Homology model of CAR/SRC-1. Helices are numbered according to the common fold of nuclear receptors. The circle indicates the position of the ligand binding pocket.

flexibility was observed for the loop connecting H9 and H10/11. However, this loop is located at the LBD surface far away from the ligand binding pocket and the AF-2 domain (Figs. 4.2 and 4.4).

4.2.2   Constitutive Activity

In both models the AF-2 domain remained closely attached to the LBD during the entire MD simulation (Fig. 4.4). It is not expected that a large movement of the AF-2 domain occurs at the time range of the performed MD simulation, however the high stability of the LBD-AF-2 interaction is in close agreement with the experimental data (Andersin et al., 2003; Frank et al., 2004). In the CAR model, strong interactions can be observed between AF-2 and LBD which could be reproduced in several MD simulations using slightly different starting geometries of the homology model. The aliphatic AF-2 amino acids Leu343 and Ile346 contact with hydrophobic or aromatic residues of the LBD (Val199,
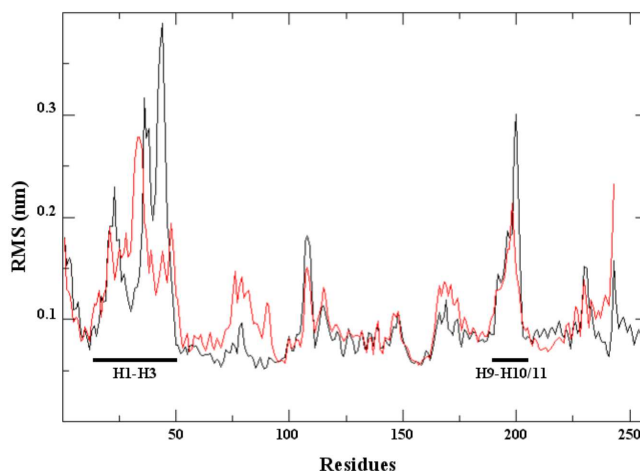
Figure 4.3: Analysis of the MD simulation. RMS-fluctuation within the CAR LBD backbone region during MD simulations (color codes: CAR, red; CAR/SRC-1, black). Loops spanning from H1 to H3 (residues 11-50) and from H9-H10/11 (residues 190-203) show high flexibility in both models.

Tyr326, Ile330; see Fig. 4.5*A*). The importance of these van der Waals interactions have been also detected by applying the programme GRID. Using the hydrophobic methyl probe within GRID we inspected the interaction possibilities between the LBD and the AF-2 domain. For this purpose we generated an AF-2-truncated CAR model and calculated the GRID interaction fields. The calculated contour maps were then viewed superimposed on the structure of the complete CAR (Fig. 4.5*B*). Two main interaction regions were detected at the LBD-AF-2 interface region by GRID. The location and size of the calculated contour maps is in close agreement with the position of Leu343 and Ile346 from the AF-2 domain.

Tyr326 is surrounded by a cluster of aromatic or hydrophobic residues (Val199, His203, Phe234, Phe238, Ile330) that seem to fix the position of Tyr326 side chain. During the MD simulation, a transient hydrogen bond between Tyr326 and Asn165 is formed. Phe161 points into the ligand binding pocket and interacts with Tyr224. Each of the described interactions between AF-2 and the LBD is also observed in the CAR/SRC-1 model. However, the AF-2 domain is positioned closer to the LBD (Figs. 4.4 and 4.5*A*) allowing contacts between Ile346 and Tyr326 on H10/11. Tyr326 moved towards Leu343 resulting in stronger van
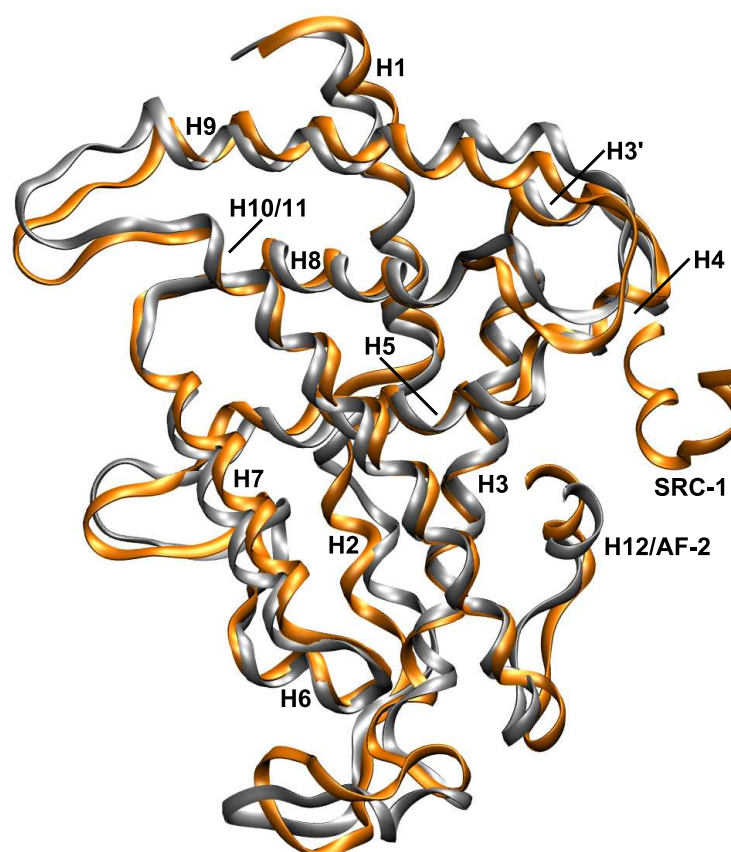
Figure 4.4: Comparison of CAR with CAR/SRC-1. Superposition of representative frames from the MD simulation (color codes: CAR: grey; CAR/SRC-1: orange). Structural differences between both models are observed mainly in loop regions. In the model with bound SRC-1 the AF-2 helix is slightly oriented.

der Waals interactions (Table 4.1).

The ligand binding pocket of the CAR is constituted by the helices H3, H5 to H7, H10/11 and the beta-sheet connecting helix H5 and H6. The AF-2 domain assumes the active conformation, forming a lid over the binding cavity which is significantly smaller (630 Å^3 before and 480 Å^3 after MD simulation) compared to the binding cavity of the PXR receptor (1294 Å^3). As observed for PXR and VDR, the CAR ligand binding pocket is mainly composed of hydrophobic residues with only a few polar residues contributing.

The hydrogen bond between Tyr326 and Asn165 appeared to be more stable as indicated from distance plots (data not shown), resulting in a permanent interaction between H3 and H10/11 in the presence of SRC-1 peptide. Phe161 was reoriented towards the interface between LBD and AF-2 where it forms
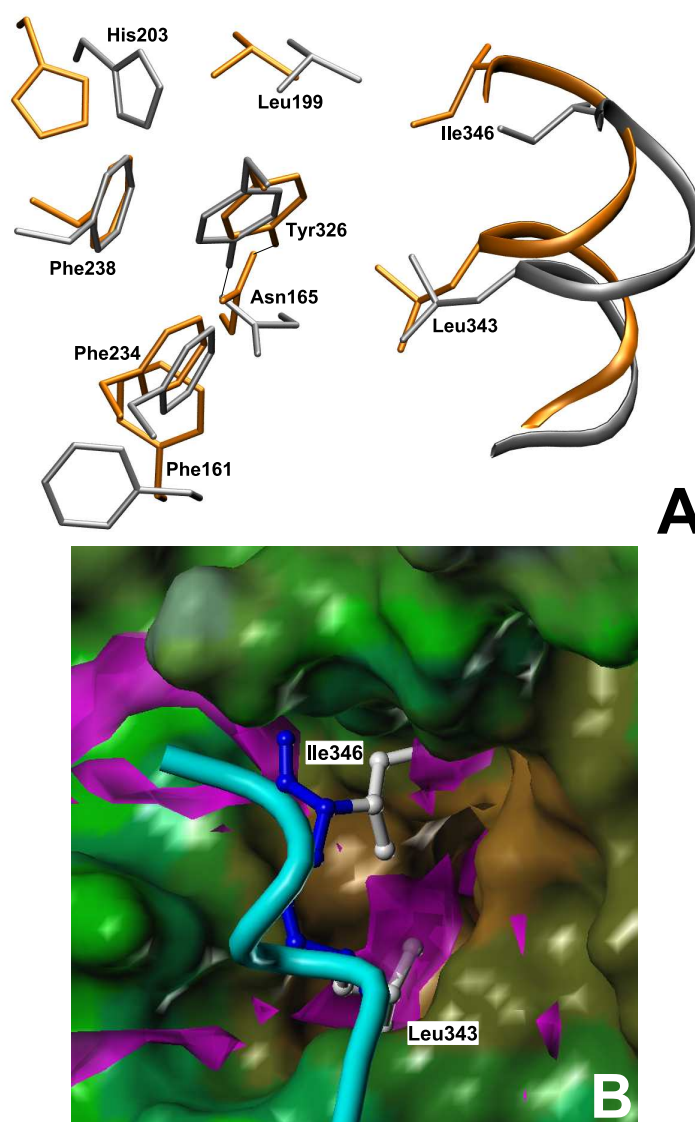
Figure 4.5: The constitutive activity. (**A**) Superposition of CAR (grey) and CAR/SRC-1 (orange). In both models AF-s (tube representation) interacts via Leu343 with Tyr326 and Ile330 (not shown) located on H10/11 and via Ile346 with Val199 on H4. Various surrounding amino acids stabilise the position of Tyr326 (Val199, His203, Phe234, Phe238 and Ile330). Additionally a hydrogen bond is formed between Asn165 and Tyr326 which has been found to be more stable upon SRC-1 binding. (**B**) Favourable regions of interactions between the GRID methyl probe and the AF-2 truncated LBD (coloured magenta, contour level -2.5 kcal/mol). Only the MOLCAD surface of the LBD is shown, coloured according to the lipophilic potential (blue polar, brown lipophilic). The position of the two hydrophobic residues Leu343 and Ile346 from the AF-2 helix (coloured cyan) is in close agreement with the obtained GRID results.

contacts to LBD residues (Asn165, Phe234, Tyr326) as well as AF-2 (Leu343) and the preceding loop (Met339).

	CAR	CAR/SRC-1			Mut.	Mut.
		Empty	Clotr.	TMPP	Phe238Ala	Phe243Ala
Tyr326-Leu343	5.2	4.7	3.9	4.1	9.1	7.9
Y326-I346	7.3	4.5	4.8	4.6	8.6	8.9

Table 4.1: Observed distances (in Å) between Tyr326 from the LBD and Leu343/Ile346 from the AF-2 domain in the individual models (see text for further explanation).

4.2.3 Co-Activator Binding

Activation of nuclear receptors requires the binding of co-activators such as SRC-1. The known crystallographic and experimental data reveal that nuclear receptors possess specific interaction patterns for the binding of co-activators (Darimont et al., 1998; Feng et al., 1998) which seem to be essential for their function. Thus reproducing the interactions between LBD and co-activator seems to be a prerequisite for a reliable homology model. In crystal structures the interaction domain of the co-activator adopts an alpha helical form and contains the LxxLL motif which interacts with hydrophobic residues located within a groove formed by the helices H3, H4 and AF-2. In the CAR/SRC-1 model, the hydrophobic groove is formed by 11 residues (shown in Fig. 4.6*A*). Leu342 (AF-2) is fixed in a hydrophobic pocket formed by LBD and SRC-1 (Fig. 4.6*B*). The helix dipole of the SRC-1 peptide is known to be stabilised by two conserved amino acids interacting with its N- and C-terminal residues that form the so-called "charge-clamp" (Nolte et al., 1998; Darimont et al., 1998). This conserved interaction pattern could also be reproduced for the CAR/SRC-1 model where these residues match Lys177 (H3) and Glu345 (AF-2) (Fig. 4.6*C*). Lys177 forms a hydrogen bond with the backbone carbonyl group of Leu693 of SRC-1, whereas Glu345 interacts with the backbone amide groups of Ile689 and Leu690. Additional hydrogen bonds are also formed between His687 of SRC-1 and Lys195 (H4) and between His687 and Glu345. Furthermore, a hydrogen bond between a histidine residue on the co-activator and the conserved lysine

(Lys177) is present in other nuclear receptors (Watkins et al., 2003a; Xu et al., 1999; Gampe Jr et al., 2000). This interaction is also observed in the homology model and persisted during the entire MD simulation.
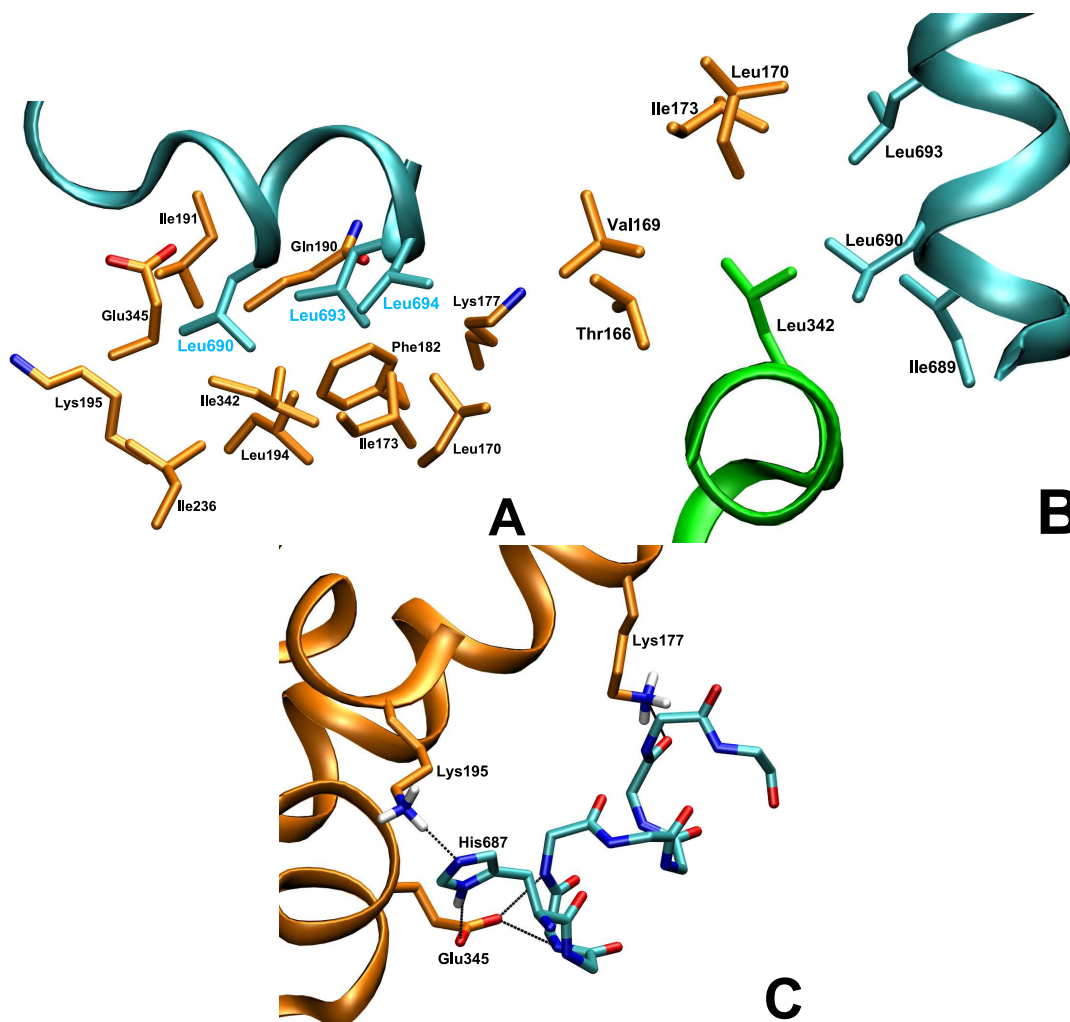
Figure 4.6: Interactions between LBD and SRC-1 (**A**) Several residues from H3, H3', H4 and AF-2 (carbon atoms in orange) form a hydrophobic groove to which SRC-1 (cyan) can bind. Leucines from the LxxLL motif on SRC-1 are shown explicitly. (**B**) Binding of SRC-1 to the LBD fixes Leu342 from AF-2 (green) in a hydrophobic pocket formed by several amino acids from LBD (orange) and SRC-1 (cyan). (**C**) The helix dipole of SRC-1 is stabilised by Lys177 (H3) and Glu345 (AF-2) located on the LBD (orange) forming the so-called "charge clamp". Lys177 interacts via a hydrogen bond with the backbone carbonyl group of Leu693, whereas Glu345 interacts with the backbone nitrogen atoms of Ile689 and Leu690. Also hydrogen bonds are formed between His687 (SRC-1) and Lys195 (H4) as well as His687 and Glu345. Parts of the LBD have been removed to show the interactions more clearly.

4.2.4 Docking Studies

In order to analyse the activation mechanism of CAR upon agonist binding, the CAR/SRC-1 model was taken for docking studies using program GOLD (CCDC, Cambridge, UK). The docking runs yielded 19 favoured poses for clotrimazole grouped into two clusters that differed only slightly from each other. For TMPP 18 poses were found which were grouped into two clusters. The docking pose with the highest fitness score was selected for each ligand and investigated further by MD simulations. Clotrimazole is bound deeply in the ligand binding pocket (Fig. 4.7*A*). No direct contact between clotrimazole and AF-2 domain is observed. The aromatic side chains of the ligand mainly interact with several aromatic residues of the binding pocket (Phe112, Phe161, Phe234 and Tyr326). During the 2.5 ns MD simulation the position of clotrimazole did not change significantly. The distance between Tyr326 and Leu343 (AF-2) was decreased (Table 4.1). Phe161 moved towards the interface between LBD and AF-2 establishing van der Waals interactions with Met339. The ligand binding pocket was widened by a minor movement of the β4-strand in combination with small side chain reorientations: Residues His160 and Tyr224 maintained their hydrogen bond but were pushed towards the outside of the LBD whereas Phe243 is now pointing away from the ligand binding pocket. These events increased the size of the pocket resulting in a volume of 750Å3. TMPP shows a different binding mode (Fig. 4.7*B*). It is located much closer to the interface between LBD and AF-2 than clotrimazole. One of the methylphenyl groups directly interacts with Leu343 and Ile346 from AF-2. The two remaining methylphenyl groups point into the ligand binding pocket, both interacting with Phe161 and pushing it deeper into the pocket. As a result, the distance to Met339 is increased as compared to the ligand-free receptor. Additional van der Waals interactions of the methylphenyl groups with Phe234 and Tyr326 could be observed. The phosphate group formed strong hydrogen bonds to both Asn165 and Tyr326 that remained stable during the entire MD simulation. Due to steric effects, TMPP provoked a reorientation of Val169 which resulted in an interaction with Ile346 from AF-2 (Fig. 4.7*B*). In agreement with the simulation of clotrimazole, the distance between Tyr326 and Leu343 (AF-2) was also decreased in the simulation of TMPP (Table 4.1). The size of the ligand binding pocket increased to a final volume of 560Å^3.
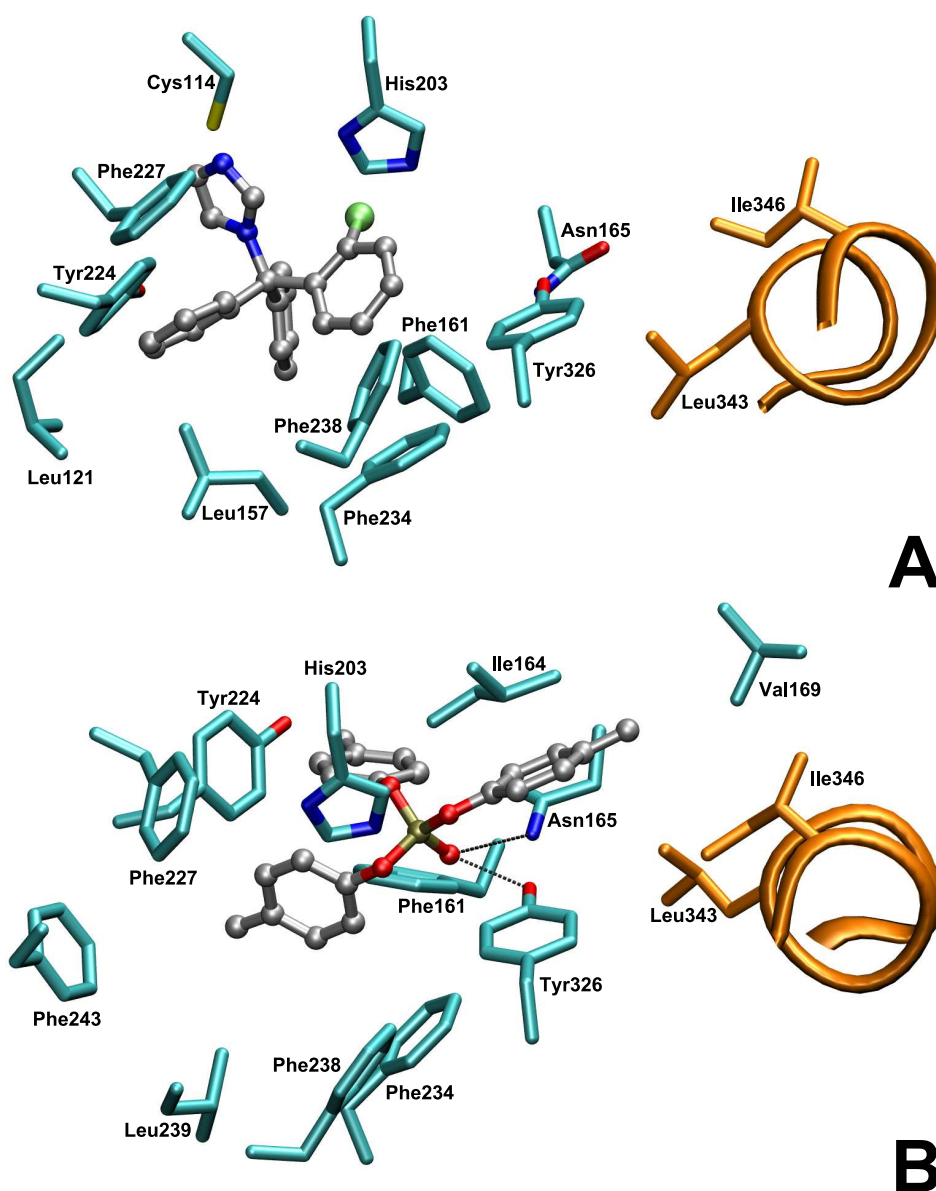
Figure 4.7: Binding modes of CAR activators: (**A**) Clotrimazole (carbon atoms in grey) is positioned deeply in the LBD without any contact to the AF-2 helix (orange). Main interactions are observed with aromatic amino acids (Phe112, Phe161, Phe234 and Tyr326). Phe161 has moved towards the interface LBD/AF-2 enabling interaction with Met339 from the loop connecting H10/11 and AF-2. (**B**) TMPP (carbon atoms in grey) is located close to the AF-2/LBD interface establishing van der Waals interactions with Leu343 and Ile346 from AF-2. Additionally, hydrogen bonds with residues Asn165 and Tyr326 are observed. Phe161 is surrounded by two methylphenyl groups resulting in reorientation into the ligand binding pocket.

4.2.5 Mutagenesis Studies

Mutagenesis studies have shown that the single point mutation Phe238Ala results in a loss of CAR basal activity (Fig. 4.8*A*) indicating a prominent role in the mechanism of constitutive activity. Also the Phe243Ala mutation reduces the basal activity significantly. In order to prove the consistency between the experimental and theoretical studies, both mutants were modeled and analysed by MD simulations. During the simulations, the Phe238Ala mutation provoked rotation of the side chain of Tyr326 (around χ_1) resulting in a more buried conformation of Tyr326 within the LBD (Fig. 4.8*B*). As a result, the interactions of Tyr326 with Leu343 and Ile346 were disrupted (Table 4.1). The hydrogen bond between Asn165 and Tyr326 was also lost. Phe161 shows weak interactions with Leu343 and is in close contact to Met339.

The mutation Phe243Ala forced H7 towards the beta-sheet by more than 3Å, and H10/11 was pushed in the same direction. These movements resulted in a pronounced reorientation of several residues in the ligand binding pocket (Fig. 4.8*C*). Tyr326 now points more deeply in the ligand binding cavity and contact with AF-2 is lost (Table 4.1). The conformation of Phe238 changed resulting in loss of any stabilising effects on Tyr326. The hydrogen bonds between Tyr326 and Asn165 as well as between His160 and Tyr224 were destroyed, provoking a rotation of the side chain of Tyr224 side chain into the ligand binding pocket.
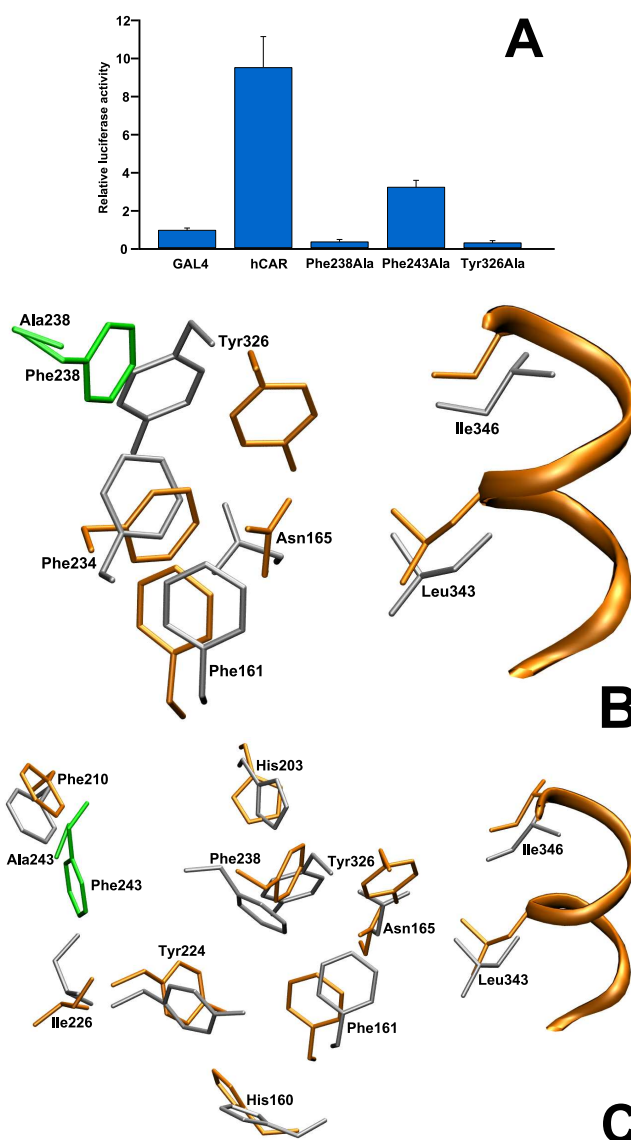
Figure 4.8: Site-directed Mutagenesis: (**A**) Relative luciferase activity measured for CAR wildtype and three independent point mutations. Wildtype CAR is constitutively active whereas mutation of Phe238, Phe243 and Tyr326 results in decrease or loss of basal activity (**B**) Mutation of Phe238 to alanine (both in green): During the MD simulations Tyr326 changed its position pointing now into the ligand binding pocket. Contacts between Tyr326 and AF-2 are disrupted and the hydrogen bond with Asn165 is lost. (**C**) Mutation of Phe243 to alanine (both in green) results in side chain reorientation of several residues. Due to displacement of H10/11 Tyr326 now points into the ligand binding pocket. The hydrogen bond between Tyr224 and His160 is abolished resulting in rotation of Tyr224 into the ligand binding pocket.

4.3 Discussion

The AF-2 domain located at the C-terminal end is thought to be responsible for the activation of nuclear receptors. Supposedly, agonists and antagonists induce conformational changes of AF-2 that subsequently results in formation of a complex with co-activators or co-repressors, respectively. In case of the agonist-occupied binding site, AF-2 covers the ligand binding pocket like a lid. In contrast to other nuclear receptors, CAR has substantial constitutive activity, but its structural basis is not yet clear. Using MD simulations of a homology model of CAR, we found evidence for a potential activation mechanism based on specific van der Waals interactions between the LBD and the AF-2 domain which contribute to AF-2 remaining anchored to the LBD. Of several unique residues involved in these interactions, Tyr326 has been found to be of special importance. A cluster of hydrophobic and aromatic residues around Tyr326 fix the side chain to enable its van der Waals interactions with AF-2 residues. This has also been demonstrated by the calculated molecular interaction fields with GRID (Fig. 4.5*B*). Phe238, which is located in close proximity, seems to prevent rotation of the Tyr326 side chain around its χ_1 angle and thus blocking Tyr326 from the ligand binding pocket. Val199, His203, Phe234 and Ile330 are positioned above and below the plane of the Tyr326 side chain thereby preventing rotation around χ_2. In addition the orientation of Tyr326 is stabilised via a hydrogen bond with Asn165. Assisted by its surrounding residues, Tyr326 emerges as a central interaction partner for AF-2 and keeping it closely attached to the LBD. This interaction pattern seems to be unique among related nuclear receptors and thus provides a convincing explanation for the constitutive activity of CAR. The critical role of Tyr326 in constitutive activity was borne out experimentally: the Tyr326Ala mutant had lost its basal activity (Fig. 4.8*A*). Van der Waals interactions between LBD and AF-2 seem to be a common feature for constitutive activity as seen in crystal structures of murine LRH-1 (Liver Receptor Homologue 1) and human ERR3 (Estrogen Related Receptor 3) (Greschik et al., 2002; Sablin et al., 2003).

Further results are supported by data from site-directed mutagenesis: Replacement of Phe238 for alanine reduced the basal activity significantly, and during the MD simulation, the position of Tyr326 changed and its van der Waals contacts with AF-2 were disrupted. The weaker interaction between LBD and

AF-2 might lead to a reorientation of AF-2, disruption of the SRC-1 binding site, and finally to the reduction of basal activity. The CAR activity can also be modulated by residues more distant from the LBD/AF-2 interface: the mutation Phe243Ala resulted in a modification of the overall shape of the ligand binding pocket that caused a displacement of H10/11 and a subsequent reorientation of Tyr326. Thus, we could explain the experimentally observed loss of basal activity on a structural level with models of mutated receptors. This suggests that the current models and methods are applicable to generation of further hypotheses and experimental testing to elucidate CAR ligand specificity and mechanism of activation.

It has recently been reported that replacement of either Leu343 or Ile330 by Ala reduces the basal activity significantly (Frank et al., 2004). This is in accordance with our observations where Leu343 is the only amino acid from AF-2 that permanently shows van der Waals interactions with the LBD. Upon mutation to alanine, the contact to Tyr326 would be disrupted and the remaining interactions between Ile346 and LBD might not be sufficient to keep AF-2 anchored to the active conformation. Mutation of Ile330 to alanine would not only reduce the hydrophobic surface area and the number of potential interaction partners for Leu343 but also destabilise the position of Tyr326, because Ile330 is one of the amino acids that restrain the side chain of Tyr326. This stabilisation might be reduced by its mutation to alanine resulting in an increased flexibility for Tyr326 which might have implications on the stability of interactions with AF-2.

