From DEPARTMENT OF BIOSCIENCES AND NUTRITION
Karolinska Institutet, Stockholm, Sweden

MOLECULAR DYNAMIC STUDIES OF NUCLEAR RECEPTORS LIGAND BINDING DOMAIN

Sofia Burendahl

Karolinska Institutet

Stockholm 2009

Published by Karolinska Institutet. Printed by E-print Oxtorgsgatan 9-11, 111 57 Stockholm.

ISBN 978-91-7409-459-6

To my family

ABSTRACT

Nuclear Receptors (NR) function as transcription factors that regulate genes that affect processes like reproduction, development and metabolism. The NRs are activated upon a given signal, which can be a ligand binding or a chemical modification. When activated, the receptors perform a conformational change that opens up interaction surfaces on the ligand binding domain, where coactivators can bind and transcription of the target gene can start. In this thesis the key events for NR activation, the ligand binding/unbinding mechanism and interactions with cofactors, are studied. Furthermore the communication between the ligand binding pocket and the cofactor interaction surface was investigated.

To gain insight on the ligand unbinding mechanism we performed molecular dynamics (MD) studies on two NRs. The unbinding mechanism from Retinoic Acid Receptor (RAR) and Estrogen Receptor (ER) was studied with modified MD methods, random acceleration MD and steered MD. In the RAR study 4 unbinding pathways of the ligand retinoic acid were obtained, where one of the pathways were more likely than the others. Thus ligand binding could be obtained without major conformational changes on the receptor structure. In the ER study, three different ligands unbinding from ER α and β was studied. The results showed that an ER agonist or selective agonist could unbind from the receptor without causing major conformational changes, while a slightly more bulky antagonist could not. Thus NR agonist and antagonist would use different unbinding mechanisms. The results from the ER simulations also showed variance in pathway preference between the different ligands. Differences between the ligands and receptor subtypes might therefore also effect the unbinding and hence influence ligand selectivity.

When the NR is activated an interaction surface becomes available and cofactors with a conserved motif can bind. In MD studies of Liver Receptor Homologue 1 (LRH-1) and Liver X Receptor (LXR), the interaction between different cofactor peptides and the receptors were characterized. In the LRH-1 study, a specific interaction from an aspartate to the motif was identified while the interactions between LXR and cofactor peptides showed a less specific binding. Thus specificity between LXR-cofactors should be found in other factors.

Cofactor interactions were also studied in the context of ligand binding. For LRH-1, a bound ligand to the receptor caused different effects on the receptor- cofactor peptides interaction. This indicates a communication pathway between the ligand and the cofactor peptide, an allosteric communication. Allosteric signaling is difficult to study, to do so we used a modified MD technique, anisotropic thermal diffusion method. With this method we were able to identify an allosteric signaling pathway from a coactivator peptide through LXR to the ligand in the ligand binding pocket.

LIST OF PUBLICATIONS

I. Carlsson, P., **Burendahl, S**. and Nilsson, L., *Exploring the Unbinding of Retinoic Acid from the Retinoic Acid Receptor using Random Expulsion Molecular Dynamics*, 2006, Biophys. J., 91: 3151-3161

II. **Burendahl, S**., Danciulescu, C. and Nilsson, L. *Ligand unbinding from the estrogen receptor: A computational study of pathways and ligand specificity*, 2009, submitted manuscript

III. **Burendahl, S**., Treuter, E., and Nilsson, L. *Molecular Dynamics simulation of human LRH-1: the impact of ligand binding in a constitutively active nuclear receptor*, 2008, Biochemistry, 47(18):5205-15

IV. **Burendahl, S**., and Nilsson, L. *Computational studies of molecular interactions and communication beteween LXR cofactors and ligand binding*, 2009, manuscript

PAPER NOT INCLUDED IN THIS THESIS

Monné, M., Han, L., Schwend, T., **Burendahl, S**., and Jovine, L. *Crystal structure of the ZP-N domain of ZP3 reveals the core fold of animal egg coats*, Nature, 2008, 456: 653-657

CONTENTS

LIST OF ABBREVIATIONS

ATD	Anisotropic Thermal Diffusion
AF-1	Activation Function 1
AF-2	Activation Function 2
CHARMM	Chemistry at HARvard Macromolecular Mechanics
DBD	DNA Binding Domain
DNA	Dioxyribonuclei Acid
ER	Estrogen Receptor
GR	Glucocorticoid Receptor
H1, H2, etc.	Helix 1, Helix 2, etc
HAT	Histone Acetyltransferase
HDAC	Histone Deacetyltransferase Activity
HMT	Histone argenine metyltransferase
HRE	Hormone Response Element
LBD	Ligand Binding Domain
LBP	Ligand Binding Pocket
LXR	Liver X Receptor
LRH-1	Liver Receptor Homologue 1
MD	Molecular Dynamics
MM	Molecular Mechanics
NR	Nuclear Receptor
PPAR	Peroxisome Proliferator Activated Receptor
QM	Quantum Mechanics
RAMD	Random Acceleration Molecular Dynamics
RAR	Retinoic Acid Receptor
RNA	Ribonuclei Acid
RXR	Retinoic X Receptor
SMD	Steered Molecular Dynamics
SNRM	Selective Nuclear Receptor Modulator
S1/S2	β-Sheet1 / β-Sheet2
vdW	van der Waals

INTRODUCTION

Nuclear receptors (NR) are a protein family that regulates gene transcription. In the cells the NRs play an important role as they are involved in the cell development, differentiation, homeostasis and metabolism (reviewed in (1-3)). Due to their central role in many cellular events the NRs have been identified as a drug target for cancer, diabetes, hormone resistance syndromes and cardiovascular diseases. The growing number of the population affected by problems controlled by NR have placed the NRs as the second largest gene family for FDA (U.S. Food and Drug Administration) approved drugs and the NR glucocorticoid receptor (GR) as one of the targets with the most number of approved drugs (reviewed in (4)).

The NRs action is controlled by a signal, often small molecules called ligands. The ligand binds the receptor´s ligand binding domain (LBD) which induces a number of events and results in a change of the target protein production. Since drugs often are small organic molecules, they can bind the NRs and resemble their natural ligand. In such ways, a synthetic molecule can induce a biological response in the cell.

The mechanism for NRs action in the cell has been studied with various experimental techniques. Molecular biology experiments typically relate a change in the environment or in one of the components tested, to the amount of target protein produced. Doing so, it is possible to determine how the tested variables affect the biological activity of the receptor. Thus molecular biology studies of the NRs have highly contributed to the understanding of what factors influence the mechanism of the receptors. To get a more detailed explanation of the action of NRs, one can use structural biology experiments. The structural biologist can determine the 3D structure of a protein and with help from the molecular biologist, point out more exactly which part of the protein that is responsible for the activity. Protein 3D structures are like photos of the protein in its stable conformations. This can be compared to taking very few snapshots of a person during one day. The most likely motif would be the person sleeping or working, however we all know there is a lot of action going on between those two points. Similarly structural studies of a protein are not fast enough to catch all the events, but often just show a fraction of the actual action that the proteins undertake. To get more information on the continuous movement of the proteins, one has to use theoretical models and computer calculations. Computer simulations of the molecules movements, also called dynamics, are a useful tool for the understanding of the function of proteins. The first molecular dynamics (MD) simulation of a large molecule was performed in 1977 (5) and since then the technique has significantly contributed to the understanding of protein motions.

In this thesis MD simulations have been used to understand how the NRs LBD performs it actions. The first two papers concern the ligand binding mechanism of the Retinoic Acid Receptor (RAR) and Estrogen Receptors (ERs). Paper III characterizes how cofactor peptides bind to the Liver Receptor Homologue 1 (LRH-1) receptor and how this binding is influenced by the presence of a ligand. The last paper IV, studies the cofactor binding interactions, the ligand binding and the communication that occurs between these two sites in Liver X Receptor (LXR).

NUCLEAR RECEPTOR´S STRUCTURE

In the human genome 48 NRs have been indentified some of which are very similar resulting in 28 different ligand binding sites[1] (reviewed in (6)). Most of these NRs are activated by ligand binding but for some receptors no ligands are yet identified and these are called orphan receptors. The NR activation might in such case take place by other events like chemical modification of the receptor. The ability of ligand dependent activation has been suggested to been evolved during evolution when increased specificity was required although the opposite mechanism cannot be out ruled (3,7). No NRs have not been identified in fungi, plants or unicellular eukaryotes suggesting that NRs appeared during animal (metazoan) evolution (8). Evolutionary data indicate a two step gene duplication where the first wave led to the present NR subfamilies and the second wave occurred in vertebrates where prologue groups were formed within the subfamilies. Thus most of the NRs subfamilies seem to be ancient since they have homologues in lower metazoan species. The number of NRs identified in metazoan species varies greatly from 21 in the fly *Drosophila melanogaster* to 270 in the worm *Caenorhabditis elegans* (9).

NUCLEAR RECEPTOR SUPERFAMILY

All of the known NRs belong to the NR superfamily which comprises the steroid receptor family and nonsteroid receptor family. In more details the NR families have been divided into 6 subfamilies based on evolutionary analysis (10). Receptor subtypes are products from individual genes while receptor isoforms are produced from alternative splicing and/or promoter usage. Subfamily 1 is formed by thyroid hormone receptors (TRs), retinoic acid receptors (RARs), vitamin D receptor (VDR) and peroxisome proliferator-activated receptors (PPARs) as well as some orphan receptors. The second subfamily contains the retinoic X receptor (RXR) among others, which plays an important role as heterodimer partner for other receptors. The third subfamily is formed by the steroid receptors GR, androgen receptor (AR), progesterone receptor (PR), ERs and the highly related orphan estrogen-related receptors (ERRs). Subfamily 4, 5 and 6 contains orphan receptors like steroidogenic factor 1 (SF-1) and LRH-1. The additional subfamily 0 includes the small heterodimeric partner (SHP) and dosage-sensitive sex reversal (DAX-1) which are not full length receptors.

DOMAIN STRUCTURE

The NRs share a common domain organization named from A-F (Figure 1) (reviewed in (11)). The N-terminal A/B region often varies in length and sequence between the family members. The D region is a linker between the evolutionary conserved DNA Binding Domain (DBD) (C region) and the LBD (E region). The C-terminal F region is not present in all NRs and its function is unknown. The NRs contain regions responsible for the transcriptional activity. The first one found in the variable A/B region, is constitutively active and referred to as the activation function 1 (AF-1). The second activation function 2 (AF-2) is located on the LBD and has a ligand dependent activation. Both the DBD and the LBD have surfaces responsible for interactions with

[1] Assuming that subtypes have the same natural ligand.

another NR, dimerization surfaces, and nuclear localization signal. Most of the NRs display this domain organization except some which only contain regions similar to the DBD or LBD.

Figure 1. Schematic view of the NR domain organization.

The DBD is the most conserved region of the NRs and it mediates DNA binding. The structure of the DBD has been solved for several NRs and is conserved within the family (Figure 2). The domain consists of two "zinc fingers", held together by 4 cysteines arranged around a zinc ion. The DBD spans about 60-70 amino acids, where the ones responsible for binding to the core DNA recognition motif are present at the first zinc finger. The recognition helix of the first zinc finger binds to the DNA major groove often in dimer complex with another DBD. The amino acids responsible for dimerization are found at the second finger.

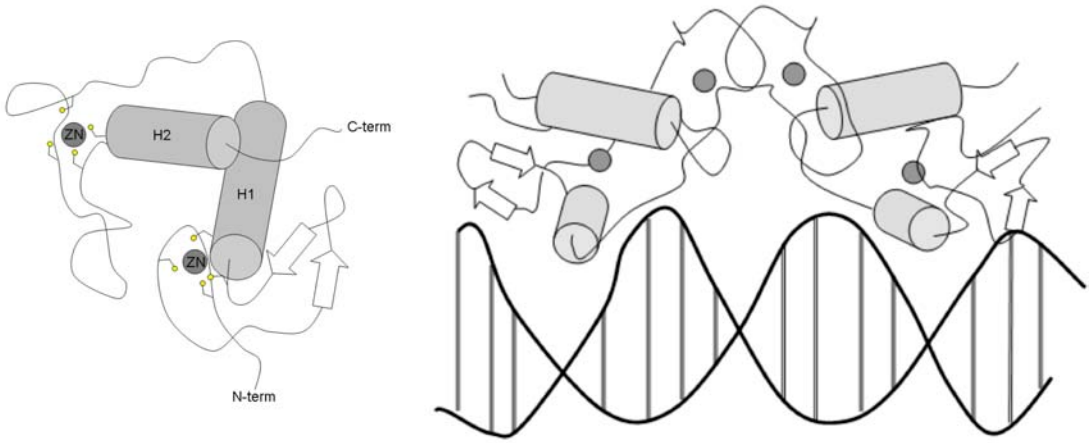

Figure 2. NR DBD structure consists of two zinc fingers each one coordinated around a zinc ion by 4 cysteins (lines with spheres) (left picture). The DBD often binds to the DNA as a dimer, interacting with the first finger to the DNA major groove (right picture). The pictures are based on the structure of ER (pdb id 1HCQ).

