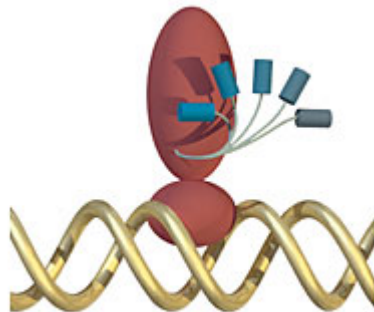


NUCLEAR RECEPTORS

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

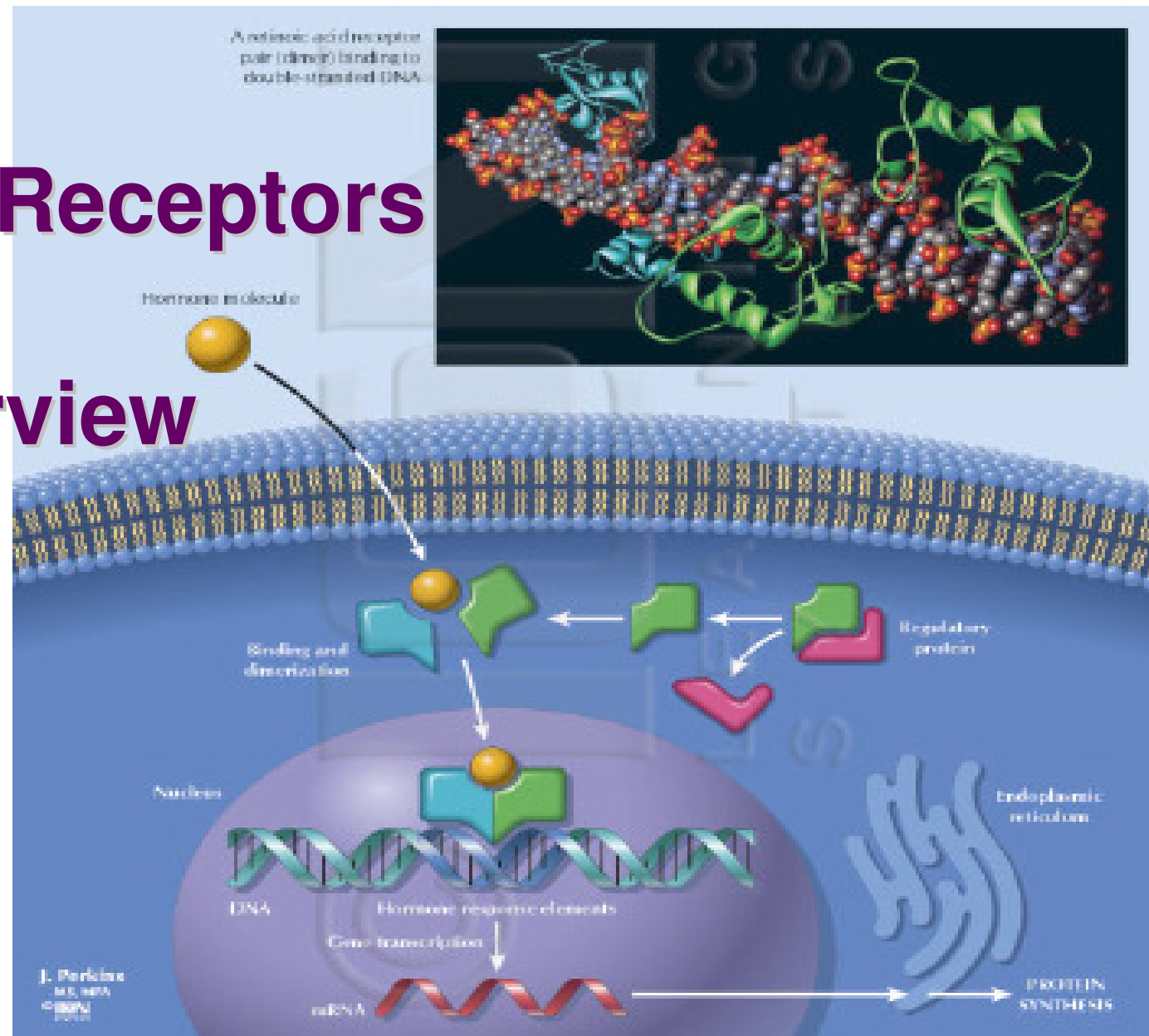
PART II DNA Binding Domain

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-bound structure](#)
- [Interaction NR-coactivator](#)
- [Conservation of residues](#)
- [Antagonist-bound structure](#)

Nuclear Receptors

Overview



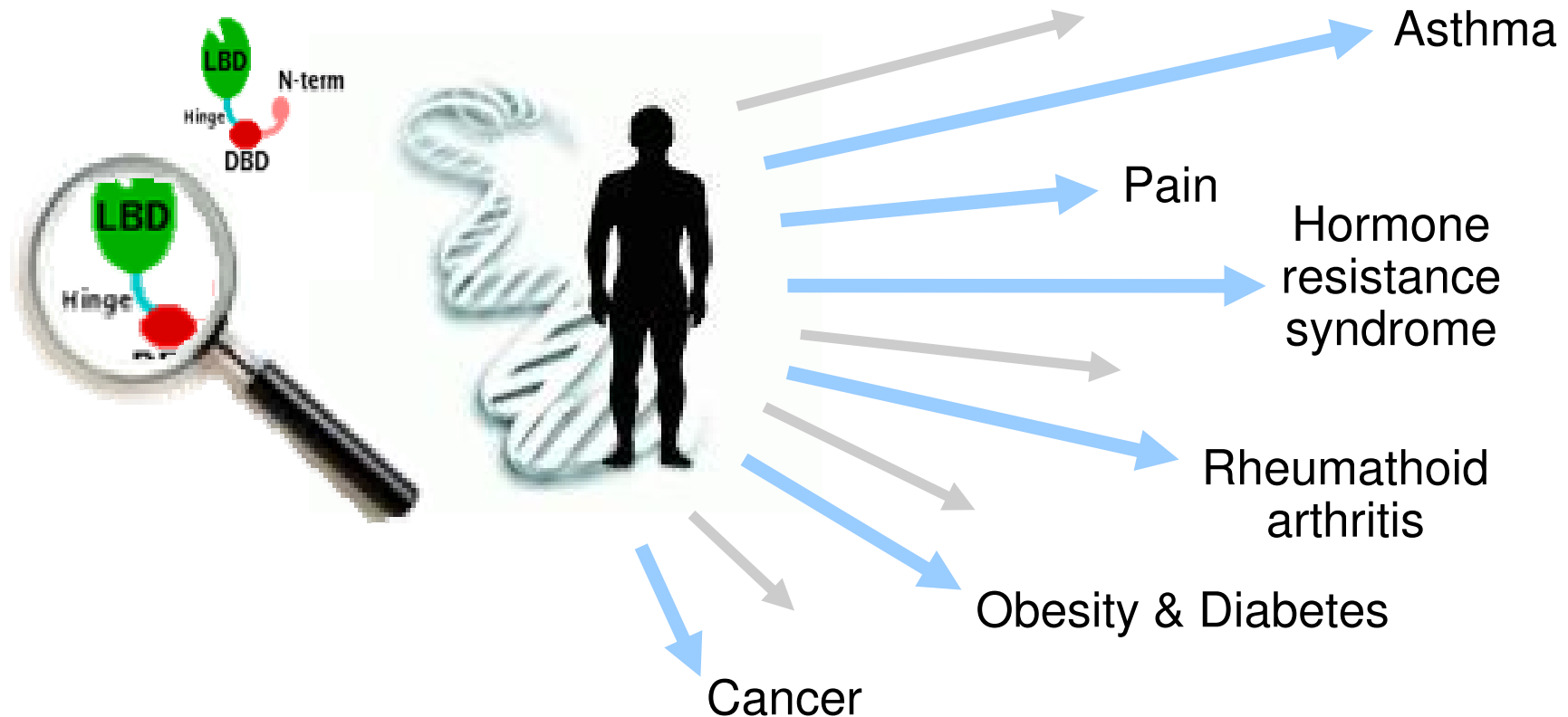
Introduction

- 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.



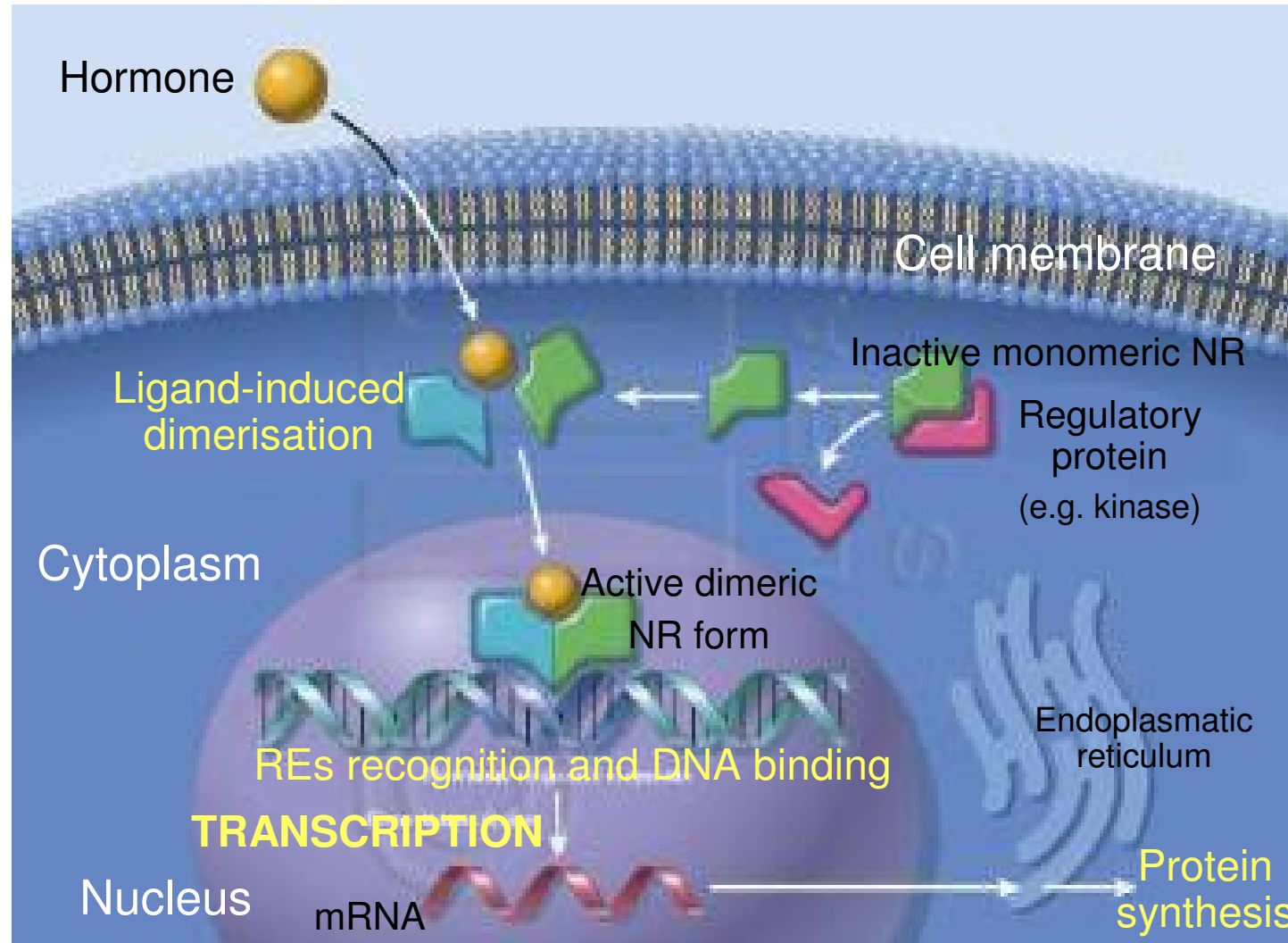
Scientific interest

- 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.



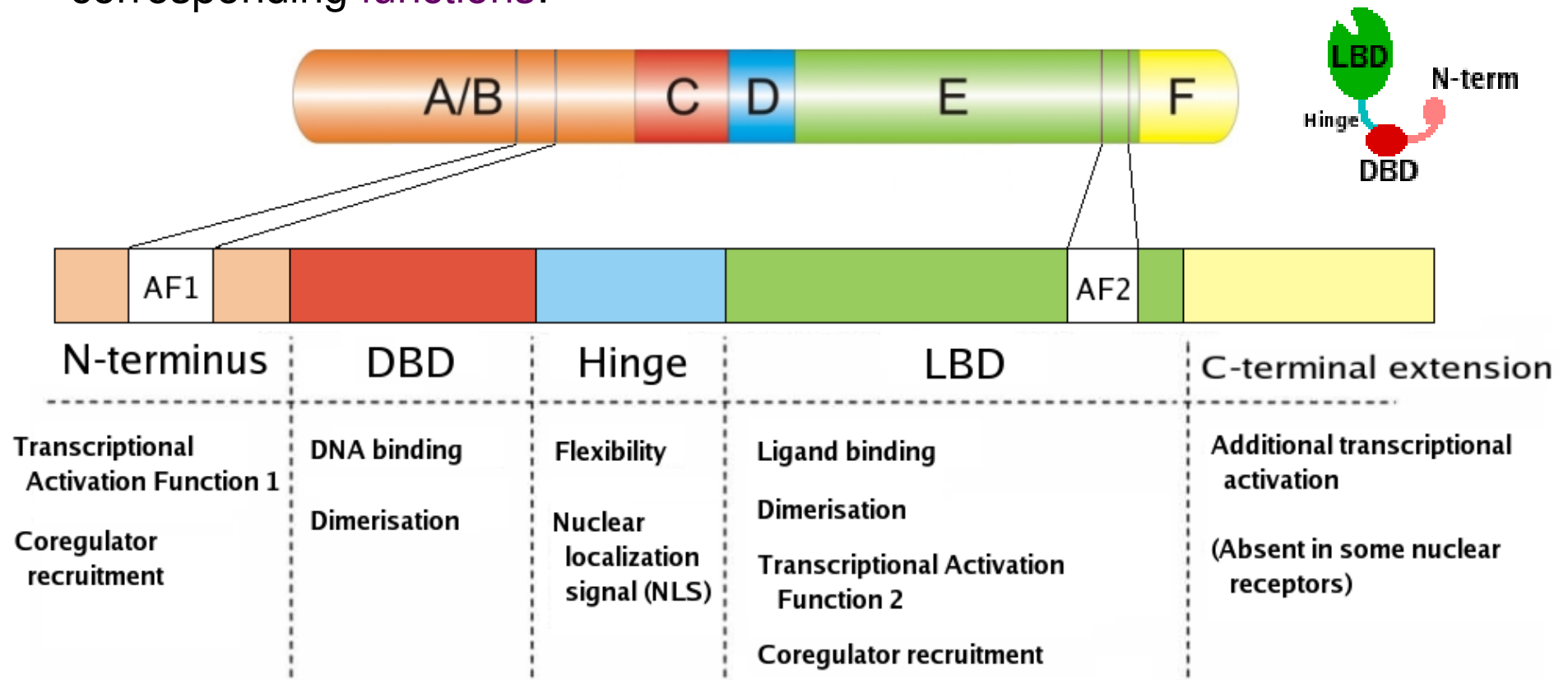
Homeostasis
Cellular
differentiation
Physiology

...



Primary structure

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

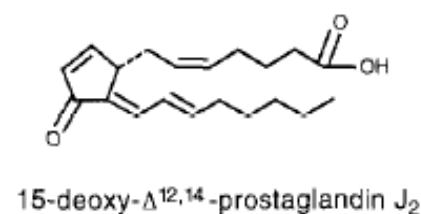
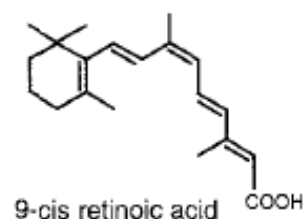
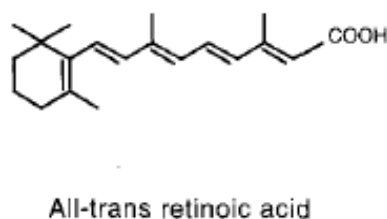
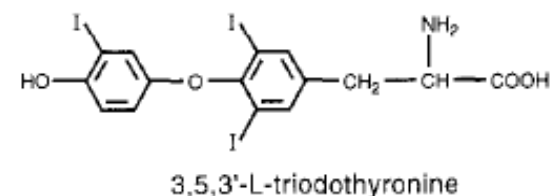
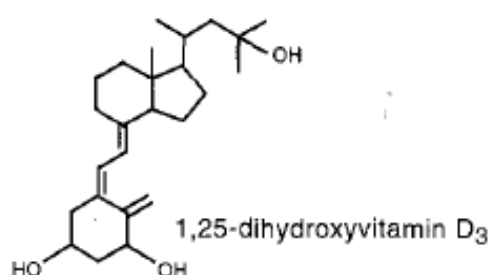
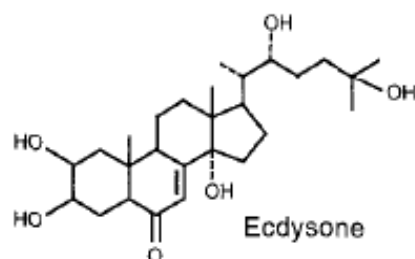
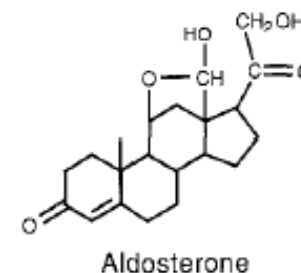
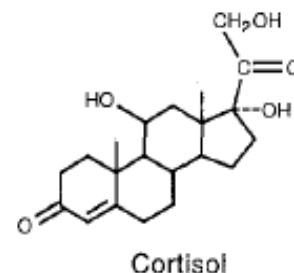
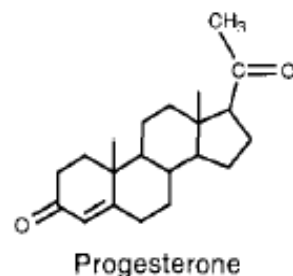
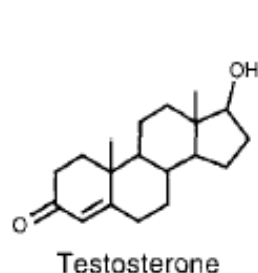
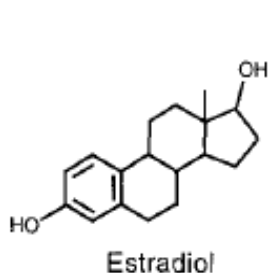


- DNA Binding Domain (DBD) and Ligand Binding Domain (LBD) are significant **conserved regions**, but DBD is the most one.



Ligands

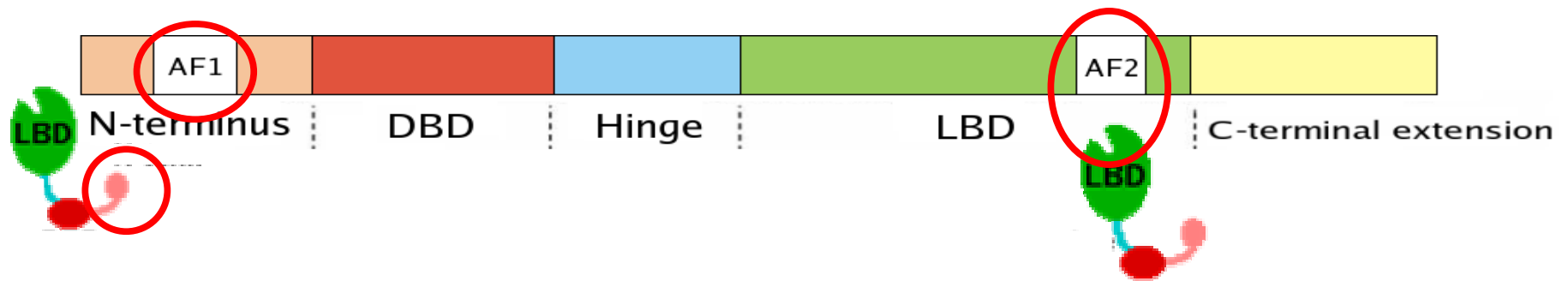
- **Lipophilic substances** such as endogenous hormones, vitamins A and D, drugs, and xenobiotic endocrine disruptors:



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.