In the agonist-bound structures of PXR and VDR, van der Waals interactions between the ligand and the AF-2 domain have been observed (Figs. 4.9*A&B*) (Rochel et al., 2000; Watkins et al., 2003a). These are believed to maintain AF-2 attached to the LBD enabling co-activator binding when a ligand is present. We have shown that the single residue Tyr326 in CAR makes van der Waals contacts to AF-2. Tyr326 thereby accommodates a corresponding position relative to AF-2 as the ligands do in PXR and VDR structures (Fig. 4.9*C*). We therefore infer that the constitutive activity of CAR may result from a "molecular mimicry" of a bound agonist. Co-Activator binding is essential for the ability to activate transcription. As a prerequisite, the AF-2 domain has to adopt a position that, together with residues of the LBD, allows formation of this hydrophobic groove. We could show that the AF-2 domain in CAR is able to form this hydrophobic groove even in absence of any bound agonist.
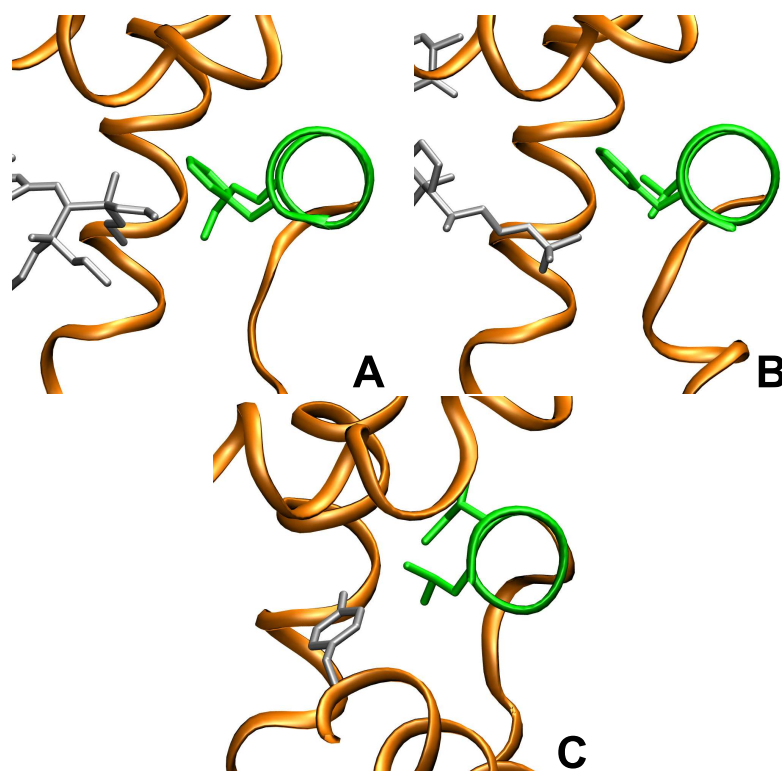
Figure 4.9: "Molecular mimicry": Interaction between the AF-2 (green) and the LBD domain (orange) for PXR (**A**), VDR (**B**) and CAR (**C**). Ligands for PXR and VDR as well as the corresponding amino acid in CAR, Tyr326, are coloured in grey. PXR and VDR exhibit van der Waals interactions between the bound ligand and residues located on AF-2 (green). In CAR a bound ligand is mimicked by Tyr326 that might lead to constitutive activity of CAR.

Due to limited flexibility of the AF-2 domain the hydrophobic groove does not show a strictly defined geometry in absence of the SRC-1 peptide. Thus, binding of SRC-1 seems to induce limited alterations on parts of the hydrophobic groove; especially the AF-2 domain is slightly reoriented. As a result, specific interactions between the LxxLL motif of SRC-1 and LBD/AF-2 occur which are conserved among nuclear receptors (Darimont et al., 1998; Feng et al., 1998). Existing van der Waals interactions between the LBD and AF-2 are reinforced and additional contacts are introduced which might contribute to keep AF-2 in the new position.

Additionally, we observed a novel hydrogen bond between the LBD (Gln331) and the ultimate C-terminal residue Ser348 in the presence of SRC-1. The influence on basal activity of human CAR remains to be elucidated. Localisation of

AF-2 in closer proximity to the LBD results in a small rotation of Tyr326 towards AF-2 which might be responsible for rotation of Phe161 towards the interface between LBD and AF2. Based on our observations we propose a cooperative binding mode for SRC-1. The hydrophobic groove and additional residues involved in SRC-1 binding (*e.g.* Lys177, Lys195) exhibit considerable flexibility resulting in a weak initial binding of SRC-1. After reorientation of several residues within the SRC-1 binding site (*e.g.* Phe161, Tyr326), specific interactions with AF-2 and LBD are established. The AF-2 domain is stabilised in this new position through novel interactions with the LBD that further enhances SRC-1 binding. Thermal denaturation experiments performed with PXR have shown that the overall stability of the LBD increases upon co-activator binding (Watkins et al., 2003a). Our results suggest, that this might be due to an enhanced interaction between LBD and AF-2.

Applying docking procedures binding modes for structurally diverse ligands in the CAR binding site were obtained. For both clotrimazole and TMPP one favourable binding mode was proposed, respectively. A similar conformation of clotrimazole in a CAR model has been described in a previous study (Xiao et al., 2002). Both agonists used in this study have shown to interact with amino acids surrounding Tyr326 leading to further stabilisation of the tyrosine side chain. As a result the distance between Tyr326 and Leu343 is decreased compared to the empty receptor (Table 4.1). This gives reason to propose an increase of the van der Waals interactions between LBD and AF-2 which might keep AF-2 in its new position facilitating SRC-1 binding and further leading to increased CAR activity. Although the adopted binding modes of clotrimazole and TMPP are quite different, both ligands induce comparable structural changes that result in a further increase of CAR activity. Based on the observations for clotrimazole and TMPP binding a general mechanism of the action of agonists could be proposed that is based on a further stabilisation of favourable side chain conformations as previously described for the activation mechanism.

The function of CAR as a xenosensor requires recognition of a diverse set of ligands. Thus the ability of the binding site to adapt to a variety of ligands is essential. Upon agonist binding, the ligand binding pocket is able to expand up to 80 %. Increasing the size of the cavity has been also reported for PXR complexed with hyperforin (Watkins et al., 2003b). During our simulations we

observed two regions of moderate flexibility upon ligand binding. In contrast to PXR, structural adaptions took place within parts of the beta-sheet (β4-strand) and a residue located at the interface LBD/AF-2 (Val169). The smaller ligand spectrum of CAR compared to PXR might thus be due to the significantly smaller ligand binding pocket and the limited flexibility of regions located therein.

Several homology models of the human and mouse CAR LBD have been generated up to now which have given first insights into ligand binding and interactions between LBD and AF-2 (Xiao et al., 2002; Dussault et al., 2002; Jacobs et al., 2003). These models differ in some respect from the presented one. The ligand binding pocket is much larger (1150-1170 Å^3 vs 480 Å^3) than observed in our model. This might be due to the selection of PXR as only template for the model generation and different orientation of side chains within the binding pocket (Xiao et al., 2002; Dussault et al., 2002). In contrast to previous studies we performed MD simulations in order to account for the dynamic behaviour of this complex system. The simulations gave insight into the formation of several new interactions that have been found to be critical for CAR activity and that have been supported by the described mutagenesis studies.

4.4 Conclusions

Our simulations have given new insights into the molecular basis of the constitutive activity of CAR. We proposed an activation mechanism based on specific van der Waals interactions between residues from the LBD and AF-2 domain. Functional consequences of LBD mutations could be reproduced or at least explained on a structural level. Nevertheless, information on the mechanism of action of antagonists is still lacking. Modelling of CAR in complex with co-repressors might reveal the structural basis of CAR inactivation and will create a basis for the development of specific CAR inhibitors.

Chapter 5

The Ligand Specificity of Human CAR (II)

5.1   Introduction

Nuclear receptors (NRs) are ligand-inducible transcription factors that govern many physiological processes. Receptors for steroid hormones, thyroid hormone (TR) and retinoic acid (RAR) are critical for cellular differentiation and development while many "adopted orphan" receptors are metabolic sensors regulating cholesterol, fatty acid, and glucose turnover (Mangelsdorf et al., 1995; Chawla et al., 2001; Aranda and Pascual, 2001). The most important structural component found in NRs is the ligand-binding domain (LBD). The NRs display great selectivity for ligands due to distinct sizes, contours and lipophilicities of their ligand-binding pockets (LBP) that are caused by variations of the lining residues. Ligand binding induces conformational changes in the LBD. These changes allow the NR to recruit co-activators needed for histone acetylation, or co-repressors required for histone deacetylation at the NR target gene (Renaud and Moras, 2000; Weatherman et al., 1999). In most NRs studied, the position of the helix 12 determines which type of co-regulator is able to bind to the LBD. After binding of an agonist, helix 12 is stabilised in an orientation where it forms a hydrophobic co-activator surface together with helices 3 and 4, and the conserved "charge clamp" residues glutamate (helix 12) and lysine (helix 3) are essential for co-activator binding and ligand-induced NR activation. This co-activator surface overlaps with the binding site for co-repressors that is

composed of helices 3 and 4. Antagonist binding stabilises another orientation of helix 12 which prevents co-activator recruitment and promotes co-repressor binding (Steinmetz et al., 2001).

The human constitutive androstane receptor (CAR, NR1I3) belongs to the nuclear receptor subfamily 1I together with pregnane X (PXR, NR1I2) and vitamin D receptors (VDR, NR1I1) (Robinson-Rechavi et al., 2002; Moore et al., 2002). CAR and PXR share some ligands including a variety of xenobiotics, steroid hormones and prescription drugs. These two receptors are important activators of overlapping sets of genes that code for cytochrome P450, conjugative enzymes and transport proteins (Ueda et al., 2002; Honkakoski et al., 1998b; Wei et al., 2000; Willson and Kliewer, 2002; Honkakoski et al., 2003; Maglich et al., 2002; Rosenfeld et al., 2003), and their induction can lead to adverse drug effects or harmful drug-drug interactions (Rodrigues and Lin, 2001). CAR also protects against the toxicity of endogenous cholestatic bile acids and bilirubin by enhancing their metabolism (Guo et al., 2003; Huang et al., 2003). Important hepatic enzymes that are involved in fatty acid oxidation and energy metabolism such as squalene epoxidase and phosphoenolpyruvate carboxykinase 1 are repressed by CAR (Ueda et al., 2002; Yamamoto et al., 2003).

Unlike most NRs, human and mouse CAR have high constitutive activity in the absence of any added ligand (Baes et al., 1994; Forman et al., 1998). This seems to be due to the constitutive interaction of CAR with several co-activators including SRC-1, GRIP-1, PGC-1 and TIF2 (Forman et al., 1998; Tzameli et al., 2000; Min et al., 2002; Jyrkkärinne et al., 2003; Shiraki et al., 2003; Frank et al., 2004). The prototypic cytochrome P450 inducer phenobarbital stimulates translocation of CAR, but has not been shown to bind CAR LBD unlike many other activators (Willson and Kliewer, 2002; Honkakoski et al., 1998b). The differences in ligand specificity between human and mouse CAR are significant. For example, mouse CAR is inhibited by a limited set of steroids related to 3α-androstenol (ANDR) (Forman et al., 1998; Jyrkkärinne et al., 2003) whereas TCPOBOP (Tzameli et al., 2000), chlorpromazine (Wei et al., 2002) and 17α-ethynyl-$3,17\beta$-estradiol (EE2) (Mäkinen et al., 2003) are potent activators of mouse CAR. However, they do not activate human CAR at all (Honkakoski et al., 2003) and EE2 even inhibits it (Jyrkkärinne et al., 2003). Thus, of the few known modulators of human CAR, EE2 and ANDR are re-

pressors (Jyrkkärinne et al., 2003; Kawamoto et al., 2000; Mäkinen et al., 2002), whereas 6-(4-chlorophenyl)imidazo-[2,1-b][1,3]thiazole-5-carbaldehyde *O*-(3,4-dichloro-benzyl)oxime (CITCO) and tri-(*p*-methylphenyl) phosphate (TMPP) are activators (Maglich et al., 2003; Honkakoski et al., 2004). The effect of clotrimazole on human CAR is not clear: the effects range from repressing (Moore et al., 2000) to no effect (Toell et al., 2002) to activating (Mäkinen et al., 2002). A likely explanation for this discrepancy is the use of different cell lines in co-transfection assays. It is known that the responses of selective modulators of estrogen and progesterone receptors vary due to differential content and selection of NR co-regulators in cell lines (Smith et al., 1997; Liu et al., 2002). In our studies with HEK293 cells, clotrimazole has consistently activated human CAR (Jyrkkärinne et al., 2003; Mäkinen et al., 2002; Honkakoski et al., 2004).

The crystal structure of human CAR LBD has not been reported, and partly for that reason the molecular determinants for its ligand specificity remain obscure. We wished to study the importance of the amino acid residues forming the LBP of human CAR. In order to select the critical residues for mutation analysis, homology models of human CAR LBD were created. Twenty-two LBP amino acids were identified and alanine scanning mutagenesis was carried out. The activities and ligand specificities of these mutants were tested in mammalian activity assays. The yeast two-hybrid assay was applied to study the interactions between the mutated LBDs and the steroid receptor coactivator-1 (SRC-1) and the nuclear receptor co-repressor (NCoR). Then, to gain insight into the molecular mechanisms of ligand binding, improved CAR models were constructed, and the ligands were docked and run through molecular dynamics (MD) simulations. Various ligand derivatives were used to complement the mutation analysis. These studies identified, for the first time, several amino acids that *i*) contribute to the high basal activity of human CAR, *ii*) control the ligand selectivity, and *iii*) explain some of the species differences in CAR ligand specificity.

5.2 Results

5.2.1 Modulation of Human CAR Activity

Fig. 5.1*A* shows that TMPP and clotrimazole enhanced the basal activity of GAL4-human CAR LBD by 2.3- and 1.8-fold, while EE2 and ANDR inhibited it by 40% and 30%, respectively. These results were expected from data on full-length human CAR (Mäkinen et al., 2003; Kawamoto et al., 2000; Mäkinen et al., 2002; Honkakoski et al., 2004; Moore et al., 2000). None of the chemicals changed the activity of GAL4 only. In competition assays (Figs. 5.1*B&C*), the presence of TMPP and clotrimazole shifted the inhibition curves of EE2 and ANDR to the right, with 5-10-fold increases in apparent IC_50 values. Therefore, the inhibitors and activators compete for the same or overlapping binding site within the human CAR LBD. Because the ligand response is a net result from interactions of human CAR with associated co-regulators, Y2H assays with NRID peptides from SRC-1 and NCoR were set up. There was a very strong basal interaction between human CAR and the SRC-1 NRID (Fig. 5.1*D*) which was not much affected by TMPP, increased 35% by clotrimazole and decreased 10-20% by inhibitors ANDR and EE2. The very weak basal interaction between human CAR and the NCoR NRID was not affected by TMPP, modestly increased (3-fold) by clotrimazole and strongly enhanced (over 12-fold) by the inhibitors ANDR and EE2. Our data support the results that human CAR inhibitors act by promoting NCoR interaction and activators slightly increase SRC-1 interaction (Jyrkkärinne et al., 2003; Mäkinen et al., 2002; Honkakoski et al., 2004).
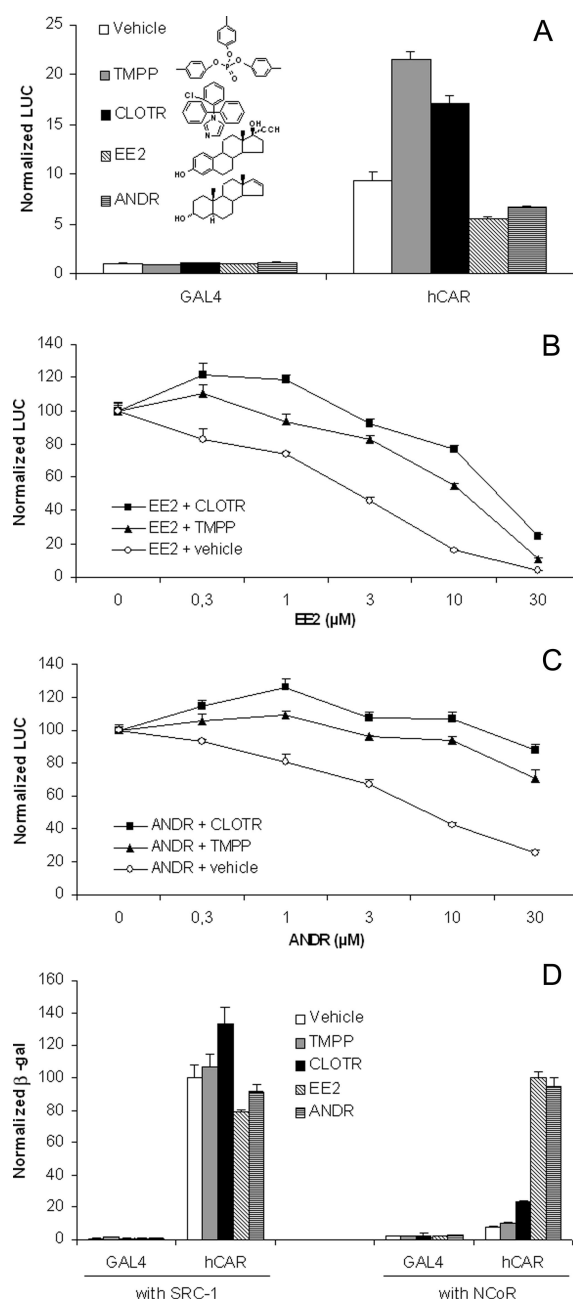
Figure 5.1: Modulation of human CAR activity. (**A**), Reporter activities of empty GAL4 and GAL4-human CAR LBD (hCAR) in HEK293 cells after exposure to indicated modulators TMPP (10 mM), clotrimazole (2 mM), EE2 (10 mM) and ANDR (10 mM). (**B**) and (**C**), Dose-response curves of inhibitors EE2 (**(B)**) and ANDR (**(C)**) in the absence or presence of activators TMPP (10 mM) and clotrimazole (CLOTR) (2 mM). (**D**), The modulator-elicited responses in recruitment of SRC-1 (left panel) and NCoR peptides (right panel) with empty GAL4 and GAL4-human CAR LBD (hCAR) in yeast cells. The data are expressed as mean +/- SEM of at least three independent experiments.

5.2.2 Homology Models of Human CAR

The initial homology model was built on VDR template to identify residues that form the human CAR LBP. Fig. 5.2 shows 22 residues that were selected for the subsequent mutation analysis. Similarly to PXR, most of the residues lining the pocket (15/22) are hydrophobic, and polar residues (N165, C202, H203, T209, Y224, Y326, Q329) were rather evenly distributed. The four modulators were then docked into this model and MD simulations were carried out. However, the number of possible conformations obtained was too high for any meaningful interpretation of the functional data.


```
                108   124 129 134 157   178 187   208 216   224 226 232 236   251 256   267 278        229 308          335 341 347
         N─┤ H1 ├┤H2├┤  H3  ├┤ H4/H5 ├┤  β  ├┤H6├┤ H7 ├┤ H8 ├┤  H9  ├┤ H10/11 ├┤H12├─C
                    x      x xx x        x xxx x       x  x      x x  xx                              x xxx x x
            F129┘ ┌F161┘       V199┘      F217┘             ┌─F243                      Y326┘       ┌─L343
                  I164───      C202┘      Y224─            ├──L242                      Q329┤       ├──M340
                  N165───      H203──              ├──F238                      I330┘       ┤──I333
                  M168────     L206 ──     ┌T209                 ├──F234
```

```
                       110        120        130
                        .          .          .
    hCAR 106      LSKEQEELIRTLLGAHTRHMGTMFEQFVQFRPP----------------------------------------
    hVDR 120   LRPKLSEEQQRIIAILLDAHHKTYDPTYSDFCQFRPP----------------------------------------
    hPXR 142      GLTEEQRMMIRELMDAQMKTFDTTFSHFKNFRLPGVlssgcelpeslqapSREEAAKWSQVRKDLCSIKVSLQ
                140        150        160        170        180        190        200
                 .          .          .          .          .          .          .
    hCAR 139   ------AHLFIHHQPLPTLAPVLPLVTHFADINTFMVLQVIKFTKDLPVFRSLPIEDQISLLKGAAVEICHIVLNT
    hVDR 157   ------VRVNDGGGSVTLELSQLSMLPHLADLVSYSIQKVIGFAKMIPGFRDLTSEDQIVLLKSSAIEVIMLRSNE
    hPXR 215   LRGEDGSVWNYKPPADSGGKEIFSLLPHMADMSTYMFKGIISFAKVISYFRDLPIEDQISLLKGAAFELCQLRFNT
                210        220        230        240        250        260        270        280
                 .          .          .          .          .          .          .          .
    hCAR 209   TFCLQTQNFLCG--PLRYTIEDGARVGFQVEFLELLFHFHGTLRKLQLQEPEYVLLAAMALFSPDRPGVTQRDEID
    hVDR 278   SFTMDDMSWTCGNQDYKYRVSDVTKAGHSLELIEPLIKFQVGLKKLNLHEEEHVLLMAICIVSPDRPGVQDAALIE
    hPXR 291   VFNAETGTWECG--RLSYCLEDTAG-GFQQLLLEPMLKFHYMLKKLQLHEEEYVLMQAISLFSPDRPGVLQHRVVD
                290        300        310        320        330        340
                 .          .          .          .          .          .
    hCAR 283   QLQEEMALTLQSYIKGQQRRPRDRFLYAKLLGLLAELRSINEAYGYQIQHI---QGLS-AMMPLLQEICS
    hVDR 354   AIQDRLSNTLQTYIRCRHPPPgshLLYAKMIQKLADLRSLNEENSKQYRCLSFQPECSMKLTPLVLEVFGneis
    hPXR 364   QLQEQFAITLKSYIECNRPQPAHRFLFLKIMAMLTELRSINAQHTQRLLRI---QDIHPFATPLMQELFGItgs
```

Figure 5.2: Sequence alignment of human CAR. Top, The first and the last residue of each helix and the β-sheet in the human CAR model are shown in the graph, and the positions of mutated LBP residues are indicated by an X . Bottom, The alignment of human CAR, VDR, and PXR sequences. The human CAR residues 106-128, 158-223, 249-307, and 334-337 were modelled based on human PXR residues 143-165, 240-305, 330-388, and 415-418, respectively. The template for human CAR residues 129-157, 224-248, 308-333 and 340-348 was taken from hVDR residues 147-164 plus 216-226, 295-319, 397-404 and 415-423, respectively. The mutated CAR residues are shown in bold and hydrophobic residues are underlined. The VDR and PXR residues lining the LBPs are shown in bold and italics. Residues marked lowercase are missing in the template VDR and PXR structures.

This indicated that more careful selection and validation of the protein model is of utmost importance. After these initial models were completed, the human PXR crystal coordinates (Watkins et al., 2003a) became available, and improved knowledge of co-regulator binding of other NRs enabled us to create more advanced models with and without NCoR or SRC-1 NRID peptides (see Figs. 5.3*A&B*). In order to account for the structural differences between PXR and CAR (Moore et al., 2002), the VDR structure (Rochel et al., 2000) was used as scaffold for the modelling of the region connecting helices 1 and 3, and helices 6 and 7. The helices 10 through 12 were also built upon the VDR template. In the final model, the LBP is formed by helices 3, 5-7 and 10/11 and the β-sheet. The 22 residues used in mutation analysis are lining the LBP as well (Figs. 5.3*C&D*). The evaluation of the models indicated that they were of high stereochemical quality. In the Ramachandran plots, 86.4-92.2% of the ϕ/ψ torsion angles were within the favorable region for all CAR models, the Profiles-3D scores were close (89.1-99.1%) to values expected for a high quality model of corresponding size, and the protein folds remained stable during MD simulations. Finally, only one or two preferred conformations of ligands (Table 5.1) with excellent scoring values emerged.

These CAR LBD models (Fig. 5.3*A*) suggested that even without ligand, helix 12 adopts the active conformation due to hydrophobic interactions of L343 with LBD residues Y326 and I330, and of I346 with residue V199. Of importance is the hydrogen bonding between N165 and Y326 that will stabilise Y326 in a position that allows interactions with L343 in helix 12 (Fig. 5.3*B*). The central role of residues N165, V199, Y326, I330 and L343 for basal activity was confirmed by dramatic decreases in activity and SRC-1 interaction by their mutation to alanine (see Figs. 5.4*A&C*). The presence of SRC-1 seemed to further stabilise helix 12 in the active conformation and push it closer to the LBD in such a way that Y326 could now contact with I346 (Figs. 5.3*A&B*). Thus, helix 12 adapts to SRC-1 binding by improving the fit between the NRID peptide and the LBD. Another example of this adaptive fit was that Q331 (helix 11) made a hydrogen bond to S348 (helix 12) only in the presence of SRC-1 (Fig. 5.3*B*), suggesting its involvement in proper aligning of helix 12. Subsequent mutagenesis of Q331A decreased the basal activity of human CAR by about 70% (data not shown). More detailed MD studies on the structural basis of constitutive activity are described in another paper (Windshügel et al., 2005). In the presence of NCoR, helix 12 could not acquire the active conformation (Fig. 5.3*A*).

Figure 5.3: Homology models of human CAR. (**A**), Superposition of LBDs in different CAR models. The C-terminal helix 12 remains in its active conformation without (grey) or with SRC-1 (orange). In presence of SRC-1 peptide, helix 12 is pushed closer to the LBD. Helix 12 is displaced from its original position by the NCoR peptide (cyan). Here, only the C-terminal part of CAR is shown for clarity. NRID peptides were SLTERHKILHRLLQE (SRC-1) and NLGLEDIIRKALMGSFDDK (NCoR) with core sequences underlined. (**B**), Schematic representation of contacts between helix 12 and the rest of LBD. The lines indicate interactions in the presence (solid) or absence (dotted) of SRC-1 peptide. (**C**) and (**D**), Detailed views residues lining the LBP. Parts of the contributing structural elements are shown as tubes. LBP residues that have been mutated are shown in black, other residues are kept grey. (**C**), Side view of the LBP, lined by mainly hydrophobic amino acids. The hydrogen bond between N165 and Y326 is indicated. (**D**), Rotation about 90 degrees, view from the beta-sheet into the LBP.

Significantly, the size of the LBP in the final model (480 Å^3) was smaller than that of crystallised PXR (1294 Å^3) (Watkins et al., 2003a) and of previous CAR models (1150-1170 Å^3) that were built on PXR but not validated by MD simulations (Frank et al., 2004; Dussault et al., 2002; Jacobs et al., 2003). The small size of LBP is due to protruding, bulky aromatic (e.g., F129, F161, F217, Y224, F234,

| Model | Poses | Clusters | Poses in | Max. RMSD | | Distance |
Ligand			cluster(s)	Cluster 1	Cluster 2	betw. clusters
CAR only						
TMPP	11	2	6; 5	0.38	0.47	1.41-1.81
TPP	11	2	7; 6	0.73	0.98	1.44-1.94
Clotr.	19	2	18; 1	0.93	-	4.39-5.15
EE2	23	2	15; 8	0.39	0.44	7.54-7.82
E2	11	2	9; 2	0.47	0.18	7.64-7.69
Mestr.	9	1	9	0.43	-	-
CAR/SRC-1						
TMPP	18	2	12; 6	0.91	0.41	1.54-1.81
TPP	30	2	27; 3	0.63	0.94	1.53-2.03
Clotr.	19	2	11; 8	0.36	0.57	1.65-2.10
CAR/NCoR						
ANDR	9	1	9	0.44	-	-
EE2	9	1	9	0.23	-	-

Table 5.1: Docking of ligands into the final homology models.

F238 and Y326) and hydrophobic (e.g., I164, M168, V199, I330) residues (Figs. 5.3*C&D*). The smaller size of CAR pocket relative to PXR is also consistent with the fact that fewer and smaller chemicals may act as ligands for CAR than for PXR (Willson and Kliewer, 2002; Honkakoski et al., 2003).

5.2.3 Basal Activities of Human CAR Mutants

The activities of the 22 mutants were measured and compared to those of the wild-type human CAR (Figs. 5.4*A&B*). Interestingly, the basal activities (white columns) of 16 mutants were decreased to 10% or less of the wild-type activity. Our interpretation is that mutation of aromatic (F161, F234, F238, Y326) or hydrophobic (C202, I164, M168, I330, I333, M339) residues that protrude into the LBP to alanine will create more space and reorganise the surrounding LBP residues. This reorganisation would decrease the basal activity in most cases via the central residue Y326, the position of which is crucial for the stabilisation

of helix 12 in active conformation. Second, at least three other residues (V199, N165 and L343) are also involved in stabilisation of helix 12 (see above); accordingly, their mutation also decreased the basal activity. Third, three mutants (F129A, F217A and Y224A) also had low basal activity, but they could not be activated in mammalian cells (Fig. 5.4*A*) or could not elicit a SRC-1 or NCoR response in yeast (Figs. 5.4*C&D*) like the other 19 mutants did. Residues F217 and Y224 form a wall in the LBP (Fig. 5.3*C*), and their change to alanine is likely to disrupt protein folding locally. Finally, the remaining six mutants (H203A, L206A, T209A, L242A, F243A and Q329A) retained 30-80% of the wild-type activity. These residues were located on the top and back of LBP (Figs. 5.3*C&D*), and apart from H203, their side chains were not clearly projected to the center of LBP or the central residue Y326. Therefore, their larger distance from Y326 was consistent with the finding that their mutations could not profoundly attenuate the basal activity via influence on position of Y326.