The structure of the LBD has also been determined for several members of the NR family. A canonical domain structure has been elucidated from the results, where the receptors are arranged in a three-layered anti-parallel helical sandwich (Figure 3). The12 α-helices (H1 to H12) are interrupted by a β-sheet (S1 and S2) between H5 and H6.

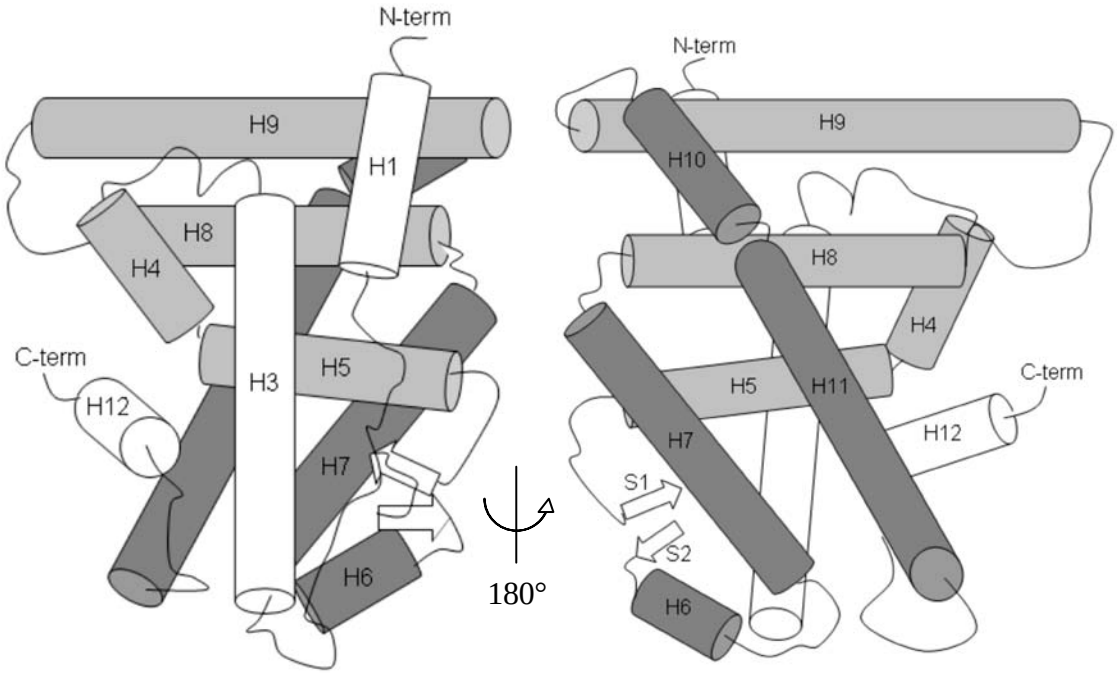

Figure 3. Schematic picture of the NR LBD canonical fold. The three layers in the anti-parallel helical sandwich are indicated with white, gray and dark gray color. The pictures are based on the structure of ERα (pdb id 1GWR).

Some of the NRs show a small deviation from this canonical structure, such as PPARγ which contains an extra helix H2'. The LBD is a multi functional domain including surfaces for protein interaction and a ligand binding pocket (LBP). The AF-2 interaction surface responsible for cofactor interactions is located at H3, H4 and H12 while the surface for interaction with another NR partner, the dimerization surface, is found at H10/11, some of H9 and the loop between H7-H8. The enclosed LBP is mostly hydrophobic and located beneath H5 in close contact with H3, H7, H11, H12 and the β-sheets. The pocket size and shape varies among the members of the NR family spanning from the large PPARγ pocket to the collapsed pocket of the orphan receptor Nuclear Receptor related 1(NURR1). It is particularly the LBP side towards the β-sheets and the presence or absence of the H2 that contributes to the variability in volume. Some of the NRs also have an additional solvent cavity, close to the LBP, which can during certain circumstances be exploited by ligands. Binding of ligand to a large LBP often results in low affinity interactions while binding to small LBPs more often cause very specific interactions between the NR and ligand. Several studies indicate LBD stabilization upon ligand binding and induction of a more compact domain structure (12,13). Structures of the ligand free LBP (apo) has also revealed some structural variations between receptors, indicating equilibrium of receptor conformations (14).

The first DBD structure was solved in 1991 (15) and the LBD structure in 1995 (16,17). However, the receptor structure of the DBD-LBD complex remained elusive until 2008, when the structure of the PPAR-RXR was solved (18). This structure indicates a central role of the PPAR LBD, contacting all the ordered domains of both proteins. Moreover the PPAR domain directs the DBD of both receptors towards the DNA. Also surprising is that all three ligands tested induced the same domain rearrangements within the molecule.

NUCLEAR RECEPTOR´S FUNCTION

MECHANISM OF ACTION

NRs can be activated by different mechanisms and the most common mechanism is the binding of a ligand (Figure 4). The ligand can be generated in three different ways; 1) the active ligand or hormone can be synthesized in a classical endocrine organ and enter the cell. 2) Alternatively the ligand might be synthesized within the target cell from a precursor or prehormone, or 3) the ligand may be a metabolite synthesized within the target cell. The NRs are often located in the nucleus even in their apo state, but some receptors like the steroid receptors are located in the cytoplasm in the absence of ligand. Ligand binding to a cytoplasmatic receptor releases the binding of heat chock proteins and translocates the receptor to the nucleus. The translocation is done by active transport trough the membrane and guided by a tag on the receptor, the nuclear localization signal. NRs located in the nucleus bind the ligand and release bound corepressors. When activated the receptor binds the DNA as a monomer or it forms a dimer complex and binds the DNA. A specific DNA sequence, the hormone response element (HRE), is recognized by the NR which binds to this sequence. The HREs are often built up by two half sites, separated by a various length of DNA sequence, and located close to the regulatory regions of the target gene. In the active conformation the NR interacts with coactivators and other transcription factors which make the DNA accessible for RNA polymerase II which starts to decode the sequence of the target gene.

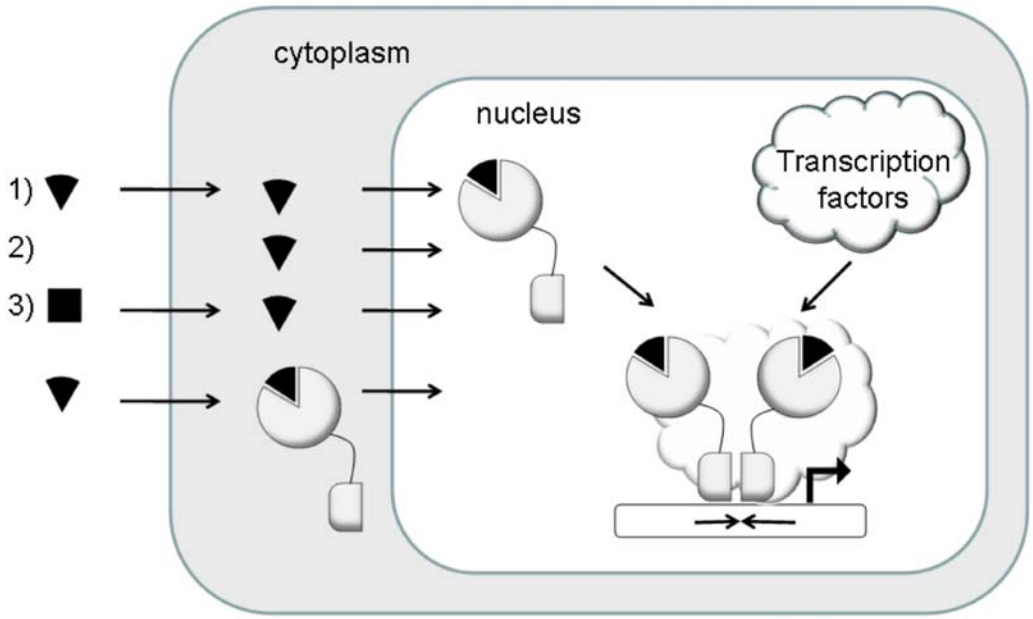

Figure 4. Ligand mediated activation mechanism of NRs. The ligand can be produced in three ways. In the cell the ligand binds the NR either in the cytoplasm or in the nucleus. Upon ligand binding the NR adopts an active conformation and can bind DNA. Together with various transcription factors the NR activates gene transcription.

MECHANISTIC CLASSES

NRs can be grouped into classes depending on the mechanism of action and subcellular distribution in the absence of a bound ligand (Figure 5) (1,19). The Class I receptors, including the steroid receptors in NR subfamily 3, bind to the ligand in the cytoplasm of the cell and then translocate to the cell nucleus. In the nucleus the homodimer bind to its HRE. The Class I receptors bind to a HRE consisting of two half sites where the second half site is an invert repeat of the first. The Class II receptors, mainly NR subfamily 1, are always located in the nucleus and form a heterodimer upon ligand binding. The heterodimer partner is usually RXR, which has been shown to interact with very different receptors. The Class II receptors show a variable DNA interaction mode where direct repeats and inverted HRE repeats can be accessed. The Class III receptors, principally NR subfamily 2, are similar to Class I except in their DNA binding mode. The Class III also forms homodimers but bind direct HRE repeats on the DNA. The Class IV receptors bind DNA as monomers or dimers, but only interact with a single DNA HRE. Most of the orphan receptors belong to this class.

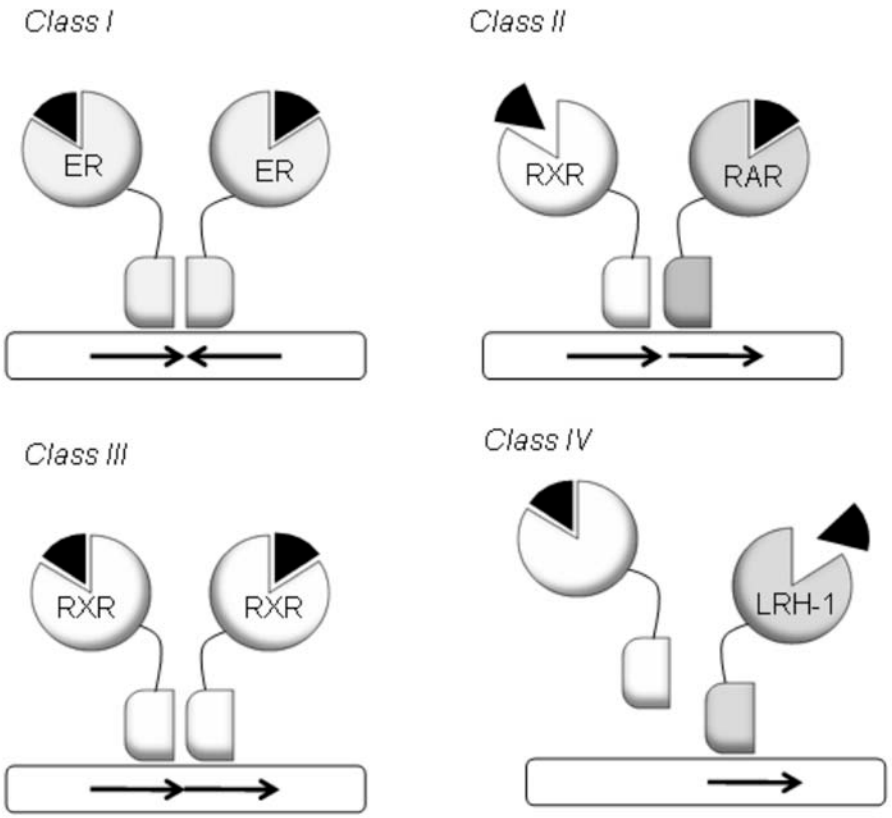

Figure 5. NRs mechanistic classes overview showing DNA binding, dimerisation and ligand binding properties. Representative examples for each group are shown.

LIGAND BINDING AND OTHER SIGNALS

NRs can bind chemically and structurally diverse molecules. The ligands are entitled after what transcriptional effect they induce. Binding of an agonist ligand induces the active conformation of the receptor and active gene transcription. Ligands that repress gene transcription are called antagonists and ligands which show different transcription modification depending on cell type and environmental factors are called selective NR modulators (SNRM) (reviewed in (20)). Binding of an antagonist causes a conformational change where the AF-2 surface is usually unavailable, due to different H12 conformation (Figure 6), and recruitment of coactivators is halted. The binding of a SNRM induces different conformational changes depending on the cellular context and thus different transcriptional activity will be achieved.

