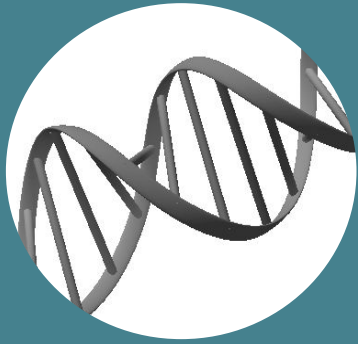
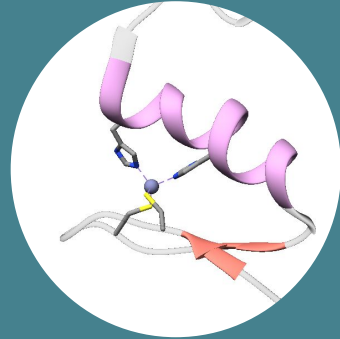
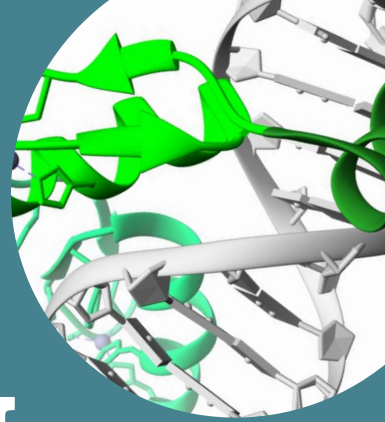




TRANSCRIPTION FACTORS



Maddalen Arizkorreta, Marta Figueras, Gal·la Garcia, Eva Martínez, Mar Prieto

INDEX

1

Introduction

2

Methodology

3

Structure analysis

3.1

c-Jun and c-Fos

3.2

Egr-1

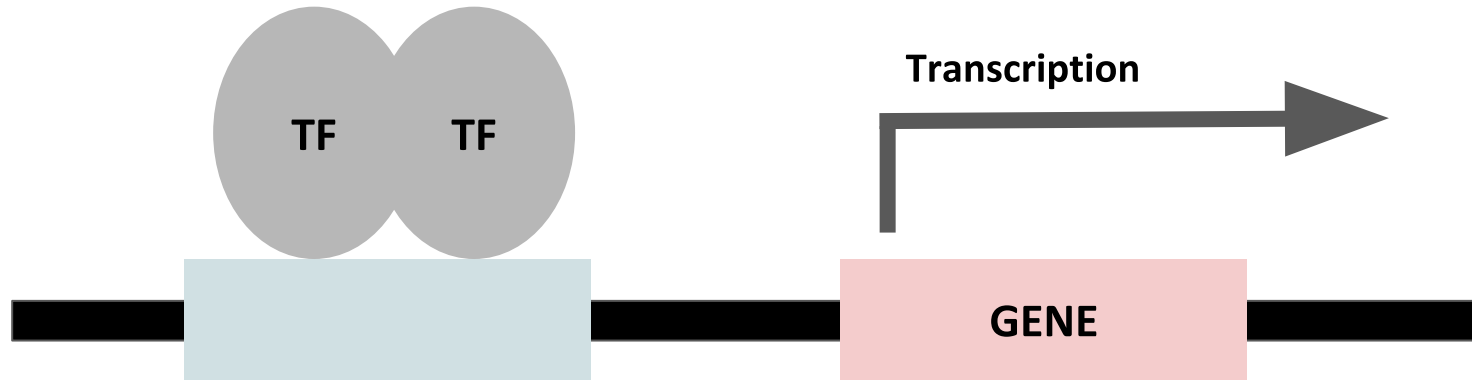
4

Conclusions



INTRODUCTION

INTRODUCTION



Transcription factors (TF) are some proteins that bind to specific DNA sequences and regulate gene expression controlling (promoting or suppressing) transcription of **DNA to mRNA**.

INTRODUCTION

Transcription factors are classified based in their sequence similarity and therefore the structure of their DNA binding domains.

FOUR DIFFERENT SUPERCLASSES

BD

BASIC DOMAINS

HTH

HELIX-TURN-HELIX

ZF

ZINC-COORDINATING DNA BINDING

BS

BETA-SCAFFOLD FACTORS WITH MINOR GROOVE CONTACTS

INTRODUCTION

BD

HTH

ZF

BS

BASIC DOMAINS

SITUATIONS:

1. Free in solution → excess of positive charges prevent from being structured
2. Interacting with DNA → α -helically folded

DOMAINS:

1. **DBD**: mediates sequence-specific DNA-binding
2. **DIMERIZATION DOMAIN**: Leucine zipper (ZIP) / helix-loop-helix (HLH) / helix-span-helix (HSH)

INTRODUCTION

BD

HTH

ZF

BS

HELIX-TURN-HELIX MOTIFS

Di-helical

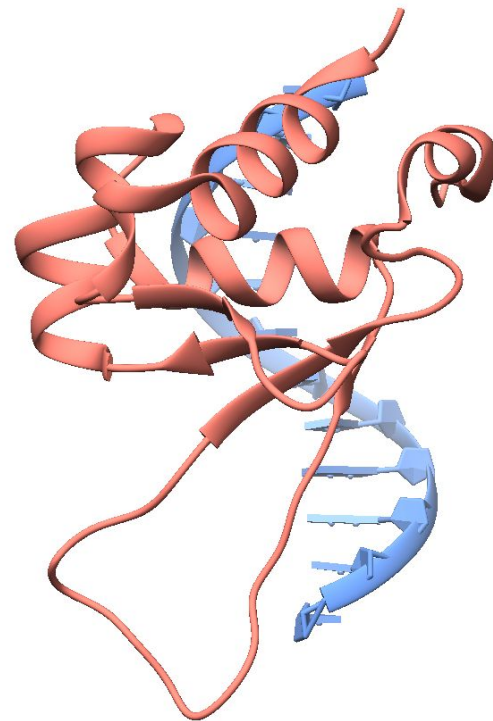
Tri-helical

Tetra-helical

Winged helix-turn-helix

Important roles in:

- DNA repair and replication
- RNA metabolism and protein-protein interactions in signaling



HSF2. PDBid: 5D8K

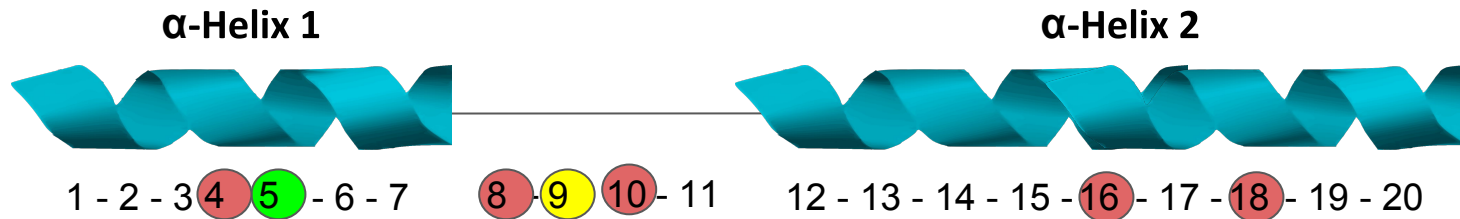
INTRODUCTION

BD

HTH

ZF

BS



 Hydrophobic

 Gly or Ala

 Gly

α -Helix 1 $\rightarrow$ stabilizes

α -Helix 2 $\rightarrow$ binds to DNA

INTRODUCTION

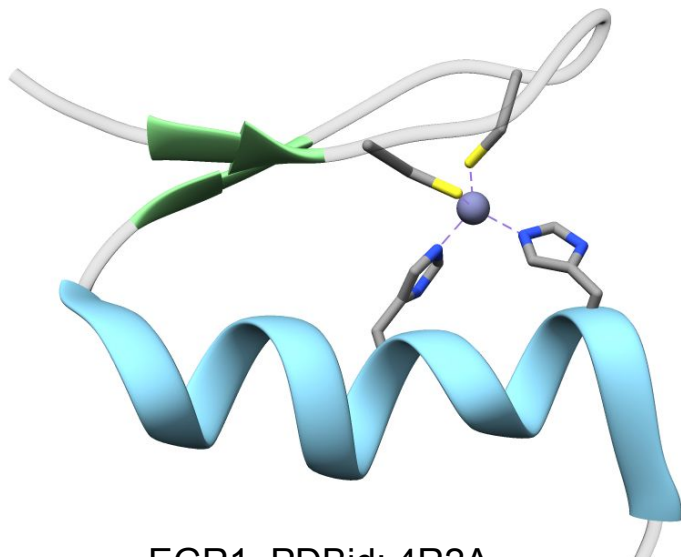
BD

HTH

ZF

BS

Zinc finger is a small protein structural motif that is characterized by the coordination of one or more **Zn ions** in order to stabilize the fold.



EGR1. PDBid: 4R2A

ZINC-COORDINATING DNA BINDING

Cys2-His2

Gag-Knuckle

Treble-clef

Zinc ribbon

Zn2/Cys6

Miscellaneous

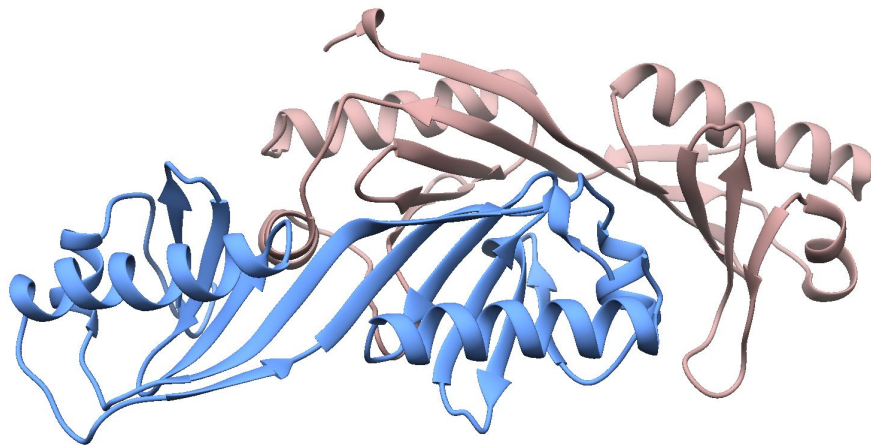
INTRODUCTION

BD

HTH

ZF

BS



Tata binding protein.
PDBid: 1PCZ

BETA-SCAFFOLD FACTORS WITH MINOR GROOVE CONTACTS

RHR	STAT
p53	MADS box
β -barrel α -helix TF	HMG-box
Heteromeric CCAAT factors	Grainyhead
Cold-shock domain factors	Runt
TATA binding proteins	



METHODOLOGY

METHODOLOGY

EVOLUTION

FINDING PROTEIN SEQUENCES



DIFFERENT SPECIES

MULTIPLE SEQUENCE
ALIGNMENT

SAME FAMILY
ALL HUMAN

MULTIPLE SEQUENCE
ALIGNMENT

STRUCTURAL ALIGNMENT

STRUCTURAL ANALYSIS



General and secondary
structure

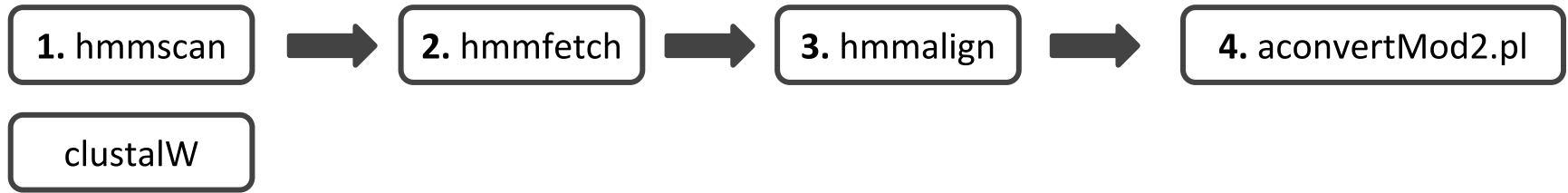
DNA-protein interactions

Other conformations

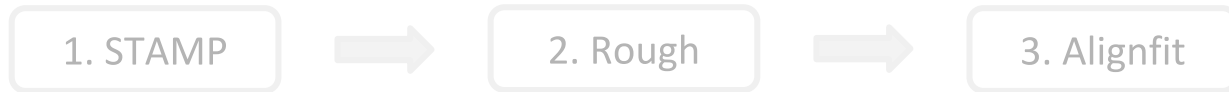
CONCLUSIONS

METHODOLOGY

MULTIPLE SEQUENCE ALIGNMENT



STRUCTURAL ALIGNMENT



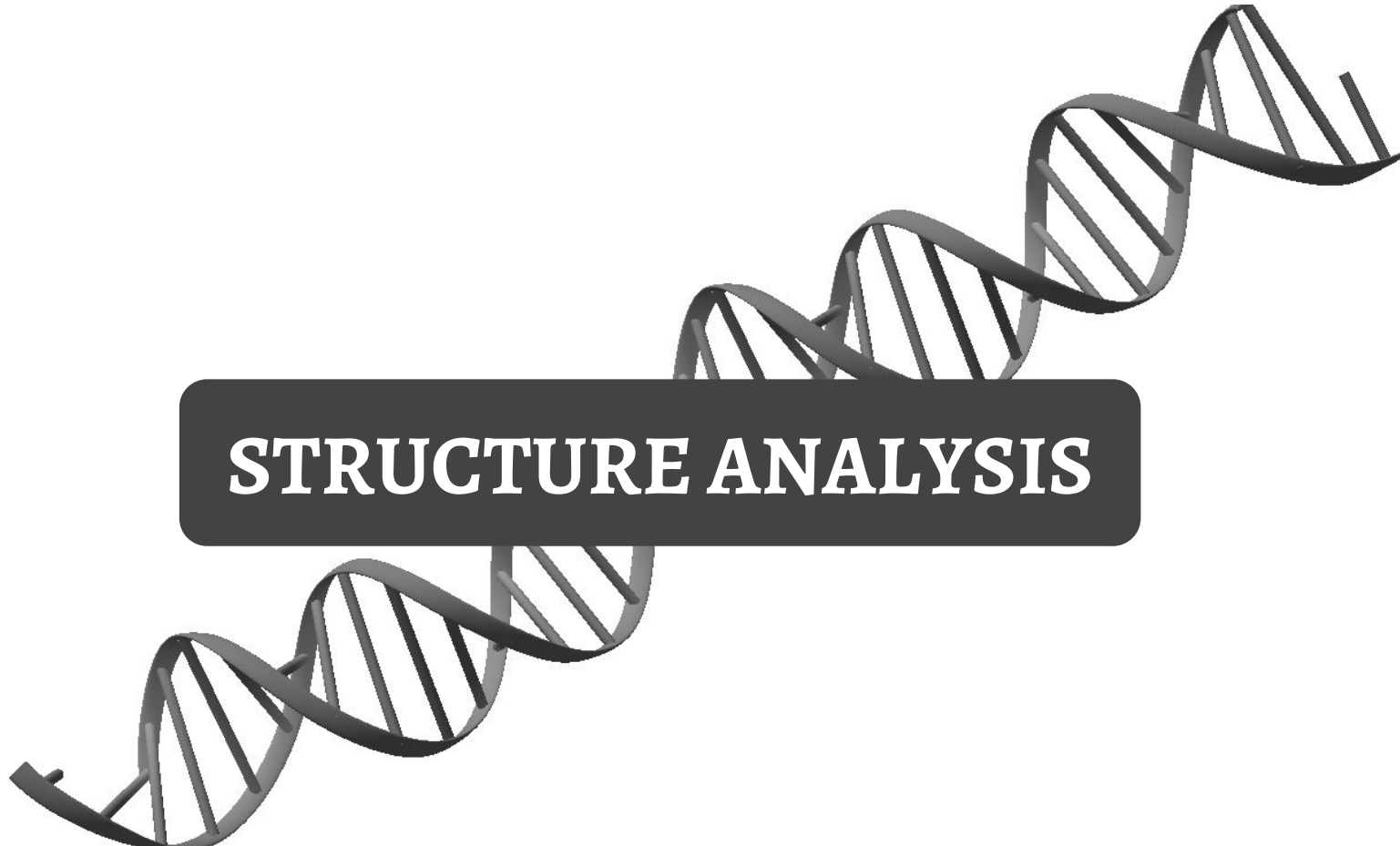
METHODOLOGY

MULTIPLE SEQUENCE ALIGNMENT



STRUCTURAL ALIGNMENT



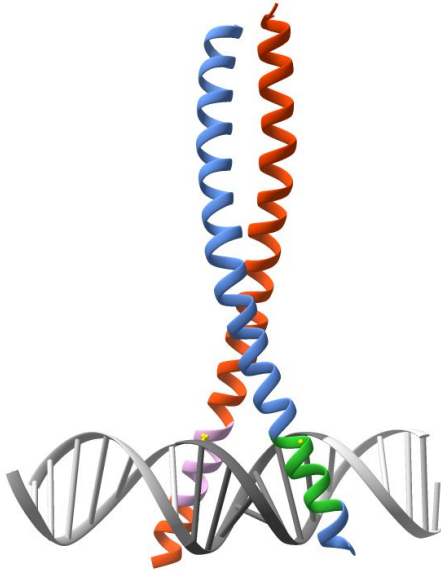


STRUCTURE ANALYSIS

Proteins

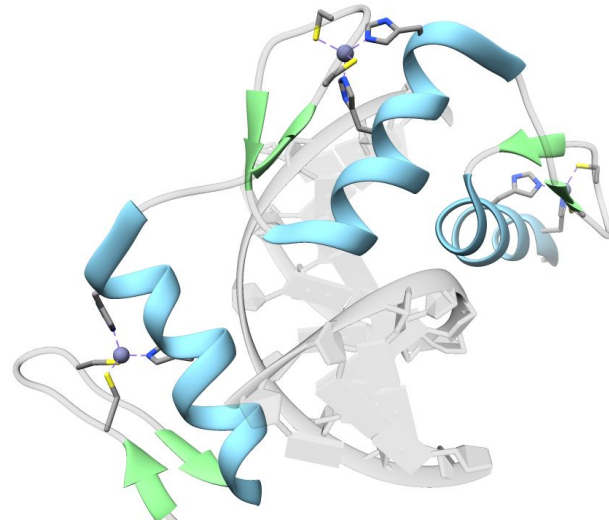
LEUCINE-ZIPPER FACTORS (bZIP)

c-FOS c-JUN
(Homo Sapiens)