The majority of mutants (13/16) with most dramatic losses in basal activity also had corresponding losses in SRC-1 interaction (Fig. 5.4*C*, white columns). Remaining three low activity mutants (N165A, C202A and I333A) showed only modest reduction in SRC-1 interaction (20-40 %). Of these three exceptions, N165A interacted strongly with NCoR even without any ligand (Fig. 5.4*D*). This increase in co-repressor binding well explains the loss of net activity of this mutant in mammalian cells.
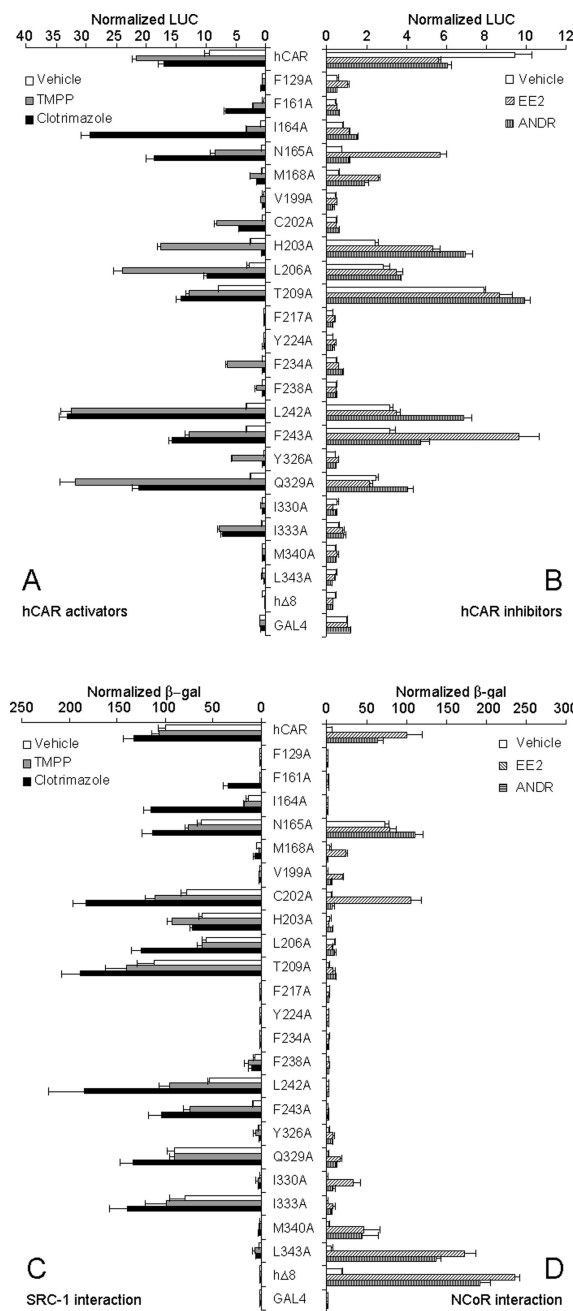
Figure 5.4: The effects of LBP mutations on human CAR activity. (**A**) and (**B**), The reporter activities were measured in HEK293 cells as in Fig. 5.1 for activators (**(A)**) or inhibitors (**(B)**), and normalised to empty GAL4 values (set at 1.0). **C** and (**D**), The same constructs were the assayed in yeast for interactions with NRID peptides SRC-1 (**(C)**) or NCoR (**(D)**). The SRC-1 and the NCoR results were normalised to wild-type CAR treated with vehicle (= 100, **C**) or with 10 micro M ANDR (=100, **D**), respectively. The data are expressed as mean +/- SEM of 3-6 independent experiments.

The reason for low basal activities of C202A and I333A was not directly apparent from the NCoR and SRC-1 recruitment. Therefore, we tested the possibility that these two mutants differed in their association with co-activators other than SRC-1. The recruitment of NRID peptides from TRAP or PGC-1 was significantly lower by C202A and I333A mutants than by the wild-type CAR (Fig. 5.5*A*), providing a basis for their low net activity in mammalian cells that are known to express a variety of co-activators. In remaining six mutants retaining about one-third or more of wild-type activity (H203A, L206A, T209, L242A, F243A and Q329A), the correspondence between SRC-1 association and co-transfection assay was good (T209A, F243A; Figs. 5.4*A&C*) or compensated for by progressively larger decreases in PGC-1 interaction (H203A, L206A, L242A; Fig. 5.5*A*), thus leaving only one mutant (Q329A) out of 22 without a clear match between co-transfection and Y2H results.

5.2.4   Docking and MD Simulation of Ligand Binding

To find determinants of ligand binding, the four modulators were docked into appropriate human CAR/NRID peptide models, and MD simulations were carried out. Docking of TMPP into CAR-SRC-1 yielded 18 favoured poses which could be grouped into two clusters which differed only slightly from each other. For clotrimazole 19 favoured poses showing also just minor differences between each other were obtained (Table 5.1). The two different binding orientations were analysed visually and compared considering their scoring values. The top-ranking pose for TMPP and clotrimazole was then further investigated by MD simulation. A representative frame of the MD simulation, calculated using NMRCLUST (Kelley et al., 1996), of CAR/SRC-1 model with bound activators is shown in Figs. 5.6*A&B*, respectively. Clotrimazole was bound deep in the LBP without any contact to helix 12. The phenyl rings of clotrimazole made contacts with residues F161, F217, Y224, F234 and Y326 while no LBP residue formed a clear hydrogen bond with the imidazole ring. During MD simulation, the distance between Y326 and helix 12 was decreased. F161 reoriented to the interface between helix 12 and LBD, interacting with the hydrophobic M339, and F243 now pointed away from the LBP. All these movements enlarged the LBP to a volume of 750 Å^3. TMPP shows a different binding mode (Fig. 5.6*A*).
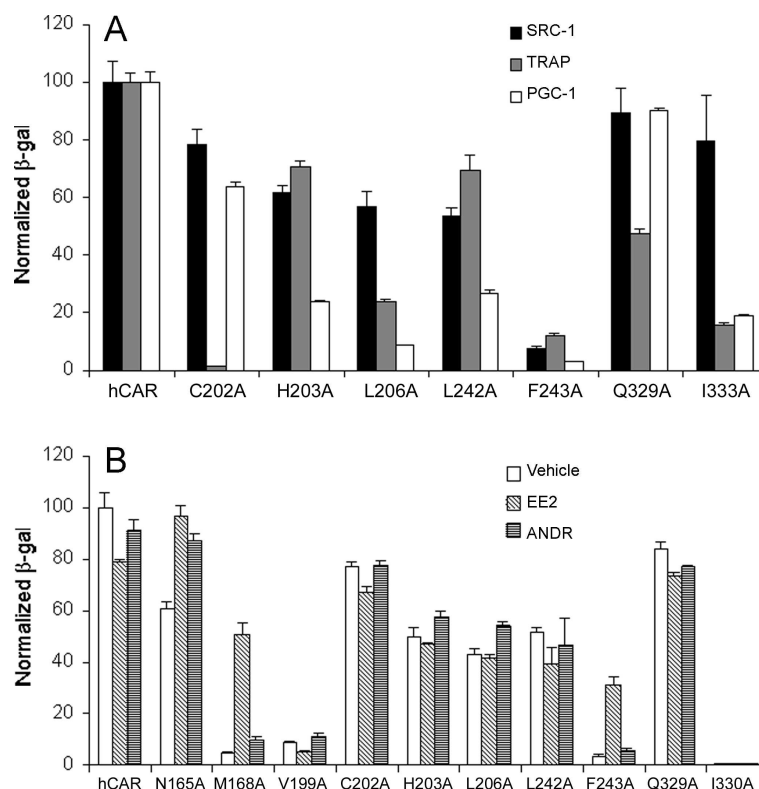
Figure 5.5: Differential association of coactivators with human CAR mutants. (**A**), The basal interaction of CAR and selected mutants with NRID peptides from coactivators SRC-1 (black), TRAP (grey) and PGC-1 (white). The assays were conducted as in Fig. 5.4*C*, and normalised to wild-type CAR (= 100). (**B**), The EE2- and ANDR-elicited interactions of CAR and selected mutants with NRID peptide from coactivator SRC-1. The assays were done as in Fig. 5.4*C*, and normalised to vehicle-treated wild-type CAR (= 100). The data are expressed as mean +/- SEM of three independent experiments.

It contacted residues I164, N165, M168, V199, C202, H203, F217 and Y224. In addition, one of the methylphenyl groups interacted directly with helix 12 residues L343 and I346. Two other methylphenyl groups contacted F234 and Y326 and also interacted with F161, pushing it deeper into the LBP. TMPP was also held in place by strong hydrogen bonds between the phosphate group and N165 and Y326 that remained stable during the entire MD simulation. Upon TMPP binding, the helix 12 moved closer to LBD, and the LBP size increased to about 560 $Å^3$. In summary, the binding of activators caused an expansion of the LBP volume, and significantly, helix 12 was packed closer to the LBD. This shift of helix 12 was expected to improve the subsequent binding of SRC-1 peptide (compare with Fig. 5.3*A*).
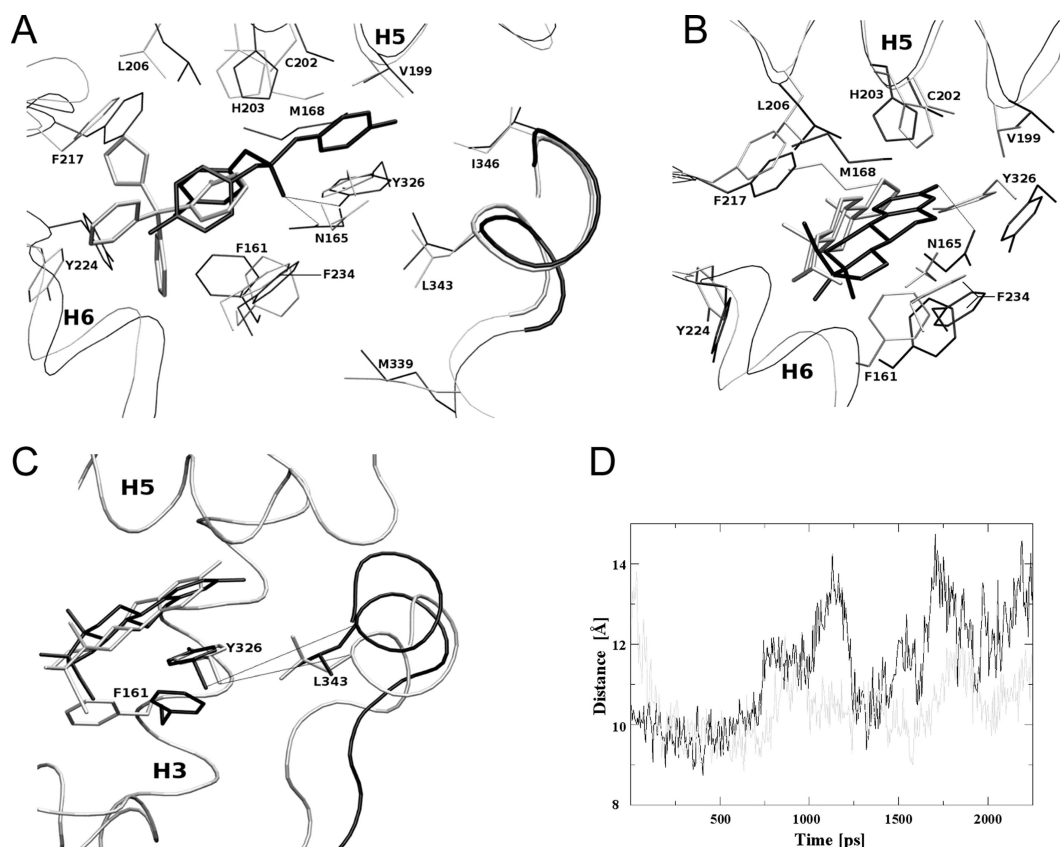
Figure 5.6: The conformations of human CAR modulators within the LBP. (**A**), Activators clotrimazole (grey) and TMPP (black) were docked and simulated in the LBP of CAR-SRC-1 model. Clotrimazole is located deep within the LBP without any hydrogen bonds or direct contacts to helix 12. TMPP interacts with helix 12 and is bound by hydrogen bonds to N165 and Y326. In both cases, ligand binding decreases the distance between Y326 and helix 12 as compared to ligand-free CAR. (**B**), Inhibitors ANDR (grey) and EE2 (black) docked in the LBP of CAR-NCoR model. A stable hydrogen bond is formed between the 3-hydroxyl group of EE2 and N165 while the 3α-hydroxyl of ANDR was not involved in hydrogen bonding. (**C**), Comparison of CAR with docked EE2 (light grey) and a representative frame from the subsequent MD simulation (black). After MD, the position of EE2 showed only a minor change. F161 was reoriented towards the interface between LBD and helix 12. The distance between the LBD and helix 12 was increased (black lines, see also **D**). Parts of the protein backbone were removed for clarity. (**D**), Distances between Cα atoms of Y326 and L343 plotted for each frame of the MD. In empty CAR (grey), this distance is about 10.7 Å. Higher values are found at the beginning of the MD due to equilibration of the model. In complex with docked EE2 (black), the distance is increased by about 2 Å with much larger fluctuations.

In the CAR/NCoR model, only one energetically favourable solution was obtained when ANDR or EE2 were docked. In both cases, the ligand was interacting with the same set of side chains: F161, N165 and M168 from helix 3, residues V199, C202, H203 and L206 from helix 5, and F217, Y224, F234 and Y326 (Fig. 5.6*B*). However, the plane and orientation of the steroids were quite opposite: the A-ring of ANDR was located in the same region as the D-ring of EE2, and their C18 methyl groups were facing helix 5 and helix 3, respectively. ANDR did not form any clear hydrogen bonds while the 3-hydroxyl of EE2 made a stable bond with N165. The conformation of EE2 suggested that EE2 might inhibit CAR by disturbing the hydrophobic interactions between helix 12 and the LBD with the protruding 3-hydroxyl group that pointed towards L343 in the CAR/SRC-1 model. MD simulation of CAR-EE2 complex revealed that binding of EE2 to CAR LBD increased the distance between Y326 and helix 12 while EE2 itself rotated by 30 degrees along its long axis and forced reorientation of F161 (Figs. 5.6*C&D*). Similar weakening of interactions between helices 11 and 12 due to reorientation of Y326 was seen in the CAR/SRC-1 model upon EE2 addition (data not shown). For ANDR inhibition, a similar mechanism could not be suggested. Nevertheless, Y326 was located much deeper in LBP, increasing its distance from helix 12 (Fig. 5.6*B*).

5.2.5 Ligand Specificities of Human CAR Mutants

Next, the responses of 22 mutants to activators were measured in co-transfection and Y2H assays (Figs. 5.4*A&C*). We first considered mutants with low basal activity: Ligand responsiveness was totally lost by only three mutants (F129A, F217A and Y224A). Four mutants (V199A, I330A, M340A, L343A) did not respond to activators but inhibitors provoked their association with NCoR (Fig. 5.4*D*). Three of these residues (V199, I330, L343) were found to keep helix 12 in the active position (see above), which explains the lack of SRC-1 recruitment by these mutants even after addition of an activator. The remaining 15 mutants had either low-to-moderate basal activity which could be increased by TMPP and/or clotrimazole to variable degrees (Fig. 5.4*A*). Of these, eight mutants (M168A, C202A, L206A, T209A, L242A, F243A, Q329A and I333A) were activated by TMPP to same or higher degree than by clotrimazole, thus resembling the preference of wild-type human CAR. These residues were located on the top and back of the LBP (Figs. 5.3*A&B*), away from the residues in contact with the activators (Fig. 5.6*A*), which probably explains why these mutations had only weak effects on activator preference. Four mutants were activated by TMPP only (H203A, F234A, F238A and Y326A). These aromatic residues form hydrophobic and/or stacking interactions with clotrimazole in MD simulations (Fig. 5.6*A*) and their replacement with alanine explains the loss of activation. Finally, three helix 3 mutants (F161A, I164A and N165A) preferred clotrimazole over TMPP (Figs. 5.4*A&C*). A likely explanation is that residues I164 and N165 interact with the TMPP phosphate and methylphenyl groups (but not with clotrimazole) and these interactions would be lost upon mutation (Fig. 5.6*A*). Because F161 interacts with both TMPP and clotrimazole, it is easy to see why the F161A mutation decreased activation and SRC-1 recruitment by both activators (Figs. 5.3*A&C*).

Responses to human CAR inhibitors were investigated next (Figs. 5.4*B&D*). Among the 16 low-activity mutants, any further inhibition was difficult to observe. However, two helix 3 mutants displayed significant activation by EE2 only (N165A) or by both inhibitors (M168A). These results were supported by Y2H assays. First, only ANDR but not EE2 further enhanced NCoR recruitment by N165A while the NCoR responses of M168A were greatly decreased

(EE2) or abolished (ANDR) (Fig. 5.4*D*). Second, both EE2 and ANDR increased interactions of these mutants with SRC-1 (Fig. 5.5*B*). In NCoR assays, three additional mutants (V199A, C202A and I330A) were responsive to EE2 but not to ANDR. Among the five mutants with moderate basal activity, none were inhibited by EE2 or ANDR anymore; in fact, there was about a 2-fold activation by either steroid (L242A and F243A) or both (H203A). Consistent with the loss of inhibition, the responses to EE2 and ANDR of these mutants in NCoR assays were attenuated by 70-90 % (Fig. 5.4*D*). SRC-1 assays with EE2 and ANDR indicated a clear-cut activation of F243A by EE2 only (Fig. 5.5*B*).

Taken together, three amino acids (N165, M168 and F243) appear to regulate the response of EE2 consistently in both types of assays while additional residues (V199, C202, I330) may contribute to recognition of ANDR. Of these residues, at least N165, M168, V199 and C202 are located close to the steroids, and residue N165 formed a hydrogen bond with EE2 in MD simulations (Fig. 5.6*B*). Because ANDR was not fixed by any hydrogen bonds, mutation of closely located residues V199, C202 and I330 to alanine may create sufficient space to completely reorient ANDR but not EE2, explaining the selectivity of these residues.

In summary, based on carefully constructed molecular models, we were able to identify, for the first time, residues critical for ligand binding and to explain the effects of mutations with all types of human CAR modulators. This was not possible with simpler, single template-based models or without MD simulations. To validate our models further, we docked TPP, a derivative of TMPP lacking the three methyl groups, into CAR models and assayed the activities of selected human CAR mutants. The applied docking procedure led to two different binding orientations (Table 5.1). One binding mode resembled the one obtained for TMPP that displayed hydrogen bonds with Y326 and N165. In the second cluster, the phosphate group of TPP was reoriented and made a hydrogen bond to H203 instead (Fig. 5.7*A*). Docking of TPP into the CAR-SRC-1 model resulted in a similar pose showing the same hydrogen bond to H203. This suggested that TMPP and TPP interact differentially with residues N165 and Y326. The N165A mutation was expected to eliminate hydrogen bonding to Y326 and increase its flexibility, thereby decreasing TMPP-induced CAR activity, but not affect TPP-induced activity very much. This prediction was borne out exactly in activation assays (Fig. 5.7*B*).
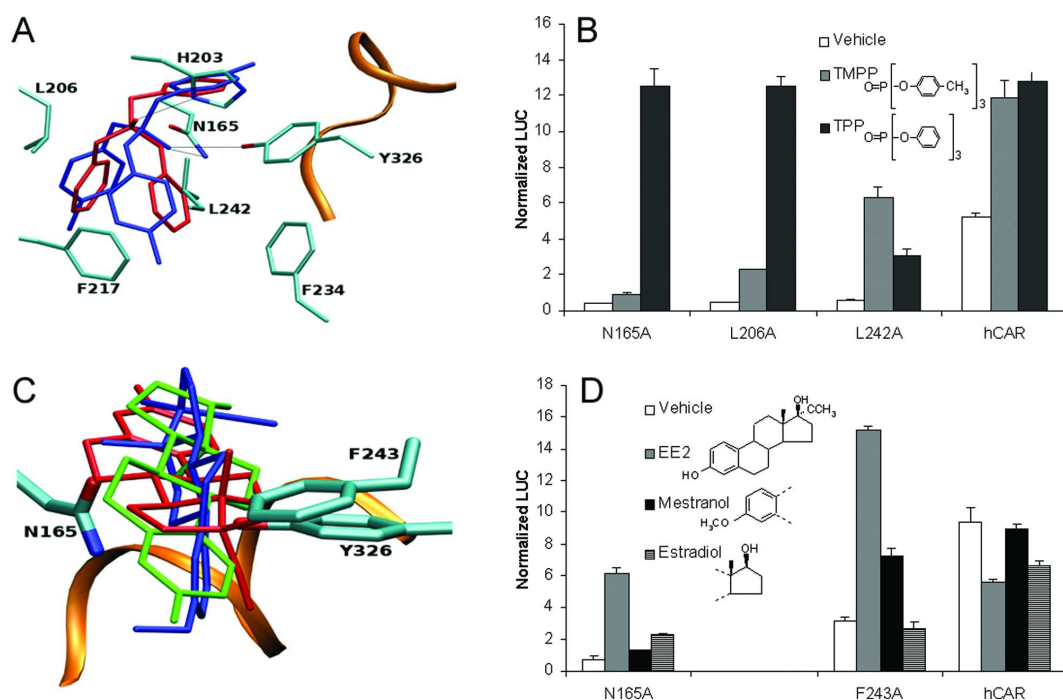
Figure 5.7: The predicted interactions of modulator derivatives within the human CAR LBP. (**A**), Docked conformations of TPP (red) and TMPP (blue) in CAR LBP. Unlike TMPP, TPP is not bound to N165 or Y326 but it formed a hydrogen bond with H203. Helix 12 is shown in orange ribbon. (**B**), The TMPP- and TPP-elicited activities of selected mutants were measured as in Fig. 5.1. (**C**), Docked conformations of E2 (green), EE2 (red) and mestranol (blue) in CAR LBP. Residues important for ligand recognition are depicted with capped sticks. (**D**), The estrogen-elicited activities of selected mutants were measured as in Fig. 5.1.

In another test, derivatives of EE2 were studied (Fig. 5.7*C*). Docking of mestranol (3-hydroxyl is methylated) into CAR resulted in just one favourable conformation whereas for E2 (lacking the 17α-ethynyl group) two different binding orientations were observed (Table 5.1). About 80% of all poses could be grouped in a major cluster in which the 17β-hydroxyl group of E2 pointed towards the helix 12/LBD interface. The minor cluster revealed a reversed orientation of the steroid showing much lower scoring values. For EE2, 23 poses in two different clusters were obtained. In the top-ranking pose, the 3-hydroxyl group of the steroid was pointing towards the helix 12/LBD interface (*i.e.*, a reversed binding mode compared to E2 and mestranol). This conformation of EE2 is similar to the one obtained when docking the ligand into the CAR/NCoR model. The

second cluster contains poses with lower scores and a conformation similar to that obtained for mestranol. Keeping in mind the top-ranking binding orientations of E2 and EE2 (Fig. 5.7C) and the above-suggested mechanism of EE2 inhibition, these calculations imply that E2 and EE2 would inhibit human CAR but mestranol would not. The experimental results confirmed this prediction (Fig. 5.7D). Because both N165 and F243 residues were thought to be important for EE2 recognition (see above), we also deduced that all these estrogens would be activators of the N165A mutant, regardless of their orientation. This would be the case due to lack of hydrogen bonding between A165 and Y326. The effect of F243A would, on the other hand, depend on how well the estrogen was fixed into the LBP by hydrophobic interactions. In co-transfections, N165A was activated by 2-6-fold by estrogens, while the activity of F243A was enhanced 2-6-fold by EE2 and mestranol but not by E2 (Fig. 5.7D). We thought that the inactivity of E2 was due to its complete reorientation within the LBP of F243A mutant but mestranol was still able to interact with and stabilise Y326 via its 17α-ethynyl group.

Previously we have found, based on mouse/human CAR chimeras, that replacement of human CAR amino acids 190-253 with corresponding mouse residues converted EE2 from an inhibitor to an activator (Jyrkkärinne et al., 2003). In this region, LBP residues that differed between the human and mouse CAR were C202L, L242I and F243L. We made these human-to-mouse mutations and analysed the effect of EE2 on CAR activity. We found that although inhibition by EE2 was reduced by C202L and L242I, only F243L mutation was clearly activated by EE2 (Fig. 5.8). In all, the critical role of F243 in EE2 recognition was supported by docking into homology models, mutagenesis studies and naturally occurring species variation.
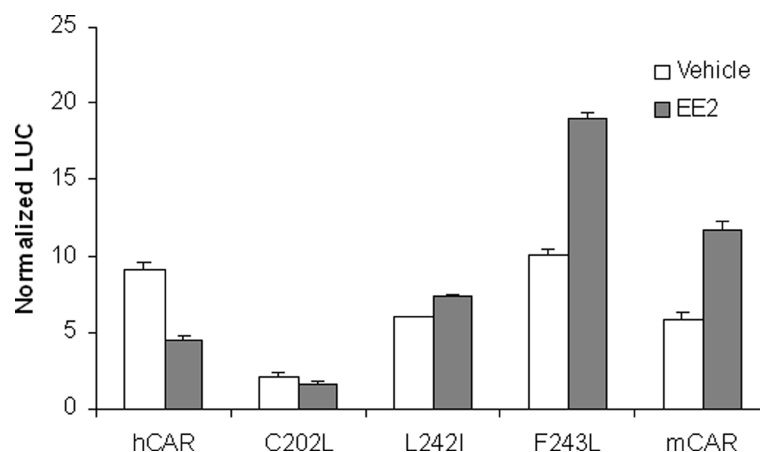
Figure 5.8: The species difference in CAR response to EE2. The reporter activities of indicated CAR constructs were transfected into HEK293 cells and treated with vehicle or 10 micro M EE2 and assayed as in Fig. 5.1. The data are expressed as mean +/- SEM of three independent experiments.

5.3 Discussion

5.3.1 Factors Contributing to Basal Activity of CAR

Our homology models of CAR indicated that residues I346 and L343 in helix 12 made hydrophobic contacts with V199 (helix 5) and with Y326 and I330 (helix 11), respectively. Furthermore, the central residue Y326 was stabilised by N165. All these interactions contributed to helix 12 acquiring the active position, because mutation of each residue to alanine caused marked decreases in the basal activity. In addition, novel contacts between helices 11 and 12 were formed upon SRC-1 binding (Q331-S348, Y326-I346). Other NRs with significant basal activity such as Nurr1 (Wang et al., 2003b), ERRγ (Greschik et al., 2002), RORβ (Stehlin et al., 2001) and LRH-1 (Sablin et al., 2003) also have similar contacts in their crystal structures (see Table 5.2). However, the interaction in each case may involve not only hydrophobic contacts but also hydrogen bonding, aromatic stacking, and even salt bridges. Even though residues corresponding to the pair I330-L343 are often seen, the interacting amino acid pairs do not uniformly match the pairs in human CAR. Nevertheless, these findings indicate that the presence of extensive contacts between helices 11 and 12 may be fundamental to basal activity of NRs in general. The possibility of such interactions has been pointed out in other homology models of CAR although no MD sim-

ulations had been conducted (Frank et al., 2004; Andersin et al., 2003). We were not able to detect any hydrogen bonds between C347 and Y326, as described for mouse and human CAR (Frank et al., 2004; Andersin et al., 2003), but we agree on the importance of a hydrophobic contact between I330 and L343 in human CAR (Frank et al., 2004).

A charge clamp between K205 in helix 4 and the negatively charged carboxyl terminus (S358) was suggested important for basal activity (Dussault et al., 2002). This notion was based on a mouse CAR model built on PXR and loss of activity due to extension of helix 12 by one helical turn (Dussault et al., 2002). We could not detect such an interaction in any of our models. The residue corresponding to K205 (K195 in human CAR) is highly conserved among NRs and it has well-documented interactions with NRID peptides in other NR crystal structures (Watkins et al., 2003a; Gampe Jr et al., 2000). In a crystal structure of an agonist-bound NR that lacks the SRC-1 peptide, the corresponding lysine residue does not contact helix 12 at all (Rochel et al., 2000). In addition, extension of helix 12 by three residues in mouse CAR did not influence the basal activity (Andersin et al., 2003). Mutation of K205 is therefore likely to decrease basal activity of mouse CAR by loss of SRC-1 binding rather than by destabilisation of helix 12. On another note, the residue N165 (helix 3) exerts its effects through Y326 while the residue V199 (helix 5) appears to contribute to stability of helix 12 our CAR and CAR/SRC-1 models. Although the role of N165 may be of unique importance to CAR, hydrophobic contacts between residues matching V199 and helix 12 are present at least in VDR (I268-F422) and PXR (F281-F429) structures (Rochel et al., 2000; Watkins et al., 2003a).

The above data indicate that LBD residues that interact with helix 12 also influence SRC-1 and/or NCoR binding. This view supports reports on both naturally occurring or alanine scanning mutants of other NRs such as RARα and TR (Collingwood et al., 1998; Côte et al., 2000; Marimuthu et al., 2002; Benko et al., 2003). For example, RARα alanine scanning (Benko et al., 2003) indicated that mutation of amino acids in close proximity to helix 12 such as L266A (V199 in CAR) and V395A (I330) or indirectly affecting helix 12 such as T233A (N165) exhibit no or inefficient release of co-repressors upon agonist binding. Even though the same residues are implied, the actual effect of corresponding mutations in CAR was opposite: N165A mutation increases co-repressor recruitment in CAR but the corresponding mutation T233A decreases it in RARα, while the

others (V199A, Q329A and I330A) cannot interact with NCoR. This difference in the direction of effect is very likely due to constitutive interaction of CAR with co-activators and of RARα with co-repressors.

NR (Ref.)	Helix 11 residue	Helix 12 residue	Nature of interaction	Corresponding hCAR residues	Corresponding interaction in CAR?
CAR (this study)	Tyr326	Leu343	Hydrophobic		
	Ile330	Leu343	Hydrophobic		When SRC-1 present
	Tyr326	Ile346	Hydrophobic		When SRC-1 present
	Gln331	Ser348	H bond		
Nurr1 (Wang et al., 2003b)	Phe574	Phe592	Hydrophobic	Gln329, Leu347	No
	Lys577	Asp589	Salt bridge	His332, Gln334	No
ERRγ (Greschik et al., 2002)	Phe435	Phe450	Aromatic stacking	Ile330, Leu343	Yes
	Phe435	Leu454	Hydrophobic	Ile330, Cys347	No
RORβ (Stehlin et al., 2001)	His423	Tyr446	H bond	Tyr326, Leu343	Yes
	Leu427	Tyr446	Hydrophobic	Ile330, Leu343	Yes
LRH-1 (Sablin et al., 2003)	Leu536	Leu551	Hydrophobic	Ile330, Leu343	Yes
hCAR/mCAR models (Frank et al., 2004)	Tyr336	Cys347	H bond	Tyr326, Cys347	No
(Andersin et al., 2003)	Leu340	Leu353	Hydrophobic	Ile330, Leu343	Yes

Table 5.2: The interactions between helices 11 and 12 in constitutively active NRs.

5.3.2 Ligand Specificity of Human CAR

To our knowledge, this is the first report that addresses the LBP residues affecting the ligand specificity of human CAR. Table 5.2 summarises the selective effects of LBP residues on ligand specificity. Of importance are the roles of helix 3 residues in regulating the preference to TMPP, the requirement of several aromatic residues for clotrimazole recognition, and finally, the distinction between EE2 and ANDR by residues C202 and F243. A further support to our models came from the finding that nature of residue 243 regulates species-specific recognition of EE2. The single mutation of F243 to the corresponding leucine in mouse CAR converted EE2 to an activator. As expected from ligand-dependent effects of mutations described in Table 5.3, it should be emphasised that F243 probably is not the only regulator of human/mouse differences. Indeed, activation by the mouse CAR-selective agonist TCPOBOP seems to depend on the residue T350 (Jyrkkärinne et al., 2003).

We noticed intriguing differences when comparing the CAR LBP to that of

Activators			Inhibitors		
TMPP	Both	Clotrimazole	EE2	Both	ANDR
Ile164	Phe161	His203	Asn165	Met168	Val199
Asn165	Met340	Phe234	Phe243		Cys202
				His203[a]	
	Val199[a]	Phe238		Leu206[b]	
	Ile330[a]			Leu209[b]	
	Leu343[a]	Tyr326		Leu242[b]	

Table 5.3: Summary of selective ligand responses affected by specific CAR LBP residues. [a] Low basal activity and unresponsiveness to activators due to their participation in LBD contacts with helix 12. [b] Significant basal activity but any inhibition and responses to NCoR were lost.

