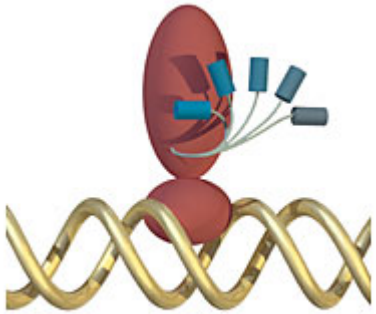
NUCLEAR RECEPTORS

Homology, function and structure

Structural Bioinformatics



Ferran Briansó
Elisenda Feliu
Núria Queralt

Contents

Contents

PART I
Nuclear Receptors

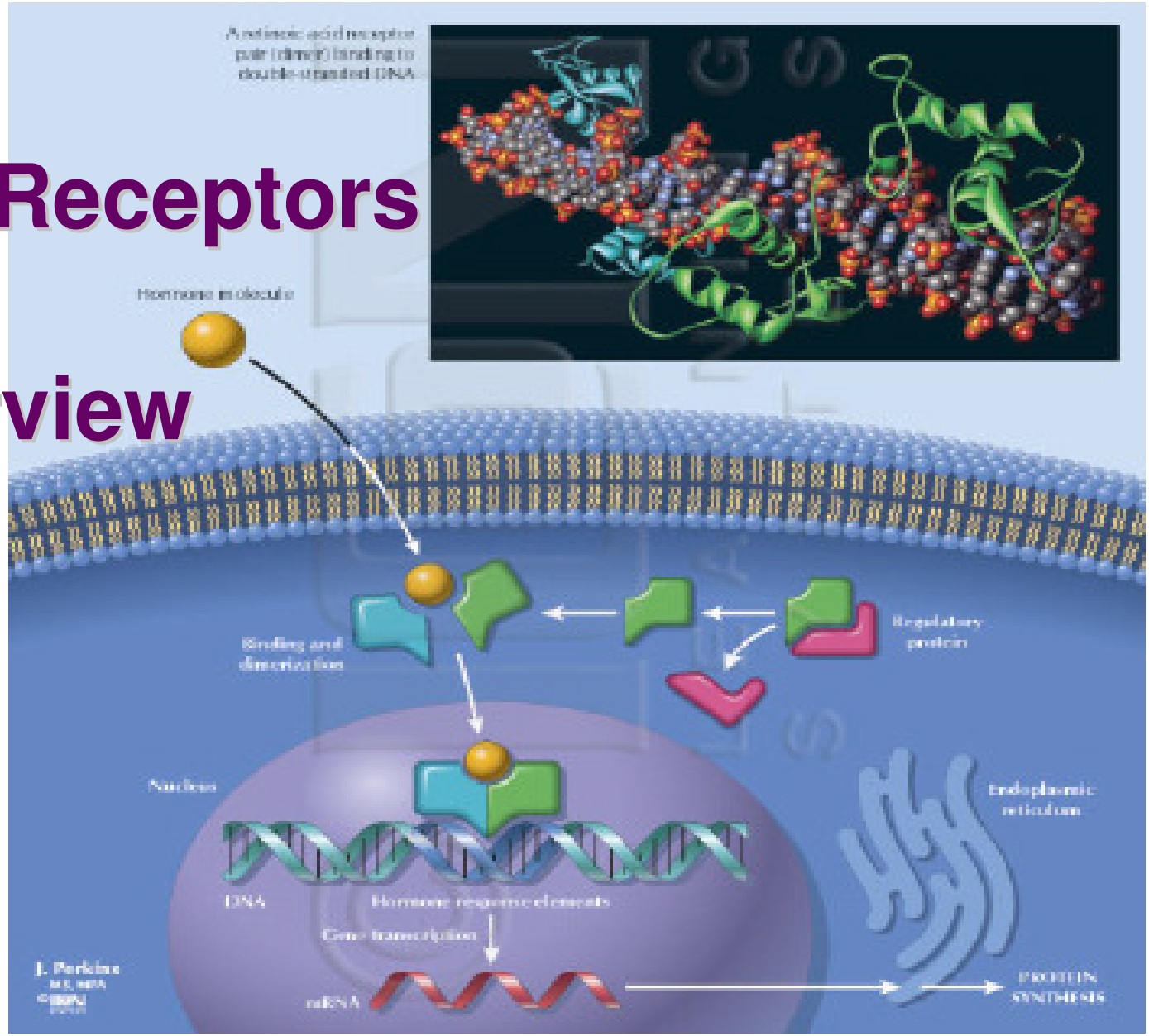
PART II
DNA Binding Domain

PART III
Ligand Binding Domain

Contents

Nuclear Receptors

Overview



Introduction

- Nuclear receptors (NRs) belong to a large superfamily that are ligand activated intracelluar 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.

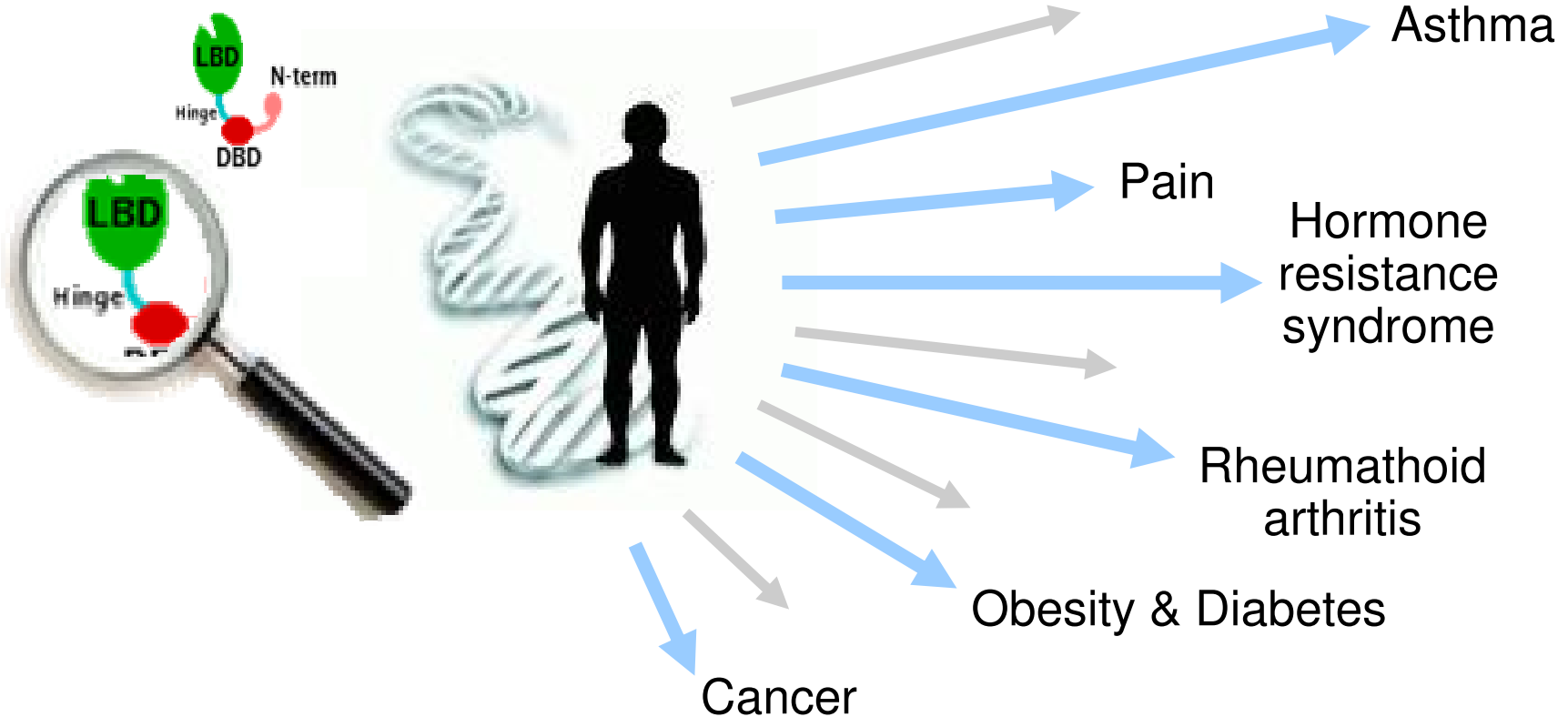
- Nuclear receptors may be classified either according to activation mechanism (type I or II), or sequence homology (NR subfamilies 0-6).

- Nowadays, there are more than 350 NR structures in the PDB.

Scientific interest

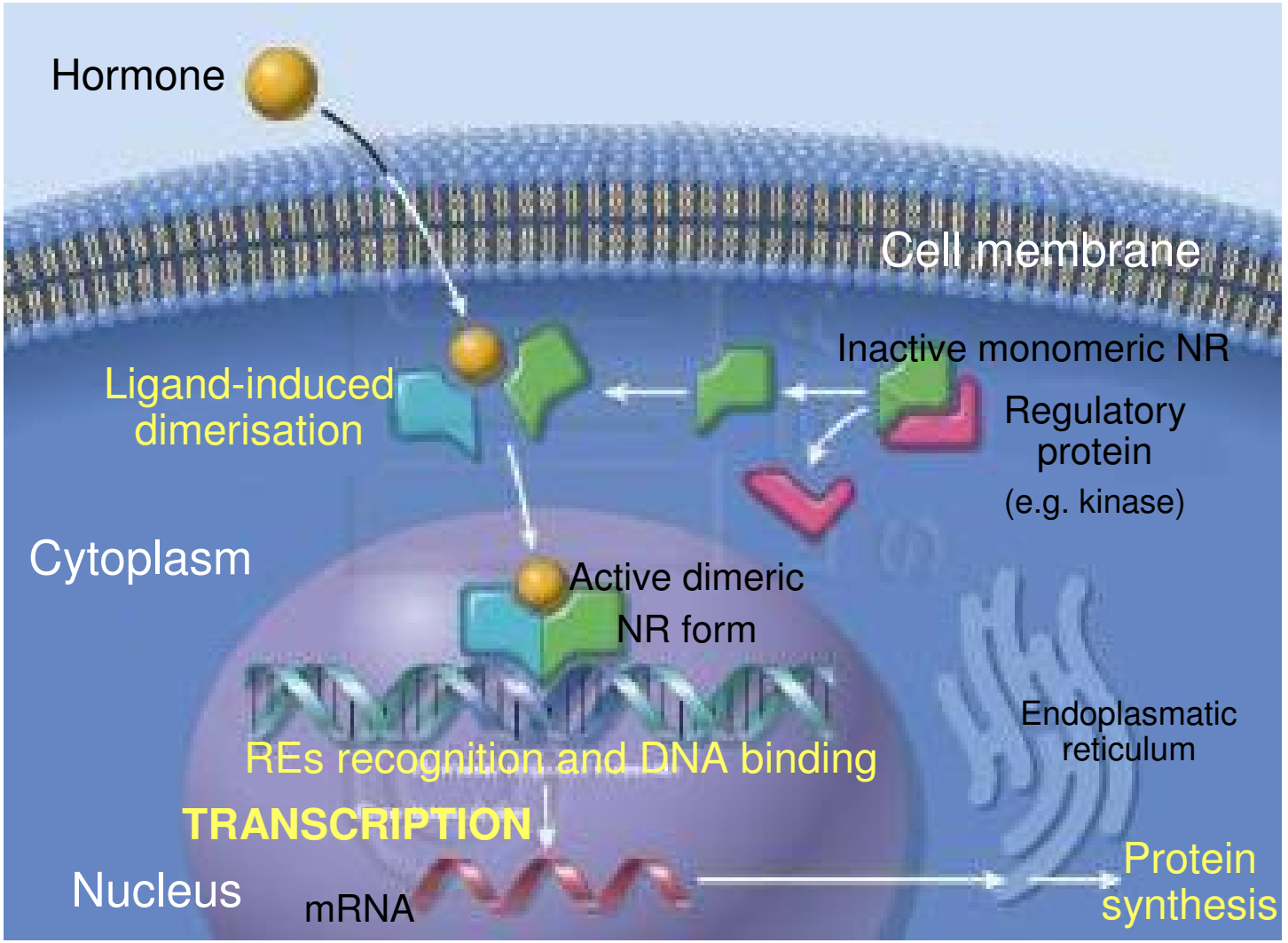
- Nuclear receptors are transcription factors involved in such important physiological functions as control of embryonic development, organ physiology, cell differentiation and homeostasis.

- 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

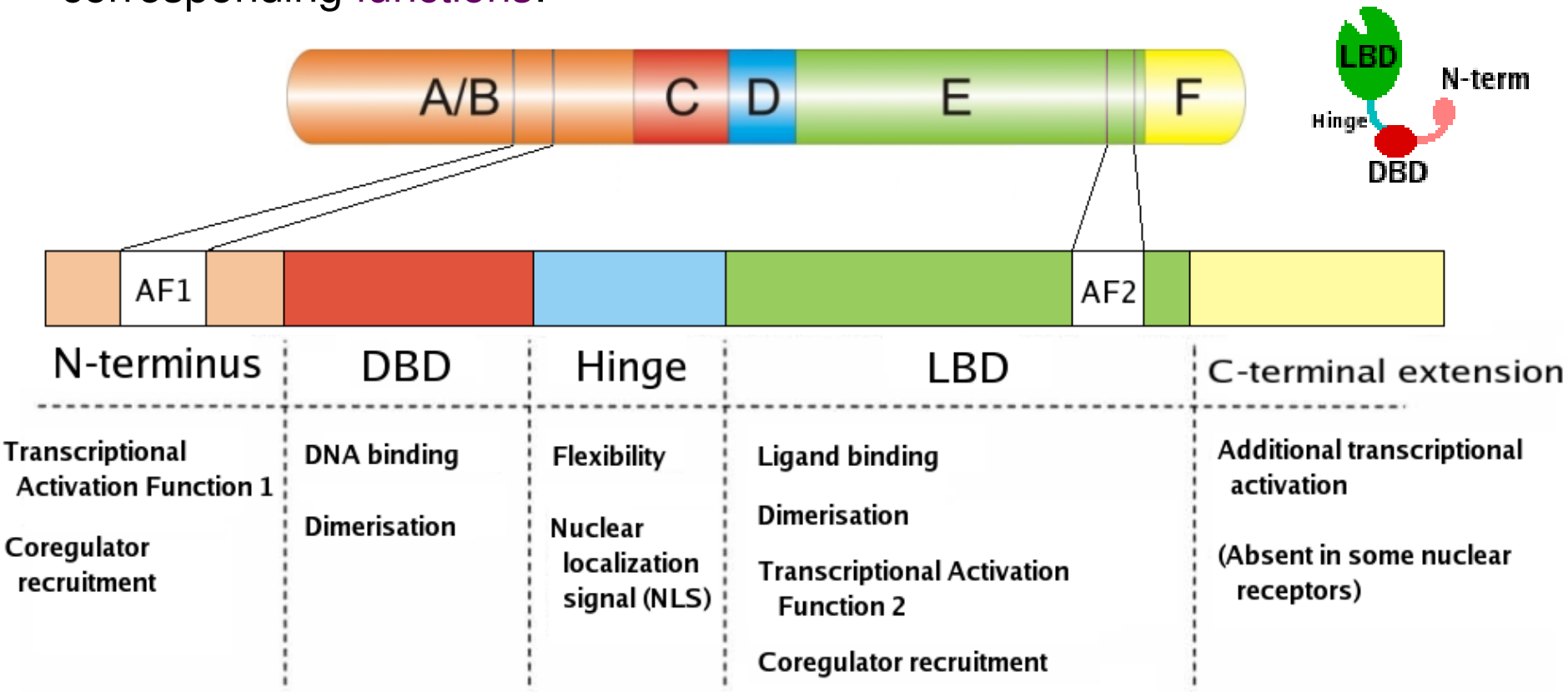
Cellular differentiation

Physiology

...

Primary structure

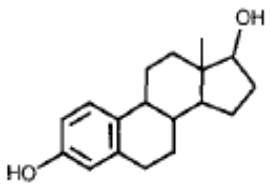
- A typical nuclear receptor contains the following domains, with corresponding functions:



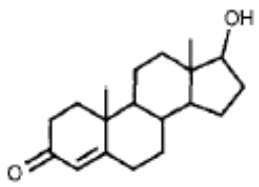
- DNA Binding Domain (DBD) and Ligand Binding Domain (LBD) are significant conserved regions, but DBD is the most one.

Ligands

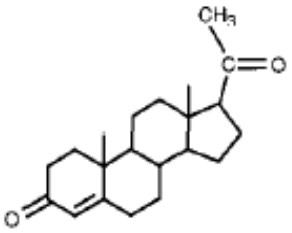
- Lipophilic substances such as endogenous hormones, vitamins A and D, drugs, and xenobiotic endocrine disruptors:



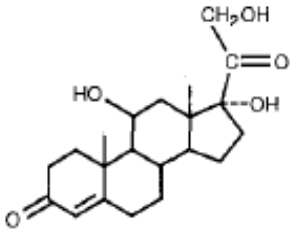
Estradiol



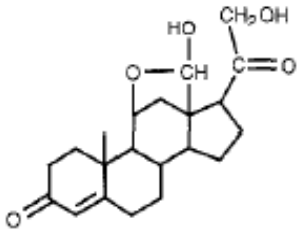
Testosterone



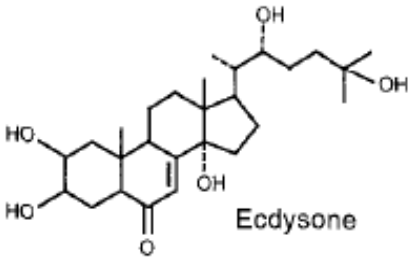
Progesterone



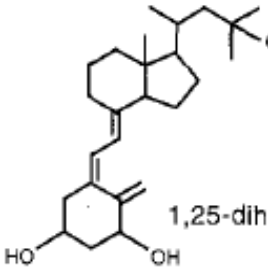
Cortisol



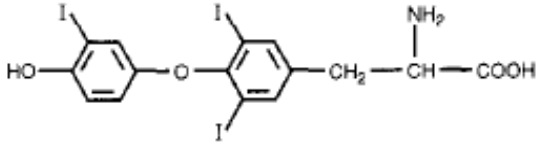
Aldosterone



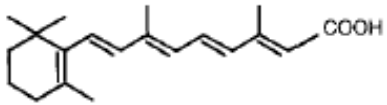
Ecdysone



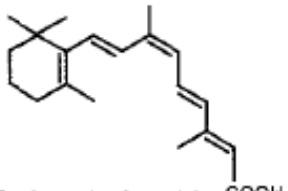
1,25-dihydroxyvitamin D_3



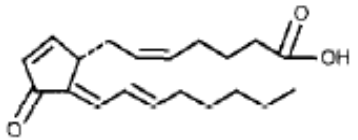
3,5,3'-L-triodothyronine



All-trans retinoic acid



9-cis retinoic acid

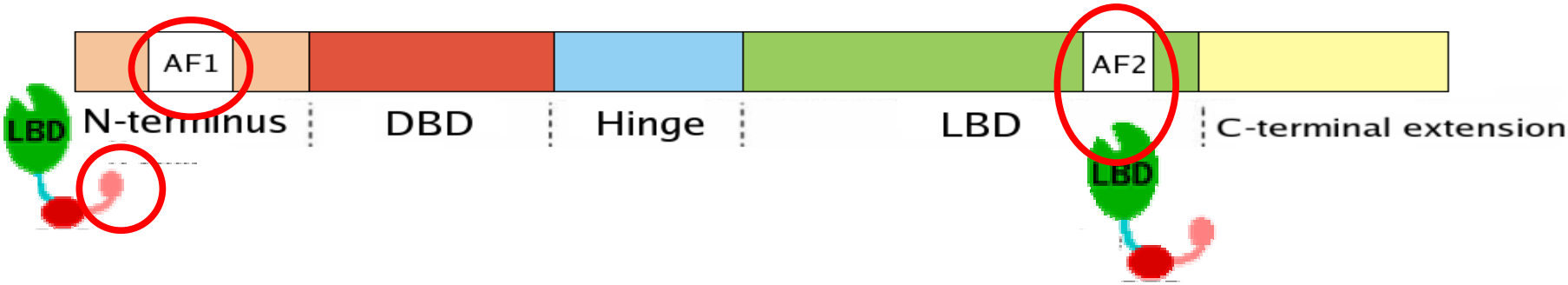


15-deoxy-$\Delta^{12,14}$-prostaglandin J_2

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.