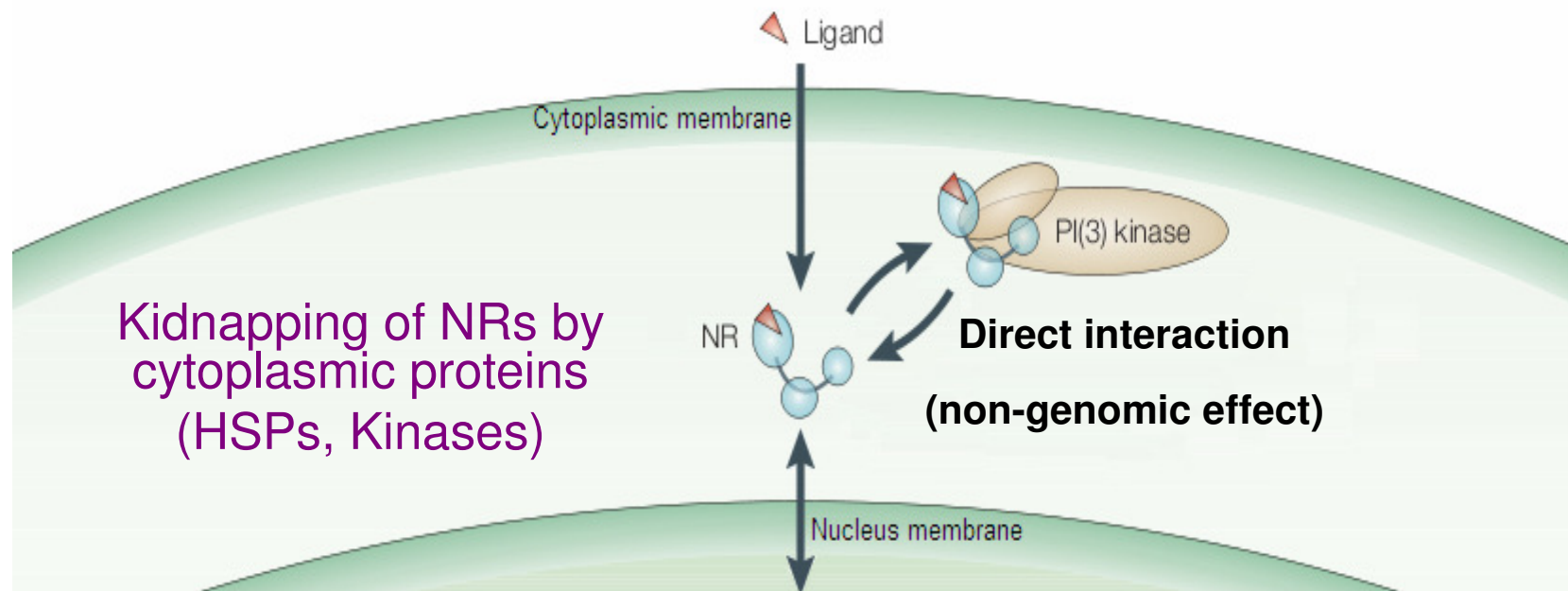
- **Activation Function-2, included in the Ligand Binding Domain**

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.

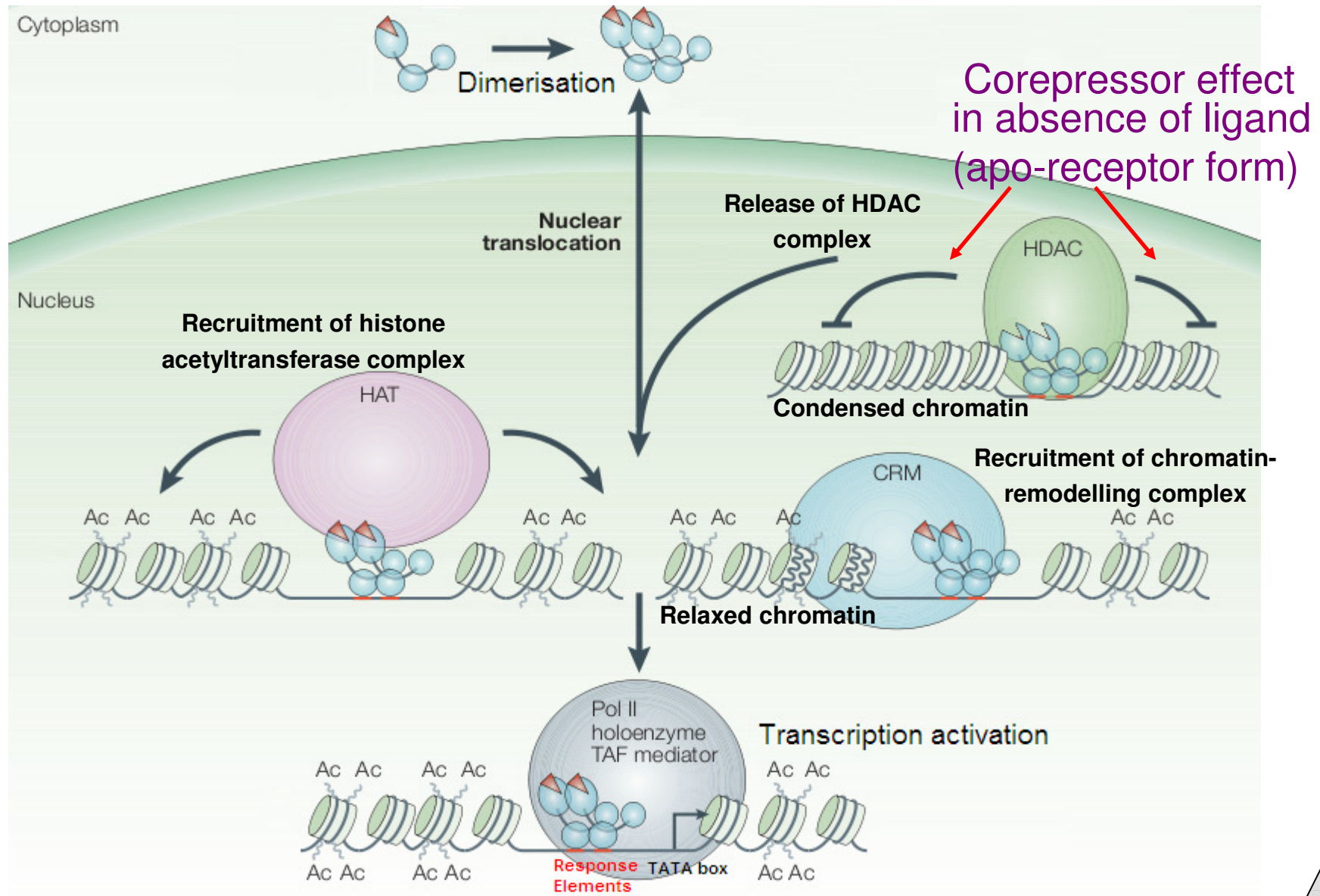


Regulation and mode of action

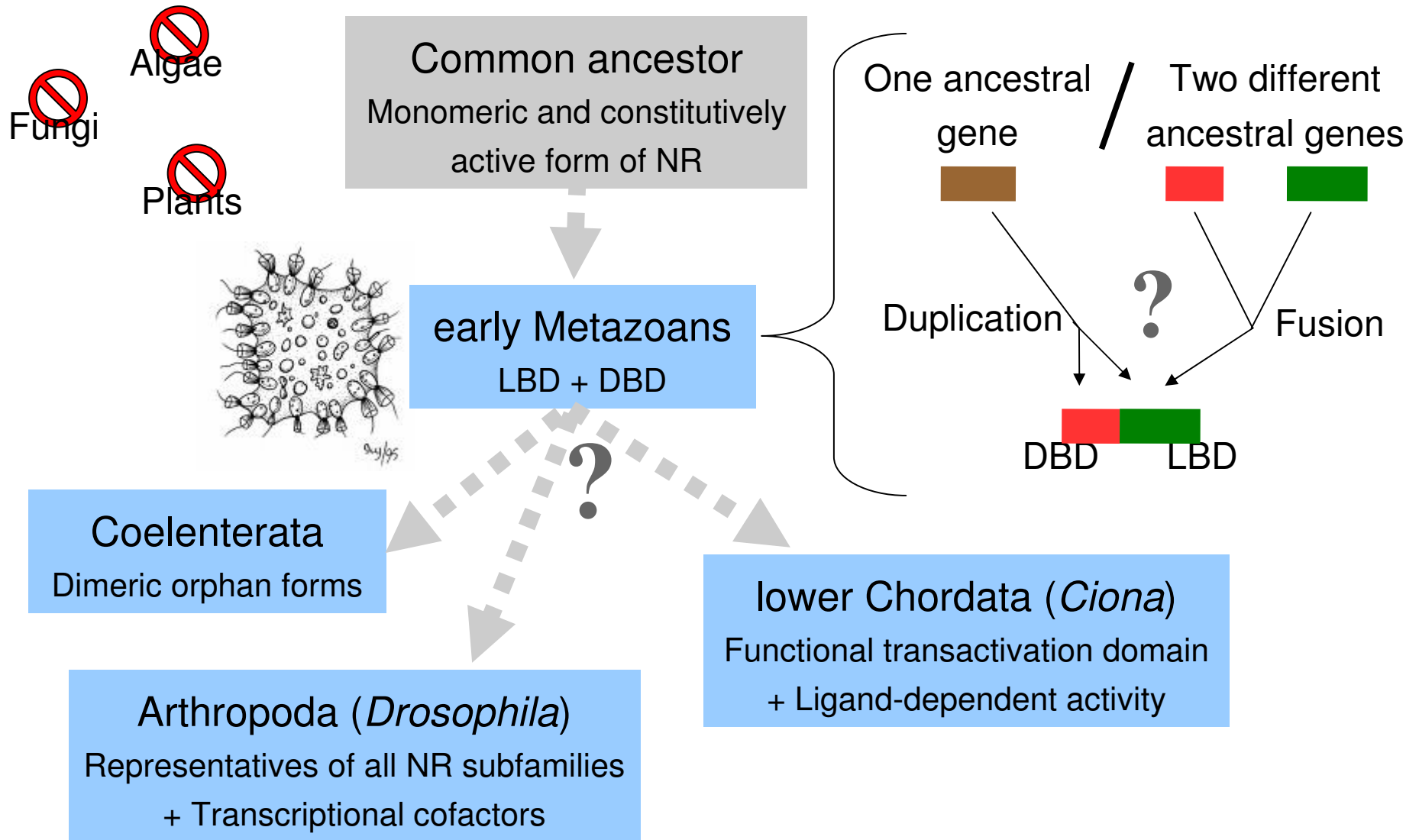
- Cytoplasmic regulation mechanism:



Regulation and mode of action



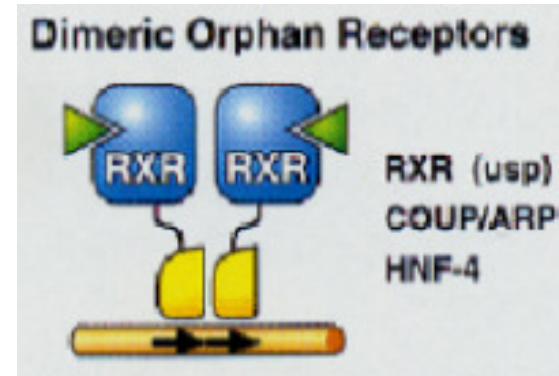
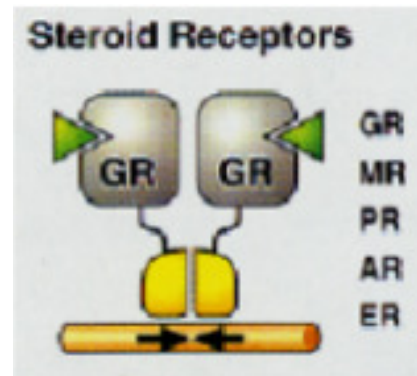
Evolution



Former classification

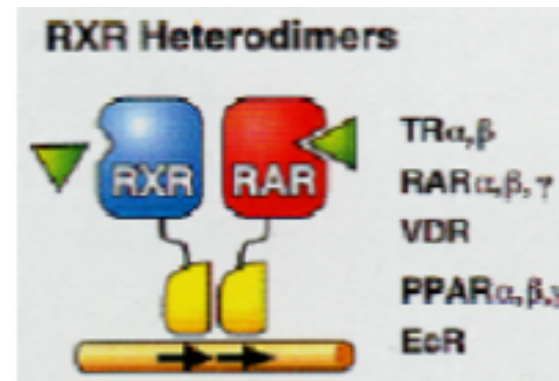
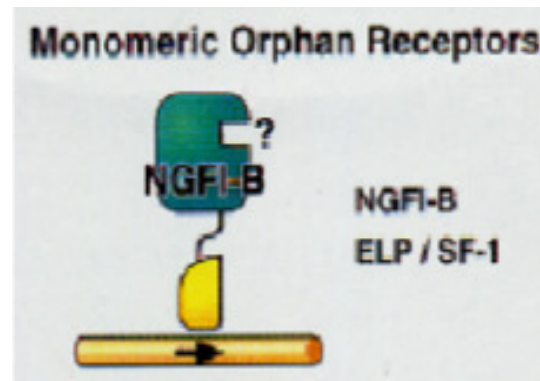
- The NR superfamily have been classified into **four subfamilies** based on their DNA-binding, ligand-binding and dimerisation properties:

Homodimeric
Ligand-dependent
Head-to-head



Homodimeric
Unknown ligand
Head-to-tail

Monomeric
Unknown ligand



Heterodimeric
Ligand-independent
Head-to-tail



Receptors & Ligands

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



Receptors & Ligands

Name	Abbreviation	Nomenclature	Ligand
ErbA2-related gene-2	EAR2	NR2F6	Orphan
Oestrogen receptor	ER α	NR3A1	Oestradiol-17 β , tamoxifen, raloxifene
	ER β	NR3A2	Oestradiol-17 β , various synthetic compounds
Oestrogen receptor-related receptor	ERR α	NR3B1	Orphan
	ERR β	NR3B2	DES, 4-OH tamoxifen
	ERR γ	NR3B3	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



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 α)
 - 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

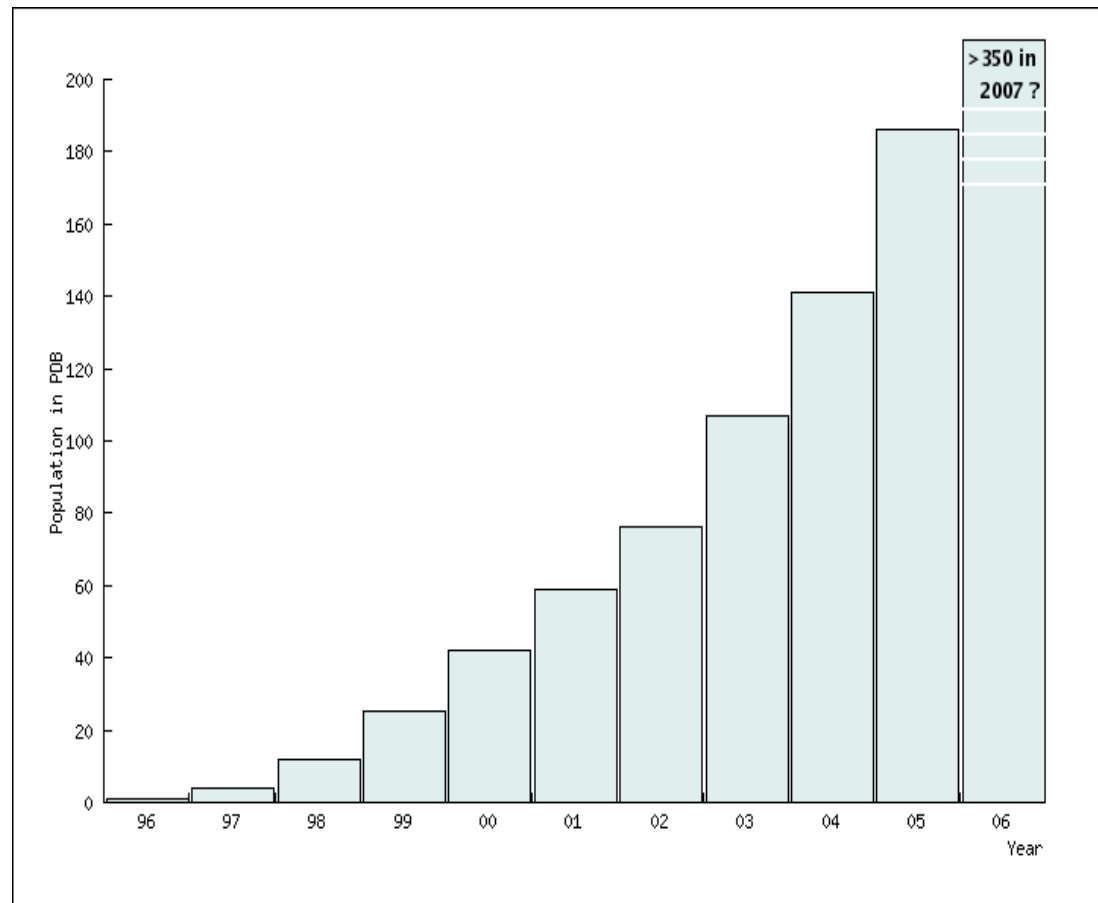
- Since isolation and cloning of the first nuclear receptor in 1985, a large number of NRs have been identified. But only a part of the current subfamilies are well represented in the Protein Data Bank.

- Structures are obtained from the LDB or the DBD, but not from the full protein.

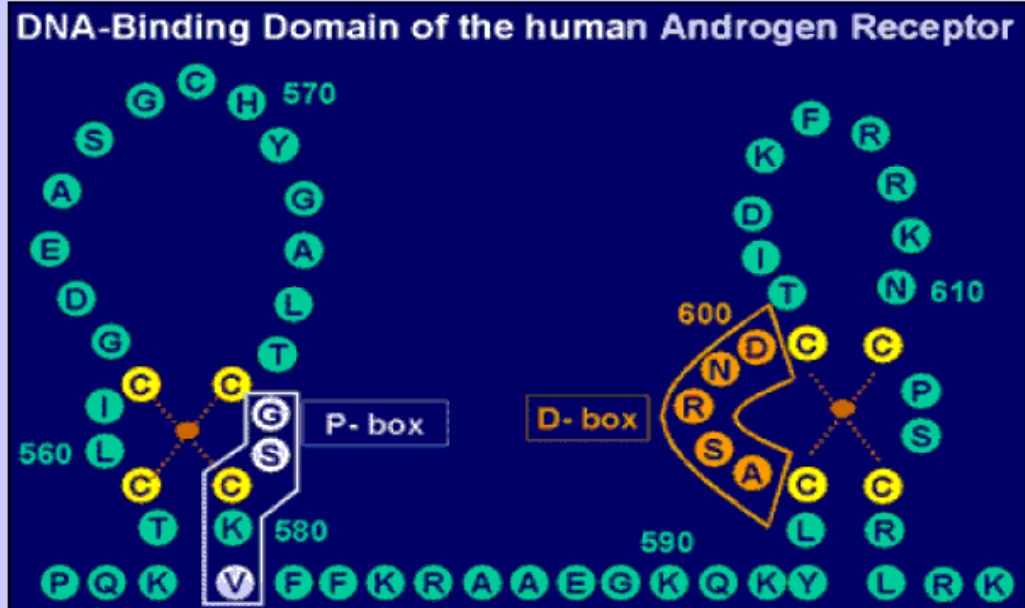
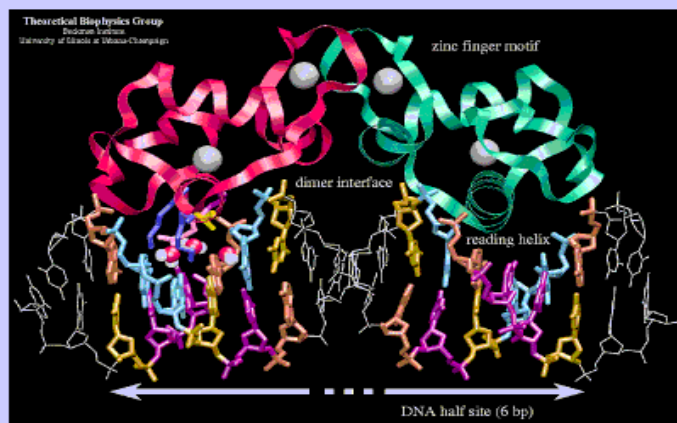
- Many new NRs are temporarily classified as Orphan receptors.

