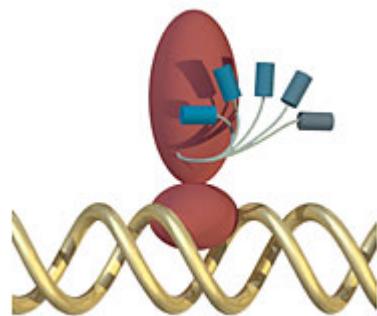


# NUCLEAR RECEPTORS

Homology, function and structure

*Structural Bioinformatics*



Ferran Briansó  
Elisenda Feliu  
Núria Queralt

# Contents

---

- PART I**
- Nuclear Receptors**
- PART II**
- DNA Binding Domain**
- PART III**
- Ligand Binding Domain**

- Introduction
- Scientific interest
- Biological context
- Primary structure
- Ligands
- Transcriptional activation functions
- Regulation and mode of action
- Evolution
- Former classification
- Receptors & Ligands
- Homology classification
- Research status

# Contents

---

**PART I**  
Nuclear Receptors

**PART II**  
**DNA Binding Domain**

**PART III**  
Ligand Binding Domain

- [DNA Binding Domain](#)
- [DBD Secondary structure](#)
- [DBD Tertiary structure](#)
- [DBD Conservation](#)
- [Characterization of the NRs](#)
- [Interactions DNA-DBD](#)
- [Conservation of the specific contacts](#)
- [Non-specific interactions](#)
- [Conservation of the non-specific contacts](#)
- [Structural conservation](#)
- [Dimerization](#)
- [Homodimerization](#)
- [Heterodimerization](#)
- [ClustalW alignment](#)

# Contents

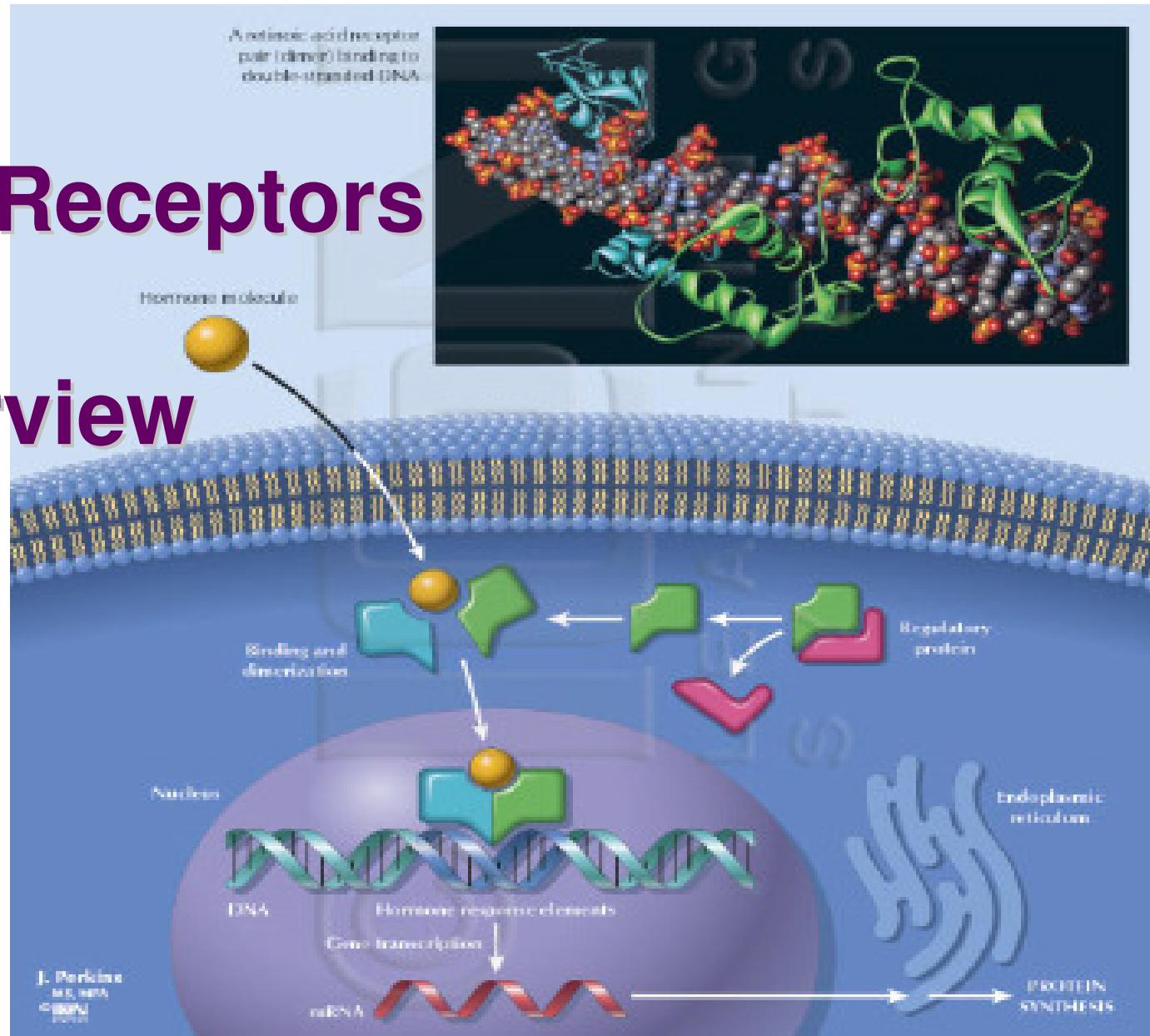
---

- PART I**  
Nuclear Receptors
- PART II**  
DNA Binding Domain
- PART III**  
**Ligand Binding Domain**

- Ligand Binding Domain
- Structural conservation through families
- LBD 3D structures
- ClustalW sequence alignment
- NR3A1 sequence and structural alignment
- Searching homologues with psi-blast
- Secondary structure
- Apo- Vs Holo-structures
- Holo-structures
- LBD helix 12 (H12)
- Agonist/antagonist-induced conformation
- Agonist-bond structure
- Interaction NR-coactivator
- Conservation of residues
- Antagonist-bond structure

# Nuclear Receptors

## Overview



# Introduction

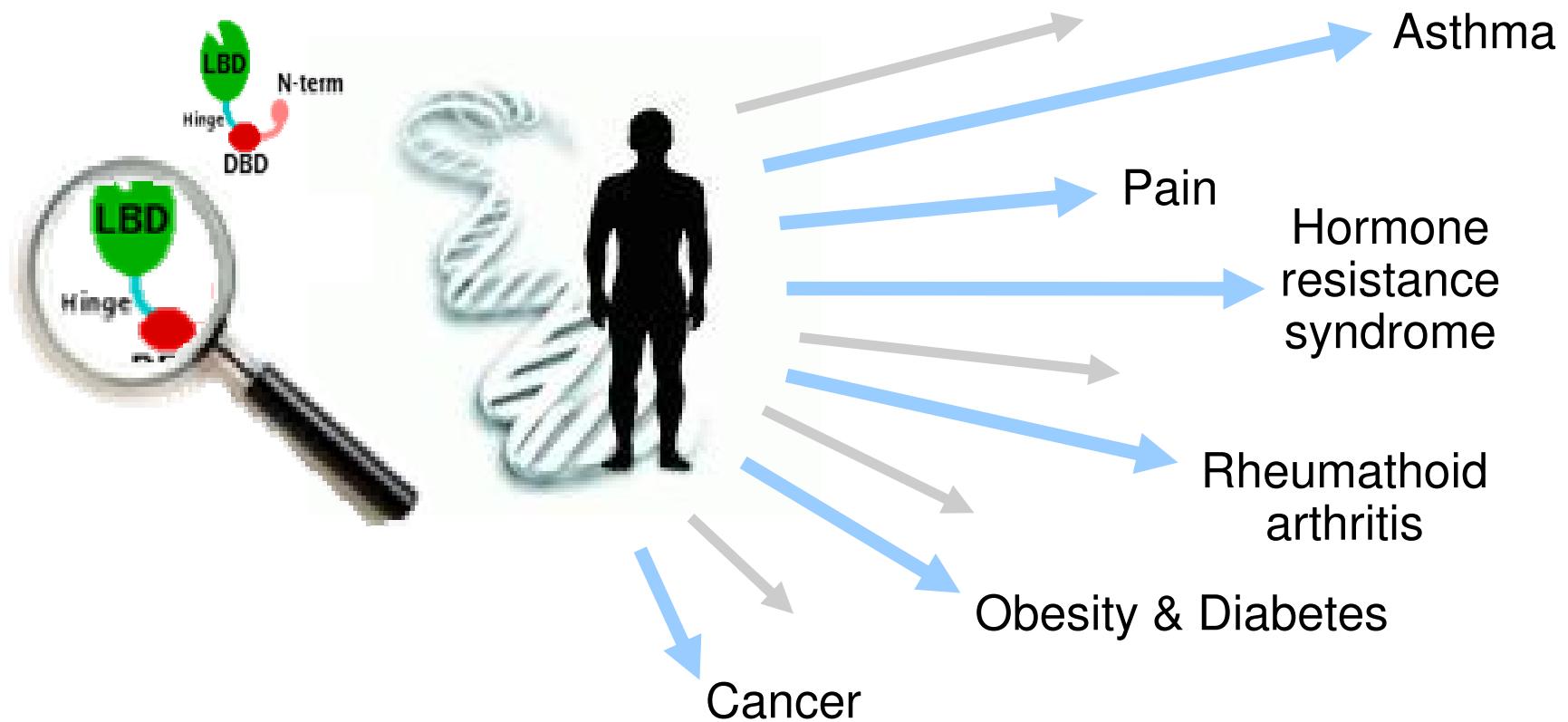
---

- Nuclear receptors (NRs) belong to a **large superfamily that are ligand activated intracellular transcription factors** which up or down regulate the expression of several genes.
- Nuclear receptors are **soluble proteins that can bind to specific DNA regulatory elements** (response elements or REs) and act as cell type- and promoter-specific regulators of transcription.
- In contrast to other transcription factors, the activity of nuclear receptors can be **modulated by binding to the corresponding ligands**, small lipophilic molecules that easily penetrate biological membranes.
- Nuclear receptors may be classified either according to **activation mechanism** (type I or II), or **sequence homology** (NR subfamilies 0-6).
- Nowadays, there are **more than 350 NR structures in the PDB**.



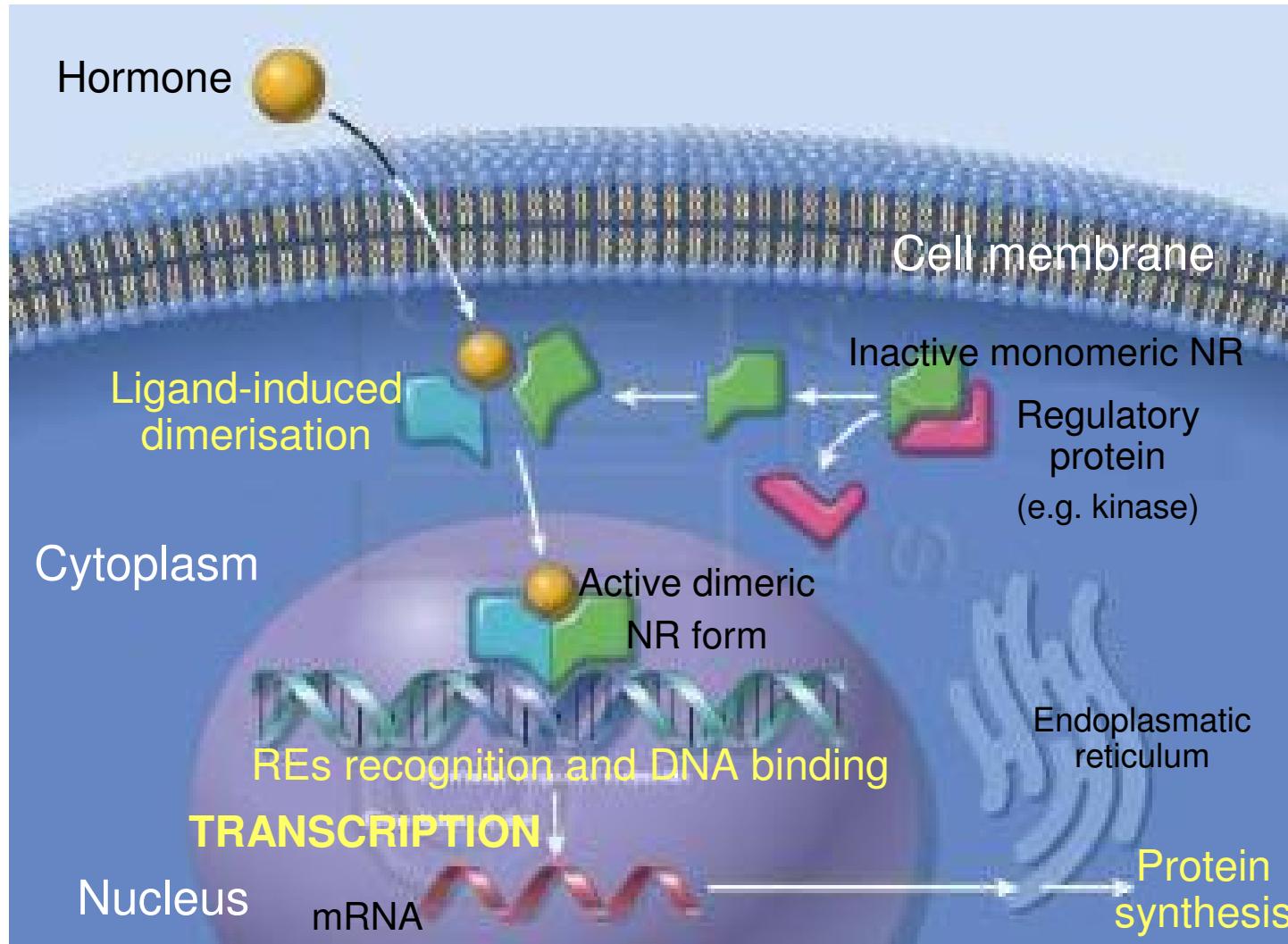
# Scientific interest

- Nuclear receptors are transcription factors involved in such important physiological functions as **control of embryonic development, organ physiology, cell differentiation** and **homeostasis**.
- Due to the role of nuclear receptors in gene expression control, members of this family are **suitable targets for new drug development**.



# Biological context

- Nuclear receptors are **key elements** for control of gene expression.

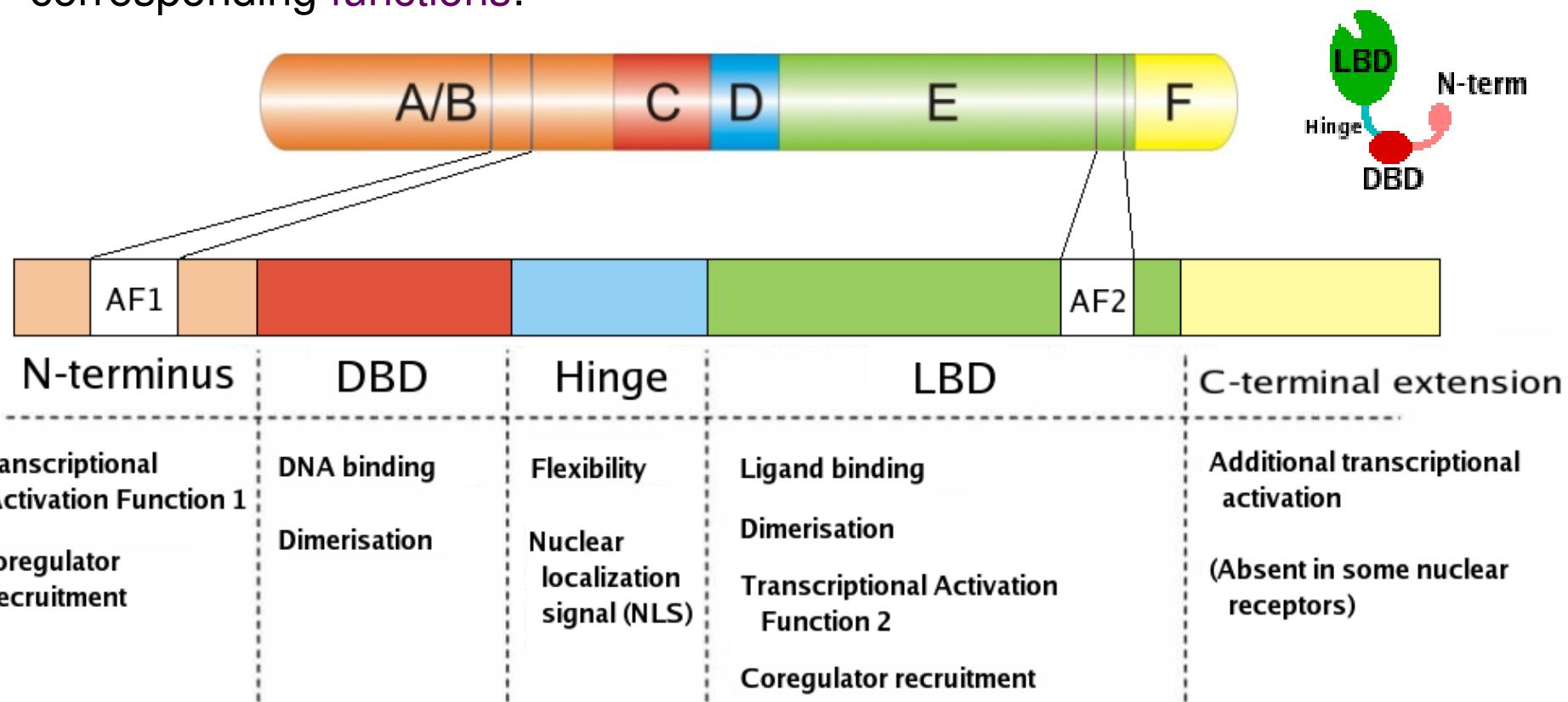


Homeostasis  
Cellular  
differentiation  
Physiology



# Primary structure

- A typical nuclear receptor contains the following **domains**, with corresponding **functions**:

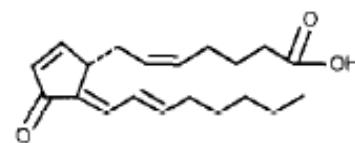
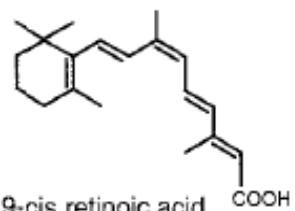
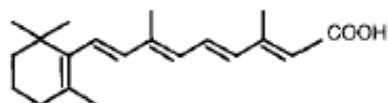
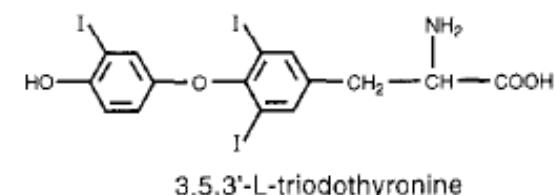
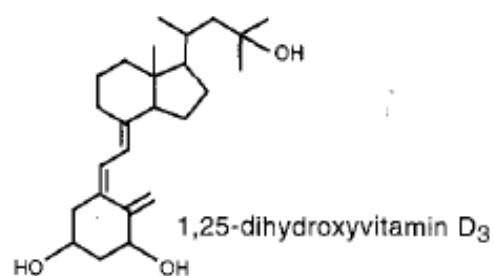
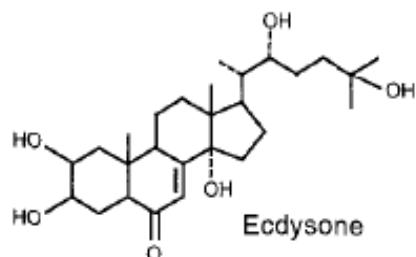
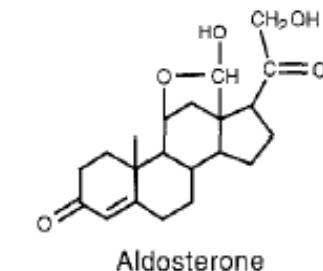
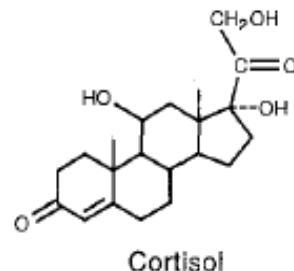
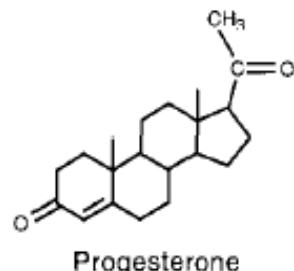
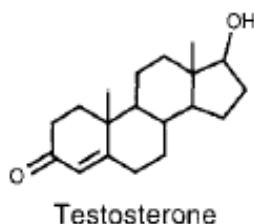
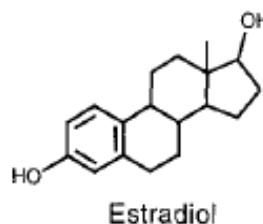


- DNA Binding Domain (DBD) and Ligand Binding Domain (LBD) are significant **conserved regions**, but DBD is the most one.



# Ligands

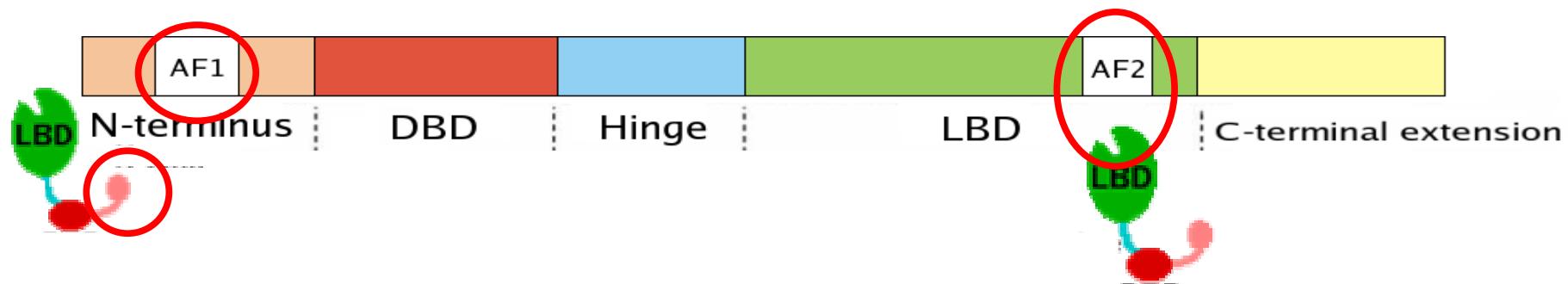
- **Lipophilic substances** such as endogenous hormones, vitamins A and D, drugs, and xenobiotic endocrine disruptors:



# Transcriptional activation functions

- **Activation Function-1, placed in the N-terminus region**

An important domain for the transcriptional activation of nuclear receptors is the ligand-independent activation function (AF-1), which generally resides in the N-terminal region of nuclear receptors. AF-1 **functions in a promoter-context and/or cell-type specific manner** and cooperates with AF-2 in the regulation of gene transcription.