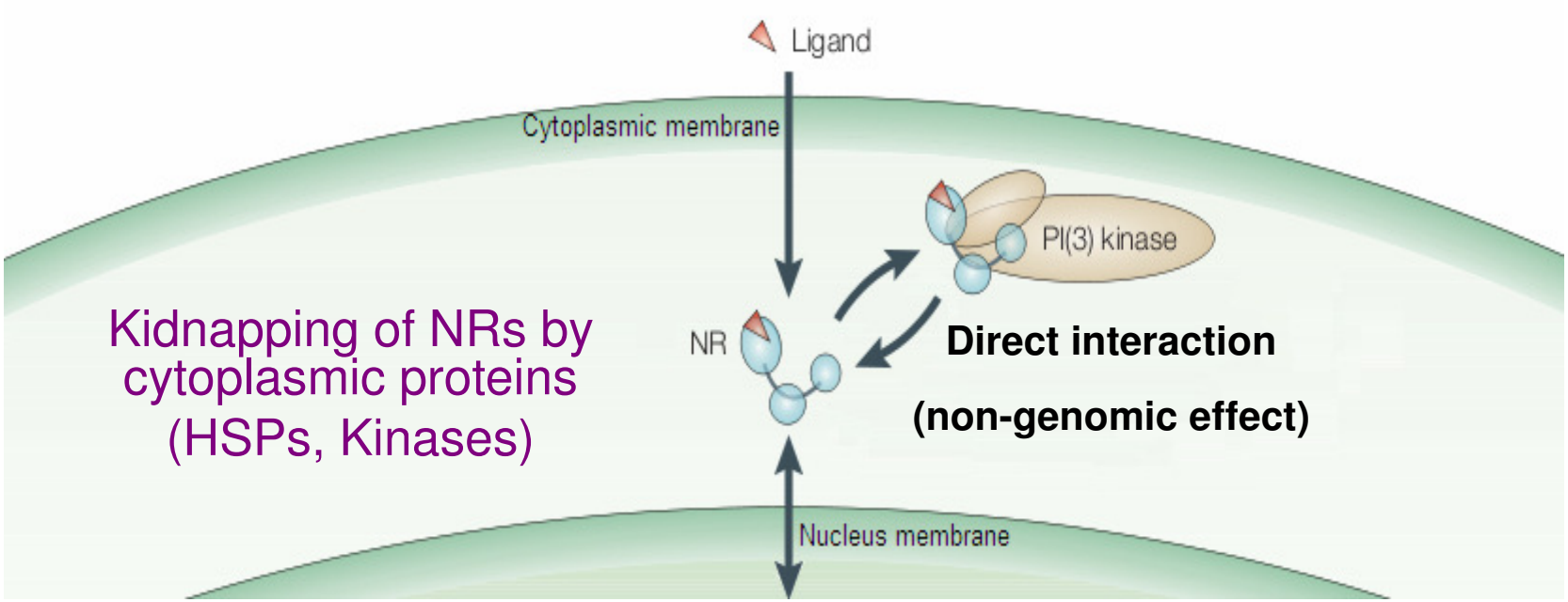


- **Activation Function-2, included in the Ligand Binding Domain**

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.

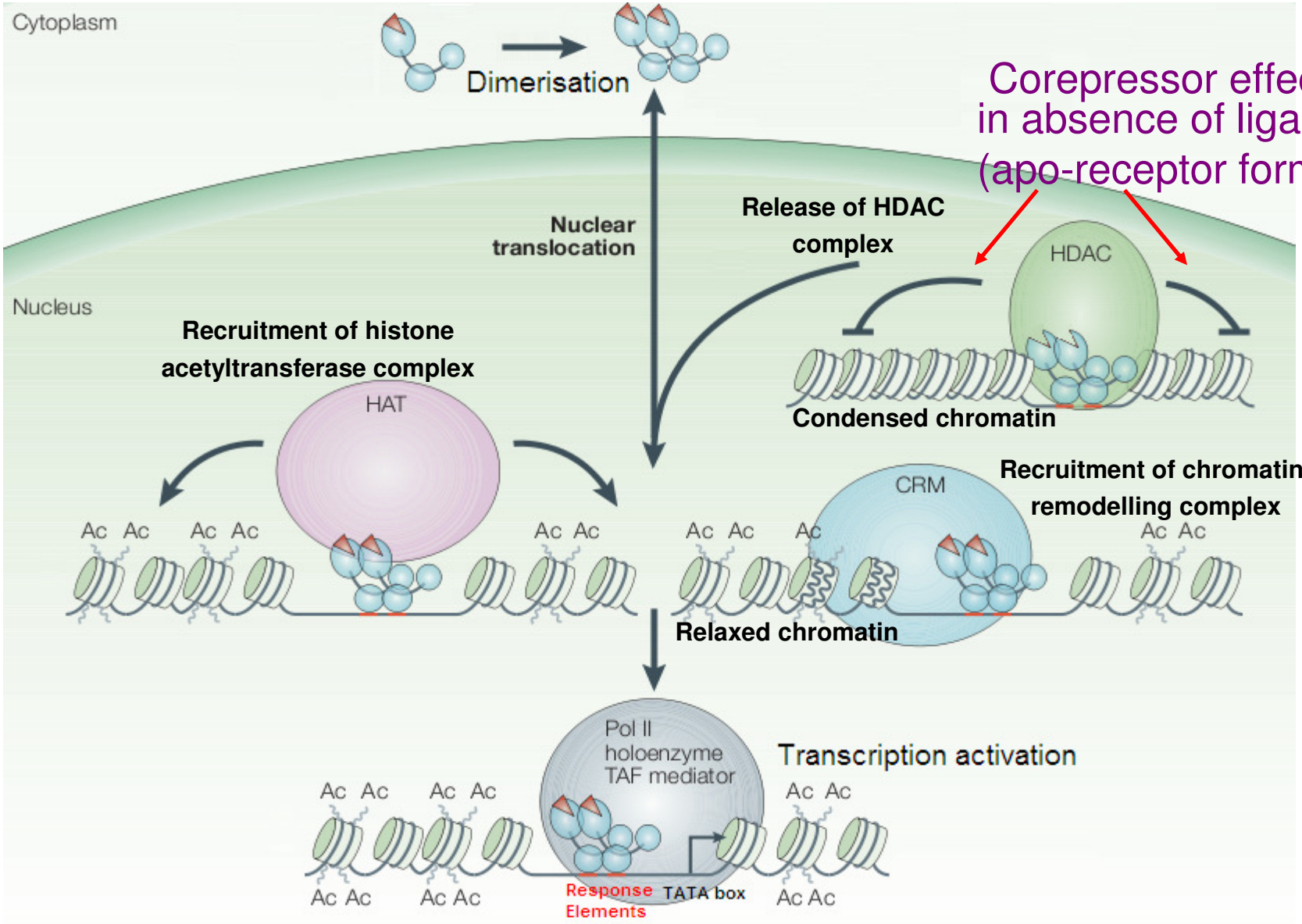
Regulation and mode of action

- Cytoplasmic regulation mechanism:



Kidnapping of NRs by cytoplasmic proteins (HSPs, Kinases)

Regulation and mode of action



Evolution

Algae

Fungi

Plants

Common ancestor
Monomeric and constitutively active form of NR

early Metazoans
LBD + DBD

One ancestral gene / Two different ancestral genes

Duplication ? Fusion

DBD LBD

?

Coelenterata
Dimeric orphan forms

lower Chordata (*Ciona*)
Functional transactivation domain
+ Ligand-dependent activity

Arthropoda (*Drosophila*)
Representatives of all NR subfamilies
+ Transcriptional cofactors

Novac & Heinzel; 2004

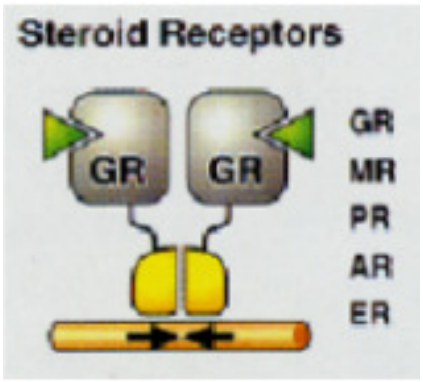
Former classification

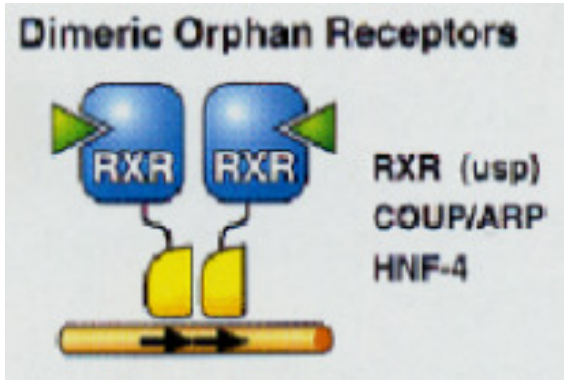
- The NR superfamily have been classified into four subfamilies based on their DNA-binding, ligand-binding and dimerisation properties:

Homodimeric

Ligand-dependent

Head-to-head





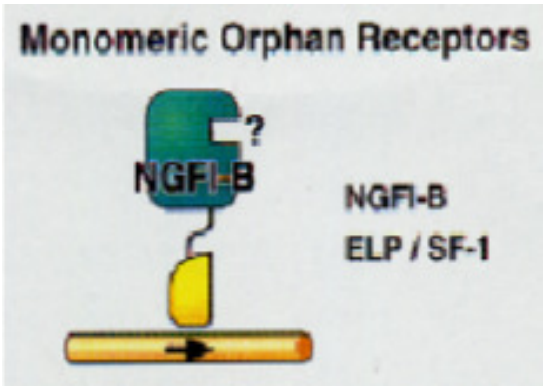
Homodimeric

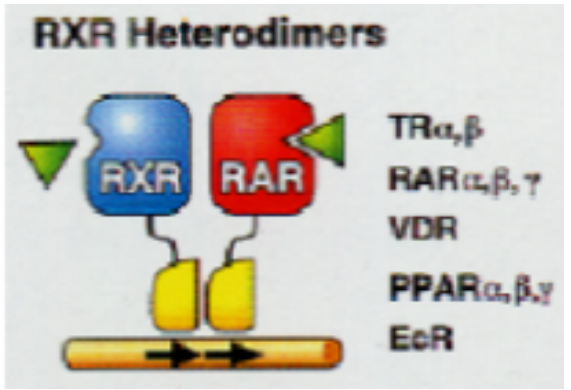
Unknown ligand

Head-to-tail

Monomeric

Unknown ligand





Heterodimeric

Ligand-independent

Head-to-tail

Receptors & Ligands

Name	Abbreviation	Nomenclature	Ligand
Thyroid hormone receptor	TRα TRβ	NR1A1 NR1A2	Thyroid hormone Thyroid hormone
Retinoic acid receptor	RARα RARβ RARγ	NR1B1 NR1B2 NR1B3	Retinoic acid Retinoic acid Retinoic acid
Peroxisome proliferator-activated receptor	PPARα PPARβ PPARγ	NR1C1 NR1C2 NR1C3	Fatty acids, leukotriene B4, fibrates Fatty acids Fatty acids, prostaglandin J2,
Reverse erbA	Rev-erbα Rev-erbβ	NR1D1 NR1D1	Orphan Orphan
RAR-related orphan receptor	RORα RORβ RORγ	NR1F1 NR1F2 NR1F3	Cholesterol, cholesteryl sulphate Retinoic acid Retinoic acid
Liver X receptor	LXRα LXRβ	NR1H3 NR1H2	Oxysterols, T0901317, GW3965 Oxysterols, T0901317, GW3965
Farnesoid X receptor	FXRα FXRβ*	NR1H4 NR1H5	Bile acids, Fexaramine Lanosterol
Vitamin D receptor	VDR	NR1I1	1,25-dihydroxy vitamin D_3, litocholic acid
Pregnane X receptor	PXR	NR1I2	Xenobiotics, PCN
Constitutive androstane receptor	CAR	NR1I3	Xenobiotics, phenobarbital
Human nuclear factor 4	HNF4α HNF4γ	NR2A1 NR2A2	Orphan Orphan
Retnoid X receptor	RXRα RXRβ RXRγ	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	ERα	NR3A1	Oestradiol-17β, tamoxifen, raloxifene
	ERβ	NR3A2	Oestradiol-17β, various synthetic compounds
Oestrogen receptor-related receptor	ERRα	NR3B1	Orphan
	ERRβ	NR3B2	DES, 4-OH tamoxifen
	ERRγ	NR3B3	DES, 4-OH tamoxifen
Glucocorticoid receptor	GR	NR3C1	Cortisol, dexamethasone, RU486
Mineralocorticoid receptor	MR	NR3C2	Aldosterone, spirolactone
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 (RXRα)
  - Group C: Testicular receptor
  - Group E: TLX/PNR
  - Group F: COUP/EAR

Homology classification

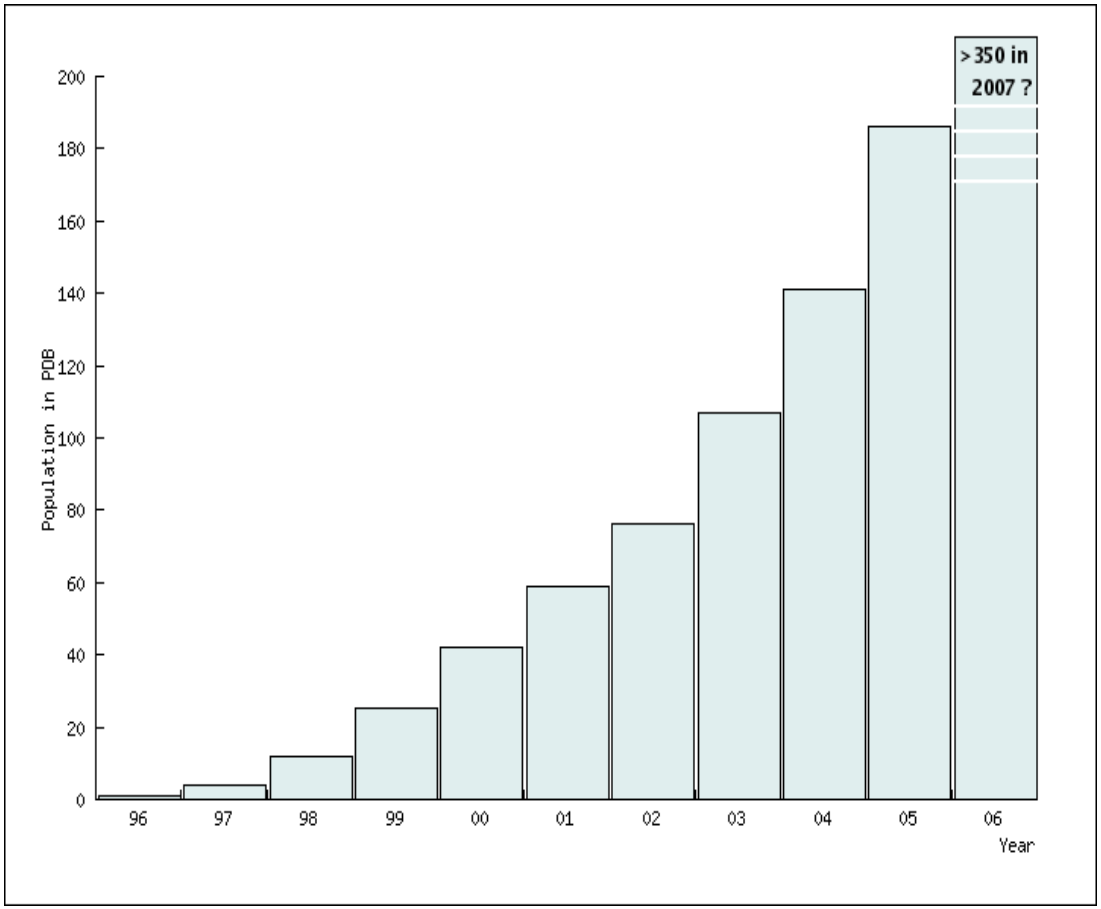
- Subfamily 3: Estrogen Receptor-like (Steroid hormone receptor)
  - Group A: Estrogen receptor (Sex hormone receptors)
  - Group B: Estrogen related receptor
  - Group C: 3-Ketosteroid receptors
- Subfamily 4: Nerve Growth Factor IB-like
  - Group A: NGFIB/NURR1/NOR1
- Subfamily 5: Steroidogenic Factor-like
  - 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.

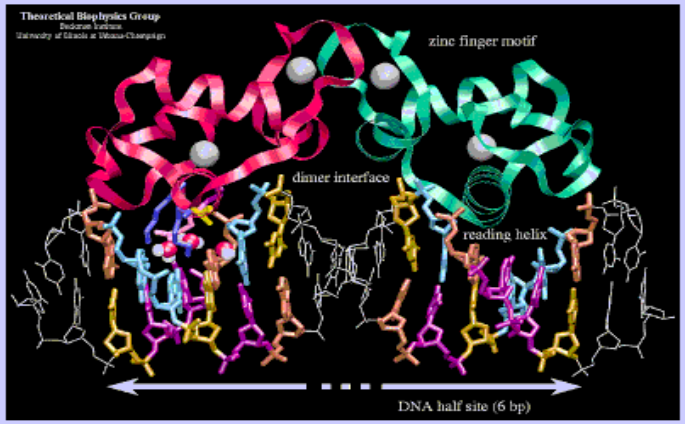
• Structures are obtained from the LDB or the DBD, but not from the full protein.