Adapted from a graphic of annual growth for Nuclear Receptor entries in the PDB, reported by FCP web page

<http://cgl.imim.es/fcp/>

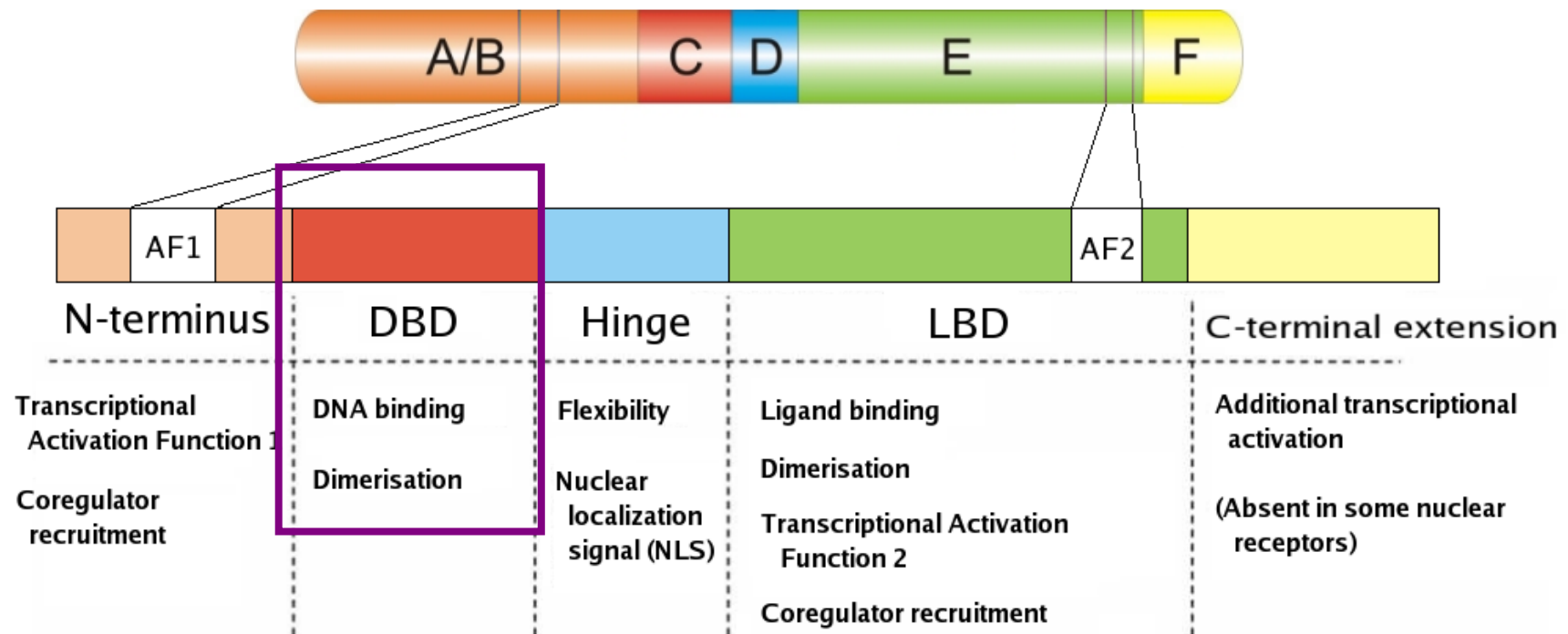


DNA Binding Domain



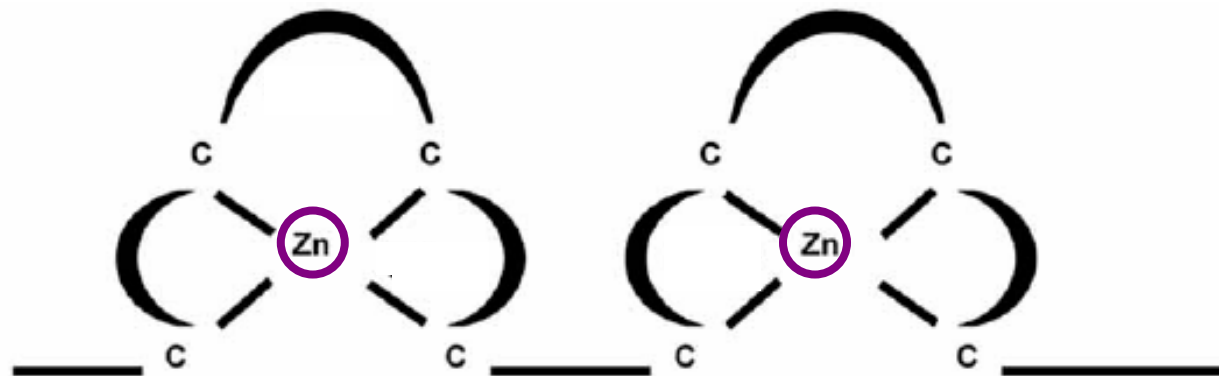
DNA Binding Domain

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

RPCLVCSDEASGCHYGVLTCGSKVFFKRAVEGQHNYLCAGR
--STTT-S---EEETTEE-**HHHHHHHHH**TS-----SSSS

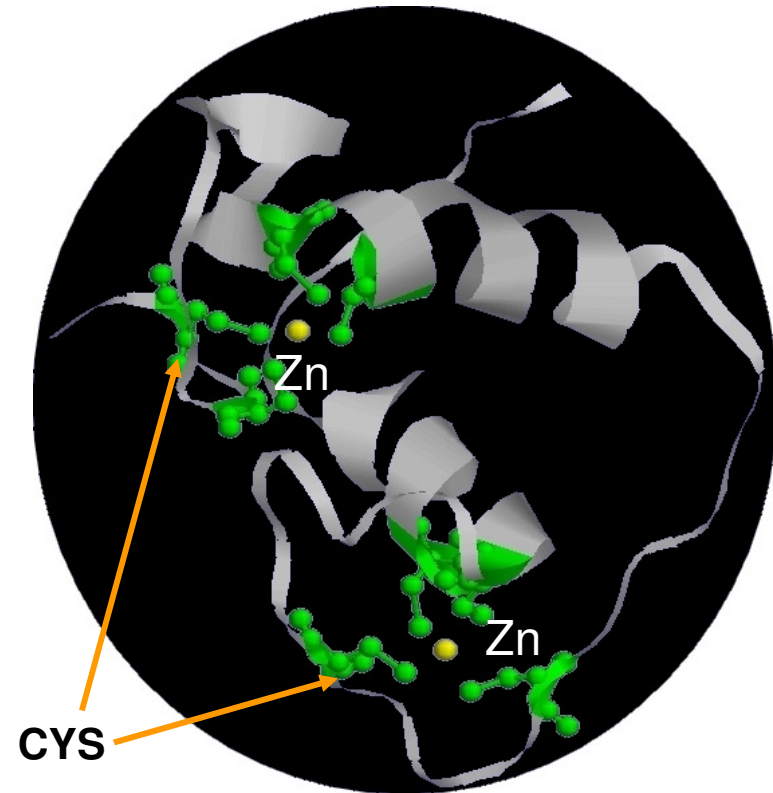
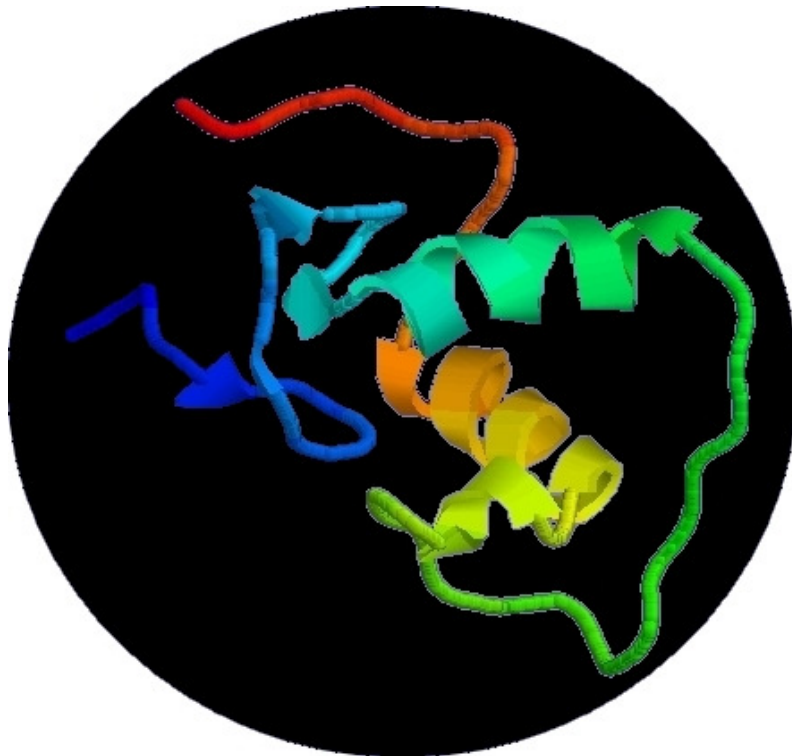
DCIIDKIRRKNCPACRYRKCLQAGMNLEAR
-----TTTTTT-**HHHHHHHHH**S-----

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

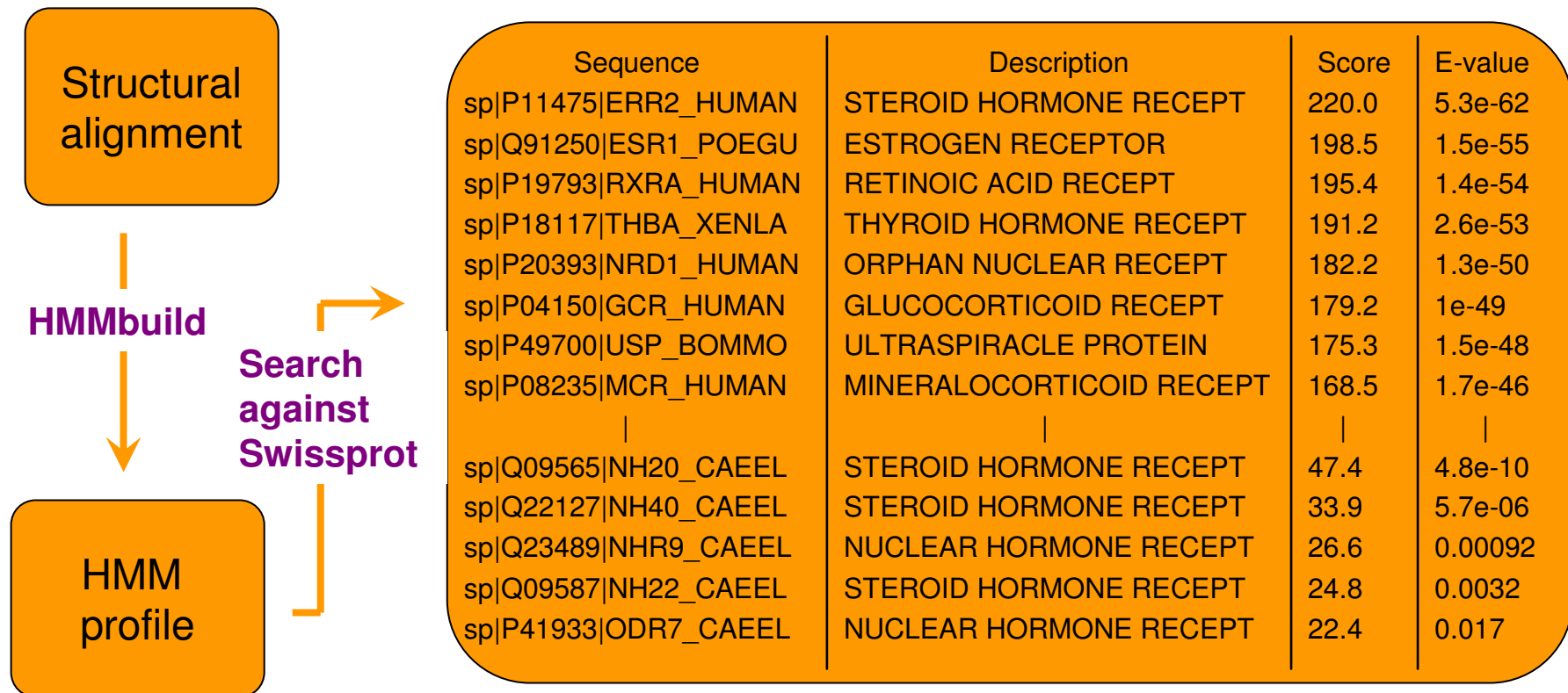
ERbeta  AIPKRLCLVCGDIASGYHYGVASCEACKAFFKRTIQGN--IEYS
ER       MKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGH--NDYM
GR       MKPARPCLVCSDEASGCHYGVLTGSCCKVFFKRAVEGQ--HNYL
RAR      -----PCFVCQDKSSGYHYGVSACEGCKGFFRRSIQKN--MVYT
RXR      -FTKHI CAICGDRSSGKHYGVYSCEGCKGFFKRTVRKD--LTYT
TR       ---DEL CVVCGDKATGYHYRCITCEGCKGFFRRTIQKNLHPSYS
VDR      ----LLCKVCGDVASGFHYGVLACEGCKGFFRRSIQQN-IQYKR
          *  . * *  . . * **      . *   **  ** . * . .

ERbeta  CPATNECEITKRRRKSCQACRFMKALKVGMLKEGVRLDRVRGGR
ER       CPATNQCTIDKNRRKSCQACRLRKCYEVGMMKG-----
GR       CAGRNDCIIDKIRRKNCPACRYRKCLQAGMNLEARKTKK-----
RAR      CHRDKNCIINKVTRNRCQYCRLOKCFEVGMSKESVRND-----
RXR      CRDNKDCLIDKRQRNRCQYCRYQKALAMGMKREAVQEERQRG--
TR       CKYEGKCVIDKVTRNQCECRFKKCIYVGMATDLVLDDSKRLAK
VDR      CLKNENCSIVRINRNRQQQCRFKKCLSVGMSRDAVRFGR-----
          *      * * .  *  *  **   * .   **
```



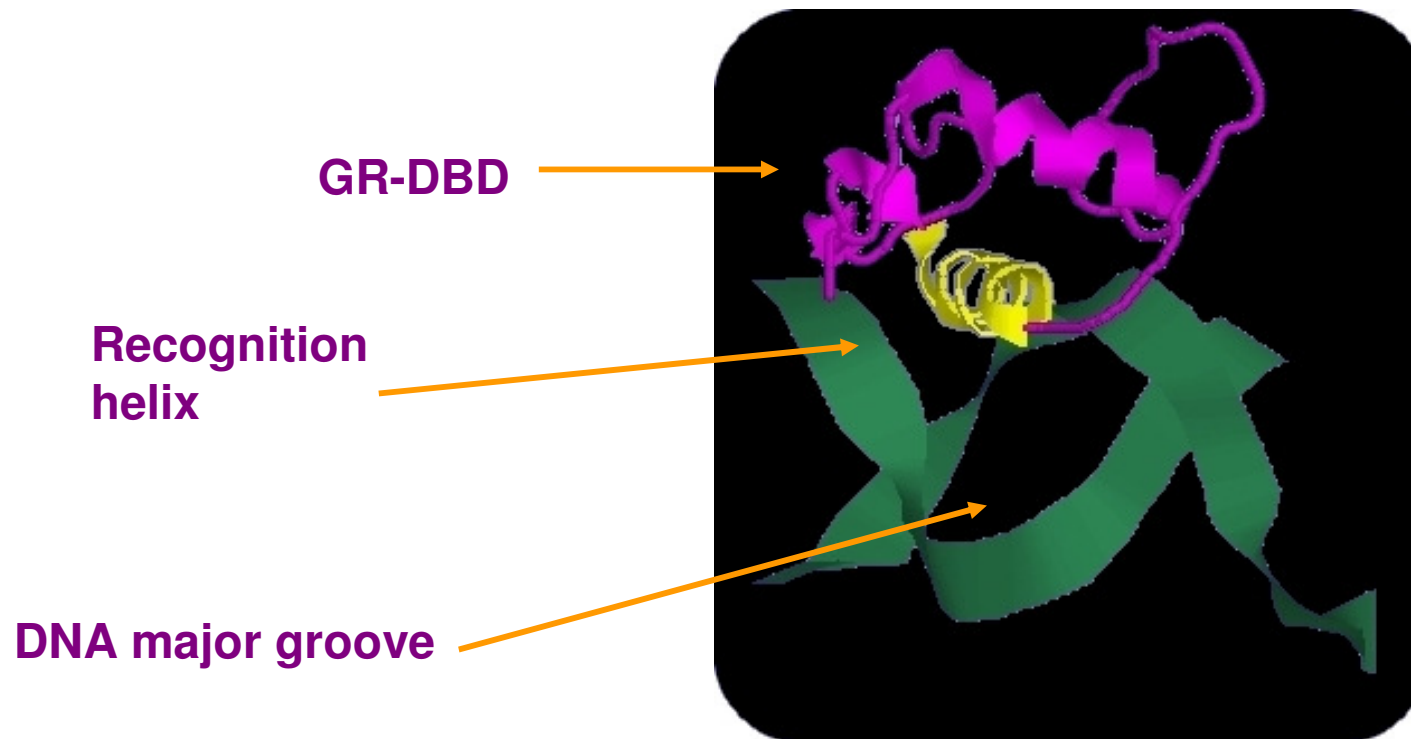
Characterization of the NRs

The DNA binding domain characterizes the family of nuclear receptors.



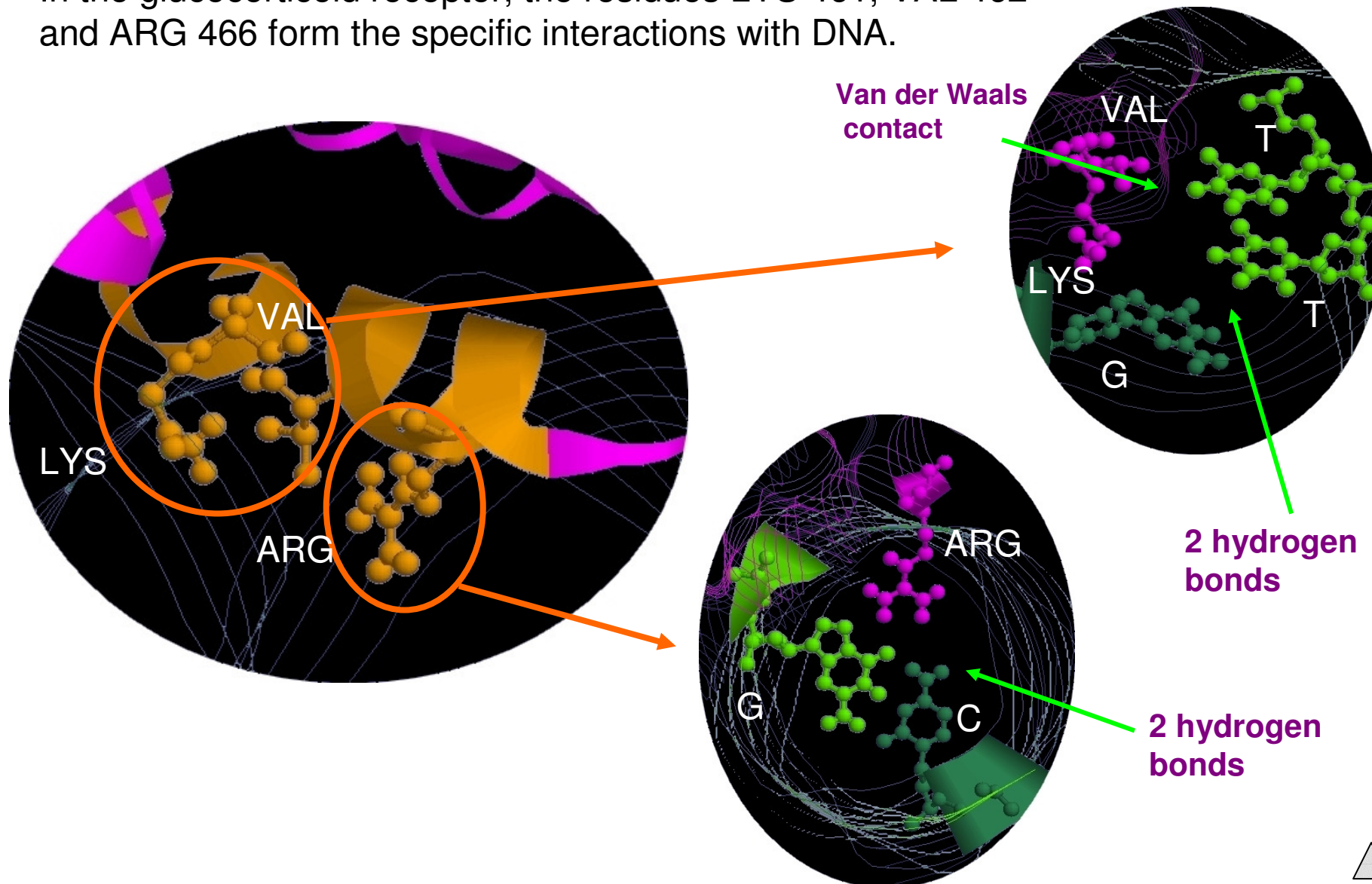
Interactions DNA - DBD

- The first α -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.



Conservation of the Specific contacts

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

CLUSTAL W multiple sequence alignment
DNA BINDING DOMAIN

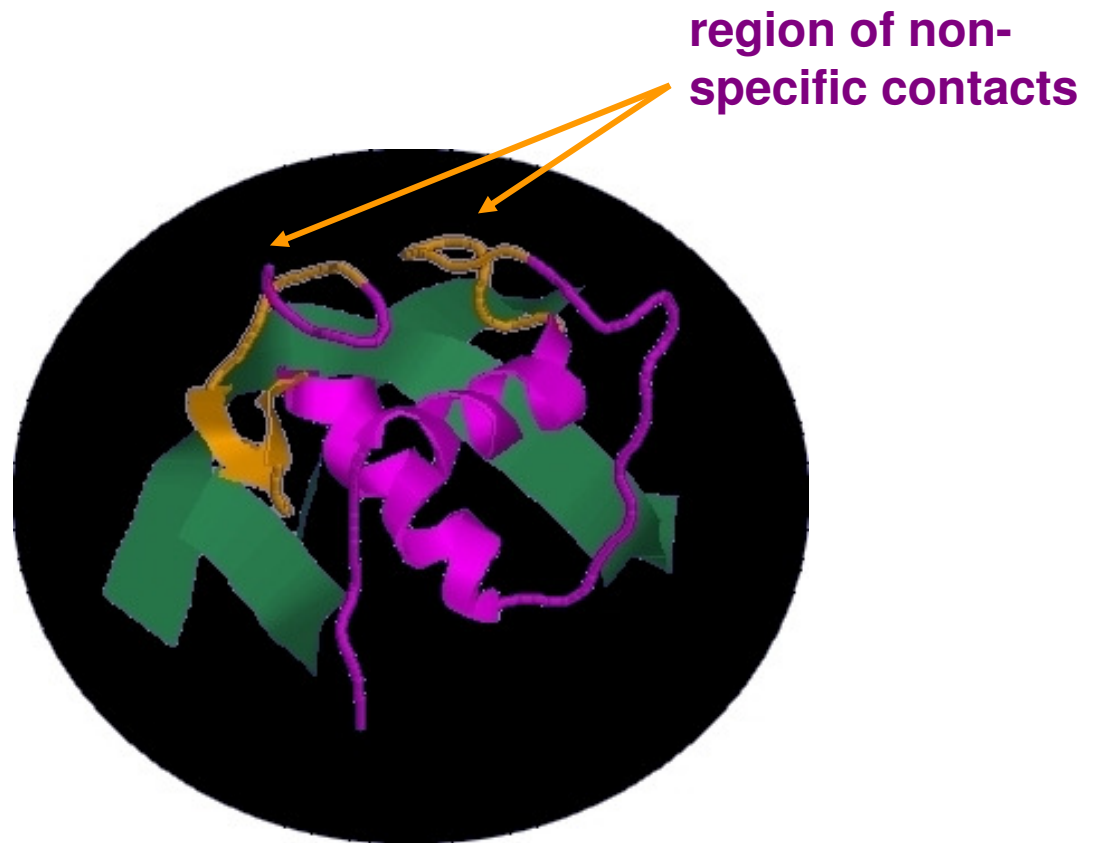
```
ERbeta  AIPKRLCLVCGDIASGYHYGVASCEACKAFFKRTIQGN--IEYS
ER      MKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGH--NDYM
GR      MKPARPCLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQ--HNYL
RAR     -----PCFVCQDKSSGYHYGVSACEGCKGFFRRSIQKN--MVYT
RXR     -FTKHICAICGDRSSGKHYGVSCEGCKGFFKRTVRKD--LTYT
TR      ---DELCVVCGDKATGYHYRCITCEGCKGFFRRTIQKNLHPSYS
VDR     ----LLCKVCGDVASGFHYGVLACEGCKGFFRRSIQQN-IQYKR
          *  . * *  . . * **      . *  **  ** . * . .
```