VDR and PXR, the templates for our homology models. The residue corresponding to the key residue Y326 in CAR contacts the ligand in PXR (H407) and in VDR (H397) structures and their mutation appears to modulate ligand selectivity (Watkins et al., 2001; Watkins et al., 2003a) or eliminate ligand binding (Yamamoto et al., 2000; Väisänen et al., 2002), respectively. The residue S237 in VDR is important for both ligands 1a,25-dihydroxyvitamin D and lithocholic

acid (Yamamoto et al., 2000; Adachi et al., 2004) but the corresponding M168 in CAR shows selectivity to CAR inhibitors only. Similarly, the helix 5 residues are important to the ligand selectivity of VDR (S274, S278) (Adachi et al., 2004) but they have only a minor impact on CAR. Finally, the residues N165 or F243 that are crucial to CAR are not implicated in VDR or PXR ligand selectivity at all. This indicates that even though most of the LBP residues may correspond spatially, their functions differ tremendously.

In conclusion, we have developed well-defined homology models for human CAR. When coupled with extensive functional analysis, these models have helped us to suggest mechanisms that contribute to the high basal activity of human CAR, to identify residues that impart selectivity to the CAR ligand recognition and finally, have given the first explanations to the wide species differences in CAR ligand specificity.

Chapter 6

Homology Model Evaluation (III)

6.1 Introduction

As ligand-dependent transcription factors, nuclear receptors (NRs) are involved in many physiological processes including development, differentiation, reproduction and metabolism (Laudet and Gronemeyer, 2001). Binding of small ligands such as steroid and thyroid hormones, fatty and bile acids, as well as retinoids induces gene expression of specific enzymes through a complex cascade culminating in modulation of the histone acetylation status and transcription rate of target genes.

The structural assembly of NRs is modular consisting of a variable N terminus, a central DNA binding domain (DBD) and an C-terminal ligand binding domain (LBD) (Robinson-Rechavi et al., 2003). From X-ray data the three-dimensional architecture of the DBD as well as the LBD is known (Hard et al., 1990; Luisi et al., 1991; Bourguet et al., 1995). The LBD shows a conserved folding pattern consisting of 12 to 14 helices arranged in a three-layered helix sandwich and a beta-sheet composed of 2 to 5 strands (Wurtz et al., 1996). Ligands are bound in a mainly hydrophobic ligand binding pocket (LBP) located between the outer layers of the helix sandwich. Size and shape of the LBP varies between different NRs ranging from 220 Å^3 (ERR3) to 1300 Å^3 (PPARγ) (Nolte et al., 1998; Greschik et al., 2002). An exception is NURR1 which shows no LBP (Wang et al., 2003b). Besides the binding pocket, the LBD also carries the ligand dependent activation function 2 (AF-2) located on the C-terminal helix 12 (H12). The position and conformation of this helix is modulated by agonist and antag-

onist binding and it defines the type of co-regulator binding that are necessary for transcriptional regulation. Co-regulators encompass coactivators such as SRC-1 (steroid receptor coactivator-1) and corepressors (e.g. SMRT) which bind nearby helix 12 and which are responsible for the recruitment of additional proteins necessary for gene regulation. Agonists induce H12 to cover and seal the LBP. The emerging hydrophobic surface composed of helices H3, H4 and H12 enables coactivator binding via specific aliphatic amino acid residues. A "charge clamp" formed by highly conserved lysine (H3) and glutamate (H12) residues stabilises the coactivator binding by interacting with the coactivator's nuclear receptor interaction domain (NRID) [7]. Binding of antagonists or inverse agonists results in transformation of H12 into a disordered conformation disrupting the coactivator binding site, thus enabling corepressor recruitment (Brzozowski et al., 1997).

The nuclear hormone receptor superfamily comprises 48 members in human. For some NRs the function or endogenous ligands are yet unknown. Therefore these NRs are designated as "orphan" receptors (Evans, 1988). CAR belongs to the sub-family NR1I of the nuclear hormone receptor superfamily including the vitamin D (VDR, NR1I1) and the pregnane X receptor (PXR, NR1I2). CAR is part of the metabolic defense in humans. In conjunction with the closely related PXR, CAR regulates the expression of metabolising enzymes upon xenobiotic stress (Wei et al., 2000). Both share a variety of ligands and regulate an overlapping set of target genes (Maglich et al., 2002; Moore et al., 2000). CAR has been shown to activate expression of cytochrome P450s (CYP3A4, 2B10, 2C9) as well as conjugating enzymes and transporters (Maglich et al., 2002; Honkakoski et al., 1998b; Goodwin et al., 2002; Ferguson et al., 2002b; Burk et al., 2005; Maher et al., 2005). Moreover, CAR has been recently found to play a significant role in bilirubin clearance and bile acid detoxification (Huang et al., 2003; Guo et al., 2003; Goodwin and Moore, 2004). In contrast to other NRs, CAR possesses a constitutive activity in absence of any ligand which can be repressed by ligands such as androstenol and androstanol (Baes et al., 1994; Forman et al., 1998). These testosterone metabolites have been identified as endogenous ligands for CAR, thereby rendering CAR an "adopted orphan receptor" (Honkakoski et al., 2004). Only few activators of CAR are known so far, including 6-(4-chlorophenyl)imidazo[2,1-b][1,3]thiazole-5-carbaldehyde *O*-(3,4-dichlorobenzyl)oxime (CITCO), 5β-pregnane-3,20-dione,

tri-p-methylphenyl phosphate (TMPP), clotrimazole and artemisinin as well as some HMG-CoA reductase inhibitors (Moore et al., 2000; Maglich et al., 2003; Honkakoski et al., 2004; Mäkinen et al., 2002; Burk et al., 2005; Kobayashi et al., 2005).

The analysis of the underlying mechanism of the CAR constitutive activity has been the objective of several studies (Frank et al., 2004; Andersin et al., 2003; Dussault et al., 2002; Windshügel et al., 2005). In order to elucidate the structural basis for constitutive activity, we generated a CAR homology model on the basis of the X-ray structures of the closely related receptors VDR and PXR. Based on MD simulations of CAR and receptor mutants as well as considering experimental site-directed mutagenesis data, we proposed a structural model explaining the constitutive activity. We observed that the basal activity of CAR is mainly dependent on van der Waals (vdW) and hydrophobic interactions between H12 and apolar residues of the LBD (Jyrkkärinne et al., 2005). Amino acids V199 (H5), Y326 (H10) and I330 (H11) were observed to interact with L343 and I346 of the helix 12. Almost simultaneously with our homology model, the X-ray structures of human (PDB code: 1XV9, 1XVP) and mouse CAR (PDB code: 1XLS, 1XNX) in complex with structurally diverse ligands were reported (Xu et al., 2004; Suino et al., 2004; Shan et al., 2004). Based on these crystal data also a hypothesis for the constitutive activity has been proposed by Xu et al. (Xu et al., 2004). VdW interactions between LBD and H12, a short helix between H11 and H12 as well as a salt bridge formed by a conserved lysine residue and the C terminus were proposed to be the key elements for maintaining the high basal activity of CAR (Xu et al., 2004; Suino et al., 2004; Moore, 2005). The available X-ray data for human CAR afforded us evaluate our homology model and the suggested activation mechanism. Moreover, MD simulations of ligand-free CAR X-ray structures were carried out to evaluate the putative mechanism of constitutive activity suggested by Xu et al. (Xu et al., 2004).

6.2 Results

The structural basis of the constitutive activity of CAR has been investigated during the last years (Frank et al., 2004; Dussault et al., 2002). From site directed mutagenesis studies, single amino acids have been detected to be essential for

the basal activity of CAR. To further unravel the structural mechanism of constitutive activity, we generated a homology model of the human CAR LBD using the X-ray structures of the related PXR and VDR as template (Windshügel et al., 2005). By means of MD simulations we proposed that the CAR basal activity is mainly achieved by specific vdW and hydrophobic interactions between helix 12 and the LBD. Particularly, Y326 emerged as a central amino acid for maintaining the basal activity. Stabilised by surrounding aliphatic and aromatic amino acid residues and a hydrogen bond with N165, the Y326 side chain shows permanent contacts with H12 that resemble agonist-H12 interactions in other NRs. Therefore, we inferred that Y326 mimics a bound receptor agonist keeping H12 in the active position even in absence of any ligand. MD simulations of CAR mutants and site-directed mutagenesis data further supported our suggested hypothesis (Jyrkkärinne et al., 2005). The now available X-ray structures of the CAR LBD gave us the opportunity to assess the quality of the homology modelling and refinement procedure.
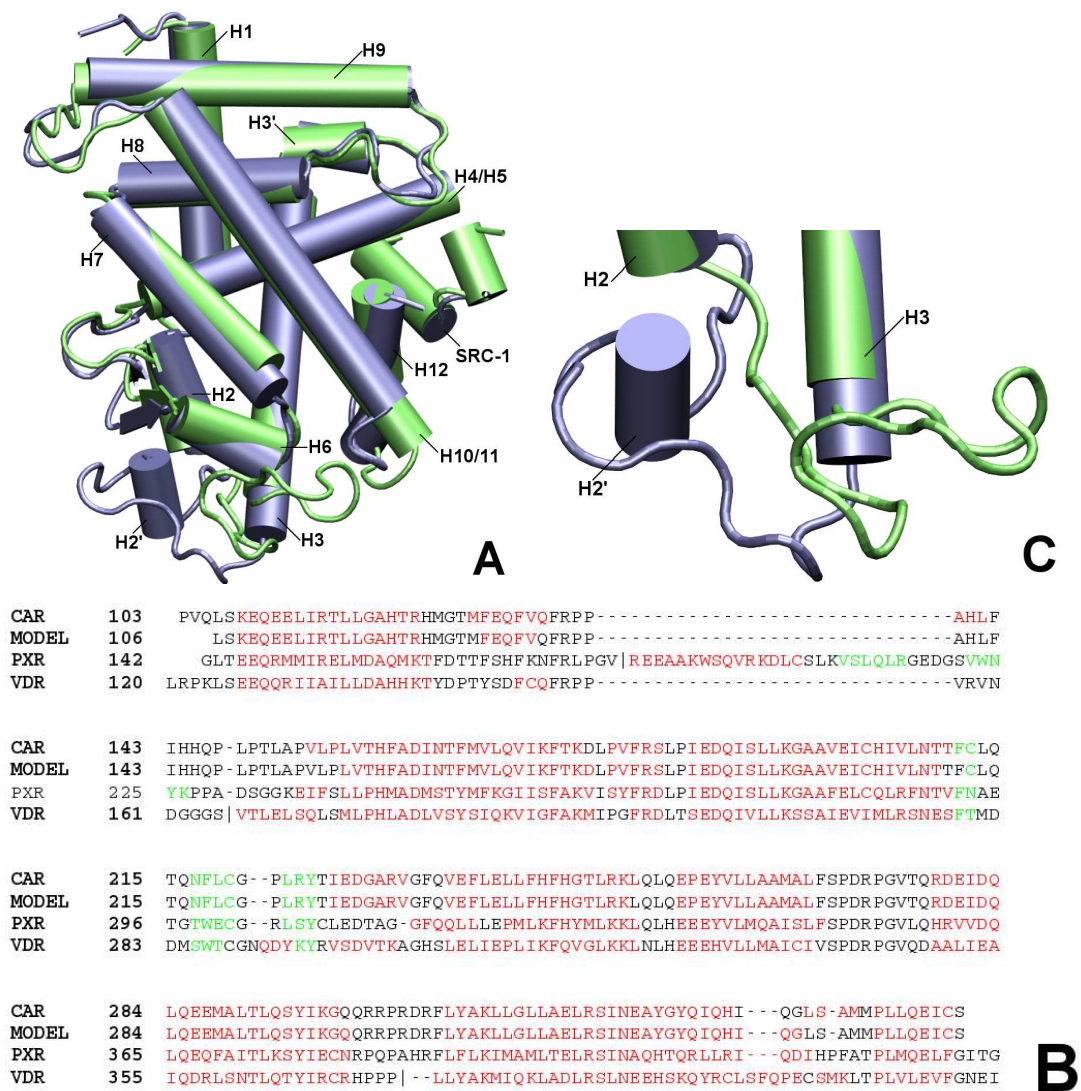
6.2.1 Quality of the Homology Model

To investigate how well the modelled structure matches the X-ray data, the CAR/SRC-1 homology model and crystal structures were superimposed on their backbone atoms (Fig. 6.1*A*). The overall arrangement of helices and loops in the model is in good agreement with the corresponding elements in the X-ray structures (see Fig. 6.1*A*) and the alignment in Fig. 6.1*B*. Solely helices H3 and H10/11 show an additional turn in the homology model. The most strikingly difference is observed in the region connecting H2 and H3. Here the X-ray structures possess an additional helix (H2′) whereas the homology model contains a flexible loop (Fig. 6.1*C*). Also the orientation of the H2-H3 loop is different compared to the corresponding element in the X-ray structures.

To assess the model accuracy, the RMSD values between superimposed model and crystallographic structures were calculated. RMSD values for the backbone atoms were found to vary between 3.4 and 4.4 Å suggesting a suboptimal modelling quality (see Table 6.1 for details). However, from a visual inspection a good overall agreement of secondary structural elements of the homology model and the X-ray structures is observed (Fig. 6.1*A*). In fact, high RMSD values originate mainly from large deviations in the H2-H3 region. Excluding this segment (amino acids 139 to 153) from the measurement, the RMSD dropped significantly to values between 1.8 and 2.2 Å (Table 6.1). Since the H2′-H3 loop is located at the protein surface and the residues located therein neither contribute to the formation of the LBP nor to the dimerisation interface it is suggested that the false prediction in this region has no major influence on the general reliability of the CAR model. This is also supported by the closely related VDR. It contains a significantly larger H2-H3 region which does not affect the main receptor function (Rochel et al., 2001).

To determine the stereochemical quality of the homology model we used two approaches, ProSa2003 and Profiles-3D (Table 6.1) (Sippl, 1993; Bowie et al., 1991). Combined Z-scores calculated with ProSa show relatively high energies for the initial model. However, they were significantly decreased in the subsequent minimisation and MD simulation. A further decrease was observed when the CAR model was simulated in the ligand-complexed form. The resulting Z-score is close to the values observed for the CAR X-ray structures (-9.70 to -10.05). In agreement with the ProSa result also the Profiles-3D approach shows

```
CAR    103    PVQLSKEQEELIRTLLGAHTRHMGTMFEQFVQFRPP------------------------------AHLF
MODEL  106      LSKEQEELIRTLLGAHTRHMGTMFEQFVQFRPP------------------------------AHLF
PXR    142      GLTEEQRMMIRELMDAQMKTFDTTFSHFKNFRLPGV|REEAAKWSQVRKDLCSLKVSLQLRGEDGSVWN
VDR    120    LRPKLSEEQQRIIAILLDAHHKTYDPTYSDFCQFRPP------------------------------VRVN


CAR    143    IHHQP-LPTLAPVLPLVTHFADINTFMVLQVIKFTKDLPVFRSLPIEDQISLLKGAAVEICHIVLNTTFCLQ
MODEL  143    IHHQP-LPTLAPVLPLVTHFADINTFMVLQVIKFTKDLPVFRSLPIEDQISLLKGAAVEICHIVLNTTFCLQ
PXR    225    YKPPA-DSGGKEIFSLLPHMADMSTYMFKGIISFAKVISYFRDLPIEDQISLLKGAAFELCQLRFNTVFNAE
VDR    161    DGGGS|VTLELSQLSMLPHLADLVSYSIQKVIGFAKMIPGFRDLTSEDQIVLLKSSAIEVIMLRSNESFTMD


CAR    215    TQNFLCG--PLRYTIEDGARVGFQVEFLELLFHFHGTLRKLQLQEPEYVLLAAMALFSPDRPGVTQRDEIDQ
MODEL  215    TQNFLCG--PLRYTIEDGARVGFQVEFLELLFHFHGTLRKLQLQEPEYVLLAAMALFSPDRPGVTQRDEIDQ
PXR    296    TGTWECG--RLSYCLEDTAG-GFQQLLLEPMLKFHYMLKKLQLHEEEYVLMQAISLFSPDRPGVLQHRVVDQ
VDR    283    DMSWTCGNQDYKYRVSDVTKAGHSLELIEPLIKFQVGLKKLNLHEEEHVLLMAICIVSPDRPGVQDAALIEA


CAR    284    LQEEMALTLQSYIKGQQRRPRDRFLYAKLLGLLAELRSINEAYGYQIQHI---QGLS-AMMPLLQEICS
MODEL  284    LQEEMALTLQSYIKGQQRRPRDRFLYAKLLGLLAELRSINEAYGYQIQHI---QGLS-AMMPLLQEICS
PXR    365    LQEQFAITLKSYIECNRPQPAHRFLFLKIMAMLTELRSINAQHTQRLLRI---QDIHPFATPLMQELFGITG
VDR    355    IQDRLSNTLQTYIRCRHPPP|--LLYAKMIQKLADLRSLNEEHSKQYRCLSFQPECSMKLTPLVLEVFGNEI
```

Figure 6.1: Comparison of the CAR homology model and X-ray structures. **(A)** Superimposition of CAR/SRC-1 X-ray structure 1XVP (chains D and H, blue) and the CAR/SRC-1 homology model after free MD simulation (green) on backbone atoms (N-CA-C). α-helices are depicted by cylinders, β-sheets by arrows, random-coil by tube. **(B)** Sequence alignment and comparison of secondary structural elements observed for the X-ray structures (CAR, PXR, VDR) and CAR homology model (MODEL). Helices are marked red whereas β-strands are marked green. **(C)** Detailed view of the region located between helices H2 and H3 where model and X-ray structures show the largest deviations.

an improvement of the model quality during the refinement process (profiles-3D scores calculated for the crystal structures range from 112 to 119).

To further evaluate the accuracy of the CAR model we compared the side

	Un-liganded CAR model (Different stages of model refinement)				CAR-ligand complexes (Docking with subsequent MD)			
Model	1	2	3	4	CLOT	TMPP	CITCO	PREG
RMSD 1	3.4	3.4	3.6	3.7	4.4	3.4	3.8	3.9
RMSD 2	1.8	1.8	1.9	1.9	2.0	2.0	1.9	2.2
Z-Score	-8.69	-9.37	-9.68	-9.44	-10.05	-9.49	-9.78	-9.63
Prof.-3D	112	111	119	116	112	115	118	115

Table 6.1: Stereochemical parameters of the CAR homology model and the model-ligand complexes. RMSD values (Å) are average values calculated between the X-ray structures 1XV9/1XVP and the model structures at different stages of the refinement (model "1" denotes the original homology model without any refinement, model "2" is obtained after energy minimisation, model "3" indicates the structure resulting from the equilibration MD and model "4" is a representative frame of the free MD; CLOT (Clotrimazole, TMPP, CITCO and PREG (5β-pregnanedione) denotes the corresponding CAR-ligand complexes derived from MD simulations). "RMSD 1" includes all backbone atoms whereas "RMSD 2" is calculated for backbone atoms excluding the segment 139 to 153.

chain orientations of the amino acids forming the LBP in the model and the X-ray structures. For this purpose 29 amino acids contributing to the accessible surface of the LBP were analysed. CAR model and X-ray structures were superimposed and the RMSD was calculated for backbone atoms (between 1.0 and 1.5 Å) and all heavy atoms (between 1.5 and 1.9 Å, depending on the stage of refinement) (Table 6.2). Figure 6.2 shows the superimposed binding pockets of the homology model and the X-ray structure 1XVP illustrating the high accuracy of the CAR model. For comparison, the RMSD values for the individual chains of the CAR X-ray structures lie in the range between 0.4 and 0.6 Å.

Besides the calculation of RMSD values also the χ_1 dihedral angles of LBP side chains were analysed and compared with those observed in the X-ray structures (Table 6.3). Dihedrals were considered as correctly predicted when the deviation compared to the values in the crystal structures was less than 10 degrees. According to this criteria, about two third of all dihedral angles were correctly predicted. The number of outliers range from 12 in the raw model to 9 in a representative frame of the free MD simulation.

	Un-liganded CAR model (Different stages of model refinement)				CAR-ligand complexes (Docking with subsequent MD)			
Model	1	2	3	4	CLOT	TMPP	CITCO	PREG
Backb.	1.1	1.1	1.1	1.0	1.2	1.5	1.4	1.4
Heavy at.	1.7	1.7	1.5	1.7	1.8	1.8	1.9	1.7

Table 6.2: Accuracy of LBP modelling. RMSD values (Å) are average values calculated between the X-ray structures 1XV9/1XVP and the model structures. For abbreviations see Table 6.1.
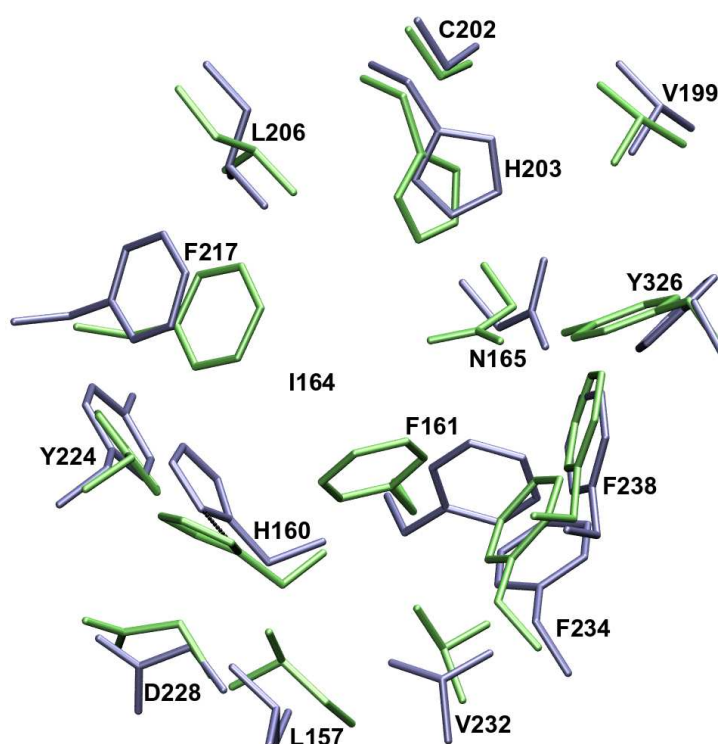
Figure 6.2: Accuracy of LBP modelling. Side chain conformations of several amino acids lining the LBP. Residues of the X-ray structure (1XVP) are coloured in blue whereas amino acids of the homology model are shown in green.

In order to identify amino acids important for the constitutive activity, a variety of CAR mutants have been experimentally studied (Frank et al., 2004; Andersin et al., 2003; Dussault et al., 2002; Windshügel et al., 2005; Feng et al., 1998). In order to elucidate the structural basis for the inactivity of some of these

Model	1	2	3	4	CLOT	TMPP	CITCO	PREG
F132	-71	-81	-71	-61	-62	**49**	**-56**	-71
L157	**-66**	**-79**	**-156**	**-155**	**-63**	**-150**	-175	**-156**
H160	174	-176	-176	178	**-158**	179	**-157**	-176
F161	**-55**	**-66**	**-144**	-177	-175	**-126**	-176	**-144**
I164	**-67**	**-46**	**-44**	**-50**	**-175**	**-58**	**-56**	**-44**
N165	**-74**	**-70**	-175	-167	-144	**-66**	-161	-175
M168	-70	-71	-73	-68	-84	-74	**-155**	-73
V169	172	**-154**	**-62**	177	-174	-175	-174	**-62**
C202	**-180**	**-175**	-69	**-177**	**177**	-75	-72	-69
H203	-70	-67	-79	-77	-69	-81	-81	-79
L206	**-71**	**-79**	-168	**-46**	-162	**-57**	-164	-168
F217	-70	-69	-63	**-51**	**-87**	-70	-62	-63
C219	-53	-55	-58	-51	-55	-56	**-172**	-58
Y224	-46	-55	-59	-62	-67	**-68**	-63	-59
T225	63	70	71	76	65	68	62	71
I226	**-68**	**-57**	**-58**	-171	**-55**	-169	**-74**	**-58**
D228	-66	-55	-53	-57	-65	-69	-66	-53
G229	-	-	-	-	-	-	-	-
V232	**-61**	**54**	**-66**	**69**	**67**	**63**	**64**	**66**
F234	-67	-67	-62	-65	-71	-71	-66	-62
F238	178	174	175	172	158	-177	177	175
L239	-72	-64	-62	-72	-65	-69	-72	-62
L242	**-146**	**-93**	**-55**	**-67**	-174	**-59**	-163	**-55**
F243	**-178**	**-175**	-87	**-68**	**-70**	**-72**	**-160**	-87
H246	-72	-67	-85	-89	-88	-74	-84	-85
I322	**-67**	-68	**62**	-68	**59**	**-59**	**-66**	**-63**
Y326	179	-178	166	180	176	**-163**	174	166
L343	**-178**	**-170**	**-57**	**-170**	**-171**	**-67**	-78	-74
FALSE	**12**	**12**	**10**	**9**	**10**	**13**	**9**	**9**

Table 6.3: LBP dihedrals. χ_1 dihedrals for residues lining the LBP. Values deviating more than 10 degrees from minimum and maximum χ_1 dihedrals in CAR X-ray structures are considered as false predicted and are highlighted by bold face/underline (abbreviations see Table 6.1).

CAR mutants we also generated models of mutated CAR and examined them by means of MD simulations (Windshügel et al., 2005). The analysis of the MD simulations showed that a variety of amino acid residues are important to maintain the interaction between the key residue Y326 and H12. For example, a reorientation of Y326 and a loss of the interaction with H12 was observed in the simulation of the F238A as well as the F243A mutant. Both mutated receptors show no or a reduced basal activity in cell-based reporter assays which is in agreement with our prediction. In the present study we analysed whether the results obtained for the CAR model could be reproduced for the CAR X-ray structures. For this purpose the receptor mutants F238A and F243A modelled from the CAR X-ray structures were analysed by means of MD simulations (5.25 ns). Consistent with the observation derived from the simulations of the homology model, the MD simulation of the F238A mutant resulted in a reorientation of the Y326 side chain (Fig. 6.3*A*). In contrast to the simulation of the F238A homology model, vdW contacts between Y326 and L343 (H12) were maintained throughout the whole simulation due to a reorientation of L343. However, the conformation of Y326 in the mutant is much less defined as in wild-type CAR indicated by a larger overall flexibility (Fig. 6.3*B*) suggesting a less stable interaction between LBD and H12. For F243A no reorientation of Y326 could be observed although the side chain flexibility is increased similar to the F238A mutant (Fig. 6.3*B*). During the MD simulations of the F243A homology model the hydrogen bond between H160 and Y224 was disrupted provoking a reorientation of the Y224 side chain into the LBP, an observation that was also observed for the F243A X-ray structure (data not shown).
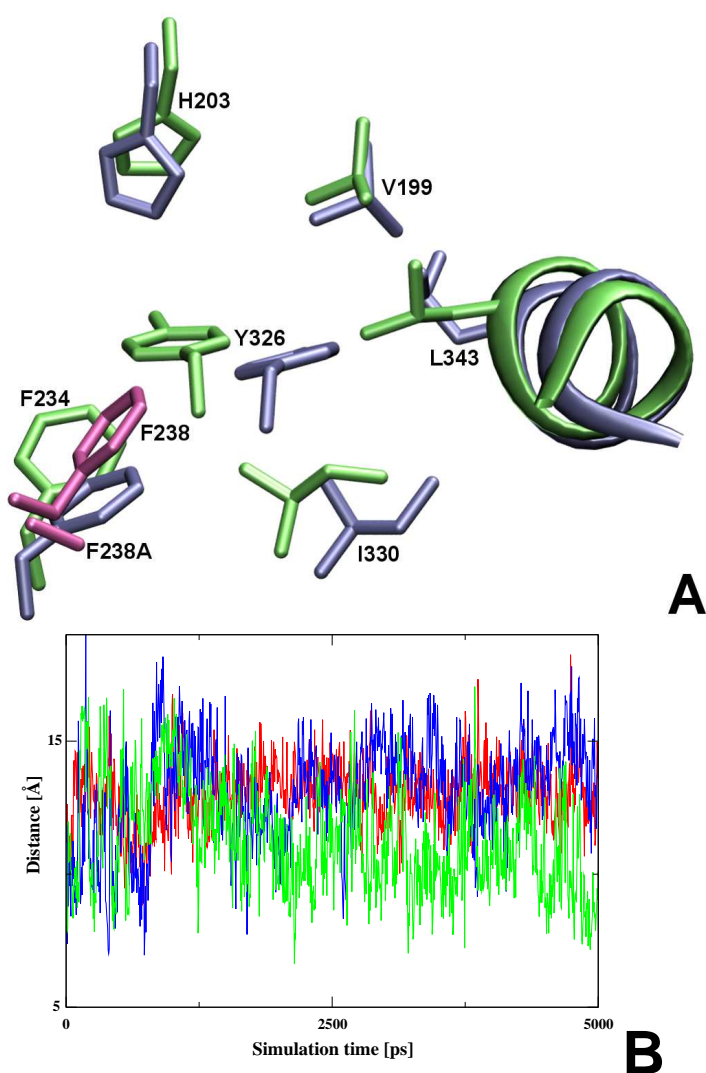
Figure 6.3: Analysis of CAR receptor mutants. **(A)** Superimposition of CAR X-ray structure 1XVP_DH (blue) and the mutant F238A (green). Side chains of F238 and F238A are shown in pink. Helix 12 is indicated by ribbon. **(B)** Course of distance between Y326 (atom CZ) and L343 (atom CA) during the MD simulation for wild-type CAR (red) and mutants F238A (blue) and F243A (green).

6.2.2 Reproducing Ligand Binding Modes

We analysed whether the docking programme GOLD is able to correctly reproduce the binding mode of the two co-crystallised ligands 5β-pregnane-dione and CITCO. For this purpose the ligands were first docked into the X-ray structures in order to determine how well the binding modes could be reproduced. The conformation obtained by docking is close to the experimentally

observed position indicated by an RMSD of 1.48 Å for CITCO and 0.65 Å for
5β-pregnanedione. The position of CITCO docked in the homology model is
almost identical to that in the crystal structures indicated by an RMSD of 1.97
Å (Fig. 6.4*A*). A subsequent MD simulation (2.25 ns) resulted in only slightly
movement of CITCO within the LBP (data not shown).
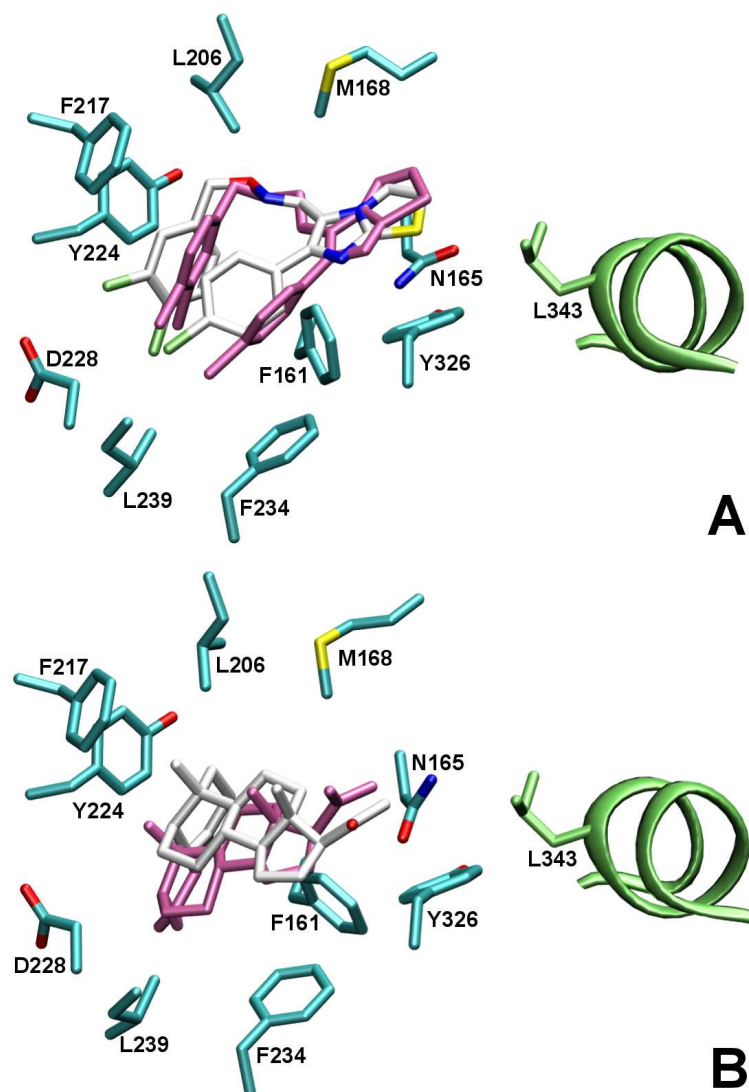
Figure 6.4: Docking results. Position of **(A)** CITCO and **(B)** 5β-pregnanedione docked
into the homology model in comparison with the positions observed in the X-ray
structures. Carbon atoms of the co-crystallized ligands are coloured white. Docking
poses are shown in pink. Helix 12 is shown as green ribbon.