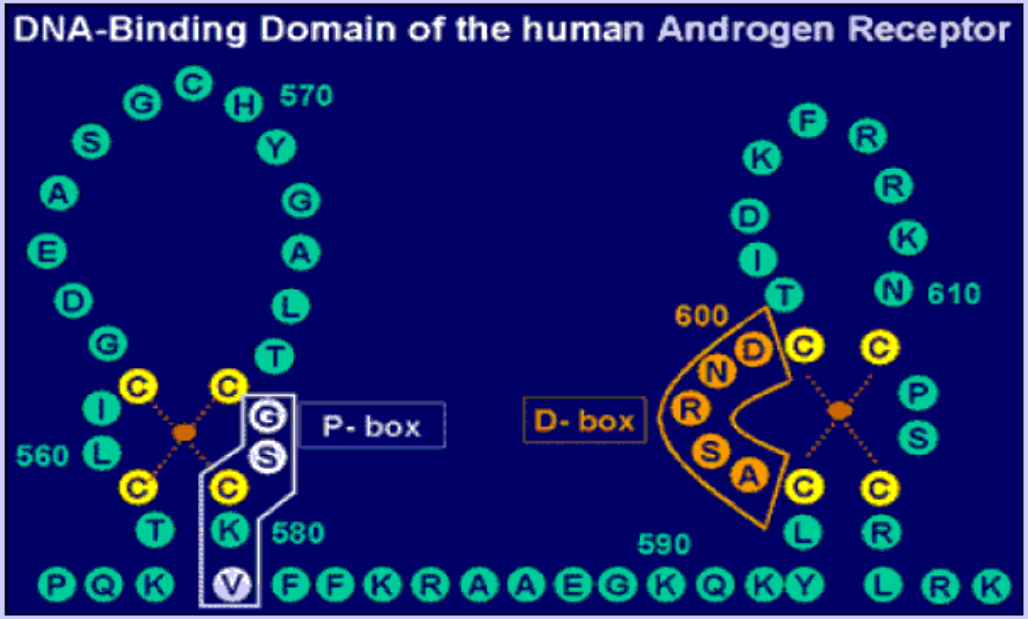
• Many new NRs are temporarily classified as Orphan receptors.



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

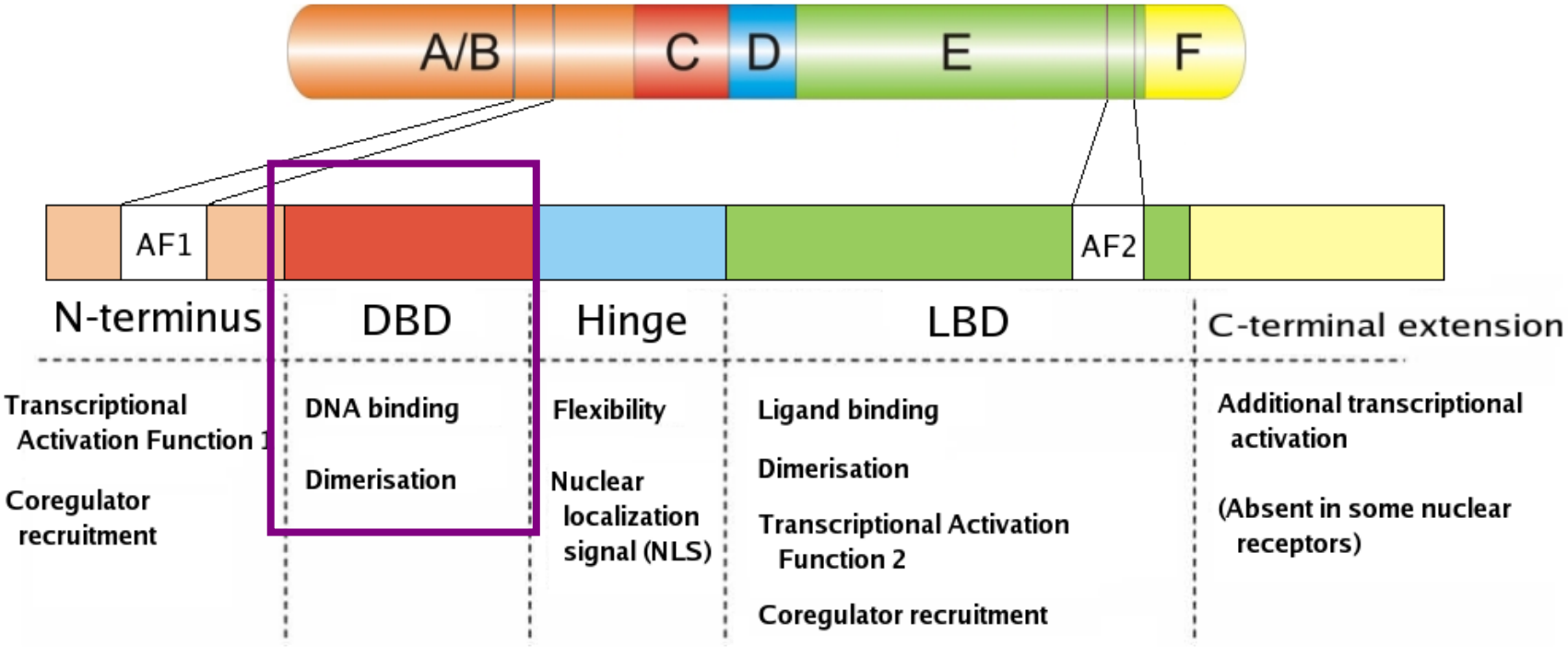
DNA Binding Domain





DNA Binding Domain

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

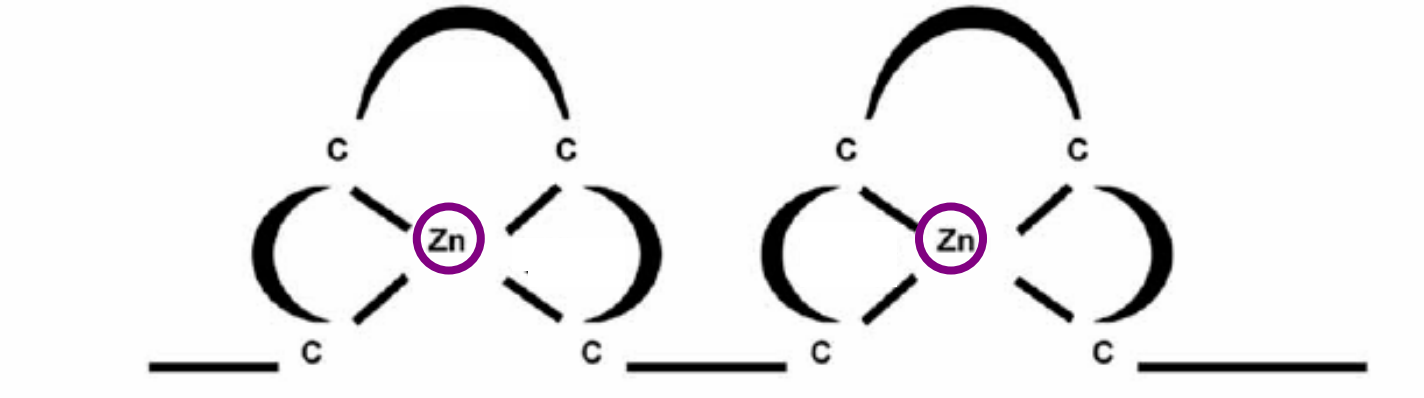


DNA Binding Domain

• The DNA Binding Domain (DBD) is a highly conserved domain in the family of Nuclear Receptors.

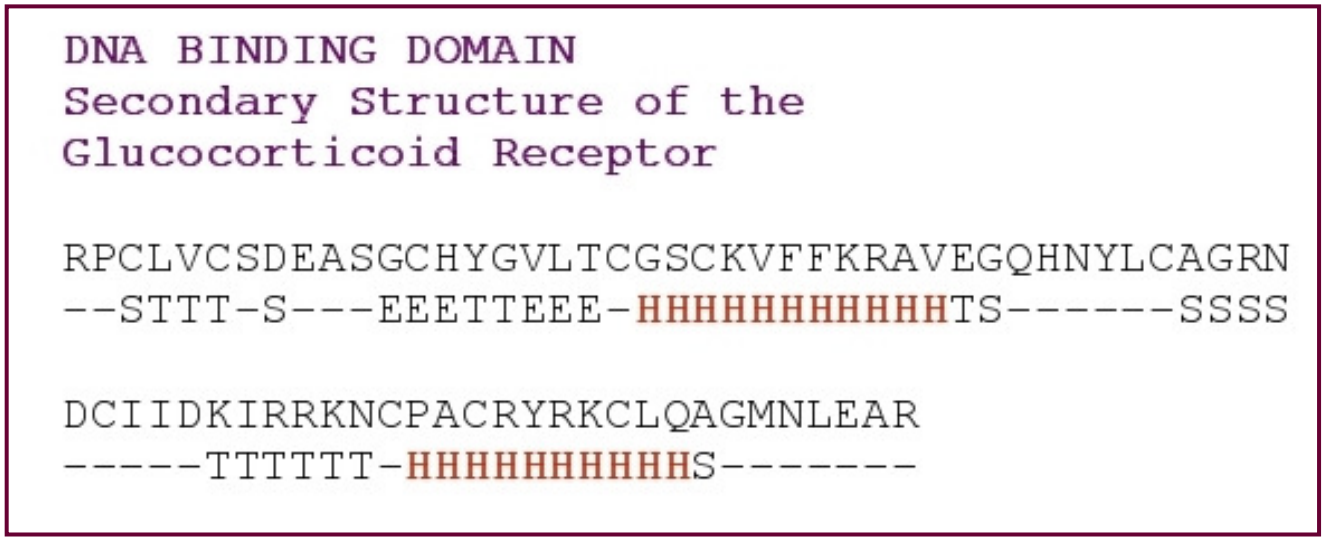
• The DBD consists of about 70 residues that bind to activating elements of DNA called hormone response elements.

• 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:

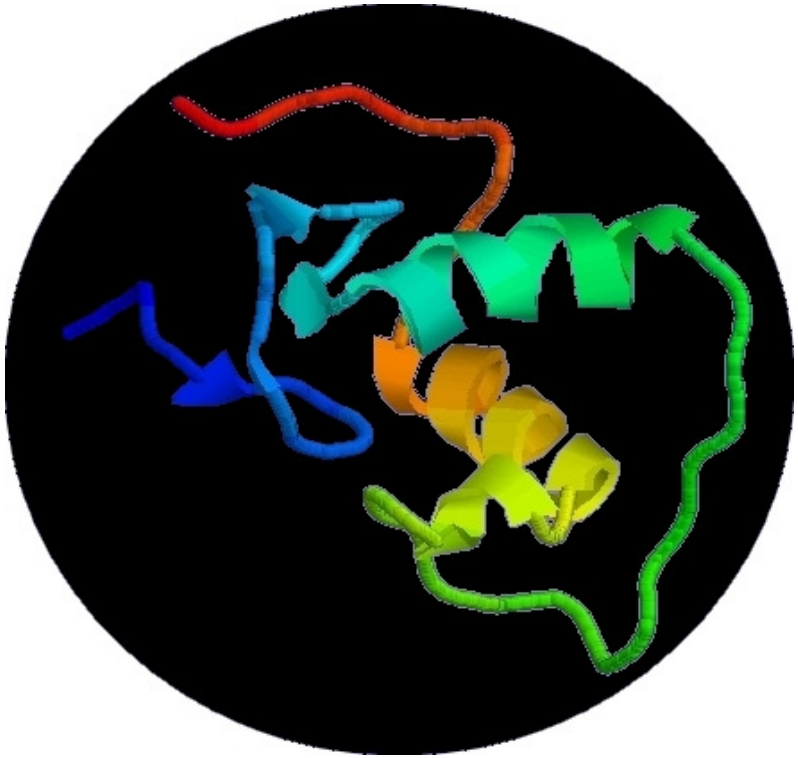


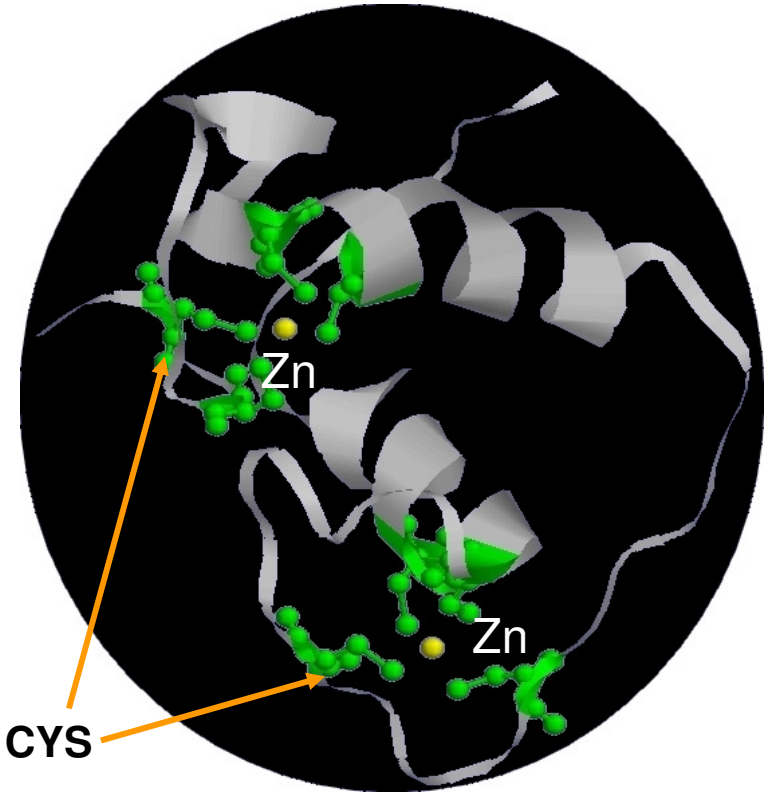
• The DNA binding domain of all nuclear receptors contains two α-helices.

• 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

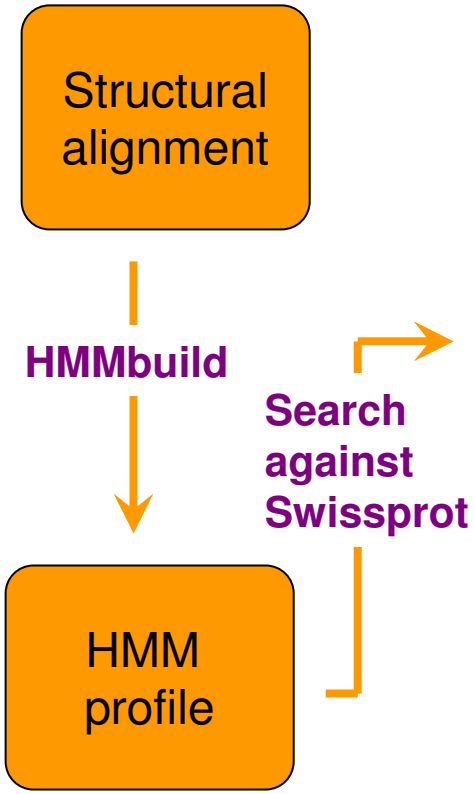
DBD - Conservation

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          MKPARPCLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQ--HNYL
RAR         -----PCFVCQDKSSGYHYGVSACEGCKGFFRRSIQKN--MVYT
RXR         -FTKHICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKD--LTYT
TR          ---DELCVVCGDKATGYHYRCITCEGCKGFFRRTIQKNLHPSYS
VDR         ----LLCKVCGDVASGFHYGVLACEGCKGFFRRSIQQN-IQYKR
                     * .* * ..* **    .*  ** **.*..