```
ERbeta  CPATNECEITKRRRKSCQACRFMKALKVGMLKEGVRLDRVRGGR
ER      CPATNQCTIDKNRRKSCQACRLRKCYEVGMMKG-----
GR      CAGRNDCIIDKIRRNKCPACRYRKCLQAGMNLEARKTKK-----
RAR     CHRDKNCIINKVTRNRCQYCRLQKCFEVGMSKESVRND-----
RXR     CRDNKDCLIDKRQRNRCQYCRYQKALAMGMKREAVQEERQRG--
TR      CKYEGKCVIDKVTRNQCECRFKKCIYVGMATDLVLDDSKRLAK
VDR     CLKNENCIVRINRNRCQQCRFKKCLSVGMSRDAVRFGR-----
          *      * * . * * **  * .  **
```



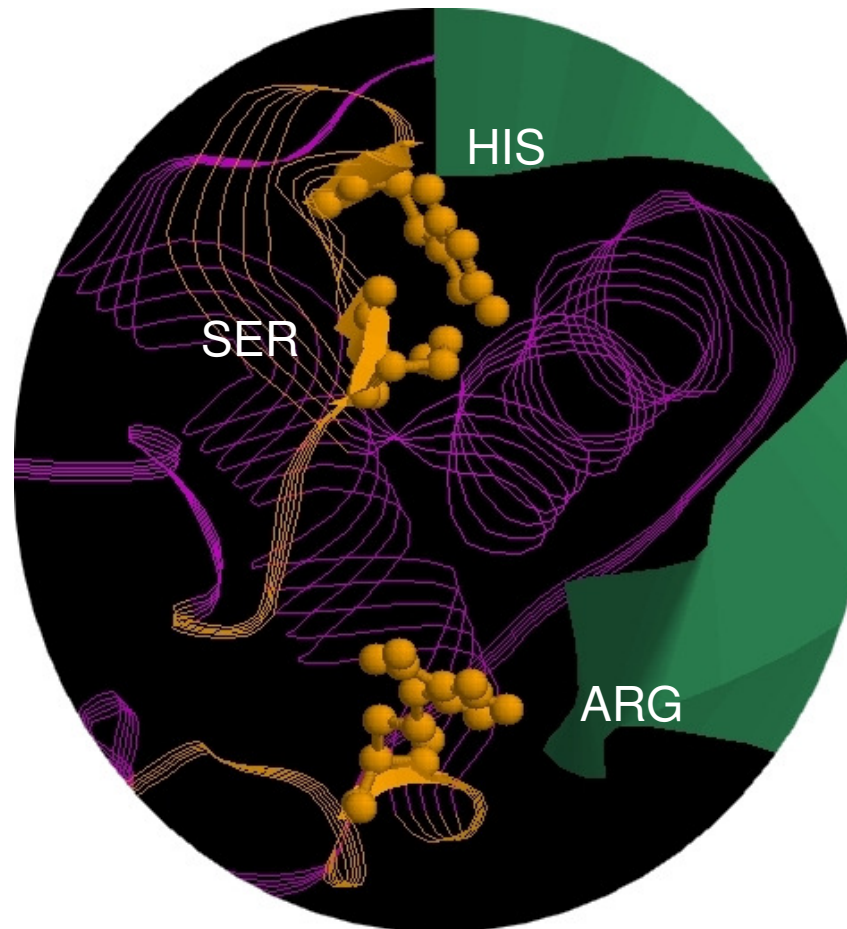
Non-specific Interactions

- 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	SGYHYG	VASCEACKAFFKRT	IQGN--IEYS
ER	MKETRYCAVCNDYA	SGYHYG	VWVSCGCKAFFKRS	IQGH--NDYM
GR	MKPARPCLVCSDEA	SGCHYG	VLTCGSCKVFFKRAVEGQ	--HNYL
RAR	-----PCFVCQDKS	SGYHYG	VSAEGCKGFFRRS	IQKN--MVYT
RXR	-FTKHICAI	CGDRS	SGKHYG	VYSCEGCKGFFKRTVRKD--LTYT
TR	---DEL	CVVCGDKA	TGYHYR	CITCEGCKGFFRRTIQKNLHPSYS
VDR	---LLCKV	CGDVA	SGFHYG	VLACEGCKGFFRRSIQQN-IQYKR
		* . * * . . * **	. * ** ** . * . .	
ERbeta	CPATNECEI	TKRR	RKSCQACRFMKALKVGMLKEGVRLDRVRGGR	
ER	CPATNQCTI	DKNR	RKSCQACRLRKCYEVGMMKG-----	
GR	CAGRNDCI	IDKIR	RKNCPACRYRKCLQAGMNLEARKTKK-----	
RAR	CHRDKNCI	IINKVT	RNRCQY	CRLQKCFEVGMSKESVRND-----
RXR	CRDNKDCL	IDKRQ	RNRCQY	CRYQKALAMGMKREAVQEERQRG--
TR	CKYEGKCV	IDKVT	RNQCQE	CRFKKCIYVGMATDLVLDDSKRLAK
VDR	CLKNENC	SIVRIN	RNRCQQ	CRFKKCLSVGMSRDAVREGR-----
		* * * . * **	* . **	



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-IPKRLCLVCGDIASGYHYGVASCEACKAFFKRTIQG--NI-EYS
RXR      -F-TKHICAI CGDRSSGKHYGVYSCEGCKGFFKRTVRK-D-L-TYT
GR       -MKPARPCLVCSDEASGCHYGVLTCEGCKVFFKRAVE-G-QH-NYL
VDR      -----LLCKVCGDVASGFHYGVLAEGCKGFFRRTSIQQ-N-IQYKR
RAR      -----PCFVCQDKSSGYHYGVSAEGCKGFFRRTSIQKN-M-V-YT
TR       ----DEL CVVCGDKATGYHYRCITCEGCKGFFRRTIQKNLHPS-YS

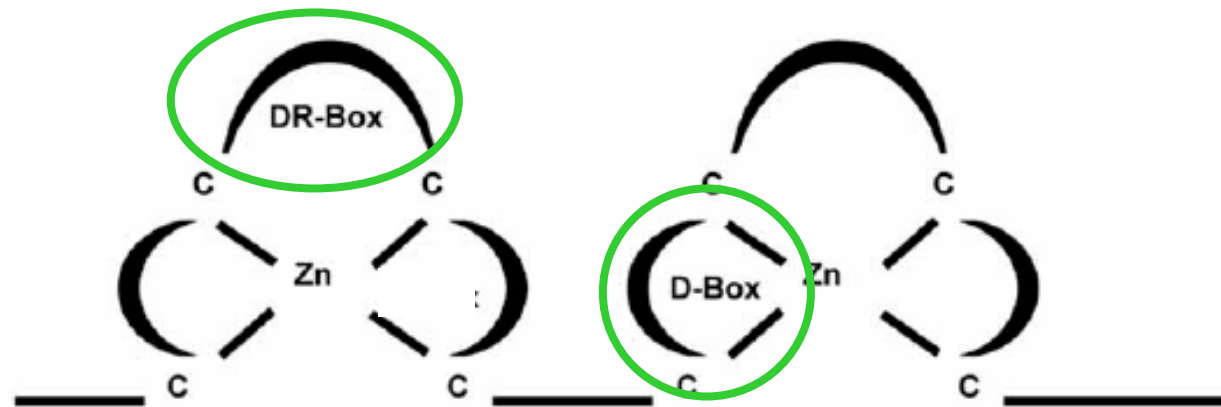
ER      CPATNQCTI--DKNRRKSCQACRLRKCYEVGMM-KG-----
ERbeta  CPATNECEITKRR--KSCQACRFMKALKVGMLKE-G-V-RLDRVR
RXR      CRDNKDCLIDKRQR--NRCQYCRYQKALAMGMKREAVQEER-Q---
GR       CAGRNDCIIDKIR--KNCPACRYRKCLQAGMNLE--A-R-KT---
VDR      CLKNENC SIVRINR--NRCQQCRFFKKCLSVGMSRDA-V-R-F----
RAR      CHRDKNCIINKVTR--NRCQYCR LQKCFEVGMSKES-V-R-N----
TR       CKYEGKCVIDKVTR--NQCQECRFKKCIYVGMATDL-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 β turn in the dimer.

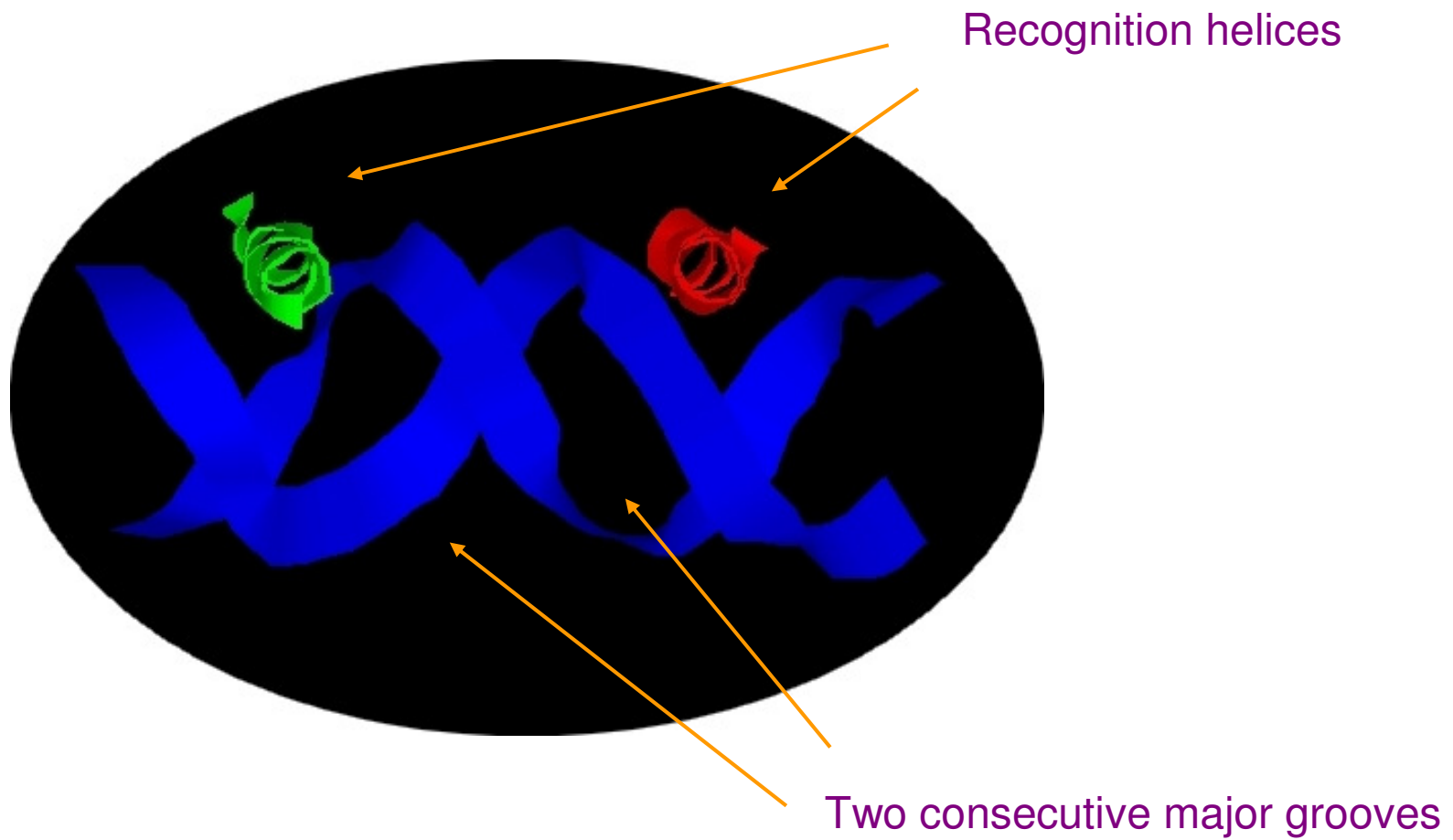


Dimerization

- The two nuclear receptors in the dimer bind to the DNA backbone through **specific and non-specific interactions**, as describe 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**.

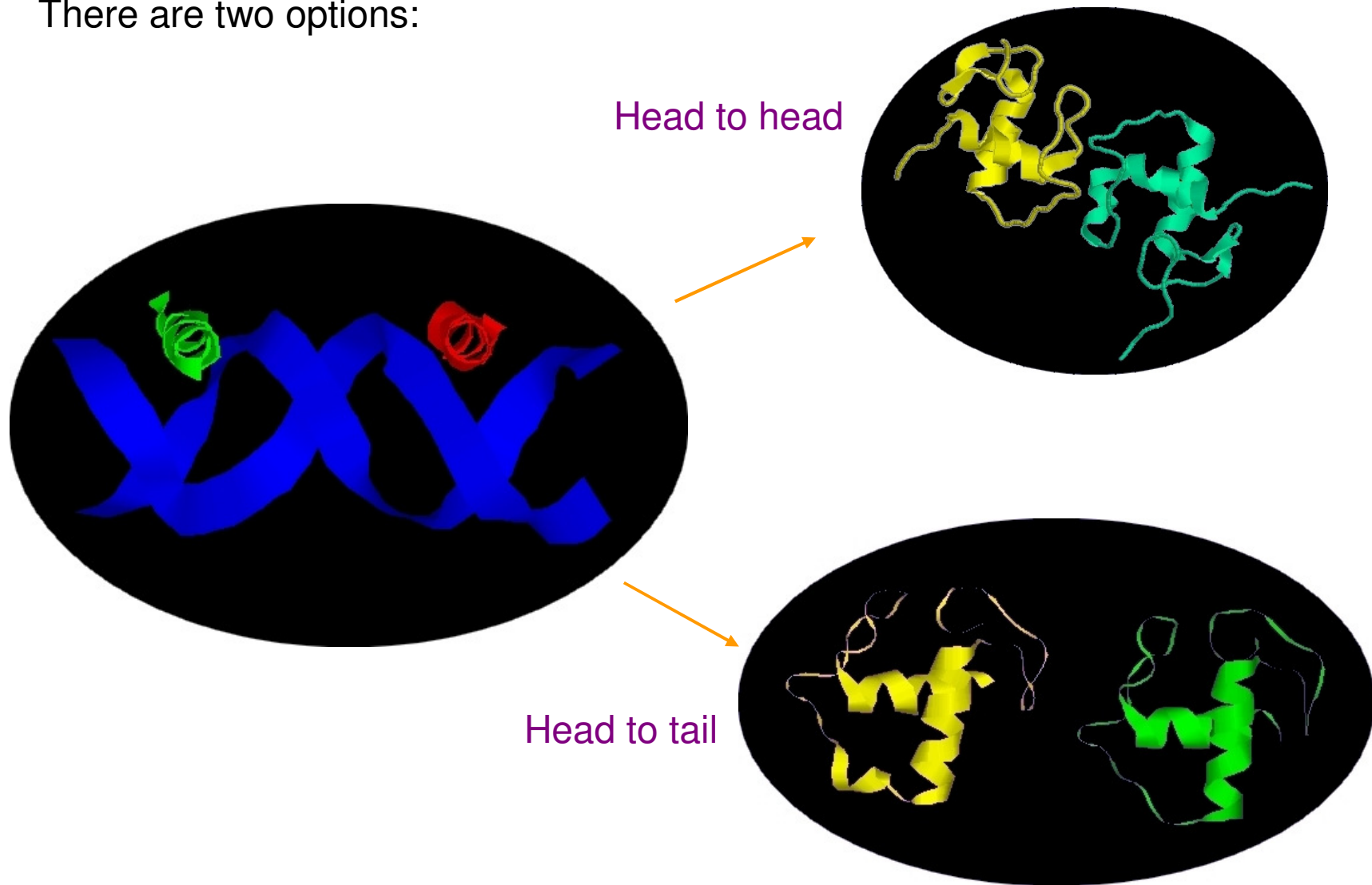


Dimerization



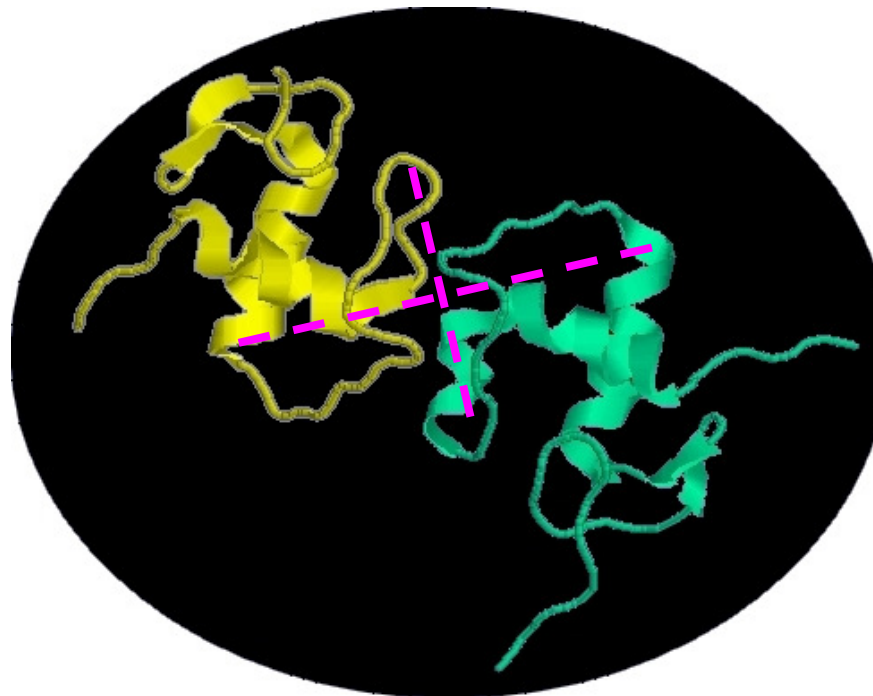
Dimerization

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:

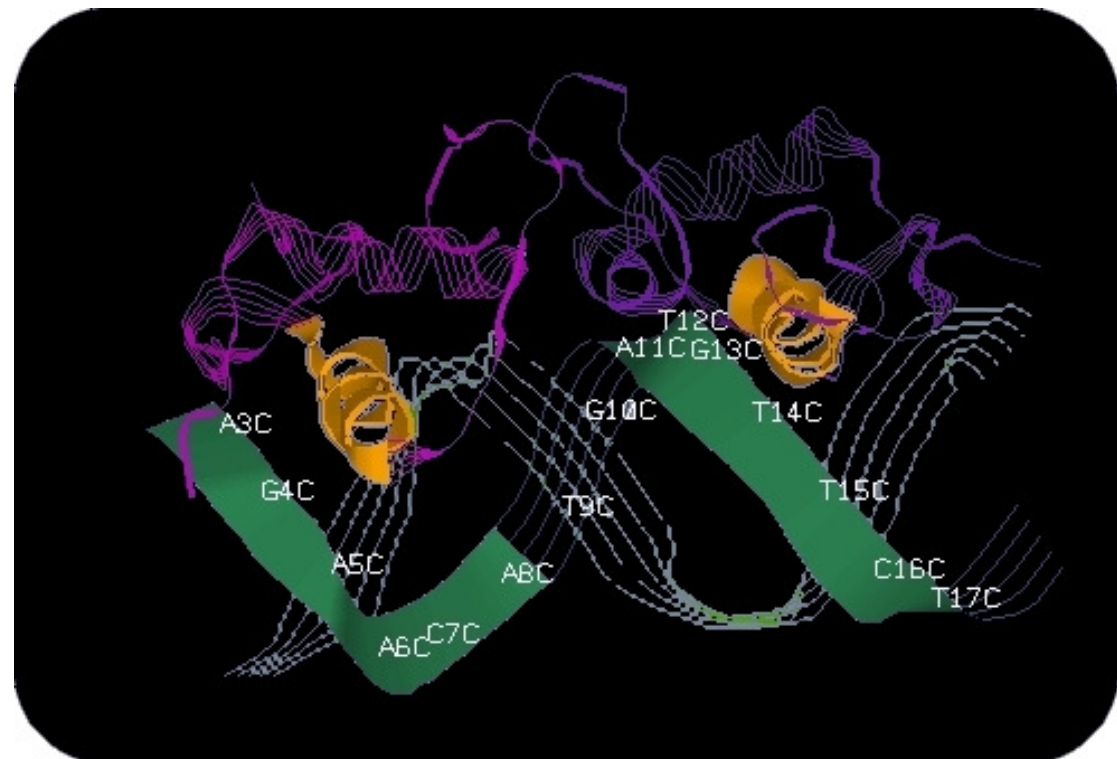


Homodimerization

- 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_{xx}TGTTCT 3'
3' TCTTGTxxxACAAGA 5'

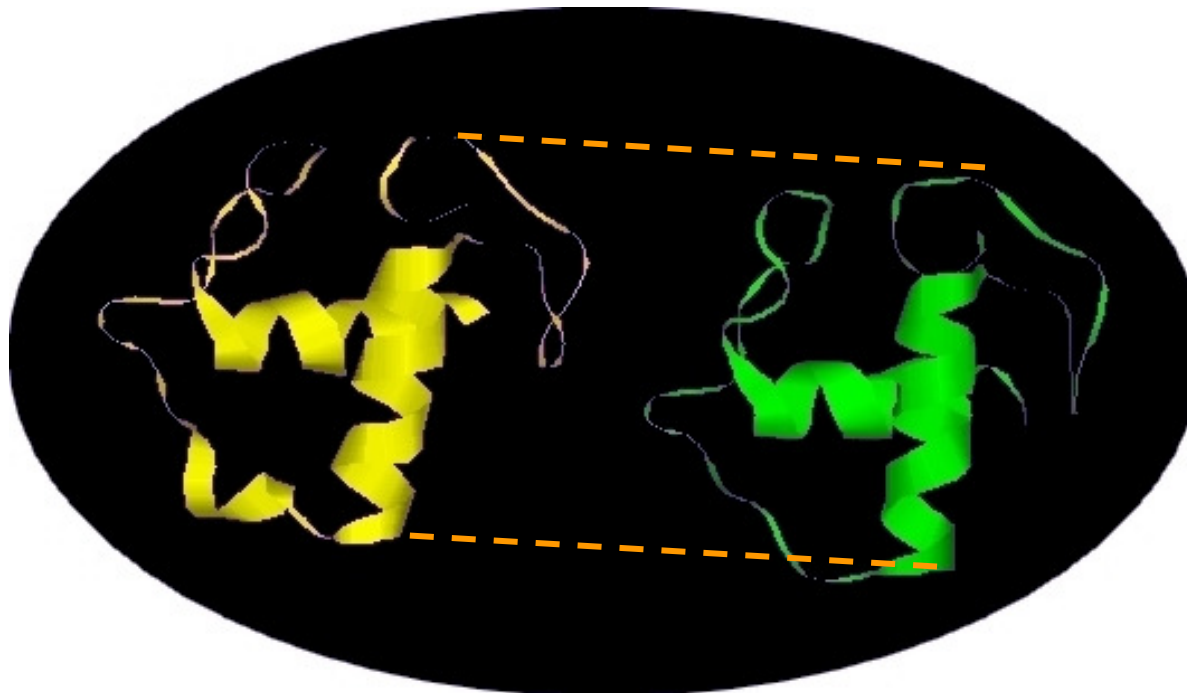


AGAACA xxx TGTTCT



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:

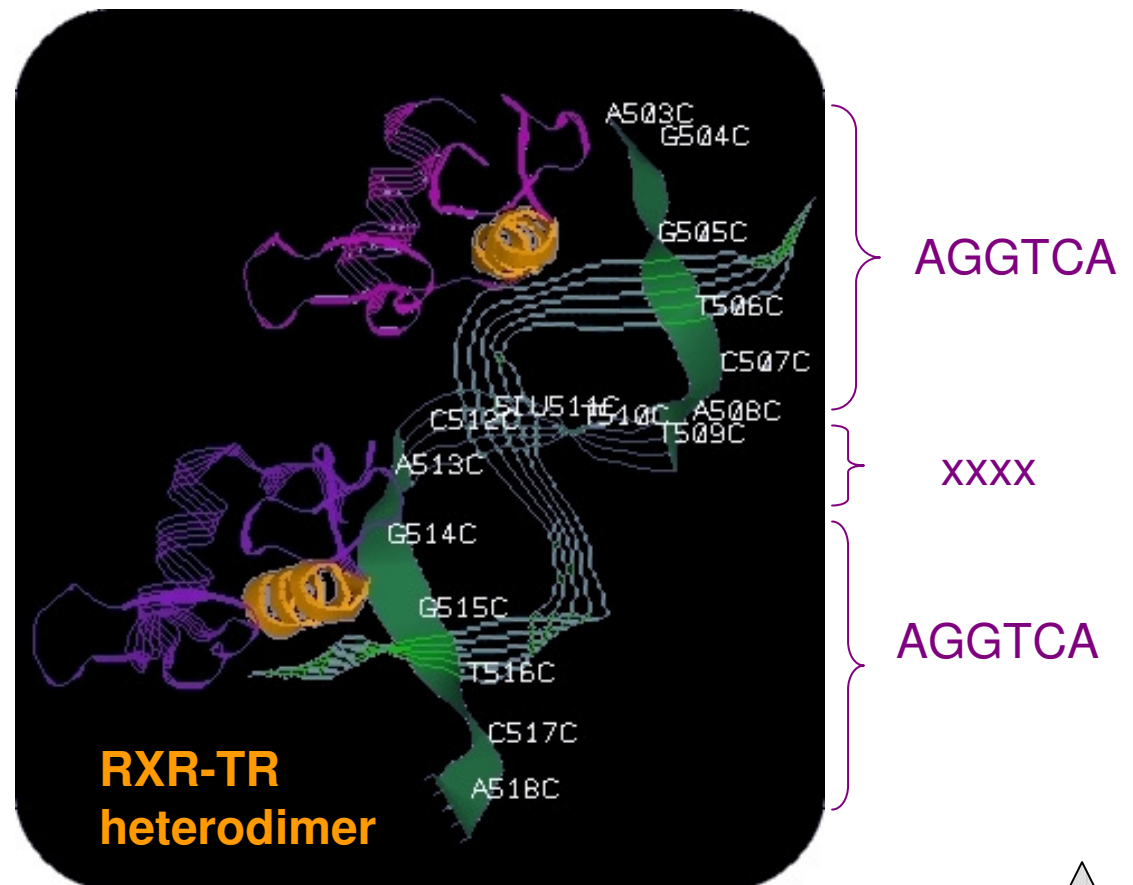


Heterodimerization

- 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:
AGGTCAxxxAGGTCA
TCCAGTxxxTCCAGT
- RXR-TR:
AGGTCAxxxxAGGTCA
TCCAGTxxxxTCCAGT
- RXR-RAR:
AGGTCAxxxxAGGTCA
TCCAGTxxxxTCCAGT



ClustalW alignment

ClustalW alignment with one member of each nuclear receptor family.