The experimentally determined position of 5β-pregnanedione could not be re-produced well by the GOLD docking (RMSD 4.0 Å) due to the deviated side chain conformation of Y224 in the homology model which prevents a correct placement of the ligand. To analyse whether the presence of 5β-pregnanedione will change the Y224 conformation, an MD simulation (2.25 ns) of the CAR model in complex with the docked ligand was carried out. The simulation showed a reorientation of Y224, thus leading to a orientation of 5β-pregnanedione that is consistent with the crystal structure (RMSD 2.0 Å) (Fig. 6.4*B*).

Reorientation of side chains upon ligand binding has also been observed during the MD simulation of the CAR model complexed with either clotrimazole or TMPP (Windshügel et al., 2005). Two amino acids residues of the LBP emerged as highly flexible: F161 as well as Y224 were found to adopt different conformations during the simulations. The conformational flexibility of side chains facing toward the binding pocket allows the adoption to structurally diverse ligands. MD simulations of the CAR X-ray structure in complex with 5β-pregnanedione did not show any significant side chain reorientation of F161 whereas during the simulation of ligand-free CAR X-ray F161 adopts different conformations confirming the observation from the homology model (data not shown). The other amino acids lining the LBP did not change their orientation during the MD simulations significantly.

6.2.3   The Basis for Constitutive Activity

Based on the CAR crystal structures a hypothesis for the structural basis of the constitutive activity has been deduced by Xu and coworkers (Xu et al., 2004). The authors stated that CAR basal activity is mainly achieved by three structural features (Fig. 6.5*A*):

First, a hydrophobic barrier formed by residues F161, N165, F234 and Y326 is interacting with H12 keeping it in the active conformation. Second, an additional helix (termed as "helix X") located between H11 and H12 is supposed to orient H12 in its active position and additionally contacts the hydrophobic barrier. Third, the missing C-terminal extension of H12 - observed in other NRs - is considered to allow the formation of a salt bridge between K195 on helix 5 and the C-terminal free carboxylate further stabilising the active position of H12.
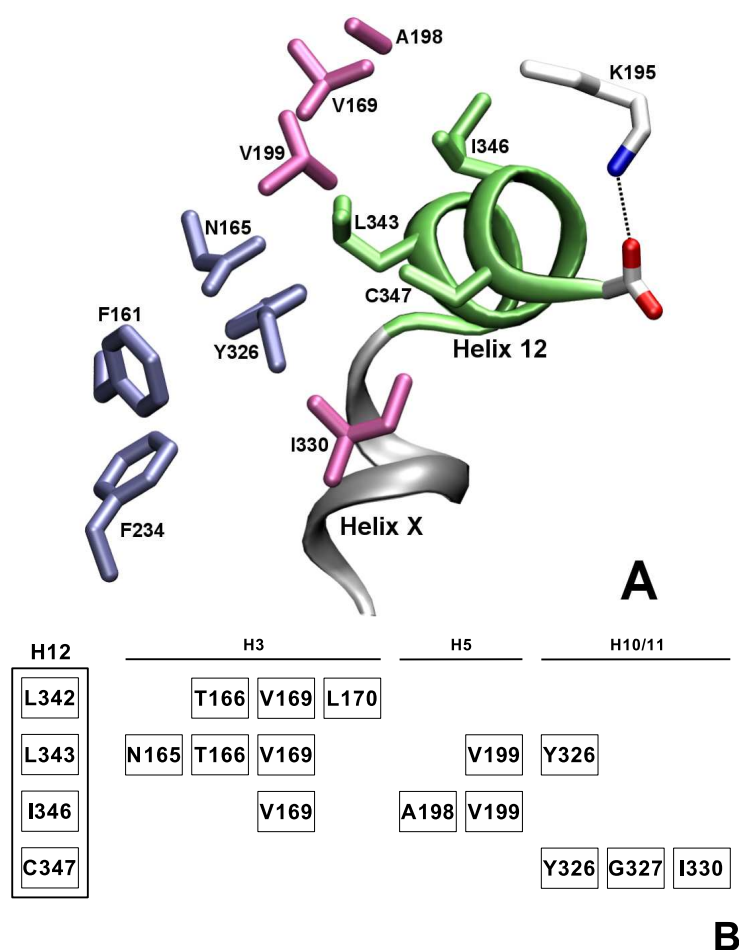
Figure 6.5: Interaction between the LBD and helix H12. **(A)** The proposed hydrophobic barrier formed by residues N165, F161, F234 and Y326 is coloured in blue. Other residues involved in vdW interactions between LBD and H12 are coloured pink. "Helix X" is shown in grey. The salt bridge connecting H5 and the C terminus (S348) is depicted by dotted lines. H12 is shown in green. **(B)** Schematic view of amino acid residues contributing to vdW interactions between H12 and the residues of the LBD. For each amino acid residue located on H12 the corresponding interaction partner on helices 3, 5 and 10/11 is given.

We analysed the homology model if there is an agreement with the features detected in the X-ray structure. Furthermore, MD simulations of the CAR X-ray structures were used to evaluate the structural features proposed to contribute to basal activity.

A hydrophobic barrier is also existent in the homology model including Y326 as key element. Additionally, other vdW interactions between LBD and H12 can be observed in the model involving amino acids V199, I330 and I346. Further

amino acids which contribute to the interaction of the LBD with H12 can be observed in the X-ray structures as well in the homology model (Fig. 6.5*A*). The interaction map in Figure 6.5*B* schematically depicts all vdW contacts between amino acid residues of H12 and their corresponding interaction partners located on H3, H5 and H10/11. Among those the influence on constitutive activity has already been confirmed for most residues by site-directed mutagenesis (Frank et al., 2004; Jyrkkärinne et al., 2005).

  The salt bridge between K195 and S348 detected in the CAR X-ray structures (Fig. 6.5*A*) is not observed in the homology model. Instead, K195 makes a hydrogen bond to H687 located on the SRC-1 NRID whereas the C terminus (S348) interacts with Q331 on H11. The importance of Q331 for the constitutive activity has already been revealed by site-directed mutagenesis (Jyrkkärinne et al., 2005). Mutation of Q331 to alanine results in a decrease of the basal activity by about 70 percent. To elucidate the stability of the salt bridge in CAR X-ray structures, MD simulations (5.25 ns) of CAR/SRC-1 complexes were carried out. Ligands were removed in order to eliminate any effect on the LBD. During the simulation the salt bridge between K195 and the C terminus (S348) was found to be not stable. The distance plot clearly indicates an early separation of the two amino acids (Fig. 6.6*A*). Instead, similar to the CAR/SRC-1 homology model, K195 is involved in a hydrogen bond with H687 located on the SRC-1 NRID (Fig. 6.6*B*). Additionally, a transient hydrogen bond formation was observed between Q331 and S348 (Fig. 6.6*C*).
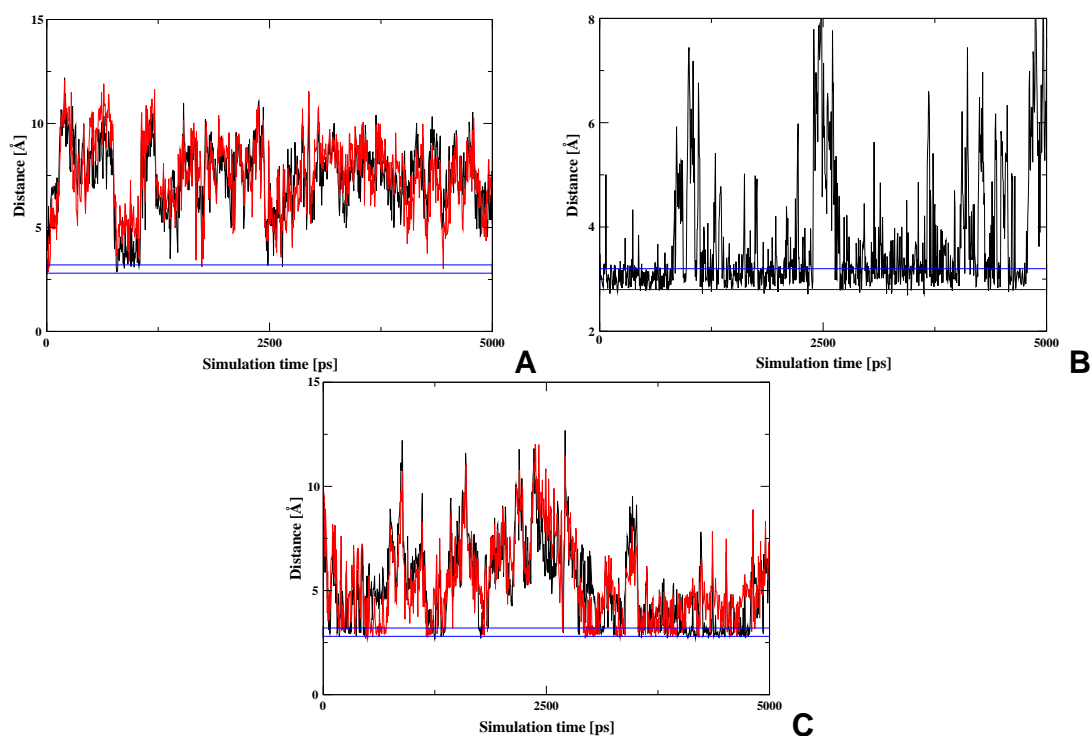
Figure 6.6: Stability of the salt bridge K195-S348. Distance plots for amino acid pairs during the MD simulation. Blue lines define the area of optimum hydrogen bond and salt bridge distances, respectively. **(A)** Course of distance between the K195 ammonium group (atom NZ) and the C terminus. Plots are shown for both oxygens of the carboxy-terminal group (atom O1: red; atom O2: black). **(B)** Distance plot for the ammonium group of K195 (atom NZ) and H687 of the coactivator SRC-1 (atom NE). **(C)** Transient hydrogen bond formation between Q331 (atom NE2) and both oxygens of the C terminus (atom O1: red; atom O2: black).

6.2.4 The Role of Helix X

Since CAR has been crystallised only in complex with agonistss and not in un-liganded form it is not clear whether the so-called "helix X" is an essential feature for maintaining the basal activity or rather formed upon agonist binding. Therefore we investigated the stability of the "helix X" by means of MD simulations of un-liganded CAR (without coactivator NRID). During the simulation time of 10 ns the "helix X" did not show any unfolding. This is not an absolute evidence for the stability of "helix X" since helix unfolding generally occurs on a much longer time scale. However, visual inspection of a variety of NR X-ray structures revealed that the "helix X" can also be observed among other nuclear receptors.

Altogether 26 X-ray structures in the Protein Databank show a comparable helical element (Table 6.4). In case of VDR, which shows no basal activity, all crystal structures complexed with an agonist possess this "helix X" (Fig. 6.7*A*). The same holds true for the retinoid acid-related orphan receptor α and β (RORα/β). It has been speculated that in CAR the single amino acid between "helix X" and H12 limits the conformational freedom of H12 whereas in VDR this segment comprises two amino acids thus resulting in a larger flexibility of H12. In contrast to VDR, the segment in constitutively active ROR consists of a phenylalanine and a proline residue that has been proposed to allow only limited flexibility. However, we found that in VDR the flexibility of the segment (L414, T415) is restricted by hydrogen bonds to R154 (located on the loop connecting H2 and H3) and S235 (H3).

CAR	VDR	TRα/β	GR	RORα/β
1XV9,1XVP, 1XLS	1DB1, 1IE8, 1IE9, 1RJK, 1RK3, 1RKG, 1RKH, 1S19, 1TXI	1N46, 1NAV, 1NAX, 1NQ0, 1NQ2, 1R6G 1Y0X	1M2Z, 1P93	1K4W, 1N4H, 1N83, 1NQ7, 1S0X

Table 6.4: Occurrences of helix 11' in NR X-ray structures (PDB entry) containing a helix 11' (helix X) motif.

Another important observation is the fact that the thyroid receptor (TR) shows variability in the H11-H12 region. Among the 13 X-ray structures inspected, seven show a "helix X"-like element whereas six do not possess an helical segment (Fig. 6.7*B*). As detected for ROR, the loop connecting "helix X" and H12 in TR comprises a phenylalanine and proline residue. Additionally, the C-terminal amino acid of the "helix X" in TR (L450) makes a hydrogen bond to T273 located on helix H3.

Another example for the observation of the "helix X" is the glucocorticoid receptor (GR, 1P93). When GR is co-crystallised with dexamethasone two of the four monomers in the PDB structure show an "helix X" whereas the others do not.
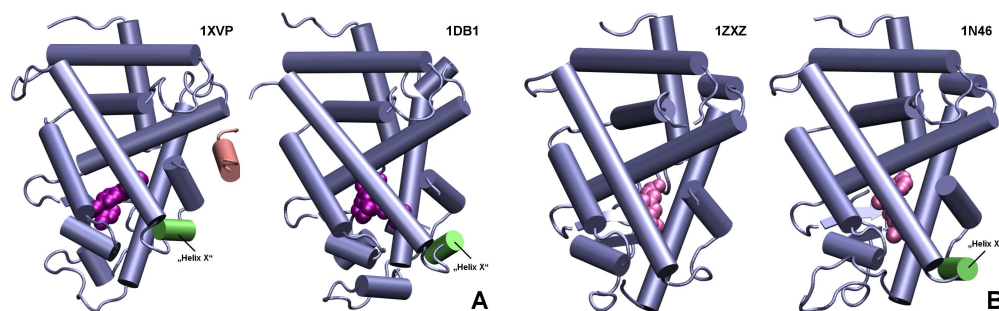
Figure 6.7: Observation of the "helix X" motif in other NRs. **(A)** Result of the visual analysis for "helix X"-like motifs in other NRs. Besides CAR (displayed PDB entry 1XVP) also all VDR X-ray structures crystallised so far contain this helical element (displayed PDB entry 1DB1). "Helix X" is coloured green and the SRC-1 coactivator NRID is shown in pink. **(B)** Comparison of TRβ X-ray structures with structurally diverse ligands. In complex with T3 (Triiodothyronine) no helical structure between H10/11 and H12 is obseved (PDB entry 1ZXZ), whereas in complex with a thyromimetic compound a "helix X"-like structure (coloured green) is found (PDB entry 1N46). Bound ligands are coloured magenta.

6.3 Discussion

In the last few years several homology models of CAR have been published by us and others. The models were used to analyse the structural features of the constitutive activity or the ligand binding (Maglich et al., 2003; Frank et al., 2004; Dussault et al., 2002; Windshügel et al., 2005; Xiao et al., 2002; Jacobs et al., 2003). Since the coordinates of the individual models have not been published a direct structural comparison is difficult. However, we tried to make a general comparison of the different CAR models (see also Table 6.5).

From the published data and our own results it can be stated that CAR models based on the single template PXR generally result in an overall unfavourable architecture due to the structural deviations of PXR from the common NR topology (Dussault et al., 2002). Models based on two templates represent a more sophisticated approach in which the VDR X-ray structure is used to model the problematic regions (Windshügel et al., 2005; Xiao et al., 2002). Refinement procedures applied during the modelling process usually result in favourable stereochemical parameters (e.g. percentage of amino acids in the most favoured region of the Ramachandran plot). This was also observed for the different stages during the refinement process of our models. However, it must be

stated that this is not a measure for the agreement with crystallographic data. Another important feature of the homology models, especially if they are used for docking or structure-based drug design, is the accuracy of the binding pocket. For example, the CAR models generated by Dussault et al. and Xiao et al. were found to contain huge cavities whereas in our model the LBP is about 700 Å^3 smaller and much closer to the actual pocket size observed in the CAR X-ray structures (see Table 6.5 for details). The large difference in LBP volume is a consequence of the usage of the PXR structure as modelling template (cavity size PXR 1280-1544 Å^3, depending on the ligand) as well as the manual adjustment of side chain conformations (Dussault et al., 2002; Xiao et al., 2002).

	Dussault et al.	Xiao et al.	Windshügel et al.	CAR X-ray	
Organism	Mouse	Human	Human	Human	Mouse
Modelling Template(s)	PXR	PXR/VDR	PXR/VDR		
Ram. plot (% most fav.)	82.6	90	86.4	83.3-86.4	83.3-84.1
LBP size [Å^3]	1150	1170	480	675	525

Table 6.5: Comparison of CAR homology models with available X-ray data.

The release of the CAR X-ray structures gave us the possibility for a detailed structural comparison of the homology model with the experimental data in order to evaluate the quality of the model . Based on the comparison we could show that the applied modelling procedure yielded a reliable CAR model. Not only the assignment and position of the secondary structural elements are in close agreement with the X-ray data but also, more importantly, the side chain conformations of residues constituting the LBP were predicted with high accuracy. This is reflected by low RMSD values and the fraction of correctly predicted χ_1 dihedral angles. The good reproduction of the binding mode for the two agonists 5β-pregnanedione and CITCO further emphasises the high degree of consistency between model and crystallographic data.
Comparing the intermediate stages of the refinement process with the available

crystallographic data revealed a slight improvement of the model quality. Considering all calculated parameters, the CAR model derived from the constrained MD simulation is closest to the X-ray structures whereas free MD simulations decreased the model accuracy. This observation is in agreement with results obtained by Flohil et al. (Flohil et al., 2002). However, it has to be considered that CAR specific features such as the hydrogen bonds connecting N165 and Y326 as well as Q331 and S348 are not formed during energy minimisation or constrained MD, respectively.

MD simulations of CAR-ligand complexes slightly impaired the LBP accuracy in terms of RMSD, whereas the stereochemical parameters are comparable to the other models (Profiles-3D scores) or were slightly improved (ProSa Z-score). Results from the measurement of the χ_1 dihedral suggest that ligand binding induces a reorientation of LBP side chains in order to facilitate optimal protein-ligand interactions. This is indicated by a higher fraction of correctly predicted χ_1 dihedrals in case of CITCO and 5β-pregnanedione compared to the result obtained for TMPP and clotrimazole. However, an increased number of χ_1 outliers for TMPP does not allow to infer a worse LBP quality. Compared to CITCO and 5β-pregnanedione, TMPP is structurally different and therefore it can also be expected for this ligand that binding induces an adoption of the side chains. The mutational studies carried out for CAR support the hypothesis that the structural basis for the constitutive activity is mainly dependent on vdW and hydrophobic interactions between H12 and the LBD. Several vdW contacts between H12 and its interaction partners H3, H5 as well as H10/11 have also been detected in the CAR X-ray structures, where H12 is in contact with a hydrophobic segment consisting of altogether nine amino acid residues. Furthermore, the hydrogen bond between N165 and Y326 sustains the structural integrity of the LBD-H12 interface. The importance of several amino acids for constitutive activity has already been confirmed by site-directed mutagenesis studies by us and others (Frank et al., 2004; Jyrkkärinne et al., 2005). Any alanine mutant reduced the basal activity dramatically. Another amino acid residue which is involved in vdW interactions with H12, V169, is a promising target for further experimental studies. We expect that the mutation of V169 to an alanine residue will also significantly reduce the constitutive activity of CAR.

Amino acids located on H12 have been already analysed by site-directed mutagenesis. The mutations L342A and L343A resulted in significantly lower basal

activity which is in agreement with our theoretical observations (Jyrkkärinne et al., 2005). In contrast the mutation C347A showed only slight decrease of basal activity indicating a less important role for maintaining the constitutive activity in human CAR (Frank et al., 2004). The high sensitivity of the CAR basal activity for single point mutations indicates a relatively weak interaction of H12 with the LBD allowing an easy shift from the active into the inactive H12 conformation. This can be easily achieved by the introduction of single alanine mutations. Xu et al. have discussed the role of F161 as interaction partner for the "helix X" and thus as component of the hydrophobic barrier (Xu et al., 2004). From the MD simulations of the CAR homology models and X-ray structures, F161 emerged as an amino acid with high flexibility. During any MD simulations done for ligand-free CAR X-ray structures, F161 adopts a conformation pointing into the LBP and thus having no vdW interactions with "helix X". In contrast, in CAR complexed with the agonists CITCO or 5β-pregnanedione F161 is reoriented towards the LBD-H12 interface. In the X-ray structure of murine CAR complexed with the agonist 1,4-bis[2-(3,5-dichloro-pyridyloxy)]benzene (TCPOBOP), the corresponding F171 points into the LBP, similar to the conformation observed during MD simulations for ligand-free human CAR X-ray structures (Suino et al., 2004). These observations clearly indicate the high flexibility of F161 and the possibility to adopt different conformations depending on the bound ligand.

Moreover, the observed flexibility of F161 might be a hint for the structural basis of activating CAR beyond its basal activity. In contrast to agonists showing direct contact with H12 (such as TCPOBOP does in murine CAR) other agonists might interact indirectly with H12 by reorienting F161 towards this helix. In this conformation, F161 is able to restrict the movement of Y326 via its bulky aromatic side chain. Results from the MD simulations show that agonist binding moves Y326 closer towards H12 in cooperation with F161 thus reinforcing vdW interactions.

We observed in the MD simulations that the salt bridge connecting K195 and the C-terminal S348 is not stable. K195 rather was found to form a hydrogen bond to H687 located on the SRC-1 NRID. The C terminus in turn has been observed to have transient interactions with Q331. Site-directed mutagenesis has already revealed the importance of Q331 for the basal activity of CAR. Experimental data concerning K195 and S348 are controversial. Andersin et al. (Andersin

et al., 2003) reported that extension of H12 by three residues did not influence the basal activity whereas in other studies extension of H12 resulted in loss of basal activity which has been interpreted as confirmation of the K195-C terminus interactions (Frank et al., 2004; Lempiäinen et al., 2005). However, it cannot be excluded that elongation of H12 by one turn disrupts the potential hydrogen bond between Q331 and the C terminus leading to destabilisation of H12. Furthermore, CAR X-ray data suggests that a C-terminal extension might result in a steric clash with H5 resulting in displacement of H12 and CAR inactivation. In another report mutation of K195 to alanine impaired binding of coactivator GRIP1 but not SRC-1 to the thyroid hormone receptor (Feng et al., 1998). Supershift assays for CAR showed that mutation K195A disrupts the interaction of CAR with the coactivator TIF-2. Interestingly, in this experiment CITCO was able to recover the interaction with TIF-2 (Frank et al., 2004). From the available experimental data, K195 emerges as important residue for coactivator recruitment. Whether this is achieved by direct contact with the coactivator or by keeping H12 in the active position via interacting with the C terminus is not clear so far. It cannot be excluded that K195 not only stabilises coactivator binding but simultaneously constitutes interactions with H12 as observed in the crystal structures of other NRs (e.g. PXR and PPAR).

For mouse CAR additional interactions between S337 and the C terminus have been described and discussed as being involved in the stabilisation of the active conformation of helix H12 (Suino et al., 2004; Moore, 2005). In human CAR the corresponding amino acid is glycine which is not able to interact with H12. The visual inspection of the murine CAR X-ray structure (PDB code 1XLS) also revealed no such interaction.

"Helix X" has been assigned a major role for the CAR basal activity. However, several crystal structures deposited in the Protein Databank were found to contain a corresponding helical motif and for most of them no constitutive activity has been reported. Similar helical elements in the H11-H12 region can be detected in the X-ray structures of RORα and β, TRα and β, VDR as well as GR. These observations stress the question whether the occurrence of a single turn helix can serve as an explanation for the basal activity of CAR. According to the common annotation of helices in nuclear receptors, "helix X" should be rather termed helix 11' as it has been done at first in RORβ (Stehlin et al., 2001). The flexibility in the H11'-H12 region has been used as structural explanation

for the different activation profiles of CAR and RORα/β (constitutive activity) in contrast to that of VDR (no constitutive activity) (Xu et al., 2004). However, the stated higher flexibility of the H11'-H12 region in VDR is limited by hydrogen bonds stabilising the amino acids of this segment. As described before, the loop connecting H11' and H12 in ROR and TR is identical, but in contrast to ROR, TR activates gene expression in a ligand dependent manner and shows no constitutive activity. The additional hydrogen bond that further stabilises H11' in TR also contradicts the hypothesis of Xu et al. that a rigid H11-H12 region is a key element for constitutive activity (Xu et al., 2004). The occurrence of H11' in TR (and GR) dependent on the bound ligand suggests that the formation of H11' in TR (and GR) is modulated upon ligand binding. Whether the formation of H11' is required for coactivator binding remains unknown and has to be examined by further studies in the future.

6.4 Conclusions

In this study we could show that the applied homology modelling procedure enabled the generation of a CAR model that is in good agreement with the now available X-ray crystallographic data. Particularly, the side chain conformations of amino acids lining the LBP were predicted with high accuracy. MD simulations of the CAR model and X-ray structures revealed a high flexibility of F161 that might play a role in the adjustment to structurally diverse ligands. The salt bridge between K195 and the C terminus observed in the CAR X-ray structures was found to be unstable during MD simulations. The examination of available NR X-ray structures showed that a "helix X"-corresponding element can also be observed in receptors possessing no constitutive activity. The obtained results do not support the interpretation of helix 11' (helix X) as an essential structural feature for CAR basal activity. The crystal structure of ligand-free CAR will clarify whether helix 11' ("helix X") is a permanent feature or - as we predict - is rather formed upon agonist binding.

Chapter 7

Virtual Screening

CAR agonists are regarded as potential therapeutics in cholestasis or neonatal jaundice. Ligand-activated CAR induces gene expression of metabolising enzymes finally leading to degradation of bilirubin or bile acids, respectively. Currently, only few compounds are known to increase the transcriptional activity of human CAR of which some have been already applied or are discussed as potential therapeutics (Chang and Waxman, 2006). In order to find new and structurally different human CAR agonists that might be used as new lead compounds for the treatment of hyperbilirubinemia and bile acid disorders, virtual screening procedures were carried out.

7.1   3D Database Search

At first a database search was conducted using the LeadQuest database (Tripos Inc., St. Louis, USA) containing roughly 85000 molecules. Compounds violating Lipinski's "rule of five" were excluded. Within the UNITY module of Sybyl a search query was defined consisting of the Conolly surface of the LBP (Connolly, 1983) and two sites of hydrophobic character (Figure 7.1*A*). The tolerance of the LBP surface was set to 0.5 Å. The selection of the hydrophobic features with a tolerance of 1.0 Å was based on both, observed vdW interactions between the CAR homology model and the docked ligands Clotrimazole and TMPP as well as favourable sites of lipophilic interactions within the LBP that were evaluated by calulcating the lipophilic potential of the LBP (Ghose et al., 1998). Active compounds were expected to fit into the features of the

search query whereas inactive molecules would not (Figure 7.1*B*). Altogether 9709 compounds were found to fit into the spatial environment of the search query.
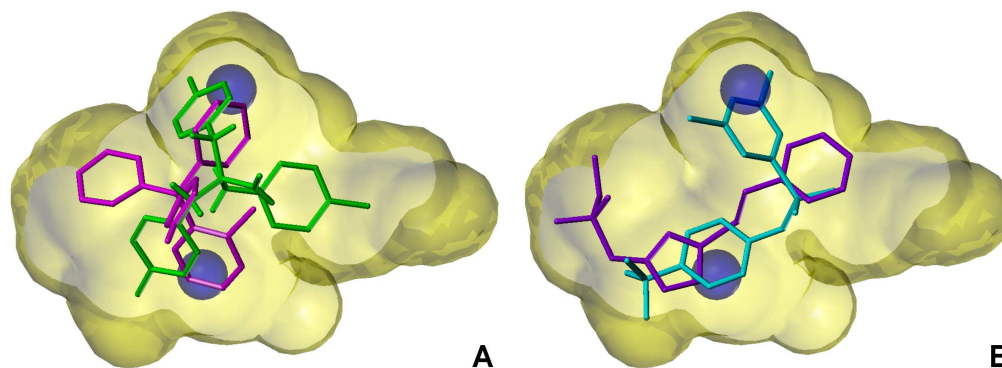
Figure 7.1: (A) The UNITY search query defined by the LBP volume (yellow) and two hydrophobic features (blue spheres) placed on the basis of information obtained from docked ligands Clotrimazole (magenta) and TMPP (green) as well as the lipophilic potential of the LBP (not shown).**(B)** Conformation of an active (cyan) and inactive (violet) compound within the LBP.

7.2 Molecular Docking

The 9709 compounds that emerged from the 3D database search were subjected to a molecular docking procedure carried out using GOLD version 2.2 (CCDC, Cambridge, UK). GoldScore was selected as scoring function. All molecules were docked within a sphere of 20 Å radius that was defined around the side chain atom CD1 of amino acid L206. For each ligand a maximum number of 10 conformations was allowed. Ligand docking was terminated once the first three conformations were all within an RMSD of 1.5 Å of each other.

In order to take the coherence of the GoldScore and the molecular weight into account, the GoldScore value was corrected by dividing the original score over the square-root of the number of heavy atoms (Pan et al., 2003).

Additionally, all conformations obtained from the GOLD docking were re-scored by applying an external scoring function (X-Score) (Wang et al., 2002). The results were likewise corrected by the method mentioned before.

Finally, the 50 best ranked conformations resulting from corrected GoldScore

and X-Score were re-docked and -scored using the genetic algorithm implemented in AUTODOCK version 3.0.5 (Morris et al., 1998).

From each docking/scoring approach the 20 top-ranked results were selected for their effect on CAR activation in a cell-based reporter assay (Jyrkkärinne et al., 2003; Mäkinen et al., 2003). By excluding multiply listed compounds altogether 64 compounds were suggested for further experimental investigation of which 54 were purchased.