ZINC-COORDINATING DNA-BINDING DOMAINS

Early Growth Factor 1 (EGR1)
(Homo Sapiens)

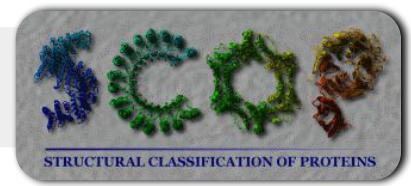




c-FOS and c-JUN

c-FOS c-JUN

CLASSIFICATION



ROOT	Scop
CLASS	Coiled-coil proteins
FOLD	Parallel coiled-coil
SUPERFAMILY	Leucine zipper domain (bZIP)
FAMILY	Leucine zipper domain (bZIP)

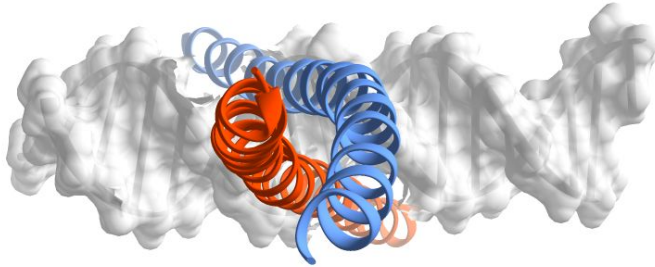
c-FOS c-JUN b-ZIP FAMILY



Essential roles from early development to tumorigenesis

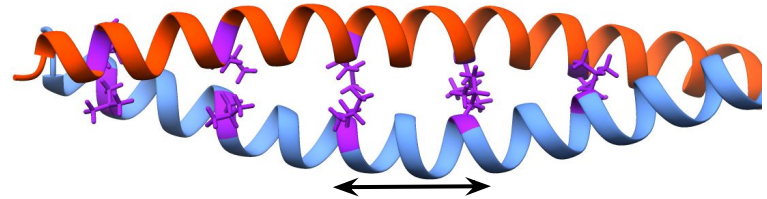
40-80 aa long with two motifs:

Basic region → allows DNA binding



Tightly bound to the major groove of the DNA, like a pair of forceps

Leucine zipper → allows dimerization and DNA binding



Every two turns (7 residues)
Interdigitate with Leucines from the other helix

c-FOS c-JUN

BIOLOGICAL FUNCTION



AP-1 TFs

Are **involved in**:

1. Cell proliferation
2. Survival
3. Apoptosis
4. Transformation
5. Oncogenesis

Belong to **Jun**, **Fos**
and **ATF** subfamilies

→ They form dimers

CRE

TRE

5'-TGAC/GTCA-3'

Heterogeneous AP-1 complexes fulfill distinct roles in cells of different lineage

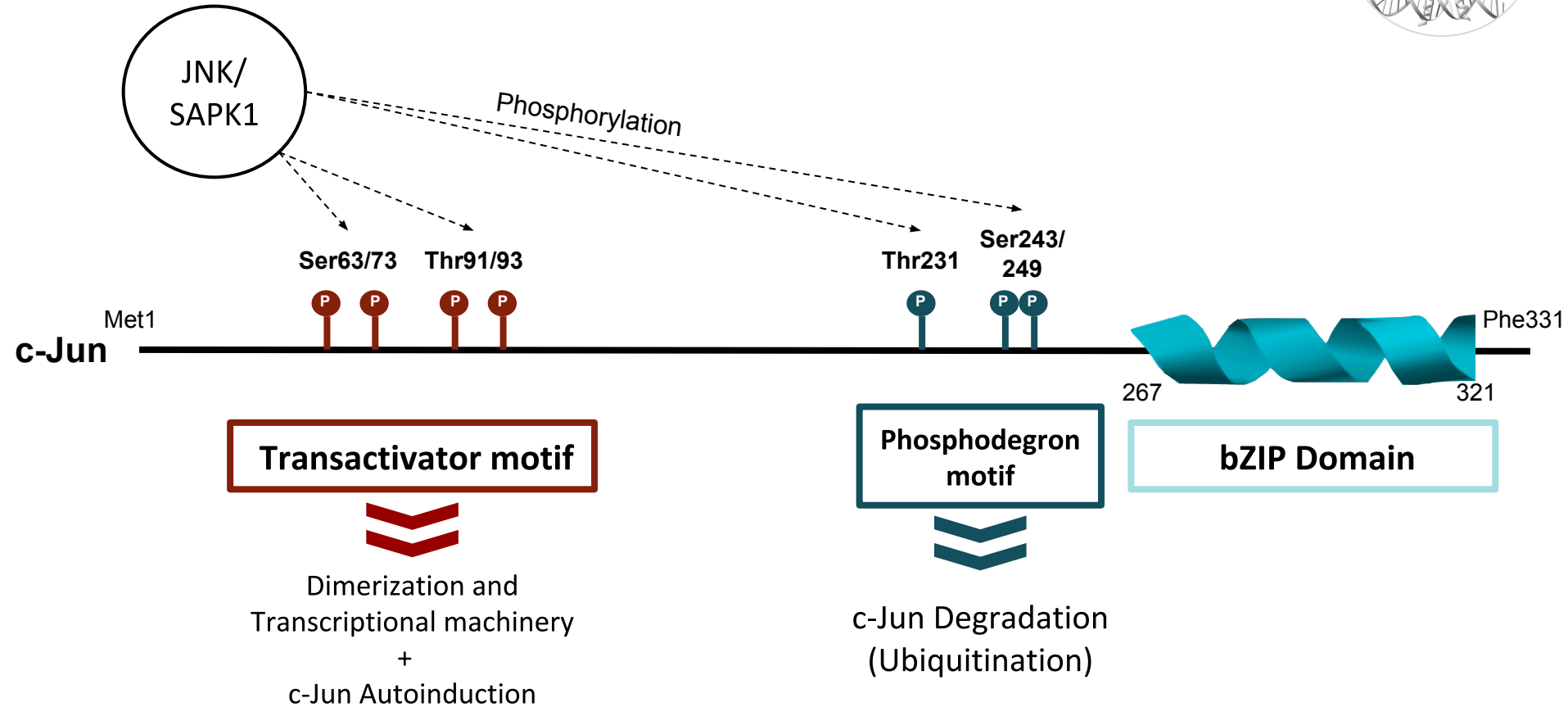
c-FOS/ c-JUN

Are regulated by MAPK pathways → **c-Jun needs to be phosphorylated** to bind to DNA and form heterodimers with c-Fos

Induce **cell proliferation** since they are **proto-oncogenes**

c-FOS c-JUN

ACTIVATION

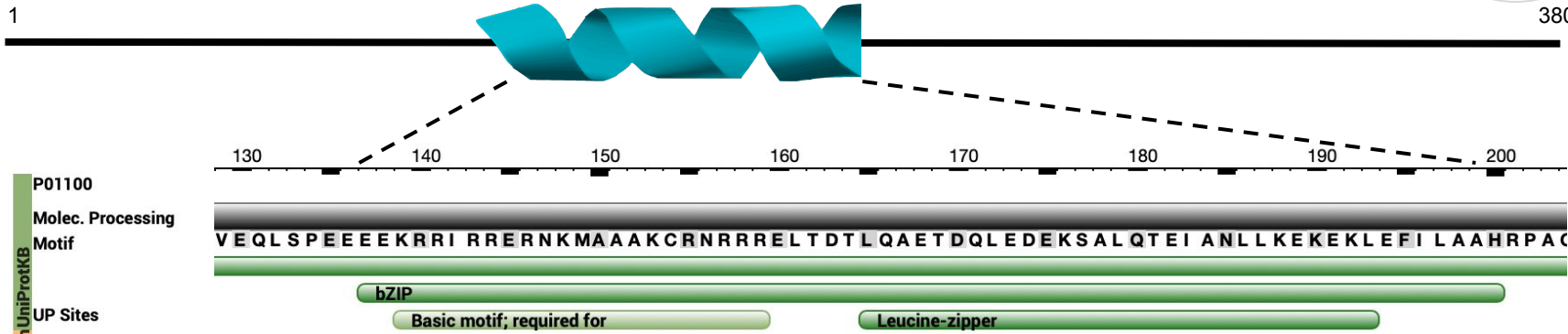


c-FOS c-JUN

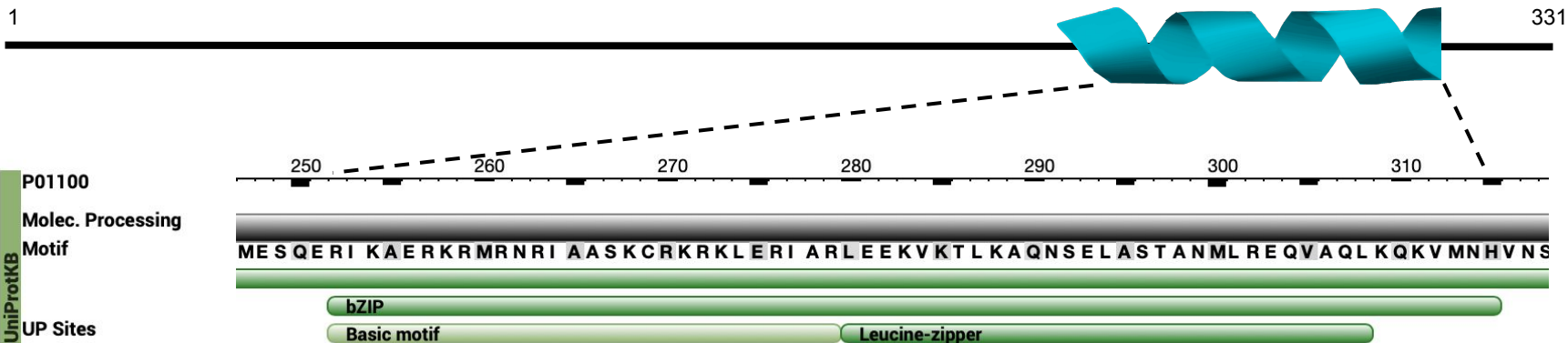
GENERAL STRUCTURE



c-FOS



c-JUN

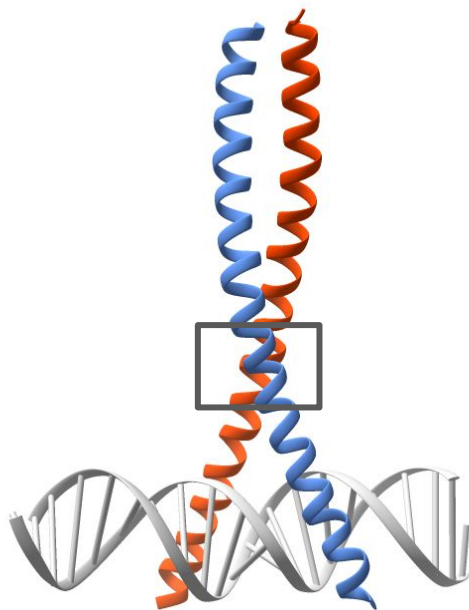


c-FOS c-JUN

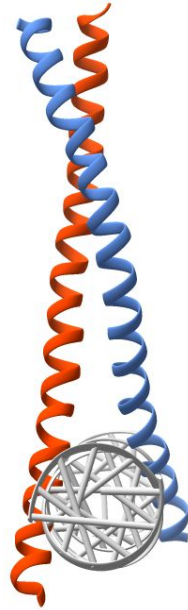
GENERAL STRUCTURE



COMPLEX I



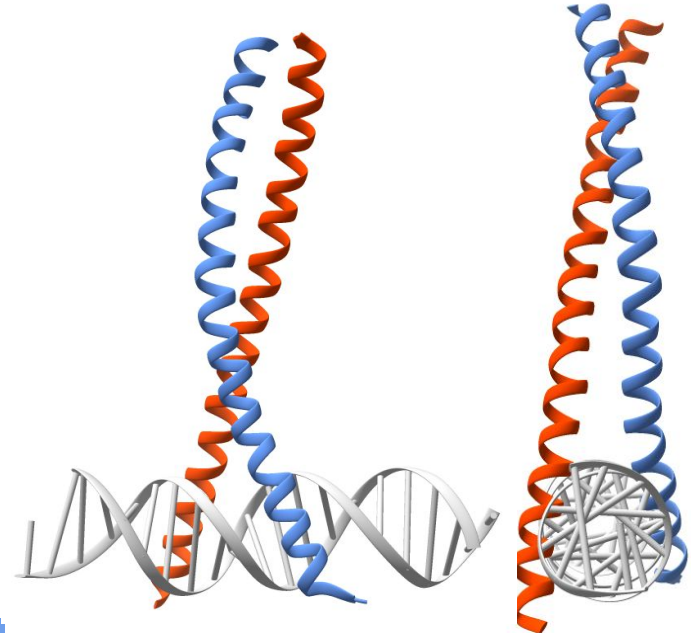
5'- TGACTCA -3'



c-FOS

c-JUN

COMPLEX II



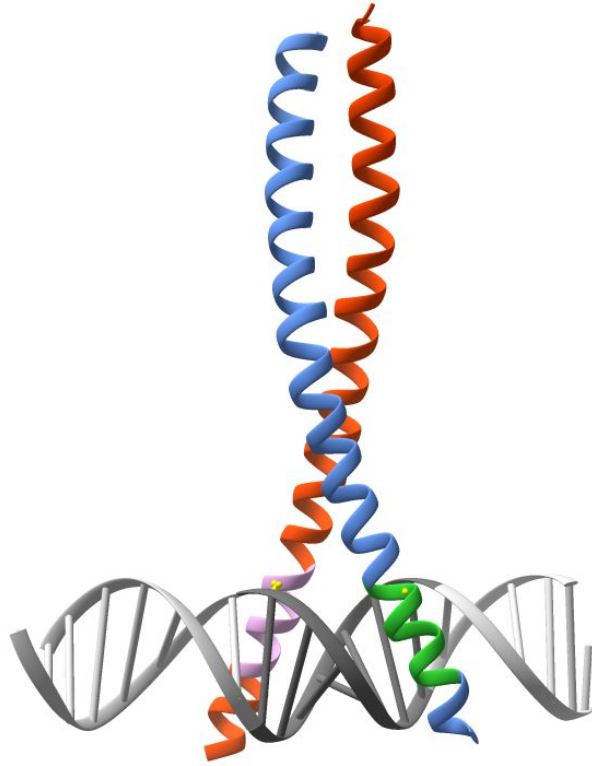
5'- TGAGTCA -3'

c-FOS c-JUN

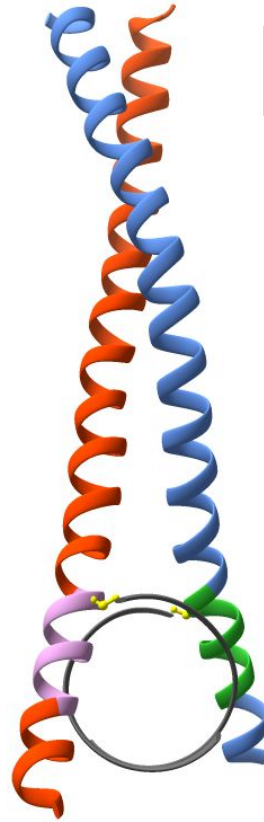
GENERAL STRUCTURE



DNA-protein interactions



5'- **T**G**A**C**T**C**A** -3'
3'- A**C****T**G**A**G**T** -5'