ERbeta      CPATNECEITKRRRKSCQACRFMKALKVGMLKEGVRLDRVRGGR
ER          CPATNQCTIDKNRRKSCQACRLRKCYEVGMMKG-----------
GR          CAGRNDCIIDKIRRKNCPACRYRKCLQAGMNLEARKTKK-----
RAR         CHRDKNCIINKVTRNRCQYCRLQKCFEVGMSKESVRND------
RXR         CRDNKDCLIDKRQRNRCQYCRYQKALAMGMKREAVQEERQRG--
TR          CKYEGKCVIDKVTRNQCQECRFKKCIYVGMATDLVLDDSKRLAK
VDR         CLKNENCSIVRINRNRCQQCRFKKCLSVGMSRDAVRFGR-----
            *     * * .  *  *  **  *.   **
```

Characterization of the NRs

The DNA binding domain characterizes the family of nuclear receptors.

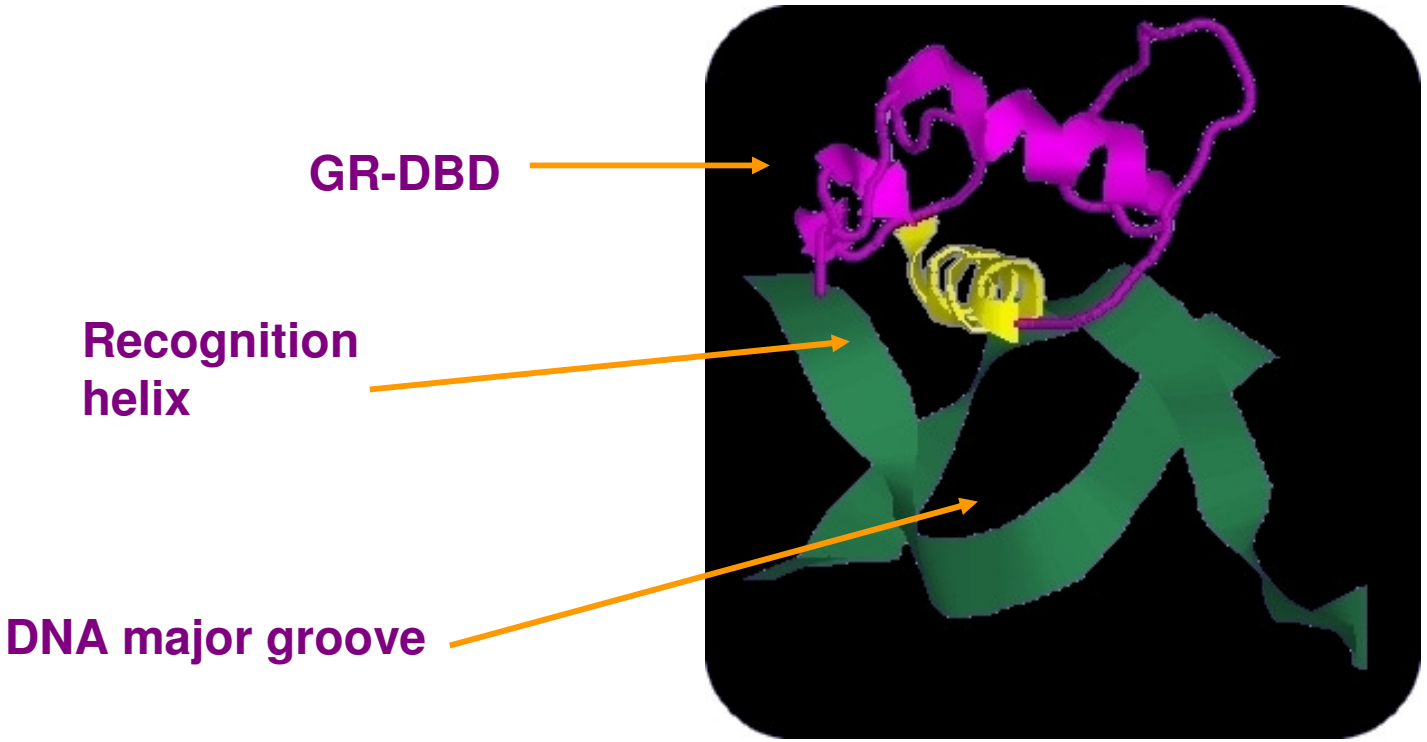


Sequence	Description	Score	E-value
sp|P11475|ERR2_HUMAN	STEROID HORMONE RECEPT	220.0	5.3e-62
sp|Q91250|ESR1_POEGU	ESTROGEN RECEPTOR	198.5	1.5e-55
sp|P19793|RXRA_HUMAN	RETINOIC ACID RECEPT	195.4	1.4e-54
sp|P18117|THBA_XENLA	THYROID HORMONE RECEPT	191.2	2.6e-53
sp|P20393|NRD1_HUMAN	ORPHAN NUCLEAR RECEPT	182.2	1.3e-50
sp|P04150|GCR_HUMAN	GLUCOCORTICOID RECEPT	179.2	1e-49
sp|P49700|USP_BOMMO	ULTRASPIRACLE PROTEIN	175.3	1.5e-48
sp|P08235|MCR_HUMAN	MINERALOCORTICOID RECEPT	168.5	1.7e-46
|	|	|	|
sp|Q09565|NH20_CAEEL	STEROID HORMONE RECEPT	47.4	4.8e-10
sp|Q22127|NH40_CAEEL	STEROID HORMONE RECEPT	33.9	5.7e-06
sp|Q23489|NHR9_CAEEL	NUCLEAR HORMONE RECEPT	26.6	0.00092
sp|Q09587|NH22_CAEEL	STEROID HORMONE RECEPT	24.8	0.0032
sp|P41933|ODR7_CAEEL	NUCLEAR HORMONE RECEPT	22.4	0.017

Interactions DNA - DBD

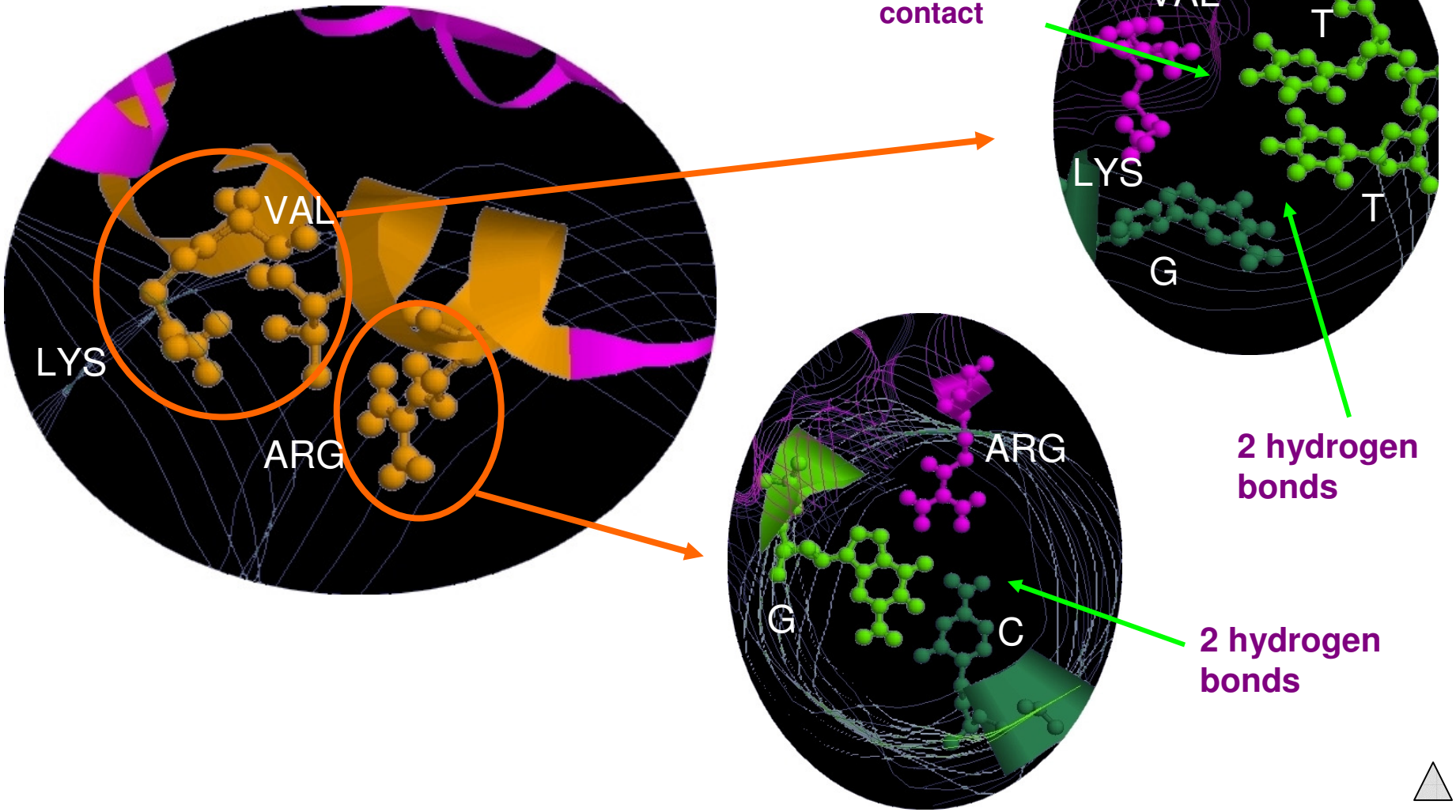
• The first α-helix in the zinc motif forms sequence-specific interactions with the edge of the bases in the major groove of one DNA strand.

•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

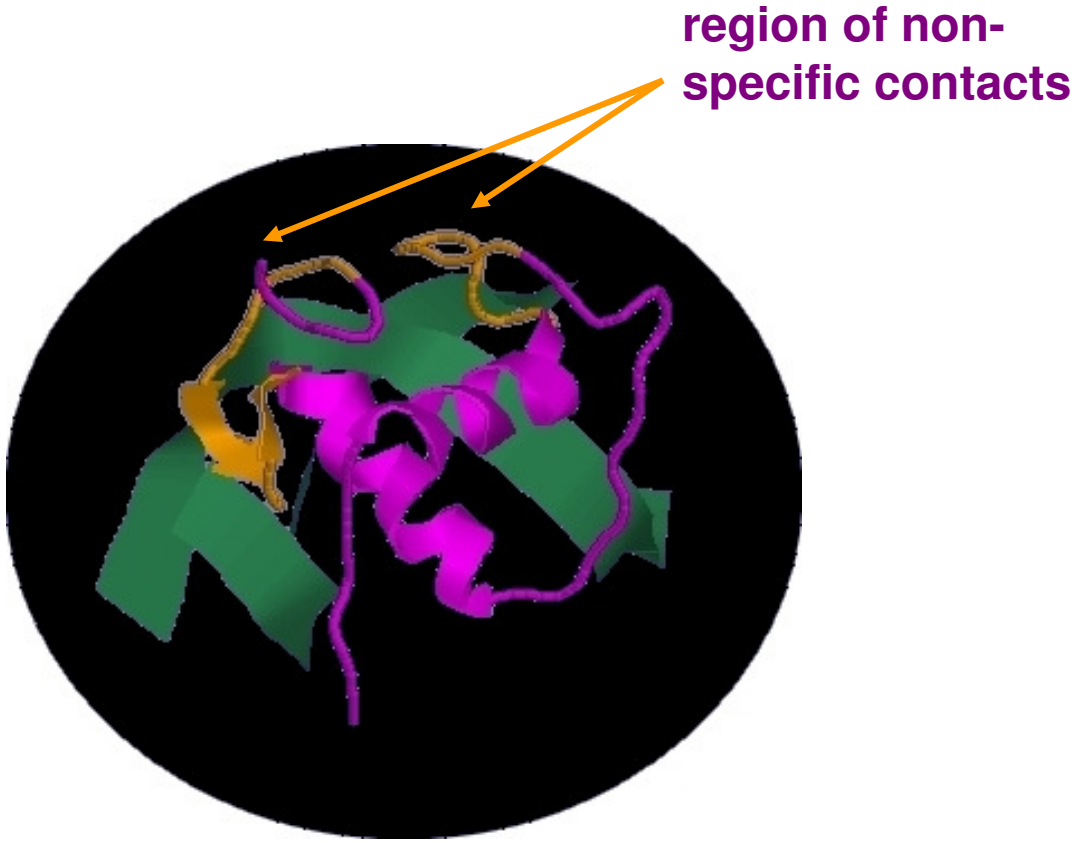
- The LYS and ARG are conserved in the nuclear receptor family:

```
CLUSTAL W multiple sequence alignment
DNA BINDING DOMAIN

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

ERbeta      CPATNECEITKRRRKSCQACRFMKALKVGMLKEGVRLDRVRGGR
ER          CPATNQCTIDKNRRKSCQACRLRKCYEVGMMKG----------
GR          CAGRNDCIIDKIRRKNCPACRYRKCLQAGMNLEARKTKK-----
RAR         CHRDKNCIINKVTRNRCQYCRLQKCFEVGMSKESVRND------
RXR         CRDNKDCLIDKRQRNRCQYCRYQKALAMGMKREAVQEERQRG--
TR          CKYEGKCVIDKVTRNQCQECRFKKCIYVGMATDLVLDDSKRLAK
VDR         CLKNENCSIVRINRNRCQQCRFKKCLSVGMSRDAVRFGR-----
            *     * * .  *  *  **  *.   **
```

Non-specific Interactions

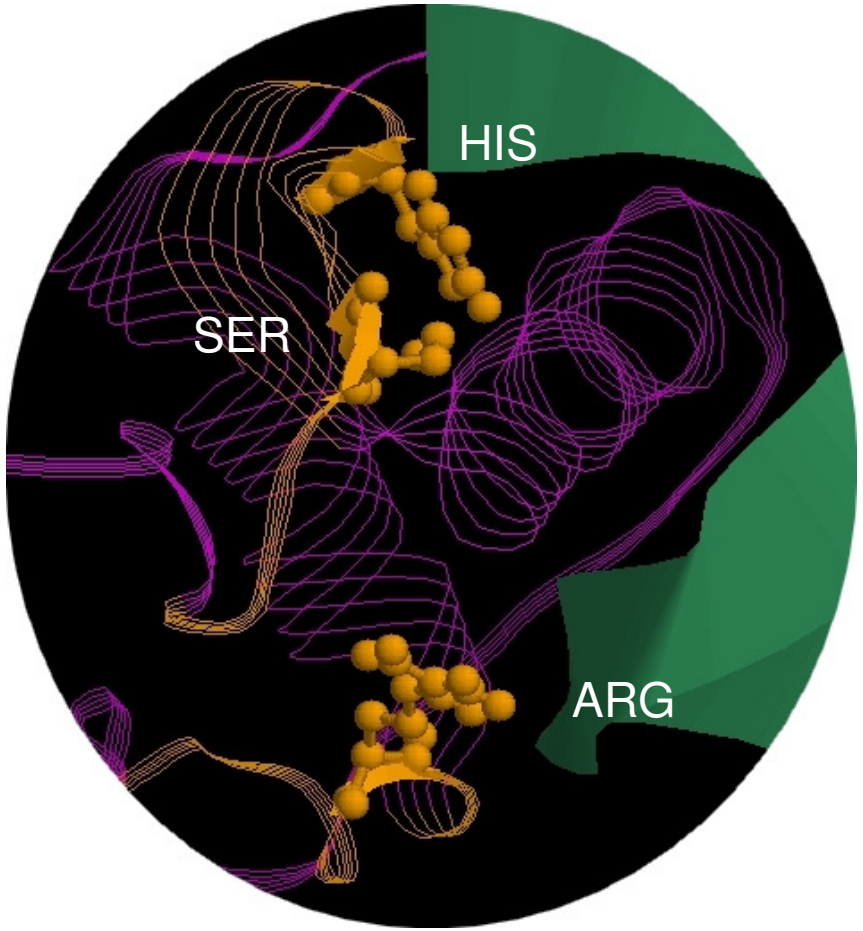
• The recognition helix is positioned and oriented in the major groove by a number of non-specific interactions between the phosphate groups and protein side chains.

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

ERbeta      CPATNECEITKRRRKSCQACRFMKALKVGMLKEGVRLDRVRGGR
ER          CPATNQCTIDKNRRKSCQACRLRKCYEVGMMKG----------
GR          CAGRNDCIIDKIRRKNCPACRYRKCLQAGMNLEARKTKK-----
RAR         CHRDKNCIINKVTRNRCQYCRLQKCFEVGMSKESVRND------
RXR         CRDNKDCLIDKRQRNRCQYCRYQKALAMGMKREAVQEERQRG--
TR          CKYEGKCVIDKVTRNQCQECRFKKCIYVGMATDLVLDDSKRLAK
VDR         CLKNENCSIVRINRNRCQQCRFKKCLSVGMSRDAVRFGR-----
            *     * * .  *  *  **  *.   **
```