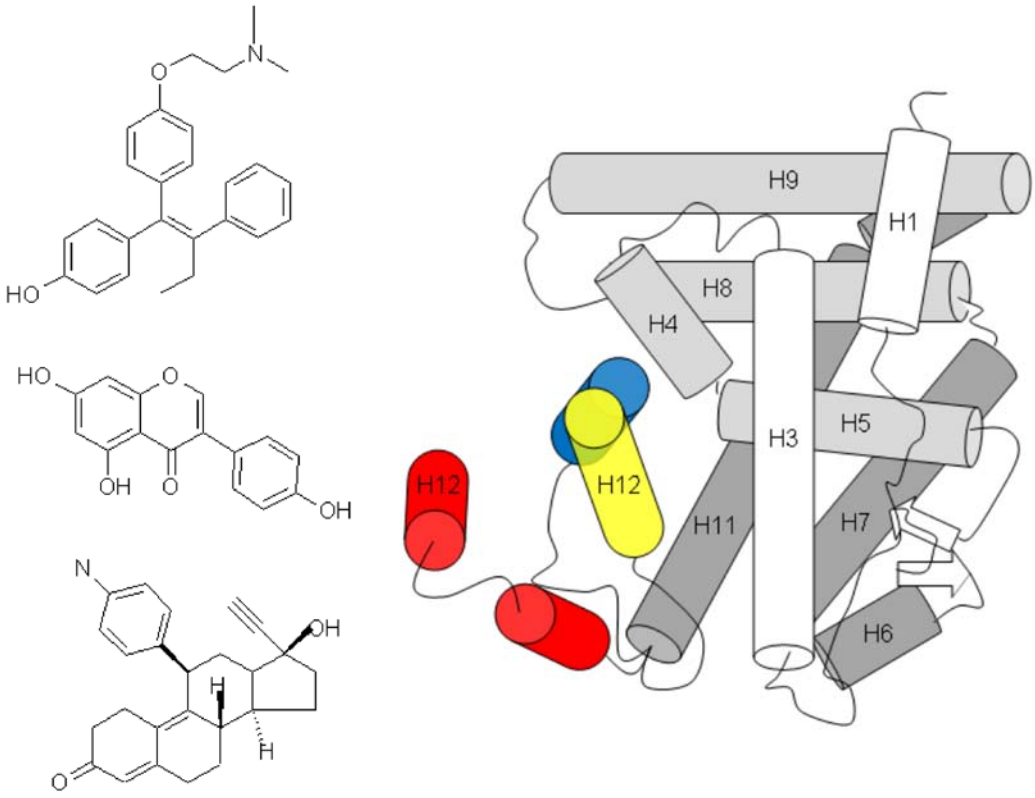

Figure 6. Conformation of NR LBDs bound to antagonist or SNRM. The ligands induce different H12 positions, making the AF-2 region unavailable for coactivator binding. The schematic picture were drawn after the structure of ERα bound to the antagonist 4-hydroxytamoxifen (blue H12) (pdb id 3ERT), ERβ bound to the SNRM genistein (yellow H12) (pdb id 1QKM) and GR bound to the antagonist mifepristone (red H12) (pdb id 1NHZ). Most of the LBD structure is preserved (H1-H11) without larger conformational changes, while the H12 positions deviates depending on ligand bound. Ligand structure are shown to the left starting with the 4-hydroxytamoxifen on top, followed by genistein and on the bottom mifepristone.

Activation of NRs is often mediated by ligand binding but sometimes by other mechanisms. Ligand-independent activation mechanisms are used by the true orphan receptors, for which no natural ligand exists but also as an alternative mechanism for NRs like ER (21-23). Ligand independent NR activation is often a consequence of

cross-talk with other signal-transduction pathways. Such cross-talk has been observed to act by multiple mechanisms including a high level of complexity.

Activation through a chemical modification has been observed for both ligand binding and orphan receptors. Common posttranscriptional modifications are receptor phosphorylation, methylation and ubiquitylation which alter the receptor function (reviewed in (22,23)). The modification can occur on the DBD or at the LBD but the majority of modified residues are found in the A/B region. For most NRs phosphorylation at the A/B region by Mitogen-activated protein kinase (MAPK) or Akt helps the recruitment of coactivators or of components of the transcription machinery and thus it increase the response to the ligand. But phosphorylation can also terminate the ligand response due to its close linkage to the phophorylation dependent ubiquitylation which target the protein for degradation by the 26S proteasome. Thus posttranscriptional modifications show a fine tuning between NR activation and degradation. Recent results have suggested that the phosphorylation/ubiquitylation signal for protein suicide might be incorporated in the mechanism of NR gene transcription. Experimental studies show that the steroid receptors with coactivators cycle on and off the promoter on the DNA many times indicating a brief DNA interaction (reviewed in (24)). Such cyclic interaction between the transcription complex and the DNA has been suggested to be directed by changes in the chromatin structure and acetylation/ methylation/ phosphorylation of the NRs, their cofactors and histones. Further on, phosphorylation or ubiquitylation of the NR cofactors can work as an efficient mechanism to clear out cofactors, permitting new ones to bind and allowing transcriptional elongation with a higher speed.

Another type of cross-talk occurs between the NRs and other transcription factors. This can be exemplified by the mutual gene repression between GR and either AP-1 or the nuclear factor κB (reviewed in (25,26)). In this case both of the transcriptional activators can repress the transcription of the other one, without binding to the DNA or change the other transcriptional activators DNA binding properties. The mechanism of such event remains elusive.

A third type of cross-talk is the so called ‘non-genomic’ action observed for several NRs (reviewed in (27,28)). This was observed for ERs since the effect of some estrogens are so rapid that they cannot depend on the activation of RNA and protein synthesis. The actions are believed to be mediated through membrane associated ERs and connected to protein-kinase cascades. 17-estradiol has been shown to activate the MAPK and phophoinositol 3– kinase signaling pathway in various cell types. Membrane bound ERs also associate with a variety of signal molecules like G proteins and can activate membrane tyrosine kinase receptors. However, the non-genomic action of estrogens are cell specific, thus the response to estrogens depend on the signal transduction molecules and the targets in the target cell.

COFACTORS: COACTIVATORS AND COREPRESSORS

NRs modify transcription activity by recruitment of positive or negative regulatory proteins, coactivators or corepressors. Some of these cofactors are used only by the NRs but many of them are used in similar manners by numerous other transcription factors. Coactivation or corepression are not an intrinsic feature of a given cofactor but can depend on the target gene and the cell-type specific context (reviewed in (29)). As an example a coregulator can function either as a coactivator or corepressor depending on the cellular environment.

Full activation of NRs is achieved by synergism between the two activation function regions AF-1 and AF-2 (30). The AF-1 region varies in sequence and length from 23 amino acids in VDR up to 602 amino acids in mineralocorticoid receptor. Structural studies with nuclear magnetic resonance and circular dichroism spectroscopy reveal a disordered AF-1 domain. Secondary structure formation has been suggested as an important step in the AF-1 activation mechanism similar to what has been observed for other transcription factors (31-33). In contrast to the disordered AF-1 region, the AF-2 region in the LBD is highly structured. The AF-2 surface is built up by H3, H4 and the correct position of H12. The position of H12 is crucial and small deviations might shift a preference for coactivator binding to corepressor binding, modifying the transcriptional activity. The active conformation of H12, exposing the hydrophobic AF-2 region allows interaction with a LXXLL motif (where L is a leucine and X any amino acid) present in most coactivators (34) (Figure 7).

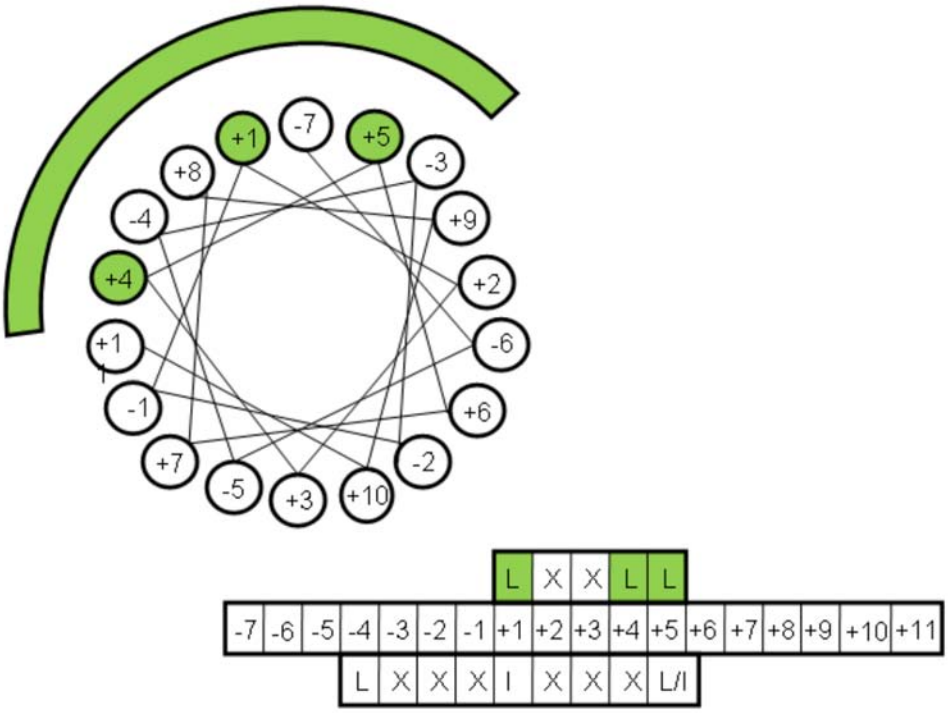

Figure 7. Schematic picture of the NR cofactor interactions. The conserved LXXLL motif is folded with the hydrophobic leucines (green spheres in the helical wheel) towards the protein surface (green half circle). The numbering in cofactors starts with +1 at the first leucine as indicated in the table. Below the numbering row the corepressor motif is indicated.

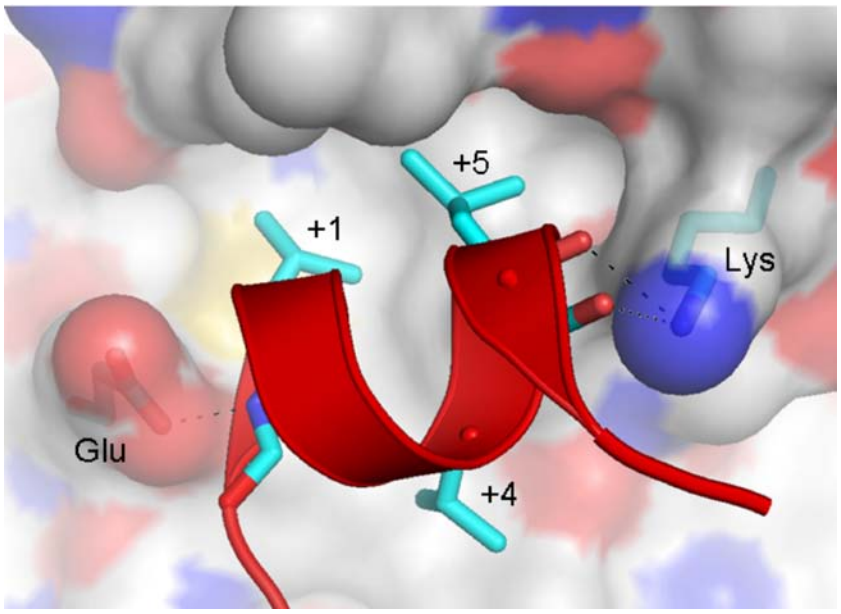

Figure 8. Interaction of ER AF-2 region and coactivator with the conserved LXXLL motif (pdb id 2J7X). The conserved leucines in the LXXLL motif are shown in sticks folded towards the protein. The coactivator backbone makes hydrogen bond interactions to a glutamic acid and lysine of the receptor, the charge clamp.

The short LXXLL motif forms a helical structure, folding the leucines towards the receptor hydrophobic surface (Figure 7, helical wheel). Structural studies of NRs and coactivator peptides have revealed electrostatic interaction between a receptor glutamic acid and lysine sidechains and the coactivator backbone, so called charge clamp (35) responsible for the coactivator orientation during binding (Figure 8). Additionally specific interactions between combinations of NR and coactivators can be found within or adjacent to the LXXLL motif (36). Until recently only structures of a cofactor peptides was available, but in 2008 the structure of the first full length coactivator DAX-1 and LRH-1 was solved (37), indicating that also interaction further away from the core motif might contribute to the strength of the interaction and specificity in NR-cofactor binding.

The NRs coactivators can be divided into 5 major groups (reviewed in (23)). The first group is members of the protein subfamily with molecular weight 160kDa (p160) which mainly serve as adapters recruiting other complexes. The second group is the histone acetyltransferases (HAT), the third group histone argenine metyltransferases (HMT), the fourth group nucleosome remodeling complexes and the last fifth group is mediator complexes which bridge NRs and the basal transcription machinery. The order and sequential recruitment of these coactivators have recently been determined and defined as a ‘transcriptional clock’ (38). Studies of ER identify ER and the p160 nuclear receptor coactivator (NCoA) family to be the first ones recruited to the DNA. Secondly HATs and HMTs are recruited to remodel the chromatin and make the DNA more accessible. At the last step the mediators are bound, facilitating the arrival of RNA Polymerase II and general transcription factors. However the order of coactivator recruitment might differ from one target gene to another indicating that one NR can use multiple programs for gene activation depending on the promotor context.