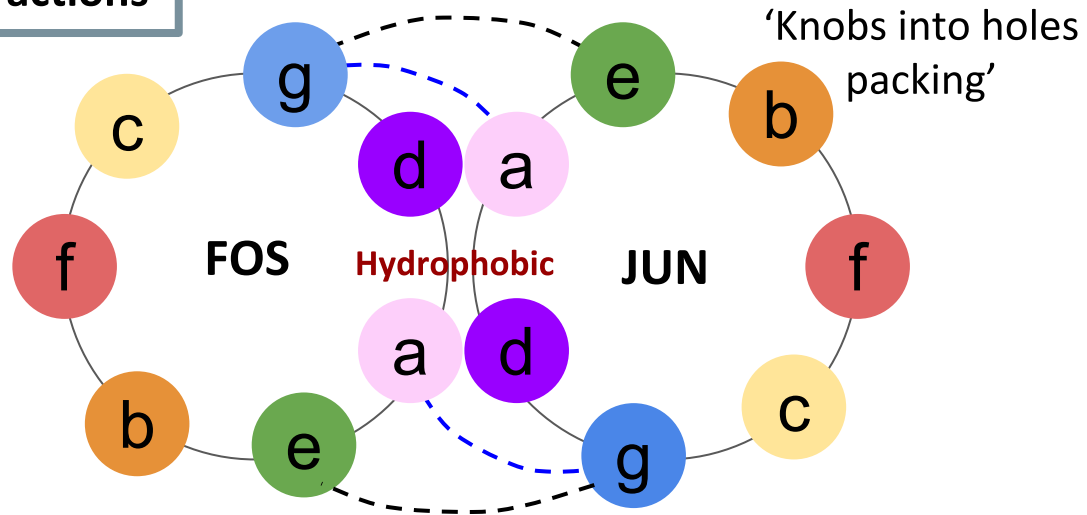
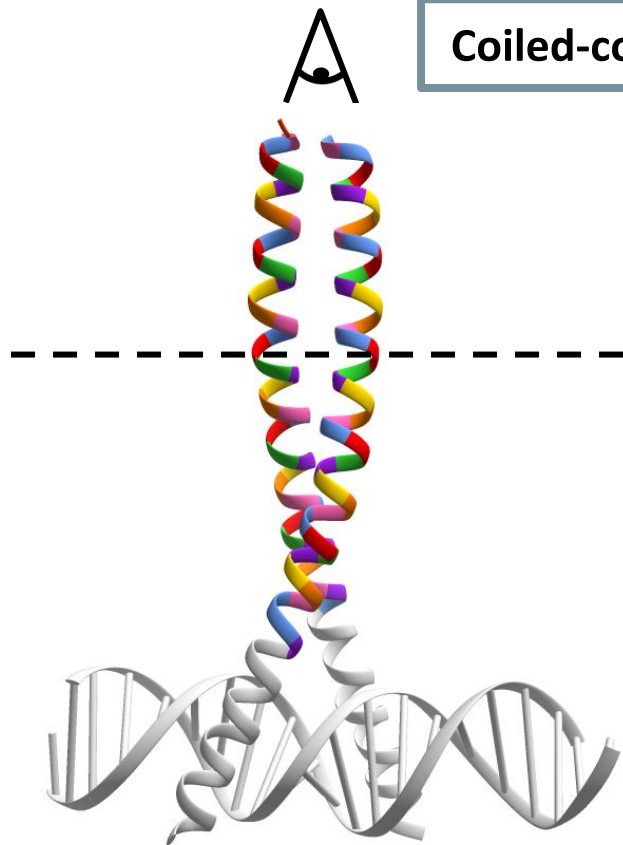
-  FOS
-  JUN
-  DNA-binding Motif-F
-  DNA-binding Motif-J
-  Conserved Serine/Cysteine
-  AP-1 site

c-FOS c-JUN

GENERAL STRUCTURE



Coiled-coil interactions

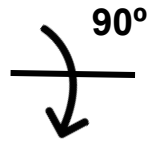
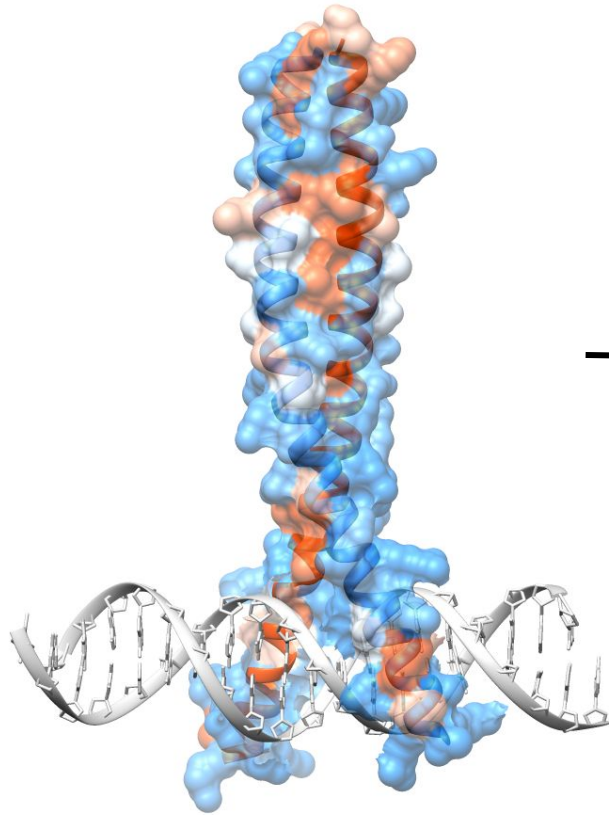


	g	a	b	c	d	e	f	g	a	b	c	d	e	f	g	a
FOS	L	T	D	T	L	Q	A	E	T	D	O	L	E	D	E	K
JUN	R	I	A	R	L	E	E	K	V	K	T	L	K	A	Q	N

Sequence alignment showing FOS and JUN amino acid sequences. Red boxes highlight specific residues (L, T, D, T, L, Q, A, E, T, D, O, L, E, D, E, K, R, I, A, R, L, E, E, K, V, K, T, L, K, A, Q, N, S, E, L, A, S, T, A, N, M, L, R, E, Q, V, A, Q, L, K, Q, K, V). A blue 'X' marks a mismatch between FOS (E) and JUN (N) at position 14. A dashed line connects the 'g' beads (L and R) and the 'd' beads (D and T) between the two sequences. A black bar at the bottom indicates a region of high sequence similarity.

c-FOS c-JUN

IMPORTANT RESIDUES AND INTERACTIONS

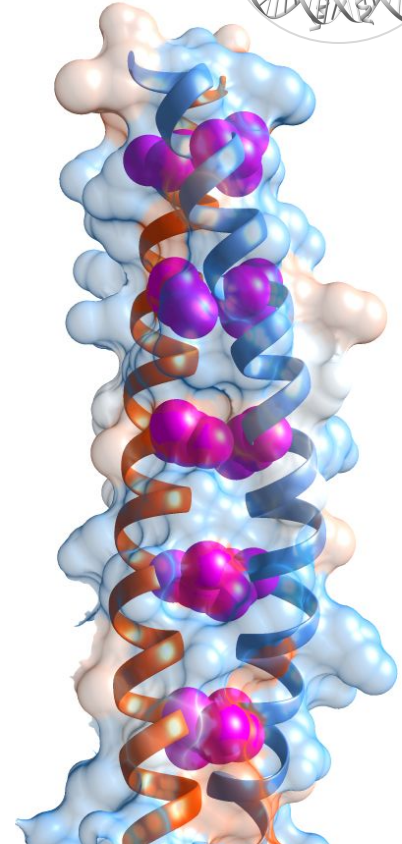
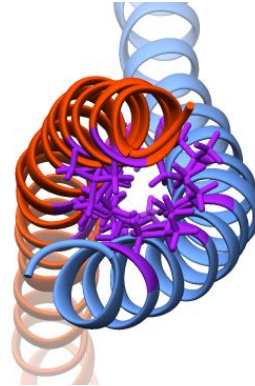


Hydrophilic



Hydrophobic

HYDROPHOBIC CORE



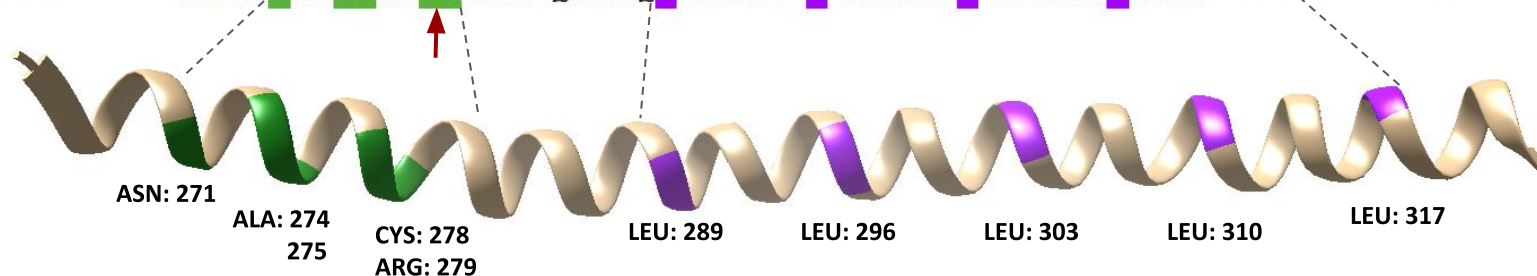
c-FOS c-JUN

SECONDARY STRUCTURE - EVOLUTION AMONG DIFFERENT SPECIES (MSA)

CLUSTAL W(1.60) multiple sequence alignment



```
FOS_MOUSE  ---KRRIRRERNKMAAAKCRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLE---fil-
FOS_HUMAN  ---KRRIRRERNKMAAAKCRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLE---fil-
FOS_DANRE  -----RRERSKMAAAKCRNRRRELTDTLQAETDQLEDEKSALQNDIANLLKEKERL-----efil
FOS_BOVIN  ---KRRIRRERNKMAAAKCRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLE---fil-
FOS_RAT    ---KRRIRRERNKMAAAKCRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLE---fil-
FOS_PIG    ---KRRIRRERNKMAAAKCRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLE---fil-
FOS_CHICK  ---KRRIRRERNKMAAAKCRNRRRELTDTLQAETDQLEEEKSALQAEIANLLKEKEKLE---fil-
FOS_CAT    ---KRRIRRERNKMAAAKCRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLE---fil-
FOS_SHEEP  -----ERNKMAAAKCRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLE---fil-
GCN4_YEAST -----KRARNTAAARRSRARKLQRMKQLEDKVEELL SKNYHLENEVARLKKLV-----
```



CONSENSUS MOTIF

LEUCINE ZIPPER DOMAIN

c-FOS c-JUN

SECONDARY STRUCTURE - SAME FAMILY STRUCTURES (STAMP)



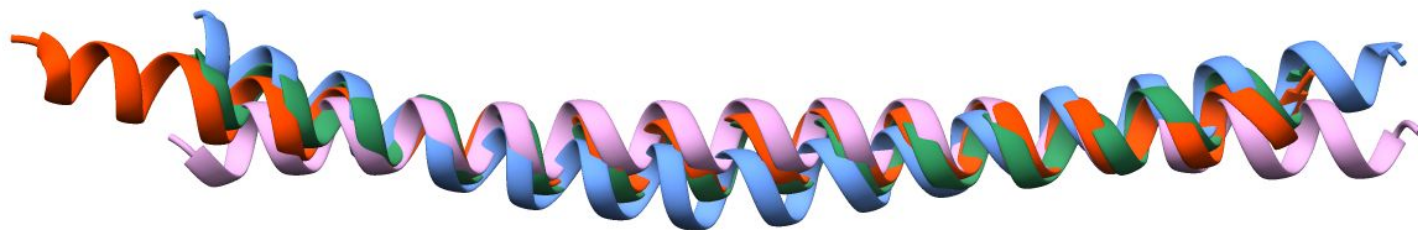
Sequence alignment

CLUSTAL W(1.60) multiple sequence alignment

```
1CI6A      -----EQNKTAATRYRQKKRAEQEALTGECKELEKKNEALKERADSLAKEIQYLKDL--ieev
1fosE      ----KRRIRRERNKMAAAKSRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLE----fila
1fosF      -----RKRMRNRRIAASKSRKRKLERIARLEEKVKTLKAQNSELASTANMLREQVAQLKQKVM----
5zkoA      ----KREVRLMKNREAARECRRKKKEYVKCLENRVAVLENQNKTLIEELKALKDLY-----c---
```

Structural alignment

```
1ci6      EQNKTAATRYRQKKRAEQEALTGECKELEKKNEALKERADSLAKEIQYLKDLIEEV-----
1fosF      RKRMRNRRIAASKSRKRKLERIARLEEKVKTLKAQNSELASTANMLREQVAQLKQKVM-----
1fosE      ----KRRIRRERNKMAAAKSRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLEFILA
5zko      ----KREVRLMKNREAARECRRKKKEYVKCLENRVAVLENQNKTLIEELKALKDLYC-----
```



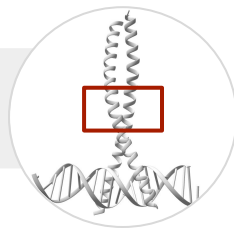
1ci6 → Atf-4
1fosF → c-JUN
1fosE → c-FOS
5zko → CREB1

Score 8.50

RMSD 1.90

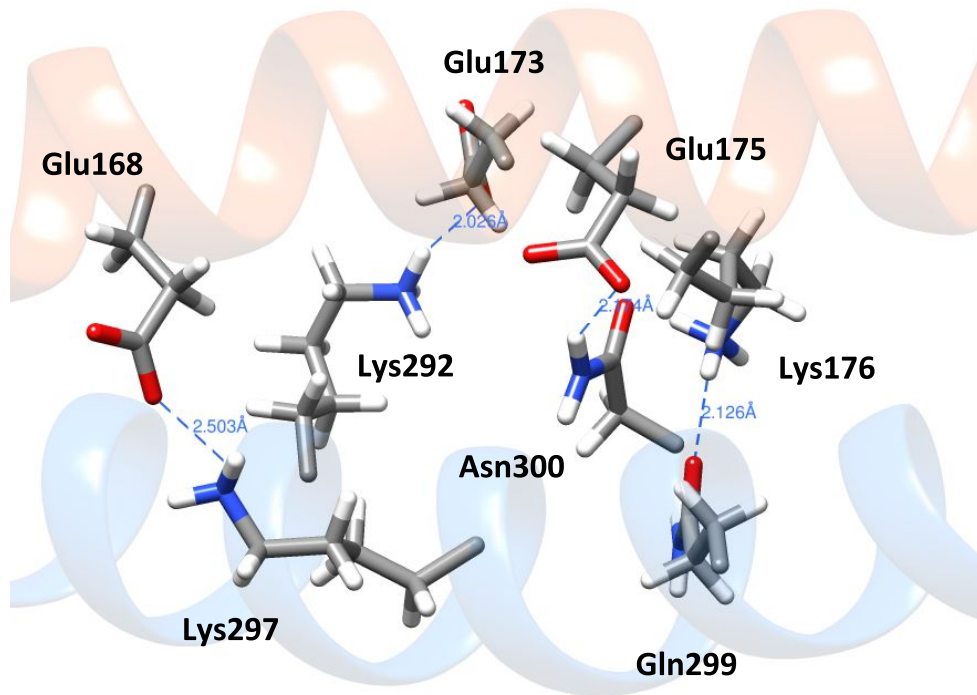
c-FOS c-JUN

IMPORTANT RESIDUES AND INTERACTIONS



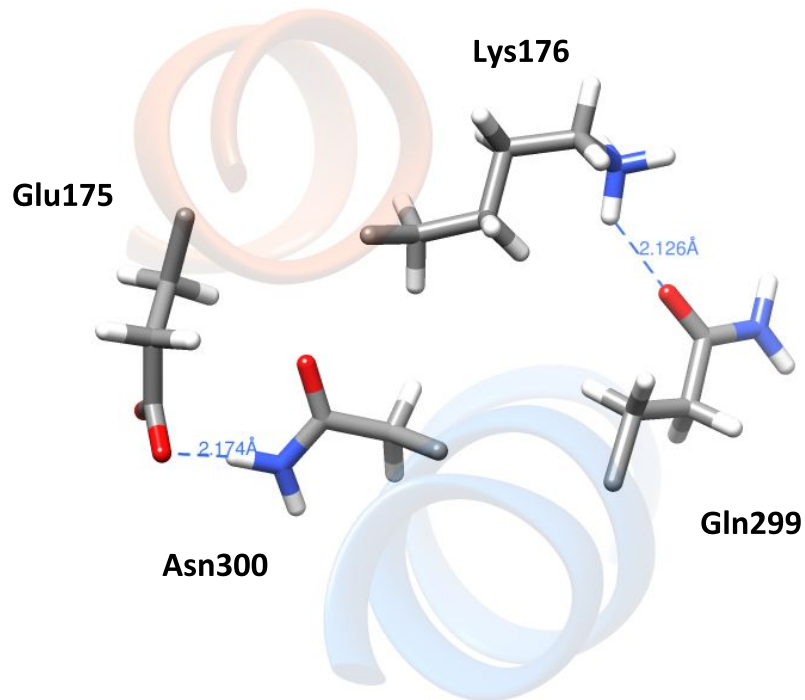
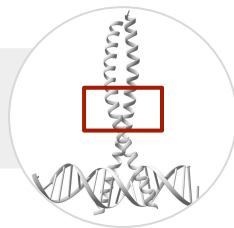
Coiled-coil interactions