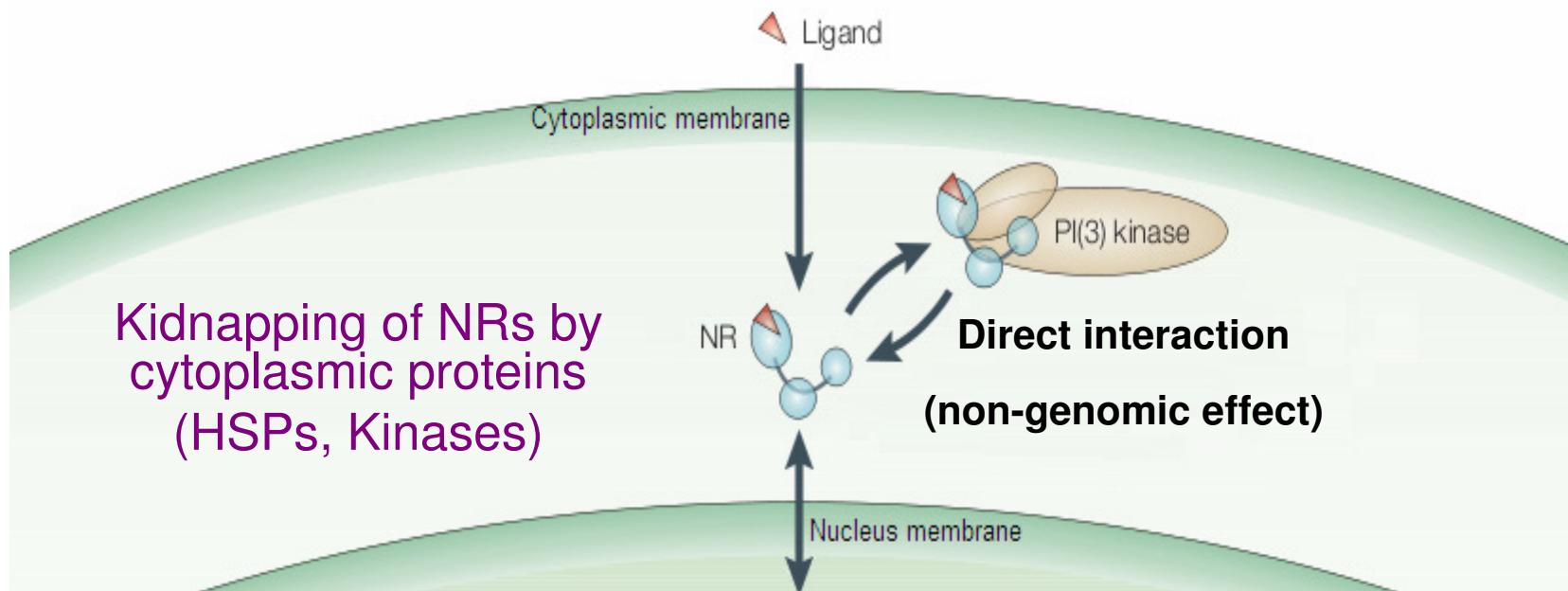
- **Activation Function-2, included in the Ligand Binding Domain**

The ligand-dependent activation function (AF-2) is the **key region for NR-ligand interaction**. AF-2 makes the function of agonist/antagonist ligand response, **changing the LDB conformation** and regulating the gene transcription, according to each case and with participation of other coregulatory elements.

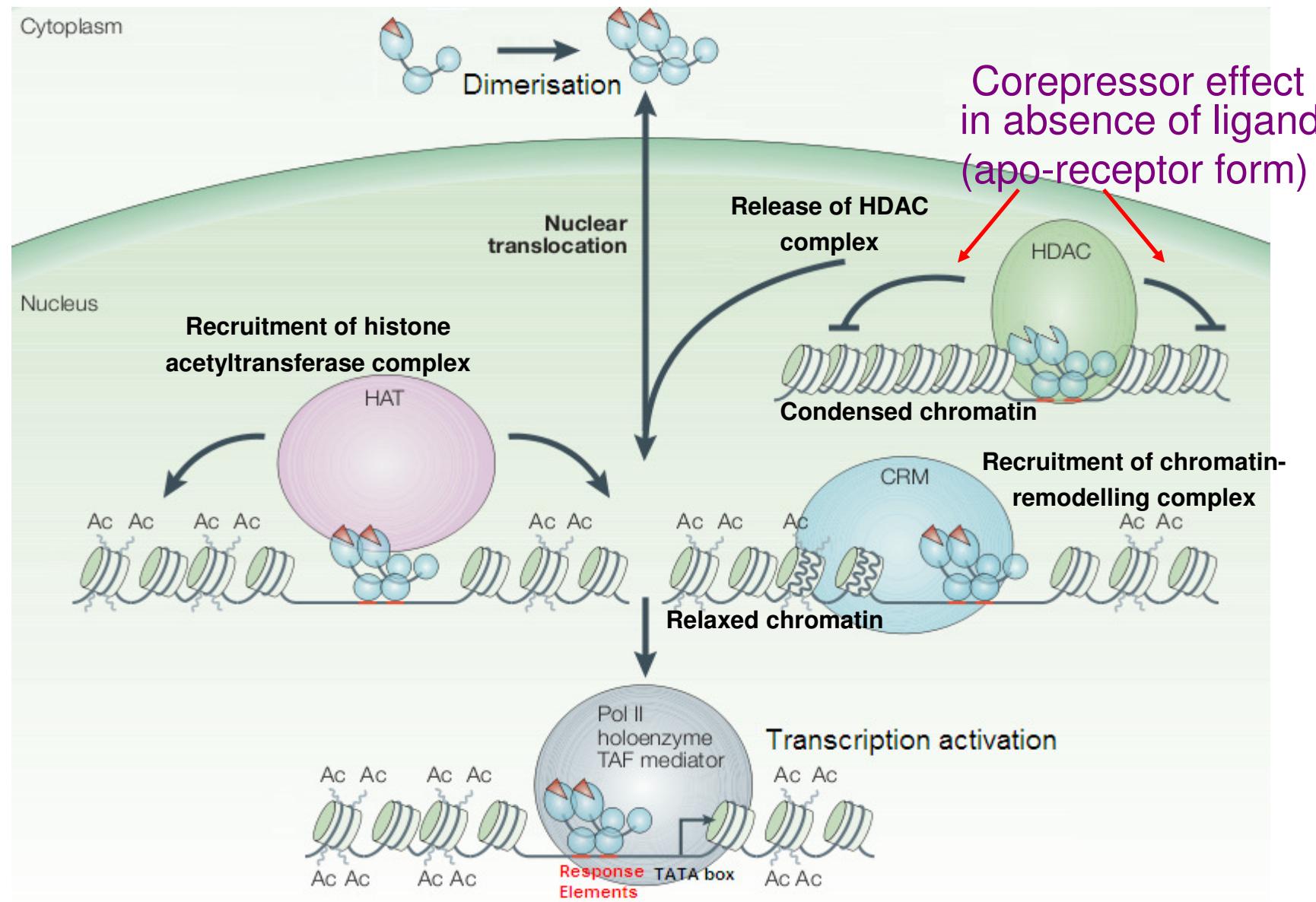


# Regulation and mode of action

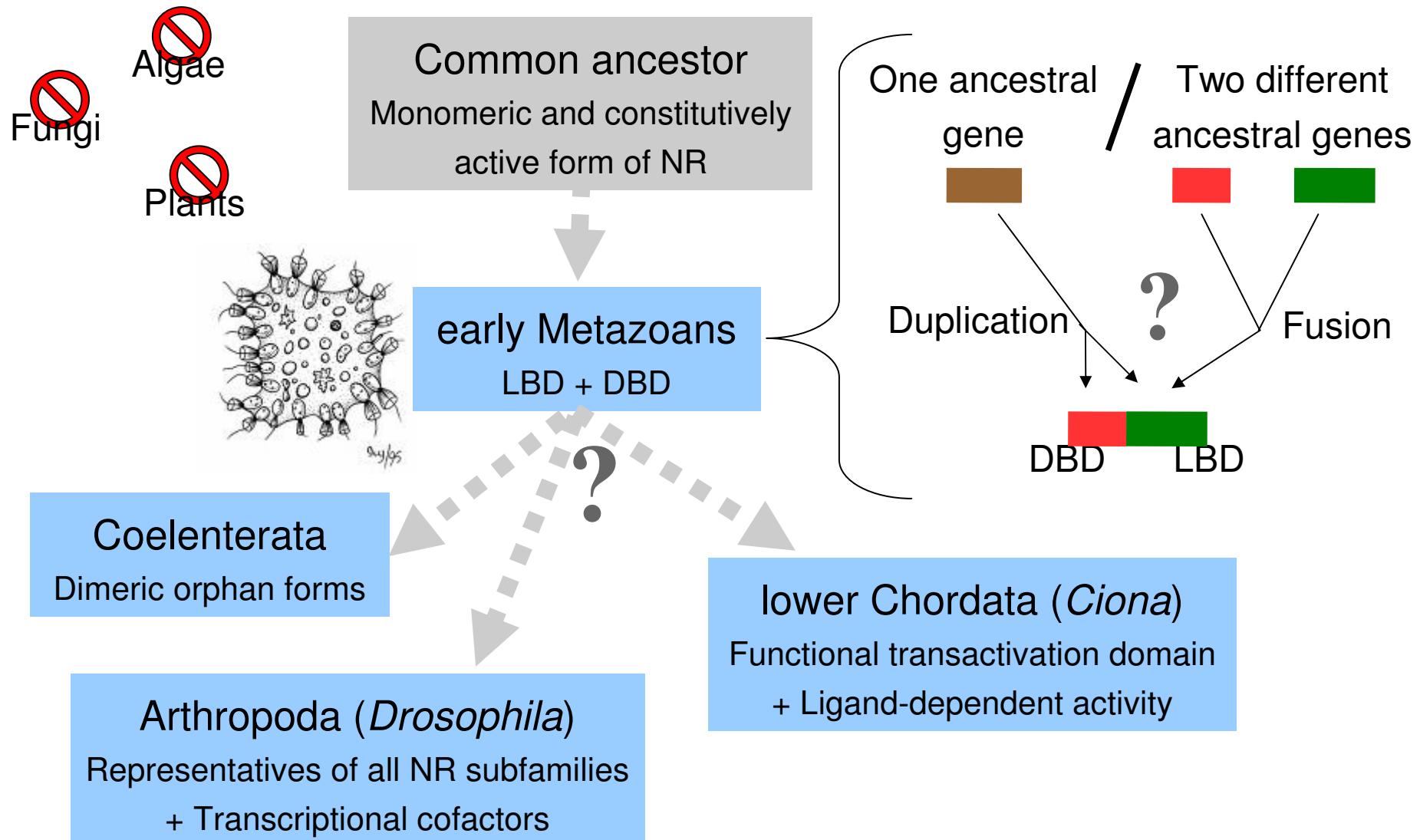
- Cytoplasmic regulation mechanism:



# Regulation and mode of action

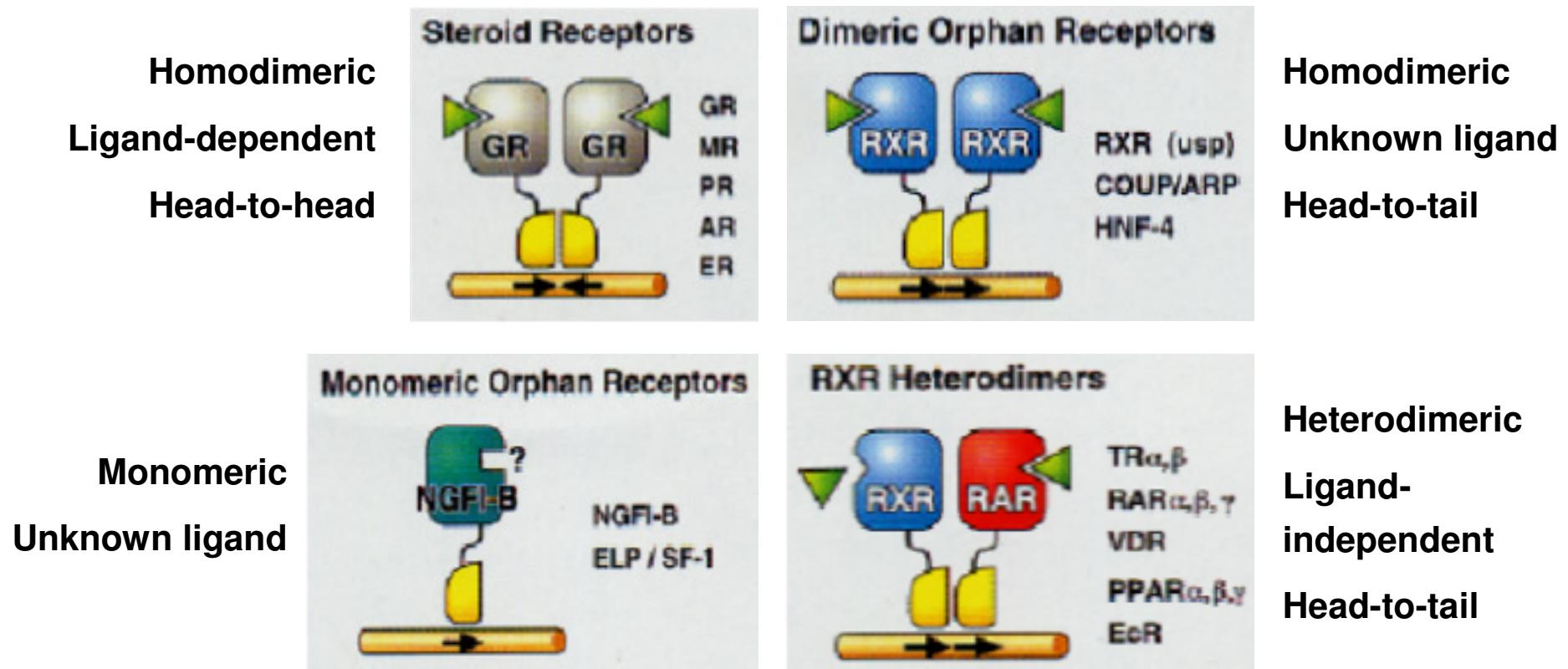


# Evolution



# Former classification

- The NR superfamily have been classified into **four subfamilies** based on their DNA-binding, ligand-binding and dimerisation properties:



# Receptors & Ligands

Name	Abbreviation	Nomenclature	Ligand
Thyroid hormone receptor	TR $\alpha$	NR1A1	Thyroid hormone
	TR $\beta$	NR1A2	Thyroid hormone
Retinoic acid receptor	RAR $\alpha$	NR1B1	Retinoic acid
	RAR $\beta$	NR1B2	Retinoic acid
	RAR $\gamma$	NR1B3	Retinoic acid
Peroxisome proliferator-activated receptor	PPAR $\alpha$	NR1C1	Fatty acids, leukotriene B4, fibrates
	PPAR $\beta$	NR1C2	Fatty acids
	PPAR $\gamma$	NR1C3	Fatty acids, prostaglandin J2,
Reverse erbA	Rev-erb $\alpha$	NR1D1	Orphan
	Rev-erb $\beta$	NR1D1	Orphan
RAR-related orphan receptor	ROR $\alpha$	NR1F1	Cholesterol, cholesteryl sulphate
	ROR $\beta$	NR1F2	Retinoic acid
	ROR $\gamma$	NR1F3	Retinoic acid
Liver X receptor	LXR $\alpha$	NR1H3	Oxysterols, T0901317, GW3965
	LXR $\beta$	NR1H2	Oxysterols, T0901317, GW3965
Farnesoid X receptor	FXR $\alpha$	NR1H4	Bile acids, Fexaramine
	FXR $\beta$ *	NR1H5	Lanosterol
Vitamin D receptor	VDR	NR1I1	1,25-dihydroxy vitamin D <sub>3</sub> , lithocholic acid
Pregnane X receptor	PXR	NR1I2	Xenobiotics, PCN
Constitutive androstane receptor	CAR	NR1I3	Xenobiotics, phenobarbital
Human nuclear factor 4	HNF4 $\alpha$	NR2A1	Orphan
	HNF4 $\gamma$	NR2A2	Orphan
Retinoid X receptor	RXR $\alpha$	NR2B1	Retinoic acid
	RXR $\beta$	NR2B2	Retinoic acid
	RXR $\gamma$	NR2B3	Retinoic acid
Testis receptor	TR2	NR2C1	Orphan
	TR4	NR2C2	Orphan
Tailless	TLL	NR2E2	Orphan
Photoreceptor-specific nuclear receptor	PNR	NR2E3	Orphan
Chicken ovalbumin upstream promoter-transcription factor	COUP-TFI	NR2F1	Orphan
	COUP-TFII	NR2F2	Orphan



# Receptors & Ligands

Name	Abbreviation	Nomenclature	Ligand
ErbA2-related gene-2	EAR2	NR2F6	Orphan
Oestrogen receptor	ER $\alpha$	NR3A1	Oestradiol-17 $\beta$ , tamoxifen, raloxifene
	ER $\beta$	NR3A2	Oestradiol-17 $\beta$ , various synthetic compounds
Oestrogen receptor-related receptor	ERR $\alpha$ ERR $\beta$ ERR $\gamma$	NR3B1 NR3B2 NR3B3	Orphan DES, 4-OH tamoxifen DES, 4-OH tamoxifen
Glucocorticoid receptor	GR	NR3C1	Cortisol, dexamethasone, RU486
Mineralocorticoid receptor	MR	NR3C2	Aldosterone, spiro lactone
Progesterone receptor	PR	NR3C3	Progesterone, medroxyprogesterone acetate, RU486
Androgen receptor	AR	NR3C4	Testosterone, flutamide
NGF-induced factor B	NGFIB	NR4A1	Orphan
Nur related factor 1	NURR1	NR4A2	Orphan
Neuron-derived orphan receptor 1	NOR1	NR4A3	Orphan
Steroidogenic factor 1	SF1	NR5A1	Orphan
Liver receptor homologous protein 1	LRH1	NR5A2	Orphan
Germ cell nuclear factor	GCNF	NR6A1	Orphan
DSS-AHC critical region on the chromosome, gene 1	DAX1	NR0B1	Orphan
Short heterodimeric partner	SHP	NR0B2	Orphan

Gronemeyer, Gustafsson & Laudet; 2004



# Homology classification

---

***Nuclear Receptor Nomenclature Committee*** current classification:

- Subfamily 1: **Thyroid Hormone Receptor-like**
  - Group A: Thyroid hormone receptor (Thyroid hormone)
  - Group B: Retinoic acid receptor (Vitamin A and related compounds)
  - Group C: Peroxisome proliferator-activated receptor
  - Group D: Rev-erb
  - Group F: Retinoid-related orphan receptor
  - Group H: Liver X receptor-like
  - Group I: Vitamin D receptor-like
- Subfamily 2: **Retinoid X Receptor-like**
  - Group A: Hepatocyte nuclear factor-4 (HNF4)
  - Group B: Retinoid X receptor (RXR $\alpha$ )
  - Group C: Testicular receptor
  - Group E: TLX/PNR
  - Group F: COUP/EAR



# Homology classification

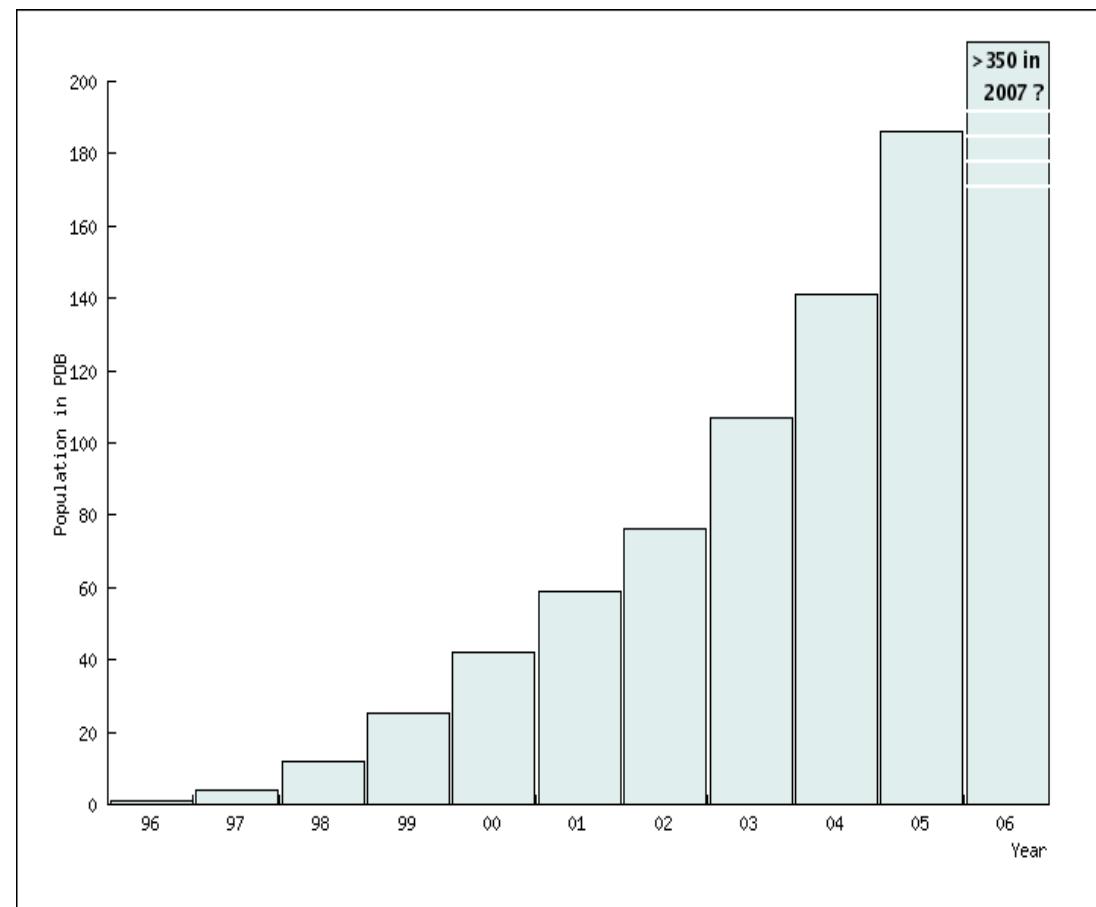
- Subfamily 3: **Estrogen Receptor-like** (Steroid hormone receptor)
  - Group A: Estrogen receptor (Sex hormone receptors)
  - Group B: Estrogen related receptor
  - Group C: 3-Ketosteroid receptors
- Subfamily 4: **Nerve Growth Factor IB-like**
  - Group A: NGFIB/NURR1/NOR1
- Subfamily 5: **Steroidogenic Factor-like**
  - Group A: SF1/LRH1
- Subfamily 6: **Germ Cell Nuclear Factor-like**
  - Group A: GCN1
- Subfamily 0: **Miscellaneous**
  - Group B: DAX/SHP



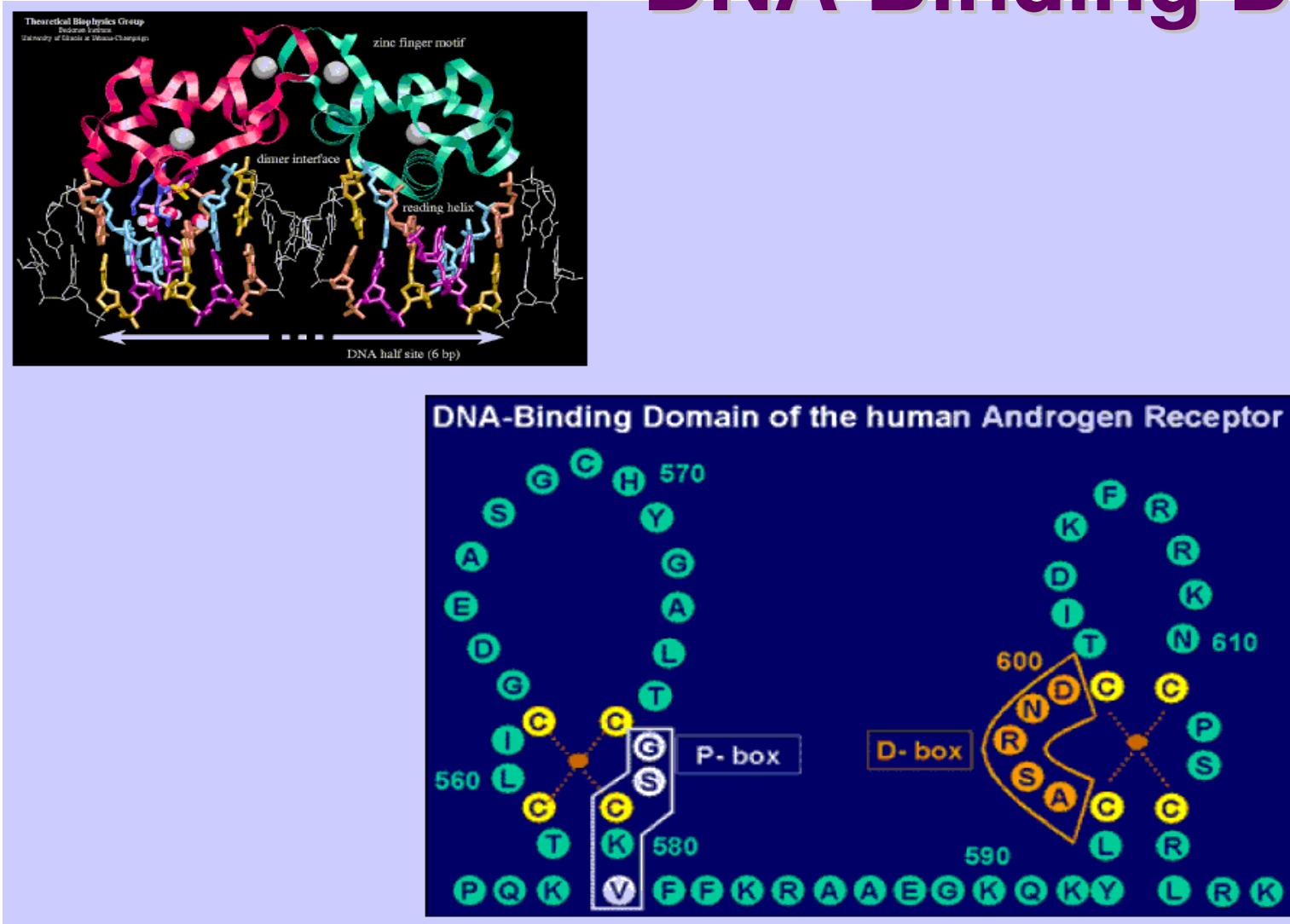
# Research status

- Since isolation and cloning of the first nuclear receptor in 1985, a large number of NRs have been identified. But only a part of the current subfamilies are well represented in the Protein Data Bank.
- Structures are obtained from the LDB or the DBD, but not from the full protein.
- Many new NRs are temporarily classified as Orphan receptors.