Structural conservation

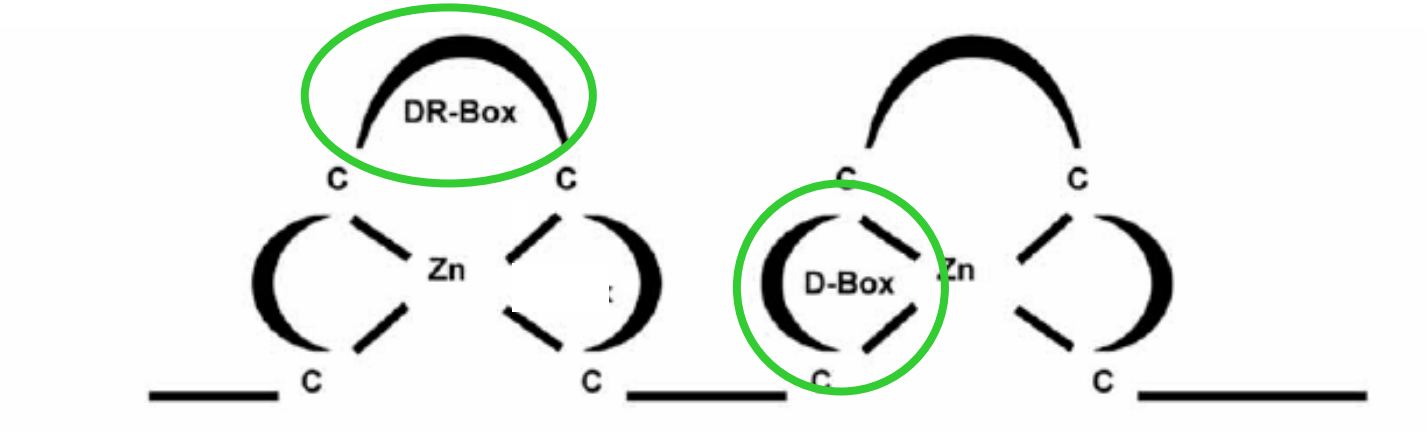
Conservation of the zinc finger cysteines, the specific contact residues and the non-specific contacts in a structural alignment:

```
STAMP - STRUCTURAL SUPERPOSITION - DNA BINDING DOMAIN

ER          -MKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRS--IQGHN-DYM
ERbeta      A-IPKRLCLVCGDIASGYHYGVASCEACKAFFKRTIQG--NI-EYS
RXR         -F-TKHICAICGDRSSGKHYGVYSCEGCKGFFKRTVRK-D-L-TYT
GR          -MKPARPCLVCSDEASGCHYGVLTCGSCKVFFKRAVE-G-QH-NYL
VDR         -----LLCKVCGDVASGFHYGVLACEGCKGFFRRSIQQ-N-IQYKR
RAR         ------PCFVCQDKSSGYHYGVSACEGCKGFFRRSIQKN-M-V-YT
TR          ----DELCVVCGDKATGYHYRCITCEGCKGFFRRTIQKNLHPS-YS

ER          CPATNQCTI--DKNRRKSCQACRLRKCYEVGMM-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--NRCQYCRLQKCFEVGMSKES-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

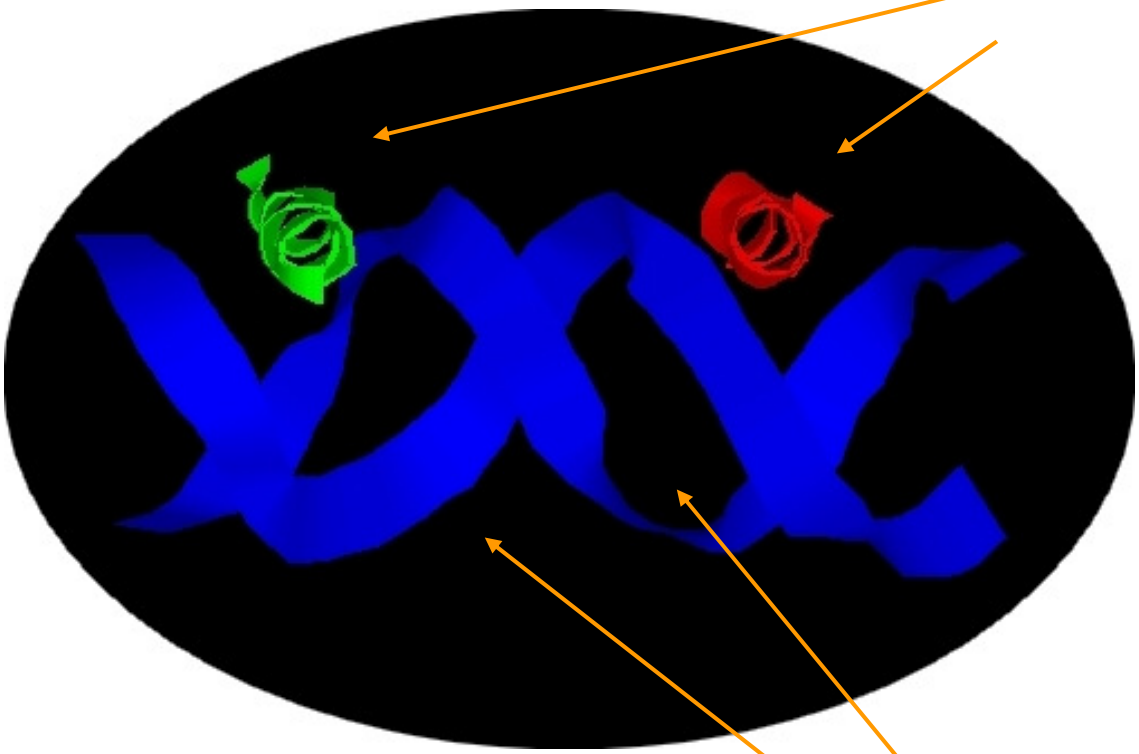
- The two nuclear receptors in the dimer bind to the DNA backbone through specific and non-specific interactions, as describe above.

- 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

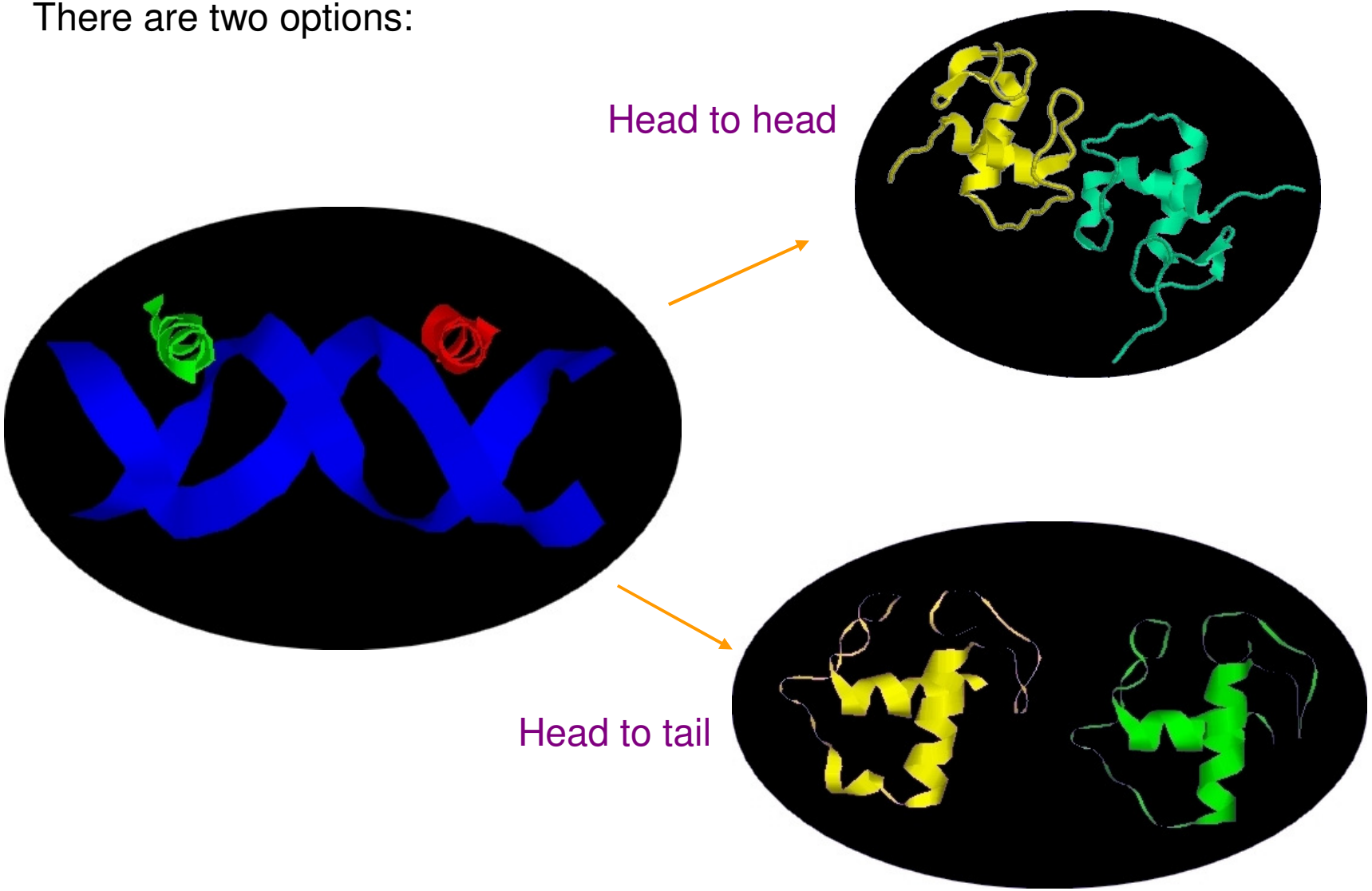


Recognition helices

Two consecutive major grooves

Dimerization

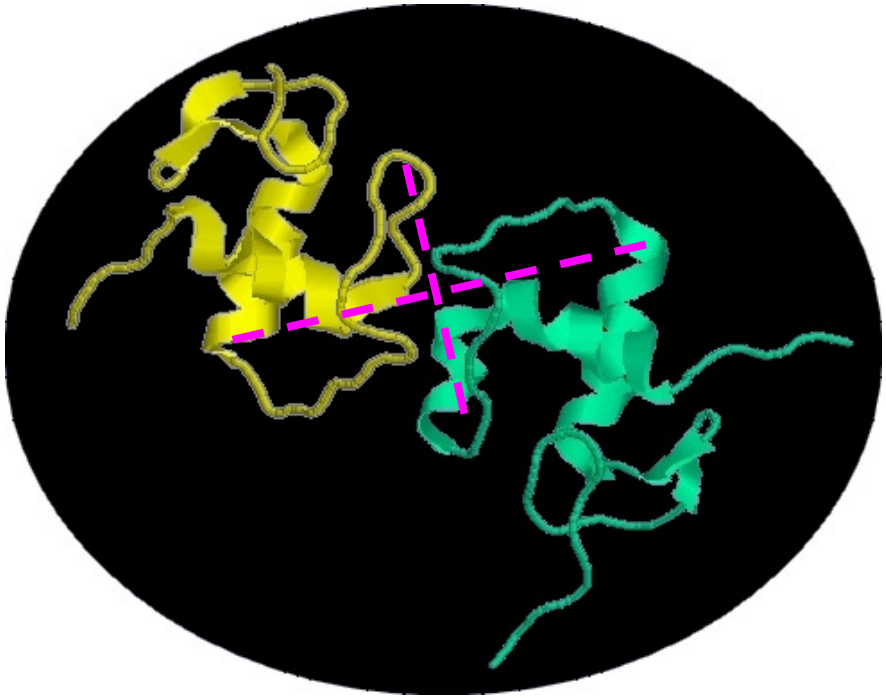
There are two options:



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:

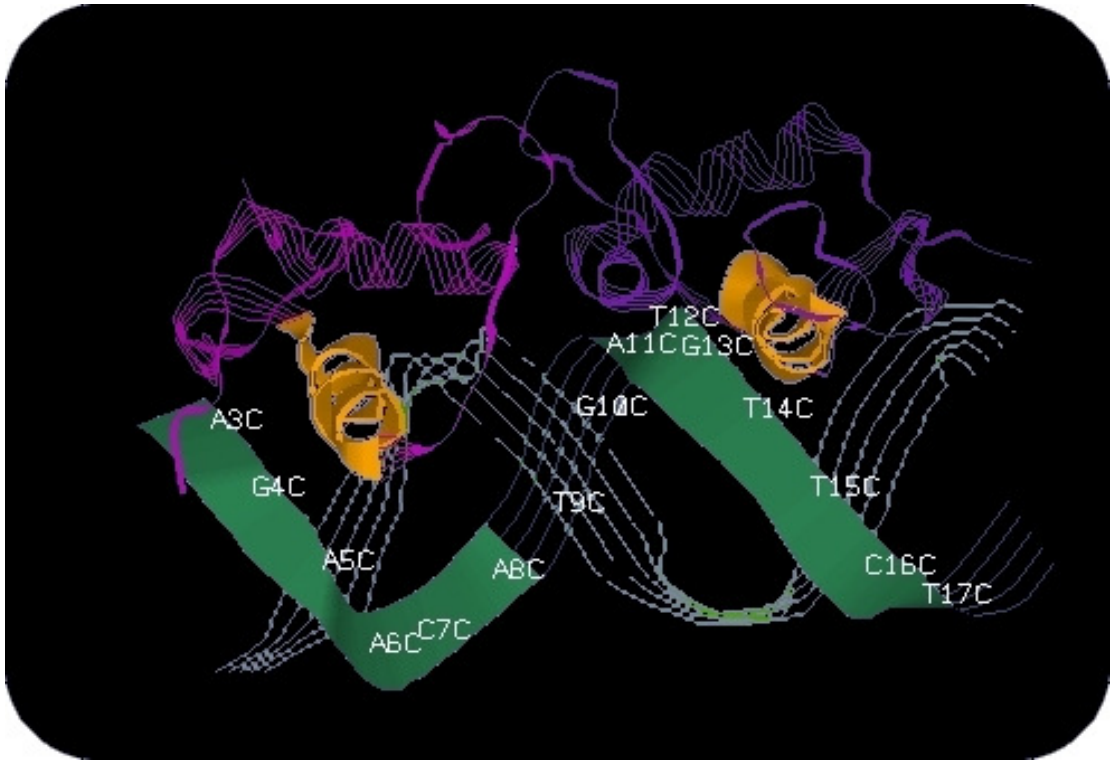


Homodimerization

• 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’ AGAACAxxxTGTTCT 3’
3’ TCTTGTxxxACAAGA 5’

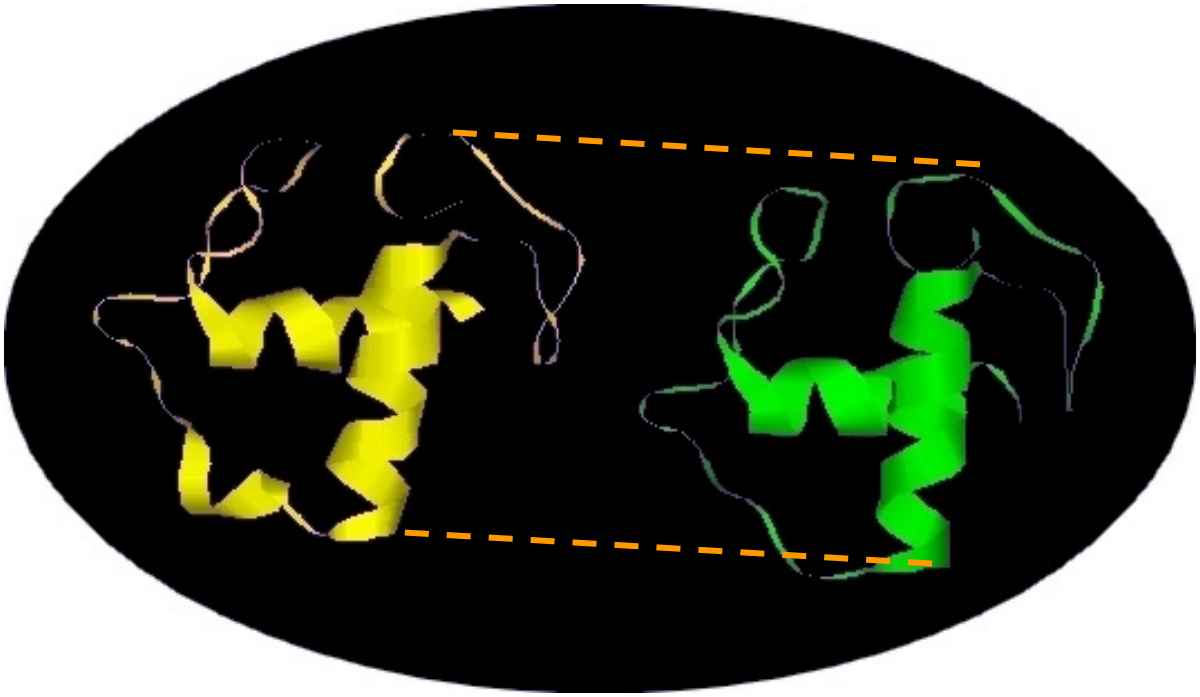


AGAACA xxx TGTTCT

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.

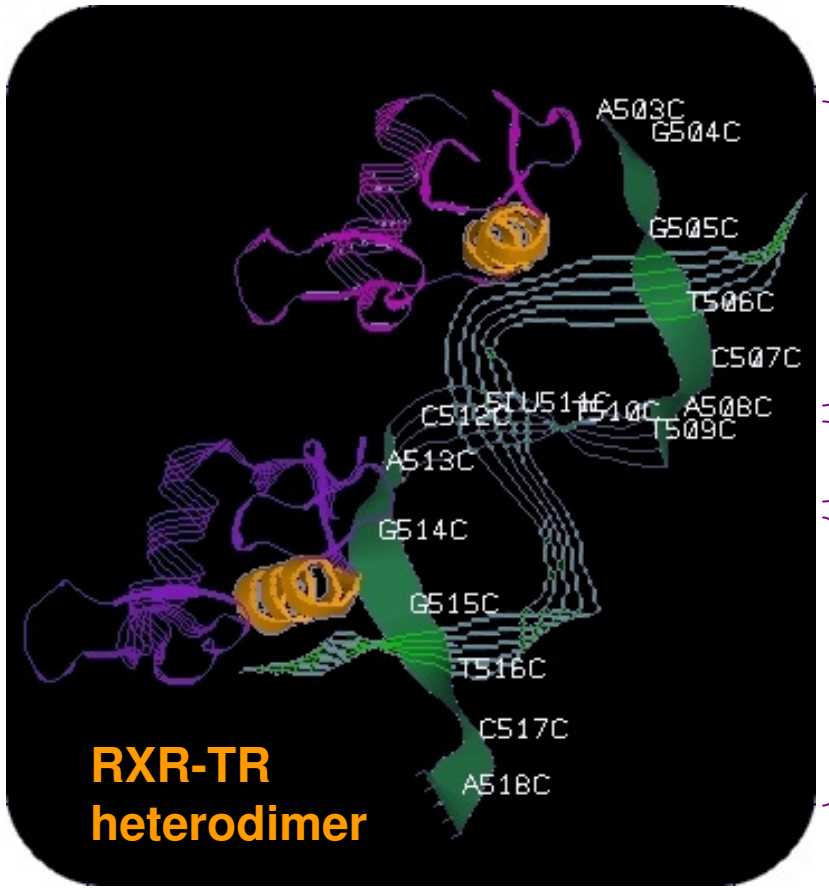
• The spacing between the direct repeats is diferent for each receptor dimer combination, and hence it determines the DNA specificity of each RXR heterodimer.

Response elements:

- RXR-VDR:
AGGTCAxxxAGGTCA
TCCAGTxxxTCCAGT

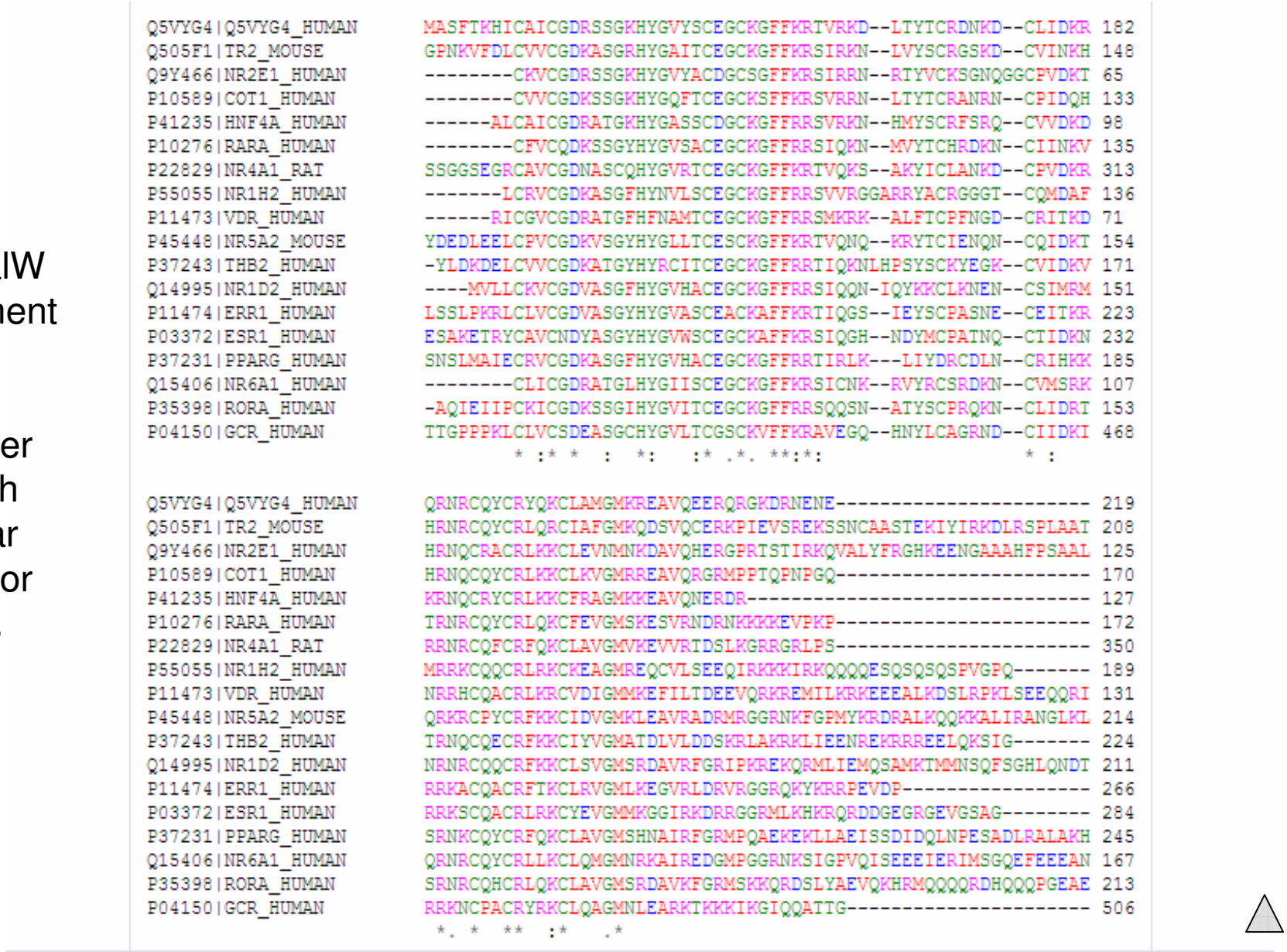
- RXR- TR:
AGGTCAxxxxAGGTCA
TCCAGTxxxxTCCAGT

- RXR-RAR:
AGGTCAxxxxAGGTCA
TCCAGTxxxxTCCAGT



ClustalW alignment

ClustalW alignment with one member of each nuclear receptor family.



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.

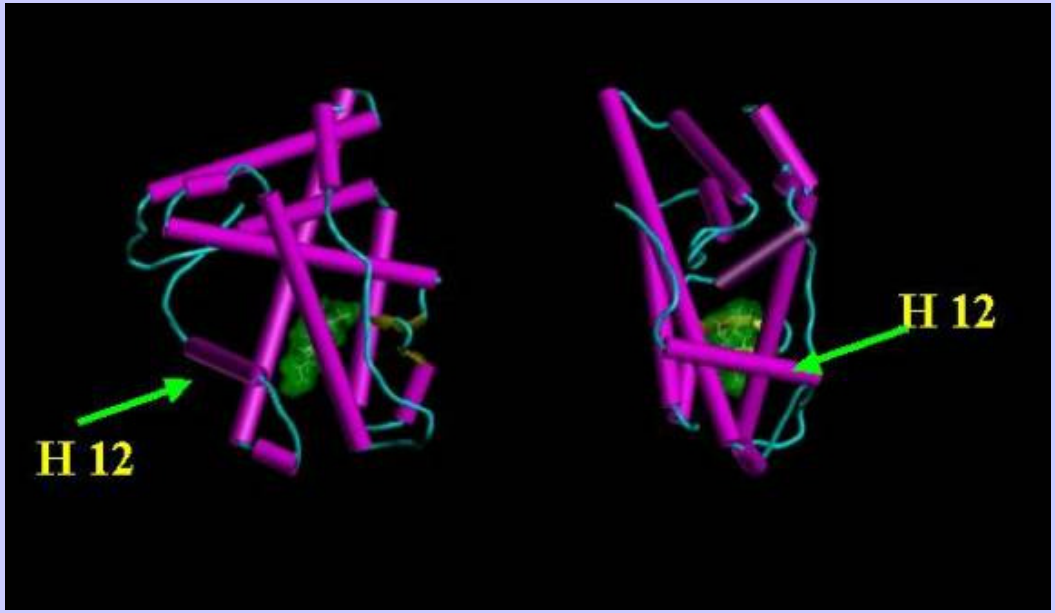
•The DBD is a highly conserved domain in the family of nuclear receptors.

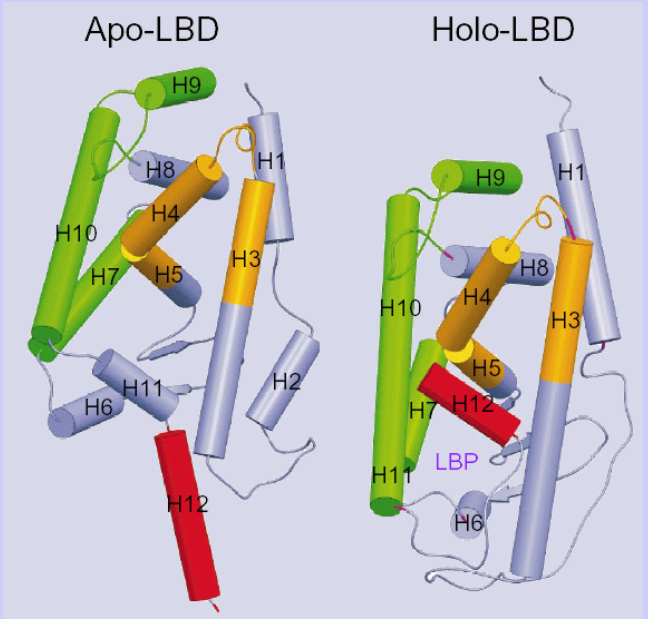
•The cysteine residues and the residues that form the specific and non-specific interactions with DNA are conserved in the whole family.

•The DBD characterizes the family of nuclear receptors.

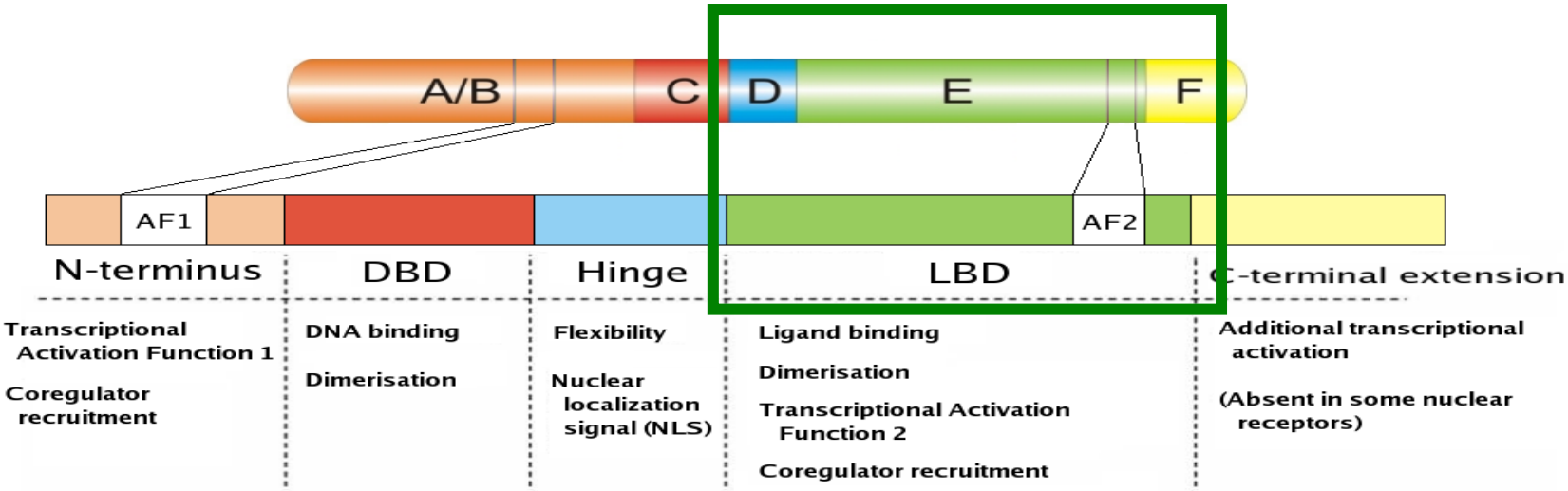
•Depending on the type of dimerization (homo or hetero), there are two type of DNA recognition: invers or direct repeats.

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 diplays 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)
- 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

. **APO: NR1C3 (PPAR)**

. **HOLO(+): NR1A2 (TR)**

. **HOLO(-): NR1C1 (PPAR)**

Subfamily 2: Retinoid X Receptor-like

- Group A: Hepatocyte nuclear factor-4 (HNF4)
- Group B: Retinoid X receptor (RXRα)
- Group C: Testicular receptor
- Group E: TLX/PNR
- Group F: COUP/EAR

. **APO: NR2B1 (RXR)**

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

. **HOLO(-): ?**

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

- Group A: Estrogen receptor (Sex hormone receptors; sex hormones: Estrogen)
- Group B: Estrogen related receptor
- Group C: 3-Ketosteroid receptors

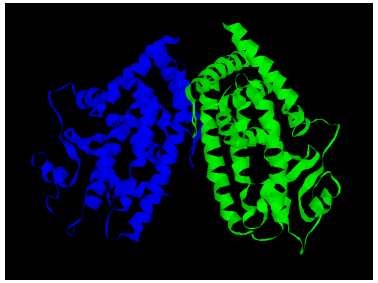
. **APO: NR3B3**

. **HOLO(+): NR3A1**

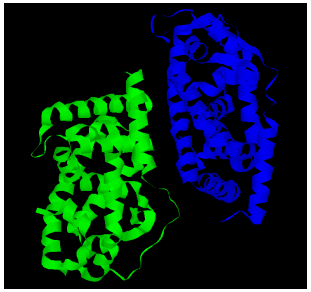
. **HOLO(-): NR3A1**

Structural conservation through families

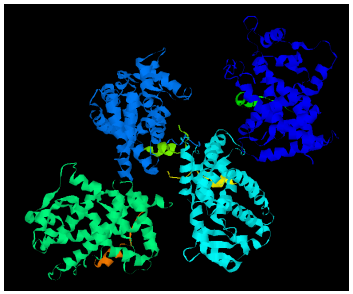
Subfamily 1: Thyroid Hormone Receptor-like



APO: 1prg.pdb (R=2.4 Å)



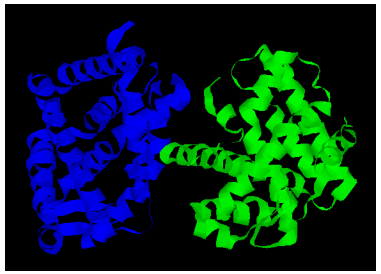
HOLO(+): 1n46.pdb (R=2.2 Å)



HOLO(-): 1kkq.pdb (R=3.0 Å)

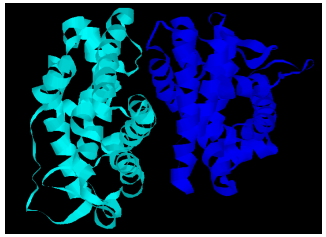
Subfamily 2: Retinoid X Receptor-like

APO: 1lbd.pdb (R=2.4 Å)

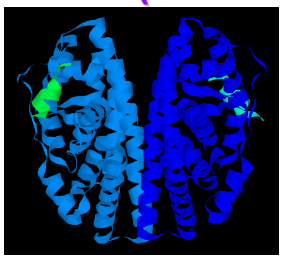


HOLO(+): 1fby.pdb (R=2.25 Å)

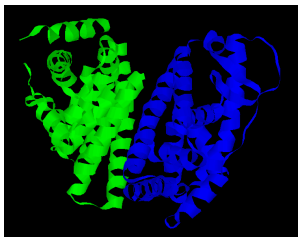
Subfamily 3: Estrogen Receptor-like (Steroid hormone receptor)



APO: 1kv6.pdb (R=2.7 Å)



HOLO(+): 3erd.pdb (R=2.03 Å)



HOLO(-): 1err.pdb (R=2.6 Å)

LBD 3D structures

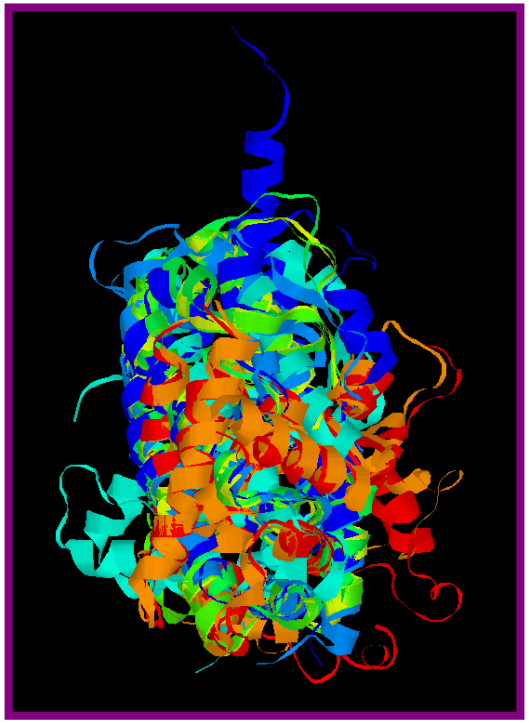
- The 3D structures of crystallised LBDs superimposed showed that the overall structures of the LBDs of different nuclear receptors are similar, revealing a canonical fold for the nuclear receptor LBD.

STAMP code:

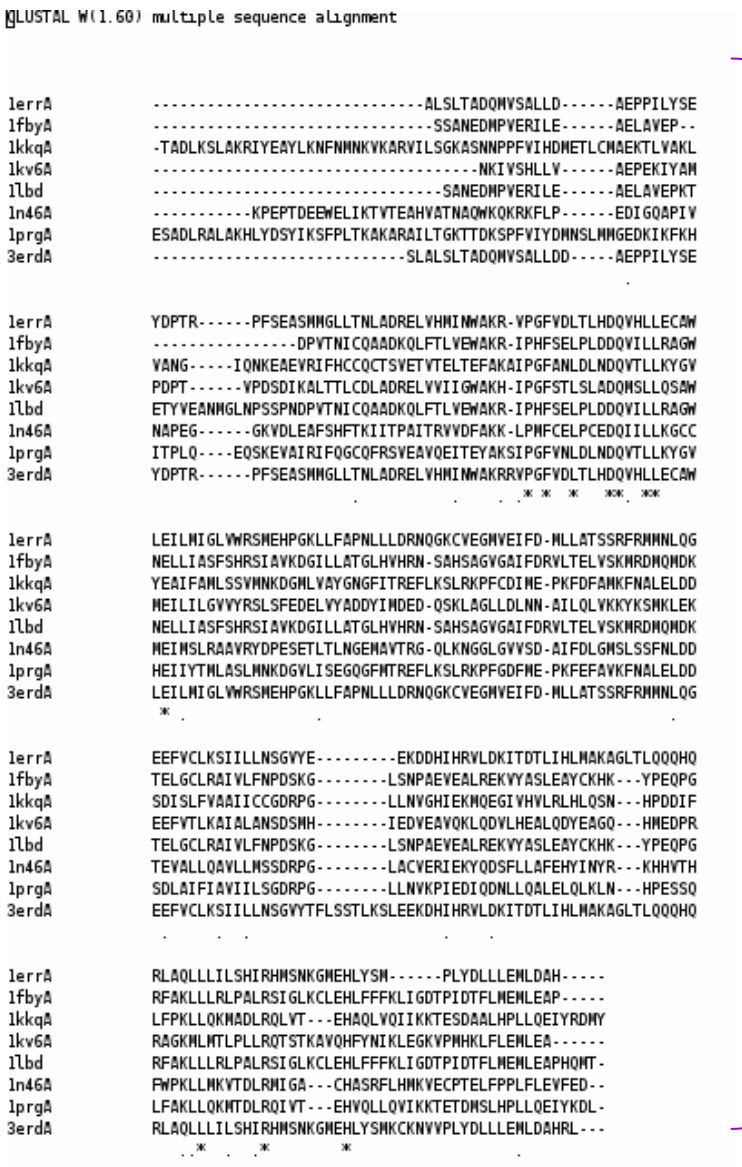
Alignment score Sc: 1.60

Alignment length Lp: 324

RMSD: 3.83



ClustalW sequence alignment

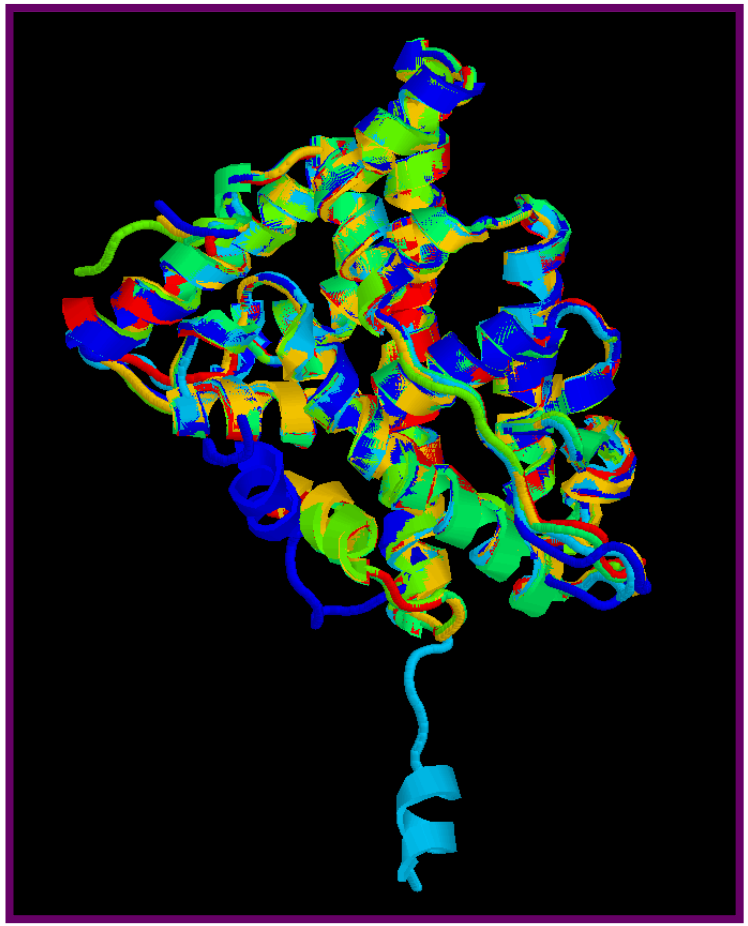


- Homology or Remote homology?

All human

NR3A1sequence and structural alignment