Hydrogen bonds	
Glu175 - Asn300	2.174 Å
Lys176 - Gln299	2.126 Å
Salt bridges	
Glu168 - Lys297	2.503 Å
Glu173 - Lys292	2.026 Å

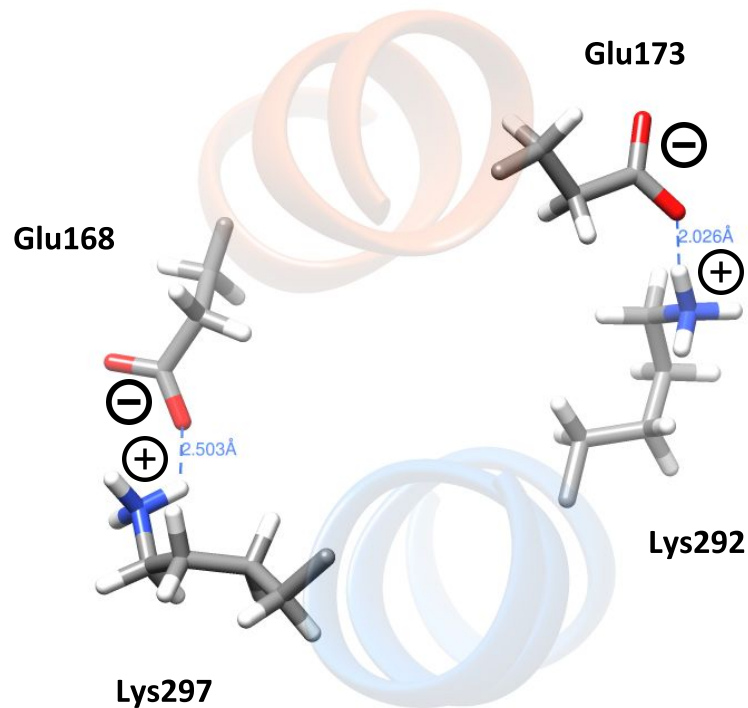


c-FOS c-JUN

IMPORTANT RESIDUES AND INTERACTIONS



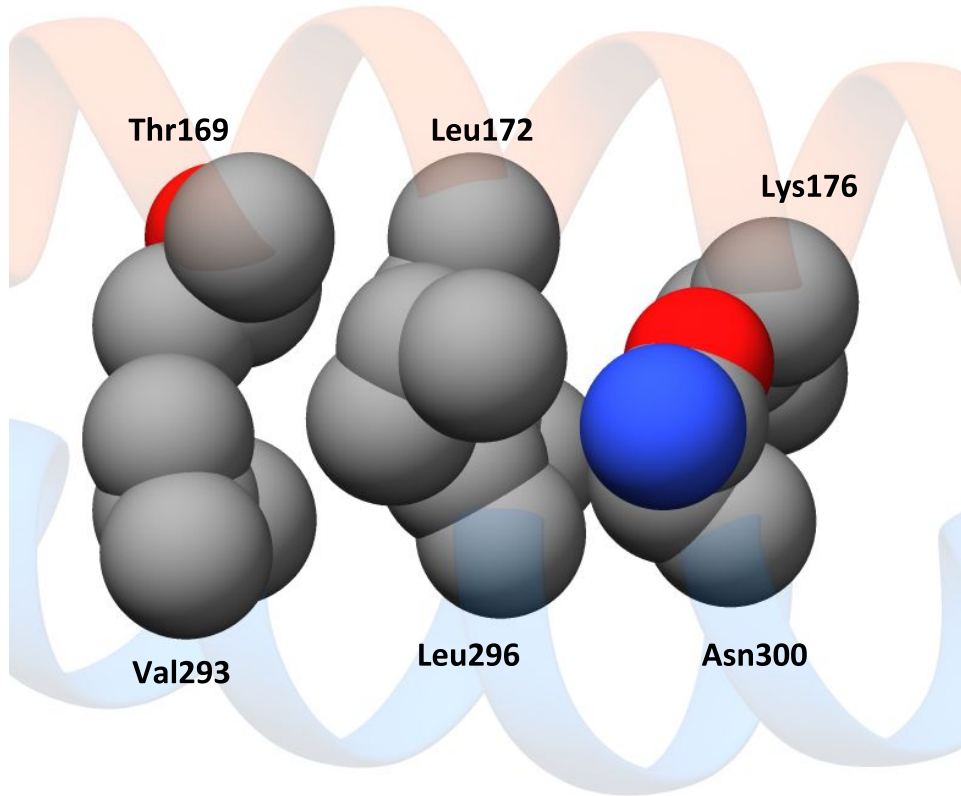
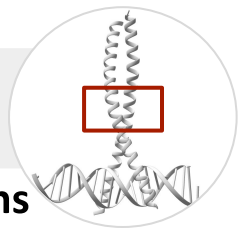
Hydrogen bonds



Salt bridges

c-FOS c-JUN

IMPORTANT RESIDUES AND INTERACTIONS



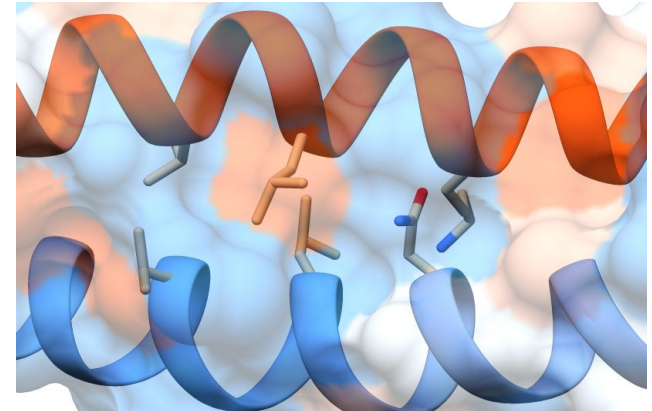
Coiled-coil interactions

Hydrophobic interactions

Thr169 - Val293

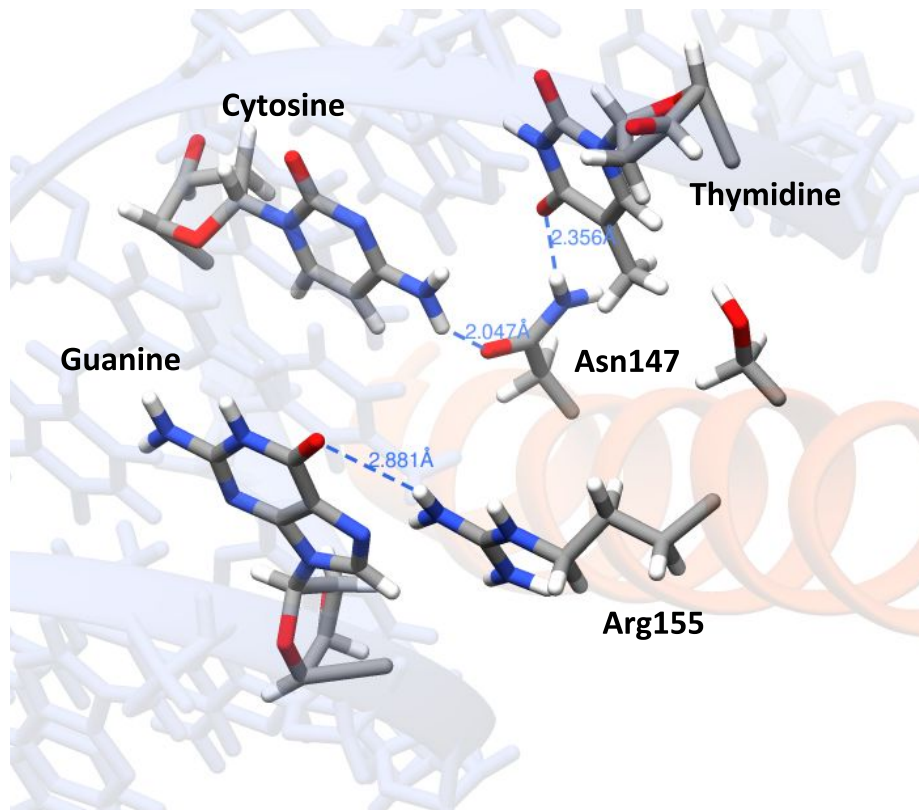
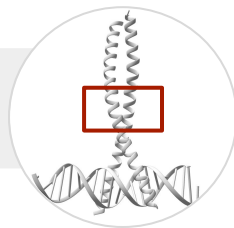
Leu172 - Leu296

Lys176 - Asn300



c-FOS c-JUN

IMPORTANT RESIDUES AND INTERACTIONS

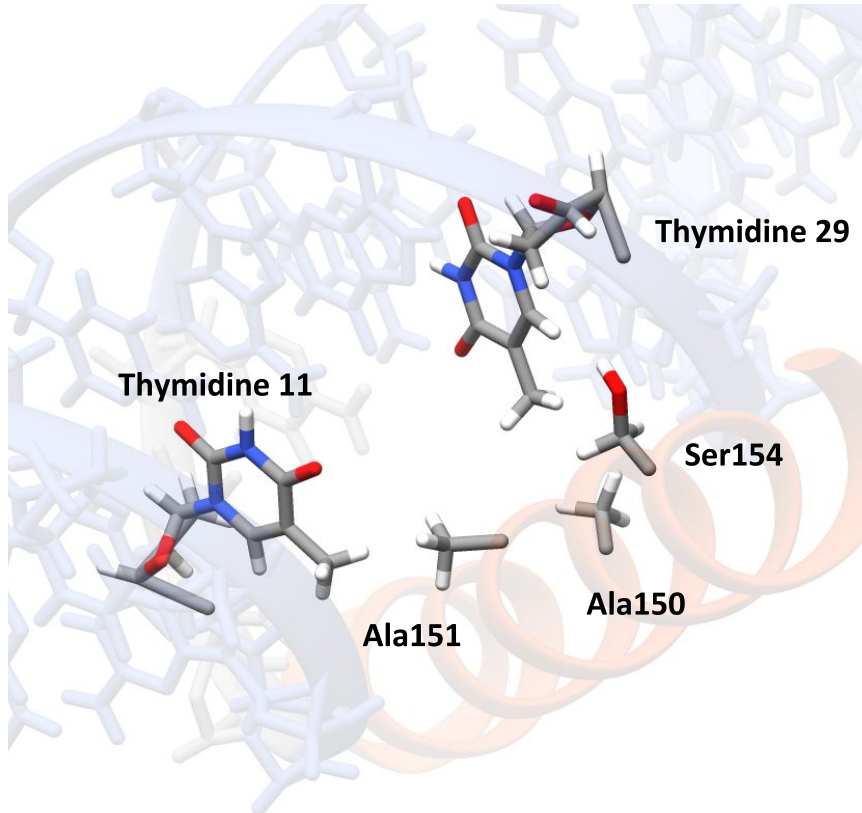
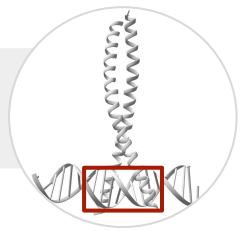


c-Fos (complex I) coiled-coil interactions

Hydrogen bonds	
Arg155 - Gua (A)	2.881 Å
Asn147 - Cyt (A)	2.047 Å
Asn147 - Thy (B)	2.356 Å

c-FOS c-JUN

IMPORTANT RESIDUES AND INTERACTIONS

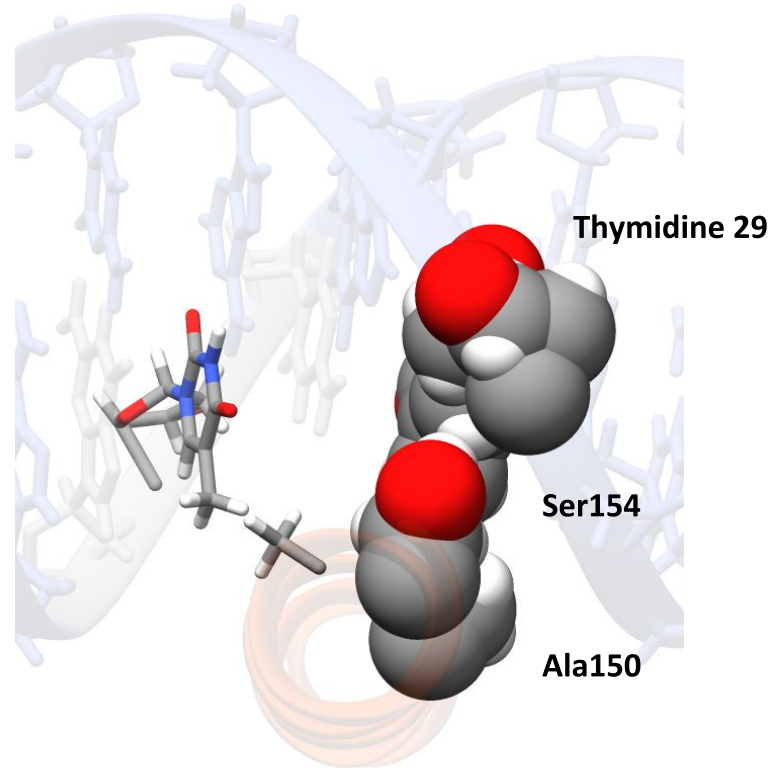
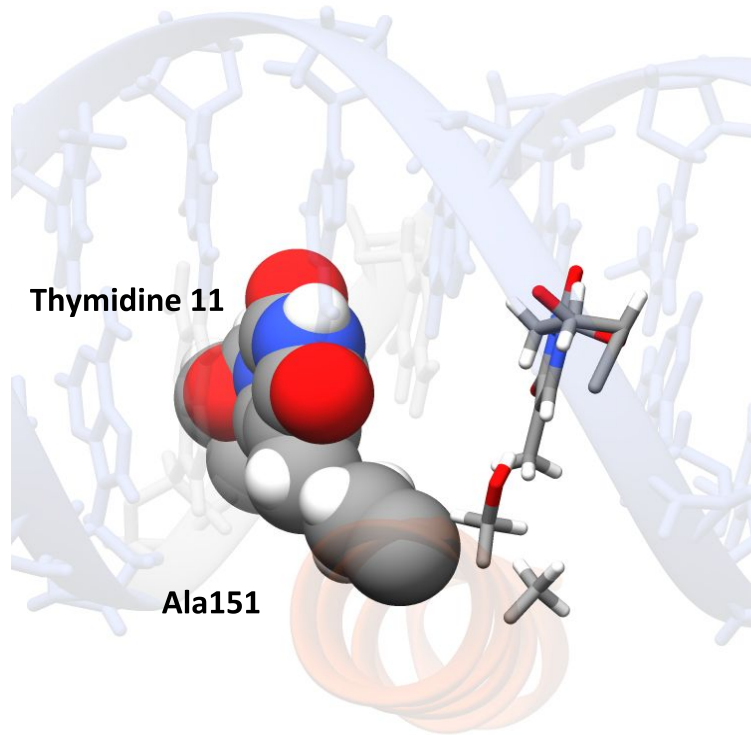
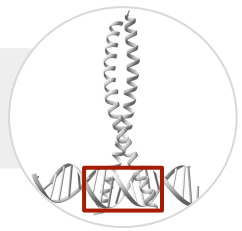


c-Fos (complex I) DNA interactions

Van der Waals contact
Ala150 - Thy29 (B)
Ser154 - Thy29 (B)
Ala151 - Thy11 (A)