```
Q5VYG4|Q5VYG4_HUMAN      MASFTKHICAICGDRSSGKHYGVYSCGCKGFFKRTVRKD--LTYTCRDNKD--CLIDKR 182
Q505F1|TR2_MOUSE          GPNKVFDLCVVGCDKASGRHYGAITCEGCKGFFKRSIRKN--LVYSCRGSKD--CVINKH 148
Q9Y466|NR2E1_HUMAN        -----CKVCGDRSSGKHYGVYACDGCSGFFKRSIRRN--RTYVCKSGNQGGCPVDKT 65
P10589|COT1_HUMAN          -----CVVCGDKSSGKHYGQFTCEGCKSFFKRSVRRN--LTYTCRANRN--CPIDQH 133
P41235|HNF4A_HUMAN        -----ALCAICGDRATGKHYGASSCDGCKGFFRRSVRKN--HMYSRFSRQ--CVVDKD 98
P10276|RARA_HUMAN          -----CFVCQDKSSGYHYGVSAACGCKGFFRRSIQKN--MVTYCHRDKN--CIINKV 135
P22829|NR4A1_RAT          SSGGSSEGRCAVCGDNASCQHYGVRTCEGCKGFFKRTVQKS--AKYICLANKD--CPVDKR 313
P55055|NR1H2_HUMAN        -----LCRVCGDKASGFHYNVLSCEGCKGFFRRSVVRGGARRYACRGGGT--CQMDAF 136
P11473|VDR_HUMAN          -----RICGVCGRATGFHFNAMTCEGCKGFFRSMKRRK--ALFTCPFNGD--CRITKD 71
P45448|NR5A2_MOUSE        YDEDLEELCPVCGDKVSGYHYGLLTCEGCKGFFKRTVQNN--KRYTCIENQN--CQIDKT 154
P37243|THB2_HUMAN          -YLDKDELGVVCGDKATGYHYRCITCEGCKGFFRRTIQKNLHPSYSCKYEGK--CVIDKV 171
Q14995|NR1D2_HUMAN        ----MVLCKVCGDVASGFHYGVHACEGCKGFFRRSIQNN--IQYKKCLKNEN--CSIMRM 151
P11474|ERR1_HUMAN          LSSLPKRLCLVCGDVASGYHYGVASCEACKAFFKRTIQGS--IEYSCPASNE--CEITKR 223
P03372|ESR1_HUMAN          ESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGH--NDYMCPATNQ--CTIDKN 232
P37231|PPARG_HUMAN        SNSLMAIECRVCGDKASGFHYGVHACEGCKGFFRRTIRLK---LIYDRCDLN--CRIHKK 185
Q15406|NR6A1_HUMAN        -----CLICGDRATGLHYGIISCEGCKGFFKRSICNK--RVYRCSRDKN--CVMSRK 107
P35398|RORA_HUMAN          -AQIEIIPCKICGDKSSGIHYGVITCEGCKGFFRRSQQSN--ATYSCPRQKN--CLIDRT 153
P04150|GCR_HUMAN          TTGPPPKLCLVCSDEASGCHYGLTCGSKVFFKRAVEGQ--HNYLCAGRND--CIIDKI 468
                          * : * * : * : * : * : * :
Q5VYG4|Q5VYG4_HUMAN      QRNRCQYCRYQKCLAMGMKREAVQEERQRGKDRNENE----- 219
Q505F1|TR2_MOUSE          HRNRCQYCRLLQRCIAFGMKQDSVQCERKPIEVSRKSSNCAASTEKIYIRKDLRSPLAAT 208
Q9Y466|NR2E1_HUMAN        HRNQCRACRLKKCLEVNMMNKDAVQHERGPRTSTIRKQVALYFRGHKEENGAAAHFPSAAL 125
P10589|COT1_HUMAN          HRNQCRQYCRLLKKCLKVGMRREAVQGRMPPTQPNPGQ----- 170
P41235|HNF4A_HUMAN        KRNQCRQYCRLLKKCFRAGMKKEAVQNERDR----- 127
P10276|RARA_HUMAN          TRNRCQYCRLLQKCFEVMGMSKESVRNDRNKKKKEVPKP----- 172
P22829|NR4A1_RAT          RRNRCQFCRFQKCLAVGMVKEVVRTDSLKGRRGRPLPS----- 350
P55055|NR1H2_HUMAN        MRRKCCQQCRLLRKCKEAGMREQCVLSEEQIRKKKIRKQQQQESQSQSQSPVGPQ----- 189
P11473|VDR_HUMAN          NRRHCQACRLKRCVDIGMMKEFILTDEEVQRKREMILKRKEEALKDSLRLPKLSEEQQRI 131
P45448|NR5A2_MOUSE        QRKRCQYCRFKKCIDVGMKLEAVRADMRGGRNKFQPMYKRDRALKQQKALIRANGLKL 214
P37243|THB2_HUMAN          TRNQCECRFKKCIYVGMATDLVLDDSKRLAKRKLIEENREKRRREELQKSIG----- 224
Q14995|NR1D2_HUMAN        NRNRCQQCRFKKCLSVGMSRDVAFGRIPKREKQRMLEMQSAMKTMNSQFSGLQNDT 211
P11474|ERR1_HUMAN          RRAQACQACRFTKCLRVGMLKEGVRLDRVRGGRQKYKRRPEVDP----- 266
P03372|ESR1_HUMAN          RRAQACQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQDDGEGRGVGSAG----- 284
P37231|PPARG_HUMAN        SRNRCQYCRFQKCLAVGMSHNAIRFGRMPQAEKEKLLAEISSDIDQLNPESADLRALAKH 245
Q15406|NR6A1_HUMAN        QRNRCQYCRLLKCLQMGMRNKAIREDMGPMGRNKSIGPVQISEEEIERIMSGQEFEEAN 167
P35398|RORA_HUMAN          SRNRCQHCRLLQKCLAVGMSRDVAFGRMSKKQRDSLYAEVQKHRMQQQQRDHQQQPGAE 213
P04150|GCR_HUMAN          RRAQCPACRYRKCLQAGMNLARKTKKKIKGIQQATTG----- 506
                          * . * * * : * . *
```

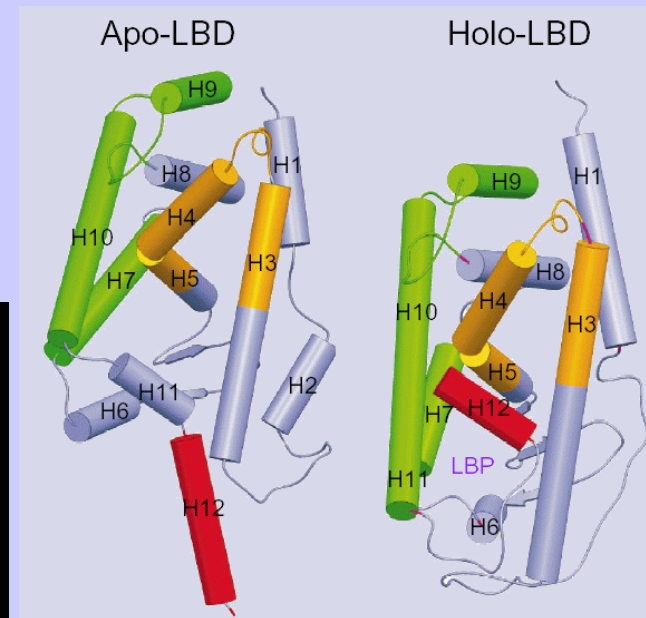
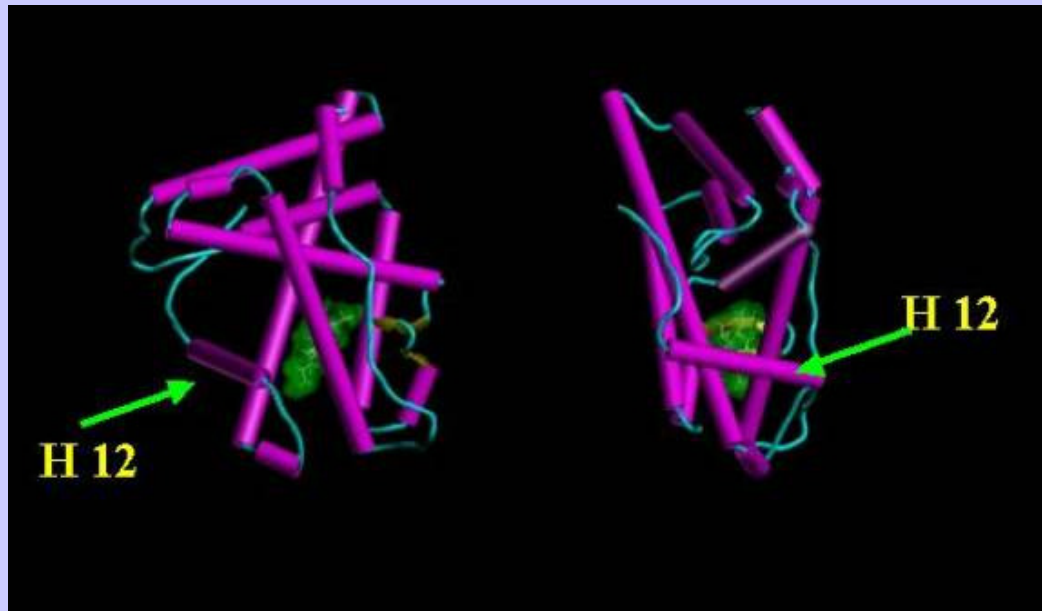


Summary

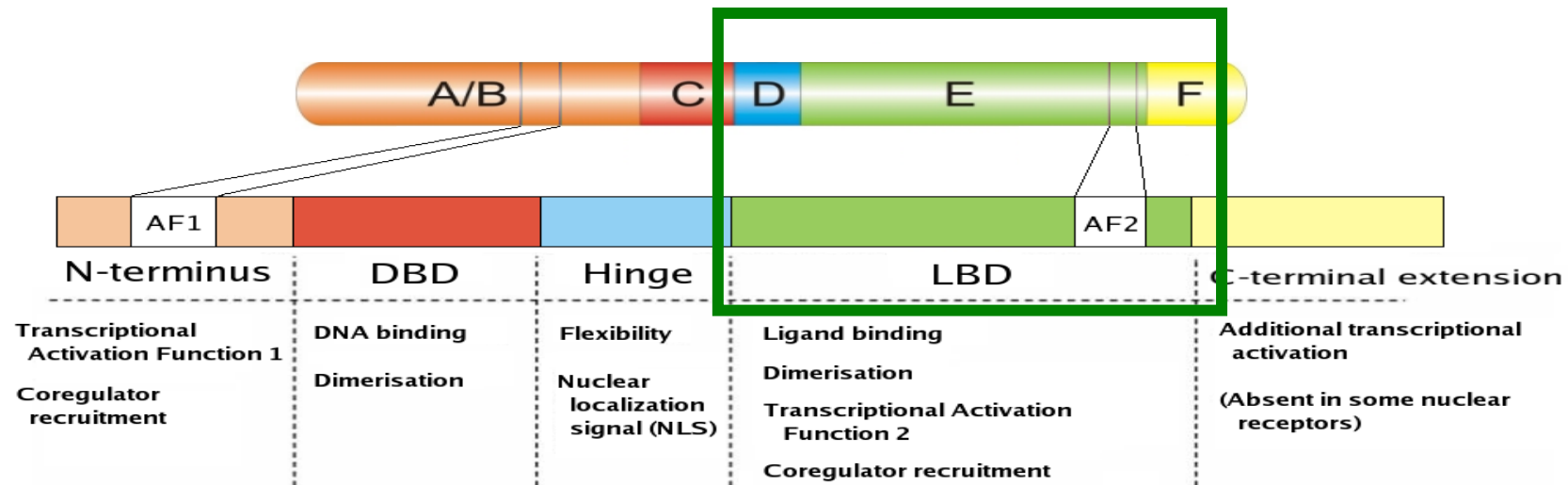
- The DBD consists of about 70 residues that bind to activating elements of DNA called hormone **response elements**.
- In the DBD there are **two zinc containing regions**. Each region binds a zinc atom through four cysteine residues.
- The DBD is a **highly conserved** domain in the family of nuclear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

. **HOLO(+): NR2B1 (RXR)**

Group C: Testicular receptor

. **HOLO(-): ?**

Group E: TLX/PNR

Group F: COUP/EAR

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

. **APO: NR3B3**

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

. **HOLO(+): NR3A1**

Group B: Estrogen related receptor

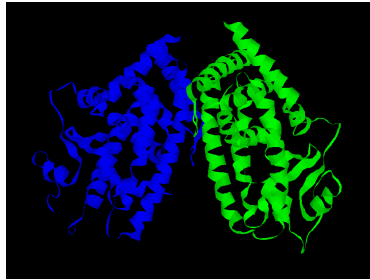
. **HOLO(-): NR3A1**

Group C: 3-Ketosteroid receptors

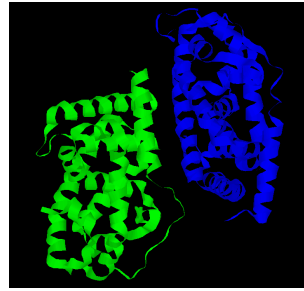


Structural conservation through families

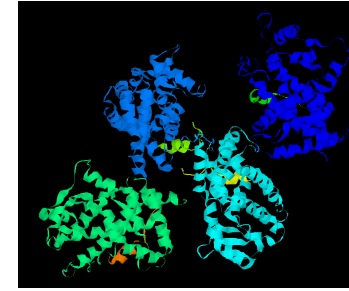
Subfamily 1: Thyroid Hormone Receptor-like



APO: 1prg.pdb (R=2.4 Å)



HOLO(+): 1n46.pdb (R=2.2 Å)

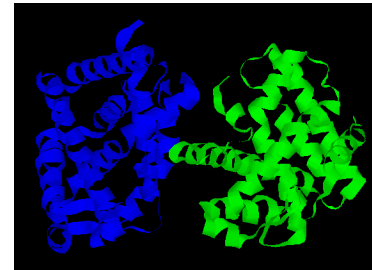


HOLO(-): 1kkq.pdb (R=3.0 Å)

Subfamily 2: Retinoid X Receptor-like

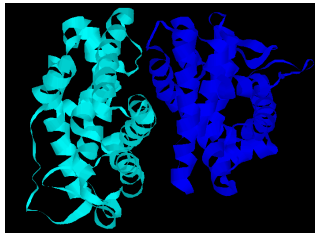


APO: 1lbd.pdb (R=2.4 Å)

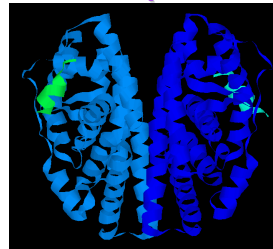


HOLO(+): 1fby.pdb (R=2.25 Å)

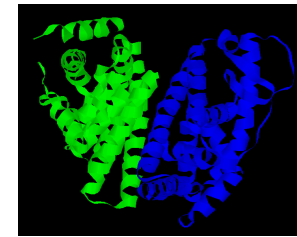
Subfamily 3: Estrogen Receptor-like (Steroid hormone receptor)



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