7.3   Re-Docking into X-ray Structures

The release of CAR X-ray structures provided additional protein input structures for virtual screening. Both human CAR X-ray structures resolved in complex with either 5β-pregnane-3,20-dione or CITCO were tested for their adequacy in virtual screening procedures. Thus, eight known CAR ligands were docked into all structures and grouped into clusters of similar conformations. In case of crystal structure 1XV9 compounds with few rotatable bonds (especially steroids) were found in just few, slightly different conformations whereas for more flexible residues (more than 4 rotatable bonds) significantly more conformations were found (see Table 7.1). For X-ray structure 1XVP the result is almost the other way round. This is most probably due to a slight adaption of the 1XV9 LBP to the steroid framework of 5β-pregnane-3,20-dione which in turn favours one optimum geometry for steroid-like compounds but fails to find a distinct conformation for more flexible molecules. By contrast X-ray structure 1XVP is complexed with a more bulky compound and appears to be suited not only for screening steroids but also for molecules having several rotatable bonds. This may be a result of the X-ray structure refinement which was conducted in absence of CITCO due to an ambiguous electron density for the co-crystallised ligand (Xu et al., 2004). Therefore it is likely that the LBP conformation is close to that of a ligand-free receptor and not adapted to a bound ligand. As the database used in virtual screening contains structurally diverse compounds the crystal structure 1XVP was selected for the virtual screening.

A database containing the 500 top-ranked compounds from the initial virtual screening approach was re-docked into both CAR monomers of the PDB entry 1XVP by applying GOLD version 2.2 and GoldScore as scoring function with

| Ligand	rotatable bonds	number of clusters	
		1XV9	1XVP
ANDR	1	2	1
EE2	2	1	2
PREG	1	1	2
CITCO	5	3	4
CLOT	4	2	4
TCPOBOP	4	5	1
TMPP	5	7	5
TPP	5	5	2

Table 7.1: Comparison of docking results obtained for crystal structures of human CAR. For both X-ray structures the number of clusters is given for each compound. Abbreviations: ANDR, 5α-androstan-3α-ol; EE2, 17α-ethynyl-3,17β-estradiol; PREG, 5β-pregnane-3,20-dione; CLOT, clotrimazole.

settings as described before.

From the top-ranked molecules a total of 66 compounds were selected according to their GoldScore values. Altogether 30 compounds were purchased and tested for CAR activation as described before.

7.4 Results

In the screening study that made use of the homology model a compound was considered as CAR agonist when at least 50 % of the Clotrimazole-induced CAR activity was reached. Actually, most compounds were found to have significantly lower activating effects on CAR and some even showed potential inhibiting effects. Figure 7.2 summarises the impact of all tested compounds on transcriptional activity by CAR. Only six molecules fulfilled the pre-defined criterion of CAR agonists corresponding to a hit rate of 11 %. Interestingly, all active compounds are represented by *N*-substituted or *N,N*-disubstituted sulfonamides. The molecular weight of all ligands varies between 300 and 400 Da.
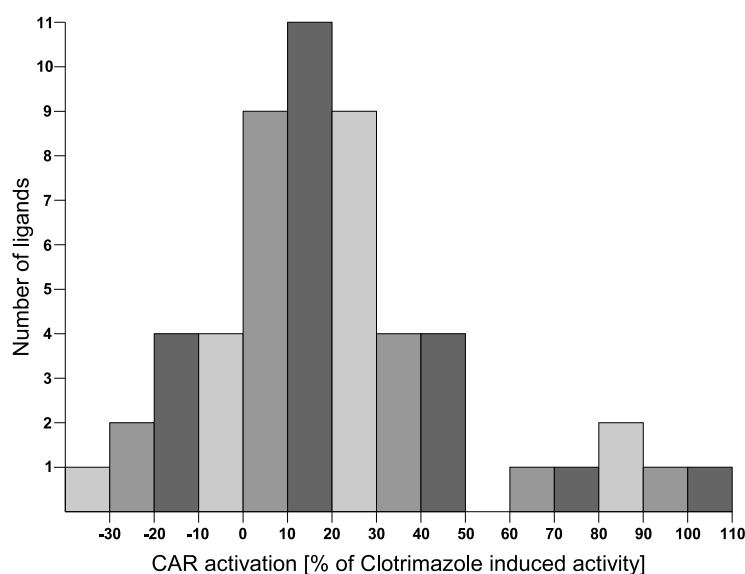
Figure 7.2: Number of tested compounds and their effect on CAR transactivation.

In the screening approach based on the CAR X-ray structure compounds activating CAR by at least 50 % above the basal activity rate were considered as agonists due to the strong CAR activation induced by Clotrimazole in the reporter assays. According to this criterion, 19 compounds were identified as CAR agonists corresponding to a hit rate of 63 %. The extent of CAR activation differs considerably spanning about one order of magnitude varying from about 60 to more than 800 % (Figure 7.3).

Chemically the hit molecules can be divided into *N*-substituted and *N,N*-disubstituted sulfonamides (10) as well as 4-thiazolidinone derivatives (9). In contrast to the first VS approach that made use of the homology model the active compounds show a significantly larger structural variability indicated by a large diversity in the side chains of the core structure. The range of the molecular weight of the compounds is almost identical with that observed for the initial VS procedure varying between 315 and 415 Da.

The overlap of top-ranked molecules in both virtual screening procedures is relatively small encompassing altogether only 11 compounds of which most were revealed as inactive.
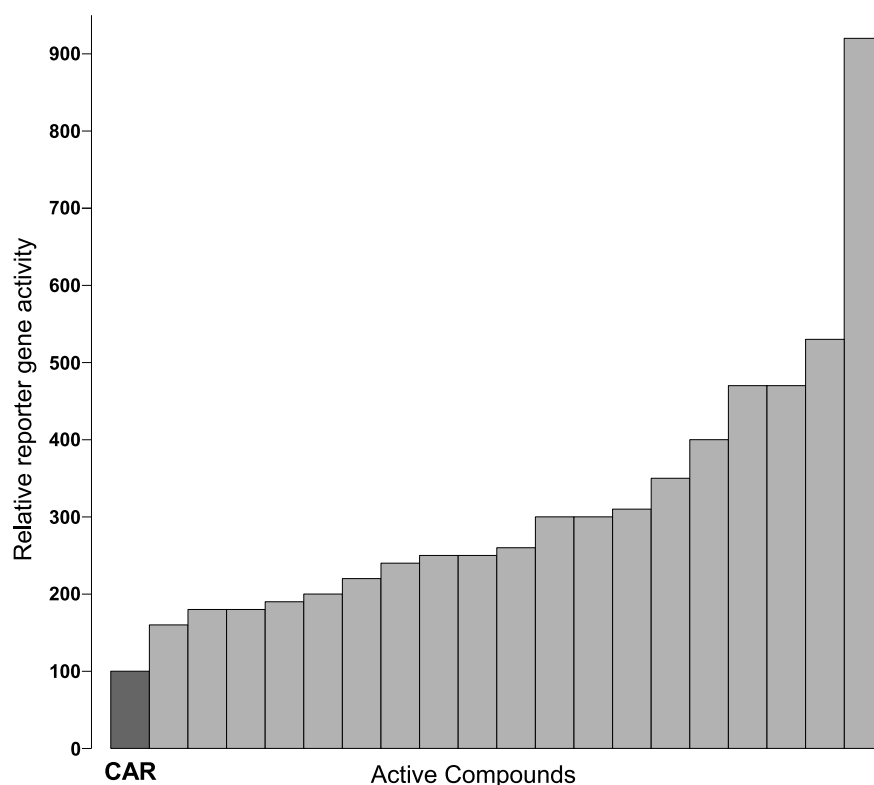
Figure 7.3: Activity of novel CAR agonists found by the VS study based on the CAR X-ray structure. The basal activity of un-liganded CAR is shown by the dark grey bar.

7.5 Discussion

The difference in the success rates of both VS approaches is remarkable. Even though the overall LBP structure of the homology model is similar to that of the X-ray structures (as described in detail in Chapter 6), small deviations can lead to large discrepancies as shown for the different docking results obtained from comparing X-ray structures 1XV9 and 1XVP (See Table 7.1).

The initial 3D database search failed to identify structurally more diverse compounds because the search query with two hydrophobic features was most probably too rigid. Thus the screening resulted in an enriched library comprising almost exclusively sulfonamides and 4-thiazolidinone derivatives.

In contrast, docking programmes sample the whole LBP and thus are capable to identify structurally more diverse ligands. In order to create an enriched compound library covering a broad range of structurally diverse molecules, a

more general search query has to be defined including also information from newly discovered CAR agonists. Furthermore information of the characteristics of the LBP will be considered (*i.e.* amino acids identified as important from functional assays).

As the LBP structures and characteristics of ligand binding by classical endocrine receptors are quite different from that of CAR (which possesses few specific hydrogen donors/acceptors and a predominance of aromatic residues) parameters of currently available virtual screening programmes must be modified. Biological datasets of the now available large number of CAR agonists will support the evaluation and improvement process of this procedure.

Chapter 8

Conclusions & Outlook

Understanding protein function not only requires experimental data from mutagenesis studies or ligand activation assays but also information about the protein structure itself. Within the scope of the present work, three-dimensional structures of the constitutive androstane receptor ligand binding domain were elaborated by means of homology modelling and MD simulations. The main goal was to gain insights into the molecular mechanism of the constitutive activity of CAR and to characterise amino acids contributing to formation of the LBP in terms of ligand and species specificity. Furthermore, *in silico* virtual screening approaches were applied in order to identify new CAR ligands.

In contrast to X-ray crystallography and NMR spectroscopy homology modelling techniques represent a theoretical approach and information derived thereof may suffer from missing experimental data. By extensive application of functional *in vitro* assays the plausibility of the results obtained from homology modelling was confirmed and used for further optimisation of the CAR models.

Furthermore, the phenotype of experimental mutational assays could be explained on a structural level by the models and, in return, important interactions within the structures were approved by mutagenesis studies.
Amino acids important for ligand and also for species specificity could be selected from the models and their importance for ligand function was revealed by *in vitro* assays.

The reliability of the model structure was further confirmed by a direct comparison with CAR X-ray data which were released last year. The comparison revealed a high degree of consistency even within the ligand binding pocket. Indeed, evaluation of the modelling quality in terms of structural similarity with CAR X-ray crystals is limited. It must be kept in mind that 3D structures obtained from X-ray crystallography represent a protein conformation in an unphysiological environment that may suffer from packing effects in the protein crystal. Actually, crystal structures represent also structural models that fit into the electron density map obtained from the diffraction pattern. The resolution determines how well the position of atoms can be assigned. In this respect, CAR crystal structures can be considered as mid or low resolution structures (2.6-3 Å) with exceedingly high overall B-factors.

Although a prediction for the molecular mechanism underlying the constitutive activity could be derived from the models that is in agreement with data from *in vitro* assays, the homology model turned out to be of limited adequacy in VS procedures as indicated by a considerably lower success rate compared to a VS approach that made use of CAR X-ray structures.

Based on an homology model of the inactive conformation also a first idea of the mechanism of ligand-dependent CAR inactivation has been proposed. This has to be further evaluated in the future by the generation of a more sophisticated model based on the murine CAR crystal structure complexed with the inverse agonist 5α-androst-16-en-3α-ol. The new model will serve as initial point of VS efforts in order to identify and develop novel CAR antagonists that may offer proceedings in therapy of obesity.

On the basis of the applied *in silico* VS approach altogether 25 compounds were revealed as novel CAR agonists. represented by sulfonamides and thiazolan-4-one derivatives. Thus, members of two classes of CAR activators were retrieved that significantly expand the number of known activators. Together with novel CAR agonists identified by a ligand-based approach (Antti Poso, personal communication), the newly discovered activating compounds will provide further optimisation in the construction of a more sophisticated search query. This may lead to a large and diverse pool of CAR agonists for the development of novel therapeutics of hyperbilirubinemia and bile acid disorders.

Chapter 9

Summary

The first part of this work describes the elucidation of the structural basis for the constitutive activity of the CAR. Based on a homology model of the LBD a potential molecular mechanism underlying the constitutive activity is presented. According to the results, the structural basis of constitutive activity is achieved by specific vdW interactions between the LBD and H12 with Y326 as central amino acid. Structure comparisons with LBD/ligand complexes of closely related NRs indicate the role of Y326 as an amino acid mimicking a bound agonist. Results from experimental and virtual mutagenesis studies strongly support the suggested mechanism. Homology models also allowed to suggest a mechanism how agonist binding activates CAR beyond its basal activity.

The second part of this work encompasses the functional characterisation of the ligand binding pocket. By alanine scanning mutagenesis studies the proposed mechanism underlying the constitutive activity could be further supported. Moreover, amino acids critical for ligand binding and specificity could be identified by means of docking studies and experimental testings. Based on a homology model of inactive CAR, a molecular basis for inverse agonists could be suggested. Finally, the models allowed an explanation for the species differences in ligand specificity observed between mouse and human CAR that is consistent with experimental data.

In the third part of this work, the homology model is compared with now available CAR X-ray structures. The accuracy of the modelled LBD and, in particular, the ligand binding pocket has been assessed. By detailed analysis it could be demonstrated that the model is in good agreement with structures obtained by X-ray crystallography. Especially the side chain conformations of the ligand binding pocket could be predicted with high accuracy. This was further approved by the reproduction of binding modes of co-crystallised ligands in the model with high consistency. Moreover, it could be revealed that the proposed structural mechanism for constitutive activity derived from the X-ray structures is problematic.

Finally, the model has been successfully used in the systematic search for novel CAR agonists. By virtual screening methods using the homology model as well as CAR X-ray structures altogether 25 compounds could be revealed as novel agonists that comprise *N*-substituted and *N,N*-disubstituted sulfonamides as well as 4-thiazolidinone derivatives.

Bibliography

Abagyan, R., Frishman, D., and Argos, P. (1994). Recognition of distantly related proteins through energy calculations. *Proteins*, 19:132–140.

Acevedo, M. L. and Kraus, W. L. (2004). Transcriptional activation by nuclear receptors. *Essays Biochem.*, 40:73–88.

Adachi, R., Shulman, A. I., Yamamoto, K., Shimomura, I., Yamada, S., Mangelsdorf, D. J., and Makishima, M. (2004). Structural determinants for vitamin D receptor response to endocrine and xenobiotic signals. *Mol. Endocrinol.*, 18:43–52.

Ajay, A. and Murcko, M. A. (1995). Computational methods to predict binding free energy in ligand-receptor complexes. *J. Med. Chem.*, 38:4953–4967.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.*, 215:403–410.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Madden, D. J. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Basic local alignment search tool. *Nucleic Acids Res.*, 25:3389–3402.

Andersin, T., Väisänen, S., and Carlberg, C. (2003). The critical role of carboxy-terminal amino acids in ligand-dependent and -independent transactivation of the constitutive androstane receptor. *Mol. Pharmacol.*, 17:234–246.

Aranda, A. and Pascual, A. (2001). Nuclear hormone receptors and gene expression. *Physiol. Rev.*, 81:1269–1304.

Arnold, K. A., Eichelbaum, M., and Burk, O. (2004). Alternative splicing affects the function and tissue-specific expression of the human constitutive androstane receptor. *Nucl. Recept.*, 2:1.

Ashburner, M., Chihara, C., Meltzer, P., and Richards, G. (1974). Temporal control of puffing activity in polytene chromosomes. *Cold Spring Harbor Symp. Quant. Biol.*, 38:655–662.

Baes, M., Gulick, T., Choi, H. S., Martinoli, M. G., Simha, D., and Moore, D. D. (1994). A new orphan member of the nuclear hormone receptor superfamily that interacts with a subset of retinoic acid response elements. *Mol. Cell. Biol.*, 14:1544–1552.

Bajorath, J. (2002). Integration of virtual and high-throughput screening. *Nat. Rev. Drug Discov.*, 1:882–894.

Baker, A. R., McDonnell, D. P., Hughes, M., Crisp, T. M., Mangelsdorf, D. J., Haussler, M. R., Pike, J. W., Shine, J., and O'Malley, B. W. (1988). Cloning and expression of full-length cDNA encoding human vitamin D receptor. *Proc. Natl. Acad. Sci. USA*, 85:3294–3298.

Baker, K. D., Shewchuk, L., Kozlova, T., Makishima, M., Hassell, A., Wisely, B., Caravella, J. A., Lambert, M. H., Reinking, J. L., Krause, H., Thummel, C. S., Willson, T. M., and Mangelsdorf, D. J. (2003). The drosophila orphan nuclear receptor dhr38 mediates an atypical ecdysteroid signaling pathway. *Cell*, 113:731–742.

Beato, M., Herrlich, P., and Schütz, G. (1995). Steroid hormone receptors: many actors in search of a plot. *Cell*, 83:851–857.

Belandia, B., Latasa, M. J., Villa, A., and Pascual, A. (1998). Thyroid hormone negatively regulates the transcriptional activity of the beta-amyloid precursor protein gene. *J. Biol. Chem.*, 273:30366–30371.

Benko, S., Love, J. D., Beladi, M., Schwabe, J. W. R., and Nagy, L. (2003). Molecular determinants of the balance between co-repressor and co-activator recruitment to the retinoic acid receptor. *J. Biol. Chem.*, 278:43797–43806.

Berendsen, H. J. C., Postma, J. P. M., DiNola, A., and Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys*, 81:3684–3690.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J. (1981). Interaction models for water in relation to protein hydration. In *Intermolecular Forces*, pages 331–342. D. Reidel Publishing Company, Dordrecht.

Berendsen, H. J. C., van der Spoel, D., and van Drunen, R. (1995). GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.*, 91:43–56.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindylov, I. N., and Bourne, P. E. (2000). The protein data bank. *Nucleic Acids Research*, 28:235–242.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. J., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977). The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.*, 112:535–542.

Berry, M., Methzger, D., and Chambon, P. (1990). Role of the two activating domains of the oestrogen receptor in the cell-type and promoter-context dependent agonistic activity of the anti-oestrogen 4-hydroxytamoxifen. *EMBO J.*, 9:2811–2818.

Bertilsson, G., Heidrich, J., Svensson, K., Aman, M., Jendeberg, L., Sydow-Backman, M., Ohlsson, R., Postlind, H., Blomquist, P., and Berkenstam, A. (1998). Identification of a human nuclear receptor defines a new signaling pathway for CYP3A induction. *Proc. Natl. Acad. Sci. USA*, 95:12208–12213.

Blumberg, B., Kang, H., Boldao Jr., J., Chen, H., Craig, A. G., Moreno, T. A., Umesono, K., Perlman, T., Robertis, E. M. D., and Evans, R. M. (1998a). BXR, an embryonic orphan nuclear receptor activated by a novel class of endogenous benzoate metabolites. *Genes Dev.*, 12:1269–1277.

Blumberg, B., Sabbagh Jr, W., Juguilon, H., Bolado Jr, J., van Meter, C. M., Ong, E. S., and Evans, R. M. (1998b). SXR, a novel steroid and xenobiotic-sensing nuclear receptor. *Genes dev.*, 12:3195–3205.

Bock, K. W., Lilienblum, W., Fischer, G., Schirmer, G., and Bock-Henning, B. S. (1987). The role of conjugation reactions in detoxication. *Arch. Toxicol.*, 60:22–29.

Boeckmann, B., Bairoch, A., Apwiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., Martin, M. J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S., and Schneider, M. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.*, 31:365–370.

Borst, P. and Elferink, R. O. (2002). Mammalian ABC transporters in health and disease. *Annu. Rev. Biochem.*, 71:537–592.

Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H., and Moras, D. (1995). Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-α. *Nature*, 375:377–382.

Bower, M. J., Chen, F. E., and Dunbrack Jr., R. L. (1997). Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool. *J. Mol. Biol.*, 267:1268–1282.

Bowie, J. U., Lüthy, R., and Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 253:164–170.

Bränden, C. I. and Jones, T. A. (1990). Between objectivity and subjectivity. *Nature*, 343:687–689.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J. Comp. Chem*, 4:187–217.

Brown, A. J. (2001). Therapeutic uses of vitamind analogues. *Am. J. Kidney Dis.*, 38:S3–S19.

Brzozowski, A. M., Pike, A. C., Dauter, Z., Hubbard, R. E., Bonn, T., Engstrom, O., Ohman, L., Greene, G., Gustafsson, J. A., and Carlquist, M. (1997). Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature*, 389:753–758.

Burchell, B., Coughtrie, M., Jackson, M., Harding, D., Fournel-Gigleux, S., Leakey, J., and Hume, R. (1989). Development of human liver UDP-glucuronosyltransferases. *Dev. Pharmacol. Ther.*, 13:70–77.

Burk, O., Arnold, K. A., Geick, A., Tegude, H., and Eichelbaum, M. (2005). A role for constitutive androstane receptor in the regulation of human intestinal MDR1 expression. *Biol. Chem.*, 386:503–513.

Burk, O., Koch, I., Raucy, J., Hustert, E., Eichelbaum, M., Brockmöller, J., Zanger, U. M., and Wojnowski, L. (2004). The induction of cytochrome p450 3A5 (CYP3A5) in the human liver and intestine is mediated by the xenobiotic sensors pregnane x receptor (PXR) and constitutively activated receptor (CAR). *J. Biol. Chem.*, 279:38379–38385.

Caldwell, J. (1982). Conjugation reactions in foreign-compound metabolism: definition, consequences, and species variations. *Drug Metab. Rev*, 13:745–777.

Canutescu, A. A., Shelenkov, A. A., and Dunbrack Jr., R. L. (2003). A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci.*, 12:2001–2014.

Carlberg, C. (1995). Mechanisms of nuclear signalling by vitamin D_3. interplay with retinoid and thyroid hormone signalling. *Eur. J. Biochem.*, 231:517–527.

Chandler, V. L., Maler, B. A., and Yamamoto, K. R. (1983). DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter hormone responsive in vivo. *Cell*, 33:489–499.

Chang, T. K. and Waxman, D. J. (2006). Synthetic drugs and natural products as modulators of constitutive androstane receptor (CAR) and pregnane x receptor (PXR). *Drug Metab. Rev.*, 38:51–73.

Charifson, P. S., Corkery, J. J., Murcko, M. A., and Walters, W. P. (1999). Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.*, 42:5100–5109.

Chawla, A., Repa, J. J., Evans, R. M., and Mangelsdorf, D. J. (2001). Nuclear receptors and lipid physiology: opening the x-files. *Science*, 294:1866–1870.

Chen, J. D. and Evans, R. M. (1995). A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature*, 377:387–388.

Cherrington, N. J., Hartley, D. P., Li, N., Johnson, D. R., and Klassen, C. D. (2002). Organ distribution of multidrug resistance proteins 1, 2, and 3 (mrp1, 2, and 3) mRNA and hepatic induction of mrp3 by constitutive androstane receptor activators in rats. *J. Pharmacol. Exp. Ther.*, 300:97–104.

Choi, H. S., Chung, M., Tzameli, I., Simha, D., Lee, Y. K., Seol, W., and Moore, D. D. (1997). Differential transactivation by two isoforms of the orphan nuclear hormone receptor CAR. *J. Biol. Chem.*, 272:23565–23571.

Chothia, C. and Lesk, A. M. (1986). The relation between the divergence of sequence and structure in proteins. *EMBO J.*, 5:823–826.

Chrencik, J. E., Orans, J., Moore, L. B., Xue, Y., Peng, L., Collins, J. L., Wisely, G., Lambert, M. H., Kliewer, S. A., and Redinbo, M. R. (2005). Structural disorder in the complex of human pregnane X receptor and the macrolide antibiotic rifampicin. *Mol. Endocrinol.*, 19:1125–1134.

Claessens, F. and Gewirth, D. T. (2004). DNA recognition by nuclear receptors. *Essays Biochem.*, 40:59–72.

Clark, R. D., Strizhev, A., Leonard, J. M., Blake, J. F., and Matthew, J. B. (2002). Consensus scoring for ligand/protein interactions. *J. Mol. Graph. Model.*, 20:281–295.

Claudel, T., Sturm, E., Kuipers, F., and Staels, B. (2003). The farnesoid X receptor: a novel drug target? *Expert Opin. Investig. Drugs*, 13:1135–1148.

Claussen, H., Buning, C., Rarey, M., and Lengauer, T. (2001). FlexE: efficient molecular dockung considering protein structure variations. *J. Mol. Biol.*, 308:377–395.

Collingwood, T. N., Wagner, R. J., Matthews, C. H., Clifton-Bligh, R. J., Gurnell, M., Rajanayagam, O., Agostini, M., Fletterick, R. J., Beck-Peccoz, P., Reinhardt, W., Binder, G., ranke, M. B., Hermus, A., Hesch, R. D., Lazarus, J., Newrick, P., Parfitt, V., Raggatt, P., de Zegher, F., and Chatterjee, V. K. (1998). A role for helix 3 of the TRbeta ligand-binding domain in coactivator

recruitment identified by characterization of a third cluster of mutations in resistance to thyroid hormone. *EMBO J.*, 17:4760–4770.

Committee, N. R. N. (1999). A unified nomenclature system for the nuclear receptor superfamily. *Cell*, 97:161–163.

Conneely, O., Maxwell, B., Toft, D., Schrader, W., and O'Malley, B. (1987). The A and B forms of the chicken progesterone receptor arise by alternate initiation of translation of a unique mRNA. *Biochem. Biophys. Res. Commun.*, 149:493–501.

Conney, A. H., Davison, C., Gastel, R., and Burns, J. J. (1960). Adaptive increases in drug-metabolizing enzymes induced by phenobarbital and other drugs. *J. Pharmacol. Exp. Ther.*, 130:1–8.

Connolly, M. L. (1983). Solvent-accessible surfaces of proteins and nucleic acids. *Science*, 221:709–713.

Constanciel, R. and Contreras, R. (1984). Self-consistent field theory of solvent effects representation by continuum models - introduction of desolvation contribution. *Theoretica Chimica Acta*, 65:1–11.

Côte, S., Zhou, D., Bianchini, A., Nervi, C., Gallagher, R. E., and Miller Jr., W. H. (2000). Altered ligand binding and transcriptional regulation by mutations in the PML/RARα ligand-binding domain arising in retinoic acid-resistant patients with acute promyelocytic leukemia. *Blood*, 96:3200–3208.

Cramer, R. D., Patterson, D. E., and Bunce, J. D. (1988). Comparative molecular field analysis (coMFA). 1. effect of shape on binding of steroids to carrier protein. *J. Am. Chem. Soc.*, 110:5959–5967.

Danielian, P. S., White, R., Lees, J. A., and Parker, M. G. (1992). Identification of a conserved region required for hormone dependent transcriptional activation by steroid hormone receptors. *EMBO J:*, 11:1025–1033.

Darden, T. A., York, D., and Pedersen, L. (1993). Particle mesh ewald: An nlog(n) method for ewald sums in large systems. *J. Chem. Phys.*, 98:10089–10092.

Darimont, B. D., Wagner, R. L., Apriletti, J. W., Stallcup, M. R., Kushner, P. J., Baxter, J. D., Fletterick, R. J., and Yamamoto, K. R. (1998). Structure and specificity of nuclear receptor-coactivator interactions. *Genes Dev.*, 12:3343–3356.

Dennery, P. A. (2002). Pharmacological interventions for the treatment of neonatal jaundice. *Semin. Neonatol.*, 7:111–119.

Dennery, P. A., Seidman, D. S., and Stevenson, D. K. (2001). Neonatal hyperbilirubinemia. *N. Engl. J. Med.*, 344:581–590.

Deroo, B. J. and Korach, K. S. (2006). Estrogen receptors and human disease. *J. Clin. Invest.*, 116:561–570.

Dilworth, F. J. and Chambon, P. (2001). Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. *Oncogene*, 20:3047–3054.

Diwan, B. A., Lubet, R. A., Ward, J. M., Hrabie, J. A., and Rice, J. M. (1992). Tumor-promoting and hepatocarcinogenic effects of 1,4-bis[2-(3,5-dichloropyridyloxy)]benzene (TCPOBOP) in DBA/2NCr and C57BL/6NCr mice and an apparent promoting effect on nasal cavity tumors but not on hepatocellular tumors in F344/NCr rats initiated with n-nitrosodiethylamine. *Carcinogenesis*, 13:1893–1901.

Dobson, A. D., Conneely, O. M., Beattie, W., Maxwell, B. L., Mak, P., Tsai, M. J., Schrader, W. T., and O'Malley, B. W. (1989). Mutational analysis of the chicken progesterone receptor. *JBC*, 264:4207–4211.

Doman, T. N., McGovern, S. L., Witherbee, B. J., Kasten, T. P., Kurumbail, R., Stallings, W. C., Conolly, D. T., and Shoichet, B. K. (2002). Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1b. *J. Med. Chem.*, 45:2213–2221.

Duan, Y. and Kollman, P. A. (1998). Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*, 282:740–744.

Ducharme, M. P., Bernstein, M. L., Granvil, C. P., Gehrke, B., and Wainer, I. W. (1997). Phenytoin-induced alteration in the n-dechloroethylation of ifosfamide stereoisomers. *Cancer Chemother. Pharmacol.*, 40:531–533.

Dunbrack Jr., R. L. (1999). Comparative modeling of CASP3 targets using PSI-BLAST and SCWRL. *Proteins*, 3:81–87.

Dussault, I., Lin, M., Hollister, K., Fan, M., Termini, J., Sherman, M. A., and Forman, B. M. (2002). A structural model of the constitutive androstane receptor defines novel interactions that mediate ligand-independent activity. *Mol. Cell. Biol.*, 22:5270–5280.

Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V., and Mee, R. P. (1997). Empirical scoring functions: I. the development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.*, 11:425–445.

Engh, R. A. and Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst. A*, 47:392–400.

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995). A smooth particle mesh ewald potential. *J. Comp. Phys.*, 103:8577–8592.

Evans, R. M. (1988). The steroid and thyroid hormone receptor superfamily. *Science*, 240:889–895.

Ewald, P. (1921). Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik*, 97:4309–4315.

Fawell, S. E., Lees, J. A., White, R., and Parker, M. G. (1990). Characterization and colocalization of steroid binding and dimerization activities in the mouse estrogen receptor. *Cell*, 60:953–962.

Feng, W., Ribeiro, R. C. J., Wagner, R. L., Nguyen, H., Apriletti, J. K., Fletterick, R. J., Baxter, J. D., Kushner, P. J., and West, B. L. (1998). Hormone-dependent coactivator binding to a hydrophobic cleft on nuclear receptors. *Science*, 280:1747–1749.

Ferguson, S. S., Chen, Y., LeCluyse, E. L., Negishi, M., and Goldstein, J. A. (2005). Human CYP2C8 is transcriptionally regulated by the nuclear receptors constitutive androstane receptor, pregnane X receptor, glucocorticoid receptor, and hepatic nuclear factor 4α. *Mol. Pharmacol*, 68:747–757.

Ferguson, S. S., Chen, Y., Negishi, M., and Goldstein, J. A. (2002a). Identification of constitutive androstane receptor and glucocorticoid receptor binding sites in the CYP2C19 promoter. *Mol. Pharmacol.*, 64:316–324.

Ferguson, S. S., LeCluyse, E. L., Negishi, M., and Goldstein, J. A. (2002b). Regulation of human CYP2C9 by the constitutive androstane receptor: discovery of a new distal binding site. *Mol. Pharmacol.*, 62:737–746.

Filikov, A. V., Mohan, V., Vickers, T. A., Griffey, R. H., Cook, P., Abagyan, R. A., and James, T. L. (2000). Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR. *J. Comput. Aided Mol. Des.*, 14:593–610.