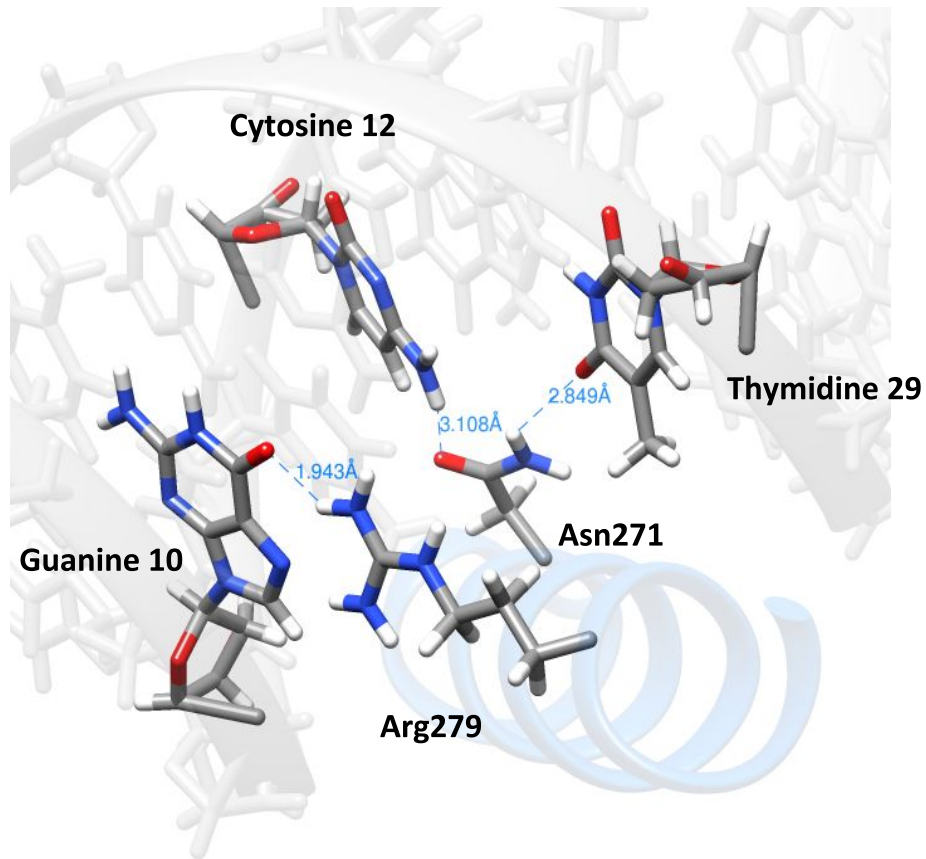
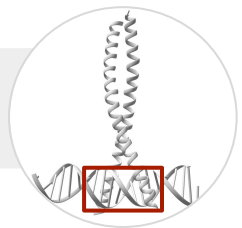
c-FOS c-JUN

IMPORTANT RESIDUES AND INTERACTIONS



c-FOS c-JUN

IMPORTANT RESIDUES AND INTERACTIONS

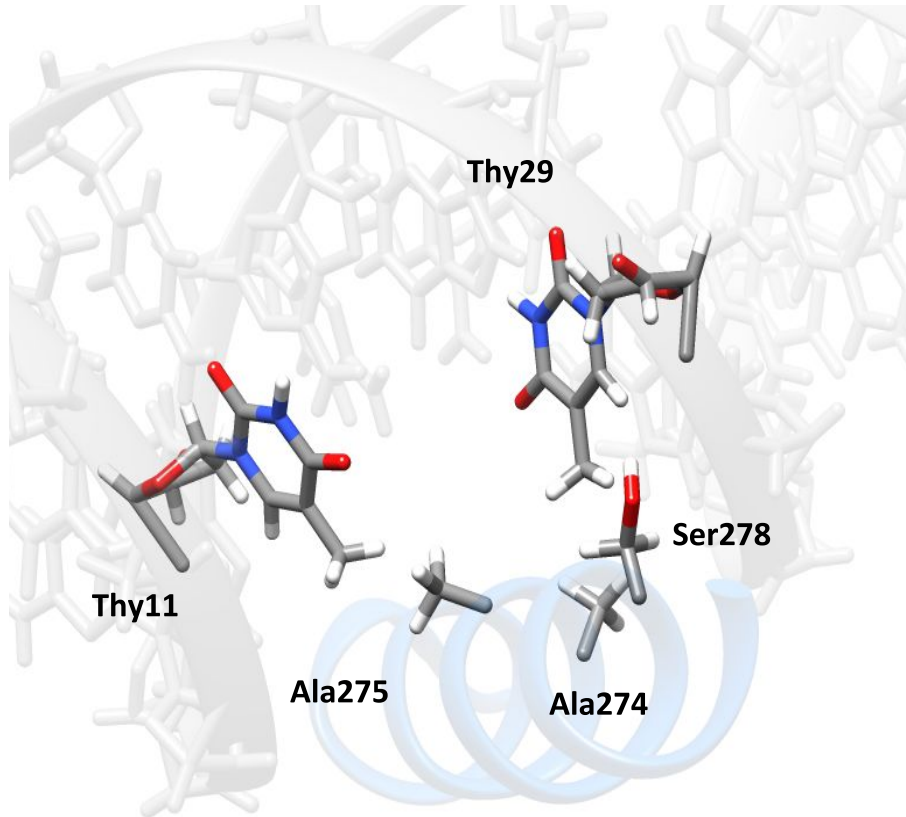
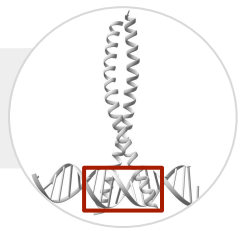


c-Jun (complex II) DNA interactions

Hydrogen bonds	
Asn271 - Thy29 (D)	2.849 Å
Asn271 - Cyt12 (C)	3.108 Å
Arg279 - Gua10 (C)	1.943 Å

c-FOS c-JUN

IMPORTANT RESIDUES AND INTERACTIONS

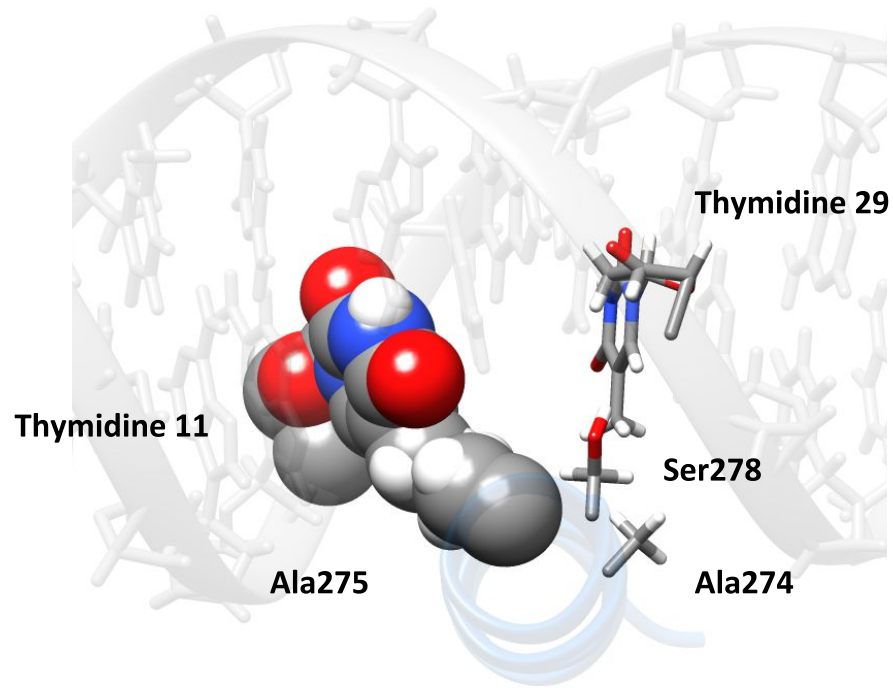
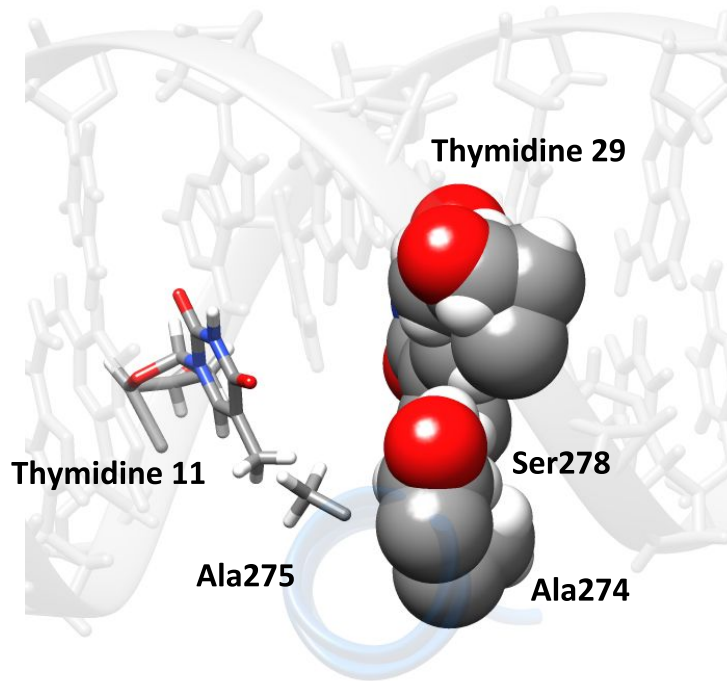


c-Jun (complex II) DNA interactions

Van der Waals contact
Ala275 - Thy29 (D)
Ser278 - Thy29 (D)
Ala274 - Thy11 (C)

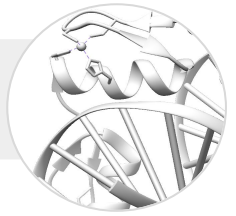
c-FOS c-JUN

IMPORTANT RESIDUES AND INTERACTIONS

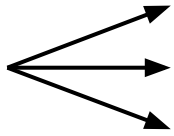




Early Growth Factor (EGR1)

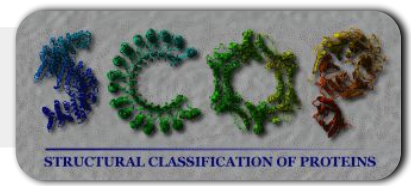


EGR1 Regulates

1. Response to 
 - Growth factors
 - DNA damage
 - Ischemia and hypoxia
2. The expression proteins involved in **inflammatory processes**
3. Biosynthesis of luteinizing hormone (LHB) in the pituitary
4. The amplitude of the expression rhythms of **clock genes**
5. Expression of p53/TP53 and TGFB1 → **prevent tumor formation**

EGR1/Zif268

CLASSIFICATION



ROOT

Scop

CLASS

Small proteins

FOLD

beta-beta-alpha zinc fingers

SUPERFAMILY

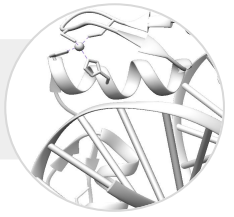
beta-beta-alpha zinc fingers

FAMILY

C2H2 finger

EGR1/Zif268

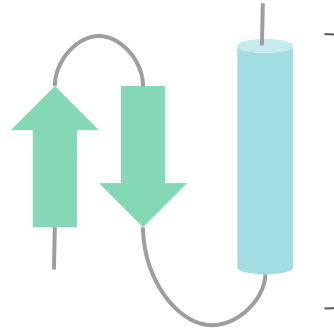
CYS2-HIS2 LIKE FOLD



Three tandem Zinc fingers.

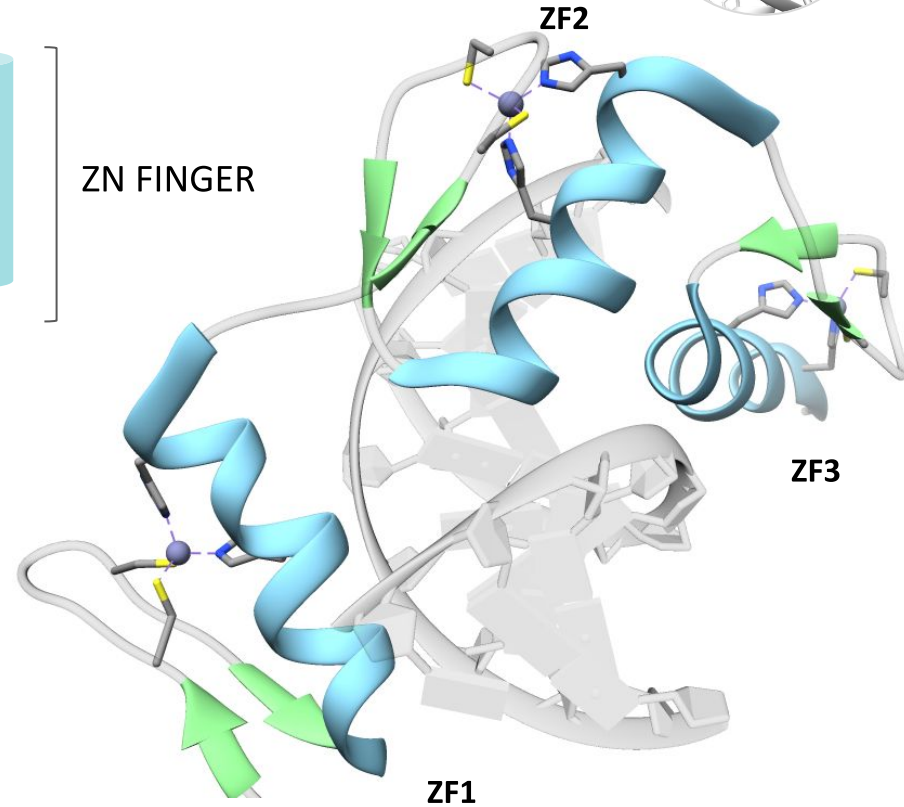
$\beta\beta\alpha$ fold:

- antiparallel beta sheet
- alpha helix



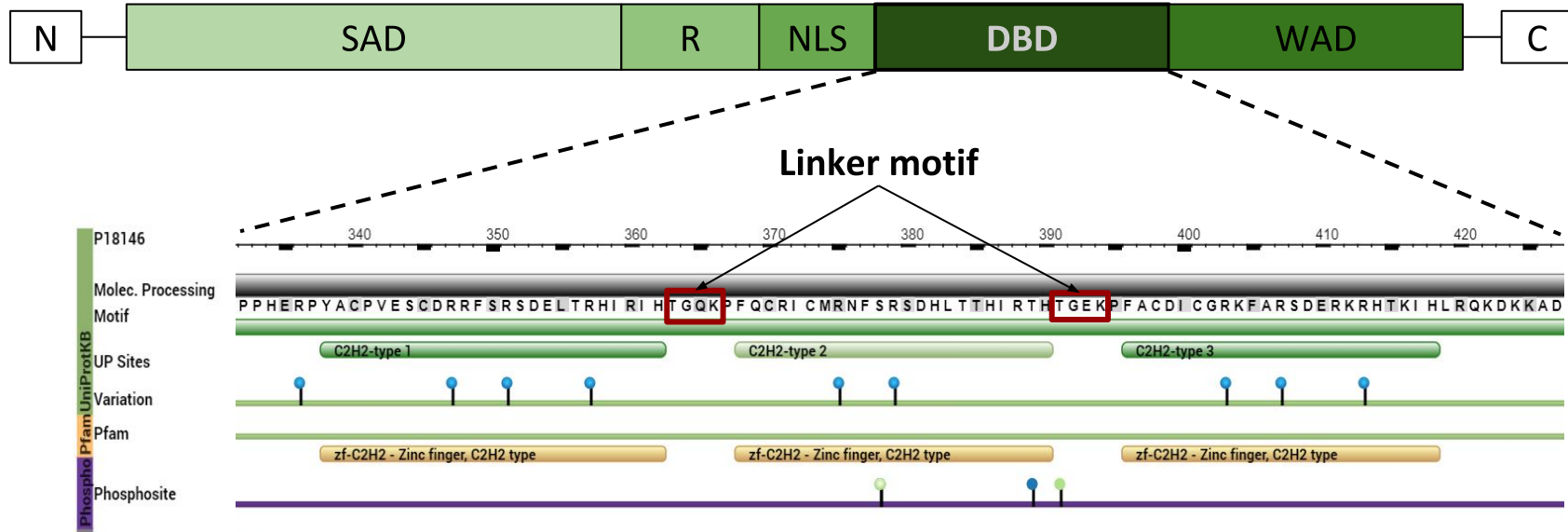
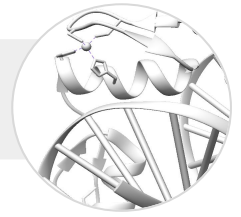
Each zinc finger interacts with a **zinc ion**.

Fits into the major groove of B-DNA.