```
CLUSTAL W(1.60) multiple sequence alignment


1g50A           ---NSLALSLTADQMVSALLDAEPPILY-SEYDPTRPFSEASMMGLLTNLADRELVHMIN
1gwqA           SKKNSLALSLTADQMVSALLDAEPPILY------SEPFSEASMMGLLTNLADRELVHMIN
1qktA           ---NSLALSLTADQMVSALLDAEPPILYS-EYDPTRPFSEASMMGLLTNLADRELVHMIN
1l2iA           ----SLALSLTADQMVSALLDAEPPILYSSEYDPTRPFSEASMMGLLTNLADRELVHMIN
1a52A           -----LALSLTADQMVSALLDAEPPILYSEEYDPTRPFSEASMMGLLTNLADRELVHMIN
3erdA           ----SLALSLTADQMVSALLDDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMIN
                     ***************   * .       . ************************

1g50A           WAKR-VPGFVDLTLHDQVHLLEC-AWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKC
1gwqA           WAKR-VPGFVDLTLHDQVHLLEC-AWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKC
1qktA           WAKR-VPGFVDLTLHDQVHLLES-AWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKS
1l2iA           WAKRRVPGFVDLTLHDQVHLLECCAWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGK-
1a52A           WAKR-VPGFVDLTLHDQVHLLEC-AWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKC
3erdA           WAKRRVPGFVDLTLHDQVHLLEC-AWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKC
                **** *****************. ***********************************

1g50A           VEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHR
1gwqA           VEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHR
1qktA           VEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHR
1l2iA           VEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVY---------LEEKDHIHR
1a52A           VEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHR
3erdA           VEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHR
                ****************************************         *********

1g50A           VLDKITDTLIHLMAKAGLTLQQQHQR-LAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
1gwqA           VLDKITDTLIHLMAKAGLTLQQQHQR-LAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
1qktA           VLDKITDTLIHLMAKAGLTLQQQHQR-LAQLLLILSHIRHMSNKGMEHLYSMKSKNVVPL
1l2iA           VLDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSSHIRHMSNKGMEHLYSMKCKNVVPL
1a52A           VLDKITDTLIHLMAKAGLTLQQQHHERLAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
3erdA           VLDKITDTLIHLMAKAGLTLQQQHQR-LAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
                ************************.     **.. ******************.******

1g50A           YDLLLEMLDAHRLH-
1gwqA           YDLLLEMLDAHR---
1qktA           YDLLLEMLDAHRLHA
1l2iA           YDLLLEMLDAH----
1a52A           YDLLLEML-------
3erdA           YDLLLEMLDAHRL--
                ********
```



* STAMP RMSD: 0.44

- Great identity in sequence and structure -> Homology

All human

Searching homologues with psi-blast

- We perfomed the search with THYROID RECEPTOR β (NR1A2)

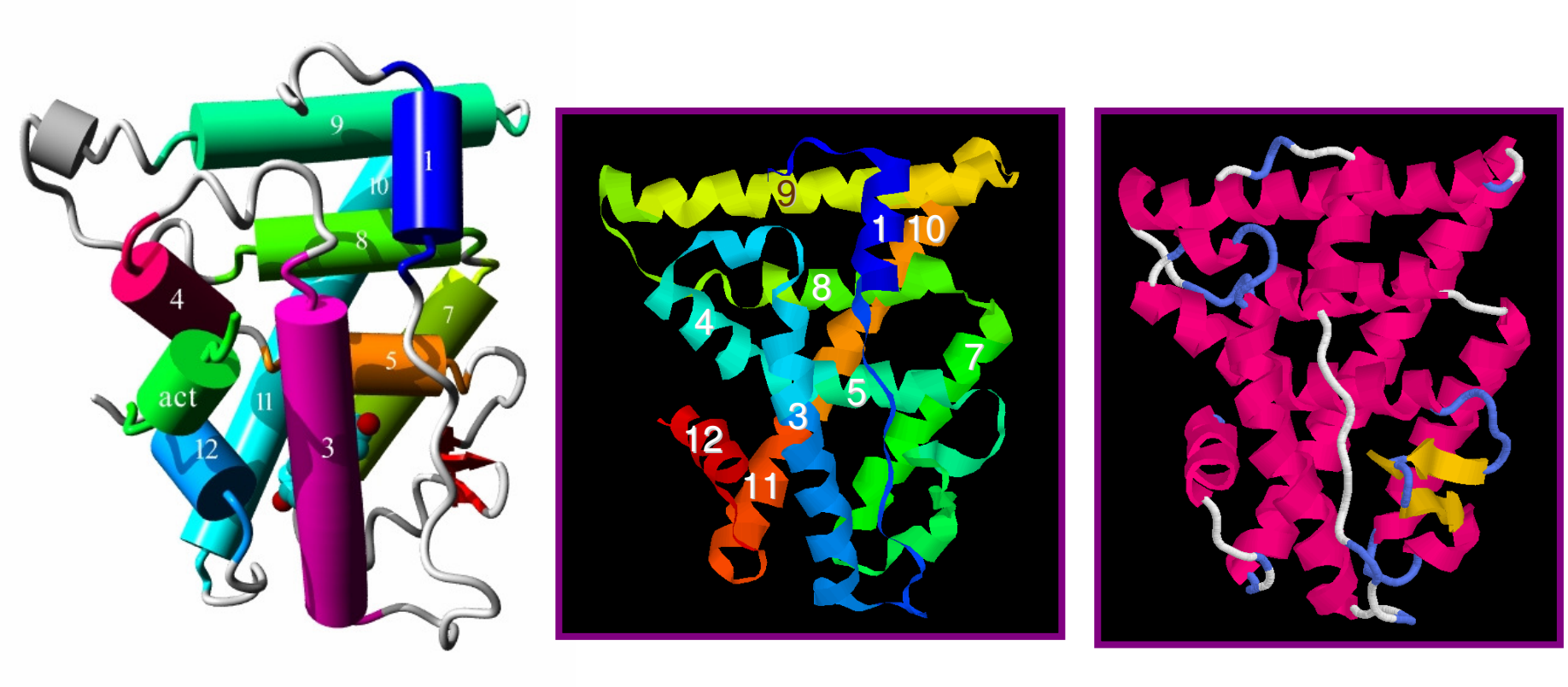
```
Results from round 3