Adapted from a graphic of annual growth for Nuclear Receptor entries in the PDB, reported by FCP web page  
<http://cgl.imim.es/fcp/>

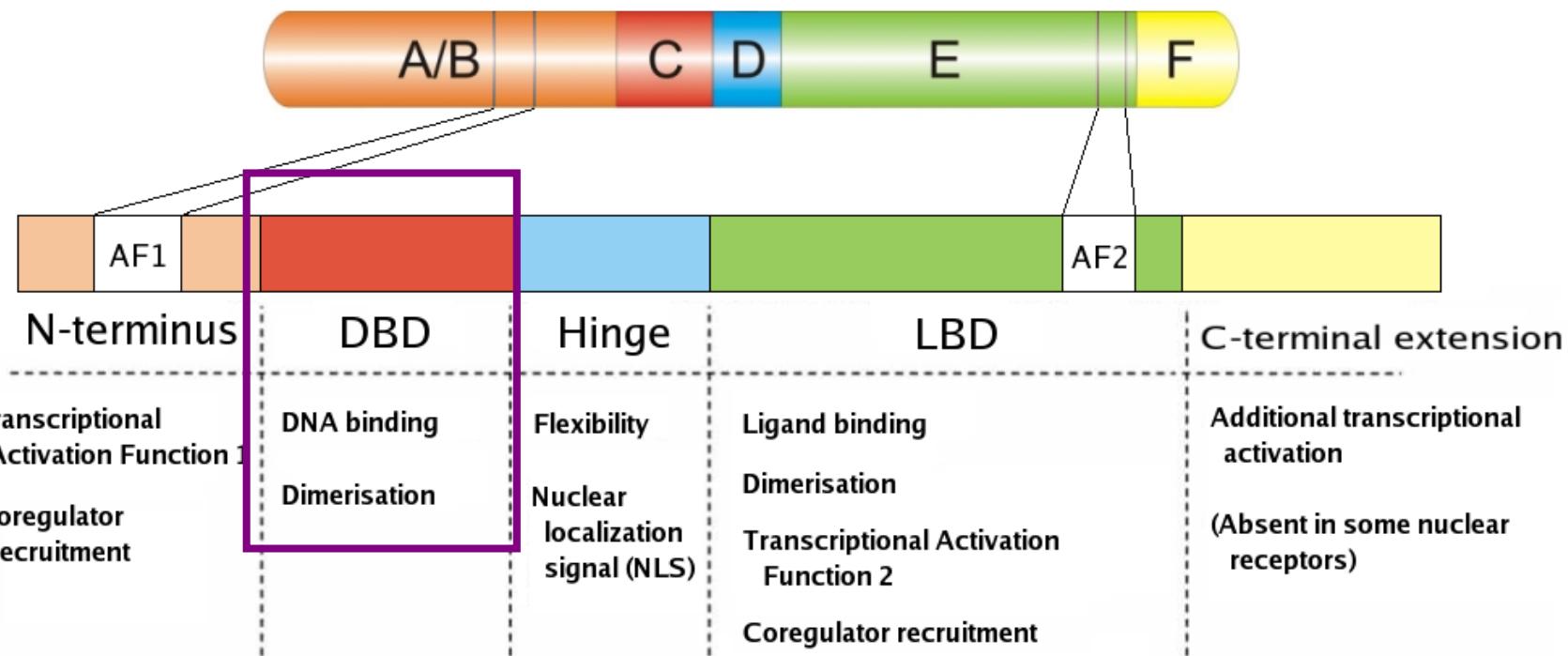


# DNA Binding Domain



# DNA Binding Domain

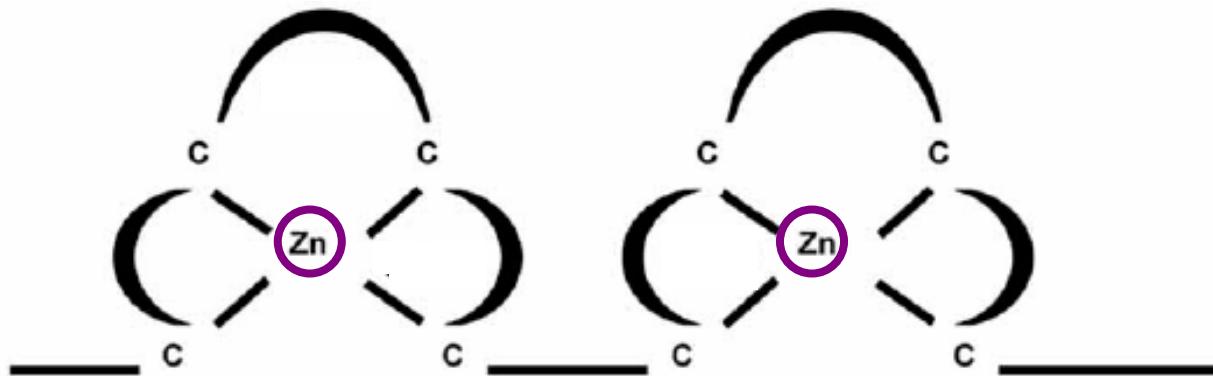
In this part we center ourselves in the structure and sequence of the DNA binding domain.



# DNA Binding Domain

---

- The DNA Binding Domain (DBD) is a highly conserved domain in the family of Nuclear Receptors.
- The DBD consists of about 70 residues that bind to activating elements of DNA called hormone response elements.
- In the DBD there are two zinc containing regions. Each region binds a zinc atom through four cysteine residues:



# DBD - Secondary Structure

- For the Glucocorticoid Receptor, the **secondary structure** of the DNA binding domain is as follows:

**DNA BINDING DOMAIN**  
**Secondary Structure of the**  
**Glucocorticoid Receptor**

RPCLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQHNYLCAGR  
--STTT-S---EEETTEEE-**HHHHHHHHHH**TS-----SSSS

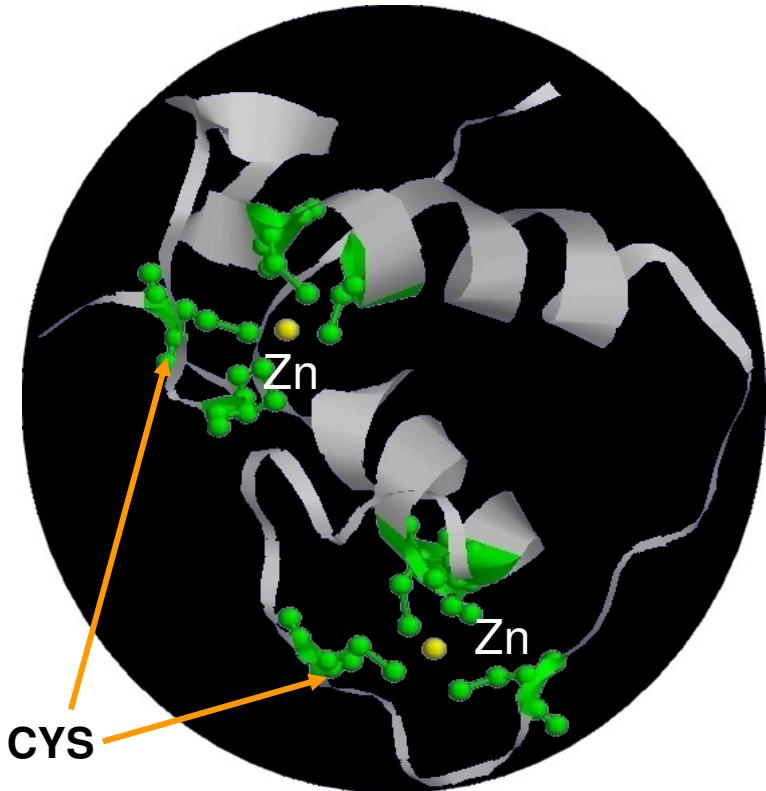
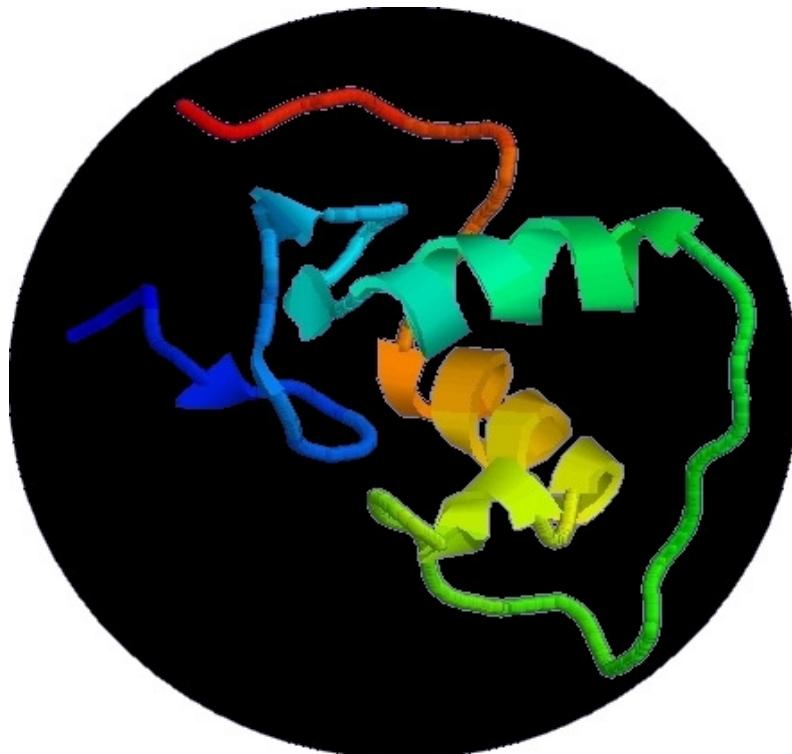
DCIIDKIRRKNCPACRYRKCLQAGMNLEAR  
-----TTTTT-**HHHHHHHHHH**S-----

- The DNA binding domain of all nuclear receptors contains **two**  $\alpha$ -helices.
- For each zinc motif, the second pair of cysteine zinc ligands, initiates an  $\alpha$ -helix.



# DBD - Tertiary Structure

The two zinc motives are interwoven into a **single globular domain**, with extensive interactions between them.



The hydrophobic sides of the two  $\alpha$ -helices pack against each other to form a compact **core with a hydrophobic interior**.



# DBD - Conservation

---

The DNA binding domain is **highly conserved** in the family of nuclear receptors.

Superposition of nuclear receptors with known structure:

Subfamily 3:

- Estrogen receptor
- Estrogen receptor  $\beta$
- Glucocorticoid receptor

Subfamily 2:

- Retinoid X receptor

Subfamily 1:

- Retinoid acid receptor
- Thyroid hormone receptor
- Vitamin D receptor



# DBD - Conservation

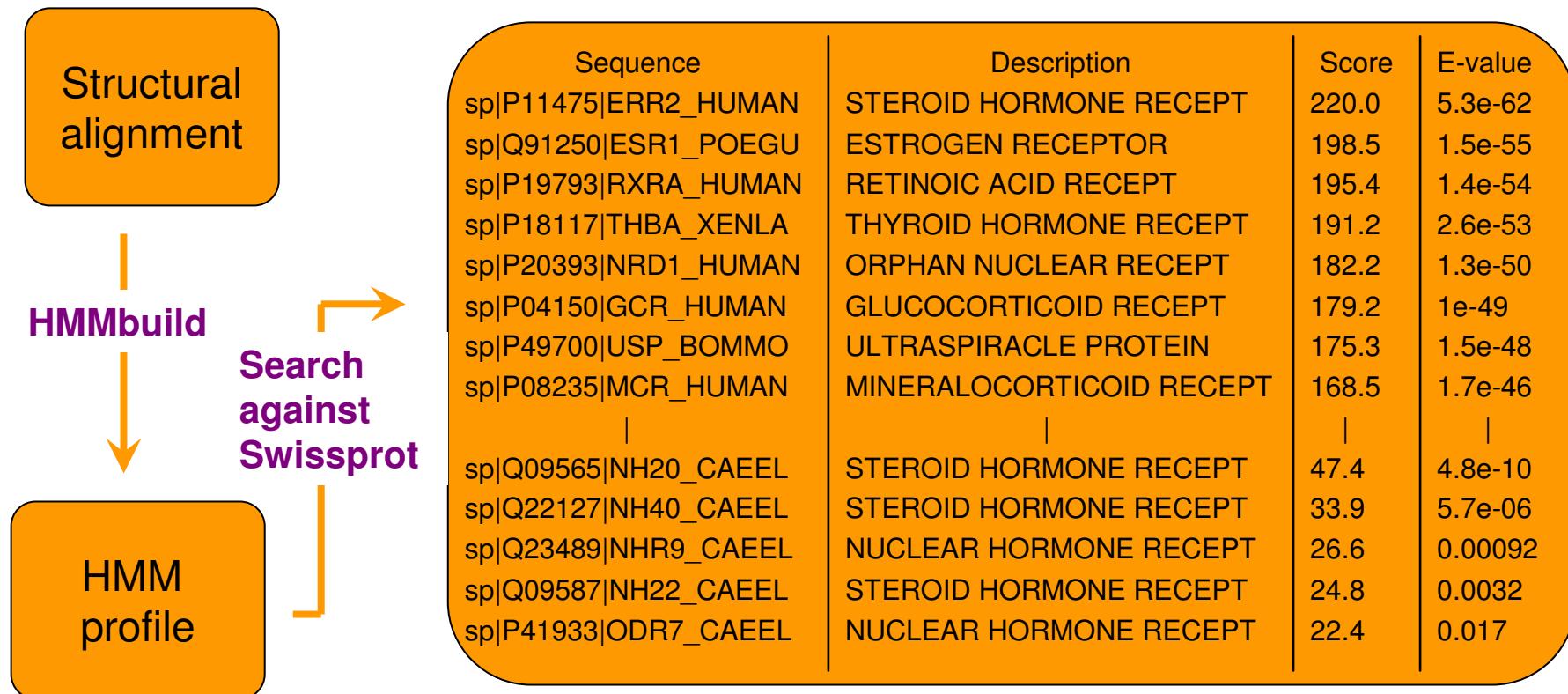
A **clustal multiple alignment** with the same sequences gives the following.  
Observe the conservation of the **cysteine residues** involved in the zinc motives.

CLUSTAL W multiple sequence alignment	
DNA BINDING DOMAIN	
ERbeta	AIPKRL <b>CLVCGDIASGYHYGVASCEACKAFFKRTI</b> QGN--IEYS
ER	MKETRY <b>CAVCNDYASGYHYGVWSCEGCKAFFKRSI</b> QGH--NDYM
GR	MKPAPR <b>CLVC</b> CSDEASGCHYGVL <b>TCGSCKVFFKRAVEGQ</b> --HNYL
RAR	---- <b>PCFVCQDKSSGYHYGVSA</b> CEG <b>CKGFFRRSI</b> QKN--MVYT
RXR	-FTKHI <b>CAICGDRSSGKHYGVYSC</b> EG <b>CKGFFKRTVRKD</b> --LTYT
TR	---DEL <b>CVVCGDKATGYHYRCIT</b> CEG <b>CKGFFRRTI</b> QKNLHPSYS
VDR	---LL <b>CKVCGDVASGFHYGVL</b> ACEG <b>CKGFFRRSI</b> QQN-IQYKR * . * * . * * . * * * . * .
ERbeta	<b>CPATNECEITKRRRKSCQACRFMKALKVGMLKEGVRLDRVRGGR</b>
ER	<b>CPATNQC</b> TIDKNRRKSCQACRLRKYEVGMMKG-----
GR	<b>CAGRND</b> CIIDKIRRKN <b>CPACRYRKCLQAGMNLEARKTKK</b> -----
RAR	<b>CHRDKN</b> CIIINKVTRNRCQY <b>CRYRKLQKCFEVGMSKESVRND</b> -----
RXR	<b>CRDNKDC</b> LIDKRQRNRCQY <b>CRYQKALAMGMKREAVQEERQG</b> --
TR	<b>CKYEGKC</b> VIDKVTRNQCQ <b>E</b> CRFKKCIVYGMATDLVLD <span style="font-size: 0.8em;">SKRLAK</span>
VDR	<b>CLKNEN</b> CIVRINRNRCQQ <b>CRF</b> KKCLSVGMSRDAVRFGR----- * * . * * . * * . * .



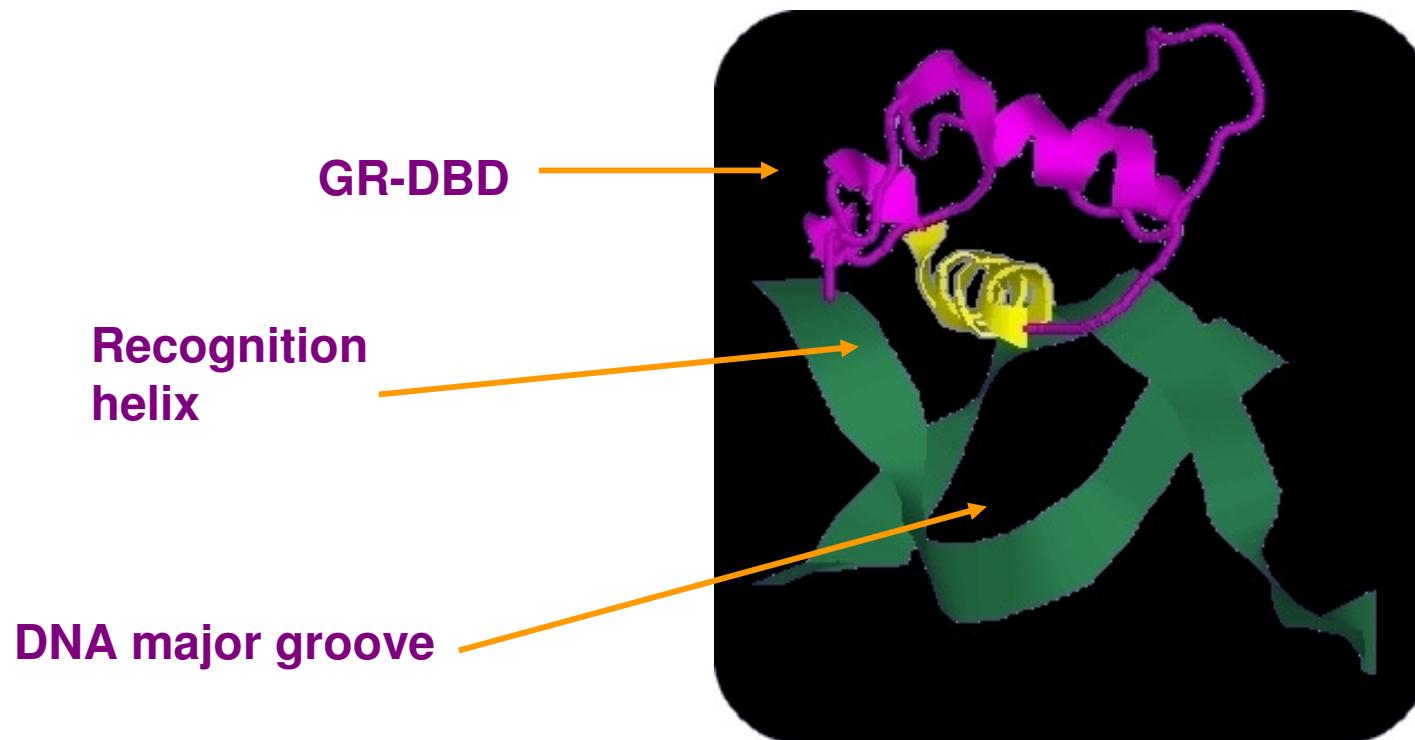
# Characterization of the NRs

The DNA binding domain characterizes the family of nuclear receptors.



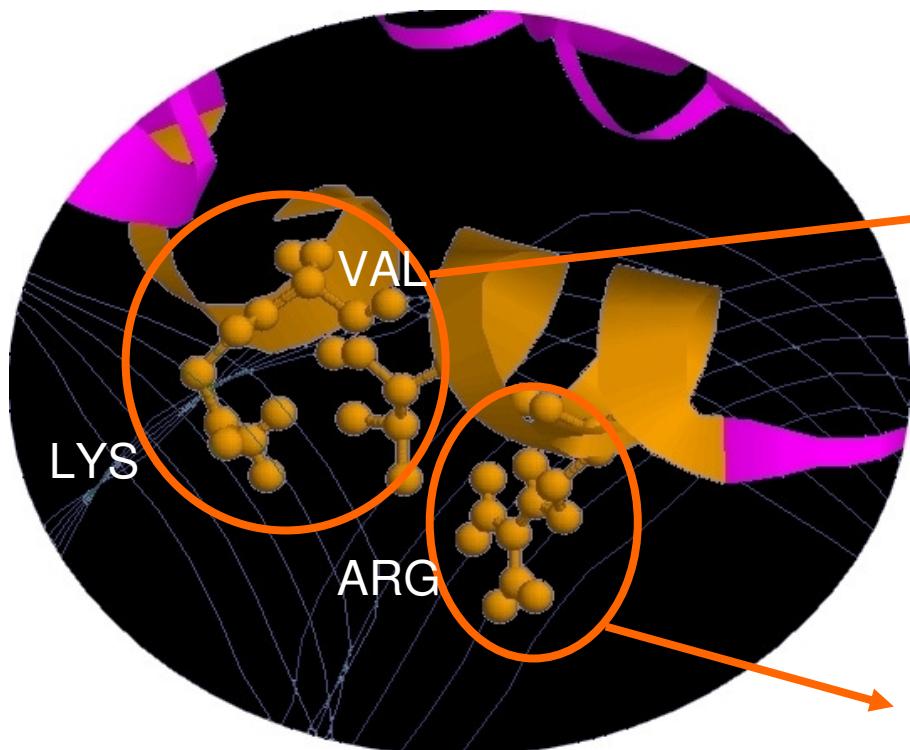
# Interactions DNA - DBD

- The first  $\alpha$ -helix in the zinc motif forms **sequence-specific** interactions with the edge of the bases in the major groove of one DNA strand.
- This helix is called the **recognition helix**.