EGR1/Zif268

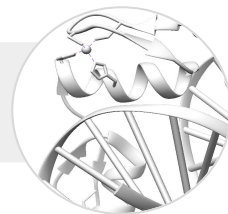
GENERAL STRUCTURE



SAD = strong activation domain; R = repressor domain; NLS = nuclear localization site; DBD = DNA binding domain; WAD = weak activation domain

EGR1/Zif268

SECONDARY STRUCTURE - SAME FAMILY STRUCTURES



Sequence alignment

```
ZBTB38 -----pyacelcakqfqspstlkmhmr
Gli      etdcrwdgcsqefdsqeqlvhhinsehihgerkefvchwggcsrelrpfkaqymlvvhmr
WT1      -----
Egr1     -----
```

ZINC FINGER 1

ZINC FINGER 2

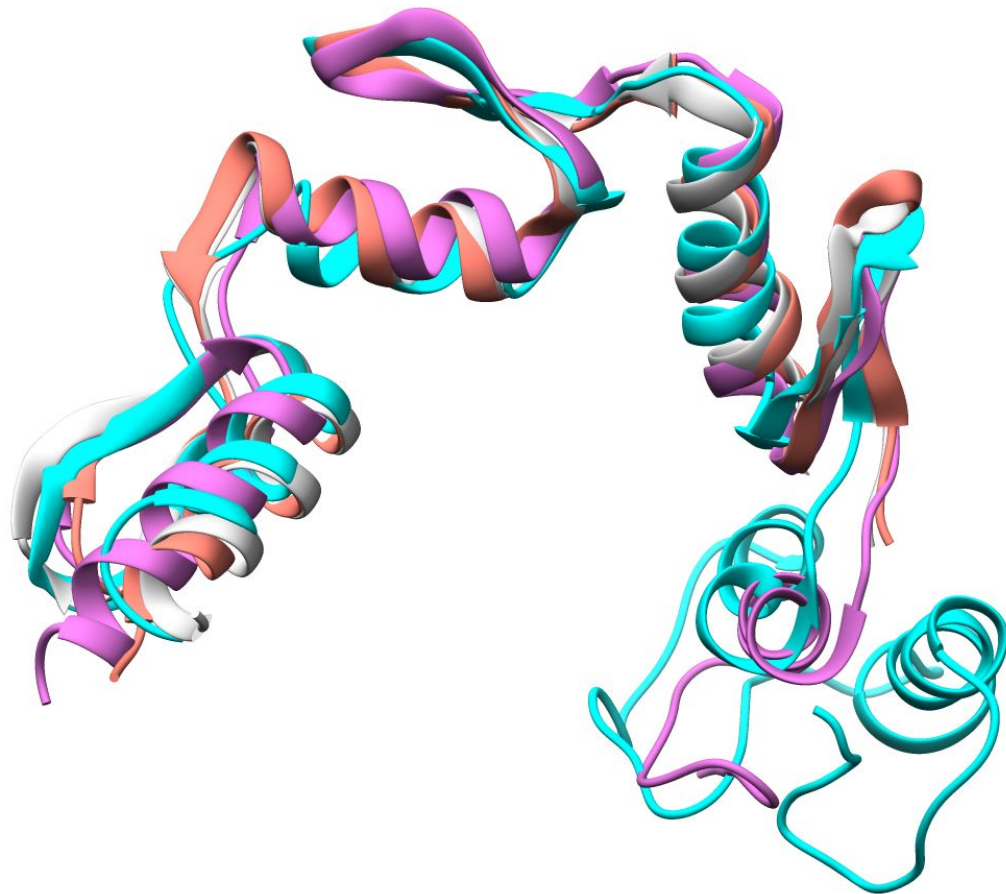
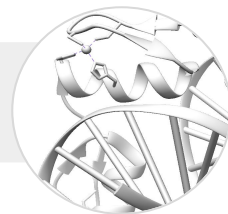
```
ZBTB38 chtgekpYQCK--TCGRCF SVQGNLQKHERIHlgk fvcqycnkaftlnetlkiherih
Gli      rhtgekpHKCTFeGCRKSYSRLENLKTHLRSHtgekpymcehegcskafsnasdrakhqn
WT1      ----ekpYQCDFkDCERRFSRSDQLKRHQRRHtgvkpfqcktcqrkfsrsdhlkthtrth
Egr1     ----erpYACPVeSCDRRFSRSELT RHIRIHtgqkpfqcrimrnfsrsdhlththirth
```

ZINC FINGER 3

```
ZBTB38 tgekryhcqfcfqrfllylstkrnheqrhireh---
Gli      rthsnekpypvcklpgctkrytdpsslrkhvktvhg
WT1      tgekp fscrwp scqkkfarsdelvrhnhmhqr---
Egr1     tgekp facdic grkfarsderkrhtkihlr----
```

EGR1/Zif268

SECONDARY STRUCTURE - SAME FAMILY STRUCTURES

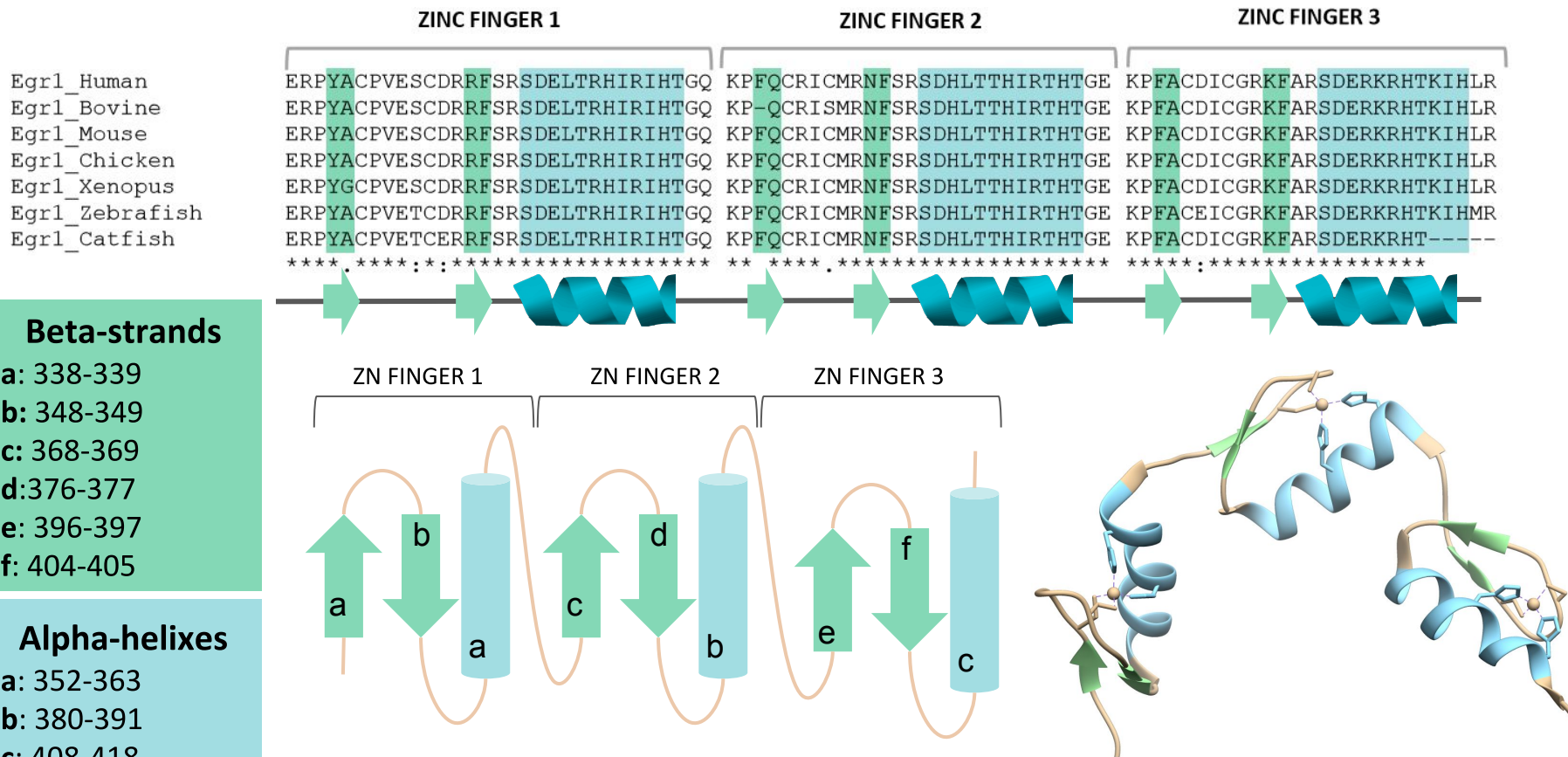


Structural alignment

- 2gli → Zinc Finger GLI2 (Gli)
- 6e93 → ZBTB38
- 4r2e → Wilms Tumor Protein (WT1)
- 4r2a → Early Growth Factor (Egr1)

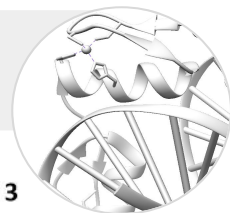
Sc 5.87
RMSD 1.83

EVOLUTION AMONG DIFFERENT SPECIES - SECONDARY STRUCTURE

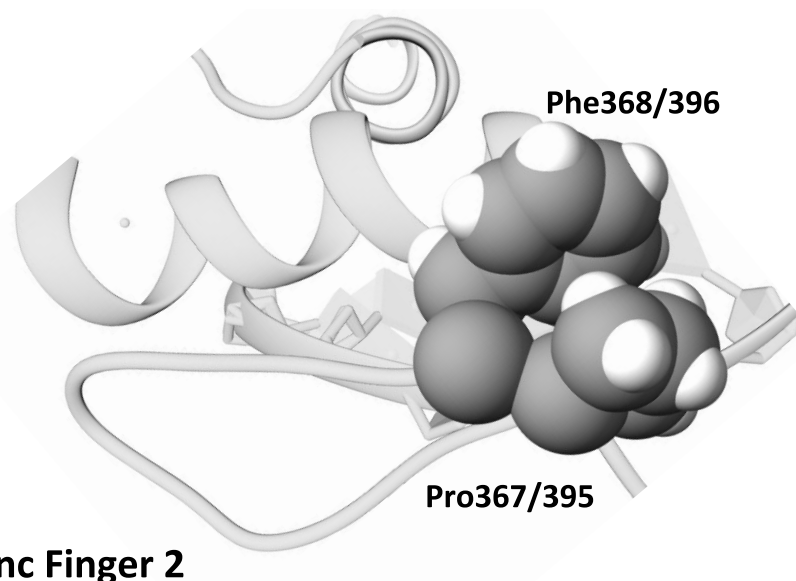
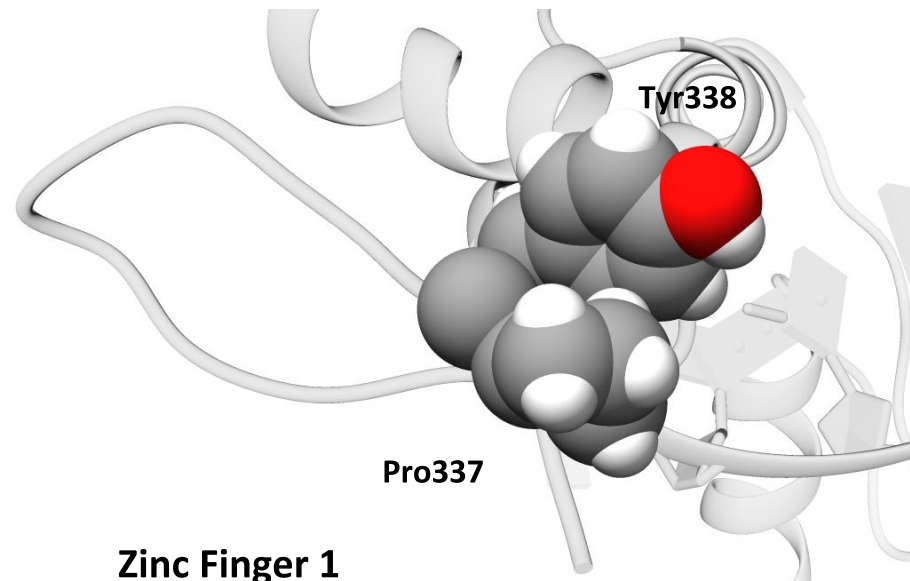


EGR1/Zif268

EVOLUTION AMONG DIFFERENT SPECIES - IMPORTANT RESIDUES



	ZINC FINGER 1	ZINC FINGER 2	ZINC FINGER 3
	338	367	395
Egr1_Human	ERP YACPVESCDRRFSR SDELTRHIRIHTGQ	KPFQCRICMRNFSRSDHLTTHIRHTGE	KPFACDICGRKFARSDEKRHTKIHLR
Egr1_Bovine	ERP YACPVESCDRRFSR SDELTRHIRIHTGQ	KP-QCRISMRNFSRSDHLTTHIRHTGE	KPFACDICGRKFARSDEKRHTKIHLR
Egr1_Mouse	ERP YACPVESCDRRFSR SDELTRHIRIHTGQ	KPFQCRICMRNFSRSDHLTTHIRHTGE	KPFACDICGRKFARSDEKRHTKIHLR
Egr1_Chicken	ERP YACPVESCDRRFSR SDELTRHIRIHTGQ	KPFQCRICMRNFSRSDHLTTHIRHTGE	KPFACDICGRKFARSDEKRHTKIHLR
Egr1_Xenopus	ERP YGCPVESCDRRFSR SDELTRHIRIHTGQ	KPFQCRICMRNFSRSDHLTTHIRHTGE	KPFACDICGRKFARSDEKRHTKIHLR
Egr1_Zebrafish	ERP YACPVETCDRRFSR SDELTRHIRIHTGQ	KPFQCRICMRNFSRSDHLTTHIRHTGE	KPFACEICGRKFARSDEKRHTKIHMR
Egr1_Catfish	ERP YACPVETCERRFSR SDELTRHIRIHTGQ	KPFQCRICMRNFSRSDHLTTHIRHTGE	KPFACDICGRKFARSDEKRHT----
	..*:*****	** ****.*****	*****:*****