```
1errA      -----ALSLTADQWYSALLD-----AEPPILYSE
1fbyA      -----SSANEDMPVERILE-----AELAVEP--
1kkqA      -TADKSLAKRIYEAYLKNFNMNKVKARVILSGKASNNPPFVIHDMETLCMAEKTLYAKL
1kv6A      -----NKIVSHLLV-----AEPEKIYAM
1lbd       -----SANEDMPVERILE-----AELAVEPKT
1n46A      -----KPEPTDEEWELIKTYTEAHVATNAQWKQKRLP-----EDIGQAPIY
1prgA      ESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMNGEDKIKFKH
3erdA      -----SLALSLTADQWYSALLD-----AEPPILYSE
```

```
1errA      YDPTR-----PFSEASHMGLLTNLADRELVHMINMAKR-VPGFVDLTLDQVHLLLECAW
1fbyA      -----DPVTNICOAADKQFTLVEMAKR-IPHFSELPDDQVILLRAGW
1kkqA      YANG-----IQNKEAEVRI FHCCOCTSVETVTELEFAKAI PGFANLDNDQVTLKYG
1kv6A      PDPT-----VPDSDIKALTTLCDLADRELVVIGMAKH-IPGFSTLSLADQMSLLQSAW
1lbd       ETYVEANMGLNPSSPNDPVTNICOAADKQFTLVEMAKR-IPHFSELPDDQVILLRAGW
1n46A      NAPEG-----GKVDLEAFSHFTKIITPAITRVYDFAKK-LPMFCELPCEQIILLKGCC
1prgA      ITPLO---EQSKEYAIRIFQGCQFRSVEAVQEITEYAKSIPGFVNLNDQVTLKYG
3erdA      YDPTR-----PFSEASHMGLLTNLADRELVHMINMAKRVPGFVDLTLDQVHLLLECAW
```

```
1errA      LEILMIGLWRSMEHPGKLLFAPNLLDRNOGKCEGHWEIFD-MLLATSSRFMMNLQG
1fbyA      NELLIASFSHRSIAYKDGILLATGLHVHRN-SAHSAVGAI FDRVLTLYSKMRDNQMDK
1kkqA      YEAI FANLSSVMNKDGNLYAYGNGFITREFLKSRLKPFCDIME-PKDFAMKFNALDLD
1kv6A      MEILILGVYYSLSFEDELVYADYINDED-QSKLAGLLDLNN-AILQVKKYKSMKLEK
1lbd       NELLIASFSHRSIAYKDGILLATGLHVHRN-SAHSAVGAI FDRVLTLYSKMRDNQMDK
1n46A      MEIWSLRAAVRYDPESETLTNGEMAVTRG-QLKNGGLGVVSD-AIFDLGMSLSSFNLD
1prgA      HEIITYNLASLWMDGYLISEGQGFHTREFLKSRLKPFQDFME-PKFEFAYKFNALDLD
3erdA      LEILMIGLWRSMEHPGKLLFAPNLLDRNOGKCEGHWEIFD-MLLATSSRFMMNLQG
```

```
1errA      EEFVCLKSIILLNSGVYE-----EKDDHIHRVLDKITDTLIHMAKAGTLQOQHQ
1fbyA      TELGCLRAIVLFNPD SKG-----LSNPAEVEALREKYYASLEAYCKHK---YPEQPG
1kkqA      SDISLFVAAIICCGDRPG-----LLNVGHIEKNQEGIVHYLRHLQSN---HPDDIF
1kv6A      EEFVTLKAIALANDSMH-----IEDVEAVOKLQDVLEALQDYEAGQ---HMDPR
1lbd       TELGCLRAIVLFNPD SKG-----LSNPAEVEALREKYYASLEAYCKHK---YPEQPG
1n46A      TEVALLQAVLLMSSDRPG-----LACVERIEKYQDSFLAFEHYINR---KHHVTH
1prgA      SDAIFIAVILSGDRPG-----LLNVKPIEDIQDNLQALELQKLN---HPESQ
3erdA      EEFVCLKSIILLNSGVYFSLSTLKSLEEKDHIHRVLDKITDTLIHMAKAGTLQOQHQ
```

```
1errA      RLAQLLLILSHIRHMSNKGHEHLYSH-----PLYDLLLEMLDAH-----
1fbyA      RFAKLLRLPALRSI GLKCLEHLFFFKLI GOTPIDFLMEMLEAP-----
1kkqA      LFPKLLQKMDLRQLVT---EHAQLVQIIKKTESDAALHPLLQEIYRDNY
1kv6A      RAGKMLHTLPLLRQSTKAVQHFYNIKLEGKVPNHKLFLENLEA-----
1lbd       RFAKLLRLPALRSI GLKCLEHLFFFKLI GOTPIDFLMEMLEAPHMT-
1n46A      FWPKLLNKVTDLRMIGA---CHASRFLHMKVECTELFPFLFLEVFED--
1prgA      LFAKLLQKNTDLROI VT---EHVOLLQVIKKTETDMSLHPLLQEIYKDL-
3erdA      RLAQLLLILSHIRHMSNKGHEHLYSMKCKNVVPLYDLLLEMLDAHRL---
```

- Homology or Remote homology?

All human



NR3A1 sequence and structural alignment

CLUSTAL W(1.60) multiple sequence alignment

```

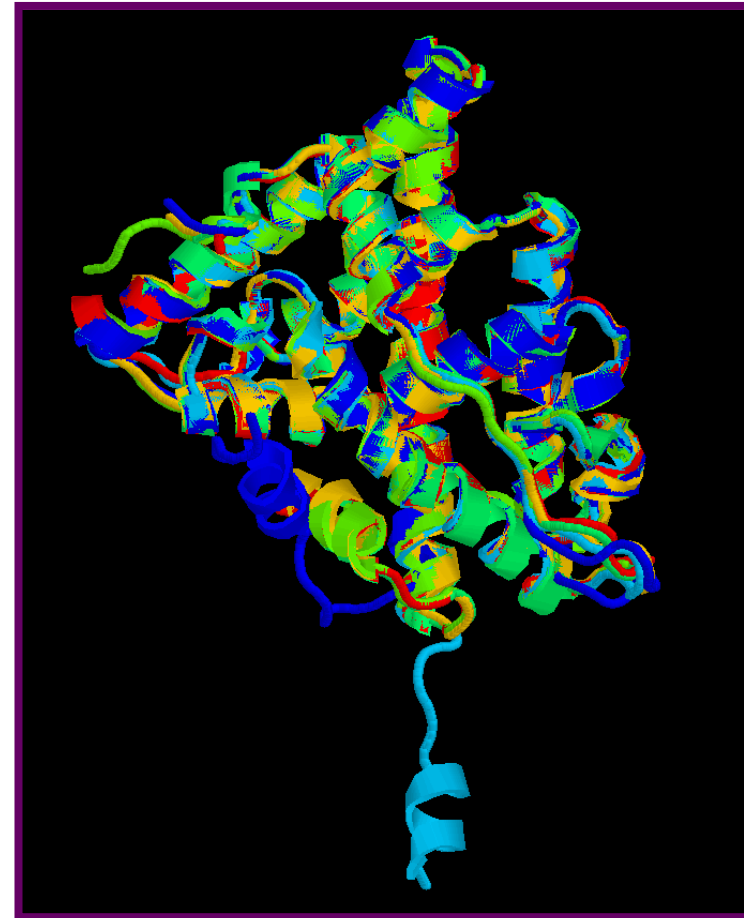
1g50A      ---NSLALSLTADQMVSALLDAEPPILY-SEYDPTRPFSEASMMGLLTNLADRELVHMIN
1gwqA      SKKNSLALSLTADQMVSALLDAEPPILY-----SEPFSEASMMGLLTNLADRELVHMIN
1qktA      ---NSLALSLTADQMVSALLDAEPPILYS-EYDPTRPFSEASMMGLLTNLADRELVHMIN
1l2iA      ----SLALSLTADQMVSALLDAEPPILYSSEYDPTRPFSEASMMGLLTNLADRELVHMIN
1a52A      ----LALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMIN
3erdA      ----SLALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMIN
            *****      *      ,      *****

1g50A      WAKR- VPGFVDLTLDQVHLLC- AWLEILMIGLVWRSMHPGKLLFAPNLLLDNRQGGK
1gwqA      WAKR- VPGFVDLTLDQVHLLC- AWLEILMIGLVWRSMHPGKLLFAPNLLLDNRQGGK
1qktA      WAKR- VPGFVDLTLDQVHLLC- AWLEILMIGLVWRSMHPGKLLFAPNLLLDNRQGGK
1l2iA      WAKRRVPGFVDLTLDQVHLLC- AWLEILMIGLVWRSMHPGKLLFAPNLLLDNRQGGK
1a52A      WAKR- VPGFVDLTLDQVHLLC- AWLEILMIGLVWRSMHPGKLLFAPNLLLDNRQGGK
3erdA      WAKRRVPGFVDLTLDQVHLLC- AWLEILMIGLVWRSMHPGKLLFAPNLLLDNRQGGK
            **** *****      ,      *****

1g50A      VEGMVEIFDMLLATSSRFMMNLQGEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHR
1gwqA      VEGMVEIFDMLLATSSRFMMNLQGEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHR
1qktA      VEGMVEIFDMLLATSSRFMMNLQGEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHR
1l2iA      VEGMVEIFDMLLATSSRFMMNLQGEFVCLKSIILLNSGVY-----LEEKDHIHR
1a52A      VEGMVEIFDMLLATSSRFMMNLQGEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHR
3erdA      VEGMVEIFDMLLATSSRFMMNLQGEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHR
            *****

1g50A      VLDKITDTLIHMAKAGTLQQQHQR- LAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
1gwqA      VLDKITDTLIHMAKAGTLQQQHQR- LAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
1qktA      VLDKITDTLIHMAKAGTLQQQHQR- LAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
1l2iA      VLDKITDTLIHMAKAGTLQQQHQR- LAQLLLILSSHIRHMSNKGMEHLYSMKCKNVVPL
1a52A      VLDKITDTLIHMAKAGTLQQQHHERLAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
3erdA      VLDKITDTLIHMAKAGTLQQQHQR- LAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
            *****      ,      ,      *****

1g50A      YDLLLLMLDAHRLH-
1gwqA      YDLLLLMLDAHR-
1qktA      YDLLLLMLDAHRLHA
1l2iA      YDLLLLMLDAH-
1a52A      YDLLLLML-
3erdA      YDLLLLMLDAHRL-
            *****
    