Unliganded NRs in the nucleus are often bound to a corepressor like Nuclear Receptor Corepressor (NCoR) and the silencing mediator of retinoid and TR (SMRT) (reviewed in (39)). The corepressor binds the LBD with a conserved LXXXIXXXL/I motif, similar to the LXXLL motif but extended at the N-terminal end (Figure 7). Structural studies show that the motif adopts an amphipathic α-helical conformation. In the apo conformation of the NRs H12 is out of the active conformation, leaving a larger area for the corepressor bind to. Thus the corepressor binds to a site similar as the coactivators, but lacking the H12 interaction. Mutations of the receptor´s hydrophobic core abolish interaction with corepressors, and amino acids adjacent to the core have also been shown to be important for corepressor binding (reviewed in (3)). Most corepressors also recruit a histone deacteyltransferase activity (HDAC) protein.
The binding of an antagonist or SNRM induces a LBD conformation with the AF-2 region unavailable (Figure 6). Full antagonist NR structure induces a LBD conformation with the H12 relocated from the AF-2 region, but SNRMs induced structures often put the H12 in a position close to the coactivator binding cleft. Thereby the binding of coactivators is prevented, but partly also the binding of the corepressors. The longer helical corepressor motif seems to be the key for corepressor interaction, where some binding might occur although the binding site is partly occupied. Full antagonism activation might thus only be achieved when the H12 is located out of the cleft.
Corepressor binding can also occur to ligand bound receptors and compete with the coactivators by displacing them (reviewed in (40)). In this atypical repression the receptor structure and the amino acids in both the NR and the corepressor are the key for interaction. These corepressors contain a LXXLL motif as the coactivators and can thus bind at the AF-2 region. Furthermore they act by recruiting HDACs as the classical repressors NCoR and SMRT. The first corepressor binding the agonist receptor form discovered was the receptor interacting protein 140.

ALLOSTERIC SIGNALING

The NRs contains structurally distinct sites which are functionally coupled. The cofactor binding sites AF-1 and AF-2, the LBP and the dimer interface, often act in a collaborative way indicating a communication between these sites (reviewed in (41)). The mechanism by which information is transmitted across long distances in proteins is often referred to as allosteric communication (reviewed in (42)). Such allosteric signal includes an enthalpic and an entropic component. The enthalpic component can be described as a change in structure where the proteins conformational ensemble changes and the entropic contribution constitute a change in the protein dynamics, upon the allosteric signal. Several methods have been used to gain insight of the nature and identification of allosteric signaling pathways. One of the methods; the statistic coupling analysis, uses statistic methods on protein family sequences to extract residue networks that have been preserved during evolution. Secondly MD studies including simplified protein models as in coarse grain methods and normal mode analysis have been utilized to evaluate correlated conformational fluctuations in the proteins. Experimental methods have also contributed to the understanding of allosteric communication, where nuclear magnetic resonance studies can enable residue to residue conformational fluctuations (43).

NUCLEAR RECEPTORS IN DRUG DISCOVERY

The cellular concentration of the intracellular receptors is extremely low; under normal physiological conditions rarely exceeding 0.01% of the total cellular proteins (44). Given that a single receptor can control more than100 genes, the effectiveness of the receptors action is indicated. Thus many of the receptors can be targeted for multiple disease treatment and their action can also be correlated with other NRs mechanisms. The complex schemes of NR action often affect several diseases and interact with other signal transduction pathways.

NUCLEAR RECEPTOR RELATED DISEASES

The intense research on NRs as targets for drug discovery every year results in an overwhelming amount of publications. In this chapter I mention the major disease groups and some of their associated target receptor to give a brief update on the research today. However, this is a very limited report of the field and many interesting and important NR related diseases are not covered.

Cancer

The treatment for breast cancer with the antiestrogen tamoxifen targeting the ER, has been extensively used. ERα is a very effective target for breast cancer treatment due to its expression in 70% of the breast cancers (reviewed in (45)). Today antiestrogen therapy is for example used before and after surgery. However, all patients do not respond to antiestrogen treatment, providing a need for an alternative signaling pathway to target. Studies of other NRs have revealed a potential importance of AR, GR, PPAR, RAR, RXR, VDR and EER, in breast cancer. The active NRs use both genomic and nongenomic NR pathways to activate growth factor signaling pathways that are central to breast cancer biology. The NR cofactors are also essential elements in regulation of NRs transcription. Recent results indicate their involvement in human diseases, but to what extent is not known. Interestingly, cofactors are broadly over- or underexpressed in human cancers (reviewed in (46)). For instance, in lung cancers 60% of the cofactors are overexpressed, 38% in breast cancer and 43% in prostate cancer. One can therefore probably assume that misexpression of cofactors is a common factor for cancer progression.

Metabolic diseases

Therapy for metabolic diseases including obesity, type II diabetes and metabolic syndrome, which includes several metabolic abnormalities, can be achieved by targeting NRs. The hyperlipidemic fibrates and antidiabetic thiazolidinenedions are drugs that act via PPARα and PPARγ respectively and are used in clinical practice (reviewed in (47,48)). Activating PPAR target genes that participate in fatty acid uptake through membranes, fatty acid binding in cells and fatty acid oxidation. The LXRs also show a key role in lipid metabolism since it increases the cholesterol efflux via activation of the ABCA1 (ATP-binding cassette transporter), CYP7A (Cholesterol 7 alpha-hydroxylase) and ApoE (Apolipoprotein E) genes (reviewed in (49,50)). However, activation of LXRs also increases triglyceride levels and formation of fatty

tissues (adipogenesis). Therefore identifying a SNRM activating LXRα but not LXRβ, could reduce the unwanted effects of LXR biology.

Inflammation

Several members of the NR superfamily have regulatory effects on the inflammatory process. Natural and synthetic glucocorticoid agonists are potent inhibitors of inflammation in various cell types. Although glucocorticoids are effective anit-inflammatory drugs, a prolonged use often results in side effects like hypertension, diabetes and muscular weakness. Furthermore some inflammatory diseases stop responding to treatment at some point during their development. Thus the development of new GR agonists show focus on reducing the unwanted side effects.
Activation of PPARγ and LXRs by metabolites of fatty acids and cholesterol, also target inflammatory response genes (reviewed in (51)). Since the target genes are distinct but overlapping, a combinatory mechanism with GR has been suggested. Indeed studies show that simultaneous treatment with GR and PPARγ or LXR agonists resulted in synergistic repression. Thus it is possible that anti-inflammatory actions of GR could be achieved at lower doses with less side effects by simultaneous administration of PPARγ or LXR agonist.

Cardiovascular diseases

Cardiovascular diseases are an increasing problem of the modern society, killing more Americans than cancer. The underlying cause is often an advanced state of atherosclerosis. Atherosclerosis is a chronic inflammatory response in the walls of arteries leading to hardening of the arteries, caused by build ups of lipids, plaques formation. PPARα and PPARγ regulate the plasma lipoprotein concentrations used for plaque formation and effect foam cell formation, an important step in the development of atherosclerosis (reviewed in (48)). Thus targeting the PPARs might modulate the atherosclerosis development. Synthetic LXR agonists administration also protects against atherosclerosis as observed in mice (reviewed in (52)). However, the increased plasma triglyceride levels observed upon LXR activation, results in a contradictory effect in the treatment of cardiovascular diseases.

Other diseases

A diverse span of other diseases related to the mechanism of NRs has been revealed and likely more are to come (reviewed in (53)). Most of the studies are focused on finding a ligand which binds the LBP of the LBD. In such ways, ligands for many orphan receptors have been discovered. However, some NRs are true orphan like the NURR1 receptor, which LBP is filled with hydrophobic amino acid side chains. Even so, a ligand that activates NURR1 has been found. Apparently the ligand activates the AF-1 region instead of binding to the LBP. NURR1 is mainly expressed in the central nervous system (CNS) and is indicated to be important for development, migration and survival of dopaminergic neurons. Since Parkinson´s disease results from the loss of dopaminergic neurons, NURR1 was proposed as a drug target for Parkinson´s disease.
Another orphan receptor expressed in the brain is the RAR-related orphan receptor (ROR) α, which in contrast to NURR1 holds a relative large LBP. The hormone melatonin, a drug used for treatment of sleep disorders, was reported as a natural ligand

for RORα. RORβ was also found to bind *all trans* retinoic acid, which inhibited RORβ activity. This indicate that retinoids can become tools for drug discovery for RORβ related CNS diseases.

DRUG DISCOVERY

Since the introduction of the term ´hormone´ about 100 years ago, an amazing progress has been made within life science increasing our knowledge of diseases and the mechanisms underlying them (54). Treatments for the diseases regulated by NRs have been identified and potent agonist and antagonist ligands discovered. Initially the drugs developed for the NRs aimed at the full agonist or antagonist action. Several such drugs like the anti-inflammatory drugs targeting GR or TR are potent but also induce several unwanted side-effects. Therefore today drug development targeting the NRs is mostly searching for molecules with partial or mixed agonist/antagonist function, the selective NR modulators (SNRM). These molecules usually bind the receptor with lower affinity inducing a conformational change in the receptor which allows only a certain number of cofactor to interact which results a limited biological response. Ideally a SNRM would activate the receptor in one cell-type to defeat the disease and inactivate the receptor in cell-types responsible for the unwanted side-effects. Important for the identification of SNRM, is the identification of cofactors that bind the NRs, their tissue distribution and expression profile during development (reviewed in (49)). Further on the information on how selective the SNRM promoter is will also be important for SNRM development.

Mapping of the human genome revealed no new NRs than the 48 already known. Therefore the drug development targeting NRs should first focus on the understanding of the mechanism of action and interaction network between receptors and other signaling pathways. Doing so development of new drugs will have a new platform to work from, making the costly process of drug development easier to manage.

The understanding of NRs actions have already contributed to explanations of diseases and effects observed in nature. One example is the effect that various chemicals that are accumulated in the environment have on endocrine systems. These chemicals can act as NR agonists or antagonists and interfere with the hormone synthesis and therefore disrupt endocrine networks in human and animal. This effect is called the ´endocrine disruption hypothesis´ and was presented after studies of male fertility in Scotland and Denmark (55,56). The studies indicated that the number and motility of sperms were decreased in men born after the World War II, hypothesizing that chemicals present at this time could have caused this effect. However it is young and unborn that are most sensitive to such disruption, since the correct levels and timing of hormones are crucial in the prenatal and postnatal development. In animal life, the most vulnerable group is the aquatic animals. Examples like these expand the field of NRs including molecular biologists, biochemists, epidemiologists, clinicians, zoologists and risk assessment experts to try to understand the NRs and their action.

MOLECULAR DYNAMICS SIMULATIONS

Simulations of particles have been performed on systems with sizes spanning from nanometers as in atoms, to meters as in simulations of human panic behavior in public rooms (57) and up to the distances measured in light years as in galaxies (58). The timescale on which these simulations act are almost as diverse as the size. From fractions of seconds in atoms simulations to fractions of the Universe age for galaxies. Despite such broad spans many of the simulations follow the classical Newtonian equations of motion even though sometimes corrections for other effects have to be taken into account. Moreover theoretical calculations often include approximations to some extent to simplify and speed up the computational time. The most precise calculation one can perform on molecules are the quantum mechanical (QM) calculations. In QM calculations both the structure and electron properties of the atoms are taken in account. Due to the high level of complexity QM calculations are computational expensive and limited to studies of hundreds of atoms. Thus QM calculations are usually performed on a selected part of a biomolecule while studies of larger biomolecules like proteins demands more approximate methods. Simulations of proteins are done on the molecular mechanics (MM) level, where atoms are treated as hard spheres connected to each other by harmonic springs. Molecular dynamics (MD) simulations of proteins gives detailed information on atom interaction and movement, but also show larger scale motions like protein dynamics.

SIMULATING PROTEINS

MD calculations of proteins have contributed to the understanding of how proteins work and behave in the cells. An example among many others (reviewed in (59,60)) is the MD study of the Src family of tyrosine kinases (61). Experiments on this system have shown that a dephoshorylation of a tyrosine at a position far off from the active site could shift the protein from an inactive to active form. How this could be achieved was revealed by MD simulation which showed that the dephosphorylation changes the protein dynamics making it move more and in an uncorrelated fashion. The flexibility is therefore the key factor for the Src family of tyrosine kinases activation. The MD simulations of Src family of tyrosine kinases were then used to guide mutational experiments of the protein. Close connections between theoretical and experimental studies gain both sides. MD simulations can provide experimentalists with detailed information on atom motions as a function of time, often of large interest in biomolecular function. Experimental data is crucial for MD simulations at different levels. Most importantly, the simulation requires the knowledge of the atom positions, given in 3D structure of the molecule, to start the simulation. Furthermore experimental data is important in validation and improvements of the MD results.
In this part of the thesis I will give a brief introduction to the field of MD simulations. First I explain the basic concept of a MD calculation on proteins and then I will introduce some modified MD techniques which can be used when classical MD is not sufficient.

FORCE FIELDS

To perform a MD simulation an equation of motion and a description of the energy of the system is needed. The energy is described as a potential energy function of the coordinates, of the atoms. From the first derivate of the potential energy function with respect to the coordinates, the forces acting on the systems can be obtained. Combining the calculation of the Newtonian equation of motion and the forces acting on the systems, the motion of the atoms in time, the trajectory of the system, can be acquired.
There are several programs available for MD calculations; common ones are CHARMM (62), Amber (63) and GROMACS (64). The parameters are empirically derived and therefore set to simulate the molecular behavior as accurately as possible (65).