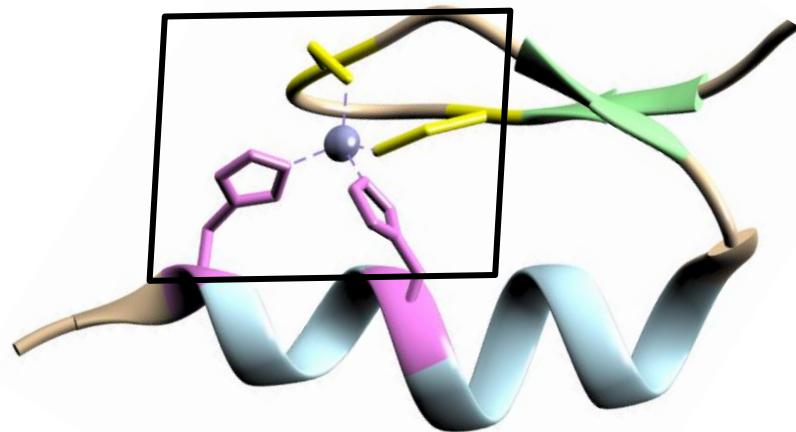
EGR1/Zif268

EVOLUTION AMONG DIFFERENT SPECIES - IMPORTANT RESIDUES



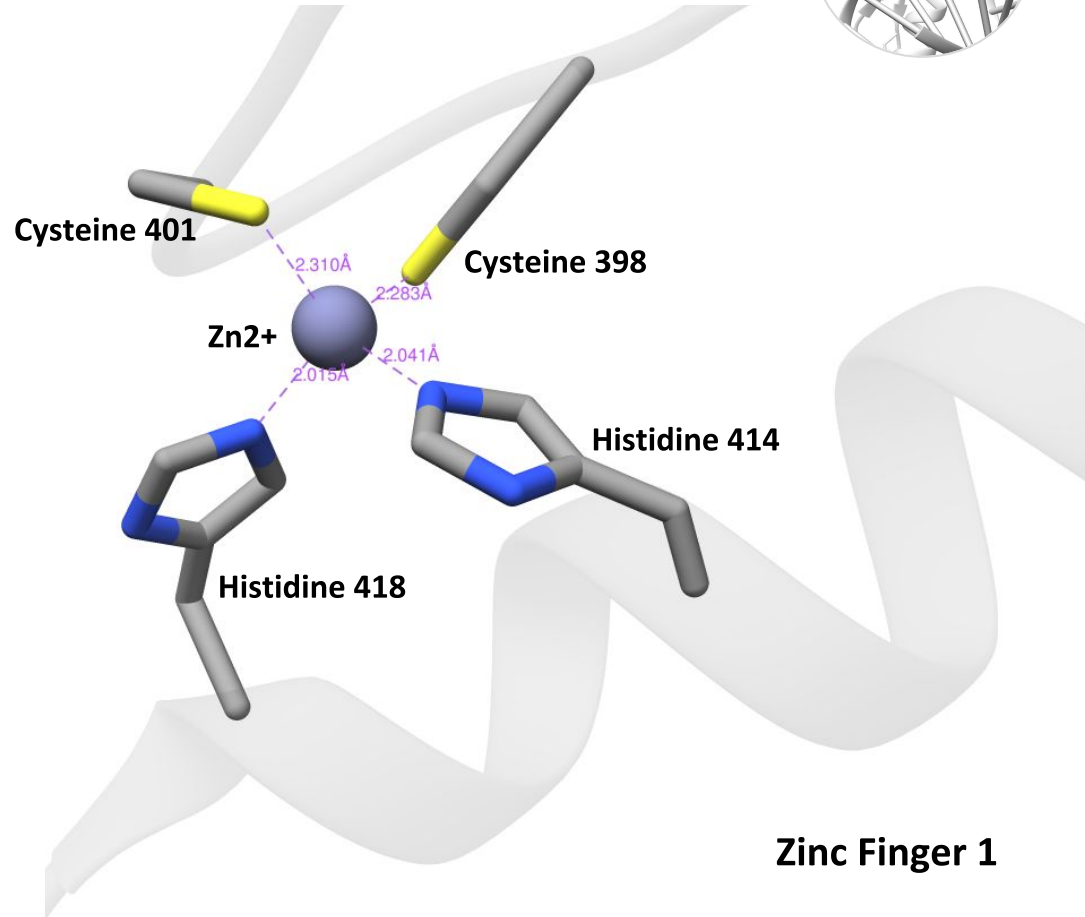
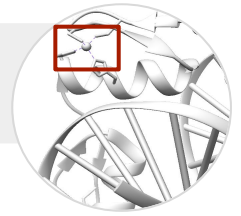
	ZINC FINGER 1	ZINC FINGER 2	ZINC FINGER 3
Egr1_Human	ERPYACPVESCDRRFSRSD EL TRHIRIHTGQ	KPFQCRICMRNFSRSDH LT THIRHTHTGE	KPFACDICGRKFARSDERKRHTKIHLR
Egr1_Bovine	ERPYACPVESCDRRFSRSD EL TRHIRIHTGQ	KP-QCRISMRNFSRSDH LT THIRHTHTGE	KPFACDICGRKFARSDERKRHTKIHLR
Egr1_Mouse	ERPYACPVESCDRRFSRSD EL TRHIRIHTGQ	KPFQCRICMRNFSRSDH LT THIRHTHTGE	KPFACDICGRKFARSDERKRHTKIHLR
Egr1_Chicken	ERPYACPVESCDRRFSRSD EL TRHIRIHTGQ	KPFQCRICMRNFSRSDH LT THIRHTHTGE	KPFACDICGRKFARSDERKRHTKIHLR
Egr1_Xenopus	ERPYGCPVESCDRRFSRSD EL TRHIRIHTGQ	KPFQCRICMRNFSRSDH LT THIRHTHTGE	KPFACDICGRKFARSDERKRHTKIHLR
Egr1_Zebrafish	ERPYACPVETCDRRFSRSD EL TRHIRIHTGQ	KPFQCRICMRNFSRSDH LT THIRHTHTGE	KPFACEICGRKFARSDERKRHTKIHR
Egr1_Catfish	ERPYACPVETCERRFSRSD EL TRHIRIHTGQ	KPFQCRICMRNFSRSDH LT THIRHTHTGE	KPFACDICGRKFARSDERKRHT-----
	****.****:*:*****	** ****.*****	*****:*****

	Cys (C)	His (H)
ZF1	Cys340 Cys345	His358 His362
ZF2	Cys370 Cys373	His386 His390
ZF3	Cys398 Cys401	His414 His418



EGR1/Zif268

EVOLUTION AMONG DIFFERENT SPECIES - IMPORTANT RESIDUES

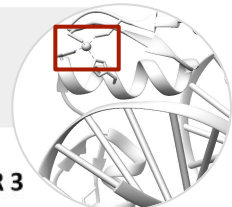


Ionic-bonding

Cys398 - Zn ²⁺	2.283 Å
Cys401 - Zn ²⁺	2.310 Å
His414 - Zn ²⁺	2.041 Å
His418 - Zn ²⁺	2.015 Å

EGR1/Zif268

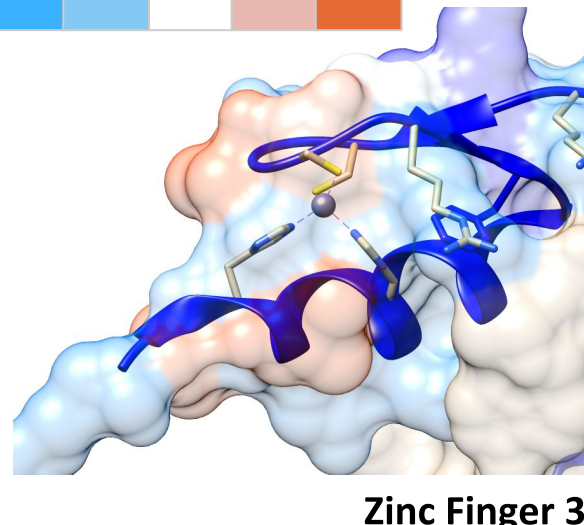
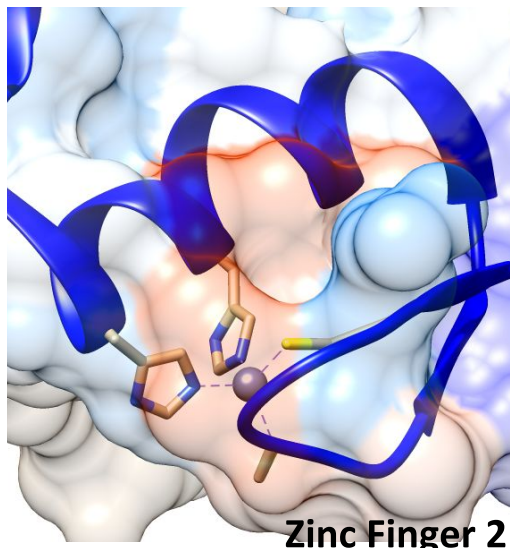
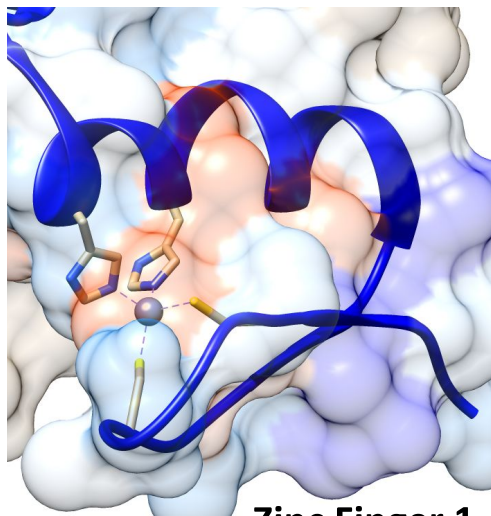
EVOLUTION AMONG DIFFERENT SPECIES - IMPORTANT RESIDUES



	ZINC FINGER 1	ZINC FINGER 2	ZINC FINGER 3
Egr1_Human	EFPPYACPVESCDRRFESRSDELTRHIRIHTGQ	KPFQCRICMRNFSRSDDLTTTHIRHTHTGE	KPFACDICGRKFFARSDEKRHTKIHLR
Egr1_Bovine	EFPPYACPVESCDRRFESRSDELTRHIRIHTGQ	KP-QCRISMNFSRSDDLTTTHIRHTHTGE	KPFACDICGRKFFARSDEKRHTKIHLR
Egr1_Mouse	EFPPYACPVESCDRRFESRSDELTRHIRIHTGQ	KPFQCRICMRNFSRSDDLTTTHIRHTHTGE	KPFACDICGRKFFARSDEKRHTKIHLR
Egr1_Chicken	EFPPYACPVESCDRRFESRSDELTRHIRIHTGQ	KPFQCRICMRNFSRSDDLTTTHIRHTHTGE	KPFACDICGRKFFARSDEKRHTKIHLR
Egr1_Xenopus	EFPPYACPVESCDRRFESRSDELTRHIRIHTGQ	KPFQCRICMRNFSRSDDLTTTHIRHTHTGE	KPFACDICGRKFFARSDEKRHTKIHLR
Egr1_Zebrafish	EFPPYACPVETCDRRFESRSDELTRHIRIHTGQ	KPFQCRICMRNFSRSDDLTTTHIRHTHTGE	KPFACEICGRKFFARSDEKRHTKIHLR
Egr1_Catfish	EFPPYACPVETCERRFESRSDELTRHIRIHTGQ	KPFQCRICMRNFSRSDDLTTTHIRHTHTGE	KPFACDICGRKFFARSDEKRHT---
	*****.*:*****	* * * * *.*****	*****.******