Flohil, J. A., Vriend, G., and Berendsen, H. J. C. (2002). Completion and refinement of 3-d homology models with restricted molecular dynamics: Application to targets 47, 58, and 111 in the CASP modelling competition and posterior analysis. *Proteins*, 48:593–604.

Forman, B. M., Tzameli, I., Choi, H. S., Chen, J., Simha, D., Seol, W., and Evans, R. M. (1998). Androstane metabolites bind to and deactivate the nuclear receptor CAR-β. *Nature*, 395:612–615.

Frank, C., Molnar, F., Matilainen, M., Lempiäinen, H., and Carlberg, C. (2004). Agonist-dependent and agonist-independent transactivations of the human constitutive androstane receptor are modulated by specific amino acid pairs. *J. Biol. Chem.*, 279:33558–33566.

Freedman, L. P., Luisi, B. F., Korszun, Z. R., Basavappa, R., Sigler, P. B., and Yamamoto, K. R. (1988). The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. *Nature*, 334:543–546.

Gampe Jr, R. T., Montana, V. G., Lambert, M. H., Miller, A. B., Bledsoe, R. K., Milburn, M. V., Kliewer, S. A., and Xu, H. E. (2000). Asymmetry in the PPARgamma/RXRalpha crystal structure reveals the molecular basis of heterodimerization among nuclear receptors. *Mol. Cell*, 5:545–555.

Gerbal-Chaloin, S., Daujat, M., Pascussi, J. M., Pichard-Garcia, L., Vilarem, M. J., and Maurel, P. (2002). Transcriptional regulation of CYP2c9 gene. role of glucocorticoid receptor and constitutive androstane receptor. *J. Biol. Chem.*, 277:209–217.

Ghose, A. K., Viswanadhan, V. N., and Wendoloski, J. J. (1998). Prediction of hydrophobic (lipophilic) properties of small organic molecules using fragmental methods: An analysis of ALOGP and CLOGP methods. *J. Phys. Chem. A*, 102:3762–3772.

Glass, C. K. and Rosenfeld, M. G. (2000). The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev.*, 14:121–141.

Go, N. and Scheraga, H. A. (1970). Ring closure and local conformational deformations of chain molecules. *Macromolecules*, 3:178–187.

Goodford, P. J. (1985). A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.*, 28:849–857.

Goodsell, D. S. and Olsen, A. J. (1990). Automated docking of substrates to proteins by simulated annealing. *Proteins*, 8:195–202.

Goodwin, B., Hodgson, E., D'Costa, D. J., Robertson, G. R., and Liddle, C. (2002). Transcriptional regulation of the human CYP3A4 gene by the constitutive androstane receptor. *Mol. Pharmacol.*, 62:359–365.

Goodwin, B. and Moore, J. T. (2004). CAR: detailing new models. *Trends Pharmacol. Sci.*, 25:437–441.

Grant, D. M. (1991). Detoxification pathways in the liver. *J. Inherit. Metab. Dis.*, 14:421–430.

Graves, A. P., Brenk, R., and Shoichet, B. K. (2004). Decoys for docking. *J. Med. Chem.*, 48:3714–3728.

Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J. M., Argos, P., and Chambon, P. (1986). Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. *Nature*, 320:134–139.

Greer, J. (1980). Model for haptoglobin heavy chain based upon structural homology. *Proc. Natl. Acad. Sci. U S A*, 77:3393–3397.

Greer, J. (1990). Comparative modelling methods: application to the family of the mammalian serine proteases. *Proteins*, 7:317–334.

Greschik, H., Wurtz, J. M., Sanglier, S., Bourguet, W., van Dorsselaer, A., Moras, D., and Renaud, J. P. (2002). Structural and functional evidence for ligand-independent transcriptional activation by the estrogen-related receptor 3. *Mol. Cell*, 9:303–313.

Guengerich, F. P. (1989). Characterization of human microsomal cytochrome p-450 enzymes. *Annu. Rev. Pharmacol. Toxicol.*, 29:241–264.

Guengerich, F. P. (1999). Cytochrome p-450 3a4: regulation and role in drug metabolism. *Annu. Rev. Pharmacol. Toxicol.*, 39:1–17.

Guo, G. L., Lambert, G., Negishi, M., Ward, J. M., Brewer Jr., H. B., Kliewer, S. A., Gonzalez, F. J., and Sinal, C. J. (2003). Complementary roles of farnesoid x receptor, pregnane x receptor, and constitutive androstane receptor in protection against bile acid toxicity. *J. Biol. Chem.*, 278:45062–45071.

Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., and Banks, J. L. (2004). Glide: a new approach for rapid, accurate docking and scoring. 2. enrichment factors in database screening. *J. Med. Chem.*, 47:1750–1759.

Handen, J. S. (2002). High-thoughput screening - challenges for the future. *Drug Discov. World*, pages 47–50.

Handschin, C. and Meyer, U. A. (2003). Induction of drug metabolism: the role of nuclear receptors. *Pharmacol. Rev.*, 55:649–673.

Hansen, C. M., Binderup, L., Hamberg, K. J., and Carlberg, C. (2001). Vitamin d and cancer: effects of 1,25(oh)2d3 and its analogs on growth control and tumorigenesis. *Front. Biosci.*, 6:D820–D848.

Hansen, T. W. and Tommarello, S. (1998). Effect of phenobarbital on bilirubin metabolism in rat brain. *Biol. Neonate*, 73:106–111.

Hansson, T., Oostenbrink, C., and van Gunsteren, W. F. (2002). Molecular dynamics simulations. *Curr. Opin. Struct. Biol.*, 12:190–196.

Hard, T., Kellenbach, E., Boelens, R., Maler, B. A., Dahlman, K., Freedman, L. P., Carlstedt-Duke, J., Yamamoto, K. R., Gustafsson, J. A., and Kaptein, R. (1990). Solution structure of the glucocorticoid receptor DNA-binding domain. *Science*, 249:157–160.

Heery, D. M., Kalkhoven, E., Hoare, S., and Parker, M. G. (1997). A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature*, 38:733–736.

Herzig, S., Hedrick, S., Morantte, I., Koo, S. H., Galimi, F., and Montminy, M. (2003). CREB controls hepatic lipid metabolism through nuclear hormone receptor PPAR-gamma. *Nature*, 426:190–193.

Hess, B., Becker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *J. Comp. Phys.*, 18:1463–1472.

Hollenberg, S. M., Weinberger, C., Ong, E. S., Cerelli, G., Oro, A., Lebo, R., Thompson, E. B., Rosenfeld, M. G., and Evans, R. M. (1985). Primary structure and expression of a functional human glucocorticoid receptor cDNA. *Nature*, 318:635–641.

Holm, L. and Sander, C. (1992). Fast and simple monte carlo algorithm for side chain optimization in proteins: application to model building by homology. *Proteins*, 14:213–223.

Höltje, H., Sippl, W., Rognan, D., and Folkers, G. (2003). *Molecular modelling. Basic principles and applications*. Wiley-VCH, Weinheim, 2nd edition.

Hong, H., Yang, L., and Stallcup, M. R. (1999). Hormone-independent transcriptional activation and coactivator binding by novel orphan nuclear receptor ERR3. *J. Biol. Chem.*, 274:22618–22626.

Honkakoski, P., Moore, R., Washburn, K. A., and Negishi, M. (1998a). Activation by diverse xenochemicals of the 51-base pair phenobarbital-responsive enhancer module in the CYP2b10 gene. *Mol. Pharmacol.*, 53:597–601.

Honkakoski, P. and Negishi, M. (1997). Characterization of a phenobarbital-responsive enhancer module in mouse p450 cyp2b10 gene. *J. Biol. Chem.*, 272:14943–14949.

Honkakoski, P., Palvimo, J. J., Penttilä, L., Vepsäläinen, J., and Auriola, S. (2004). Effects of triaryl phosphates on mouse and human nuclear receptors. *Biochem. Pharmacol.*, 67:97–106.

Honkakoski, P., Sueyoshi, T., and Negishi, M. (2003). Drug-activated nuclear receptors CAR and PXR. *Ann. Med.*, 35:172–182.

Honkakoski, P., Zelko, I., Sueyoshi, T., and Negishi, M. (1998b). The nuclear orphan receptor CAR-retinoid x receptor heterodimer activates the phenobarbital-responsive enhancer module of the CYP2b gene. *Mol. Cell. Biol.*, 18:5652–5658.

Hooft, R. W., Vriend, G., Sander, C., and Abola, E. E. (1996). Errors in protein structures. *Nature*, 381:272.

Hopfinger, A. J., REaka, A., Venkatarangan, P., Duca, J. S., and Wang, S. (1999). Construction of a virtual high throughput screen by 4d-QSAR analysis: application to a combinatorial library of glucose inhibitors of glycogen phosphorylase b. *J. Chem. Inf. Comput. Sci.*, 39:1151–1160.

Hörlein, A. J., Näär, A. M., an J. Torchia, T. H., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Söderström, M., Glass, C. K., and Rosenfeld, M. G. (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature*, 377:397–404.

Hosseinpour, F., Moore, R., Negishi, M., and Sueyoshi, T. (2006). Serine 202 regulates the nuclear translocation of constitutive active/androstane receptor. *Mol. Pharmacol.*, 69:1095–1102.

Housley, P. R., Sanchez, E. R., Danielsen, M., Ringold, G. M., and Pratt, W. (1990). Evidence that the conserved region in the steroid binding domain of the glucocorticoid receptor is required for both optimal binding of hsp90 and protection from proteolytic cleavage. a two-site model for hsp90 binding to the steroid binding domain. *J. Biol. Chem.*, 265:12778–12781.

Hsiao, P. W., Deroo, B. J., and Archer, T. K. (2002). Chromatin remodeling and tissue-selective responses of nuclear hormone receptors. *Biochem. Chell Biol.*, 80:343–351.

Huang, W., Zhang, J., Chua, S. S., Qatanani, M., Han, Y., Granata, R., and Moore, D. D. (2003). Induction of bilirubin clearance by the constitutive androstane receptor (CAR). *Proc. Natl. Acad. Sci. U S A*, 100:4156–4161.

Huang, W., Zhang, J., and Moore, D. D. (2004a). A traditional herbal medicine enhances bilirubin clearance by activating the nuclear receptor CAR. *J. Clin. Invest.*, 113:137–143.

Huang, W., Zhang, J., Washington, M., Liu, J., Parant, J. M., Lozano, G., and Moore, D. D. (2005). Xenobiotic stress induces hepatomegaly and liver tumors via the nuclear receptor constitutive androstane receptor. *Mol. Endocrinol.*, 19:1646–1653.

Huang, W., Zhang, J., Wei, P., Schrader, W. T., and Moore, D. D. (2004b). Meclizine is an agonist ligand for mouse constitutive androstane receptor (CAR) and an inverse agonist for human car. *Mol. Endocrinol.*, 18:2402–2408.

Jackson, J. P., Ferguson, S. S., Moore, R., Negishi, M., and Goldstein, J. A. (2004). The constitutive active/androstane receptor regulates phenytoin induction of cyp2c29. *Mol. Pharmacol.*, 65:1397–1404.

Jackson, T. A., Richter, C. K., Bain, D. L., Takimoto, G. S., Tung, L., and Horwitz, K. B. (1997). The partial agonist activity of antagonist-occupied steroid receptors is controlled by a novel hinge domain-binding coactivator l7/SPA and the corepressors n-cor or SMRT. *Mol. Endocrinol.*, 11:693–705.

Jacobs, M. N., Dickins, M., and Lewis, D. F. (2003). Homology modelling of the nuclear receptors: human oestrogen receptorβ (hERβ), the human pregnane-x-receptor (PXR), the ah receptor (ahr) and the constitutive androstane receptor (CAR) ligand binding domains from the human oestrogen receptor α (hERα) crystal structure, and the human peroxisome proliferator activated receptor α (PPARα) ligand binding domain from the human PPARγ crystal structure. *J. Steroid. Biochem. Mol. Biol.*, 84:117–132.

Johnson, M. S., Srinivasan, N., Sowdhamini, R., and Blundell, T. L. (1994). Knowledge-based protein modeling. *Crit. Rev. Biochem. Mol. Biol.*, 29:1–68.

Jones, D. and Thornton, J. (1993). Protein fold recognition. *J. Comput. Aided Mol. Des.*, 7:439–456.

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). A new approach to protein fold recognition. *Nature*, 358:86–89.

Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.*, 267:727–748.

Jones, T. A. and Thirup, S. (1986). Using known substructures in protein model building and crystallography. *EMBO J.*, 5:819–822.

Jorgensen, W. L. and Duffy, E. M. (2002). Prediction of drug solubility from structure. *Adv. Drug. Deliv. Rev.*, 54:355–366.

Jorgensen, W. L. and Tirado-Rives, J. (1988). The OPLS force field for proteins. energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, 110:1657–1666.

Joseph, S. B. and Tontonoz, P. (2003). LXRs: new therapeutic targets in atherosclerosis? *Curr. Opin. Pharmacol.*, 3:192–197.

Juge-Aubry, C., Pernin, A., Favez, T., Burger, A. G., Wahli, W., Meier, C. A., and Desvergne, B. (1997). DNAbinding properties of peroxisome proliferator-activated receptor subtypes on various natural peroxisome proliferator response elements. importance of the 5′-flanking region. *J. Biol. Chem.*, 272:25252–25259.

Jyrkkärinne, J., Mäkinen, J., Gynther, J., Savolainen, H., Poso, A., and Honkakoski, P. (2003). Molecular determinants of steroid inhibition for the mouse constitutive androstane receptor. *J. Med. Chem.*, 46:4687–4695.

Jyrkkärinne, J., Windshügel, B., Mäkinen, J., Ylisirniö, M., Peräkylä, M., Poso, A., Sippl, W., and Honkakoski, P. (2005). Amino acids important for ligand specificity of the human constitutive androstane receptor. *J. Biol. Chem.*, 280:5960–5971.

Kabsch, W. and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22:2577–2637.

Kallen, J., Schlaeppi, J. M., Bitsch, F., Filipuzzi, I., Schilb, A., Riou, V., Graham, A., Strauss, A., Geiser, M., and Fournier, B. (2004). Evidence for ligand-independent transcriptional activation of the human estrogen-related receptor α (errα): crystal structure of err? ligand binding domain in complex with peroxisome proliferator-activated receptor coactivator-1α. *J. Biol. Chem.*, 279:49330–49337.

Kamisako, T., Leier, I., Cui, Y., König, J., Buchholz, U., Hummel-Eisenbeis, J., and Keppler, D. (1999). Transport of monoglucuronosyl and bisglucuronosyl bilirubin by recombinant human and rat multidrug resistance protein 2. *Hepatology*, 30:485–490.

Karplus, M. and Kuriyan, J. (2005). Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U S A*, 102:6679–6685.

Karplus, M. and McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. nat. struct. biol. 9, 646-652. *Nat. Struct. Biol.*, 9:646–652.

Kast, H. R., Goodwin, B., Tarr, P. T., Stones, S. A., Anisfeld, A. M., Stoltz, C. M., Tontonoz, P., Kliewer, S., Willson, T. M., and Edwards, P. A. (2002). Regulation of multidrug resistance-associated protein 2 (ABCC2) by the nuclear receptors pregnane x receptor, farnesoid x-activated receptor, and constitutive androstane receptor. *J. Biol. Chem.*, 277:2908–2915.

Kastner, P., Krust, A., Turcotte, B., Stropp, U., Tora, L., Gronemeyer, H., and Chambon, P. (1990). Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms a and b. *EMBO J.*, 9:1603–1614.

Kawamoto, T., Kakizaki, S., Yoshinari, K., and Negishi, M. (2000). Estrogen activation of the nuclear orphan receptor CAR (constitutive active receptor) in induction of the mouse cyp2b10 gene. *Mol. Endocrinol.*, 14:1897–1905.

Kawamoto, T., Sueyoshi, T., Zelko, I., Moore, R., Washburn, K., and Negishi, M. (1999). Phenobarbital-responsive nuclear translocation of the receptor CAR in induction of the CYP2b gene. *Mol. Cell. Biol.*, 19:6318–6322.

Kelley, L. A., Gardner, S. P., and Sutcliffe, M. J. (1996). An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. *Prot. Eng.*, 9:1063–1065.

Khorasanizadeh, S. and Rastinejad, F. (2001). Nuclear-receptor interactions on DNA-response elements. *Trends Biochem. Sci*, 26:384–390.

Kim, H. J., Lee, S. K., Na, S. Y., Choi, H. S., and Lee, J. W. (1998). Molecular cloning of xSRC-3, a novel transcription coactivator from xenopus, that is related to AIB1, p/CIP, and TIF2. *Mol. Endocrinol.*, 12:1038–1047.

Kim, R. B. (2002). Transporters and xenobiotic disposition. *Toxicology*, 181-182:291–297.

Kishimoto, M., Fujiki, R., Takezawa, S., Sasaki, Y., Nakamura, T., Yamaoka, K., Kitagawa, H., and Kato, S. (2006). Nuclear receptor mediated gene regulation through chromatin remodeling and histone modifications. *Endocr. J.*, 53:157–172.

Klebe, G. (2000). Recent developments in structure-based drug design. *J. Mol. Med.*, 87:269–281.

Kliewer, S. A., Goodwin, B., and Willson, T. M. (2002). The nuclear pregnane x receptor: a key regulator of xenobiotic metabolism. *Endocr. Rev.*, 23:687–702.

Kliewer, S. A., Moore, J. T., Wade, L., Staudinger, J. L., Watson, M. A., Jones, S. A., McKee, D. D., Oliver, B. B., Willson, T. M., Zetterstrom, R. H., and Lehmann, T. P. J. (1998). An orphan nuclear receptor activated by pregnanes defines a novel steroid signaling pathway. *Cell*, 92:73–82.

Kobayashi, K., Sueyoshi, T., Inoue, K., Moore, R., and Negishi, M. (2003). Cytoplasmic accumulation of the nuclear receptor CAR by a tetratricopeptide repeat protein in hepg2 cells. *Mol. Pharmacol.*, 64:1069–1075.

Kobayashi, K., Yamanaka, Y., Iwazaki, N., Nakajo, I., Hosokawa, M., Negishi, M., and Chiba, K. (2005). Identification of HMG-coa reductase inhibitors as activators for human, mouse and rat constitutive androstane receptor. *Drug Metab. Disp.*, 33:924–929.

Koike, C., Moore, M., and Negishi, M. (2005). Localization of the nuclear receptor CAR at the cell membrane of mouse liver. *FEBS Lett.*, 579:6733–6736.

König, J., Cui, Y., Nies, A. T., and Keppler, D. (2000). A novel human organic anion transporting polypeptide localized to the basolateral hepatocyte membrane. *Am. J. Physiol. Gastrointest. Liver. Physiol.*, 278:G156–G164.

Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and dna. *Science*, 184:868–871.

Kostrubsky, V. E., Lewis, L. D., Strom, S. C., Wood, S. G., Schuetz, E. G., Schuetz, J. D., Sinclair, P. R., Wrighton, S. A., and Sinclair, J. F. (1998). Induction of cytochrome p4503a by taxol in primary cultures of human hepatocytes. *Arch. Biochem. Biophys.*, 355:131–136.

Kraus, W. L. and Wong, J. (2002). Nuclear receptor-dependent transcription with chromatin. is it all about enzymes? *Eur. J. Biochem.*, 269:2275–2283.

Krust, A., Green, S., Argos, P., Kumar, V., Walter, P., Bornert, J. M., and Chambon, P. (1986). The chicken oestrogen receptor sequence: homology with v-erba and the human oestrogen and glucocorticoid receptors. *EMBO J.*, 5:891–897.

Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., and Ferrin, T. E. (1982). A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.*, 161:269–288.

Lamba, J. K., Lamba, V., Yasuda, K., Lin, Y. S., Assem, M., Thompson, E., Strom, S., and Schuetz, E. (2004). Expression of constitutive androstane receptor splice variants in human tissues and their functional consequences. *J. Pharmacol. Exp. Ther.*, 311:811–821.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 26:283–291.

Laudet, V. (1997). Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. *J. Mol. Endocrinol.*, 19:207–226.

Laudet, V. and Gronemeyer, H. (2001). *The nuclear receptors factsbook*. Academic Press, London, UK.

Leach, A. R. (1994). Ligand docking to proteins with discrete side-chain flexibility. *J. Mol. Biol.*, 235:345–356.

Leach, A. R. (2001). *Molecular modelling. Principles and applications*. Pearson Education Ltd., Essex, UK, 2nd edition.

Leach, A. R. and Kuntz, I. D. (1990). Conformational analysis of flexible ligands in macromolecular receptor site. *J. Comp. Chem.*, 13:730–748.

Lees, J. A., Fawell, S. E., White, R., and Parker, M. G. (1990). A 22-amino-acid peptide restores DNA-binding activity to dimerization-defective mutants of the estrogen receptor. *Mol. Cell. Biol.*, 10:5529–5531.

Lehmann, J. M., McKee, D. D., Watson, M. A., Willson, T. M., Moore, J. T., and Kliewer, S. A. (1998). The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3a4 gene expression and cause drug interactions. *J. Clin. Invest.*, 102:1016–1023.

Lempiäinen, H., Molnar, F., Gonzalez, M. M., Peräkylä, M., and Carlberg, C. (2005). Antagonist- and inverse agonist-driven interactions of the vitamin d receptor and the constitutive androstane receptor with co-repressor protein. *Mol. Endocrinol.*, 19:2258–2272.

Leo, C. and Chen, J. D. (2000). The SRC family of nuclear receptor coactivators. *Gene*, 245:1–11.

Li, Y., Lambert, M. H., and Xu, H. E. (2003). Activation of nuclear receptors: a perspective from structural genomics. *Structure*, 11:741–746.

Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Deliv. Rev.*, 46:3–26.

Liu, Z., Auboeuf, D., Wong, J., Chen, J. D., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W. (2002). Coactivator/corepressor ratios modulate PR-mediated transcription by the selective receptor modulator RU486. *Proc. Natl. Acad. Sci. U S A*, 99:7940–7944.

Lo Conte, L., Ailey, B., Hubbard, T. J. P., Brenner, S. E., Murzin, A. G., and Chothia, C. (2000). SCOP: a structural classification of protein databases. *Nucleic Acids Res.*, 28:257–259.

Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R., and Sigler, P. B. (1991). Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature*, 352:497–505.

Lüthy, R., Bowie, J. U., and Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, 356:83–85.

Mackerell Jr., A. D. (2004). Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.*, 25:1584–1604.

Mader, S., Leroy, P., Chen, J. Y., and Chambon, P. (1993). Multiple parameters control the selectivity of nuclear receptors for their response elements. selectivity and promiscuity in response element recognition by retinoic acid receptors and retinoid x receptors. *J. Biol. Chem.*, 268:591–600.

Maglich, J. M., Parks, D. J., Moore, L. B., Collins, J. L., Goodwin, B., Bilin, A. N., Stoltz, C. A., Kliewer, S. A., Lambert, M. H., Willson, T. M., and Moore, J. T. (2003). Identification of a novel human constitutive androstane receptor (CAR) agonist and its use in the identification of CAR target genes. *J. Biol. Chem.*, 278:17277–17283.

Maglich, J. M., Stoltz, C. M., Goodwin, B., Hawkins-Brown, D., Moore, J. T., and Kliewer, S. A. (2002). Nuclear pregnane x receptor and constitutive androstane receptor regulate overlapping but distinct sets of genes involved in xenobiotic detoxification. *Mol. Pharmacol.*, 62:638–646.

Maglich, J. M., Watson, J., McMillen, P. J., Goodwin, B., Willson, T. M., and Moore, J. T. (2004). The nuclear receptor CAR is a regulator of thyroid hormone metabolism during caloric restriction. *J. Biol. Chem.*, 279:19832–19838.

Maher, J. M., Cheng, X., Slitt, A. L., Dieter, M. Z., and Klassen, C. D. (2005). nduction of the multidrug resistance-associated protein family of transporters by chemical activators of receptor-mediated pathways in mouse live. *Drug Metab. Disp.*, 33:956–962.

Mäkinen, J., Frank, C., Jyrkkärinne, J., Gynther, J., Carlberg, C., and Honkakoski, P. (2002). Modulation of mouse and human phenobarbital-responsive enhancer module by nuclear receptors. *Mol. Pharmacol.*, 62:366–378.

Mäkinen, J., Reinisalo, M., Niemi, K., Viitala, P., Jyrkkärinne, J., Chung, H., Pelkonen, O., and Honkakoski, P. (2003). Dual action of oestrogens on the mouse constitutive androstane receptor. *Biochem. J.*, 376:465–472.

Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schütz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., and Evans, R. M. (1995). The nuclear receptor superfamily: the second decade. *Cell*, 83:835–839.

Mannervik, B. (1985). The isoenzymes of glutathione transferase. *Adv. Enzymol. Relat. Areas Mol. Biol.*, 57:357–417.

Marimuthu, A., Feng, W., Tagami, T., Nguyen, H., Jameson, J. L., Fletterick, R. J., Baxter, J. D., and West, B. L. (2002). Tr surfaces and conformations required to bind nuclear receptor corepressor. *Mol. Endocrinol.*, 16:271–286.

Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F., and Sali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.*, 29:291–325.

McCammon, J. A., Gelin, B. R., and Karplus, M. (1977). Dynamics of folded proteins. *Nature*, 265:585–590.

McDonnell, D. P., Mangelsdorf, D. J., Pike, J. W., Haussler, M.-R., and O'Malley, B. W. (1987). Molecular cloning of complementary DNA encoding the avian receptor for vitamin d. *Science*, 235:1214–1217.

McKenna, N. J., Lanz, R. B., and O'Malley, B. W. (1999). Nuclear receptor coregulators: cellular and molecular biology. *Endocr. Rev.*, 20:321–344.

Miesfeld, R., Rusconi, S., Godowski, P. J., Maler, B. A., Okret, S., Wikstrom, A. C., Gustafsson, J. A., and Yamamoto, K. R. (1986). Genetic complementation of a glucocorticoid receptor deficiency by expression of cloned receptor cDNA. *Cell*, 46:389–399.

Min, G., Kemper, J. K., and Kemper, B. (2002). Glucocorticoid receptor-interacting protein 1 mediates ligand-independent nuclear translocation and activation of constitutive androstane receptor in vivo. *J. Biol. Chem.*, 277:26356–26363.

Molnar, F., Matilainen, M., and Carlberg, C. (2005). Structural determinants of the agonist-independent association of human peroxisome proliferator-activated receptors with coactivators. *J. Biol. Chem.*, 275:26543–26556.

Moore, D. D. (2005). CAR: Three new models for a problem child. *Cell Metab.*, 1:6–8.

Moore, L. B., Maglich, J. M., McKee, D. D., Wisely, B., Willson, T. M., Kliewer, S. A., Lambert, M. H., and Moore, J. T. (2002). Pregnane x receptor (PXR), constitutive androstane receptor (CAR), and benzoate x receptor (BXR) define three pharmacologically distinct classes of nuclear receptors. *Mol. Endocrinol.*, 16:977–986.

Moore, L. B., Parks, D. J., Jones, S. A., Bledsoe, R. K., Consler, T. G., Stimmel, J. B., Goodwin, B., Liddle, C., Blanchard, S. G., Willson, T. M., Collins, J. L., and Kliewer, S. A. (2000). Orphan nuclear receptors constitutive androstane receptor and pregnane x receptor share xenobiotic and steroid ligands. *J. Biol. Chem.*, 275:15122–15127.

Moras, D. and Gronemeyer, H. (1998). The nuclear receptor ligand-binding domain: structure and function. *Curr. Opin. Cell Biol.*, 10:384–391.

Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins*, 12:345–364.

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998). Automated docking using a lamarckian genetic algorithm and and empirical binding free energy function. *J. Comp. Chem.*, 19:1639–1662.

Muangmoonchai, R., Smirlis, D., Wang, S. C., Edwards, M., Philips, I. R., and Shephard, E. A. (2001). Xenobiotic induction of cytochrome p450 2b1 (CYP2B1) is mediated by the orphan nuclear receptor constitutive androstane receptor (CAR) and requires steroid co-activator 1 (src-1) and the transcription factor sp1. *Biochem. J.*, 355:71–78.

Muegge, I. and Martin, Y. C. (1999). A general and fast scoring function for protein-ligand interactions: a simplified potential approach. *J. Med. Chem.*, 42:791–804.

Naar, A. M., Boutin, J. M., Lipkin, S. M., Yu, V. C., Holloway, J. M., Glass, C. K., and Rosenfeld, M. G. (1991). The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. *Cell*, 65:1267–1279.

Nagpal, S., Na, S., , and Rathnachalam, R. (2005). Noncalcemic actions of vitamin d receptor ligands. *Endocr. Rev.*, 26:662–687.

Negishi, M. and Honkakoski, P. (2000). Induction of drug metabolism by nuclear receptor CAR: molecular mechanisms and implications for drug research. *Eur. J. Pharm. Schi.*, 11:259–264.

Nolte, R. T., Wisely, G. B., Westin, S., Cobb, J. E., Lambert, M. H., Kurokawa, R., Rosenfeld, M. G., Willson, T. M., Glass, C. K., and Milburn, M. V. (1998). Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-γ. *Nature*, 395:137–143.

Okey, A. B. (1990). Enzyme induction in the cytochrome P-450 system. *Pharmacol. Ther.*, 45:241–298.

Orans, J., Teotico, D. G., and Redinbo, M. R. (2005). The nuclear xenobiotic receptor pregnane x receptor: recent insights and new challenges. *Mol. Endocrinol.*, 19:2891–2900.

Owen, G. I. and Zelent, A. (2000). Origins and evolutionary diversification of the nuclear receptor superfamily. *Cell. Mol. Life Sci.*, 57:809–827.

Pan, Y., Huang, N., Cho, S., and Mackerell, A. D. (2003). Consideration of molecular weight during compound selection in virtual target-based database screening. *J. Chem. Inf. Comput. Sci.*, 43:267–272.

Pastor, M., Cruciani, G., and Watson, K. A. (1997). A strategy for the incorporation of water molecules present in a ligand binding site into a three-dimensional quantitative structure–activity relationship analysis. *J. Med. Chem.*, 40:4089–4102.

Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham 3rdand S. DeBolt, T. E., Ferguson, D., Seibel, G., and Kollman, P. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Commun.*, 91:1–41.

Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.*, 183:63–98.

Perez-Juste, G., Garcia-Silva, S., and Aranda, A. (2000). An element in the region responsible for premature termination of transcription mediates repression of c-myc gene expression by thyroid hormone in neuroblastoma cells. *J. Biol. Chem.*, 275:1307–1314.

Perola, E., Xu, K., Kollmeyer, T. M., Kaufmann, S. H., Prendergast, F. G., and Pang, Y. P. (2000). Successful virtual screening of a chemical database for farnesyltransferase inhibitor leads. *J. Med. Chem.*, 43:401–408.

Perutz, M. F., Kendrew, J. C., and Watson, H. C. (1965). Structure and function of haemoglobin. *J. Mol. Biol.*, 13:669–678.

Petkovich, M., Brand, N. J., Krust, A., and Chambon, P. (1987). A human retinoic acid receptor which belongs to the family of nuclear receptors. *Nature*, 330:444–450.