                                                                  Score     E
Sequences producing significant alignments:                       (bits)  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|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

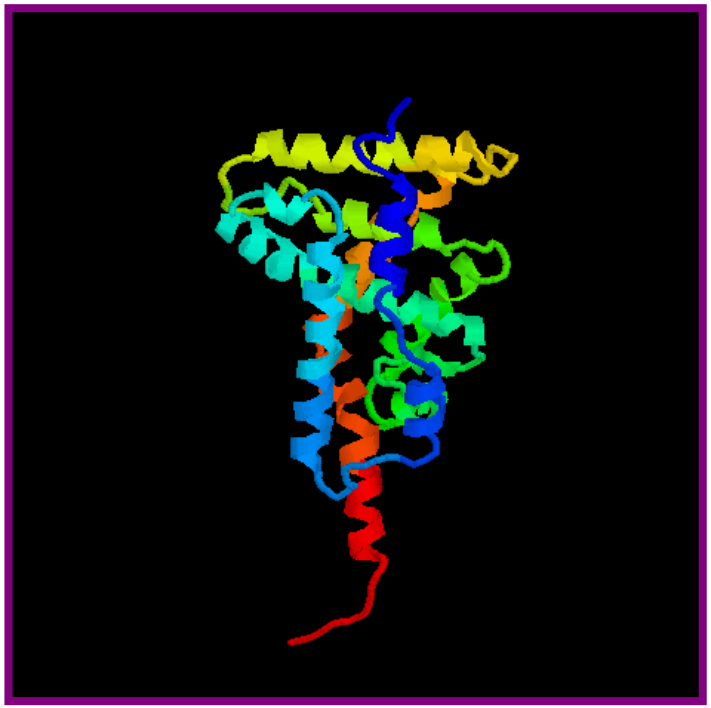
Secondary structure

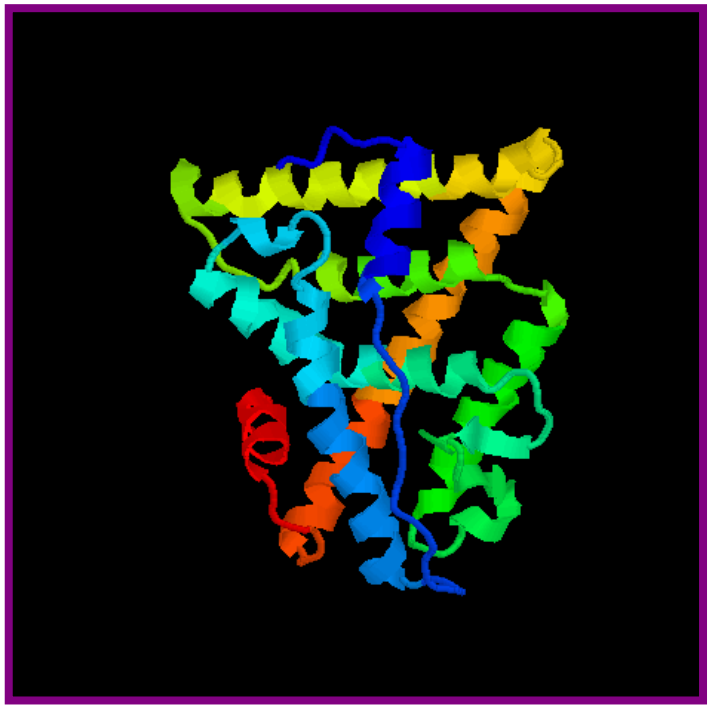
- The LBD domain forms a defined globular structure in which eleven to twelve helices are arranged together in an antiparallel, three-layered sandwich, which also includes 2-4 beta-strands.



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.



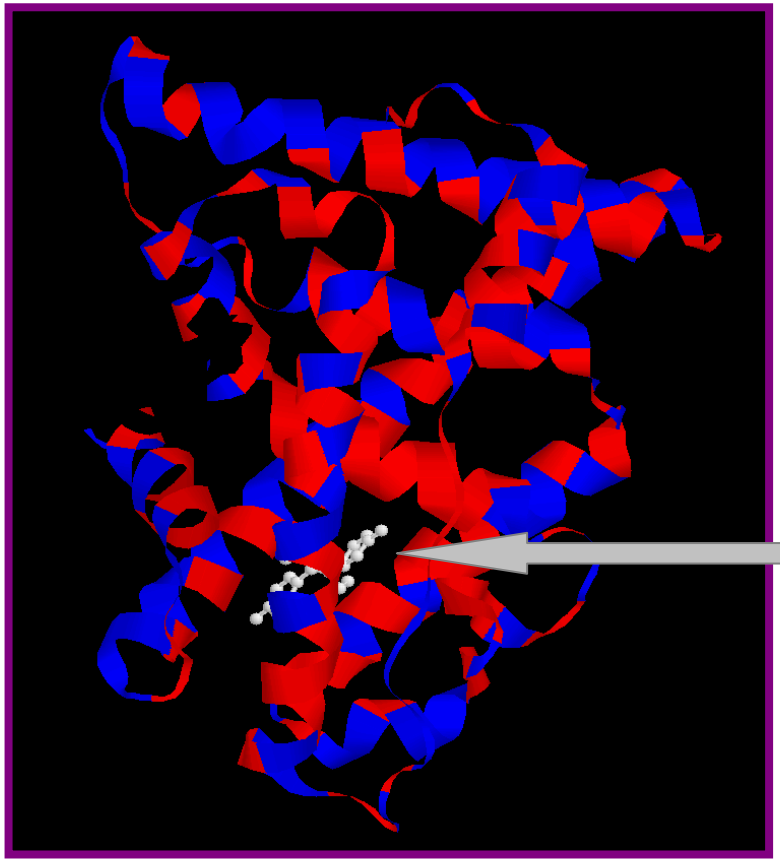


APO
Unliganded-structure

HOLO
Liganded-structure

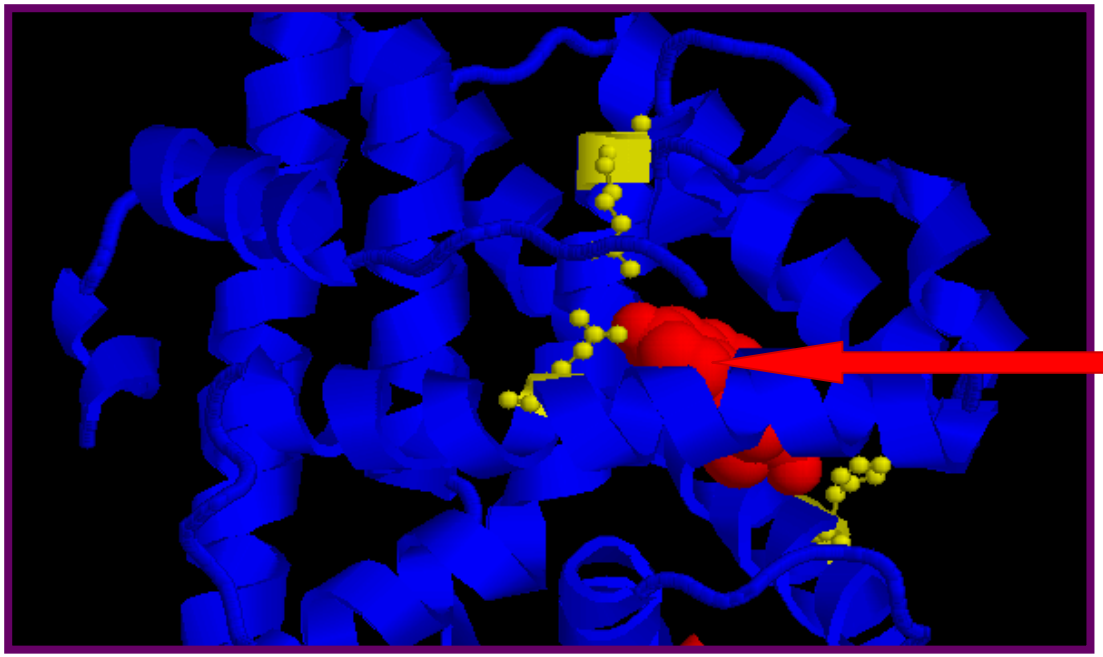
Holo-structures

- 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.

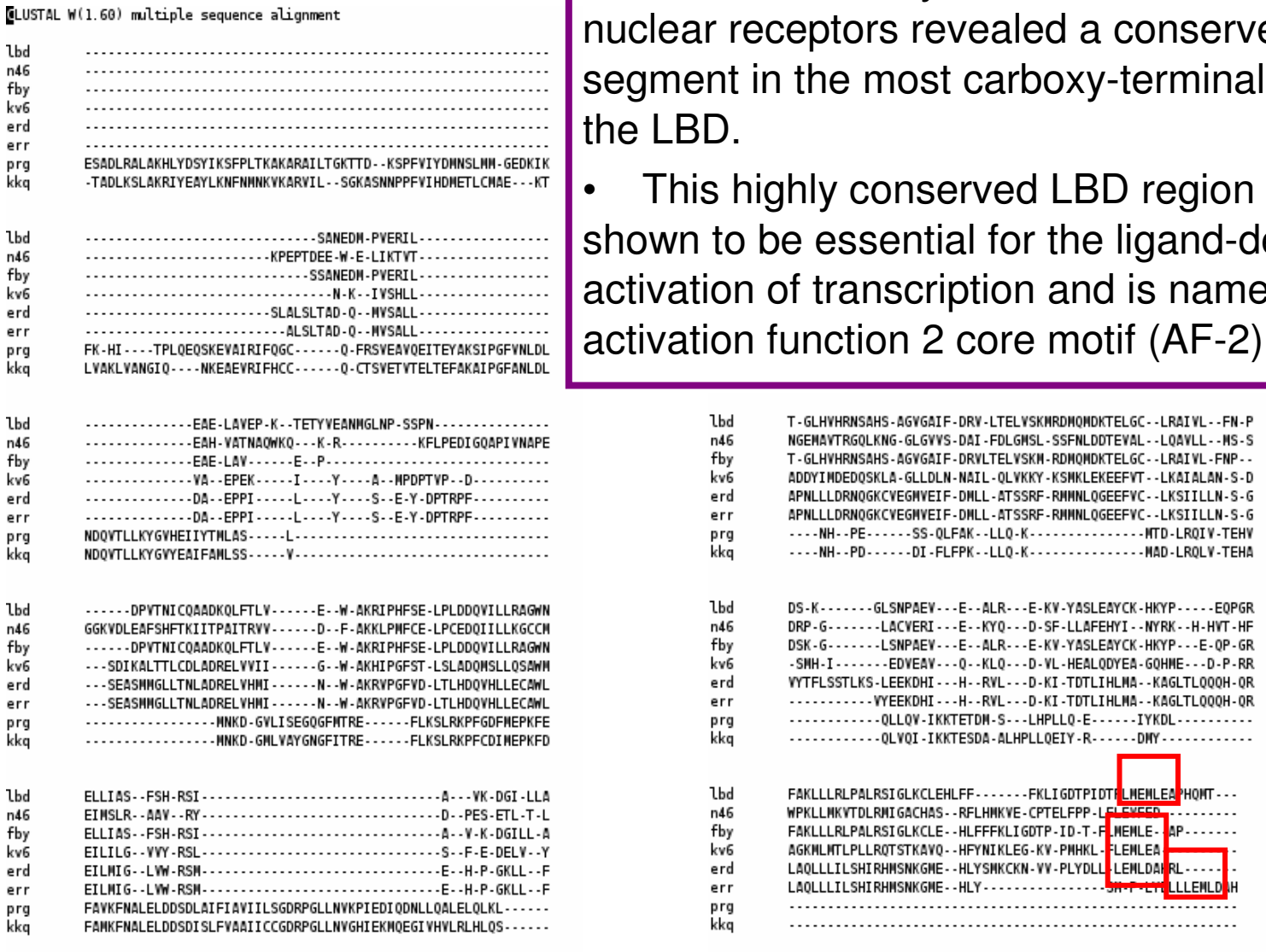


Holo-structures

- Ligand recognition is achieved through a combination of especific 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 hydroxil groups of the ligand.
  - The architecture of the pocket is rigid and only accommodate planar structures.



LBD helix 12 (H12)



- 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).

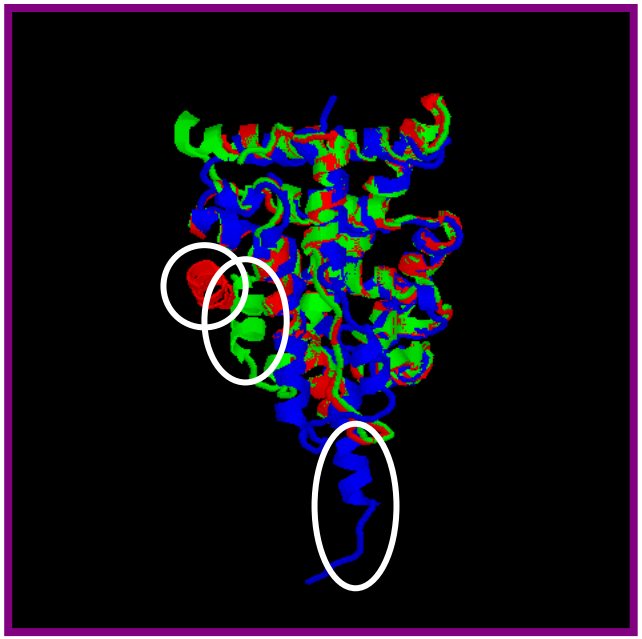
LBD helix 12 (H12)

•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--HHHHHHHHHHH-------SS---SS--HHHHHHHHHHHHHHHHHHHHHHHTTSTT

1errA.outSeq   FVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFD
1errA.outSS    GGGS-HHHHHHHHHHHHHHHHHHHHHHHTTSTTEEEEETTEEEEHHHHTTSTT-HHHHH

1errA.outSeq   MLLATSSRFRMMNLQGEEFVCLKSIILLNSGVY0EEKDHIHRVLDKITDTLIHLMAKAGL
1errA.outSS    HHHHHHHHHHHHT--HHHHHHHHHHHHHHSS----HHHHHHHHHHHHHHHHHHHHHHTT-