The potential energy of a system is described as a sum of bond stretching, bond twisting, improper potentials, potentials describing the nonbonded van der Waals (vdW) interactions and electrostatic interactions (Figure 9). The bonds and angular terms are described as harmonic potentials, the vdW term as a simple Lennard-Jones potential and the electrostatic term as a Coloumb interaction between two point charges.

$$V = V_{bonds} + V_{angles} + V_{dihedrals} + V_{improper} + V_{vdW} + V_{elec} \quad \textit{Equation 1}$$

Using the first derivate of the potential energy function gives the force acting on that particular atom at that timestep.

$$F_{ix} = -\frac{d\phi}{dx_i} \quad \textit{Equation 2}$$

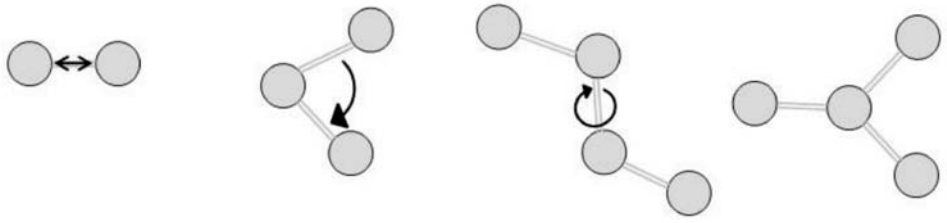

Figure 9. Potential contributions from bond stretching and angular bending between atoms. These terms are described by a harmonic potential function. The picture describes from the left, bond stretching, angular bending, dihedral bending and improper bending.

MOLECULAR DYNAMICS CALCULATIONS

The MD calculation uses Newton´s equation of motion to calculate the atom movements in time. This can be done if the atoms position and velocities are known or can be assumed and if all forces acting on each atom as a function of the atom position can be calculated. Knowing the 3D structure of the system and using the MD program´s force field, such conditions are generally fulfilled. The initial velocities are generated from a random number series, but since a complicated system very soon looses memory of the details of the initial conditions differences in the starting motions are eliminated.
The actual propagation of the coordinates can be achieved with different schemes and CHARMM can use a Verlet algorithm for this purpose. The algorithm is derived from two Taylor series expansions $t + \Delta t$ and $t - \Delta t$, where Δt is the timestep for the progress. Combining the Taylor expansion of the coordinates and the Newtons equation of motion generates the following expression for the coordinates of atom *i*.

$$x_i\,(t + \Delta t) = 2x_i(t) - x_i(t - \Delta t) + \frac{F_i(t)}{m_i}\Delta t^2 + O\Delta t^4 \qquad \textit{Equation 3}$$

From knowing two sets of coordinates, the velocities can be calculated.

$$v_i(t) = \frac{x_i(t+\Delta t) - x_i(t-\Delta t)}{2\Delta t} \qquad \textit{Equation 4}$$

Equation 3 includes the term F, which is the sum of the forces acting on that atom at that time. Thus at each timestep for each atom *equation 3* and *equation 1* have to be calculated. Doing so, the movement of the molecular system in time can be obtained.

A MD simulation can be performed under preferred environmental circumstances. Variables like temperature, pressure, solvent and salt concentration are easily manipulated and often selected to match the calculation purpose or for direct comparison to experimental data. The solvent representation is the single variable which will affect calculation time the most. Water is the most frequently used solvent and it can be represented as individual molecules (explicit representation) or as a continuum potential (implicit representation). The choice of representation depends on the system size and what kind of calculation one aims to perform. A detailed calculation demands a more accurate solvent representation and explicit solvent are recommended. The studies included in this thesis all use explicit solvent with the TIP3P model (66). This water model represent electrostatic charges at three sites, positive for the two hydrogens and negative for the oxygen, and the vdW surface represented by interaction points centered around the oxygen and hydrogens. The simulation time for a systems with explicit water are usually found between 10-100 ns today, using a 1 or 2 fs timestep (Δt) for the numerical integration (*Equation 3*).

A number of additional functions are included in the MD simulations performed by CHARMM, to speed up the calculation time or improve the accuracy of the simulation

results. Constraints on covalent bonds to hydrogens remove the highest frequency movements and allow the timestep to be increased from 0.5 or 1 fs to 2 fs using an algorithm such as SHAKE (67). Further on the computationally expensive long range vdW and electrostatic interactions can be reduced by introducing a cutoff radius, where the interactions are shifted or switched to a value close to zero (68). Since the system simulated is finite, the solvent molecules at the surface of the system will have one interaction side towards vacuum and one towards solvent. Unrealistic results created from such solvent representation might be avoided by using periodic boundary conditions (69) (Figure 10). Periodic boundary conditions image the system around itself avoiding abrupt boarders of the system. In such way the molecules of the right side will interact with the molecules on the left side of the image.

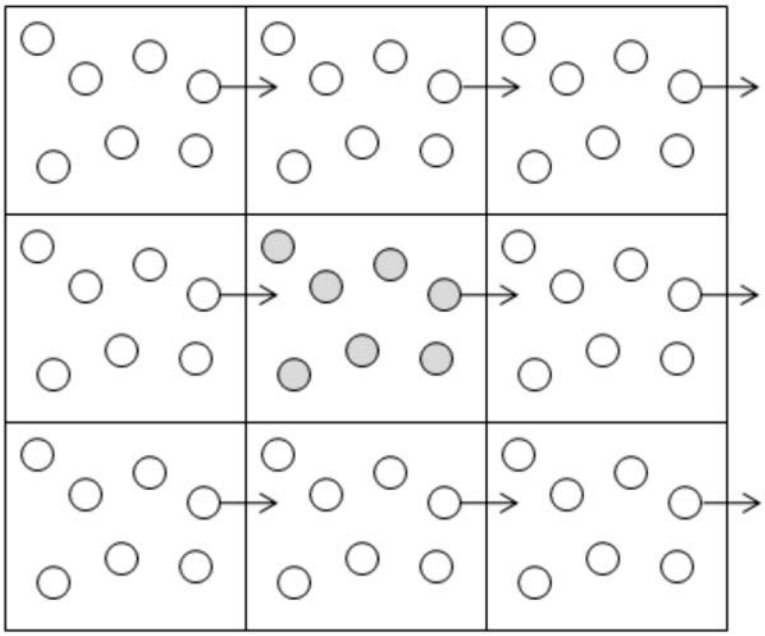

Figure 10. Schematic picture of periodic boundary conditions. The simulated system is indicated with grey spheres and its images with white spheres. The motions of the simulated system are also copied to the images as shown by arrows. Example of a motion is given by the arrow.

The obtained trajectory of the system can further be analyzed with a number of calculations. Common analysis are measurements of the change in position (root mean square deviation) or fluctuation (root mean square fluctuation) of the protein, identify hydrogen bonds or interactions between charged residues, and see how these properties change in time. Important is also to look at the trajectory with a molecular visualization program. In this thesis I have used the PyMOL tool (70) or Swiss PDB (71) for visualization of structures and production of pictures. To study the systems trajectory I have used VMD (72).

MODIFIED MOLECULAR DYNAMICS SIMULATIONS

Today biomolecular simulations are often used for studies of protein stability, dynamics and protein-protein or protein-ligand interactions which normally demand simulations on the nanosecond timescale. On this timescale vibration, rotation and translation movements can be observed, but many biochemical events take place on a timescale beyond this range. Molecular mechanisms like protein folding, domain movements or ligand binding have been shown in experiments to act on the micro second, seconds or even longer times (reviewed in (73)) Such events cannot be simulated by classical MD simulations mainly due to limited sampling of the molecules conformational space. When the simulation is started, minimization and dynamics often drive the system to the closest energy minimum (Figure 11). This minimum might be a local or the global minimum of the system. Depending on what state of the molecule one wants to study, the conformation either in the local or global minima is preferred. Most importantly however is the possibility to move between energy minima, something that can be difficult with classical MD. Several approaches have been used to improve the sampling. With the improvement of computational power, longer simulation times are possible but for now this approach is not sufficient for many studies. A second approach is to help the system over the energy barriers surrounding the energy minima. Doing so, a larger part of the conformational space can be explored.

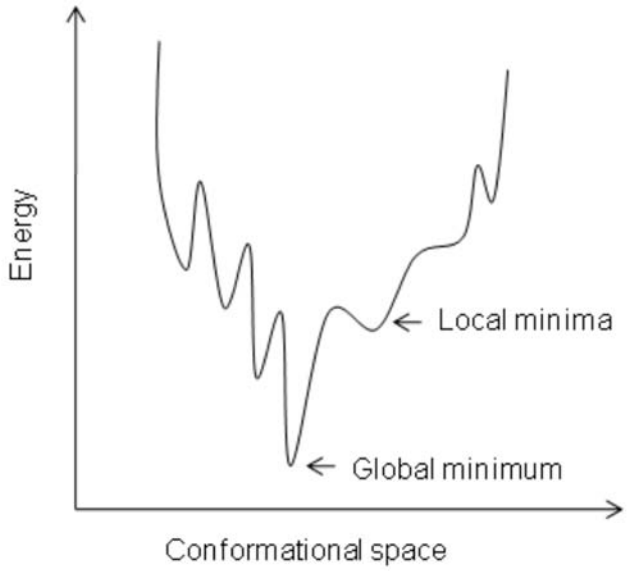

Figure 11. Schematic figure of a system´s energy function. The system simulated has several local minima and one global minimum. Therefore the probability that the simulated system will end up in a local minimum is higher.

RANDOM ACCELERATION MOLECULAR DYNAMICS

To study the mechanism of a molecule leaving its binding site a modified MD method is demanded. Several methods are available like locally enhanced sampling (74), steered MD (SMD) (75) and random acceleration MD (RAMD) (initially called Random Expulsion MD) (76). Such methods have previously been used for unbinding

studies on NRs (75,77-81). With the RAMD method the system works under the normal force field, except for a selected part, like a ligand, to which an additional force is applied. The force is given a random direction and will translate the ligand. The translation will continue until the ligand reaches an obstacle, such as a protein surface. At this point the ligand will stop and its velocity will decrease to a value close to zero. This will be detected by the RAMD method, which then gives the force a new random direction. The process is repeated and in such way the ligand is allowed to explore the binding compartment and find a more flexible part of the protein with a pathway where it can exit. The RAMD method was first used on the cytochrome p450 system, to which different ligands were bound (76). Using the RAMD method, unbinding pathways from the cytochrome p450 LBP was found. The major advantage of the RAMD method is that it demands no pre-knowledge of the pathways. The pathway search is done in an unbiased manner, giving all the possible pathways available for the system.

The pathway found in a single RAMD simulation is dependent on the initial random direction of the additional force. To obtain statistically convergent results many simulation repeats with different random direction of the force have to be employed. Typically, more than 30 repeats are used for the RAMD simulations included in this thesis. More aspects considering the RAMD variables and setup is discussed in Paper I.

STEERED MOLECULAR DYNAMICS

Similar to RAMD, SMD employ the normal force field for major parts of the system and an additional force for a selected part of the system. The selected part can be a ligand, but also single atoms or a larger part like a protein domain. For simplicity, I will use the example of a ligand in this SMD description. In SMD, the additional force is used to translate the ligand between two points in space. In practice, this is done by introducing a dummy atom to the system. A dummy atom is an atom that has no vdWs or electrostatic components and thus it cannot interact with the rest of the system, it is only used as a computational tool. The dummy atom is attached to the ligand by a harmonic spring. The dummy atom will then move between the start point and the final point with a constant velocity. When the dummy atom is translated, the attached ligand will also translate. At some point during the trajectory, the ligand might enter a narrow protein area which can cause repulsive interactions. At such point the ligand will try to slow the translation but since the dummy atom move with a constant velocity, the additional force is increased to overcome the local repulsive interactions of the ligand. Thus the force in SMD is variable and adjusts to the environmental impact on the ligand. From analysis of the SMD trajectory a force profile for the pathway can be extracted. Such profile is useful in the characterization of the pathway, describing pathway bottlenecks and to estimate the free energy of binding. However, the SMD method setup demands the knowledge of the start and final point of the translation, the pathway for translation, which in many cases is unknown. Therefore a combinatory scheme with another method e.g. RAMD, to define the pathway and then further characterize it with SMD, is preferential.

ANISOTROPIC THERMAL DIFFUSION

Another event difficult to study with classical MD is the communication between amino acids, particular at distal sites. Studies of the protein´s correlated movement or normal mode analysis (reviewed in (82)) have presented good results (83), but these methods often demand long simulations times due to the low signal-to-noise ratio. The anisotropic thermal diffusion (ATD) method increases the signal-to-noise ratio and thereby makes it possible to study molecular signaling on a picosecond timescale (84). The simulated system is first equilibrated at 300K and then cooled to 10K. At 10K the movements in the protein is almost reduced to zero. To this cooled protein a selected part, e.g. an amino acid sidechain, is heated to 300K. The heat gathered in this sidechain will propagate out to the surrounding amino acids. If the heat propagation is evenly distributed to the close environment, the heated amino acid had no clear role in any molecular communication pathway. But if the propagation of the heat is done with directionality, it is possible that this energetically favorable pathway also plays a role in molecular communication. The ATD method was first tested on the PDZ (post synaptic density protein) protein for which the intramolecular communication pathway is experimentally known. These results indicated a good correlation between experimental data, correlation covariance calculations and ATD results, suggesting that the ATD method is a good tool for studies of allosteric communications.