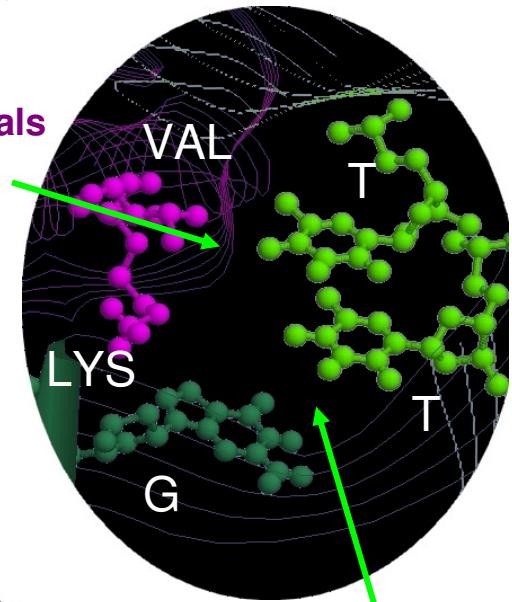


# Interactions DNA - DBD

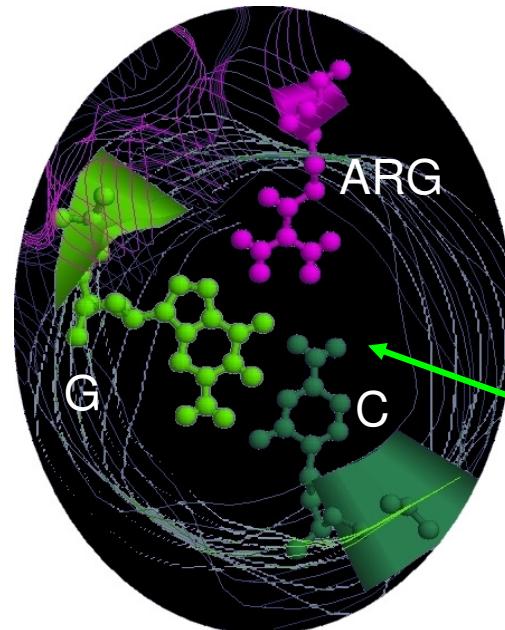
In the glucocorticoid receptor, the residues LYS 461, VAL 462 and ARG 466 form the specific interactions with DNA.



Van der Waals contact



2 hydrogen bonds



2 hydrogen bonds



# Conservation of the Specific contacts

- The LYS and ARG are conserved in the nuclear receptor family:

CLUSTAL W multiple sequence alignment  
DNA BINDING DOMAIN

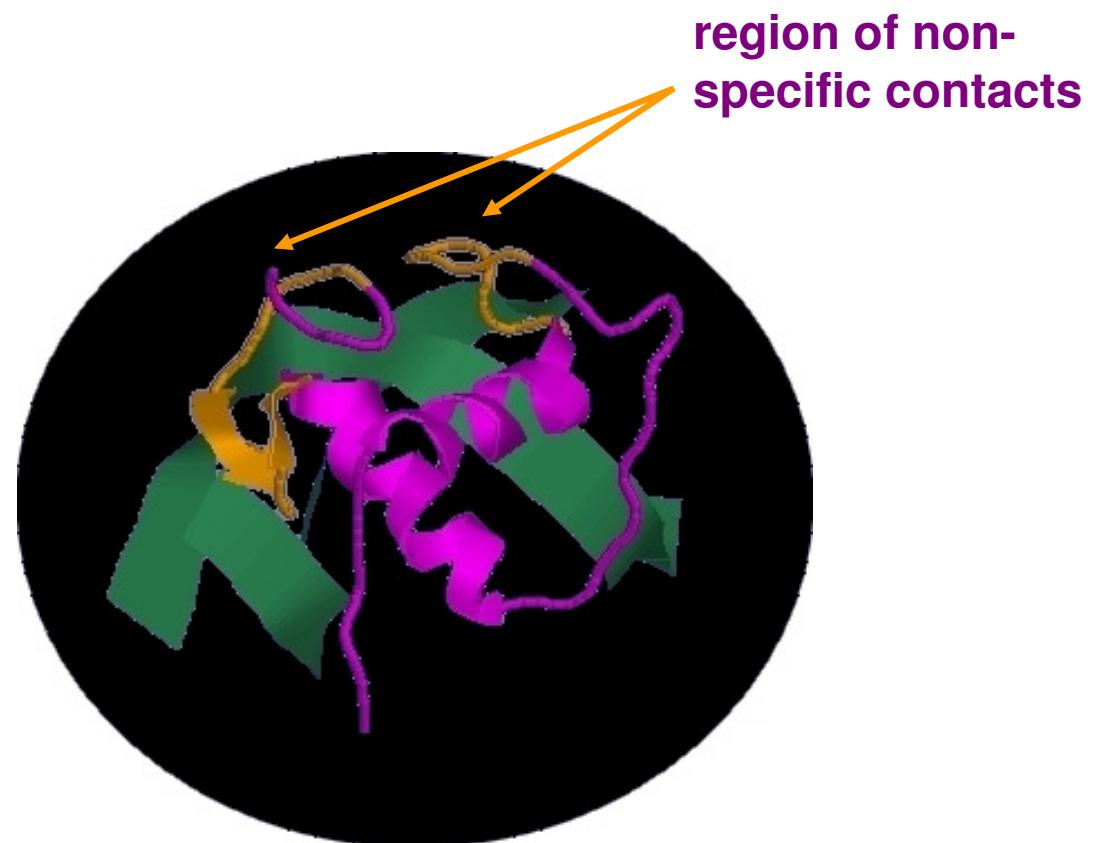
ERbeta	AIPKRLCLVCGDIASGYHYGVASCEACKAFFKRTIQGN--IEYS
ER	MKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGH--NDYM
GR	MKPAPRCLVCSDEASGCHYGVLTGSCKVFFKRAVEGQ--HNYL
RAR	-----PCFVCQDKSSGYHYGVSACEGCKGFFRRSIQKN--MVYT
RXR	-FTKHICAI CGDRSSGKHYGVYSC EGCKGFFKRTVRKD--LTYT
TR	---DELCVVCGDKATGYHYRCITCEGCKGFFRTIQKNLHPSYS
VDR	----LLCKVCGDVA SGFHYGVLACEGCKGFFRRSIQQN-IQYKR
	* . * * . * ** . * * * * . *
ERbeta	CPATNECEITKRRRKSCQACRFMKALKVGMLKEGVRLDRVRGGR
ER	CPATNQCTIDKNRRKSCQACRLRKCYEVGMMKG-----
GR	CAGRNDCCIIDKIRRKNCPACRYRKCLQAGMNLEARKTKK-----
RAR	CHRDKNCIINKVTRNRCQYCRYQKALAMGMKREAVQEERORG-----
RXR	CRDNKDCLIDKRQRNRCQYCRYQKALAMGMKREAVQEERORG-----
TR	CKYEGKCVIDKVTRNQCQE CRFKKCIYVGMA TDVLV LDDSKRLAK
VDR	CLKNENCSIVRINRNRCQQCRF KKCLSVGMSRDAVRFGR-----
	* * * . * * * * . *



# Non-specific Interactions

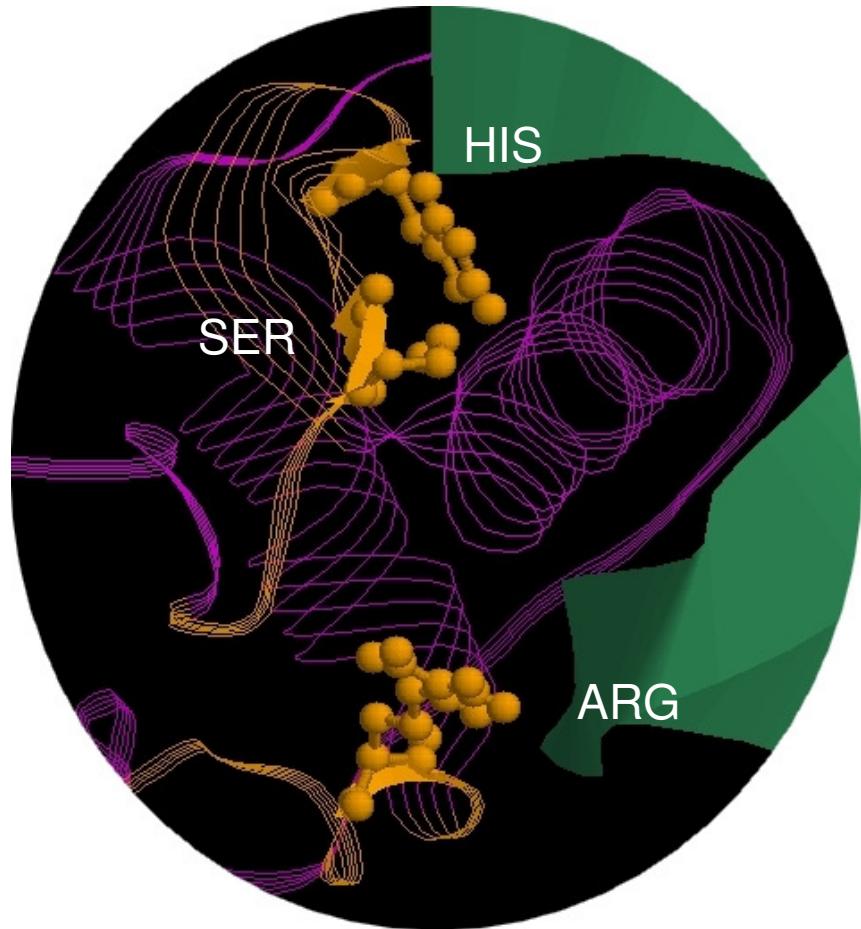
- The recognition helix is positioned and oriented in the **major groove** by a number of non-specific interactions between the phosphate groups and protein side chains.

- These contacts are made mainly by residues from the two loop regions between the second and third cysteine zinc ligands in both zinc motives.



# Non-specific Interactions

In the glucocorticoid receptor, the residues SER 448, HIS 451 and ARG 489 form hydrogen bonds to **phosphates** in the DNA backbone.



# Conservation of the Non-specific contacts

- The residues HIS 451 and ARG 489 of the glucocorticoid receptor are conserved in the nuclear receptor family.
- In the position 448, always occurs SER or THR.

## CLUSTAL W multiple sequence alignment DNA BINDING DOMAIN

ERbeta	AIPKRLCLVCGDIA	SGYHYGVASCEACKAFFKRTI	QGN--IEYS
ER	MKETRYCAVCNDYA	SGYHYGVWSCEGCKAFFKRS	IQGH--NDYM
GR	MKPAPRCLVCSDDEA	SGCHYGVLTCGSCKVFFKRAVEGQ	--HNYL
RAR	-----PCFVCQDKS	SGYHYGVSACEGCKGFFRRS	IQKN--MVYT
RXR	-FTKHICAI CGDRS	SGKHYGVY SCEGCKGFFKRTVRKD	--LTYT
TR	---DELCVVCGDKA	TGYHYRCITCEGCKGFFRRTI	QKNLHPSYS
VDR	----LLCKVCGDVA	SGFHYGVLACEGCKGFFRRS	IQQN-IQYKR
		*	.* * . * * . * * * * . . .
ERbeta	CPATNECEITKRRRKSCQACRFM	KALKVGMLKEGVRLDRVGG	R
ER	CPATNQCTIDKNRRKSCQACRLR	KCYEVGMMKG	-----
GR	CAGRNDIIDKIRRKNCPACRYR	KCLQAGMNLEARKK	-----
RAR	CHRDKNCIINKVTRNRCQYCR	LQKCFEVGMSKESVRND	-----
RXR	CRDNKDCLIDKRQRNRCQYCR	YQKALAMGMKREAQEEERQ	-----
TR	KYEGKCVIDKVTRNQCQE	CRFKKCIYVG	MATDLVLD
VDR	CLKNENCSIVRINRNRCQQCR	FKKCLSVGMSRDAVRFGR	-----
	*	.* * . * * * . *	**



# Structural conservation

Conservation of the zinc finger cysteines, the specific contact residues and the non-specific contacts in a structural alignment:

STAMP - STRUCTURAL SUPERPOSITION - DNA BINDING DOMAIN

ER	-MKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRS--IQGHN-DYM
ERbeta	A-IPKRLCLVCGDIAASGYHYGVASCEACKAFFKRTIQQ--NI-EYS
RXR	-F-TKHICAI CGDRSSGKHYGVYSC EGCKGFFKRTVRK-D-L-TYT
GR	-MKPARPCLVCSDAESGCHYGVLT CGSCKVFFKRAVE-G-QH-NYL
VDR	-----LLCKVCGDVA SGFHYGVLA CE GCKGFFRSIQQ-N-IQYKR
RAR	-----PCFVCQDKSSGYHYGVSA CE GCKGFFRSI QKN-M-V-YT
TR	-----DELCVVCGDKATGYHYRCI TCE GCKGFFRSI QKNLHPS-YS
ER	CPATNQCTI--DKNRRKSCQACRLRKYEVGMM-KG-----
ERbeta	CPATNECEITKRRR--KSCQACRFMKALKVGMLKE-G-V-RLDRVR
RXR	CRDNKDCLIDKRQR--NRCQYCRYQKALAMGMKREAVQEER-Q---
GR	CAGRNDCIIDKIRR--KNCPACRYRKCLQAGMNLE--A-R-KT---
VDR	CLKNENCSIVRINR--NRCQQCRFKKCLSVGMSRDA-V-R-F---
RAR	CHRDKNCIINKVTR--NRCQYCRQLQKCFEVGMSKES-V-R-N---
TR	CKYEGKCVIDKVTR--NQCQE CRFKKCI YVGMATDL-V-L-D---

Alignment score Sc = 7.250913

Alignment length Lp = 82

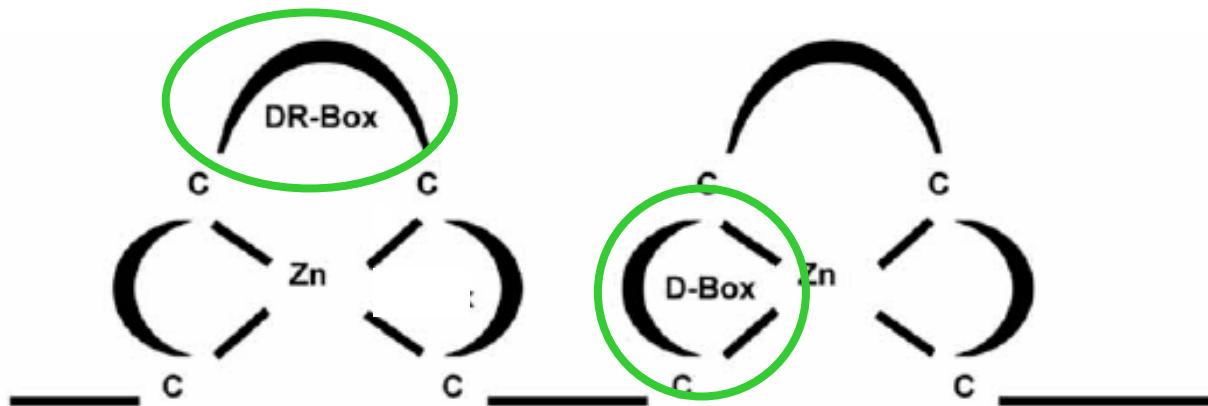
RMS deviation after fitting on 58 atoms = 1.446357



# Dimerization

---

- Recall that the nuclear receptors are active when they form **dimers**.
- In the DNA binding domain there are two dimerization sequences, the **D-box**, (the five residues between the two first cysteine zinc ligands of the second zinc motif) and the **DR-box** (residues between the second and third cysteine of the first zinc motif):



- The three dimensional structure of the DNA binding domain changes after dimerization. That is, in the monomer, the D-box is not well defined, but it is a well-defined  $\beta$  turn in the dimer.



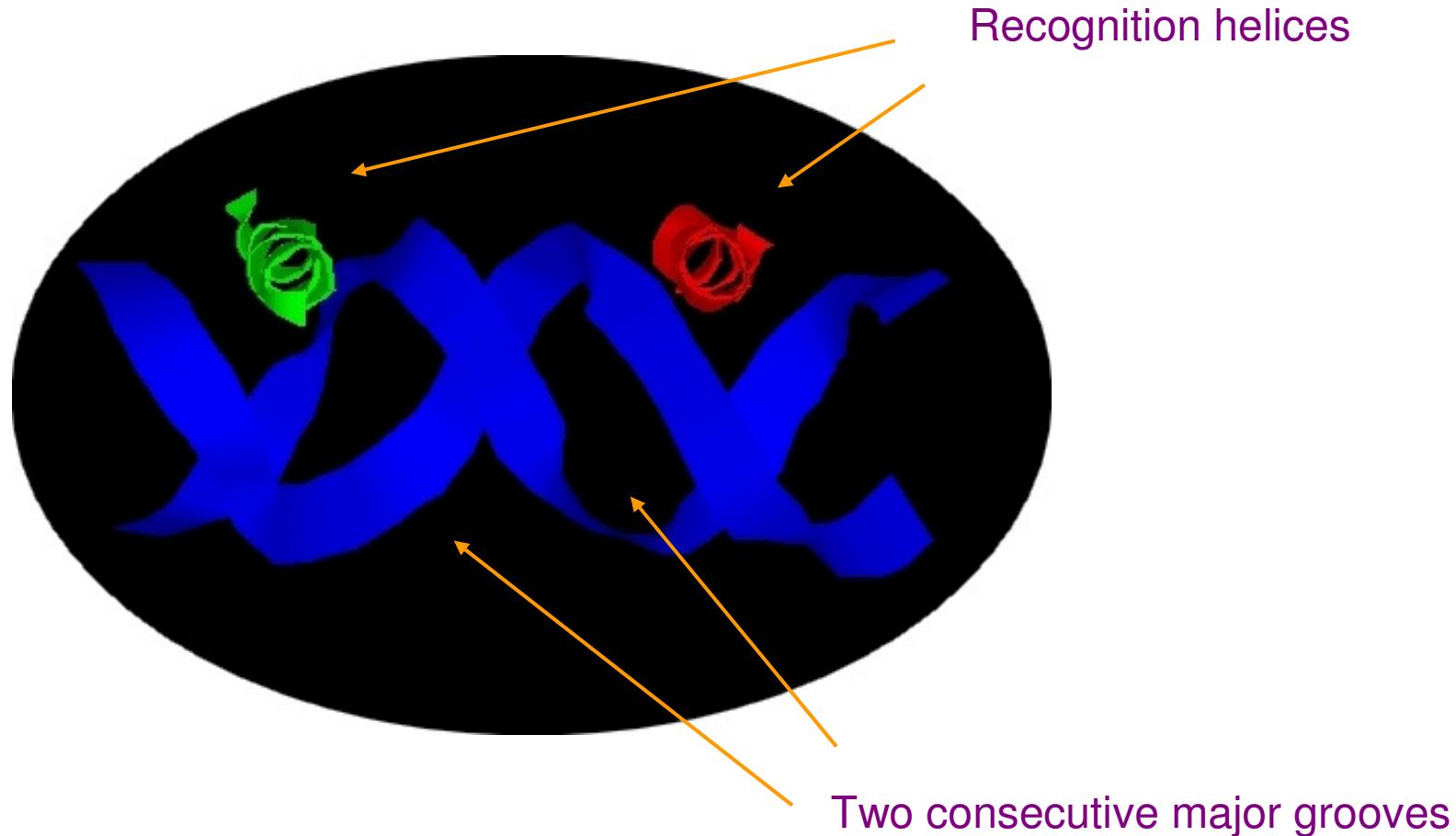
# Dimerization

---

- The two nuclear receptors in the dimer bind to the DNA backbone through **specific and non-specific interactions**, as described above.
- The recognition helix of each monomer is positioned in **two consecutive major grooves**.
- The **spacer region** between the two response elements is crucial for proper binding of the dimer receptor.
- Depending on the type of dimerization (homo or hetero), there are two types of response elements: **direct or invers**.

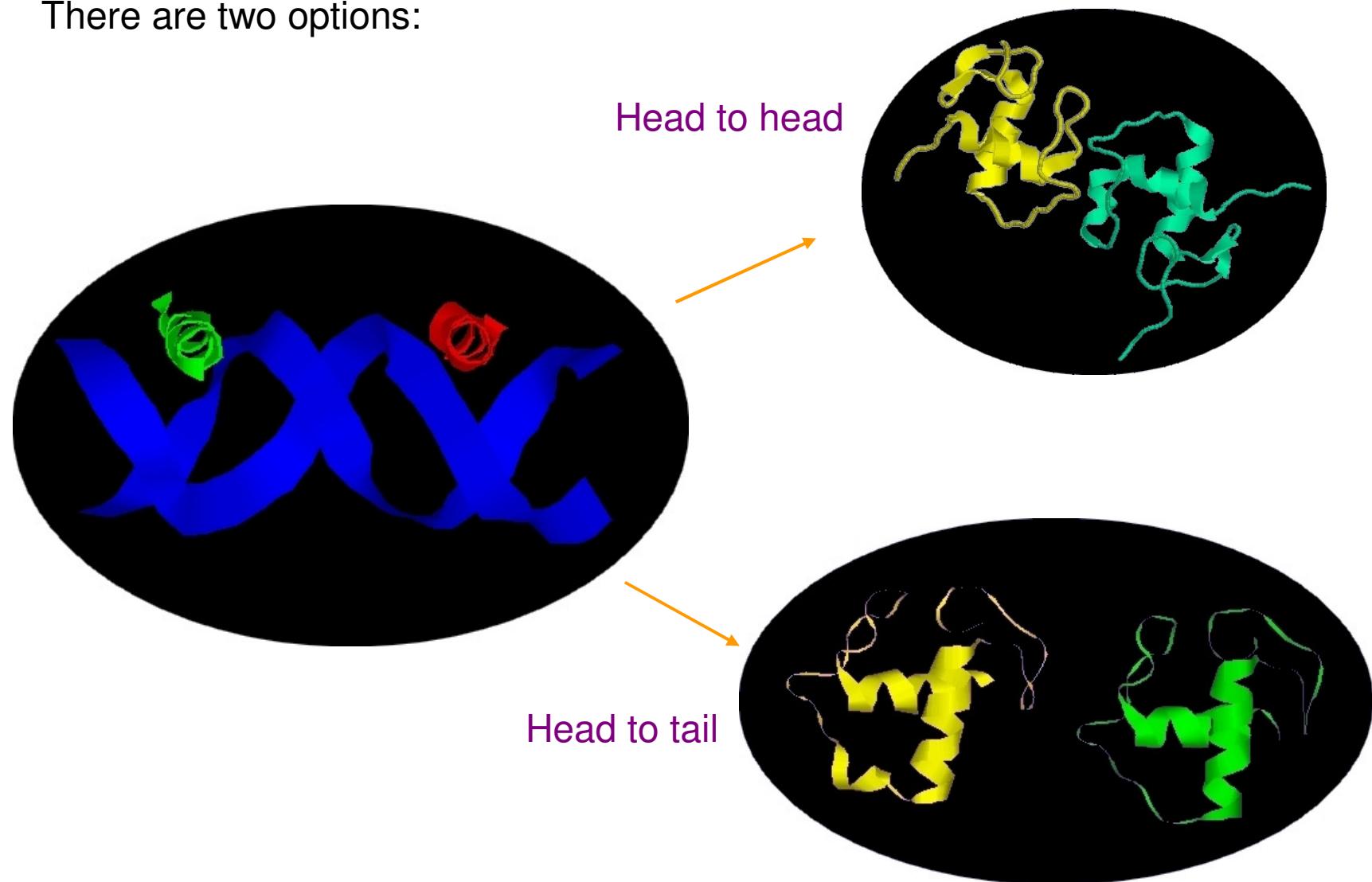


# Dimerization



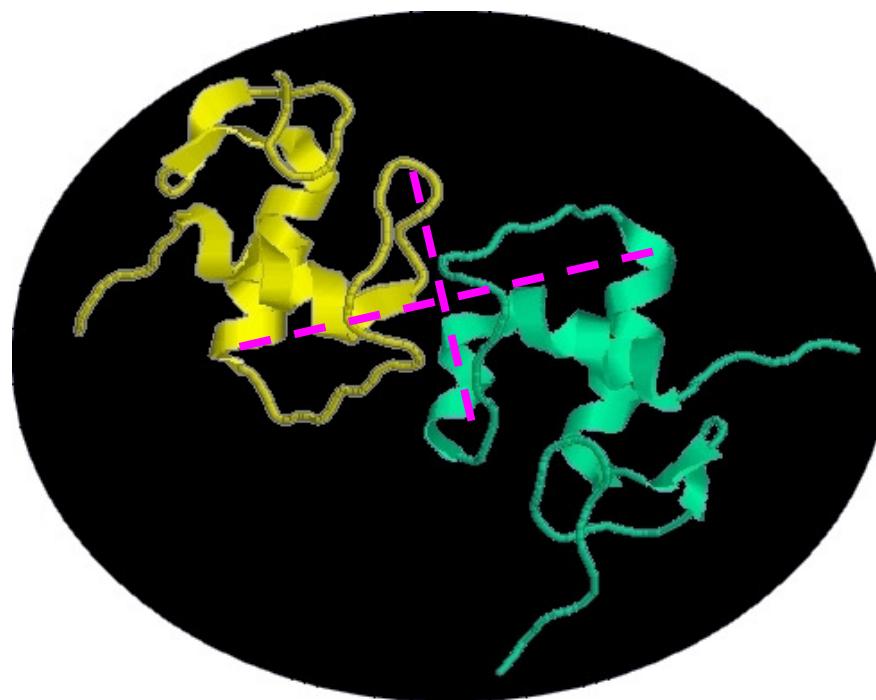
# Dimerization

There are two options:



# Homodimerization

- This is the type of dimerization of the **steroid receptor subfamily**, in particular, of the glucocorticoid receptor.
- The two monomers in the homodimer are in a “**head to head**” position, that is, they interact symmetrically:

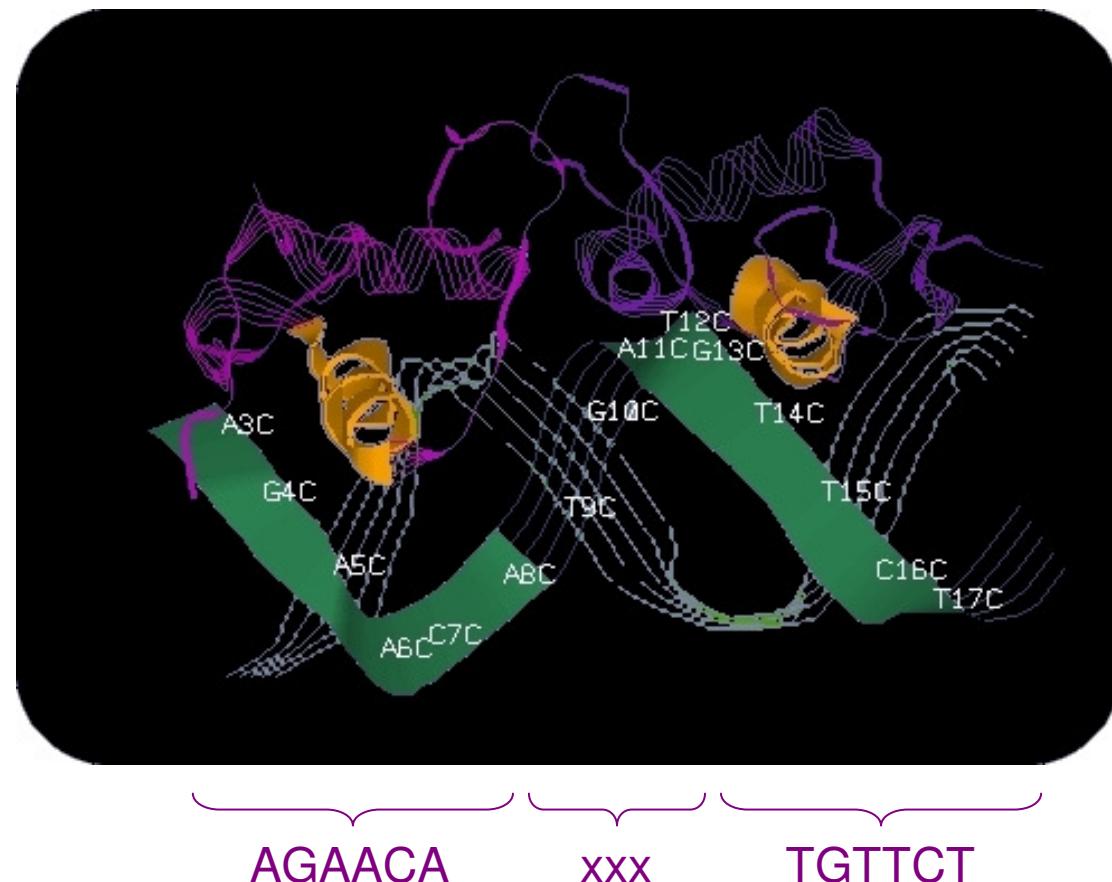


# Homodimerization

- Due to this symmetry, the homodimer recognizes response elements where the half-sites are organized in a **palindromic orientation**.

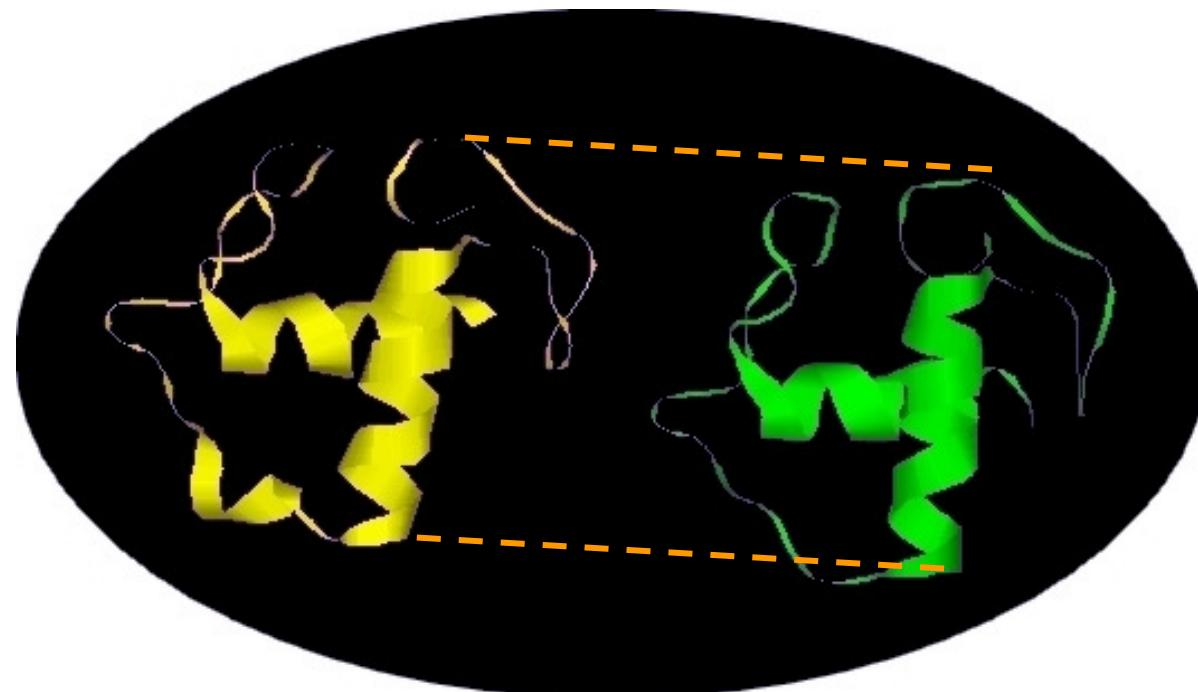
- In the glucocorticoid receptor, the response element sequence is:

5' AGAACAx<sub>xx</sub>TGTTCT 3'  
3' TCTTGTx<sub>xx</sub>ACAAGA 5'



# Heterodimerization

- Some nuclear receptors, like the vitamin D (VDR), thyroid hormone (TR) or the retinoic acid (RAR), form heterodimers with the retinoid acid receptor (RXR).
- In this case, the two monomers in the heterodimer are in a “**head to tail**” position:



# Heterodimerization

- Due to the “head to tail” position, the heterodimer binds to **direct DNA repeats**.
- The **spacing** between the direct repeats is different for each receptor dimer combination, and hence it determines the DNA specificity of each RXR heterodimer.

Response elements:

- RXR-VDR:

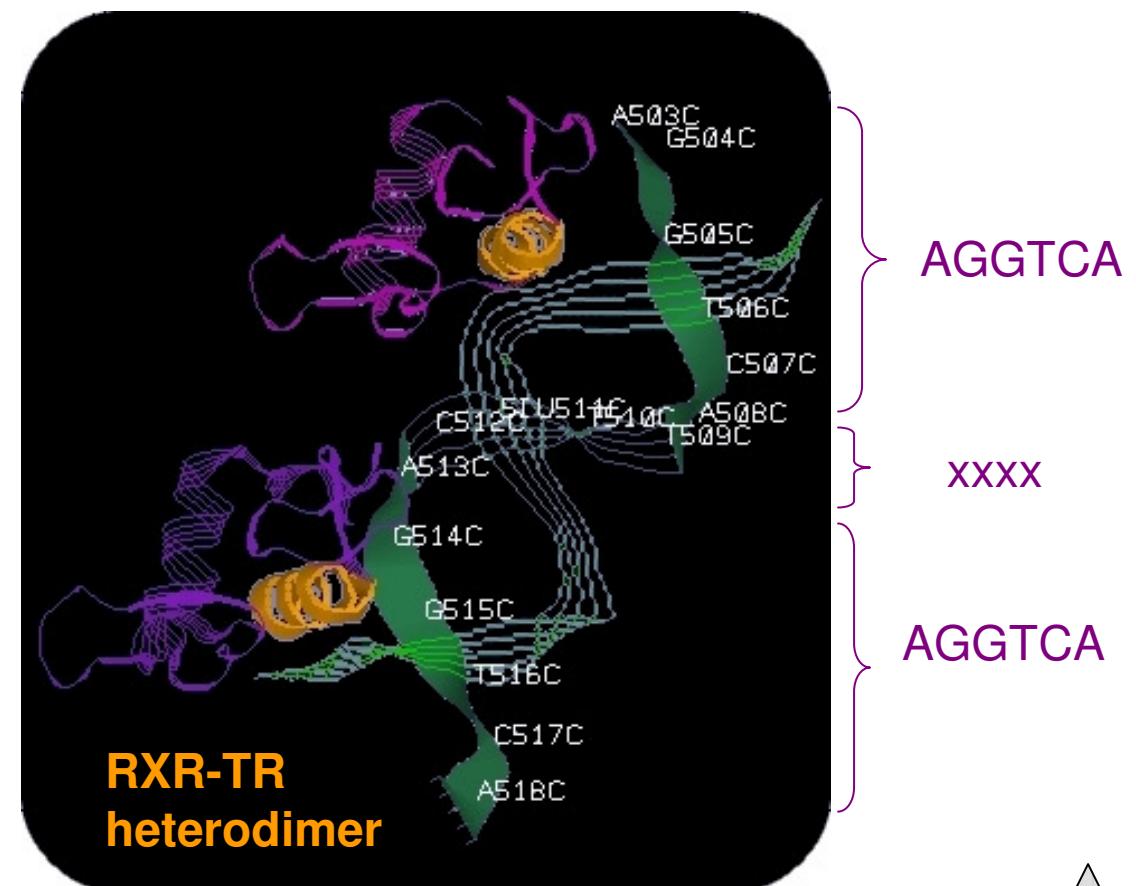
**AGGTCA**xxx**AGGTCA**  
**TCCAGT**xxx**TCCAGT**

- RXR- TR:

**AGGTCA**xxxx**AGGTCA**  
**TCCAGT**xxxx**TCCAGT**

- RXR-RAR:

**AGGTCA**xxxx**AGGTCA**  
**TCCAGT**xxxx**TCCAGT**



# ClustalW alignment

ClustalW  
alignment  
with  
one  
member  
of each  
nuclear  
receptor  
family.

Q5VYG4   Q5VYG4_HUMAN	MASFTKHICAI CGDRSSGKHYGVY SCEGCKGFFKRTVRKD--I TYTCRD NKD--CLIDKR	182
Q505F1   TR2_MOUSE	GPNKVF DLCV VCGDKASGRHYGAI TCEGCKGFFKRSIRKN--LVYSCRGSKD--CVINKH	148
Q9Y466   NR2E1_HUMAN	-----CKVCGDRSSGKHYGVYACDGCGSFFKRSIRRN--RTYVCKSGNQGGCPV DKT	65
P10589   COT1_HUMAN	-----CVVCGDKSSGKHYGQFTCEGCKSFFKRSVRNN--I TYTCRANRN--CPIDQH	133
P41235   HNF4A_HUMAN	-----ALCAICGDRATGKHYGASSCDGCKGFFRRS VRKN--HMYSCRFSRQ--CVV DKT	98
P10276   RARA_HUMAN	-----CFVCQDKSSGYHYGV SACEGCKGFFRRS I QKN--MVYTCHRD KN--CIINKV	135
P22829   NR4A1_RAT	SSGGSEGRCAVCGDNASCQHYGV RTCEGCKGFFKRTVQKS--AKYICLANKD--CPV DKT	313
P55055   NR1H2_HUMAN	-----LCRVC GDKA SGFHYNV LSCEGCKGFFRRS VV RGGARRYACRGGGT--CQMDAF	136
P11473   VDR_HUMAN	-----RICGVCGDRATGFHFNAMTCEGCKGFFRRS MKRK--ALFTCPFNGD--CRITKD	71
P45448   NR5A2_MOUSE	YDEDLEELCPVCGDKVSGYHYGLL TCESCKGFFKRTVQNQ--KRYTCIENQN--CQIDKT	154
P37243   THB2_HUMAN	-YLDKDEL CVVCGDKATGYHYRCITCEGCKGFFRRTI QKNLHPSYSCKYEGK--CVIDKV	171
Q14995   NR1D2_HUMAN	-----MVLCKVCGDVASGFHYGV HACEGCKGFFRRS I QQN--I QYKKCLKNEN--CSIMRM	151
P11474   ERR1_HUMAN	LSSLPKRLCLVCGDVASGYHYGVASCEACKAFFKRTI QGS--IEYSCPASNE--CEITKR	223
P03372   ESR1_HUMAN	ESAKETRYCAV CNDYASGYHYGV SCEGCKA FFKRS I QGH--NDYMCPATNQ--CTIDKN	232
P37231   PPARG_HUMAN	SNSLMAIECRVC GDKA SGFHYGV HACEGCKGFFRRTI RIK--LIYDRCDLN--CRIHKK	185
Q15406   NR6A1_HUMAN	-----CLICGDRATGLHYGIISCEGCKGFFRRS I CNK--RVYRC SRDKN--CVMSRK	107
P35398   RORA_HUMAN	-AQIEIIIPCKICGDKSSGIHYGVITCEGCKGFFRRSQQSN--ATYSCPRQKN--CLIDRT	153
P04150   GCR_HUMAN	TTGPPP KLCVCSDEASGCHYGV LTCGSKVFFKRAVEGQ--HNYLCAGRND--CIIDKI	468
	* : * * : * : ; * . * : * : * : * : * :	
		*
Q5VYG4   Q5VYG4_HUMAN	QRNRCQYCRYQKCLAMGMKREAVQQEERQRGKDRNENE-----	219
Q505F1   TR2_MOUSE	HRNRCQYCRYQRCIAFGMKQDSVQ CERKPIEV SREKSSNCAAS TEK IYIRKDLRSPLAAT	208
Q9Y466   NR2E1_HUMAN	HRNQCRACRCLKKCLEVN MNKDAVQHERGPRTSTIRKQVALYFRGHKEENGAAAHFPSAAL	125
P10589   COT1_HUMAN	HRNQCRQYCRYRCLKKCLKVGMRREAVQQRGRMPPTQPNPGQ-----	170
P41235   HNF4A_HUMAN	KRNQCRYCRYRCLKKCFRAGMKKEAVQNERDR-----	127
P10276   RARA_HUMAN	TRNRCQYCRYQKCFEVGMSKE SVRNDRN KKKKEVPKP-----	172
P22829   NR4A1_RAT	RRNRCQFCRFQKCLAVGMVKEVV RTDSLKGRRGR LPS-----	350
P55055   NR1H2_HUMAN	MRRK CQQC RLRK CKEAGMREQCVLSEEQ I RKKKIRKQQQQE S QSQSQSPVGPQ-----	189
P11473   VDR_HUMAN	NRRHCQACR LKRCV D I GMMK E FILT D EEV QRM KREMI LK RKEEE A L K D S I R P K L S E E Q Q R I-----	131
P45448   NR5A2_MOUSE	QRKRCPYCRFKKCI D VGMK L E A V R A D R M R G G R N K F G P M Y K R D R A I K Q Q Q K A L I R A N G L K L-----	214
P37243   THB2_HUMAN	TRNQCRQECRFK K C I Y VGMAT D L V L D D S K R L A K R K L I E E N R E K R R E E I Q K S I G-----	224
Q14995   NR1D2_HUMAN	NRNR CQQC RFK K C L S V G M S R D A V R F G R I P K R E K Q R M L I E M Q S A M K T M M N S Q F S G H L Q N D T-----	211
P11474   ERR1_HUMAN	RRKACQACRFTKCLRVGMLKEGVRLDRVRGG RQK YK R R P E V D P-----	266
P03372   ESR1_HUMAN	RRKSCQACR L R K C Y E V G M M K G I R K D R R G R M L K H K R Q R D D G E R G E V G S A G-----	284
P37231   PPARG_HUMAN	SRNKCQYCRYQKCLAVGM SHN A I R F G R M P Q A E K E K L I A E I S S D I D Q I N P E S A D I R A L A K H-----	245
Q15406   NR6A1_HUMAN	QRNRCQYCRYRLLKCLQMG M N R K A I R E D G M P G G R N K S I G P V Q I S E E E I E R I M S G Q F E E E A N-----	167
P35398   RORA_HUMAN	SRNRCQHCR L Q K C L A V G M S R D A V K F G R M S K R Q R D S I Y A E V Q K H R M Q Q Q Q R D H Q Q Q P G E A E-----	213
P04150   GCR_HUMAN	RRKNCPACRYRKCLQAGM N L E A R K T K K K I K G I Q Q A T T G-----	506

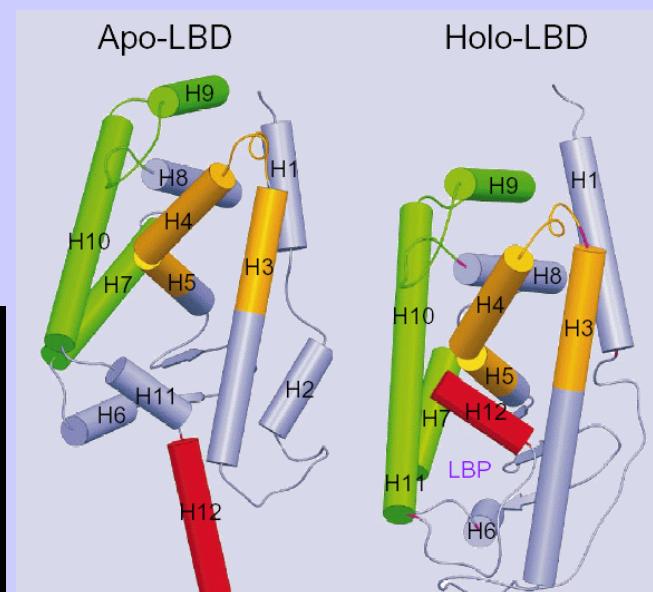
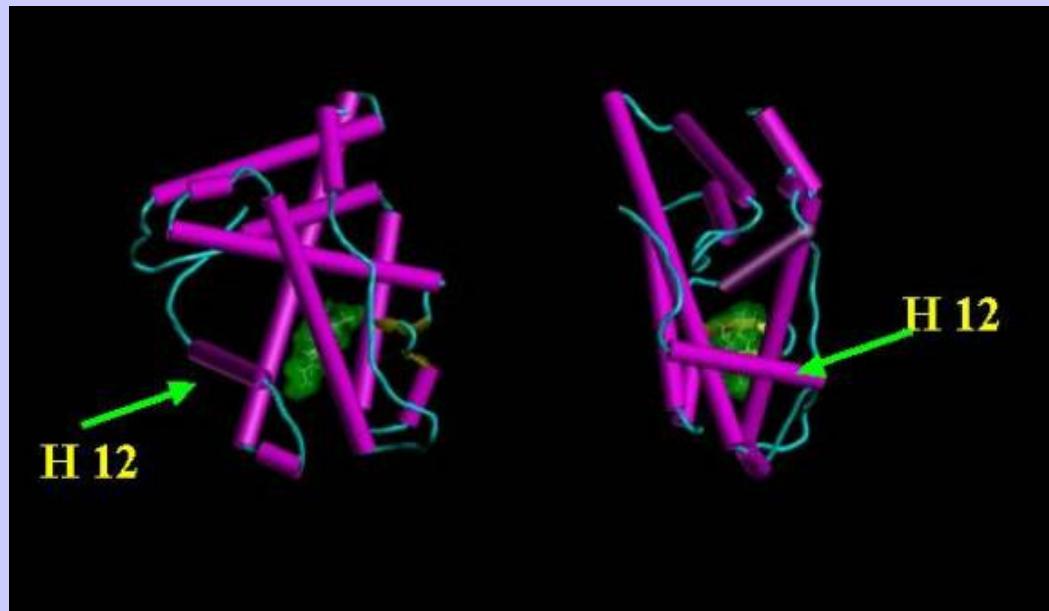


# Summary

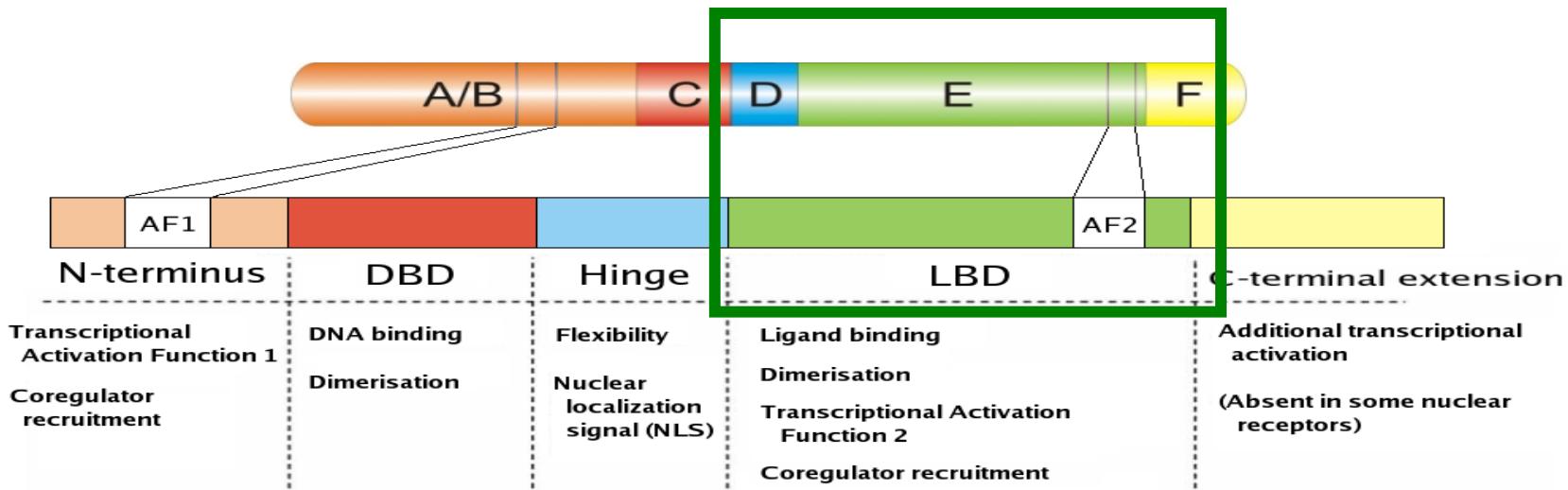
- The DBD consists of about 70 residues that bind to activating elements of DNA called hormone response elements.
- In the DBD there are two zinc containing regions. Each region binds a zinc atom through four cysteine residues.
- The DBD is a highly conserved domain in the family of nuclear receptors.
- The cysteine residues and the residues that form the specific and non-specific interactions with DNA are conserved in the whole family.
- The DBD characterizes the family of nuclear receptors.
- Depending on the type of dimerization (homo or hetero), there are two type of DNA recognition: invers or direct repeats.