```



* 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 β (NR1A2)

Results from round 3

Sequences producing significant alignments:	Score	E
Sequences used in model and found again:	(bits)	Value
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 1GWX 1GWX-B ppar-deltafragment: ligand binding domain;	239	6e-64

Psi-blast against pdb_seq

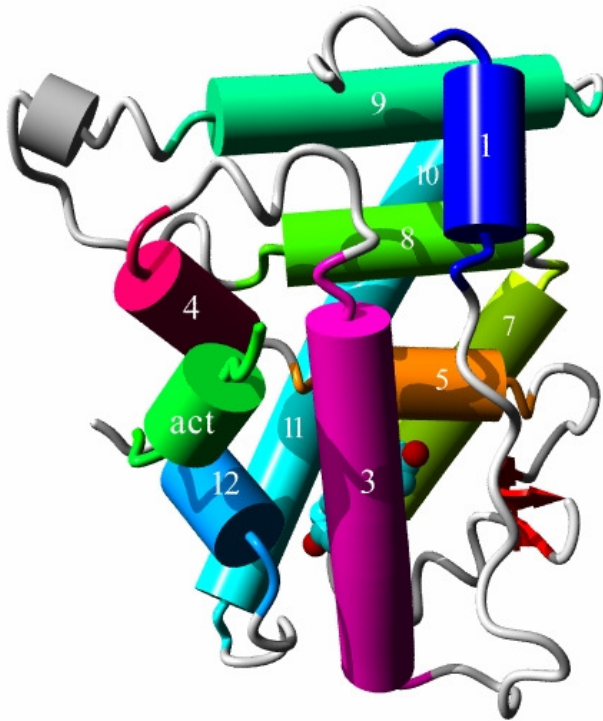


It matches with NR1B (RAR) ,
NR1I (VIT. D), NR1C (PPAR) all
groups of the same subfamily ->
homologous



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.

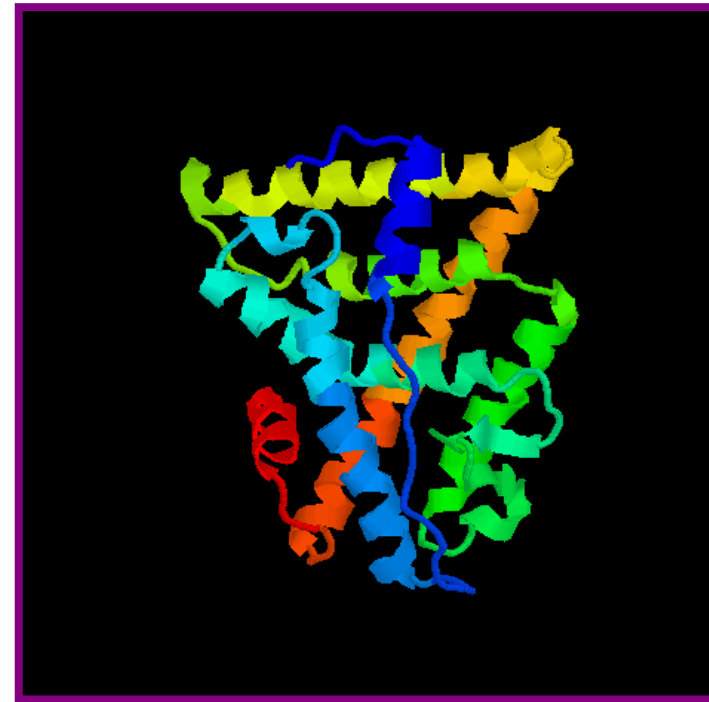


Apo- Vs Holo-structures

- 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

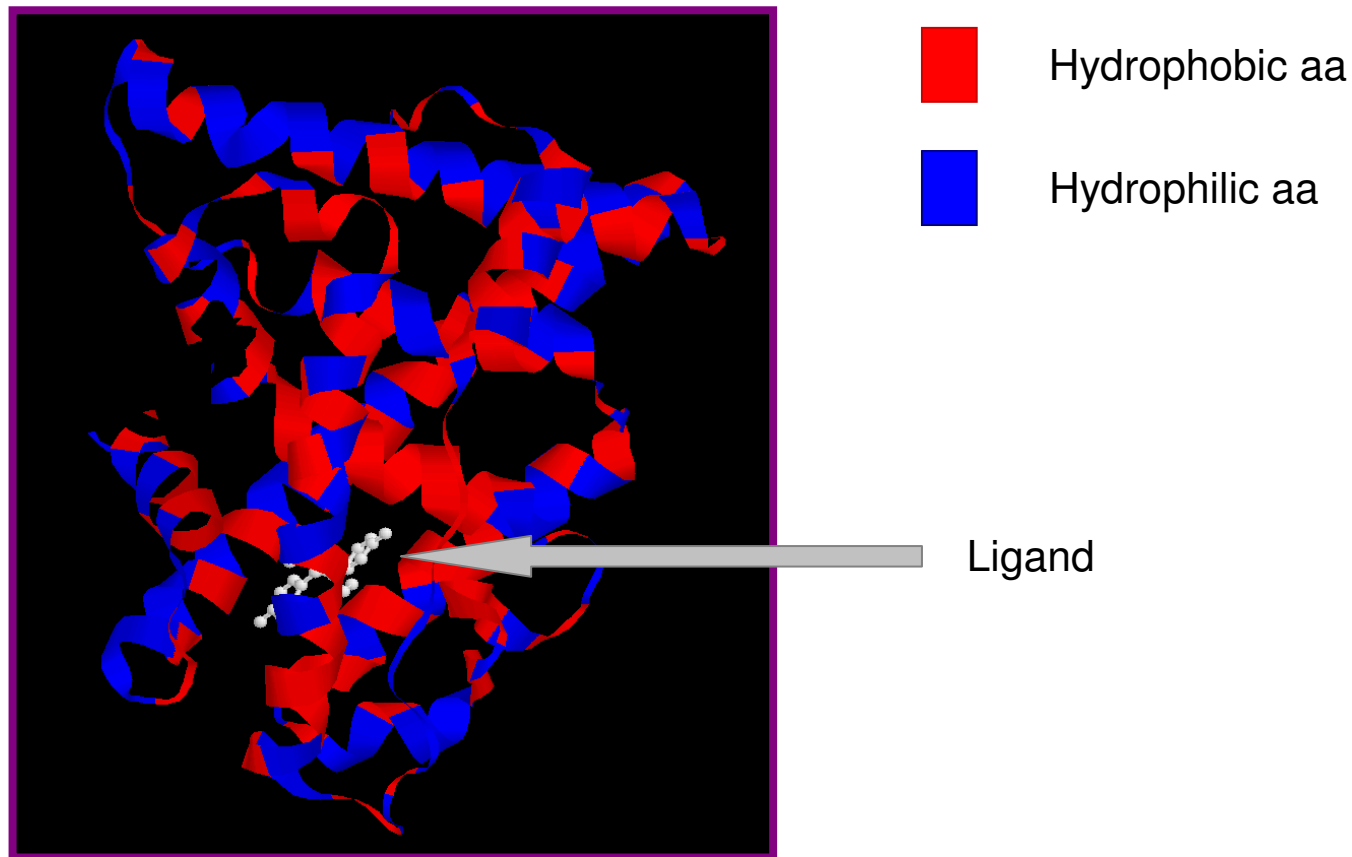


HOLO
Liganded-structure



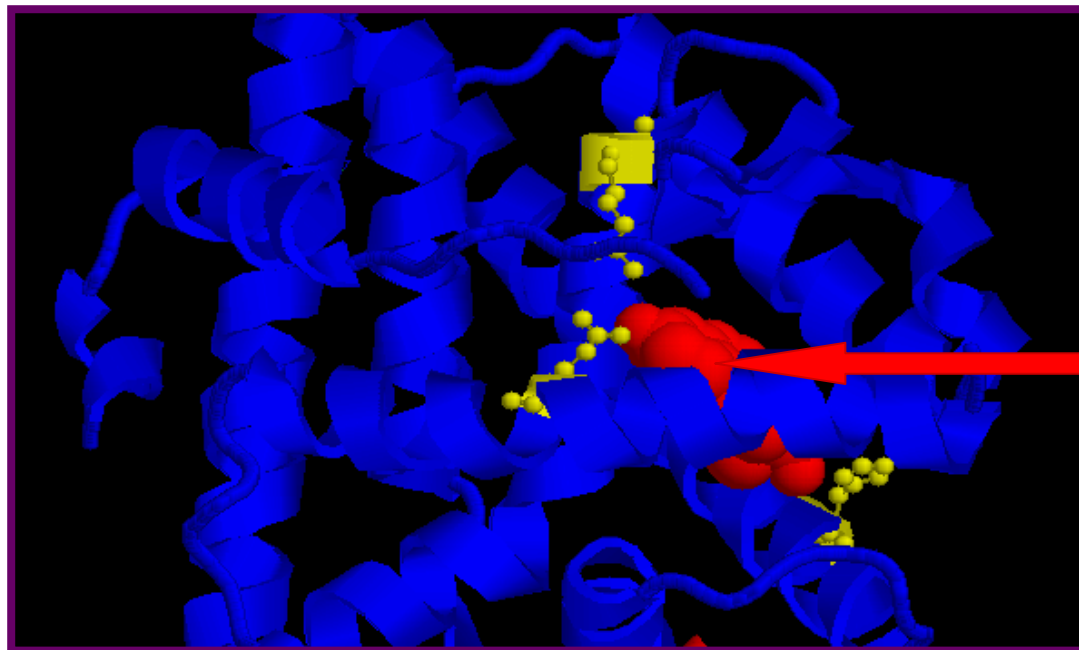
Holo-structures

- 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 Hys524 and the two hydroxyl groups of the ligand.
 - The architecture of the pocket is rigid and only accommodate planar structures.



Crucial conserved aa

Ligand



LBD helix 12 (H12)

CLUSTAL W(1.60) multiple sequence alignment

```
lbd .....
n46 .....
fby .....
kv6 .....
erd .....
err .....
prg ESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTD--KSPFVIYDMNSLNM-GEDKIK
kkq -TADLKSLAKRIYEAYLKNFNNKVKARVIL--SGKASNPPFVIHDMETLCMAE---KT
```

```
lbd .....SANEDM-PVERIL.....
n46 .....KPEPTDEE-W-E-LIKTVT.....
fby .....SSANEDM-PVERIL.....
kv6 .....N-K--IVSHLL.....
erd .....SLALSLTAD-Q--MVSALL.....
err .....ALSLTAD-Q--MVSALL.....
prg FK-HI---TPLQEQSKEVAIRIFQGC-----Q-FRSVEAQVEITEYAKSIGPVNLDL
kkq LVAKLVANGIQ---NKEAEVRIFHCC-----Q-CTSVETVTELTEFAKAIPIGFANL
```

```
lbd .....EAE-LAVEP-K--TETYVEANWLNPP-SSPN.....
n46 .....EAH-VATNAQWQ---K-R-----KFLPEDIGQAPIVNAPE
fby .....EAE-LAV-----E-P-----
kv6 .....VA--EPEK---I---Y---A--MPDPTVP--D-----
erd .....DA--EPPI---L---Y---S--E-Y-DPTRPF-----
err .....DA--EPPI---L---Y---S--E-Y-DPTRPF-----
prg NDQVTLTKYGVHEIITYMLAS---L-----
kkq NDQVTLTKYGVYEAIFAMLS---V-----
```

```
lbd .....DPVTNICQADKQLFTLV-----E-W-AKRIPHFSE-LPLDDQVILLRAGWN
n46 GGGVDLEAFSHFTKIITPAITRVV-----D-F-AKKLPWFCE-LPCEDQIILKGGCM
fby .....DPVTNICQADKQLFTLV-----E-W-AKRIPHFSE-LPLDDQVILLRAGWN
kv6 ---SDIKALTTLCDLADRELVYII-----G-W-AKHIPGFST-LSLADQMSLLQSAMN
erd ---SEASHMGLLTNLADRELVYHI-----N-W-AKRVPGFVD-LTLHDQVHLLCEAML
err ---SEASHMGLLTNLADRELVYHI-----N-W-AKRVPGFVD-LTLHDQVHLLCEAML
prg ---SEASHMGLLTNLADRELVYHI-----N-W-AKRVPGFVD-LTLHDQVHLLCEAML
kkq ---SEASHMGLLTNLADRELVYHI-----N-W-AKRVPGFVD-LTLHDQVHLLCEAML
```

```
lbd ELLIAS--FSH-RSI-----A--VK-DGI-LLA
n46 EIMSLR--AAV--RY-----D--PES-ETL-T-L
fby ELLIAS--FSH-RSI-----A--Y-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 FAYKFNALDSDSLAIFAVIILSGDRPGLLVKPIEDIQDNLLQALELQKL-----
kkq FANKFNALDSDSLAIFAVIILSGDRPGLLVGHIKQEQEIVHVLRLHLS-----
```

- 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-DRV-LTELYSKMRDMQMDKTELGC--LRAIVL--FN-P
n46 NGENAVTRGQLKNG-GLGVYS-DAI-FDLGMSL-SSFNLDDEVAL--LQAVLL--MS-S
fby T-GLHVHRNSAHS-AGVGAIF-DRVLTLYSKN-RDMQMDKTELGC--LRAIVL-FNP--
kv6 ADDYINDEDSKLA-GLLDLN-NAIL-QLVKKY-KSMKLEKEEFVT--LKAIALAN-S-D
erd APNLLDRNOGKCEVGNVEIF-DMLL-ATSSRF-RMMNLQGEFVC--LKSIIILLN-S-G
err APNLLDRNOGKCEVGNVEIF-DMLL-ATSSRF-RMMNLQGEFVC--LKSIIILLN-S-G
prg ---NH--PE-----SS-OLFAK--LLQ-K-----MTD-LRQIV-TEHV
kkq ---NH--PD-----DI-FLFPK--LLQ-K-----MAD-LRQIV-TEHA
```

```
lbd DS-K-----GLSNPAEV---E--ALR---E-KV-YASLEAYCK-HKYP-----EOPGR
n46 DRP-G-----LACVERI---E--KYQ---D-SF-LLAFEHYI--NYRK--H-HVT-HF
fby DSK-G-----LSNPAEV---E--ALR---E-KV-YASLEAYCK-HKYP---E-QP-GR
kv6 -SMH-I-----EDVEAV---Q--KLQ---D-VL-HEALQDYEA-GOHME---D-P-RR
erd VYTFSLSTLKS-LEEKDH---H--RVL---D-KI-TDTLIHLMA--KAGLTQQQH-QR
err -----VYEEKDHI---H--RVL---D-KI-TDTLIHLMA--KAGLTQQQH-QR
prg -----QLLQV-IKKTETDM-S--LHPLLQ-E-----IYKDL-----
kkq -----QLVQI-IKKTESDA-ALHPLLQEIY-R-----DMY-----
```

```
lbd FAKLLRLPALRSIGLKCLEHFF-----FKLIGDTPIDTLNENLEA-HQMT---
n46 WPKLLMKVTDLRNIGACHAS--RFLHMKVE-CPTLFPF-LFLNVEE-----
fby FAKLLRLPALRSIGLKCLE--HLFFFKLIGDTP-ID-T-FNENLE--AP-----
kv6 AGKMLMTLPLLRQTSKAVQ--HFYNIKLEG-KV-PHMKL--FLENLEA-----
erd LAQLLLILSHIRHMSNKGME--HLYSHKCKN-VV-PLYDLL-LEMLDAHRL-----
err LAQLLLILSHIRHMSNKGME--HLY-----SHF-FLEH-LLEMLD-H-----
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    ALSLIADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMINWAKRVPG
1errA.outSS     -TT--HHHHHHHHHHHH-----SS--SS--HHHHHHHHHHHHHHHHHHHHHHHHHHTTSTT
```

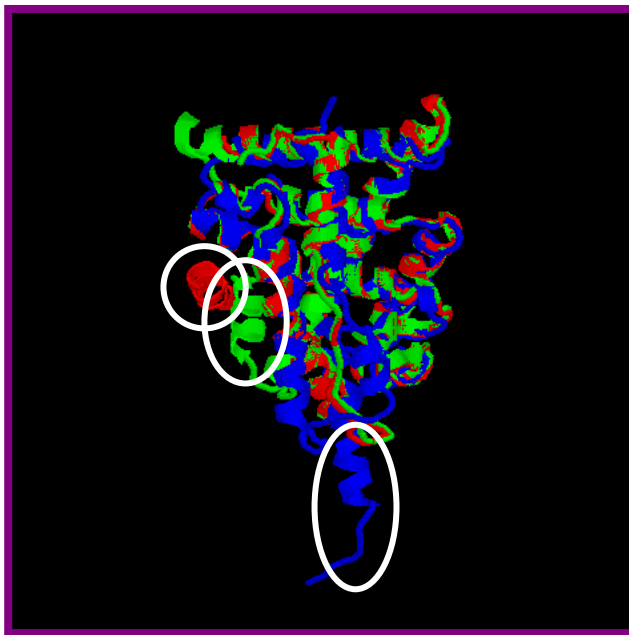
```
1errA.outSeq    FVDLTLHDQVHLLLECAWLEILMIGLVWRSMHPGKLLFAPNLLLLDRNQGKCVEGMVEIFD
1errA.outSS     GGGS- HHHHHHHHHHHHHHHHHHHHHHHHHHHHTTSTTEEEETTEEEHHHHHTTSTT - HHHHH
```

```
1errA.outSeq    MLLATSSRFMMNLQGEEFVCLKSIILLNSGVYØEEKDHIHRVLDKITDTLIHLMAKAGL
1errA.outSS     HHHHHHHHHHHHHT - - HHHHHHHHHHHHHHHHSS - - - - HHHHHHHHHHHHHHHHHHHHHHHTT -
```

[illegible]

LBD helix 12 (H12)

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



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

STAMP code:

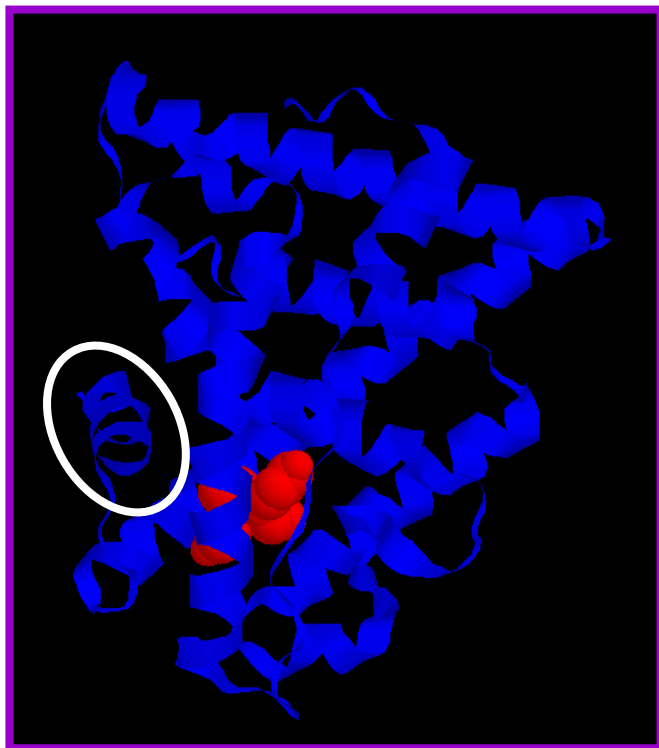
Alignment score Sc: 6.56
Alignment length Lp: 244
RMSD: 1.79