Picard, D. and Yamamoto, K. (1987). Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. *EMBO J.*, 6:3333–3340.

Ponder, J. W. and Richards, F. M. (1987). Tertiary templates for proteins. use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.*, 193:775–791.

Qatanani, M., Zhang, J., and Moore, D. D. (2005). Role of the constitutive androstane receptor in xenobiotic-induced thyroid hormone metabolism. *Endocrinology*, 146:995–1002.

Rachez, C. and Freedman, L. P. (2001). Mediator complexes and transcription. *Curr. Opin. Cell Biol.*, 13:274–280.

Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.*, 7:95–99.

Rarey, M., Kramer, B., Lengauer, T., and Klebe, G. (1996). A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.*, 261:470–489.

Remmer, H. (1958). Die beschleunigung des abbaues als ursache der gewhnung an barbiturate. *Naturwissenschaften*, 46:580–581.

Renaud, J. P. and Moras, D. (2000). Structural studies on nuclear receptors. *Cell. Mol. Life Sci.*, 57:1748–1769.

Renaud, J. P., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer, H., and Moras, D. (1995). Crystal structure of the rar-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature*, 378:681–689.

Robinson-Rechavi, M., Carpentier, A., Duffraise, M., and Laudet, V. (2002). How many nuclear hormone receptors are there in the human genome? *Trends Genet.*, 17:554–556.

Robinson-Rechavi, M., Garcia, H. E., and Laudet, V. (2003). The nuclear receptor superfamily. *J. Cell Sci.*, 116:585–586.

Robinson-Rechavi, M. and Laudet, V. (2003). Bioinformatics of nuclear receptors. *Methods Enzymol.*, 364:95–118.

Rochel, N., Tocchini-Valentini, G., Egea, P. F., Juntunen, K., Garnier, J. M., Vihko, P., and Moras, D. (2001). Functional and structural characterization of the insertion region in the ligand binding domain of the vitamin d nuclear receptor. *Eur. J. Biochem.*, 268:971–979.

Rochel, N., Wurtz, J. M., Mitschler, A., Klaholz, B., and Moras, D. (2000). The crystal structure of the nuclear receptor for vitamin d bound to its natural ligand. *Mol. Cell*, 5:173–179.

Rodrigues, A. D. and Lin, J. H. (2001). Screening of drug candidates for their drug–drug interaction potential. *Curr. Opin. Chem. Biol.*, 5:396–401.

Rosenfeld, J. M., Vargas Jr., R., Xie, W., and Evans, R. M. (2003). Genetic profiling defines the xenobiotic gene network controlled by the nuclear receptor pregnane x receptor. *Mol. Endocrinol.*, 17:1268–1282.

Rost, B. (1999). Twilight zone of protein sequence alignments. *Protein Eng.*, 12:85–94.

Roy-Chowhury, J., Wolkoff, A. W., Roy-Chowdhury, N., and Arias, I. M. (1995). *The metabolic and molecular basis of inherited diseases*, chapter Hereditary jaundice and disorders of bilirubin metabolism. McGraw-Hill, 7th edition.

Saatcioglu, F., Bartunek, P., Deng, T., Zenke, M., and Karin, M. (1993). A conserved c-terminal sequence that is deleted in v-erba is essential for the biological activities of c-erba (the thyroid hormone receptor). *Mol. Cell. Biol.*, 13:3675–3685.

Sablin, E. P., Krylova, I. N., Fletterick, R. J., and Ingraham, H. A. (2003). Structural basis for ligand-independent activation of the orphan nuclear receptor LRH-1. *Mol. Cell*, 11:1575–1585.

Sagui, C. and Darden, T. A. (1999). Molecular dynamics simulation of biomolecules: Long-range electrostatic effects. *Annu. Rev. Biophys. Biomol. Struct*, 28:155–179.

Saini, S. P., Sonoda, J., Xu, L., Toma, D., Uppal, H., Mu, Y., Ren, S., More, D. D., Evans, R. M., and Xie, W. (2004). A novel constitutive androstane receptor-mediated and CYP3a-independent pathway of bile acid detoxification. *Mol. Pharmacol.*, 65.

Sander, C. and Schneider, R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins*, 9:56–68.

Schaefer, M., Bartels, C., and Karplus, M. (1998). Solution conformations and thermodynamics of structured peptides: molecular dynamics simulation with an implicit solvation model. *J. Mol. Biol.*, 284:835–848.

Schaefer, M. and Karplus, M. (1996). A comprehensive analytical treatment of continuum electrostatics. *J. Phys. Chem.*, 100:1578–1699.

Schapira, M., Abagyan, R., and Totrov, M. (2003a). Nuclear hormone receptor targeted virtual screening. *J. Med. Chem.*, 46:3045–3059.

Schapira, M., Raaka, B. M., Das, S., Fan, L., Totrov, M., Zhou, Z., Wilson, S. R., Abagyan, R., and Samuels, H. H. (2003b). Discovery of diverse thyroid hormone receptor antagonists by high-throughput docking. *Proc. Natl. Acad. Sci. U S A*, 100:7354–7359.

Schapira, M., Raaka, B. M., Samuels, H. H., and Abagyan, R. (2000). Rational discovery of novel nuclear hormone receptor antagonists. *Proc. Natl. Acad. Sci. U S A*, 97:1008–1013.

Schwabe, J. W., Chapman, L., Finch, J. T., and Rhodes, D. (1993). The crystal structure of the estrogen receptor dna-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell*, 75:567–578.

Scott, W. R. P., Hünenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., Torda, A. E., Huber, T., Krüger, P., and van Gunsteren, W. F. (1999). The GROMOS biomolecular simulation program package. *J. Phys. Chem. A*, 103:3596–3607.

Segraves, W. A. (1991). omething old, some things new: the steroid receptor superfamily in drosophila. *Cell*, 67:225–228.

Semple, R. K., Krishna, V., Chatterjee, K., and O'Rahilly, S. (2006). Pparγ and human metabolic disease. *J. Clin. Invest.*, 116:581–589.

Shan, L., Vincent, J., Brunzelle, J. S., Dussault, I., Lin, M., Ianculescu, I., Sherman, M. A., Forman, B. M., and Fernandez, E. J. (2004). Structure of the murine constitutive androstane receptor complexed to androstenol: a molecular basis for inverse agonism. *Mol. Cell.*, 16:907–917.

Shao, D. and Lazar, M. A. (1999). Modulating nuclear receptor function: may the phos be with you. *J. Clin. Invest.*, 103:1617–1618.

Shenkin, P. S., Yarmush, D. L., Fine, R. M., Wang, H., and Levinthal, C. (1987). Predicting antibody hypervariable loop conformation. i. ensembles of random conformations for ringlink structures. *Biopolymers*, 26:2053–2085.

Sheridan, R. P., Rusinko, A., Nilakantan, R., and Venkataraghavan, R. (1989). Searching for pharmacophores in large coordinate data bases and its use in drug design. *Proc. Natl. Acad. Sci. U S A*, 86:8156–8159.

Shiau, A. K., Barstad, D., Loria, P. M., Cheng, L., Kushner, P. J., Agard, D. A., and Greene, G. L. (1998). The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell*, 95:927–937.

Shiraki, T., Sakai, N., Kanaya, E., and Jingami, H. (2003). Activation of orphan nuclear constitutive androstane receptor requires subnuclear targeting by peroxisome proliferator-activated receptor γ coactivator-1 α. a possible link between xenobiotic response and nutritional state. *J. Biol. Chem.*, 278:11344–11350.

Sippl, M. J. (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins*, 17:355–362.

Smirlis, D., Muangmoonchai, R., Edwards, M., Philips, I. R., and Shephard, E. A. (2001). Orphan receptor promiscuity in the induction of cytochromes p450 by xenobiotics. *J. Biol. Chem.*, 276:12822–12826.

Smith, C. L., Nawaz, Z., and O'Malley, B. W. (1997). Coactivator and corepressor regulation of the agonist/antagonist activity of the mixed antiestrogen, 4-hydroxytamoxifen. *Mol. Endocrinol.*, 11:657–666.

Smith, D. P., Mason, C. S., Jones, E. A., and Old, R. W. (1994). A novel nuclear receptor superfamily member in xenopus that associates with RXR, and shares extensive sequence similarity to the mammalian vitamin d3 receptor. *Nucleic Acids Res.*, 22:66–71.

Staudinger, J. L., Jones, B. G. S. A., Hawkins-Brown, D., LaTour, K. I. M. A., Liu, Y., Klassen, C. D., Brown, K. K., reinhard, J., Willson, T. M., Koller, B. H., and Kliewer, S. A. (2001). The nuclear receptor pxr is a lithocholic acid sensor that protects against liver toxicity. *Proc. Natl. Acad. Sci. USA*, 98:3369–3374.

Stehlin, C., Wurtz, J. M., Steinmetz, A., Greiner, E., Schule, R., Moras, D., and Renaud, J. P. (2001). X-ray structure of the orphan nuclear receptor RORβ ligand-binding domain in the active conformation. *EMBO J.*, 20:5822–5831.

Steinmetz, A. C., Renaud, J. P., and Moras, D. (2001). Binding of ligands and activation of transcription by nuclear receptors. *Annu. Rev. Biophys. Biomol. Struct.*, 30:329–359.

Still, W. C., Tempczyrk, A., Hawley, R. C., and Hendrickson, T. (1990). Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.*, 112:6127–6129.

Sueyoshi, T., Kawamoto, T., Zelko, I., Honkakoski, P., and Negishi, M. (1999). The repressed nuclear receptor CAR responds to phenobarbital in activating the human CYP2b6 gene. *J. Biol. Chem.*, 274:6043–6046.

Sugatani, J., Kojima, H., Ueda, A., Kakizaki, S., Yoshinari, K., Gong, Q. H., Owens, I. S., Negishi, M., and Sueyoshi, T. (2001). The phenobarbital response enhancer module in the human bilirubin UDP-glucuronosyltransferase UGT1a1 gene and regulation by the nuclear receptor CAR. *Hepatology*, 33:1232–1238.

Suino, K., Peng, L., Reynolds, R., Li, Y., Cha, J. Y., Repa, J. J., Kliewer, S. A., and Xu, H. E. (2004). The nuclear xenobiotic receptor CAR: structural determinants of constitutive activation and heterodimerization. *Mol. Cell*, 16:893–905.

Thompson, J. D., Higgins, D. G., and 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.

Tocchini-Valentini, G., Rochel, N., Wurtz, J. M., Mitschler, A., and Moras, D. (2001). Crystal structures of the vitamin d receptor complexed to superagonist 20-epi ligands. *Proc. Natl. Acad. Sci. U S A*, 98:5491–5496.

Toell, A., Kroencke, K. D., Kleinert, H., and Carlberg, C. (2002). Orphan nuclear receptor binding site in the human inducible nitric oxide synthase promoter mediates responsiveness to steroid and xenobiotic ligands. *J. Cell. Biochem.*, 85:72–82.

Tora, L., Gronemeyer, H., Turcotte, B., Gaub, M. P., and Chambon, P. (1988). The n-terminal region of the chicken progesterone receptor specifies target gene activation. *Nature*, 333:185–188.

Tora, L., White, J., Brou, C., Tasset, D., Webster, N., Scheer, E., and Chambon, P. (1989). The human estrogen receptor has two independent nonacidic transcriptional activation functions. *Cell*, 59:477–487.

Tukey, R. H. and Strassburg, C. P. (2000). Human UDP-glucuronosyl-transferases: metabolism, expression, and disease. *Annu. Rev. Pharmacol. Toxicol.*, 40:581–616.

Tzameli, I., Pissios, P., Schuetz, E. G., and Moore, D. D. (2000). The xenobiotic compound 1,4-bis[2-(3,5-dichloropyridyloxy)]benzene is an agonist ligand for the nuclear receptor CAR. *Mol. Cell. Biol.*, 20:2951–2958.

Ueda, A., Hamadeh, H. K., Webb, H. K., Yamamoto, Y., Sueyoshi, T., Ahshari, C. A., Lehmann, J. M., and Negishi, M. (2002). Diverse roles of the nuclear orphan receptor CAR in regulating hepatic genes in response to phenobarbital. *Mol. Pharmacol.*, 61:1–6.

Umesono, K. and Evans, R. M. (1989). Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell*, 57:1139–1146.

Umesono, K., Murakami, K. K., Thompson, C. C., and Evans, R. M. (1991). Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin d3 receptors. *Cell*, 65:1255–1266.

Väisänen, S., Peräkylä, M., Kärkkäinen, J. I., Steinmeyer, A., and Carlberg, C. (2002). Critical role of helix 12 of the vitamin D(3) receptor for the partial agonism of carboxylic ester antagonists.

van der Waterbeemd, H., Camenisch, G., Folkerts, G., Chretien, J. R., and Raevsky, O. A. (1998). Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and h-bonding descriptors. *J. Drug Target.*, 6:151–165.

van der Waterbeemd, H. and Jones, B. C. (2003). Predicting oral absorption and bioavailability. *Prog. Med. Chem.*, 41:1–59.

van Gunsteren, W. F. and Mark, A. E. (1998). Validation of molecular dynamics simulation. *J. Chem. Phys.*, 108:6109–6116.

Vangrevelinghe, E., Zimmermann, K., Schoepfer, J., Portmann, R., Fabbro, D., and Furet, P. (2003). Discovery of a potent and selective protein kinase ck2 inhibitor by high-throughput docking. *J. Med. Chem.*, 46:2656–2662.

Vegeto, E., Shahbaz, M. M., Wen, D. X., Goldman, M. E., O'Malley, B. W., and McDonnell, D. P. (1993). Human progesterone receptor a form is a cell- and promoter-specific repressor of human progesterone receptor b function. *Mol. Endocrinol.*, 7:1244–1255.

Visser, T. J. (1996). Pathways of thyroid hormone metabolism. *Acta Med. Austriaca*, 23:10–16.

Visser, T. J., Kaptein, E., Glatt, H., Bartsch, I., Hagen, M., and Coughtrie, M. W. (1998). Characterization of thyroid hormone sulfotransferases. *Chem. Biol. Interact.*, 109:279–291.

Visser, T. J., Kaptein, E., van Raaij, J. A., Joe, C. T., Ebner, T., and Burchell, B. (1993). Multiple UDP-glucuronyltransferases for the glucuronidation of thyroid hormone with preference for 3,3',5'-triiodothyronine (reverse t3). *FEBS Lett.*, 315:65–68.

von Itzstein, M., Wu, W. Y., Kok, G. B., Pegg, M. S., Dyason, J. C., Jin, B., van Phan, T., Smythe, M. L., White, H. F., Oliver, S. W., Colman, P. M., Varghese, J. N., Ryan, D. M., Woods, J. M., Bethell, R. C., Hotham, V. J., Cameron, J. M., and Penn, C. R. (1993). Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature*, 363:418–423.

Wagner, M., Halilbasic, E., Marschall, H. U., Zollner, G., Fickert, P., Langner, C., Zatloukal, K., Denk, H., and Trauner, M. (2005). CAR and PXR agonists stimulate hepatic bile acid and bilirubin detoxification and elimination pathways in mice. *Hepatology*, 42:420–430.

Wagner, R. L., Apriletti, J. W., McGrath, M. E., West, B. L., Baxter, J. D., and Fletterick, R. J. (1995). A structural role for hormone in the thyroid hormone receptor. *Nature*, 378:690–697.

Wang, H., Faucette, S., Moore, R., Sueyoshi, T., Negishi, M., and E., L. (2004a). Human constitutive androstane receptor mediates induction of CYP2b6 gene expression by phenytoin. *J. Biol. Chem.*, 279:29295–29301.

Wang, R., Lai, L., and Wang, S. (2002). Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput. Aided Mol. Des.*, 16:11–26.

Wang, R., Lu, Y., Fang, X., and Wang, S. (2004b). An extensive test of 14 scoring functions using the pdbbind refined set of 800 protein-ligand complexes. *J. Chem. Inf. Comput. Sci.*, 44:2114–2125.

Wang, R., Lu, Y., and Wang, S. (2003a). Comparative evaluation of 11 scoring functions for molecular docking. *J. Med. Chem.*, 46:2287–2303.

Wang, Z., Benoit, G., Liu, J., Prasad, S., Aarnisalo, P., Liu, X., Xu, H., Walker, N. P. C., and Perlmann, T. (2003b). Structure and function of nurr1 identifies a class of ligand-independent nuclear receptors. *Nature*, 423:555–560.

Wärnmark, A., Treuter, E., Wright, A. P. H., and Gustafsson, J. . (2003). Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation. *Mol. Endocrinol.*, 17:1901–1909.

Watkins, R. E., Davis-Searles, P. R., Lambert, M. H., and Redinbo, M. R. (2003a). Coactivator binding promotes the specific interaction between ligand and the pregnane x receptor. *J. Mol. Biol.*, 331:815–828.

Watkins, R. E., Maglich, J. M., Moore, L. B., Wisely, G. B., Noble, S. M., Davis-Searles, P. R., Lambert, M. H., Kliewer, S. A., and Redinbo, M. R. (2003b). 2.1 åcrystal structure of human PXR in complex with the st. john's wort compound hyperforin. *Biochemistry*, 42:1430–1438.

Watkins, R. E., Wisely, G. B., Moore, L. B., Collins, J. L., Lambert, M., Williams, S. P., Willson, T. M., Kliewer, S. A., and Redinbo, M. R. (2001). The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. *Science*, 292:2329–2333.

Waxman, D. J. and Azaroff, L. (1992). Phenobarbital induction of cytochrome p-450 gene expression. *Biochem. J.*, 281:577–592.

Weatherman, R. V., Fletterick, R. J., and Scanlan, T. S. (1999). Nuclear-receptor ligands and ligand-binding domains. *Annu. Rev. Biochem.*, 68:559–581.

Wei, P., Zhang, J., Dowhan, D. H., Han, Y., and Moore, D. D. (2002). Specific and overlapping functions of the nuclear hormone receptors CAR and PXR in xenobiotic response. *Pharmacogenomics J.*, 2:117–126.

Wei, P., Zhang, J., Egan-Hafley, M., Liang, S., and Moore, D. D. (2000). The nuclear receptor CAR mediates specific xenobiotic induction of drug metabolism. *Nature*, 407:920–923.

Whysner, J., Ross, P. M., and Williams, G. M. (1996). Phenobarbital mechanistic data and risk assessment: enzyme induction, enhanced cell proliferation, and tumor promotion. *Pharmacol. Ther.*, 71:153–191.

Willett, P. (1995). Genetic algorithms in molecular recognition and design. *Trends Biotechnol.*, 13:516–521.

Williams, M. L., Wainer, I. W., Embree, L., Barnett, M., Granvil, C. L., and Ducharme, M. P. (1999). Enantioselective induction of cyclophosphamide metabolism by phenytoin. *Chirality*, 11:569–574.

Willson, T. M. and Kliewer, S. A. (2002). PXR, CAR and drug metabolism. *Nat. Rev.*, 1:259–266.

Windshügel, B., Jyrkkärinne, J., Poso, A., Honkakoski, P., and Sippl, W. (2005). Molecular dynamics simulations of the human CAR ligand-binding domain: deciphering the molecular basis for constitutive activity. *J. Mol. Model.*, 11:69–79.

Wolber, G. and Langer, T. (2005). Ligandscout: 3-d pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model.*, 45:160–169.

Wurtz, J. M., Bourguet, W., Renaud, J. P., Vivat, V., Chambon, P., Moras, D., and Gronemeyer, H. (1996). A canonical structure for the ligand-binding domain of nuclear receptors. *Nat. Struct. Biol.*, 3:87–94.

Xiao, L., Cui, X., Madison, V., White, R. E., and Cheng, K. C. (2002). Insights from a three-dimensional model into ligand binding to constitutive active receptor. *Drug Metab. Dispos.*, 30:951–956.

Xie, W., Radominska-Pandya, A., Shi, Y., Simon, C. M., Nelson, M. C., Ong, E. S., Waxman, D. J., and Evans, R. M. (2001). An essential role for nuclear receptors SXR/PXR in detoxification of cholestatic bile acids. *Proc. Natl. Acad. Sci. U S A*, 98:3375–3380.

Xie, W., Yeuh, M. F., Radominska-Pandya, A., Saini, S. P. S., Negishi, Y., Bottroff, B. S., Cabrera, G. Y., Tukey, R. H., and Evans, R. M. (2003). Control of steroid, heme, and carcinogen metabolism by nuclear pregnane x receptor and constitutive androstane receptor. *Proc. Natl. Acad. Sci. U S A*, 100:4150–4155.

Xiong, H., Yoshinari, K., Brouwer, K. L. R., and Negishi, M. (2002). Role of constitutive androstane receptor in the in vivo induction of mrp3 and CYP2B1/2 by phenobarbital. *Drug Metab. Dispos.*, 30:918–923.

Xu, H. E., Lambert, M. H., Montana, V. G., Plunket, K. D., Moore, L. B., Collins, J. L., Oplinger, J. A., Kliewer, S. A., Gampe Jr., R. T., McKee, D. D., Moore, J. T., and Willson, T. M. (1999). Structural determinants of ligand binding selectivity between the peroxisome proliferator-activated receptors. *Proc. Natl. Acad. Sci. U S A*, 98:13919–13924.

Xu, H. E., Stanley, T. B., Montana, V. G., Lambert, M. H., Shearer, B. G., Cobb, J. E., McKee, D. D., Galardi, C. M., Plunket, K. D., Nolte, R. T., Parks, D. J., Moore,

J. T., Kliewer, S. A., Willson, T. M., and Stimmel, J. B. (2002). Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPARα. *Nature*, 415:813–817.

Xu, R. X., Lambert, M. H., Wisely, B. B., Warren, E. N., Weinert, E. E., Waitt, G. M., Williams, J. D., Collins, J. L., Moore, L. B., Willson, T. M., and Moore, J. T. (2004). A structural basis for constitutive activity in the human CAR/RXRalpha heterodimer. *Mol. Cell*, 16:919–928.

Yamamoto, K., Masuno, H., Choi, M., Nakashima, K., Taga, T., Ooizumi, H., Umesono, K., Sicinska, W., VanHooke, J., DeLuca, H. F., and Yamada, S. (2000). Three-dimensional modeling of and ligand docking to vitamin d receptor ligand binding domain. *Proc. Natl. Acad. Sci. U S A*, 97:1467–1472.

Yamamoto, K. R. (1985). Steroid receptor regulated transcription of specific genes and gene networks. *Annu. Rev. Genet.*, 19:209–252.

Yamamoto, Y., Kawamoto, T., and Negishi, M. (2003). The role of the nuclear receptor CAR as a coordinate regulator of hepatic gene expression in defense against chemical toxicity. *Arch. Biochem. Biophys.*, 409:207–211.

Yamamoto, Y., Moore, R., Goldsworthy, T. L., Negishi, M., and Maronpot, R. R. (2004). The orphan nuclear receptor constitutive active/androstane receptor is essential for liver tumor promotion by phenobarbital in mice. *Cancer Res.*, 64:7197–7200.

Yamazaki, H., Shibata, A., Suzuki, M., Nakajima, M., Shimada, N., Guengerich, F. P., and Yokoi, T. (1999). Oxidation of troglitazone to a quinone-type metabolite catalyzed by cytochrome p-450 2c8 and p-450 3a4 in human liver microsomes. *Drug Metab. Disp.*, 27:1260–1266.

Yamazaki, Y., Kakizaki, S., Horiguchi, N., Takagi, H., Mori, M., and Negishi, M. (2005). Role of nuclear receptor CAR in carbon tetrachloride-induced hepatotoxicity. *World J. Gastroenterol.*, 11:5966–5972.

Yoon, J. C., Puigserver, P., Chen, G., Donovan, J., Wu, Z., Rhee, J., Adelmant, G., Stafford, J., Kahn, C. R., Granner, D. K., Newgard, C. B., and Spiegelman, B. (2001). Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature*, 413:131–138.

Yoshinari, K., Kobayashi, K., Moore, R., Kawamoto, T., and Negishi, M. (2003). Identification of the nuclear receptor CAR:HSP90 complex in mouse liver and recruitment of protein phosphatase 2a in response to phenobarbital. *FEBS Lett.*, 548:17–20.

Zavodszky, M. I., Sanschagrin, P. C., Korde, R. S., and Kuhn, L. A. (2002). Distilling the essential features of a protein surface for improving protein-ligand docking, scoring, and virtual screening. *J. Comput. Atd. Mol. Des.*, 16:883–902.

Zelko, I., Sueyoshi, T., Kawamoto, T., Moore, R., and Negishi, M. (2001). The peptide near the c terminus regulates receptor CAR nuclear translocation induced by xenochemicals in mouse liver. *Mol. Cell. Biol.*, 21:2838–2846.

Zenke, M., Munoz, A., Sap, J., Vennström, B., and Beug, H. (1990). v-erba onco-gene activation entails the loss of hormone-dependent regulator activity of c-erba. *Cell*, 61:1035–1049.

Zhang, J., Huang, W., Chua, S. S., Wei, P., and Moore, D. D. (2002). Modulation of acetaminophen-induced hepatotoxicity by the xenobiotic receptor car. *Science*, 298:422–424.

Zhou, J. and Cidlowski, J. A. (2005). The human glucocorticoid receptor: one gene, multiple proteins and diverse responses. *Steroids*, 70:407–417.

Appendix A

Abbreviations and Units

A.1   Abbreviations

AF-1/-2	activation function-1/-2
cAMP	3′-5′-cyclic adenosine monophosphate
ANDR	3α-androstenol
AR	androgen receptor
bp	base pair
CAR	constitutive androstane receptor
CCRP	cytoplasmic CAR retention protein
CITCO	6-(4-chlorophenyl)imidazo[2,1-*b*][1,3][thiazole-5-carbaldehyde *O*-(3,4-dichlorobenzyl)oxime
CLOT	clotrimazole
CTE	C-terminal extension
CYP	Cytochrome P450
DBD	DNA binding domain
3D	three-dimensional
DR	direct repeat
EE2	17α-ethynyl-3,17β-estradiol
e.g.	exempli gratia
ER	everted repeat
et al.	et altera
Fig.	Figure
GR	glucocorticoid receptor

GRIP-1	glucocorticoid receptor interacting protein-1
GROMACS	groningen machine for chemical simulation
GROMOS	groningen molecular simulation
GST	Glutathione-S-transferase
HRE	hormone response element
i.e.	id est
LBD	ligand binding domain
LBP	ligand binding pocket
MD	molecular dynamics
NCoR	nuclear receptor corepressor
NR	nuclear receptor
NRID	nuclear receptor interaction domain
OATP2	organic anion transporter polypeptide 2
PAPS	3′-phosphoadenosine-5′-phosphosulfate
PBREM	phenobarbital responsive enhancer module
PDB	Protein Data Bank
PGC-1	PPARγ coactivator-1
PXR	pregnane X receptor
PPAR	peroxisome proliferator-activated receptor
PREG	5β-pregnane-3,20-dione
RAR	retinoid acid receptor
R&D	research and development
RMSD	root mean square deviation
ROR	retinoid acid-related orphan receptor
RXR	retinoid X receptor
SD	steepest descent
SMRT	silencing mediator of retinoid and thyroid receptors
SRC-1	steroid receptor coactivator-1
SULT	sulfotransferase
TCPOBOP	1,4-bis[2-(3,5-dichloro-pyridyloxy)]benzene
TIF2	transcriptional intermediary factor 2
TMPP	tri-(p-methylphenyl)phosphate
TPP	triphenyl phosphate
TR	thyroid hormone receptor
UGT	uridine 5′-diphosphate-glucuronosyltransferase

VDR	vitamin D receptor
vdW	van der Waals
VS	virtual screening
Y2H	yeast two-hybrid

A.2   Units

Å   Ångstrom (1 Ångstrom = 10^{-10} m)

Da   Dalton (1 Da = 1/12 of the mass of ^{12}C)

fs   femtosecond (1 fs = 10^{-15} s)

K   Kelvin (0K = -273.15 °C

kJ   kilojoule (1 kJ = 0.239 kcal)

nm   nanometer (1 nm = 10^{-9} m)

ns   Nanosecond (1 ns = 10^{-9} s)

ps   Picosecond (1 ps = 10^{-12} s)

Appendix B

Amino acids

Amino acid	three-letter code	single-letter code
Alanine	Ala	A
Arginine	Arg	R
Aspartic acid	Asp	D
Asparagine	Asn	N
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophane	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

List of Publications

I   Windshügel B., Jyrkkärinne J., Vanamo J., Poso A., Honkakoski P. and Sippl W.
Comparison of homology models and X-ray structures of the nuclear receptor CAR: Assessing the structural basis of constitutive activity.
J. Mol. Graph. Model., available online

II   Jyrkkärinne J., Windshügel B., Mäkinen J., Ylisirniö M., Peräkylä, M., Poso A., Sippl W. and Honkakoski P.
Amino acids important for ligand specificity of the human constitutive androstane receptor.
J. Biol. Chem. **280**, 5960-5971 (2005)

III Windshügel B., Jyrkkärinne J., Poso A., Honkakoski P. and Sippl W.
Molecular dynamics simulations of the human CAR ligand-binding domain: deciphering the molecular basis for constitutive activity.
J. Mol Model. **11**, 69-79 (2005)

IV   Windshügel B. and Sippl W.
In silico Proteinsimulation: Fact or Fantasy?
Bioforum **7-8**, 54-55 (2005)

V   Fischer S., Windshügel B., Horak D., Holmes K.C. and Smith J.C.
Structural mechanism of the rcovery stroke in the myosin molecular motor.
Proc. Natl. Acad. Sci. U S A **102**, 6873-6878

VI   Windshügel B.
Conformational transitions in myosin.
Diploma thesis, Ruprecht-Karls-University Heidelberg

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertationsschrift selbstständig und ohne fremde Hilfe angefertigt, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die aus ihnen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.
Die Arbeit wurde ausschliesslich der Mathematisch-Naturwissenschaftlichen Fakultät der Martin-Luther-Universität Halle-Wittenberg vorgelegt und an keiner anderen Universität oder Hochschule weder im In- und Ausland zur Erlangung des Doktorgrades eingereicht.

Halle (Saale), den 1. Juni 2006

Curriculum Vitae

Personal Information

Name	Björn Anselm Windshügel
Date of birth	January 1st 1976
Place of birth	Bietigheim, Germany

Education

Schools

1986-1990	Hariolf-Gymnasium, Ellwangen
1990-1993	Eugen-Boltz-Realschule, Ellwangen
1993-1996	Wirtschaftsgymnasium, Aalen

Universities

1996-1998	Primary studies in Biology University of Bayreuth
1998-2002	Main studies in Biology Ruprecht-Karls-University Heidelberg
2001-2002	Diploma thesis in the group of Prof. Dr. J.C. Smith Title: "Conformational Transitions in Myosin" Interdisciplinary Center for Scientific Computing Ruprecht-Karls-University Heidelberg
2002-2003	PhD student of Dr. W. Sippl in the group of Prof. Dr. H.-D. Höltje Institute for Pharmaceutical Chemistry Heinrich-Heine-University Düsseldorf
2003-2006	PhD student of Prof. Dr. W. Sippl Institute for Pharmaceutical Chemistry Martin-Luther-University Halle-Wittenberg
since 2006	Senior researcher Department of Pharmaceutical Chemistry University of Kuopio, Finland