# Ligand Binding Domain



# Ligand Binding Domain (LBD)



- This domain is encoded approximately by 250 amino acid residues in the C-terminal end of the molecule.
- This is the second best conserved region of NRs.
- This domain displays a lower degree of conservation among the various nuclear receptors than the DBD.
- The first nuclear receptor LBD structures were solved in 1995. Since then knowledge about structure and function has increased significantly.



# Structural conservation through families

## Subfamily 1: Thyroid Hormone Receptor-like

Group A: Thyroid hormone receptor (Thyroid hormone)

Group B: Retinoic acid receptor (Vitamin A and related compounds) . APO: NR1C3 (PPAR)

Group C: Peroxisome proliferator-activated receptor

. HOLO(+): NR1A2 (TR)

Group D: Rev-erb

. HOLO(-): NR1C1 (PPAR)

Group F: Retinoid-related orphan receptor

Group H: Liver X receptor-like

Group I: Vitamin D receptor-like

## Subfamily 2: Retinoid X Receptor-like

Group A: Hepatocyte nuclear factor-4 (HNF4)

. APO: NR2B1 (RXR)

Group B: Retinoid X receptor (RXR $\alpha$ )

. HOLO(+): NR2B1 (RXR)

Group C: Testicular receptor

. HOLO(-): ?

Group E: TLX/PNR

Group F: COUP/EAR

## Subfamily 3: Estrogen Receptor-like (Steroid hormone receptor)

Group A: Estrogen receptor (Sex hormone receptors; sex hormones: Estrogen)

. APO: NR3B3

. HOLO(+): NR3A1

Group B: Estrogen related receptor

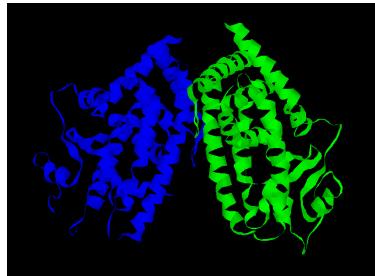
. HOLO(-): NR3A1

Group C: 3-Ketosteroid receptors

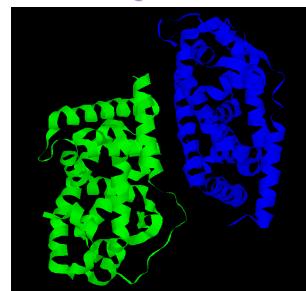


# Structural conservation through families

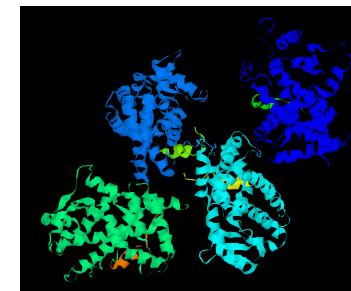
## Subfamily 1: Thyroid Hormone Receptor-like



APO: 1prg.pdb (R=2.4 Å)



HOLO(+): 1n46.pdb (R=2.2 Å)

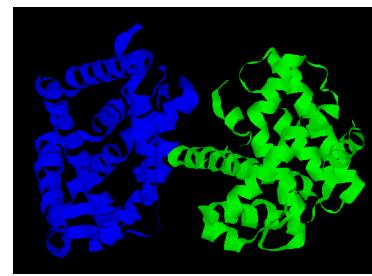


HOLO(-): 1kkq.pdb (R=3.0 Å)

## Subfamily 2: Retinoid X Receptor-like

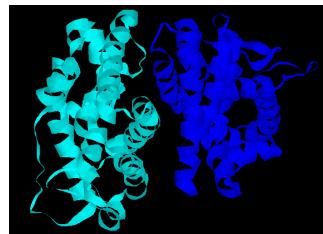


APO: 1lbd.pdb (R=2.4 Å)

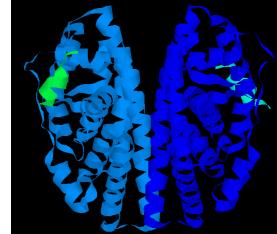


HOLO(+): 1fby.pdb (R=2.25 Å)

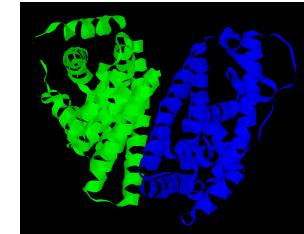
## Subfamily 3: Estrogen Receptor-like (Steroid hormone receptor)



APO: 1kv6.pdb (R=2.7 Å)



HOLO(+): 3erd.pdb (R=2.03 Å)

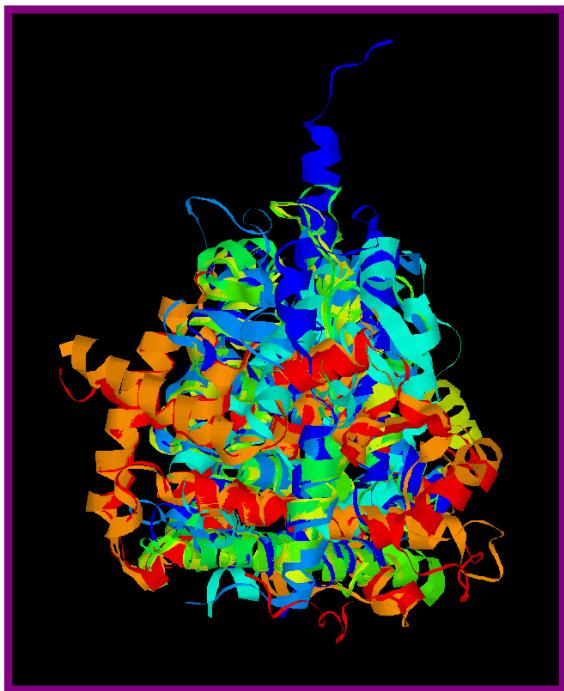


HOLO(-): 1err.pdb (R=2.6 Å)

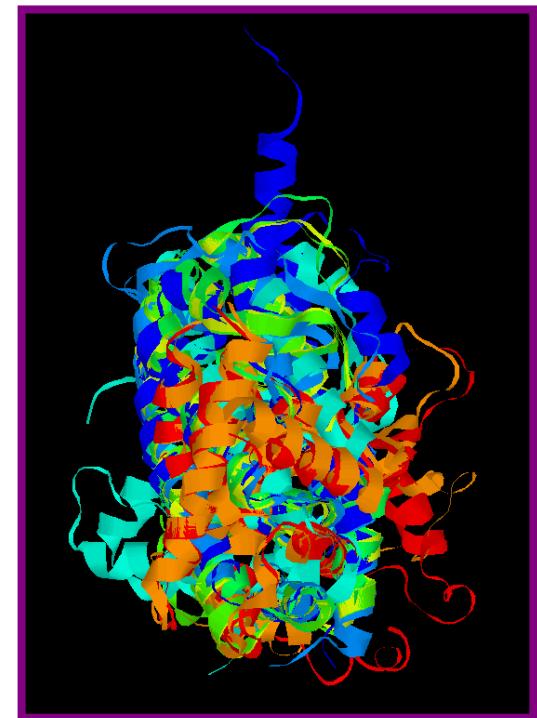


# LBD 3D structures

- The 3D structures of crystallised LBDs superimposed showed that the overall structures of the LBDs of different nuclear receptors are similar, revealing a canonical fold for the nuclear receptor LBD.



**STAMP code:**  
Alignment score Sc: 1.60  
Alignment length Lp: 324  
RMSD: 3.83



# ClustalW sequence alignment

CLUSTAL W(1.60) multiple sequence alignment

lerrA	-----	ALSLTADQWVSALDD-----AEPPILYSE
1fbvA	-----	SSANEDMPVERILE-----AELAVEP--
1kkqA	-TADLKS	LAKRIYEAYLKNFNWNKVKARVILSGKASNNPFWIHDNETLCAEKTLYAKL
1kv6A	-----	-----NKIVSHLLV-----AEPEKIYAN
1lbd	-----	-----SANEDMPVERILE-----AELAVEPKT
1n46A	-----	KPEPTDEEWELIKTVTEAHVATNAQWOKRKFLP-----EDIGOAPIV
1prgA	ESADL	RALAKHYDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDNNSLWNGEDKIKFKH
3erdA	-----	SLASLTDQWVSALDD-----AEPPILYSE
lerrA	YDPTR-----	PFSEASHWGLLTNLADRELVHMINWAKR-VPGFVDTLHDQVHLLECAW
1fbvA	-----	DPVTNICOAADKQLFTLVEWAKR-IPHFSLEPLDDQVILLRAGW
1kkqA	VANG-----	IONKEAEVRFHCCOCTSWEVTELFAKAIPGFANLDLNDQVTLKYGV
1kv6A	PDPT-----	VPDSIKALTLCOLADRELVYIIGMAHK-IPGFSTLSLADQWSLLQSAM
1lbd	ETYVEANHGLNPSSPDPVTNICOAADKOLFTLVEWAKR-IPHFSLEPLDDQVILLRAGW	
1n46A	NAPEG-----	NAPEG-KVWDFLEAFSHFTKIIITPAITRWWDFAKK-LPMFCELPCEDQIILLK6CC
1prgA	ITPLQ-----	EQSKEVAVIRIFQGCFRSVEAVQEITEYAKS1PFGFVNLDLNDQVTLKYGV
3erdA	YDPTR-----	PFSEASHWGLLTNLADRELVHMINWAKRVPGFVDTLHDQVHLLECAW
* * * * *		
lerrA	LEILMIGLWRSMEHPGKLLFAPNLLDRNQGKCVEGMVEIFD-MLLATSSRFRHMMNLQG	
1fbvA	NELLIASFHSRSIAVKDGILLATGLHVRN-SAHSAVGAIIFDRVLTLYSKHNRDQNDK	
1kkqA	YEAI	FANLSSVNNKDGMLVAYGNGFITREFLKSRLKPFCDIME-PKFDFAKFNALELODD
1kv6A	HEIIL	LGVYVRSLSFEDELYADDYDINDED-QSKLAGLDDNN-AIQLQVKKYKSHKLEK
1lbd	NELLIASFHSRSIAVKDGILLATGLHVRN-SAHSAVGAIIFDRVLTLYSKHNRDQNDK	
1n46A	HEIINSLRAAVRYDPESETLTNGEMAYTRG-QLNGGLGVYSD-AIFDGLHSLSFNDD	
1prgA	HEIYT	THLSLNNKDGVL1SEGQGMTRREFLKSRLKPFPGDFME-PKFEFAVKFNALELODD
3erdA	LEILMIGLWRSMEHPGKLLFAPNLLDRNQGKCVEGMVEIFD-MLLATSSRFRHMMNLQG	
* * * * *		
lerrA	EEFVCLKSIILNLNSGYE-----	EKDDHIIHRVLDKITDTLILHMAKAGLTLQQQHQ
1fbvA	TEL6CLRAIWLNPDSKG-----	LSNPAEVEALREKVVASLEAYCKHK--YPEQPG
1kkqA	SDISLFVAAIICCGDRPG-----	LLNVGHIEKNOEGIVHVLRLHLSN--HPDDIF
1kv6A	EEFTLKAIALANSDSMH-----	IEDVEAVOKLQDVLHEALQDYEAQG--HMEPDR
1lbd	TEL6CLRAIWLNPDSKG-----	LSNPAEVEALREKVVASLEAYCKHK--YPEQPG
1n46A	TEVALLQAVLMSSDRPG-----	LACVERIEKYQDSFLLLAFEHINYR---KHHVTH
1prgA	SDLAIFIAYIILSGDRPG-----	LLNWKPIEDIQDNLLQALELQLKLN--HPESSQ
3erdA	EEFVCLKSIILNLNSGYTFLSSTLKSLEEKDHIIHRVLDKITDTLILHMAKAGLTLQQQHQ	
* * * * *		
lerrA	RLAQ	LLLILSHIRHMSNKGMEHLYSM-----PLYDILLEMLDAH-----
1fbvA	RAFK	LLLRLPALRSIGLKCLEHFFFKLIGDTPIOTFLMENLEAP-----
1kkqA	LFP	KLLOKMDALRQLVT--EHAOLVOIIKKTESDAALHPLLQEIYRDMY
1kv6A	RAGK	MLNTPLLRQTSKAVQHFYNIKLEGKVPNHHKLFLEMLEA-----
1lbd	RAFK	LLLRLPALRSIGLKCLEHFFFKLIGDTPIOTFLMENLEAPHOMT-
1n46A	FMP	KLLNWKVTDLRMIGA--CHASRFLHMKVECPTEFLPPLFLFLEVFD-
1prgA	LFAK	KLLOKMDLROIVT--EHWVOLQVIKKTETDMSLHPLLQEIYKDL-
3erdA	RLA	QLLLILSHIRHMSNKGMEHLYSMKCKNWWPLYDILLEMDAHRL--
* * * * *		

- Homology or Remote homology?

All human



# NR3A1 sequence and structural alignment

CLUSTAL W(1.60) multiple sequence alignment

1g50A	---NSLALS LTADQM VSALL DAEPPILY -SEYDPTRPFSEASMMGLL TNLADRELVHMIN
1gwqA	SKKNSLALS LTADQM VSALL DAEPPILY -----SEPFSEASMMGLL TNLADRELVHMIN
1qktA	---NSLALS LTADQM VSALL DAEPPILY -EYDPTRPFSEASMMGLL TNLADRELVHMIN
1l2iA	---NSLALS LTADQM VSALL DAEPPILY SSEYDPTRPFSEASMMGLL TNLADRELVHMIN
1a52A	-----LALS LTADQM VSALL DAEPPILY SEYDPTRPFSEASMMGLL TNLADRELVHMIN
3erdA	----SLALS LTADQM VSALL DAEPPILY SEYDPTRPFSEASMMGLL TNLADRELVHMIN
	***** * . *****
1g50A	WAKR- VPGFV DLT LH DQV HLL EC- AWLEILM IGL WVR SM EHPG KLL FAP NLL DRN QKC
1gwqA	WAKR- VPGFV DLT LH DQV HLL EC- AWLEILM IGL WVR SM EHPG KLL FAP NLL DRN QKC
1qktA	WAKR- VPGFV DLT LH DQV HLL ES- AWLEILM IGL WVR SM EHPG KLL FAP NLL DRN QKS
1l2iA	WAKR VPGFV DLT LH DQV HLL EC- AWLEILM IGL WVR SM EHPG KLL FAP NLL DRN QKG-
1a52A	WAKR- VPGFV DLT LH DQV HLL EC- AWLEILM IGL WVR SM EHPG KLL FAP NLL DRN QKC
3erdA	WAKR VPGFV DLT LH DQV HLL EC- AWLEILM IGL WVR SM EHPG KLL FAP NLL DRN QKC
	*****
1g50A	VEGM VEIFDMLL ATSSRFRMMN LQGEEF VCL KSI ILL NSG VYTFLS STL KS LEED DH IHR
1gwqA	VEGM VEIFDMLL ATSSRFRMMN LQGEEF VCL KSI ILL NSG VYTFLS STL KS LEED DH IHR
1qktA	VEGM VEIFDMLL ATSSRFRMMN LQGEEF VCL KSI ILL NSG VYTFLS STL KS LEED DH IHR
1l2iA	VEGM VEIFDMLL ATSSRFRMMN LQGEEF VCL KSI ILL NSG VYTFLS STL KS LEED DH IHR
1a52A	VEGM VEIFDMLL ATSSRFRMMN LQGEEF VCL KSI ILL NSG VYTFLS STL KS LEED DH IHR
3erdA	VEGM VEIFDMLL ATSSRFRMMN LQGEEF VCL KSI ILL NSG VYTFLS STL KS LEED DH IHR
	*****
1g50A	VLD KIT DTLI LMA KAGL TL QQQ HQR- LAQ LLL LIL SHIR H MSN KGM EHLYS MK CKN VPL
1gwqA	VLD KIT DTLI LMA KAGL TL QQQ HQR- LAQ LLL LIL SHIR H MSN KGM EHLYS MK CKN VPL
1qktA	VLD KIT DTLI LMA KAGL TL QQQ HQR- LAQ LLL LIL SHIR H MSN KGM EHLYS MK CKN VPL
1l2iA	VLD KIT DTLI LMA KAGL TL QQQ HQR LAQ LLL LIL SHIR H MSN KGM EHLYS MK CKN VPL
1a52A	VLD KIT DTLI LMA KAGL TL QQQ HQR- LAQ LLL LIL SHIR H MSN KGM EHLYS MK CKN VPL
3erdA	VLD KIT DTLI LMA KAGL TL QQQ HQR- LAQ LLL LIL SHIR H MSN KGM EHLYS MK CKN VPL
	*****
1g50A	YD LLL EML D A H RL H-
1gwqA	YD LLL EML D A H -
1qktA	YD LLL EML D A H A
1l2iA	YD LLL EML D A H -
1a52A	YD LLL EML D -
3erdA	YD LLL EML D A H R L -
	*****



\* STAMP RMSD: 0.44

- Great identity in sequence and structure -> Homology

All human



# Searching homologues with psi-blast

- We performed the search with THYROID RECEPTOR  $\beta$  (NR1A2)

Results from round 3

Sequences producing significant alignments:	Score (bits)	E Value
Sequences used in model and found again:		
pdb 1BSX 1BSX-A thyroid hormone receptor betafragment: ligand bi...	321	1e-88
pdb 1FM6 1FM6-A retinoic acid receptor rxr-alphafragment: ligand...	278	2e-75
pdb 1FM9 1FM9-A retinoic acid receptor rxr-alphafragment: ligand...	278	2e-75
pdb 1K74 1K74-A retinoic acid receptor rxr-alphafragment: ligand...	278	2e-75
pdb 1LBD 1LBD retinoid x receptorfragment: histidine tag plus do...	277	2e-75
pdb 1G1U 1G1U-A retinoic acid receptor rxr-alphafragment: ligand...	277	2e-75
pdb 1G5Y 1G5Y-A retinoic acid receptor rxr-alphafragment: ligand...	277	2e-75
pdb 1G1U 1G1U-B retinoic acid receptor rxr-alphafragment: ligand...	277	3e-75
pdb 1G5Y 1G5Y-B retinoic acid receptor rxr-alphafragment: ligand...	277	3e-75
pdb 1G1U 1G1U-C retinoic acid receptor rxr-alphafragment: ligand...	277	3e-75
pdb 1G5Y 1G5Y-C retinoic acid receptor rxr-alphafragment: ligand...	277	3e-75
pdb 1DKF 1DKF-B retinoid x receptor-alphafragment: ligand-bindin...	271	2e-73
pdb 1DB1 1DB1-A vitamin d nuclear receptorfragment: ligand bindi...	263	5e-71
pdb 1IE8 1IE8-A vitamin d3 receptor(1,25-dihydroxyvitamin d3 rec...	263	5e-71
pdb 1IE9 1IE9-A vitamin d3 receptor(1,25-dihydroxyvitamin d3 rec...	262	7e-71
pdb 1DKF 1DKF-A retinoid x receptor-alphafragment: ligand-bindin...	262	8e-71
pdb 2LBD 2LBD retinoic acid receptor gammafragment: lbd (ligand-...	262	8e-71
pdb 3LBD 3LBD retinoic acid receptor gammafragment: lbd, ligand-...	262	1e-70
pdb 4LBD 4LBD retinoic acid receptor gammafragment: lbd, ligand-...	262	1e-70
pdb 1EXA 1EXA-A retinoic acid receptor gamma-2fragment: ligand b...	261	1e-70
pdb 1EXX 1EXX-A retinoic acid receptor gamma-2fragment: ligand b...	261	1e-70
pdb 1FCY 1FCY-A retinoic acid receptor gamma-1fragment: ligand b...	261	2e-70
pdb 1FCX 1FCX-A retinoic acid receptor gamma-1fragment: ligand b...	260	2e-70
pdb 1FCZ 1FCZ-A retinoic acid receptor gamma-1fragment: ligand b...	260	2e-70
pdb 1ILG 1ILG-A orphan nuclear receptor pxrfragment: ligand bind...	258	9e-70
pdb 1ILH 1ILH-A orphan nuclear receptor pxrfragment: ligand bind...	258	9e-70
pdb 1FBY 1FBY-A retinoic acid receptor rxr-alphafragment: ligand...	255	9e-69
pdb 1FBY 1FBY-B retinoic acid receptor rxr-alphafragment: ligand...	255	2e-68
pdb 1KKQ 1KKQ-A peroxisome proliferator activated receptorfragme...	244	2e-65
pdb 1K7L 1K7L-A peroxisome proliferator activated receptor alpha...	244	2e-65
pdb 1GWN 1GWN-B ppar-deltafragment: ligand binding domain;	239	6e-64

Psi-blast against pdb\_seq

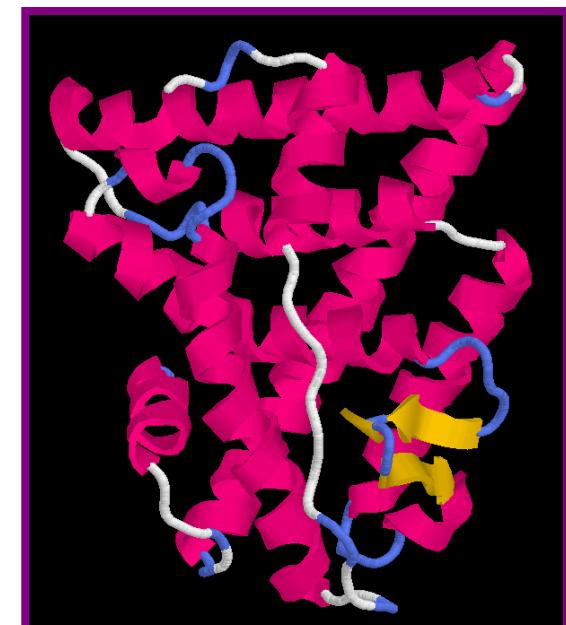
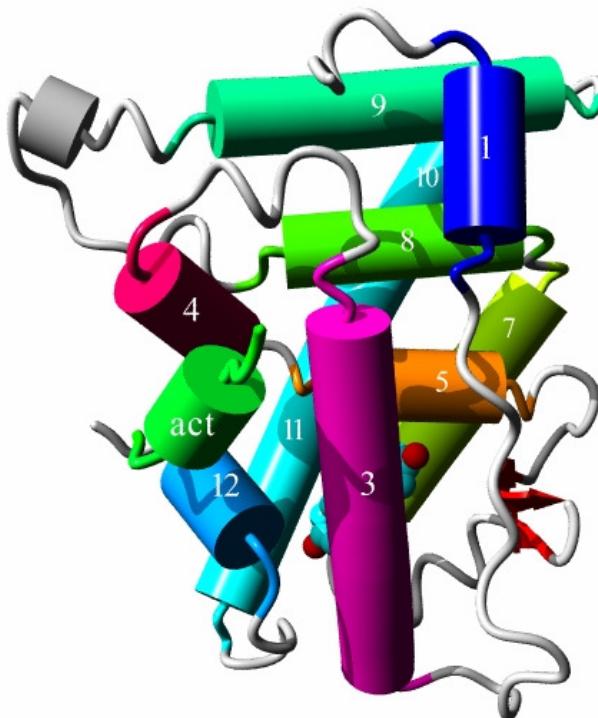


It matches with NR1B (RAR) ,  
NR1I (VIT. D), NR1C (PPAR) all  
groups of the same subfamily ->  
homologous



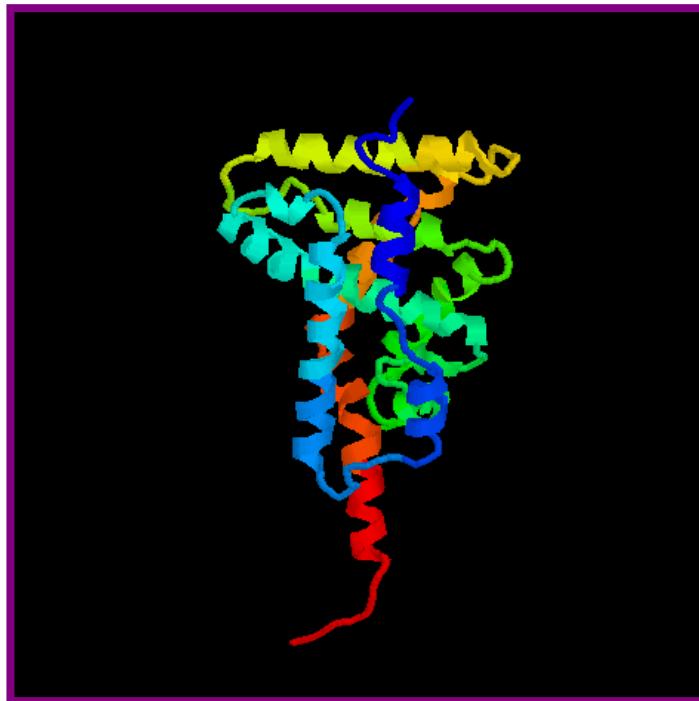
# Secondary structure

- The LBD domain forms a defined globular structure in which eleven to twelve helices are arranged together in an antiparallel, three-layered sandwich, which also includes 2-4 beta-strands.