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

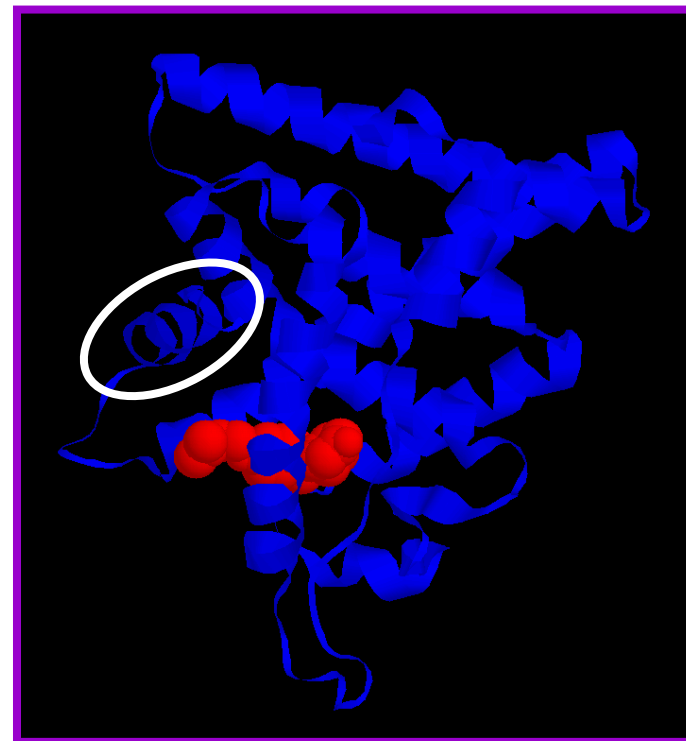


Agonist/antagonist-induced conformation

- 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

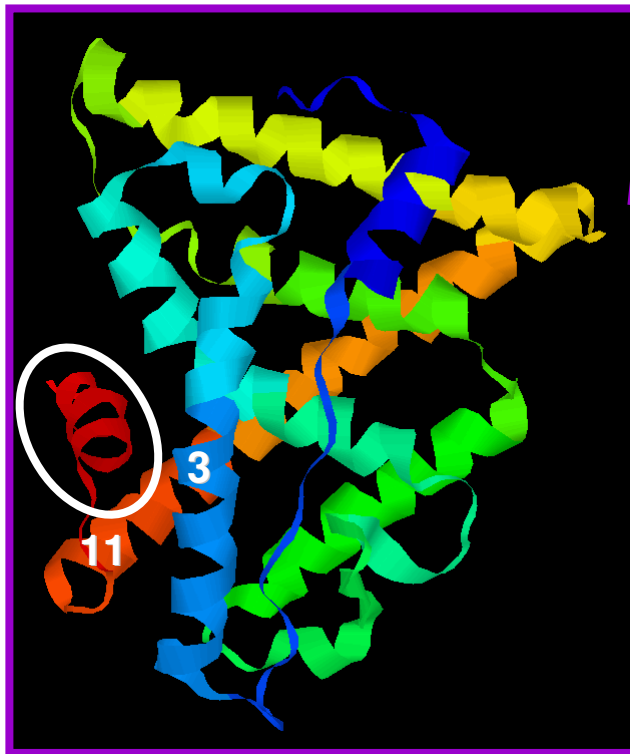


Antagonist



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

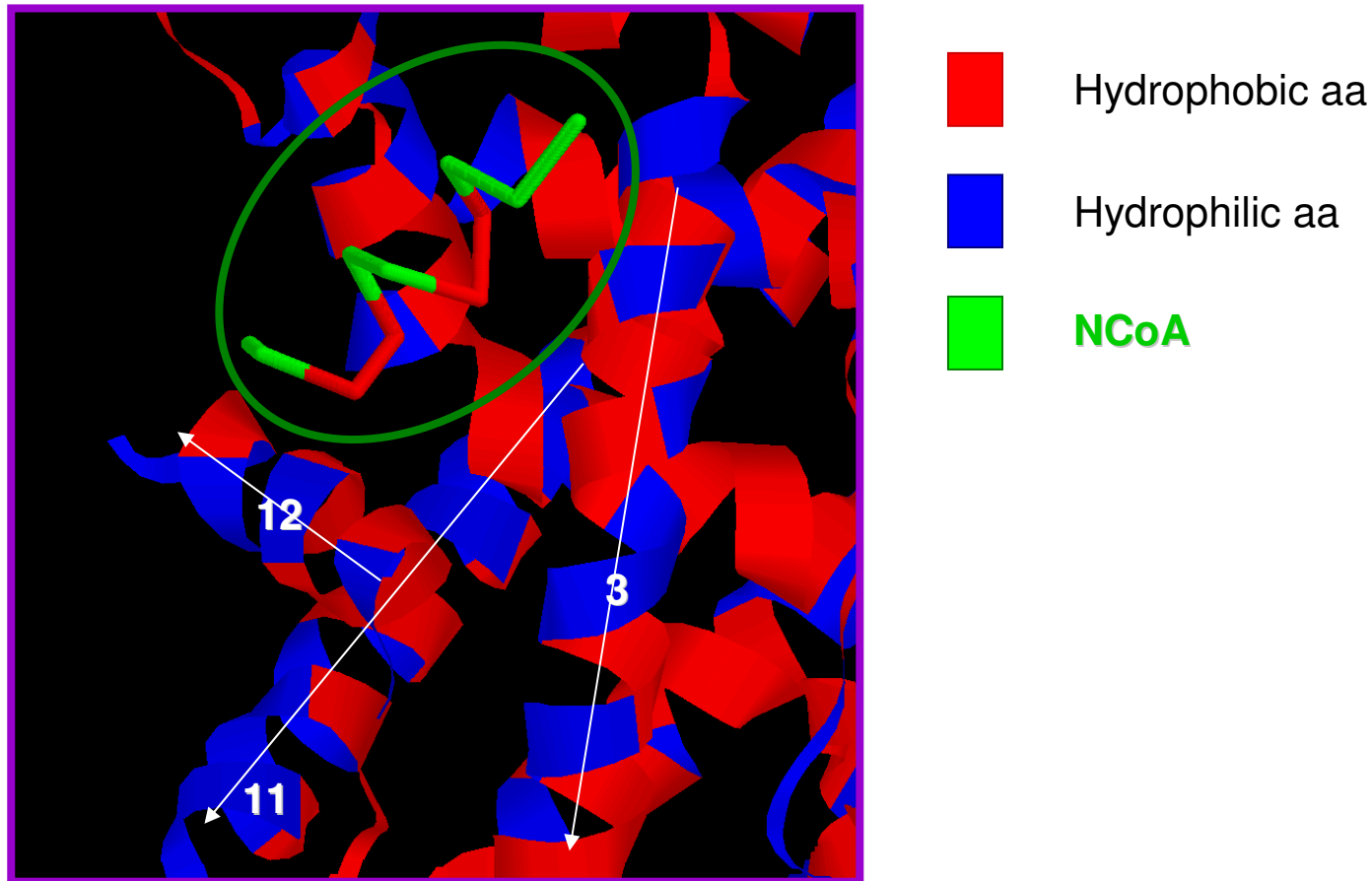


Antagonist



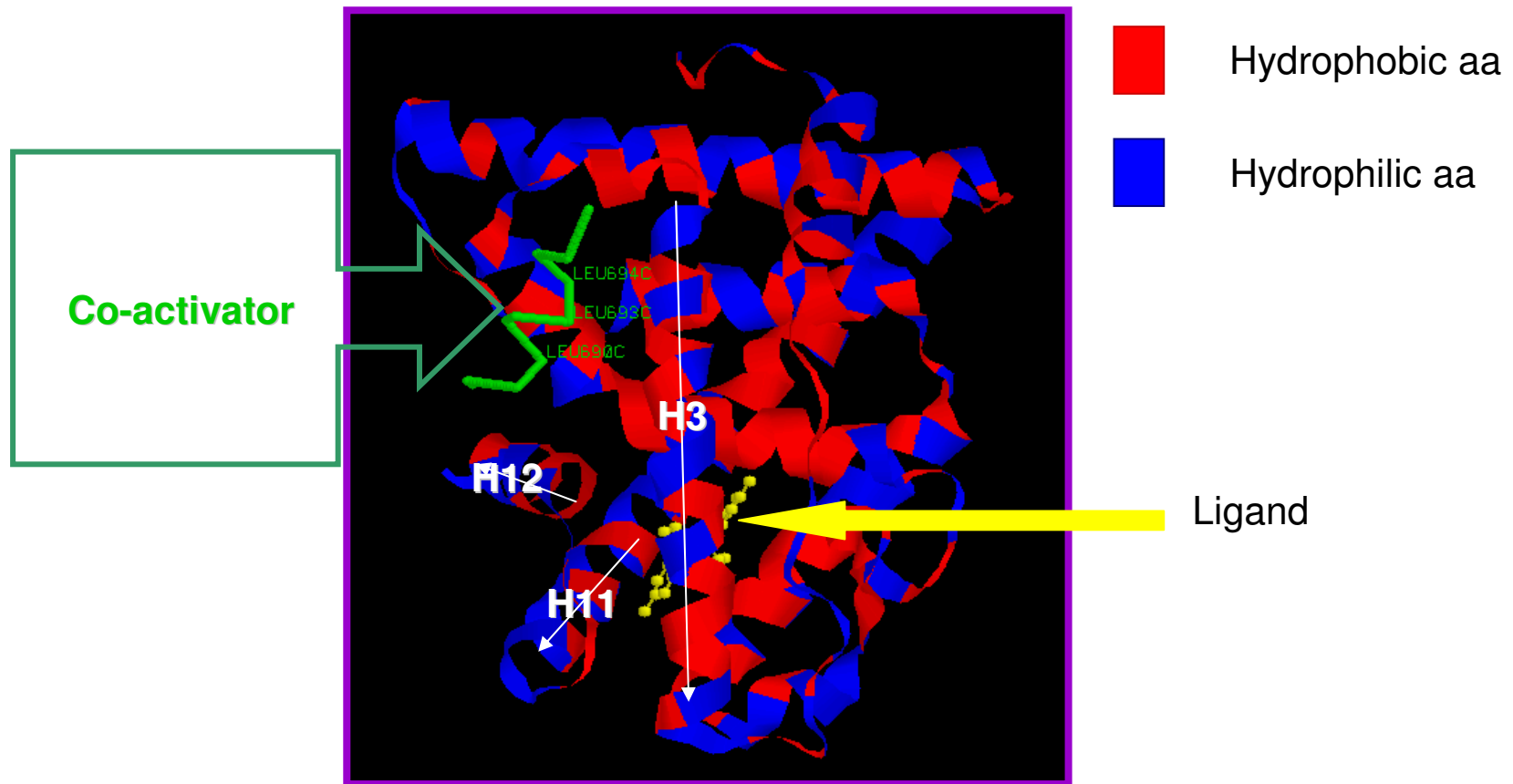
Agonist-bond structure

- 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).



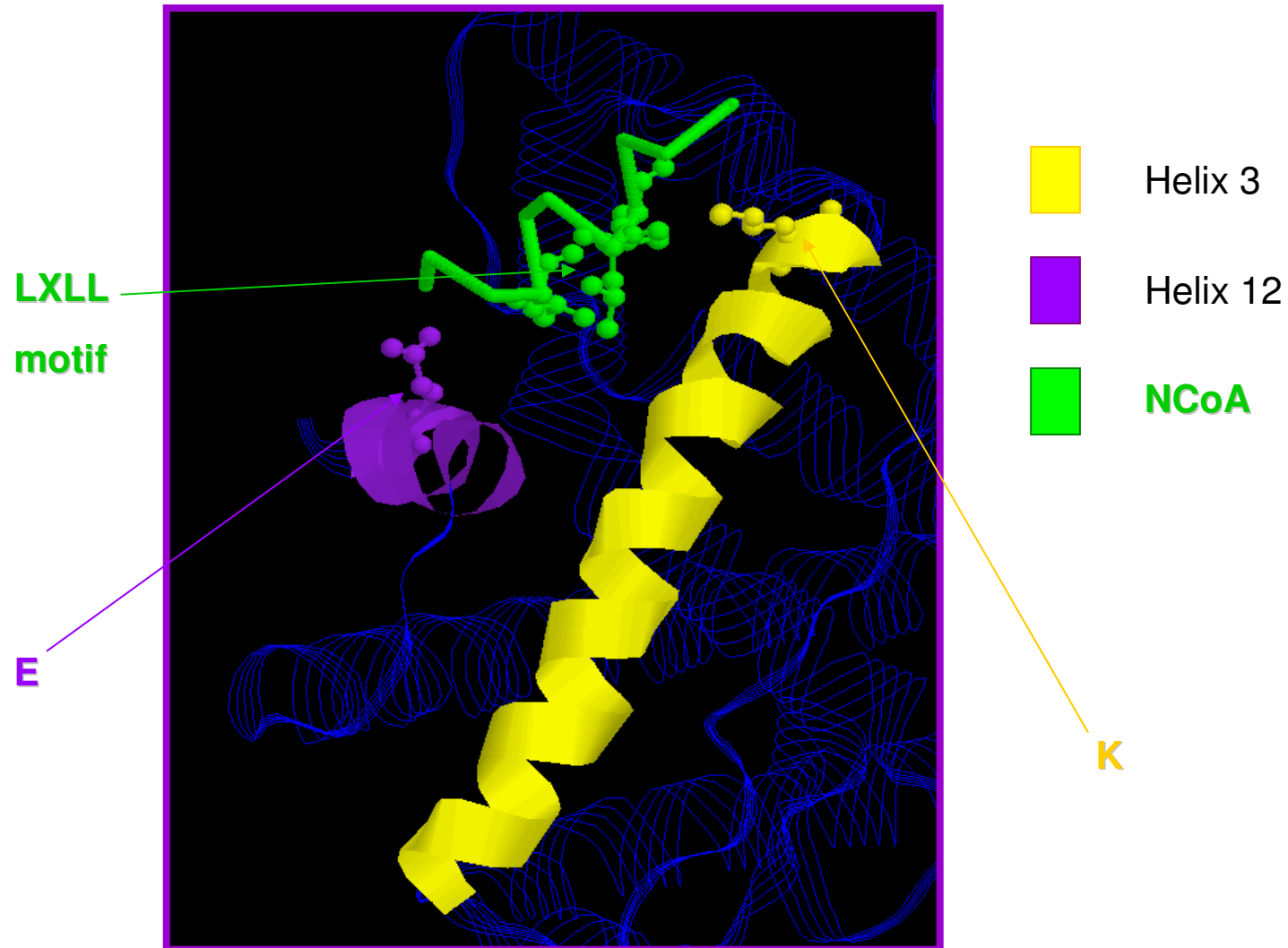
Interaction NR-coactivator

- **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 ESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTD--KSPFVIYDMNSLMN-GEDKIK
kkq -TADLKSIAKRIYEAYLKNFNNKVKARVIL--SGKASNNPFFVIHDMETLCMAE---KT
```

```
lbd .....SANEDM-PVERIL.....
n46 .....KPEPTDEE-W-E-LIKTVT.....
fby .....SSANEDM-PVERIL.....
kv6 .....N-K--IVSHLL.....
erd .....SLALSLTAD-Q--MVSALL.....
err .....ALSLTAD-Q--MVSALL.....
prg FK-HI---TPLQEQSKEVAIRIFQGC-----Q-FRSVEAVQEITEYAKSIPGFVNLDL
kkq LVAKLVANGIQ---NKEAEVRIFHCC-----Q-CTSVETVTELTEFAKAIPGFANL
```

```
lbd .....EAE-LAVEP-K--TETYVEANWNLN-SSPN.....
n46 .....EAH-VATNAQWKQ---K-R-----KFLPEDIGQAPIVNAPE
fby .....EAE-LAV-----E--P-----
kv6 .....VA--EPEK---I---Y---A--MPDPTVP--D-----
erd .....DA--EPPI---L---Y---S--E-Y-DPTRPF-----
err .....DA--EPPI---L---Y---S--E-Y-DPTRPF-----
prg NDQVTLTKYGVYHEIITMLAS---L-----
kkq NDQVTLTKYGVYEAIFAMLS---V-----
```

LYSINE

```
lbd .....DPVTNICQADKQLFTLV-----E-W-AKRIPHFSE-LPLDDQVILLRAGWN
n46 GGGVDLEAFSHFTKIITPAITRVV-----D-F-AKKLPWFCE-LPCEDQIILLKGCCH
fby .....DPVTNICQADKQLFTLV-----E-W-AKRIPHFSE-LPLDDQVILLRAGWN
kv6 ---SDIKALTTLCDLADRELVYII-----G-W-AKHIPGFST-LSLADQMSLLQSAMN
erd ---SEASHMGLLTNLADRELVYHI-----N-W-AKRIPGFVD-LTLHDQVHLLCEAML
err ---SEASHMGLLTNLADRELVYHI-----N-W-AKRIPGFVD-LTLHDQVHLLCEAML
prg ---SEASHMGLLTNLADRELVYHI-----N-W-AKRIPGFVD-LTLHDQVHLLCEAML
kkq ---MNKD-GVLISEGOGFHTRE-----FLKSLRKPFQGFMEPKFE
```

```
lbd ELLIAS--FSH-RSI-----A--VK-DGI-LLA
n46 EIMSLR--AAV--RY-----D--PES-ETL-T-L
fby ELLIAS--FSH-RSI-----A--Y-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 FAYKFNALELDDSLAIFAVIILSGDRPGLLVKPIEDIQDNLLOALELQKL-----
kkq FANKFNALELDDSDISLFYAAIICGDRPGLLVGHIEKNQEGIVHVLRLHLS-----
```

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

```
lbd T-GLHVHRNSAHS-AGVGAIF-DRV-LTELYSKMRDMQMDKTELGC--LRAIVL--FN-P
n46 NGENAVTRGQLKNG-GLGVYS-DAI-FDLGMSL-SSFNLDDEVAL--LQAVLL--MS-S
fby T-GLHVHRNSAHS-AGVGAIF-DRVLTLYSKN-RDMQMDKTELGC--LRAIVL-FNP--
kv6 ADDYINDEDSKLA-GLLDLN-NAIL-QLVKKY-KSMKLEKEEFVT--LKAIALAN-S-D
erd APNLLDRNOGKCVYEGWYEIF-DMLL-ATSSRF-RMMNLQGEFVC--LKSIIILLN-S-G
err APNLLDRNOGKCVYEGWYEIF-DMLL-ATSSRF-RMMNLQGEFVC--LKSIIILLN-S-G
prg ---NH--PE-----SS-OLFAK--LLQ-K-----MTD-LRQIV-TEHY
kkq ---NH--PD-----DI-FLFPK--LLQ-K-----MAD-LRQIV-TEHA
```

```
lbd DS-K-----GLSNPAEV---E--ALR---E-KV-YASLEAYCK-HKYP-----EOPGR
n46 DRP-G-----LACYERI---E--KYQ---D-SF-LLAFHYI--NYRK--H-HVT-HF
fby DSK-G-----LSNPAEV---E--ALR---E-KV-YASLEAYCK-HKYP---E-QP-GR
kv6 -SMH-I-----EDVEAV---Q--KLQ---D-VL-HEALQDYEA-GQHME---D-P-RR
erd VYTFLSSTLKS-LEEKDH---H--RVL---D-KI-TDTLIHLMA--KAGTLQQQH-QR
err -----VYEEKDHI---H--RVL---D-KI-TDTLIHLMA--KAGTLQQQH-QR
prg -----QLLQV-IKKTETDM-S--LHPLLQ-E-----IYKDL-----
kkq -----QLVQI-IKKTESDA-ALHPLLQEIY-R-----DMY-----
```

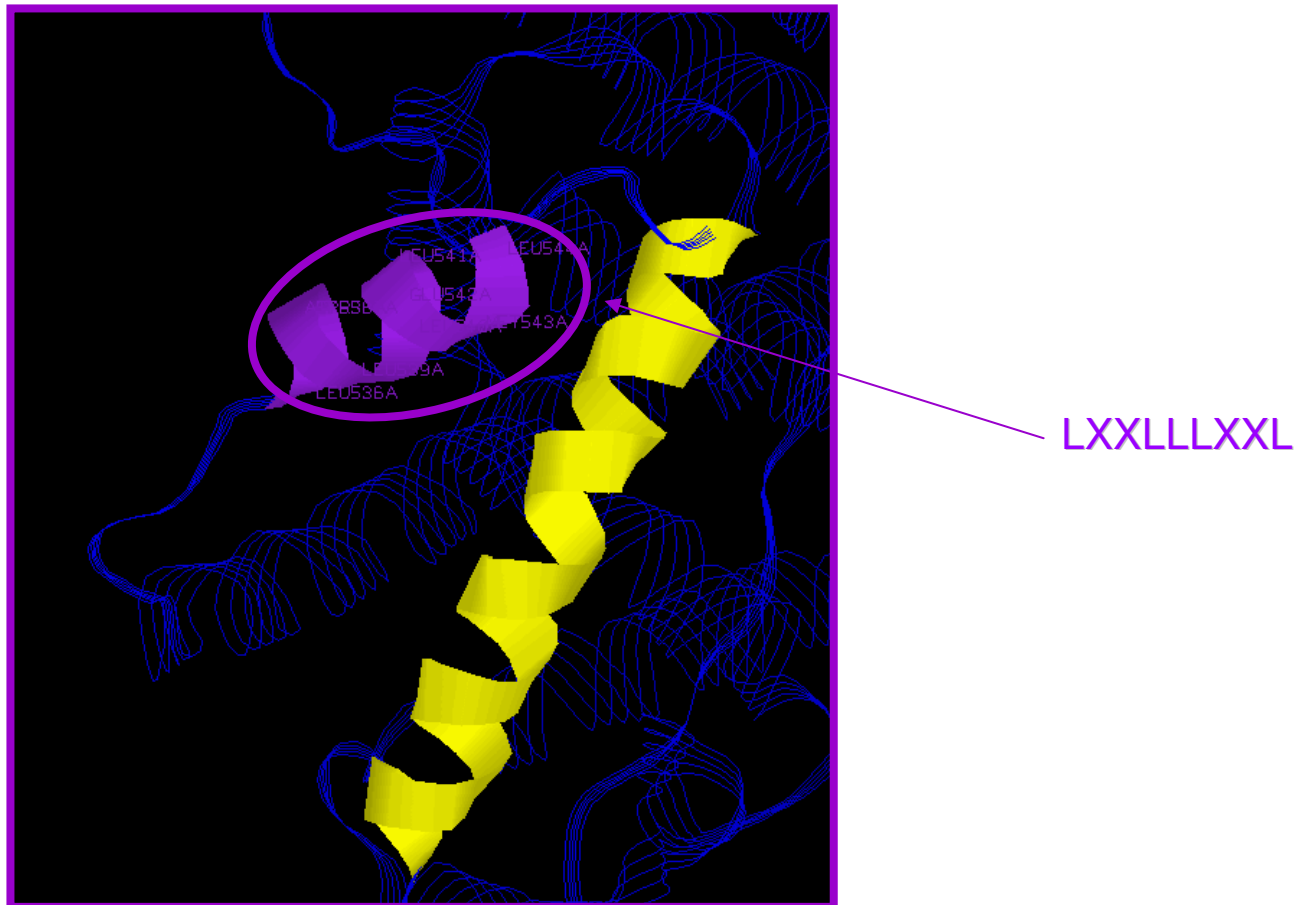
GLUTAMATE

```
lbd FAKLLLRPALRSIGLKCLELFF-----FKLIGDTPIDTFLHLEAPQMT---
n46 WPKLLMKVYDRLRMI GACHAS--RFLHMKVE-CPTELFP-LFUEYFED-----
fby FAKLLLRPALRSIGLKCLE--HLFFFKLIGDTP-ID-T-FLHLEA--AP-----
kv6 AGKMLMTLPLLRQTSKAVQ--HFYNIKLEG-KV-PHMKL-FLHLEA--
erd LAQLLILSHIRHMSNKGME--HLYSMCKN-VV-PLYDLL-LEHDAHRL-----
err LAQLLILSHIRHMSNKGME--HLY-----SF-P-LYDILLLEHDAH
prg -----
kkq -----
```



Antagonist-bond structure

- In the antagonist-bound 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 ?