SUMMARY OF MAIN RESULTS

PAPER I

Ligand binding is a key event in the action of NRs. Experimental studies of the kinetic properties of agonist and antagonist binding have revealed differences between the two ligand types, which indicate different binding and or unbinding mechanisms (85). The mechanism of binding/unbinding is so far not known on a molecular level. Structural studies of the apo state and ligand bound receptor indicates differences in the H12 position (17,86), but to what extent and for which receptors the H12 would participate in the ligand binding/unbinding mechanism is unknown.

In this study the retinoic acid (RA) unbinding from RAR was studied with the RAMD method. The RAMD method had previously not been used on highly flexible, elongated and charged ligands as RA. Therefore the first approach of this study was to optimize the RAMD protocol to such ligands. The try outs of the variables were performed in a simple 2D model system, only using a carbon atom wall and carbon atom ligands with different properties. This initial study revealed that the force application mode is crucial for ligands like RA. It is particularly the electrostatic interactions that will affect the outcome.

The results from the RAMD study on RAR showed four unbinding pathways (Figure 12). The frequency by which the RAMD method detected the pathway and the distortion on the RAR structure was used to rank the pathways. Doing so one of the pathways turned out to be more likely than the others. None of the detected pathways was found in the H12 region. This indicates that unbinding can take place from the RAR not involving larger conformational changes like translation of H12.

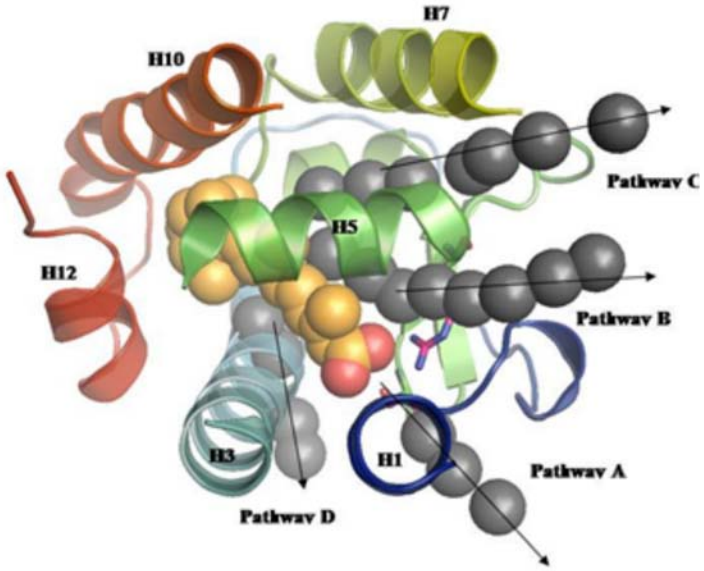

Figure 12. Unbinding pathways of RA from RAR. The initial RA (gold/red spheres) position is shown in the receptors LBP (multicolored cartoon) and its unbinding trajectory is shown as a trace of a central RA atom (grey spheres).

PAPER II

The second paper studies the ligand unbinding mechanism from the ERs with three different ligands. The aim is to find possible unbinding pathways and rank them to each other. This is particularly interesting since the ligands selected for the study span from full agonist to SNRM to antagonist. Secondly, both ER subtypes α and β are used to provide further insight on the selectivity properties observed for the SNRM. The RAMD method was used to identify all the unbinding pathways which then were further characterized by the SMD method. A third method, CAVER, which performs a search for unoccupied volumes in the static protein structure, was also employed. The results from all three methods were combined for the ranking and characterization of the pathways.

The RAMD method identified in total seven possible unbinding pathways for the agonist and SNRM in the ERs (Figure 13). Differences were observed between receptor subtypes and ligand types, indicating that specific receptor-ligand combinations can employ different unbinding pathways. In general pathways 1, 2 and 3 are the highest ranked pathways for the ERs. Interestingly, the antagonist did not unbind with the RAMD method, although the additional applied force was increased significantly. This indicates different binding/unbinding mechanism for NR agonists and antagonists, which was also suggested by experimental results (85).

The RAMD results also revealed different unbinding pathways for the SNRM in the two ER subtypes, where a pathway below the β-sheet was selected in ERα, while in ERβ the SNRM exited above the β-sheet. This preference probably arises from the different degrees of hydrophobicity between the pathways. In the pathway below the β-sheet the ligand has to pass a compartment where no polar interactions are possible. While in the pathway above the β-sheet, several polar interactions are possible, especially in ERβ that has two subtype specific polar amino acids in this region (Figure 13). The more polar environment is probably preferred by the polar SNRM studied, selecting this pathway to a higher extent. Thus specificity might be explained by differences between receptor subtypes, outside the LBP.

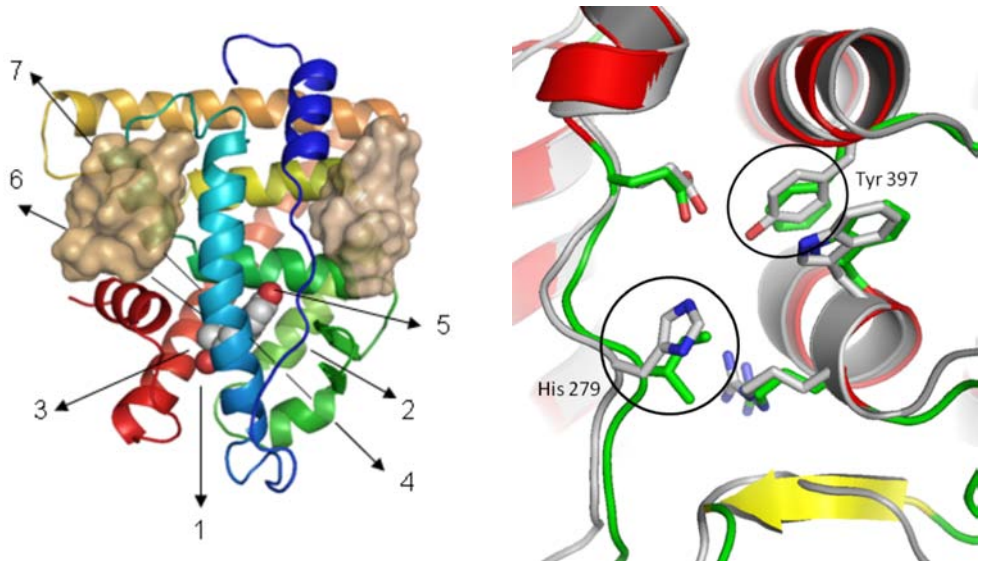

Figure 13. Ligand unbinding pathways in ERs (left picture). Pathway 2 (right picture), above the β-sheets, hold subtype specific polar amino acids for ERβ (in circle). The ERβ (grey cartoon) structure is superimposed on ERα (colored cartoon)

PAPER III

Several 3D structures of the orphan receptor LRH-1 indicated that the LBD adopts an active conformation without a ligand bound in its LBP. However, one 3D structure of the LRH-1 LBD revealed a bound phospholipid and the receptor in the active conformation. Ligand binding is an interesting aspect also in orphan receptors, mainly due to its possibilities of drug discovery. Although the biological importance for this particular phospholipid might be discussed, its presence might impact the receptor and binding to cofactor peptides. MD simulations might provide insights on how the always active LRH-1 reacts to the binding of a ligand, and if its interactions with different cofactor peptides are modified.

The results from the MD simulations of LRH-1 with or without ligand bound, in complex with four different cofactor peptides, showed that the receptor conformation was preserved. It is therefore likely that the orphan receptor LRH-1 might bind a ligand and maintain its active conformation. The receptor interactions with cofactor peptides were modified in diverse ways in the context of a ligand. Two peptide interactions were decreased when a ligand was introduced, while one interaction was unchanged and one increased with a ligand present. Thus it is likely that the regulatory effect of a bound ligand to LRH-1 is found in its interaction to cofactor peptides, as observed for many SNRMs.

The interaction between the receptor-cofactor peptides was also characterized. A specific interaction was observed between an aspartic acid sidechain from the receptor and the -4 and/or +2 amino acid on the cofactor peptide (Figure 14). Structural comparisons with other NRs indicated that such interaction might be possible for several members of the NR superfamily.

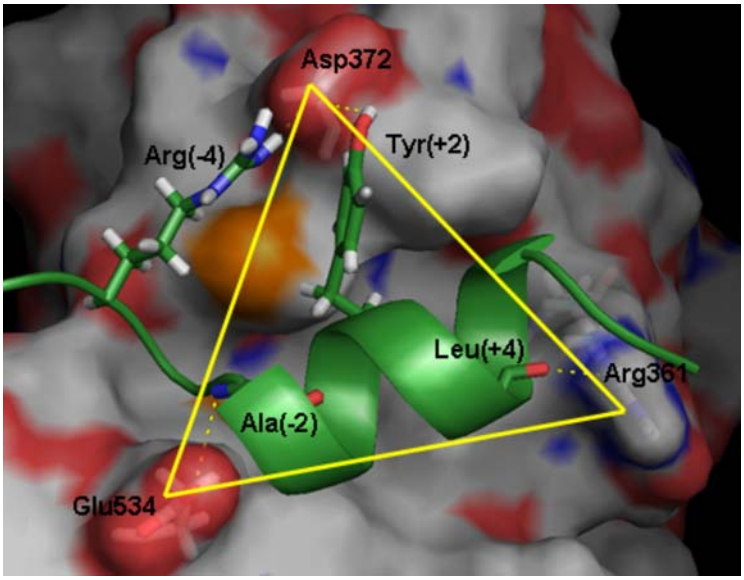

Figure 14. Interactions between LRH-1 and a cofactor peptide. The peptide (green cartoon and sticks) is oriented by the conserved charge clamp of NRs, here Glu534 and Arg361 (CPK colored surface and sticks). Additional specificity can be obtained with interactions to an aspartic acid (Asp372) here shown to the peptide´s -4 and +2 amino acids. Together with the charge clamp the aspartic acid forms a triangular shaped interaction pattern for the cofactor peptide.

PAPER IV

The LXRα is an important regulator of genes involved in metabolism and inflammation. The receptor binds various types of ligands and the flexible ligand binding pocket can adapt to diverse ligand structures. The dynamics of the flexible LBP was decreased upon cofactor binding, in the closely related pregnane X receptor (87). Flexibility of the LBP might be important in the ligand binding mechanism and in drug development. MD simulations of a ligand bound in different conformations to the LXR LBP might indicate if LBP flexibility is an important aspect for LXR as well. The communication between the LBP and the AF-2 region with the cofactor bound is an example of the allosteric signaling pathways detected in several NRs. Here we use the ATD method to identify and characterize such signal pathway. Further on the LXR interactions with nine different cofactor peptides are characterized, providing insights to the recognition and specificity between LXR and its cofactors.

The ATD method was able to identify and characterize a signal pathway between the coactivator peptide and the ligand in the LXR LBP (Figure 15). The signal transmission went through the receptor sidechain and backbone, in a cooperative manner, indicating that both these parts of the polypeptide chain are important for allosteric signaling. Interactions between the LXR and the nine cofactor peptides revealed a somewhat different interaction pattern between the receptor-cofactor complexes. Most complexes used the charge clamp interaction for interactions, sometimes with an additional specific interaction. Only in one case the specific interaction was more important than the charge clamp interaction.

The preliminary results from the flexibility of the LBP in the simulations show that the LBP dynamics is reduced or unchanged, when a cofactor peptide is bound to the receptor. This indicates that drug development targeting LXR should consider a flexible receptor structure for docking studies.

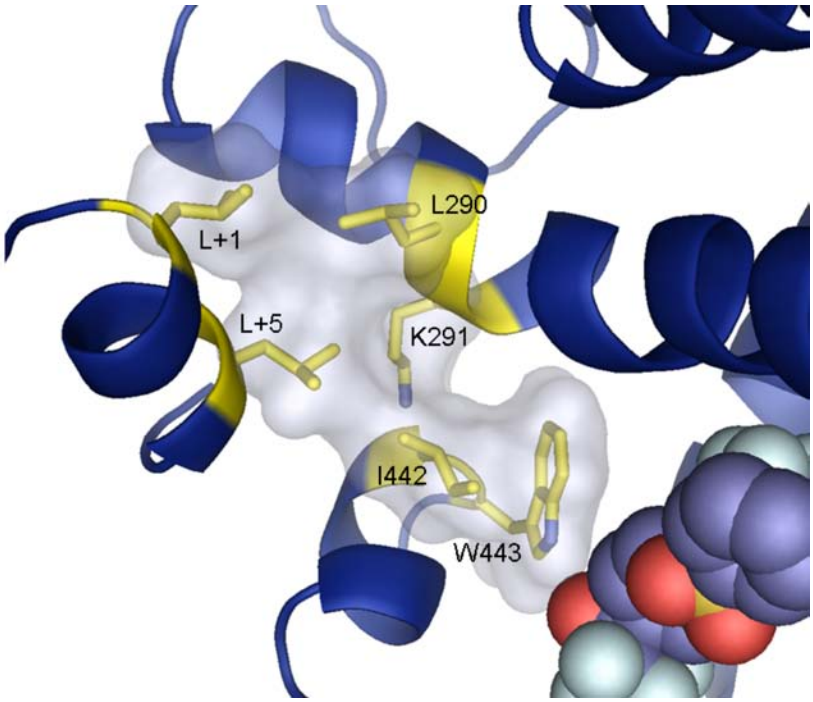

Figure 15. Communication pathway between the coactivator peptide and ligand in the LXR LBP, identified with the ATD method. Key residues for signaling are the coactivator L+1 and receptor K291. From K291, the signal might pass the H5 backbone to the right or pass over to H12 I442-W443 sidechains, on its way to the ligand.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

"Everything that living things do can be understood in terms of the jigglings and wigglings of atoms"

Richard Feynman

Although quite controversial when stated almost 50 years ago, Feynman´s statement has proven to hold in many situations. MD simulations have often been used as a tool for such studies when experimental techniques are not sufficient. However, the MD techniques can be further improved. Most important is increased sampling of the conformational space. Today several modified MD techniques like replica exchange MD (88), locally enhanced sampling MD (74) and RAMD or SMD presented in this thesis, have been developed to improve the sampling. The results from modified MD studies often bring useful insight but the interpretation of the results should be done with care. Modified MD techniques often introduce highly artificial components and the effects from these have to be taken into account when evaluating the results. One such example was observed in paper II, where the RAMD and SMD sampling turned out to be fundamentally different. The results in such study can be interpreted in a number of variant ways, where an accurate evaluation should take into account the sampling properties of the method.
Secondly sampling can be improved with simplified biomolecular models, like grid modeling. A grid model maps out the biomolecule with grid points and performs the simulation on the grid points, which significantly improves the computational time. However, much of the details are lost in such simulations. Therefore the best approach to obtain the detailed information and also improve sampling would be to increase the simulation time. Doing so, some groups have reported instabilities in the force field when simulation time exceeds 100 ns or so. Thus force field improvements and increased computational power will probably make way for much longer simulation times in the future.