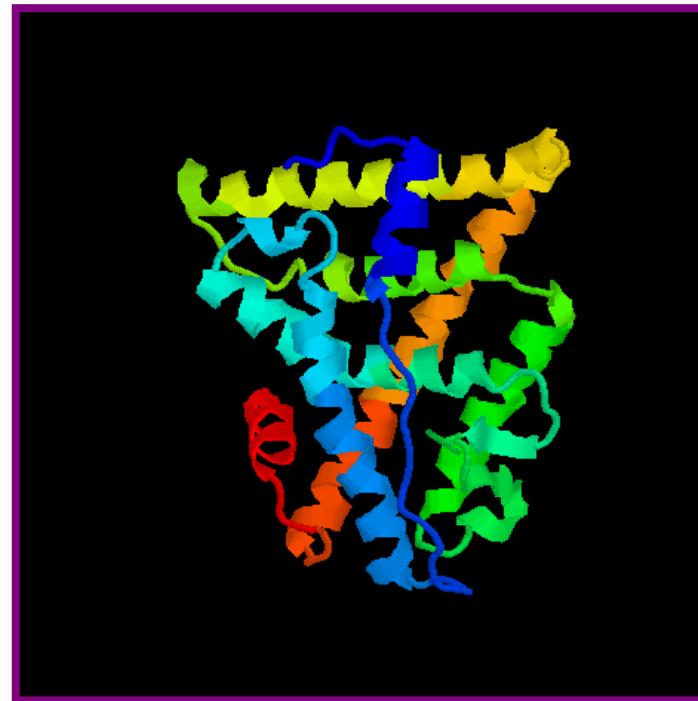


# Apo- Vs Holo-structures

- The holo-structures are more compact than the apo-structures, demonstrating that binding of ligand induces a conformational change in the LBD.



 **APO**  
Unliganded-structure



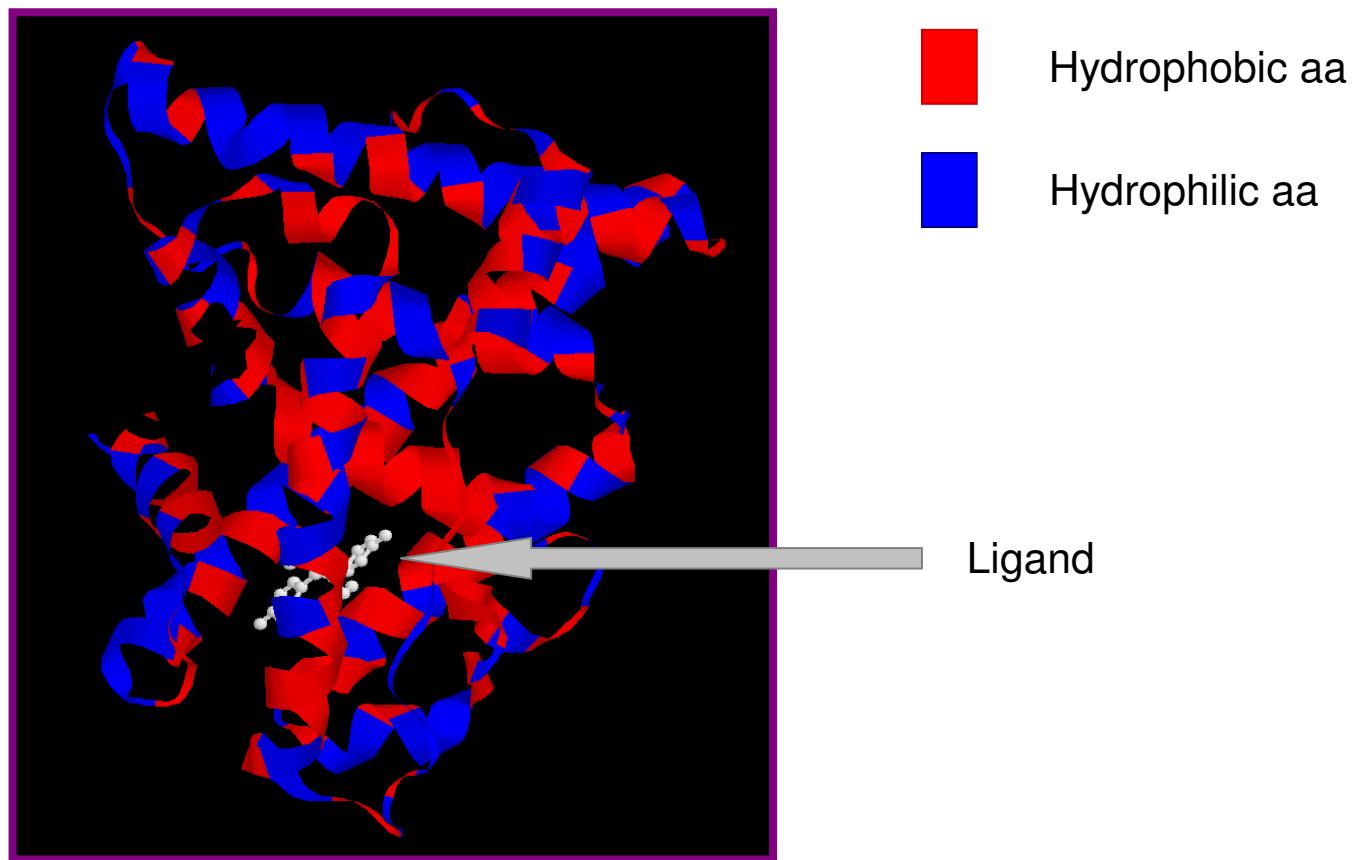
 **HOLO**  
Liganded-structure



# Holo-structures

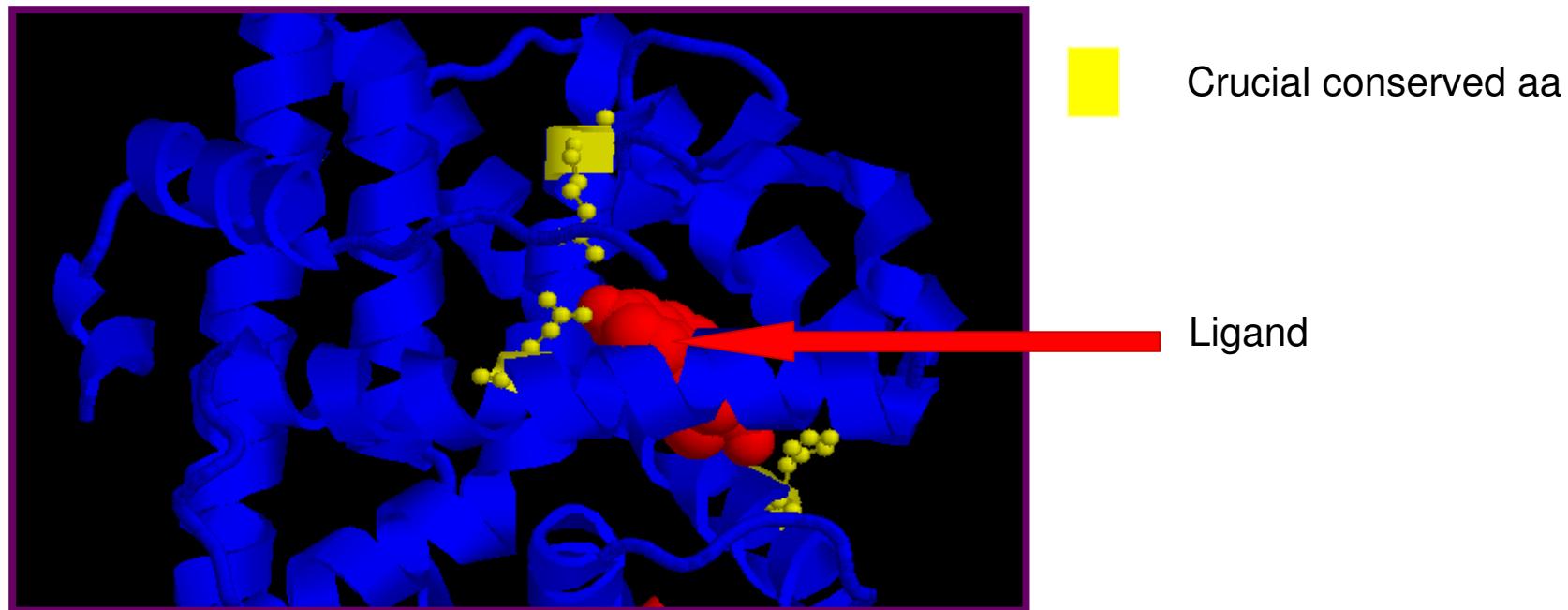
---

- In all holo-structures the ligand binds to a hydrophobic cavity buried within the core of the LBD.
- The ligand becomes an integral part of the hydrophobic core stabilising its 3D structure.



# Holo-structures

- Ligand recognition is achieved through a combination of specific hydrogen bonds and the complementarity of the binding cavity to the non-polar ligand:
  - The binding of estrogens (i.e. estradiol) to ER is by means of the key aa Glu353, Arg394 and His524 and the two hydroxyl groups of the ligand.
  - The architecture of the pocket is rigid and only accommodate planar structures.



# LBD helix 12 (H12)

CLUSTAL W (1.60) multiple sequence alignment

lbd	-----
n46	-----
fby	-----
kv6	-----
erd	-----
err	-----
prg	ESADLRALAKHYDSIYKSFLPTKAKARAILTGKTTD--KSPFVYDMMNSLMH--GEDKIK
kkq	-TADLKLAKRIYEALKNFNMNKVKARVIL--SGKASNNPPFVHDMETLCHAE---KT

lbd	-----SANEDM-PVERIL-----
n46	-----KPEPTDEE-W-E-LIKTVT-----
fby	-----SSANEDM-PVERIL-----
kv6	-----N-K-IVSHLL-----
erd	-----SLASLTAD-Q--NVSALL-----
err	-----ALSLTAD-Q--NVSALL-----
prg	FK-HI----TPLQEOKSKEVAIRIFQGC-----Q-FRSVVEAVQEITEYAKSIPGFVNLDL
kkq	LVAKLVANGIQ---NKEAEVIRIFHCC-----Q-CTSVEVTELTEFAKAIPGFVNLDL

lbd	-----EAE-LAVEP-K--TETYYEANNGLNP-SSPN-----
n46	-----EAH-VATNAWKO--K-R-----KFLPEDIGQAPIVNAPE
fby	-----EAE-LAV-----E-P-----
kv6	-----VA--EPEK-----I--Y-----A--MHDPTVP--D-----
erd	-----DA--EPPI-----L--Y-----S--E-Y-DPTRPF-----
err	-----DA--EPPI-----L--Y-----S--E-Y-DPTRPF-----
prg	NDQVTLLKYGVHEIIYITMLAS-----L-----
kkq	NDQVTLLKYGVYEAIFAMLSS-----Y-----

lbd	-----DPVTNICQAAKDQLFTLV-----E--W-AKRIPHFSE-LPLDDQVILLRAGWN
n46	GGKVDELAESHTKIIITPAITRVV-----D--F-AKKLPMFCE-LPCEDQIILLKGCM
fby	-----DPVTNICQAAKDQLFTLV-----E--W-AKRIPHFSE-LPLDDQVILLRAGWN
kv6	-----SDIKALTTLCDLADRELVII-----G--W-AKHPGFS-T-LSLADQMSLLQSAWN
erd	-----SEASMMGLLTLNADRELVHMI-----N--W-AKRPVPGFVD-LTLHDQVHILLECAWL
err	-----SEASMMGLLTLNADRELVHMI-----N--W-AKRPVPGFVD-LTLHDQVHILLECAWL
prg	-----MNKD-GVLISEGOGFMTRE-----FLKSLRKPFGDFMEPKFE
kkq	-----MNKD-GHLYVAGNGFITRE-----FLKSLRKPFCDOIIMEPKFD

lbd	ELLIAS--FSH-RSI-----A--W-K-D6I-LLA
n46	EIMSLR--AAV--RY-----D--PES-ETL-T-L
fby	ELLIAS--FSH-RSI-----A--W-K-DGILL-A
kv6	EILILG--VYV--RSL-----S--F-E-DELV--Y
erd	EILMIG--LWV--RSM-----E--H-P-GKLL--F
err	EILMIG--LWV--RSM-----E--H-P-GKLL--F
prg	FAVKFNALEDDSDLAIFIAYIILSGDRPGLLNVKPIEDIQDNLLOALEQLKL-----
kkq	FAMKFNALEDDSDISLFVAAIICCGDRPGLLNVGHIEKMQEGIVHVRLRLHQS-----

- Mutational analysis of the LBDs of several nuclear receptors revealed a conserved segment in the most carboxy-terminal part of the LBD.
- This highly conserved LBD region was shown to be essential for the ligand-dependent activation of transcription and is named activation function 2 core motif (AF-2).

lbd	T-GLHVHRNSAHS-AGVGAIF-DRY-LTELVSKHNDQMDKTELGC--LRAIVL--FN-P
n46	NGEMAYTRGQLKNG-GLGVYS-DAI-FDLGMSL-SSFNLDDTEVAL--LQAVLL--MS-S
fby	T-GLHVHRNSAHS-AGVGAIF-DRVLTTELVKM-RDQMDKTELGC--LRAIVL-FNP--
kv6	ADDYIMDEDDQSKLA-GLLDDLN-NAIL-QLVKKY--KSMKLEKEEFYT--LKAIALAN-S-D
erd	APNLLLDNRNQGKCVEGNVEIF-DMLL-ATSSRF-RMHNILQGEFFYC--LKSIIILLN-S-G
err	APNLLLDNRNQGKCVEGNVEIF-DMLL-ATSSRF-RMHNILQGEFFYC--LKSIIILLN-S-G
prg	----NH--PE-----SS-QLFAK--LLQ-K-----MTD-LRQTY-TEHY
kkq	----NH--PD-----DI-FLFPK--LLQ-K-----MAD-LRQLY-TEHA

lbd	DS-K-----GGSNPAAEV--E--ALR--E-KV-YASLEAYCK-HKYP-----EOPGR
n46	DRP-G-----LACVERI--E--KYO--D-SF-LLAFEHYI--NYRK--H-HVT-HF
fby	DSK-G-----LSNPAAEV--E--ALR--E-KV-YASLEAYCK-HKYP--E-QP-GR
kv6	-SMH-I-----EDVEAV--Q--KLO--D-VL-HEALQDYE-A-GHME--D-P-RR
erd	YTFLSSLTKS-LEEKDOI--H--RVL--D-KI-TDTLILHMA--KAGLTLQQOH-QR
err	-----VYEEKDHOI--H--RVL--D-KI-TDTLILHMA--KAGLTLQQOH-QR
prg	-----OLLOV-IKKTEDON-S--LHPPLQ-E-----IYKDL-----
kkq	-----QLVQI-IKKTESDA-ALHPLLQEIY-R-----DHY-----

lbd	FAKLLLRLPALRSIGLKCLEHLFF-----FKLIGDTPIDTLEHENLEAHQNT-----
n46	WPKLLNWKVTDLRMIGACHAS--RFLHMKWE-CPTELFPP-LFLCWEED-----
fby	FAKLLLRLPALRSIGLKCLE--HLFFFKLIGDTP-ID-T-F-LHENLE--AP-----
kv6	AGKMLMLTLLRDTSTKAVQ--HFYNIKLEG-KV-PMHKL--FLENLEA-----
erd	LAQLLLILSHIRHMSNKGME--HLYSMKCKN-VV-PLYDLL--LEMDDAHLR-----
err	LAQLLLILSHIRHMSNKGME--HLY-----SPLVYVLLLEHLDH-----
prg	-----
kkq	-----



## LBD helix 12 (H12)

- This conserved region was predicted to be an amphipathic helix which was later confirmed by the many solved LBD crystal structures.

CLUSTAL W(1.60) multiple sequence alignment

1errA.outSeq ALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMINWAKRVPG  
1errA.outSS -TT--HHHHHHHHHHHH-----SS---SS--HHHHHHHHHHHHHHHHHHHHHHHHHHHTTSTT

1errA.outSeq MLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYQEEKDHIHRVLDKITDTLIGHLMAKAGL  
1errA.outSS HHHHHHHHHHHHHT - HHHHHHHHHHHHHHHHSS - - - HHHHHHHHHHHHHHHHHHHHHHHHHHHHHTT -

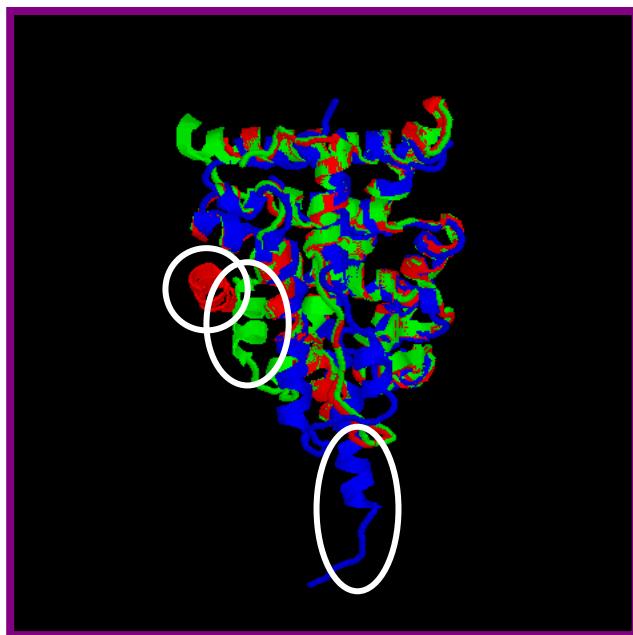
## The last Helix

## The last Helix



# LBD helix 12 (H12)

- The position of helix 12 differs in unliganded and liganded LBDs:



■ Apo - NR  
■ Holo(+) - NR  
■ Holo(-) - NR

## **STAMP code:**

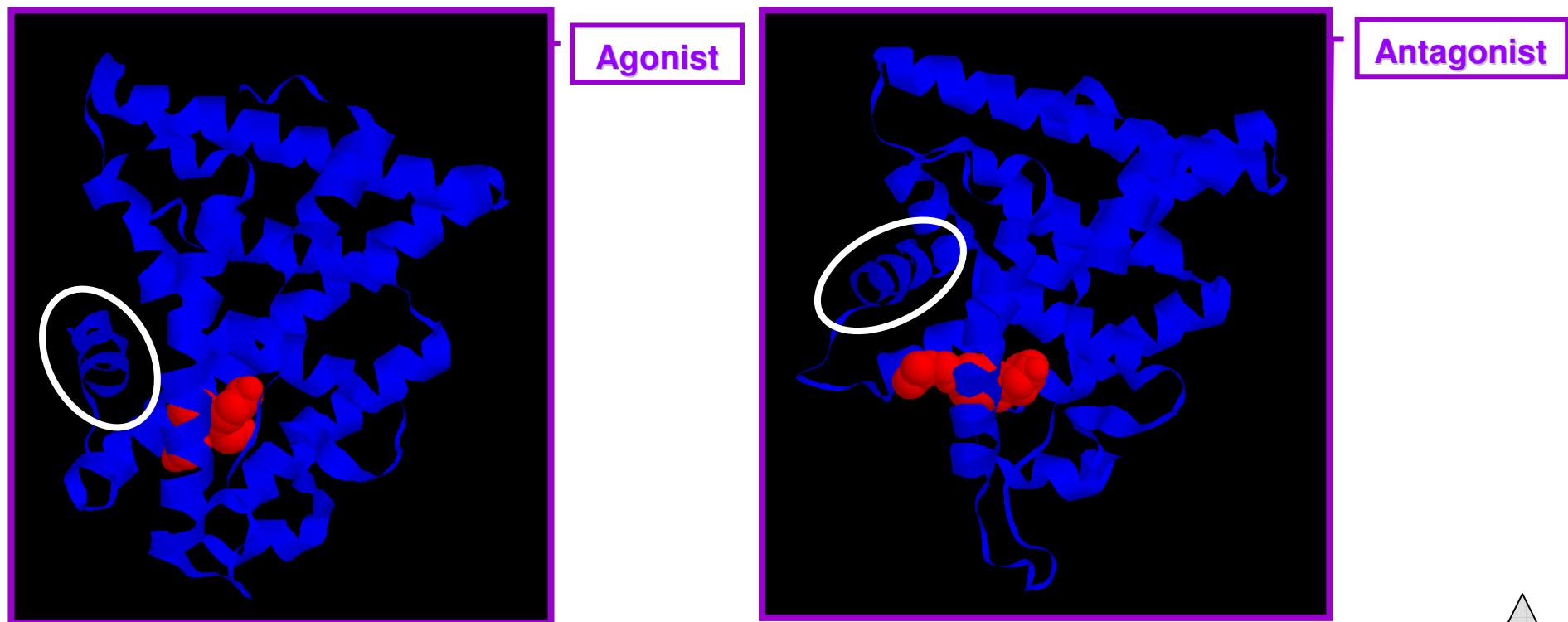
Alignment score Sc: 6.56  
Alignment length Lp: 244  
RMSD: 1.79

- This most C-terminal helix of the LBD is able to act as a molecular switch changing its position depending on ligand-binding.



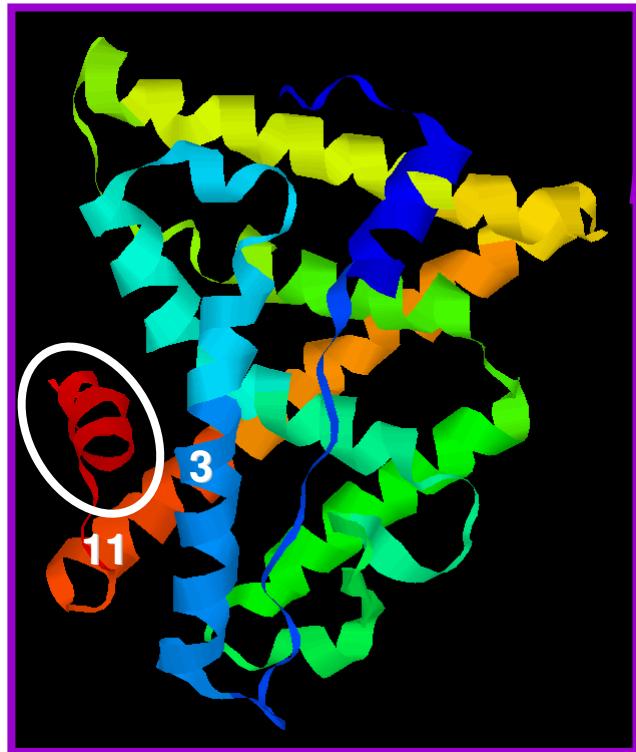
# Agonist/antagonist-induced conformation

- In holo-receptors, changes depending on which type of ligand (agonist and antagonist) is bound to the LBD:
  - **Agonists Ligands**: ligands that fit into the hormone-binding pocket and trigger conformational changes in the LBD, which are suitable for activation.
  - **Antagonists Ligands**: ligands that disrupt the basic structure of the LBD or change the position of H12 needed for binding co-activators.