1errA.outSeq   TLQQQHQRLAQLLLILSHIRHMSNKGMEHLYSMGPLYDLLLEMLDAH
1errA.outSS    -HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH--HHHHHHHHHH-
```

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

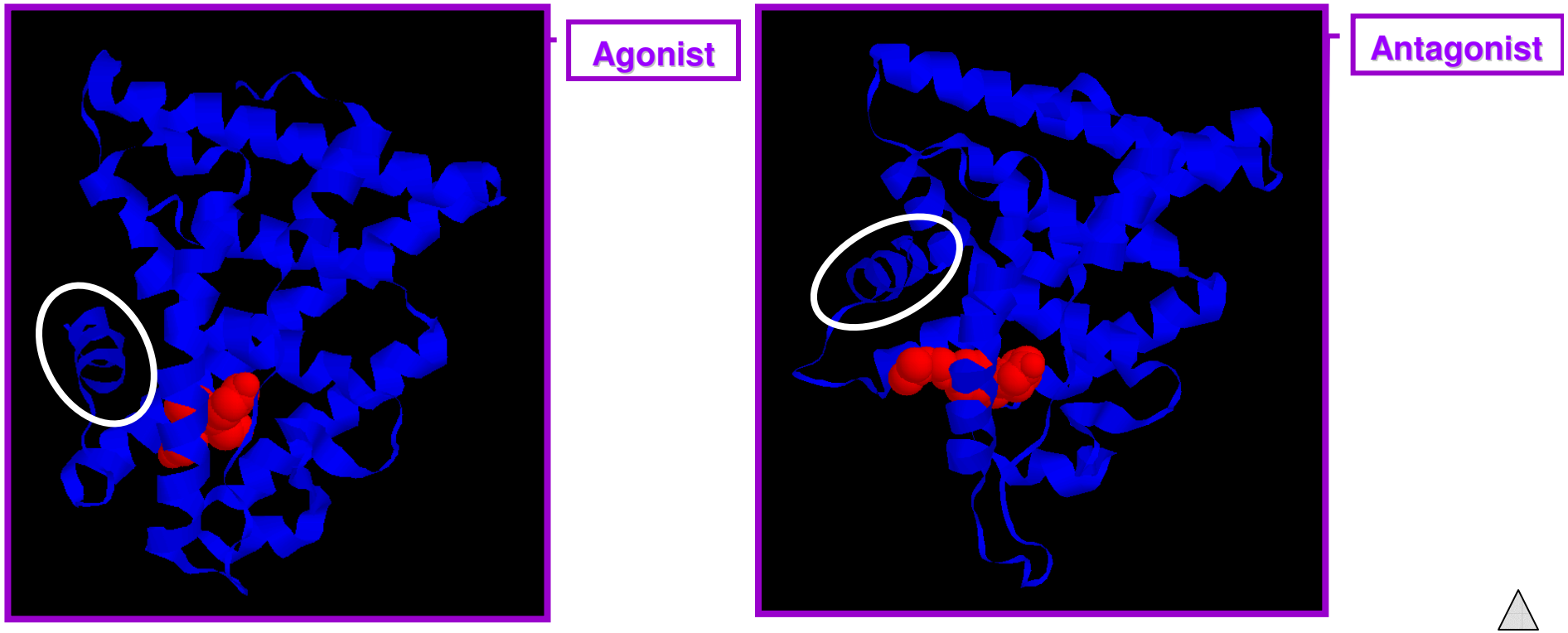
Alignment length Lp: 244

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.

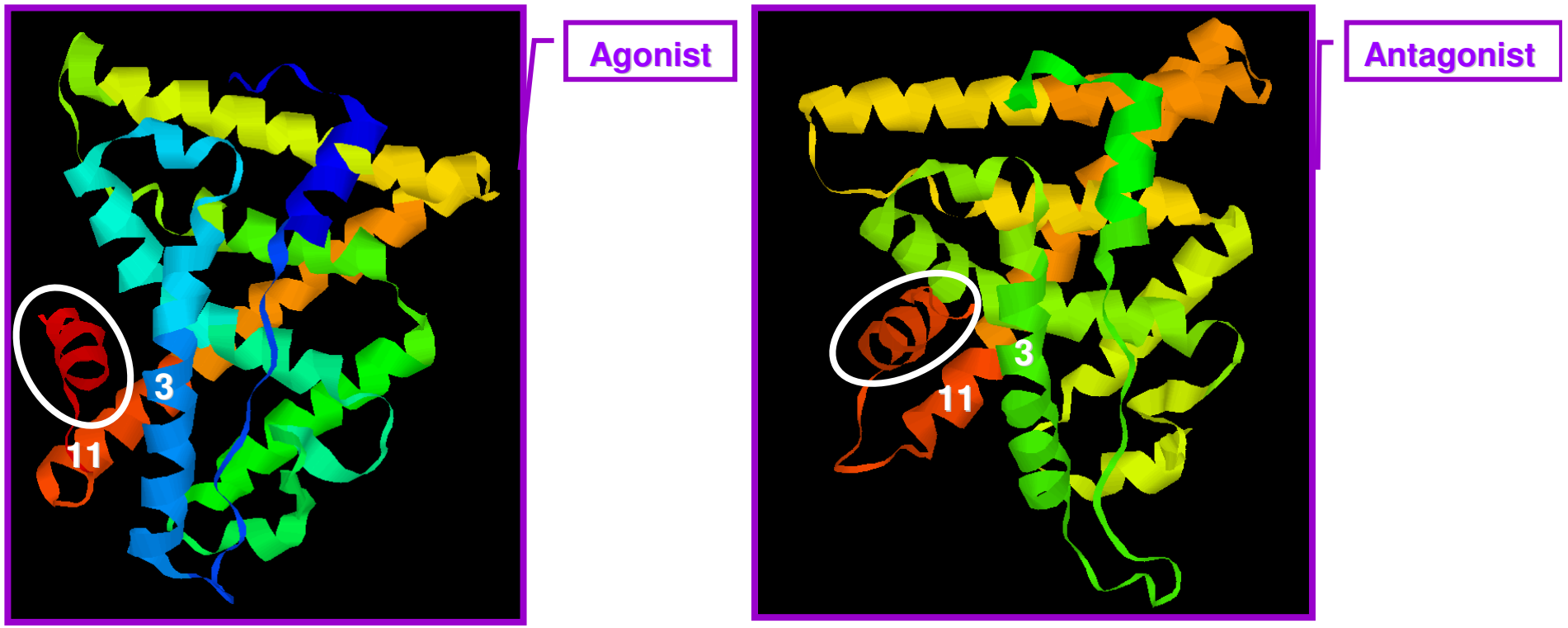
Agonist/antagonist-induced conformation

- 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.



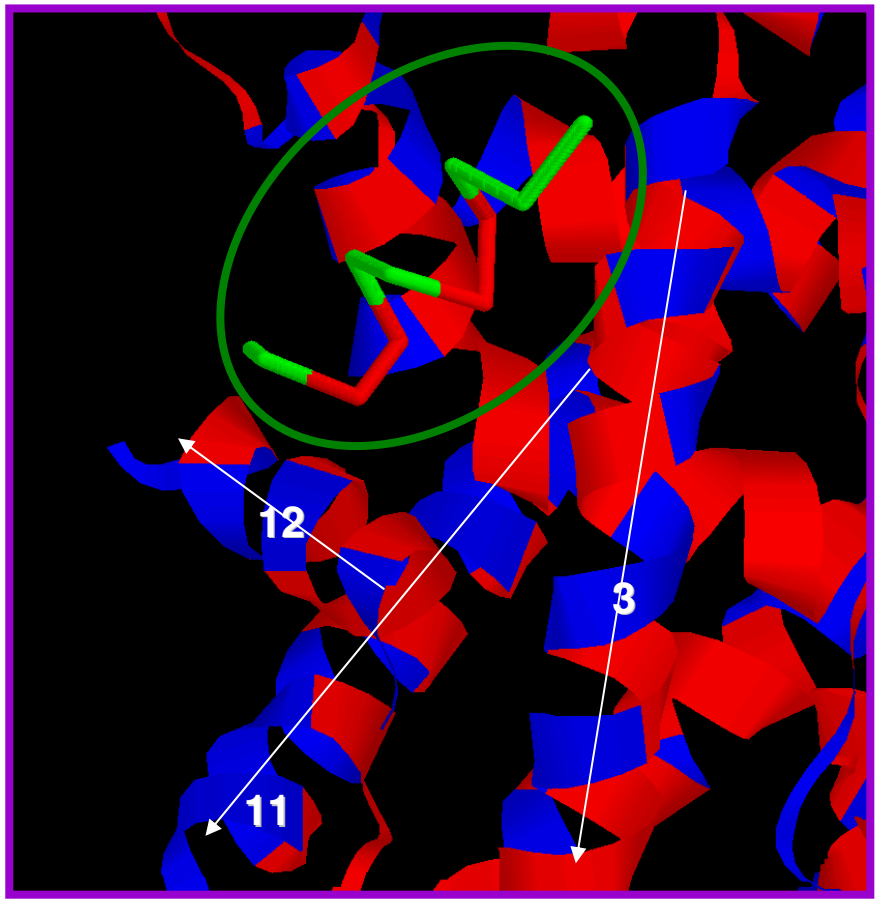
Agonist/antagonist-induced conformation

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



Agonist-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).



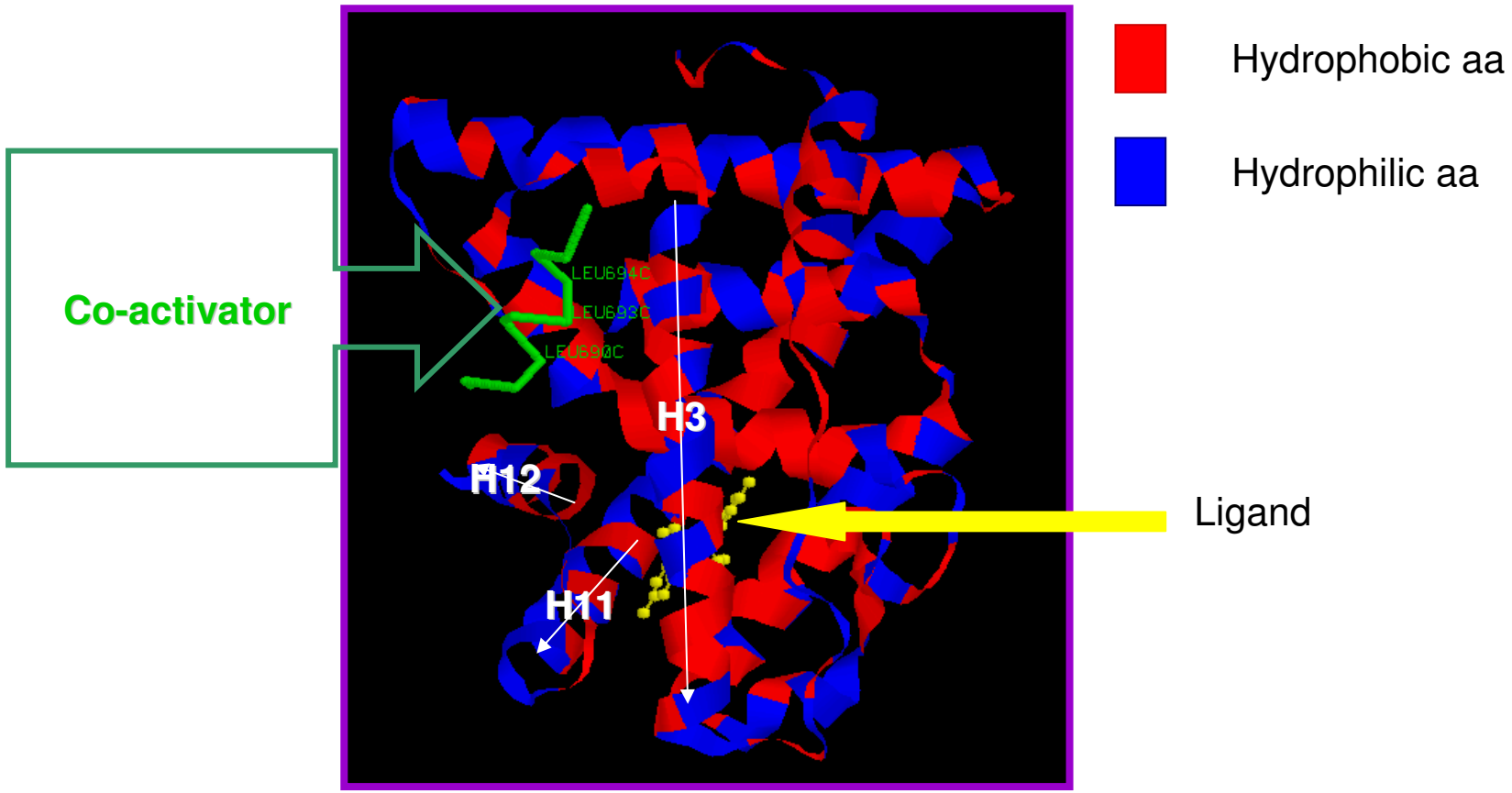
Hydrophobic aa

Hydrophilic aa

NCoA

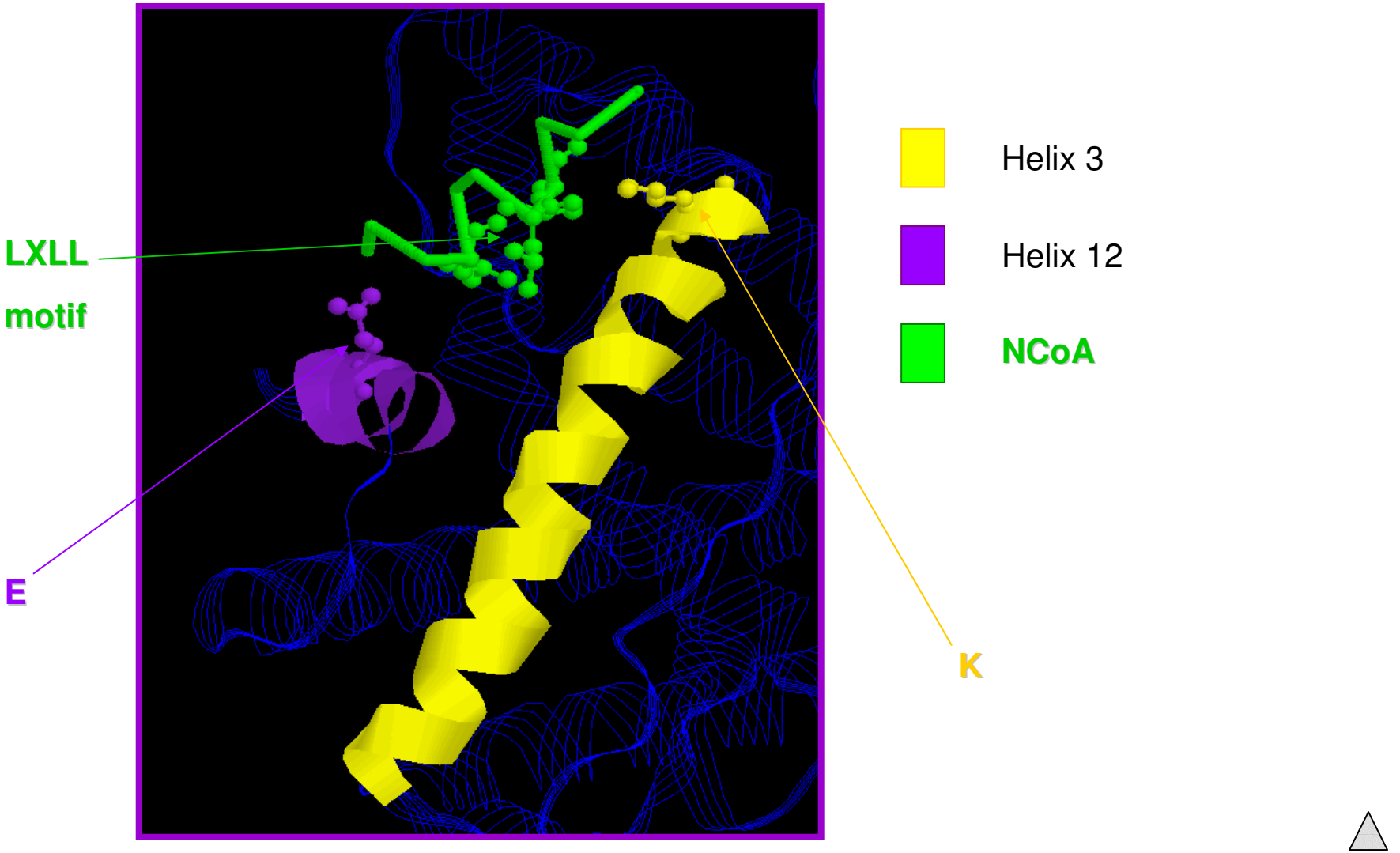
Interaction NR-coactivator

- 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.

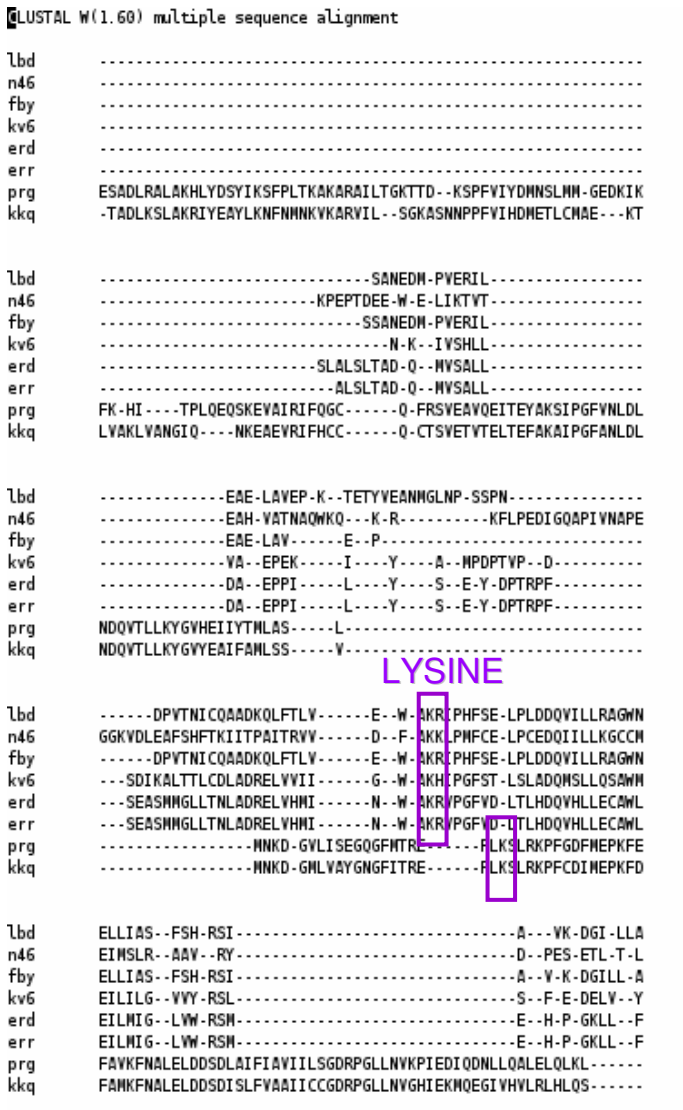


Interaction NR-coactivator

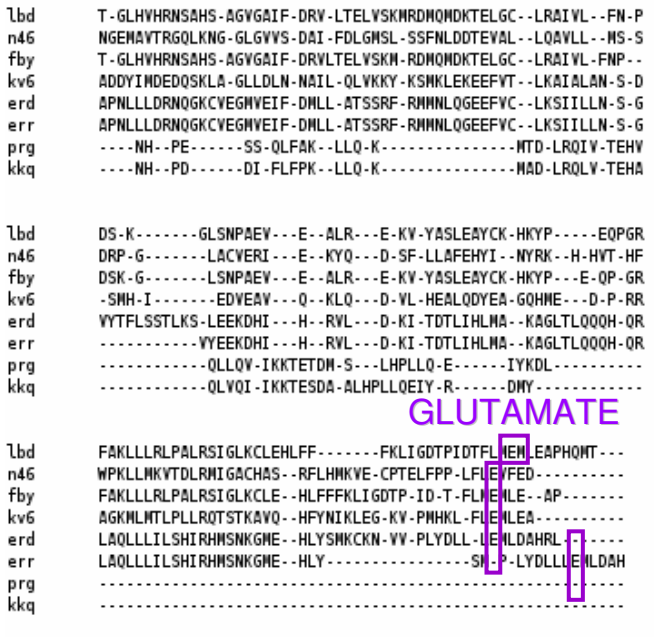
- A lysine residue at the C-terminus of H3 and a glutamate in H12 are hydrogen-bonded to the peptide bonds in the motif that stabilises the interaction: *Charge Clamp*



Conservation of residues

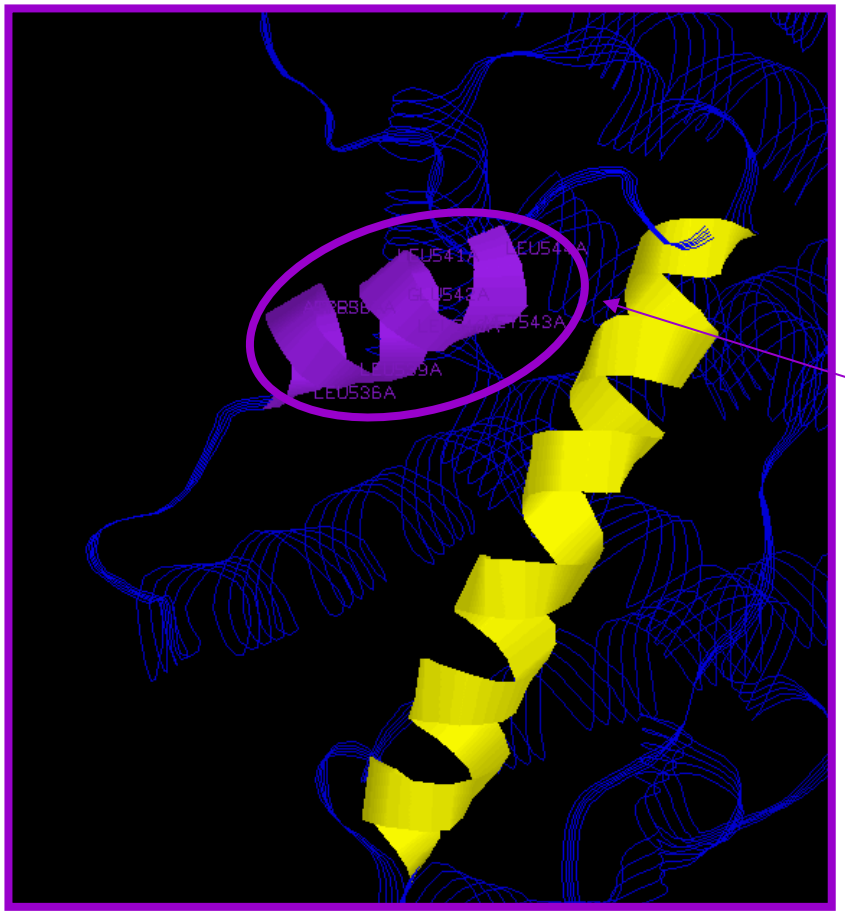


- Structural conservation for key residues to the interaction with co-activator.
- It can observe the conservation through the families.



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 ?