The MD technique can be a powerful tool in combination with other computational techniques. Combining MD and QM calculation is one such example, where precise data on binding energies from QM can be combined with the dynamic influences from MD (89). The QM/MM combination has been limited due to difficulties with the interface representation between the methods (reviewed in (90)), but improvements on this will probably increase the use for such calculations. Further on the introduction of MD into the drug discovery field has turned out to be successful (91,92). A typical screening procedure for a new drug target involves thousands or millions of compounds but only one or a few receptor structures. Introduction of flexibility to the receptor improves the prediction of the ligand binding mode significantly but also increased the computational time. Although high speed is crucial for the drug discovery research, the observed improvements when including MD justifies the time consumed. Therefore it is likely that a standard *in silico* drug development procedure will include the MD technique to some extent in the future.

One of the time consuming parts of the MD technique is the setup of the system and parameterization of ligands. The setup of the system to be simulated involves several steps of evaluation which demands structural and biochemical knowledge of the system. This limits the availability of the MD technique excluding researcher with less such experience. Attempts to automatize MD setup and parameterization are under progress and could dramatically increase the use of the MD techniques. Although the field of MD would benefit from such improvement, it is also important to stress the understanding of the method when evaluating the results.

The computational studies of NRs have so far been limited to studies of individual domains, due to lack of DBD-LBD complex 3D structures. Larger complexes including multiple domains and full length cofactors have been difficult to crystallize for X-ray crystallography, and are too big for nuclear magnetic resonance studies. However, last year pioneering results in NR structure prediction were obtained, when both a full-length cofactor-receptor complex and DBD-LBD NR heterodimer in complex with DNA were solved. These studies show unexpected interaction surfaces which undoubtedly will impact the future NR research. Presumable molecular biology progress in the mapping of the NR interaction proteins will further stress the importance of structural determination of the NR multi protein complex. A platform of knowledge of the NR interaction proteins and its interaction surfaces on the receptor provides a potential for extensive MD simulations. Such scenario might be realistic in just a few years from now. At that point the computational power is enhanced and improvements of MD sampling techniques have hopefully been made, making MD simulations a technique for the understanding of the living things, in terms of jiggling and wiggling.

POPULÄRVETENSKAPLIG SAMMANFATTNING PÅ SVENSKA

Kroppen består av hundratals olika typer av celler, som skiljer sig åt i var de är placerade och vad de ska fungera för. Cellens arbetare är ofta proteinerna, som ser till så att cellens funktioner blir utförda. Då proteinet tar emot en signal om arbete i form av en liten molekyl kallas proteinet för en receptor. Det finns flera typer av receptorer men i denna avhandling har jag studerat receptorer som har sin funktion i cellens kärna, s.k. kärnreceptorer. I cellens kärna finns även DNA´t, den genetiska koden, till vilken kärnreceptorerna kan binda. Kärnreceptorerna blir därmed en direkt länk till cellens beteende och kontrollerar flera av de viktigaste funktionerna i cellen som cellens utveckling, ämnesomsättning och död. Eftersom dessa funktioner ingår i många vanliga sjukdomar använder man kärnreceptorerna som mål för läkemedel. Exempel på sjukdomar som man kan bota eller lindra är cancer, inflamamtion, typ II diabetes, fettma och hjärt- och kärlsjukdomar. Dessa sjukdomar är idag folksjukdomar och antalet personer som drabbas ökar varje år. Detta har gjort kärnreceptorerna till den näst mest populära familjen att inrikta ett läkemedel mot, enligt USA´s motsvarighet på läkemedelsverket, FDA.

Kärnreceptorerna är en protein familj vars funktion styrs av en signal, ofta en liten molekyl en sk. ligand. Liganden binder till receptorns ligand bindande del som då ändrar sitt utseende till en aktiv form. Till den nya aktiva formen kan andra proteiner binda och receptorn kan därefter binda till DNA´t. Bindandet av liganden är en viktig händelse för receptorns aktivering. Trots det så vet man lite om hur detta går till. I de två första artiklarna har jag studerat möjliga mekanismer för hur liganden kan lämna eller binda receptorn. Resultaten visar att det krävs mycket små rörelser hos receptorn för att få en liten ligand att komma ut ur receptorn. Ligandens egenskaper, som om den gillar vatten eller fett-miljöer, avgör också vilken väg den tar. En ligand som gillar vatten kommer alltså att ta en väg där det finns vatten omkring sig och tvärtom. Att olika ligander tar olika väger kan ha betydelse för framtida läkemedels utveckling.

I den aktiva formen binder receptorn till andra proteiner, medhjälpare för protein produktionen sk. co-faktorer. Receptorn har speciella områden där dessa kan binda och interaktionen sker alltid med en konserverad del av co-faktorn. Trots att interaktionen är så väl bevarad inom hela kärnreceptor familjen kan bara vissa co-faktorer binda vissa receptorer. För att försöka förstå bindningen bättre studerade jag i artikel 3 och 4 hur olika co-faktorer binder sin receptor. För receptorn beskrivd i artikel 3 (LRH-1), visade sig att det finns kontakter utanför den konserverade delen av co-faktorerna som kan användas för att få en mer specifik binding. För receptorn (LXR), beskrivd i artikel 4, verkade det dock som att den konserverade kontakten är viktigast.

Receptorn har skilda delar för sina olika funtioner tex så binder liganden på ett ställe och co-faktorn på ett annat. Trots de åtskillda platserna finns en kommunikation mellan dess delar. Denna kommunikation kallas allosterisk signalering. Hos kärnreceptorera finns en allosterisk kommunikation mellan den ligand bindande fickan och ytan där co-faktorer binder. I artikel 4 försöker jag att hitta denna signalväg och beskriva hur denna signalering kan ske. Jag fann en möjlig väg och såg att signalering verkar ske genom både ändrad rörelse hos proteinet och ändrad form. Identifikation av allosteriska vägar gör det möjligt att bättre förstå varför vissa mutationer är farligare än andra.

ACKNOWLEDGEMENTS

During these years several persons have supported and enhanced my work and life as a PhD student, I would like to thank everyone that contributed and in special:

Prof. Lennart Nilsson, for supervising me during these years always with great patience and clear thoughts. I also want to thank you for allowing me to work freely, which has forced me to develop as a researcher and person.

My collaborators: Prof. Eckardt Treuter for the biological input on NRs and critical reading of manuscripts. Dr. Peter Carlsson for including me in the RAMD project and Dr. Cristian Danciulescu for all the work on the parameterization.

The present and former members of the molecular modeling group; Mika, Olof, Cristian, Greg and Vineet for great company at MF lunches. Jianxin, Jan, Johan, Ana and Pekka. Mattias for beeing such a nice roommate. Boel and Katarina for all the chit chats on life, science and whatever happened or not happened at the moment.

Other CSB members; Caroline, Katja, Magnus and previous members Micke, Tobbe and Linda. Kimberly, for all the pep-talk and energy! Anna-Karin, for good laughs and the accurate update on gossips.

Prof. Kurt Berndt and Doc. Lotta Hambreus for introducing me to the field of structural biology and theoretical chemistry. Your teaching at Södertörns Högskola was excellent.

Dr. Luca Jovine for including me in the ZP3 project and for bringing so much enthusiasm to the work. I really felt our work was a matter of life or inferlity!

Prof. Akiyama at TiTech, Tokyo, for kindly accepting me to your lab and generously introducing me the research in Japan. The Akiyama group and particularly Dr. Yuri Matsuzaki for nice discussions and taking care of me during my stay.

My dear and close friends; Karin, Tanja and Linda for all the great times we had, all the laughs, tears and discussions. You are absolutely outstanding! Emilia, for being a true friend and an infinite source of inspiration. Samuel, for all the long walks which were so much fun but also straightened up my head. Camilla, for discussions on complicated life-situations. Oliver, for long time friendship.

Min favorit syster yster, Nina, för att du är bäst och ditt pladder i mitt öra kittlar mig till skratt!

Mamma, för outömligt stöd och stor personlighet. Pappa och Margareta, för att ni alltid uppmuntrar mig i jobb och privatliv. Greta, för allt stöd och din uppfriskande syn på tillvaron. Maaret, för att pumps visst funkar i alla väder. Farmor, för 90 år av nyfikenhet, tänk om man fick bli som du!

Johan, för att du tror på mig när jag tappat den och fyller mina dagar med glädje. ♥

REFERENCES

1. Mangelsdorf, D. J., C. Thummel, M. Beato, P. Herrlich, G. Schutz, K. Umesono, B. Blumberg, P. Kastner, M. Mark, P. Chambon and R. M. Evans. (1995) The nuclear receptor superfamily: the second decade. *Cell* 83, 835-9.
2. Moras, D. and H. Gronemeyer. (1998) The nuclear receptor ligand-binding domain: structure and function. *Curr Opin Cell Biol* 10, 384-91.
3. Aranda, A. and A. Pascual. (2001) Nuclear hormone receptors and gene expression. *Physiol Rev* 81, 1269-304.
4. Overington, J. P., B. Al-Lazikani and A. L. Hopkins. (2006) How many drug targets are there? *Nat Rev Drug Discov* 5, 993-6.
5. McCammon, J. A., B. R. Gelin and M. Karplus. (1977) Dynamics of folded proteins. *Nature* 267, 585-90.
6. Willson, T. M. and J. T. Moore. (2002) Genomics versus orphan nuclear receptors--a half-time report. *Mol Endocrinol* 16, 1135-44.
7. Escriva, H., F. Delaunay and V. Laudet. (2000) Ligand binding and nuclear receptor evolution. *Bioessays* 22, 717-27.
8. Escriva, H., R. Safi, C. Hanni, M. C. Langlois, P. Saumitou-Laprade, D. Stehelin, A. Capron, R. Pierce and V. Laudet. (1997) Ligand binding was acquired during evolution of nuclear receptors. *Proc Natl Acad Sci U S A* 94, 6803-8.
9. Tupler, R., G. Perini and M. R. Green. (2001) Expressing the human genome. *Nature* 409, 832-3.
10. Nuclear Receptor Nomenclature Committee (1999) A unified nomenclature system for the nuclear receptor superfamily. *Cell* 97, 161-3.
11. Kumar, R. and E. B. Thompson. (1999) The structure of the nuclear hormone receptors. *Steroids* 64, 310-9.
12. Pissios, P., I. Tzameli, P. Kushner and D. D. Moore. (2000) Dynamic stabilization of nuclear receptor ligand binding domains by hormone or corepressor binding. *Mol Cell* 6, 245-53.
13. Egea, P. F., N. Rochel, C. Birck, P. Vachette, P. A. Timmins and D. Moras. (2001) Effects of ligand binding on the association properties and conformation in solution of retinoic acid receptors RXR and RAR. *J Mol Biol* 307, 557-76.
14. Lu, J., D. P. Cistola and E. Li. (2006) Analysis of ligand binding and protein dynamics of human retinoid x receptor alpha ligand-binding domain by nuclear magnetic resonance. *Biochemistry* 45, 1629-39.
15. Luisi, B. F., W. X. Xu, Z. Otwinowski, L. P. Freedman, K. R. Yamamoto and P. B. Sigler. (1991) Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature* 352, 497-505.
16. Renaud, J. P., N. Rochel, M. Ruff, V. Vivat, P. Chambon, H. Gronemeyer and D. Moras. (1995) Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature* 378, 681-9.
17. Bourguet, W., M. Ruff, P. Chambon, H. Gronemeyer and D. Moras. (1995) Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-alpha. *Nature* 375, 377-82.
18. Chandra, V., P. Huang, Y. Hamuro, S. Raghuram, Y. Wang, T. P. Burris and F. Rastinejad. (2008) Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. *Nature* 456, 350-6.
19. Stunnenberg, H. G. (1993) Mechanisms of transactivation by retinoic acid receptors. *Bioessays* 15, 309-15.
20. Nettles, K. W. and G. L. Greene. (2005) Ligand control of coregulator recruitment to nuclear receptors. *Annu Rev Physiol* 67, 309-33.
21. Gronemeyer, H., J. A. Gustafsson and V. Laudet. (2004) Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* 3, 950-64.
22. Rochette-Egly, C. (2003) Nuclear receptors: integration of multiple signalling pathways through phosphorylation. *Cell Signal* 15, 355-66.
23. Rochette-Egly, C. (2005) Dynamic combinatorial networks in nuclear receptor-mediated transcription. *J Biol Chem* 280, 32565-8.