# Agonist/antagonist-induced conformation

- In holo-receptors the position of H12 also changes depending on which type of ligand (agonist and antagonist) is bound to the LBD.



Agonist

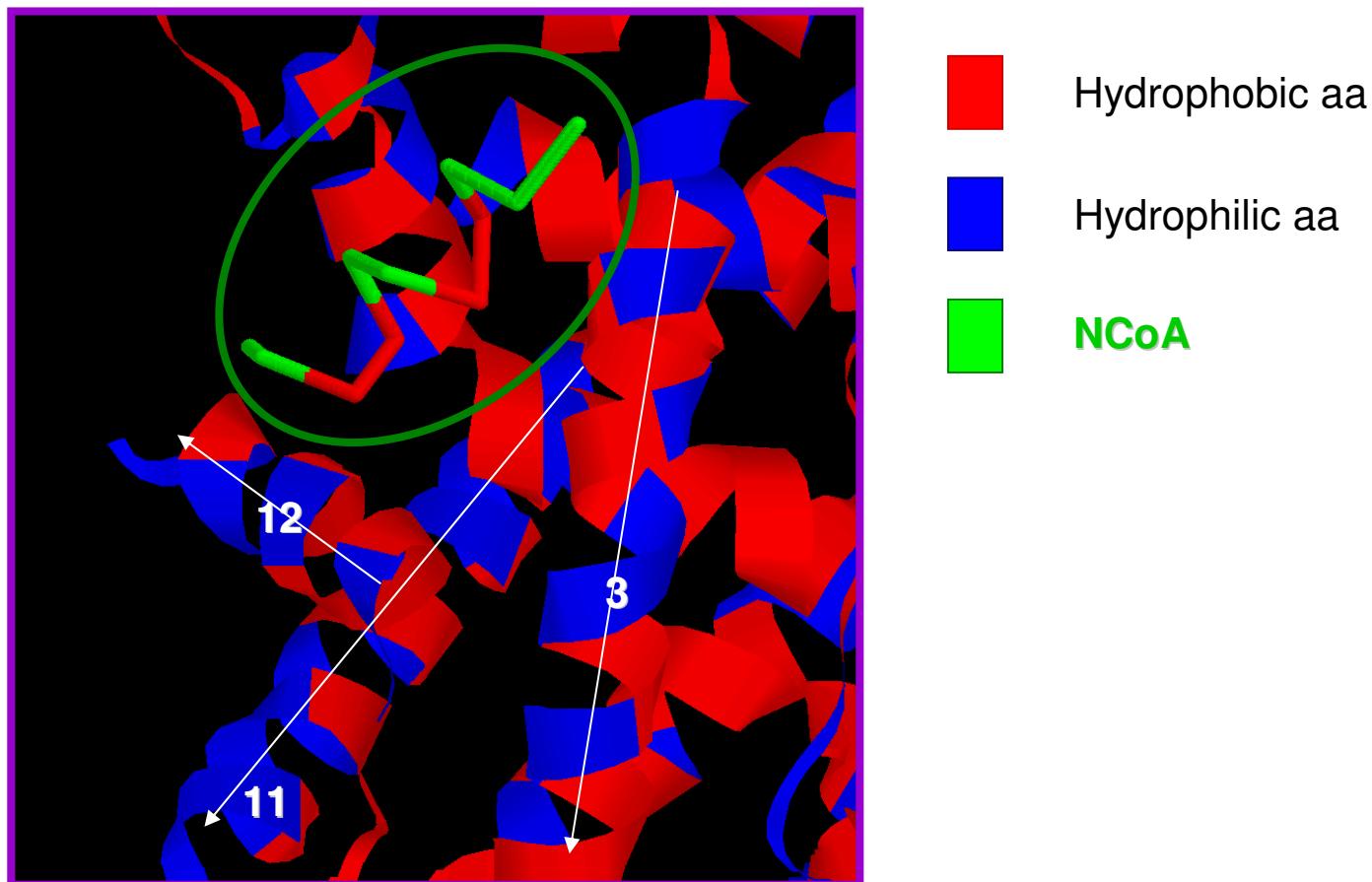


Antagonist



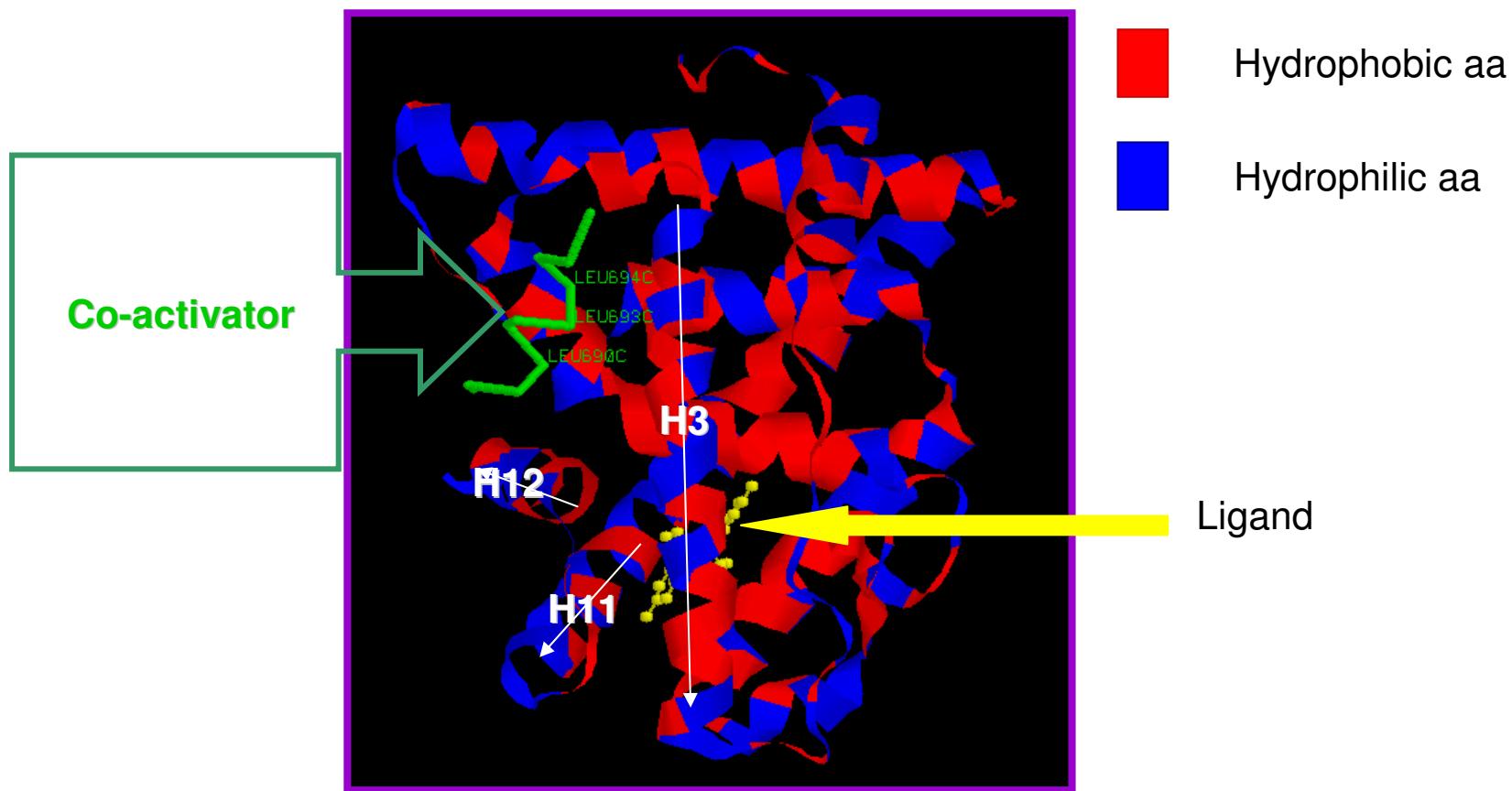
# Agonist-bond structure

- H12 localises against helices 3 and 11 forming **one side of a hydrophobic coactivator-binding surface** -> which allows recruitment of an LXXLL containing helix (the leucine-rich motif for interaction between NR co-activators (NCoAs) and NR).



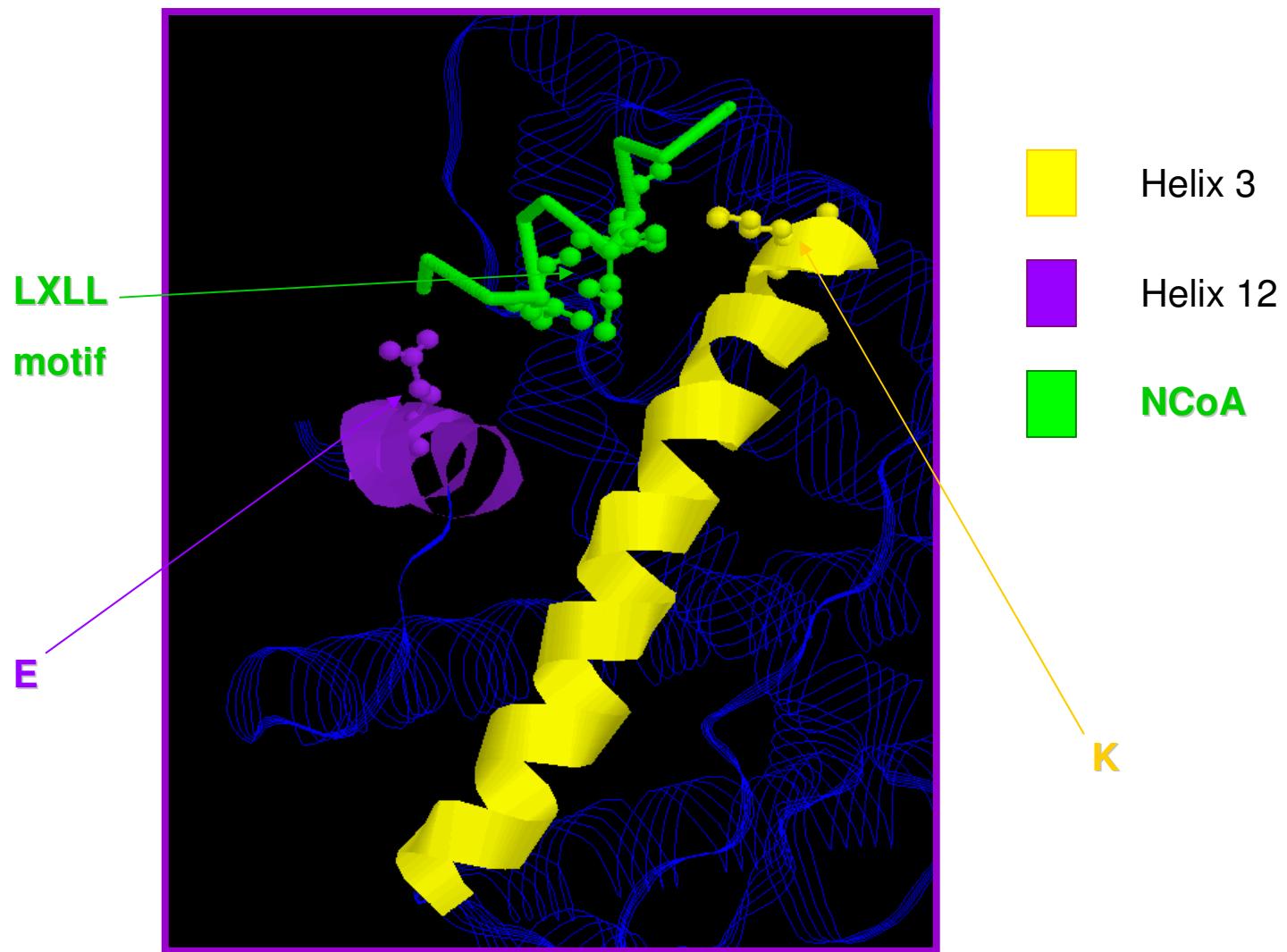
# Interaction NR-coactivator

- Interaction between NR – NCoA: The leucine-rich motif from co-activator is bond to the hydrophobic groove on the LDB by hydrophobic interactions of its leucines with the hydrophobic pocket of the receptor.



# Interaction NR-coactivator

- A **lysine** residue at the C-terminus of H3 and a **glutamate** in H12 are hydrogen-bonded to the peptide bonds in the motif that stabilises the interaction: *Charge Clamp*



# Conservation of residues

CLUSTAL W(1.60) multiple sequence alignment

```

lbd
n46
fby
kv6
erd
err
prg
kkq
ESADLRALAKHYDSIYKSFLTKAKARAILTGKTTD--KSPFVYDMMNSLMM-GEDKIK
-TADLKLAKRIYEAYLKNFNMMNKVKARYL--SGKASNNPPFVHDMETLCHAE---KT

```

```

lbd
n46
fby
kv6
erd
err
prg
kkq
-----SANEDM-PVERIL-----
-----KPEPTDEE-W-E-LIKTVT-----
-----SSANEDM-PVERIL-----
-----N-K-IVSHLL-----
-----SLASLTAD-Q-NVSALL-----
-----ALSLTAD-Q-NVSALL-----
FK-HI----TPLQEOKSKEVAIRIFQGC----Q-FRSVVEAVQEIITEYAKS1PGFVNLDL
LVAKLVANGIQ---NKEAEVIRIFHCC----Q-CTSVEVTELTEFAKA1PGFVNLDL

```

```

lbd
n46
fby
kv6
erd
err
prg
kkq
-----EAE-LAVEP-K-TETYYEANNGLNP-SSPN-----
-----EAH-VATNAWKO-K-R-----KFLPEDIGQAPIVNAPE
-----EAE-LAV-----E-P-----
-----VA-A-EPEK-----I-Y-----A-MPDPTVP--D-----
-----DA-EPII-----L-Y-----S-E-Y-DPTRPF-----
-----DA-EPII-----L-Y-----S-E-Y-DPTRPF-----
NDQVTLLKYGVHEIIYTMLAS-----L
NDQVTLLKYGVYEAIFAMLSS-----V

```

LYSINE

```

lbd
n46
fby
kv6
erd
err
prg
kkq
-----DPVTNICQAAKDQKFLTV-----E-W-AKKRPHFSE-LPLDDQVILLRAGWN
GGKVDLAEOFSHFTKIIITPAITRVV-----D-F-MKKLPMFCE-LPCEDQIILLKGCCM
-----DPVTNICQAAKDQKFLTV-----E-W-AKKRPHFSE-LPLDDQVILLRAGWN
---SDIKALTTLCDLADRELVII-----G-W-AKKRPGFST-LSLAQDQMSLLQSAWN
---SEASMMGLLTLNADRELVHMI-----N-W-AKKRPGFVD-LTLHDQVHLLECAWL
---SEASMMGLLTLNADRELVHMI-----N-W-AKKRPGFVD-LTLHDQVHLLECAWL
---NNKD-GVLISEGOGFMTRE-----FLKSLRKPFQDFMEPKF
---NNKD-GHLYVAGNGFITRE-----FLKSLRKPFQDFMEPKF

```

lysine

```

lbd
n46
fby
kv6
erd
err
prg
kkq
ELLIAS--FSH-RSI-----A--VK-D6I-LLA
EIMSLR--AAV--RY-----D-PES-ETL-T-L
ELLIAS--FSH-RSI-----A--V-K-DGILL-A
EILILG--VYY--RSL-----S-F-E-DELV-Y
EILMIG--LWV-RSM-----E-H-P-GKLL-F
EILMIG--LWV-RSM-----E-H-P-GKLL-F
FAVKFNALEDDSDLAIFIAYIILSGDRPGLLNWKPIEDIQDNLLOALEQLKL-----
FAMKFNALEDDSDISLFVAAIICCGDRPGLLNWKQEGIVHVRLRLHQS-----

```

- Structural conservation for key residues to the interaction with co-activator.
- It can observe the conservation through the families.

```

lbd
n46
fby
kv6
erd
err
prg
kkq
T-GLHVHRNSAHS-AGVGAIF-DRY-LTELVSKHNDNQMDKTELGC--LRAIVL--FN-P
NGEMAYTRGQKNG-GLGVVS-DAI-FDLGMSL-SSFNLDDTEVAL--LQAVLL--MS-S
T-GLHVHRNSAHS-AGVGAIF-DRVLTLSKHM-RDQMDKTELGC--LRAIVL-FNP-
ADDYIMDEQSKLA-GLLDLN-NAIL-QLVKKY-KSMKLEKEEFYT--LKAIALAN-S-D
APNLLLDNRQGKCYEGMVEIF-DMLL-ATSSRF-RMHNOLQGEFFYC--LKSIIILLN-S-G
APNLLLDNRQGKCYEGMVEIF-DMLL-ATSSRF-RMHNOLQGEFFYC--LKSIIILLN-S-G
-----NH--PE-----SS-QLFAK--LLQ-K-----MTD-LRQTY-TEH
-----NH--PD-----DI-FLFPK--LLQ-K-----MAD-LRQLY-TEHA

```

```

lbd
n46
fby
kv6
erd
err
prg
kkq
DS-K-----GLSNPAEV--E-ALR--E-KV-YASLEAYCK-HKYP----EOPGR
DRP-G-----LACVERI--E--KYO--D-SF-LLAFEHYI--NYRK--H-HVT-HF
DSK-G-----LSNPAEV--E-ALR--E-KV-YASLEAYCK-HKYP--E-QP-GR
-SMH-I-----EDVEAV--Q-KLQ--D-VL-HEALQDYEA-GQHME--D-P-RR
VYTFLSSTLKS-LEEKDOI--H-RVL--D-KI-TDTLILHMA--KAGLTLQQOH-QR
-----VYEEKDOI--H-RVL--D-KI-TDTLILHMA--KAGLTLQQOH-QR
-----OLQV-IKKTEDON-S--LHPPLQ-E-----IYKDL-----
-----QLVQI-IKKTESDA-ALHPLLQEIQY-R-----DMY-

```

GLUTAMATE

```

lbd
n46
fby
kv6
erd
err
prg
kkq
FAKLLLRPALRSIGLKCLEHLFF-----FKLIGDTPIDTFLHEWLEAPHQMT-
WPKLLNWKVTDLRMIGACHAS--RFLHMKWE-CPTELFPP-LFUEV/FED-----
FAKLLLRPALRSIGLKCLE--HLFFFKLIGDTP-ID-T-FLHEWLE--AP-----
AGKMLMLTLLRROTSTKAVQ--HFYNNIKLEG-KV-PMHKL-FUEWLEA-
LAQLLLILSHIRHMSNKGME--HLYSMCKN-VV-PLYDLL-LEILDAHRL-
LAQLLLILSHIRHMSNKGME--HLY-----SH-P-LYDLLLEILDAHRL-
-----SH-P-LYDLLLEILDAHRL-

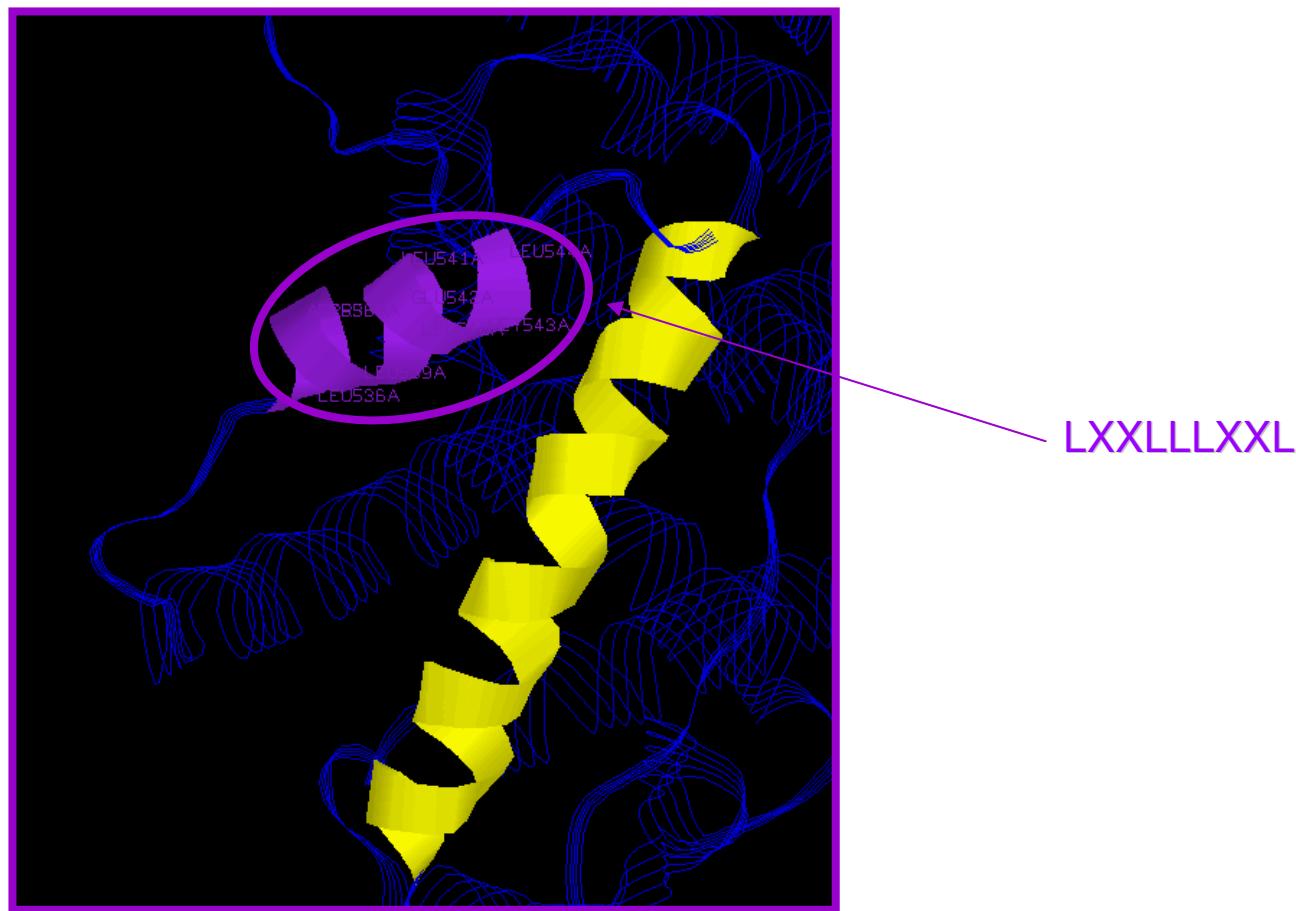
```



# Antagonist-bond structure

---

- In the antagonist-bond structures: H12 has a hydrophobic face **homologous to the LXXLL motif** that may block the interaction of co-activators and allow for co-repressor binding.



# Summary

- The C-terminal ligand-binding domain, whose overall architecture is well conserved between various family members, nonetheless diverges sufficiently to guarantee selective ligand recognition.
- The positioning of H12 is crucial for receptor activation.
- The activation of AF-2 is induced by the interaction with a ligand that changes the domain to more active conformations in the case of agonists and inactive in the case of antagonists.
- Ligand-dependent exchange of corepressors (gene repression) for coactivators (gene activation) and vice versa is the basic mechanism for nuclear receptor mediated regulation of transcription.



# References

---

- Novac, N et al.; Nuclear Receptors: Overview and Classification. Current Drug Targets, 2004.
- Gronemeyer, H. et al.; Principles for Modulation of the Nuclear Receptor Superfamily. NatureReviews, 2004.
- Mangelsdorf, D. et al.; The Nuclear Receptor Superfamily: The Second Decade. Cell, 1995.
- Tobin, J.F. et al.; Nuclear receptors as drug targets in metabolic diseases: new approaches to therapy. Trends in Endocrinology and Metabolism, 2006.
- Kurcinski, M. et al.; Steps towards flexible docking: Modeling of three-dimensional structures of the nuclear receptors bound with peptide ligands mimicking co-activators' sequences. Journal of Steroid Biochemistry & Molecular Biology, 2006.



# NUCLEAR RECEPTORS

## Homology, function and structure

# Questions ?