24. Perissi, V. and M. G. Rosenfeld. (2005) Controlling nuclear receptors: the circular logic of cofactor cycles. *Nat Rev Mol Cell Biol* 6, 542-54.
25. Kassel, O., S. Schneider, C. Heilbock, M. Litfin, M. Gottlicher and P. Herrlich. (2004) A nuclear isoform of the focal adhesion LIM-domain protein Trip6 integrates activating and repressing signals at AP-1- and NF-kappaB-regulated promoters. *Genes Dev* 18, 2518-28.
26. Herrlich, P. (2001) Cross-talk between glucocorticoid receptor and AP-1. *Oncogene* 20, 2465-75.
27. Bjornstrom, L. and M. Sjoberg. (2005) Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes. *Mol Endocrinol* 19, 833-42.
28. Safe, S. and K. Kim. (2008) Non-classical genomic estrogen receptor (ER)/specificity protein and ER/activating protein-1 signaling pathways. *J Mol Endocrinol* 41, 263-75.
29. Rosenfeld, M. G., V. V. Lunyak and C. K. Glass. (2006) Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes Dev* 20, 1405-28.
30. Warnmark, A., E. Treuter, A. P. Wright and J. A. Gustafsson. (2003) Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation. *Mol Endocrinol* 17, 1901-9.
31. Radhakrishnan, I., G. C. Perez-Alvarado, D. Parker, H. J. Dyson, M. R. Montminy and P. E. Wright. (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell* 91, 741-52.
32. Uesugi, M., O. Nyanguile, H. Lu, A. J. Levine and G. L. Verdine. (1997) Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* 277, 1310-3.
33. McEwan, I. J., K. Dahlman-Wright, J. Ford and A. P. Wright. (1996) Functional interaction of the c-Myc transactivation domain with the TATA binding protein: evidence for an induced fit model of transactivation domain folding. *Biochemistry* 35, 9584-93.
34. Heery, D. M., E. Kalkhoven, S. Hoare and M. G. Parker. (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* 387, 733-6.
35. Li, Y., M. H. Lambert and H. E. Xu. (2003) Activation of nuclear receptors: a perspective from structural genomics. *Structure* 11, 741-746.
36. McInerney, E. M., D. W. Rose, S. E. Flynn, S. Westin, T. M. Mullen, A. Krones, J. Inostroza, J. Torchia, R. T. Nolte, N. Assa-Munt, M. V. Milburn, C. K. Glass and M. G. Rosenfeld. (1998) Determinants of coactivator LXXLL motif specificity in nuclear receptor transcriptional activation. *Genes Dev* 12, 3357-68.
37. Sablin, E. P., A. Woods, I. N. Krylova, P. Hwang, H. A. Ingraham and R. J. Fletterick. (2008) The structure of corepressor Dax-1 bound to its target nuclear receptor LRH-1. *Proc Natl Acad Sci U S A* 105, 18390-5.
38. Metivier, R., G. Penot, M. R. Hubner, G. Reid, H. Brand, M. Kos and F. Gannon. (2003) Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* 115, 751-63.
39. Nagy, L., H. Y. Kao, J. D. Love, C. Li, E. Banayo, J. T. Gooch, V. Krishna, K. Chatterjee, R. M. Evans and J. W. Schwabe. (1999) Mechanism of corepressor binding and release from nuclear hormone receptors. *Genes Dev* 13, 3209-16.
40. Gurevich, I., A. M. Flores and B. J. Aneskievich. (2007) Corepressors of agonist-bound nuclear receptors. *Toxicol Appl Pharmacol* 223, 288-98.
41. Shulman, A. I., C. Larson, D. J. Mangelsdorf and R. Ranganathan. (2004) Structural determinants of allosteric ligand activation in RXR heterodimers. *Cell* 116, 417-29.
42. Swain, J. F. and L. M. Gierasch. (2006) The changing landscape of protein allostery. *Curr Opin Struct Biol* 16, 102-8.
43. Popovych, N., S. Sun, R. H. Ebright and C. G. Kalodimos. (2006) Dynamically driven protein allostery. *Nat Struct Mol Biol* 13, 831-8.

44. McDonnell, D. P., D. J. Mangelsdorf, J. W. Pike, M. R. Haussler and B. W. O'Malley. (1987) Molecular cloning of complementary DNA encoding the avian receptor for vitamin D. *Science* 235, 1214-7.
45. Conzen, S. D. (2008) Minireview: nuclear receptors and breast cancer. *Mol Endocrinol* 22, 2215-28.
46. Lonard, D. M., R. B. Lanz and B. W. O'Malley. (2007) Nuclear receptor coregulators and human disease. *Endocr Rev* 28, 575-87.
47. Gervois, P., J. C. Fruchart and B. Staels. (2007) Drug Insight: mechanisms of action and therapeutic applications for agonists of peroxisome proliferator-activated receptors. *Nat Clin Pract Endocrinol Metab* 3, 145-56.
48. Kersten, S., B. Desvergne and W. Wahli. (2000) Roles of PPARs in health and disease. *Nature* 405, 421-4.
49. Sladek, F. M. (2003) Nuclear receptors as drug targets: new developments in coregulators, orphan receptors and major therapeutic areas. *Expert Opin Ther Targets* 7, 679-84.
50. Gabbi, C., M. Warner and J. A. Gustafsson. (2009) Minireview: liver X receptor beta: emerging roles in physiology and diseases. *Mol Endocrinol* 23, 129-36.
51. Glass, C. K. and S. Ogawa. (2006) Combinatorial roles of nuclear receptors in inflammation and immunity. *Nat Rev Immunol* 6, 44-55.
52. Bensinger, S. J. and P. Tontonoz. (2008) Integration of metabolism and inflammation by lipid-activated nuclear receptors. *Nature* 454, 470-7.
53. Shi, Y. (2007) Orphan nuclear receptors in drug discovery. *Drug Discov Today* 12, 440-5.
54. Tata, J. R. (2005) One hundred years of hormones. *EMBO Rep* 6, 490-6.
55. Colborn, T., D. Dumanoski and J. Peterson Myers. 1996. Our stolen future. New York, NY, USA: Dutton.
56. Sharpe, R. M. and N. E. Skakkebaek. (1993) Are oestrogens involved in falling sperm counts and disorders of the male reproductive tract? *Lancet* 341, 1392-5.
57. Helbing, D., I. Farkas and T. Vicsek. (2000) Simulating dynamical features of escape panic. *Nature* 407, 487-90.
58. Barnes, J. (1989) Evolution of compact groups and the formation of elliptical galaxies. *Nature* 338, 123-126.
59. Dodson, G. G., D. P. Lane and C. S. Verma. (2008) Molecular simulations of protein dynamics: new windows on mechanisms in biology. *EMBO Rep* 9, 144-50.
60. Karplus, M. and J. Kuriyan. (2005) Molecular dynamics and protein function. *Proc Natl Acad Sci U S A* 102, 6679-85.
61. Young, M. A., S. Gonfloni, G. Superti-Furga, B. Roux and J. Kuriyan. (2001) Dynamic coupling between the SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation. *Cell* 105, 115-26.
62. Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus. (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4, 187-217.
63. Weiner, S. J., P. A. Kollman and e. al. (1984) A New Force-Field fpr Molecular Mechanical Simulations of Nucleic-Acids and Proteins. *Journal of the American Chemical Society* 106, 765-784.
64. van Gunsteren, A. and H. Berendsen. 1987. GROMOS (GROningen MOlecular Simulation) package. Nijenborgh 4, 9747 AG Groningen: BIOMOS bv.
65. MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* 102102, 3586-3616.
66. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein. (1983) Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* 79, 926-935.

67. Ryckaert, J. P., G. Ciccotti and H. J. C. Berendsen. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *Journal of Computational Chemistry*327-341.
68. Steinbach, P. J. and B. R. Brooks. (1994) New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular SImulations. *Journal of Computational Chemistry* 15, 667-683.
69. Brooks, C. L., 3rd, A. Brunger and M. Karplus. (1985) Active site dynamics in protein molecules: a stochastic boundary molecular-dynamics approach. *Biopolymers* 24, 843-65.
70. DeLano, W. L. 2008. The PyMOL Molecular Graphics System. http://www.pymol.org. Palo Alto, CA, USA: DeLano Scientific LLC.
71. Guex, N. and M. C. Peitsch. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* 18, 2714-2723.
72. Humphrey, W., A. Dalke and K. Schulten. (1996) VMD - Visual Molecular Dynamics. *J. Molec. Graphics* 14, 33-38.
73. Elber, R. (2005) Long-timescale simulation methods. *Curr Opin Struct Biol* 15, 151-6.
74. Elber, R. and M. Karplus. (1990) Enhanced Sampling in Molecular-Dynamics - Use of the Time-Dependent Hartree Approximation for a Simulation of Carbon-Monoxide Diffusion through Myoglobin. *Journal of the American Chemical Society* 112, 9161-9175.
75. Kosztin, D., S. Izrailev and K. Schulten. (1999) Unbinding of retinoic acid from its receptor studied by steered molecular dynamics. *Biophysical Journal* 76, 188-197.
76. Ludemann, S. K., V. Lounnas and R. C. Wade. (2000) How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms. *Journal of Molecular Biology* 303, 797-811.
77. Martínez, L., M. T. Sonoda, P. Webb, J. D. Baxter, M. S. Skaf and I. Polikarpov. (2005) Molecular Dynamics Simulations Reveal Multiple Pathways of Ligand Dissociation from Thyroid Hormone Receptors. *Biophys. J.* 89, 2011-2023.
78. Martinez, L., I. Polikarpov and M. S. Skaf. (2008) Only Subtle Protein Conformational Adaptations Are Required for Ligand Binding to Thyroid Hormone Receptors: Simulations Using a Novel Multipoint Steered Molecular Dynamics Approach. *J Phys Chem B.*
79. Sonoda, M. T., L. Martinez, P. Webb, M. S. Skaf and I. Polikarpov. (2008) Ligand dissociation from estrogen receptor is mediated by receptor dimerization: evidence from molecular dynamics simulations. *Mol Endocrinol* 22, 1565-78.
80. Blondel, A., J. P. Renaud, S. Fischer, D. Moras and M. Karplus. (1999) Retinoic acid receptor: A simulation analysis of retinoic acid binding and the resulting conformational changes. *Journal of Molecular Biology* 291, 101-115.
81. Genest, D., N. Garnier, A. Arrault, C. Marot, L. Morin-Allory and M. Genest. (2008) Ligand-escape pathways from the ligand-binding domain of PPARgamma receptor as probed by molecular dynamics simulations. *Eur Biophys J* 37, 369-79.
82. Ma, J. (2005) Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure* 13, 373-80.
83. Gaillard, T., E. Martin, E. San Sebastian, F. P. Cossio, X. Lopez, A. Dejaegere and R. H. Stote. (2007) Comparative normal mode analysis of LFA-1 integrin I-domains. *J Mol Biol* 374, 231-49.
84. Ota, N. and D. A. Agard. (2005) Intramolecular signaling pathways revealed by modeling anisotropic thermal diffusion. *J Mol Biol* 351, 345-54.
85. Rich, R. L., L. R. Hoth, K. F. Geoghegan, T. A. Brown, P. K. LeMotte, S. P. Simons, P. Hensley and D. G. Myszka. (2002) Kinetic analysis of estrogen receptor/ligand interactions. *PNAS* 99, 8562-8567.

86. Renaud, J. P., N. Rochel, M. Ruff, V. Vivat, P. Chambon, H. Gronemeyer and D. Moras. (1995) Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature* 378, 681-9.
87. Watkins, R. E., P. R. Davis-Searles, M. H. Lambert and M. R. Redinbo. (2003) Coactivator binding promotes the specific interaction between ligand and the pregnane X receptor. *J Mol Biol* 331, 815-28.
88. Sugita, Y. and Y. Okamoto. (1999) Replica-exchange molecular dynamics method for protein forlding. *Chemical Physics Letters* 314, 141-151.
89. Grater, F., S. M. Schwarzl, A. Dejaegere, S. Fischer and J. C. Smith. (2005) Protein/ligand binding free energies calculated with quantum mechanics/molecular mechanics. *J Phys Chem B* 109, 10474-83.
90. Hu, H. and W. Yang. (2008) Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods. *Annu Rev Phys Chem* 59, 573-601.
91. Amaro, R. E., R. Baron and J. A. McCammon. (2008) An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J Comput Aided Mol Des* 22, 693-705.
92. Totrov, M. and R. Abagyan. (2008) Flexible ligand docking to multiple receptor conformations: a practical alternative. *Curr Opin Struct Biol* 18, 178-84.