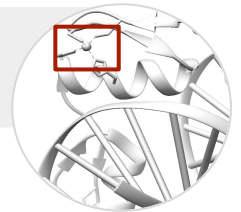
Hydrophilic

Hydrophobic

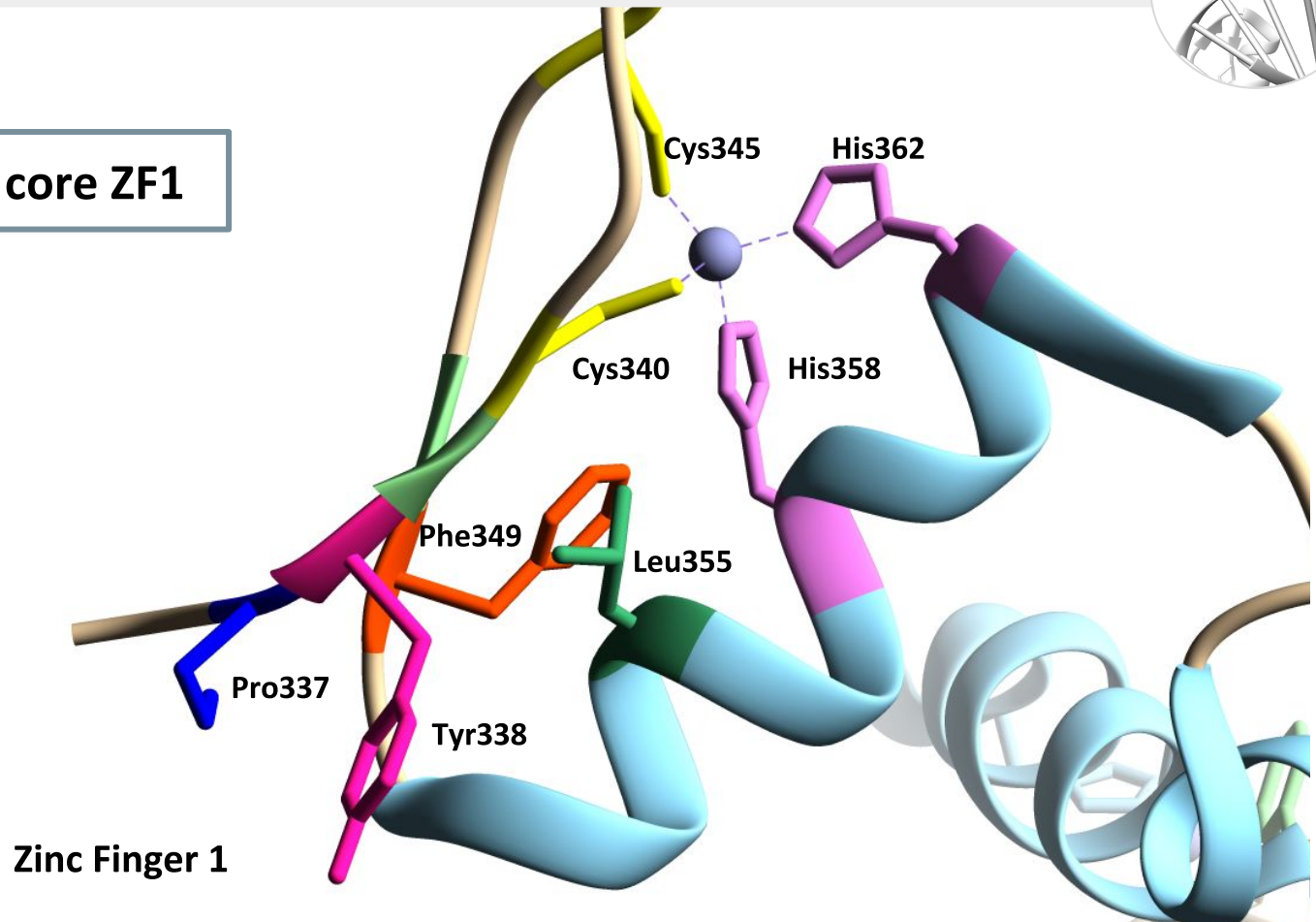


EGR1/Zif268

EVOLUTION AMONG DIFFERENT SPECIES - IMPORTANT RESIDUES

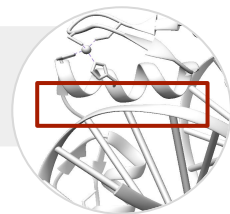


Hydrophobic core ZF1

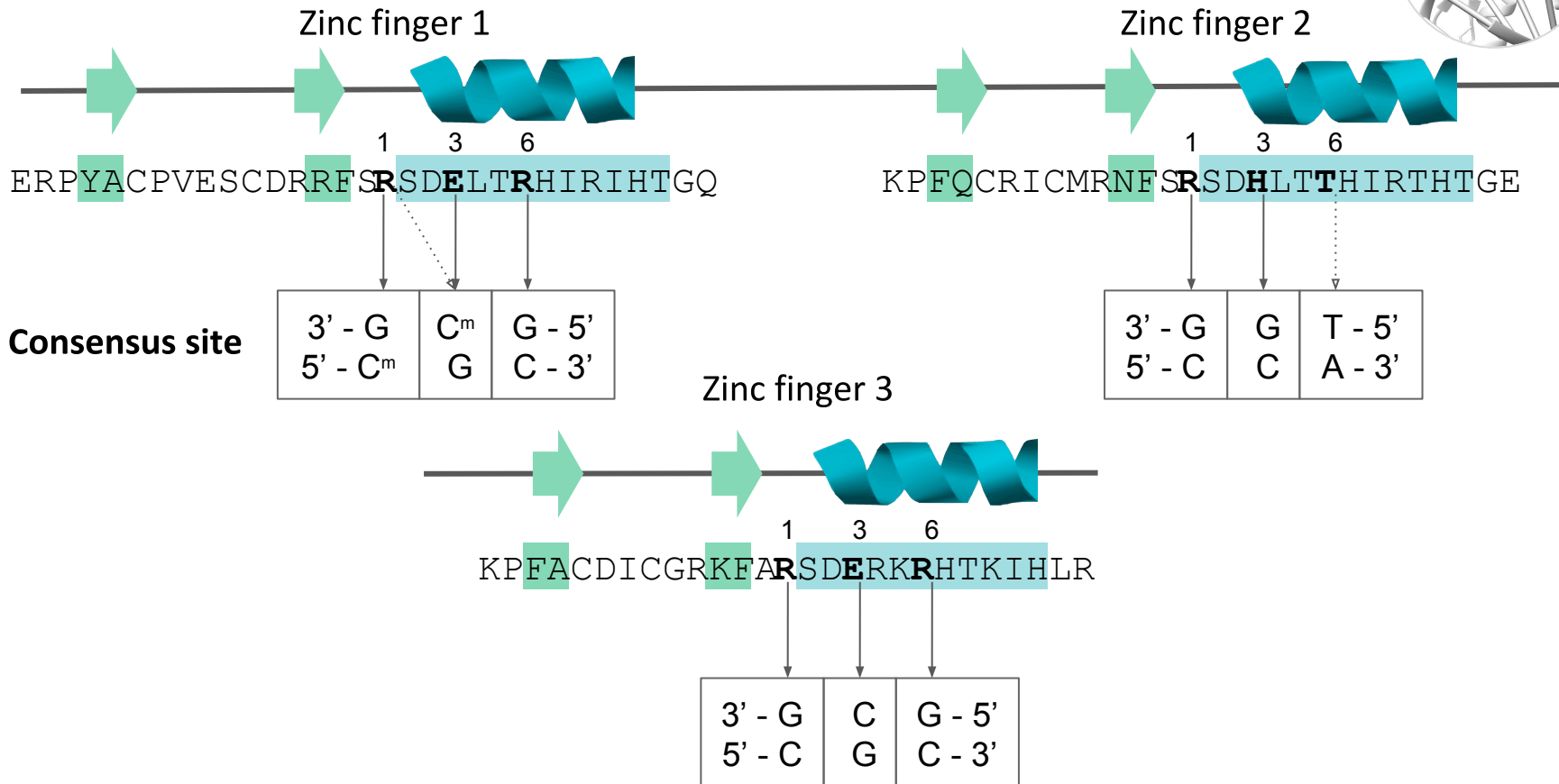


EGR1/Zif268

EVOLUTION AMONG DIFFERENT SPECIES - DNA BINDING INTERACTIONS

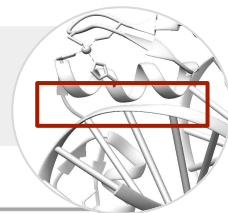


ZRE Consensus site



EGR1/Zif268

EVOLUTION AMONG DIFFERENT SPECIES - DNA BINDING RESIDUES



	ZINC FINGER 1												ZINC FINGER 2												ZINC FINGER 3																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
	351 354 357												379 382 385												407 410 413																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Egr1_Human	ERP	Y	A	C	P	V	E	S	C	D	R	R	F	S	R	S	D	E	L	T	R	H	I	R	I	H	T	G	Q	K	P	F	Q	C	R	I	C	M	R	N	F	S	R	S	D	H	L	T	H	I	R	T	H	T	G	E	K	P	F	A	C	D	I	C	G	R	K	F	A	R	S	D	E	R	K	R	H	T	K	I	H	L	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
Egr1_Bovine	ERP	Y	A	C	P	V	E	S	C	D	R	R	F	S	R	S	D	E	L	T	R	H	I	R	I	H	T	G	Q	K	P	-	Q	C	R	I	S	M	R	N	F	S	R	S	D	H	L	T	H	I	R	T	H	T	G	E	K	P	F	A	C	D	I	C	G	R	K	F	A	R	S	D	E	R	K	R	H	T	K	I	H	L	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
Egr1_Mouse	ERP	Y	A	C	P	V	E	S	C	D	R	R	F	S	R	S	D	E	L	T	R	H	I	R	I	H	T	G	Q	K	P	F	Q	C	R	I	C	M	R	N	F	S	R	S	D	H	L	T	H	I	R	T	H	T	G	E	K	P	F	A	C	D	I	C	G	R	K	F	A	R	S	D	E	R	K	R	H	T	K	I	H	L	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
Egr1_Chicken	ERP	Y	A	C	P	V	E	S	C	D	R	R	F	S	R	S	D	E	L	T	R	H	I	R	I	H	T	G	Q	K	P	F	Q	C	R	I	C	M	R	N	F	S	R	S	D	H	L	T	H	I	R	T	H	T	G	E	K	P	F	A	C	D	I	C	G	R	K	F	A	R	S	D	E	R	K	R	H	T	K	I	H	L	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
Egr1_Xenopus	ERP	Y	G	C	P	V	E	S	C	D	R	R	F	S	R	S	D	E	L	T	R	H	I	R	I	H	T	G	Q	K	P	F	Q	C	R	I	C	M	R	N	F	S	R	S	D	H	L	T	H	I	R	T	H	T	G	E	K	P	F	A	C	D	I	C	G	R	K	F	A	R	S	D	E	R	K	R	H	T	K	I	H	L	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
Egr1_Zebrafish	ERP	Y	A	C	P	V	E	T	C	D	R	R	F	S	R	S	D	E	L	T	R	H	I	R	I	H	T	G	Q	K	P	F	Q	C	R	I	C	M	R	N	F	S	R	S	D	H	L	T	H	I	R	T	H	T	G	E	K	P	F	A	C	E	I	C	G	R	K	F	A	R	S	D	E	R	K	R	H	T	K	I	H	M	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
Egr1_Catfish	ERP	Y	A	C	P	V	E	T	C	E	R	R	F	S	R	S	D	E	L	T	R	H	I	R	I	H	T	G	Q	K	P	F	Q	C	R	I	C	M	R	N	F	S	R	S	D	H	L	T	H	I	R	T	H	T	G	E	K	P	F	A	C	D	I	C	G	R	K	F	A	R	S	D	E	R	K	R	H	T	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

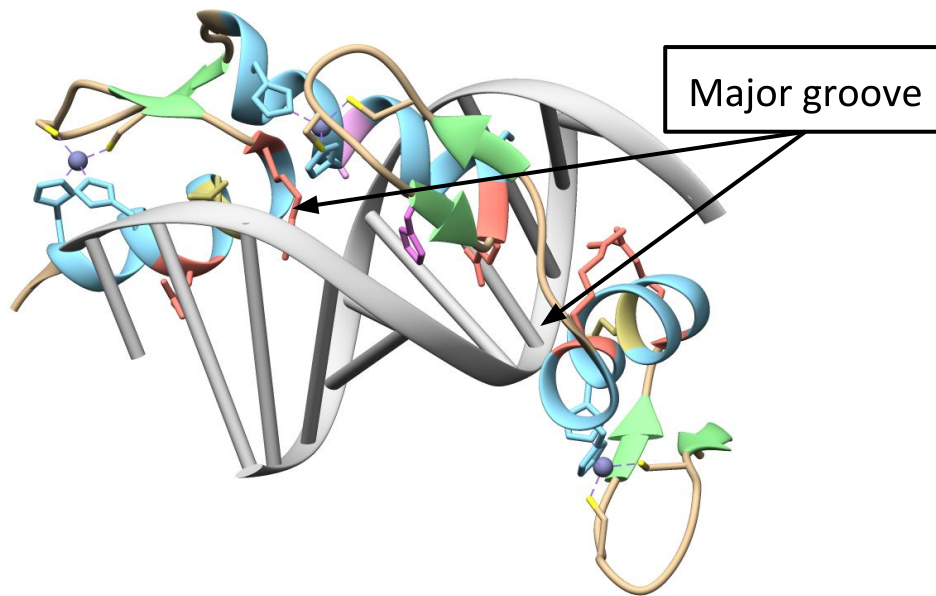
Residue Arginine (R)

Residue Glutamic Acid (E)

Residue Threonine (T)

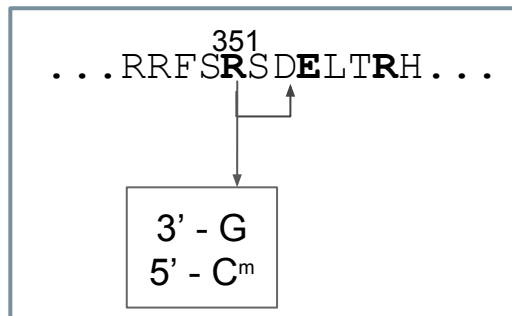
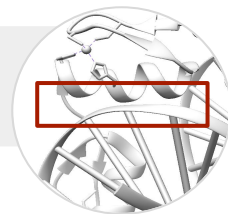
Residue Histidine (H)

ZRE Consensus site:
5'-GCG TGG GCG-3'



EGR1/Zif268

IMPORTANT RESIDUES AND INTERACTIONS

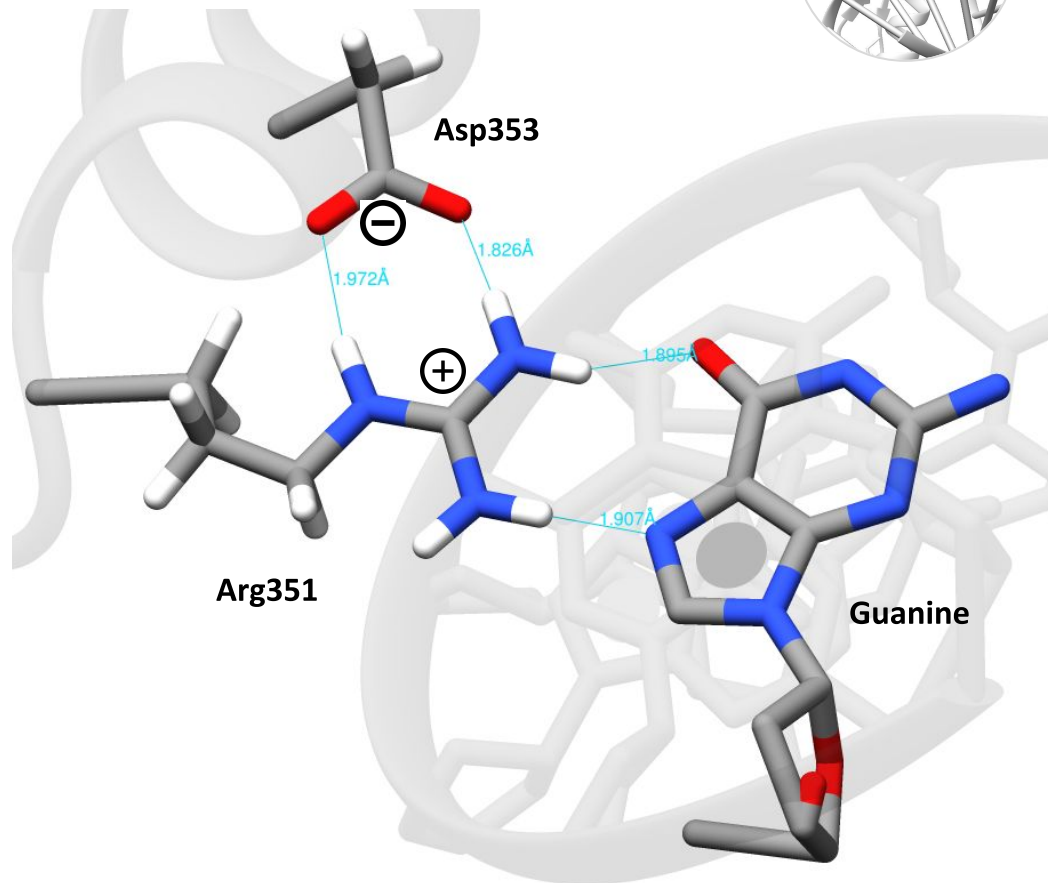


Hydrogen bond - Salt bridge

Asp353 (O) - Arg351 (H)	1.826 Å
Asp353 (O) - Arg351 (H)	1.972 Å

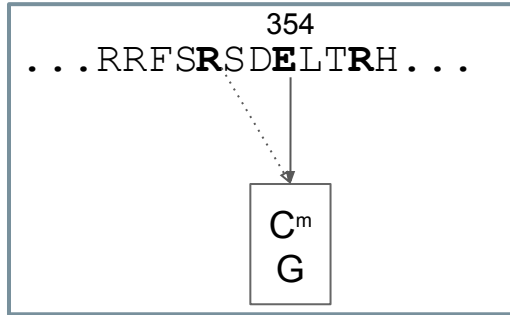
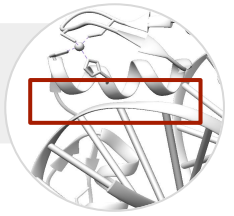
Hydrogen bonds

Arg351 (H) - DG10 (O)	1.895 Å
Arg351 (H) - DG10 (N)	1.907 Å

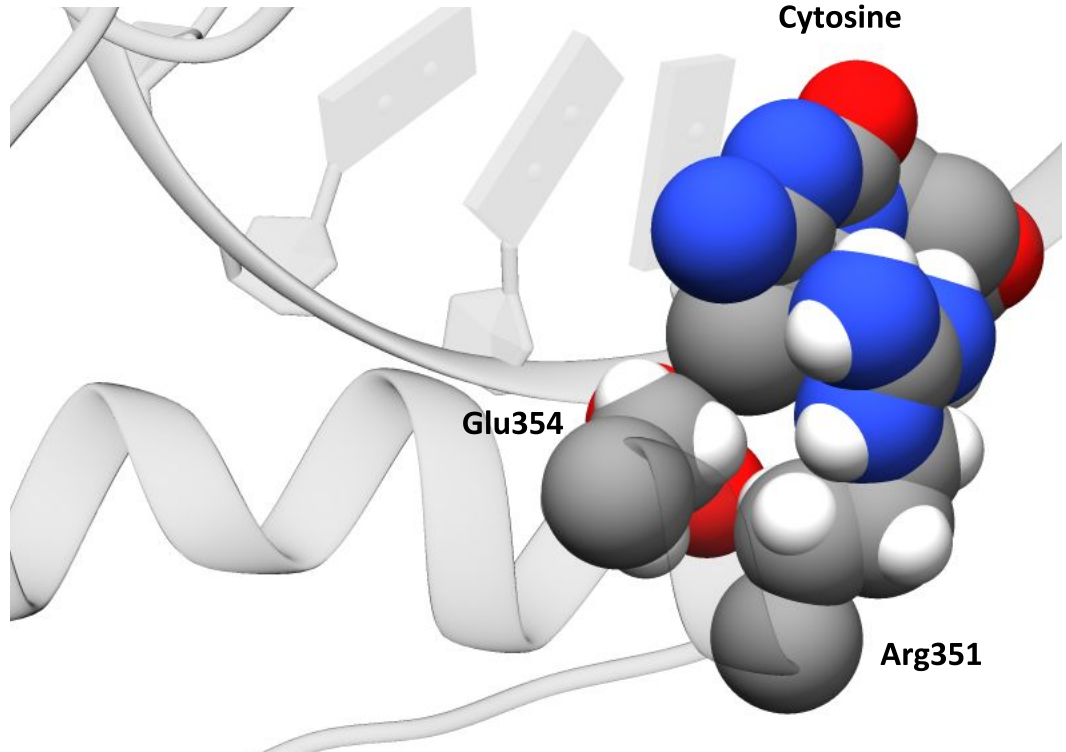


EGR1/Zif268

IMPORTANT RESIDUES AND INTERACTIONS

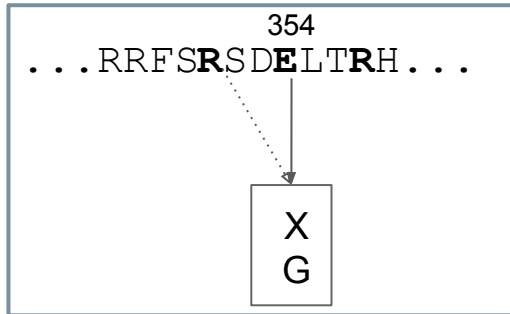


Van der Waals contact



EGR1/Zif268

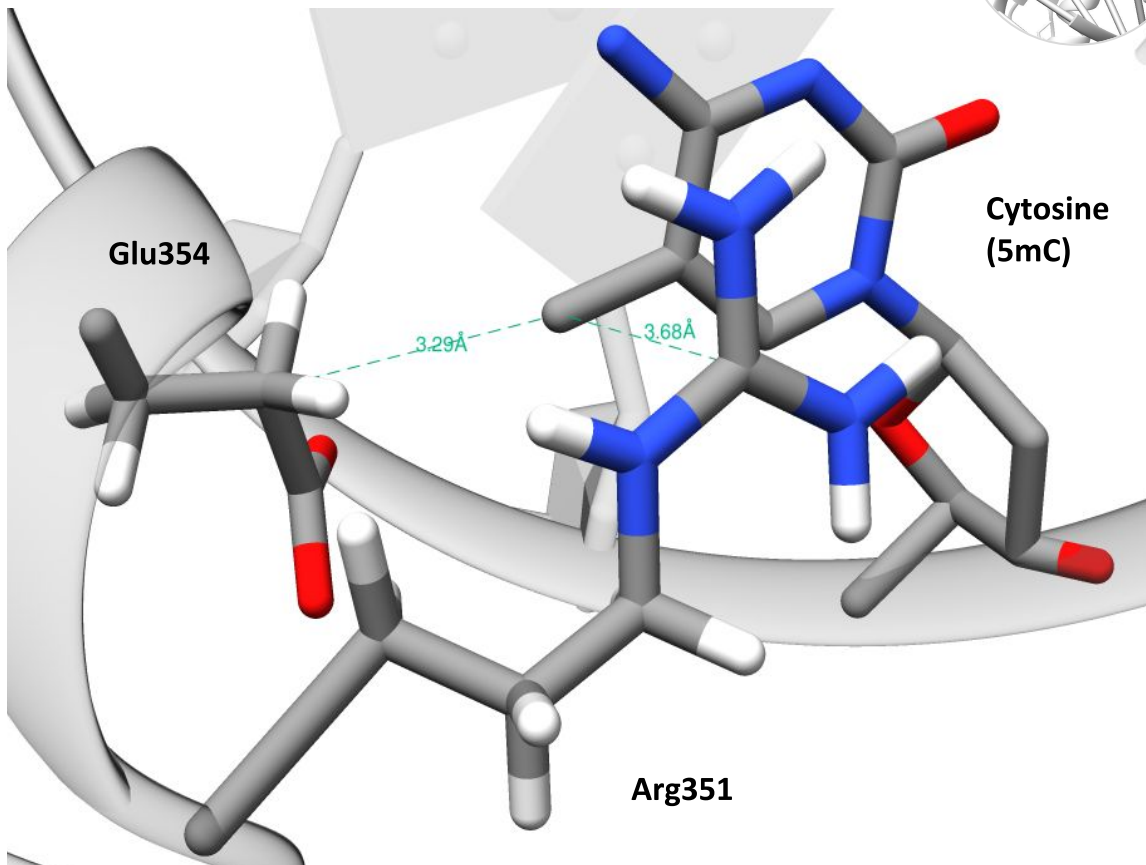
IMPORTANT RESIDUES AND INTERACTIONS



Van der Waals contacts

Arg351 - 5CM9	3.68 Å
Glu354 - 5CM9	3.29 Å

Methyl-Arg-Gua tirad

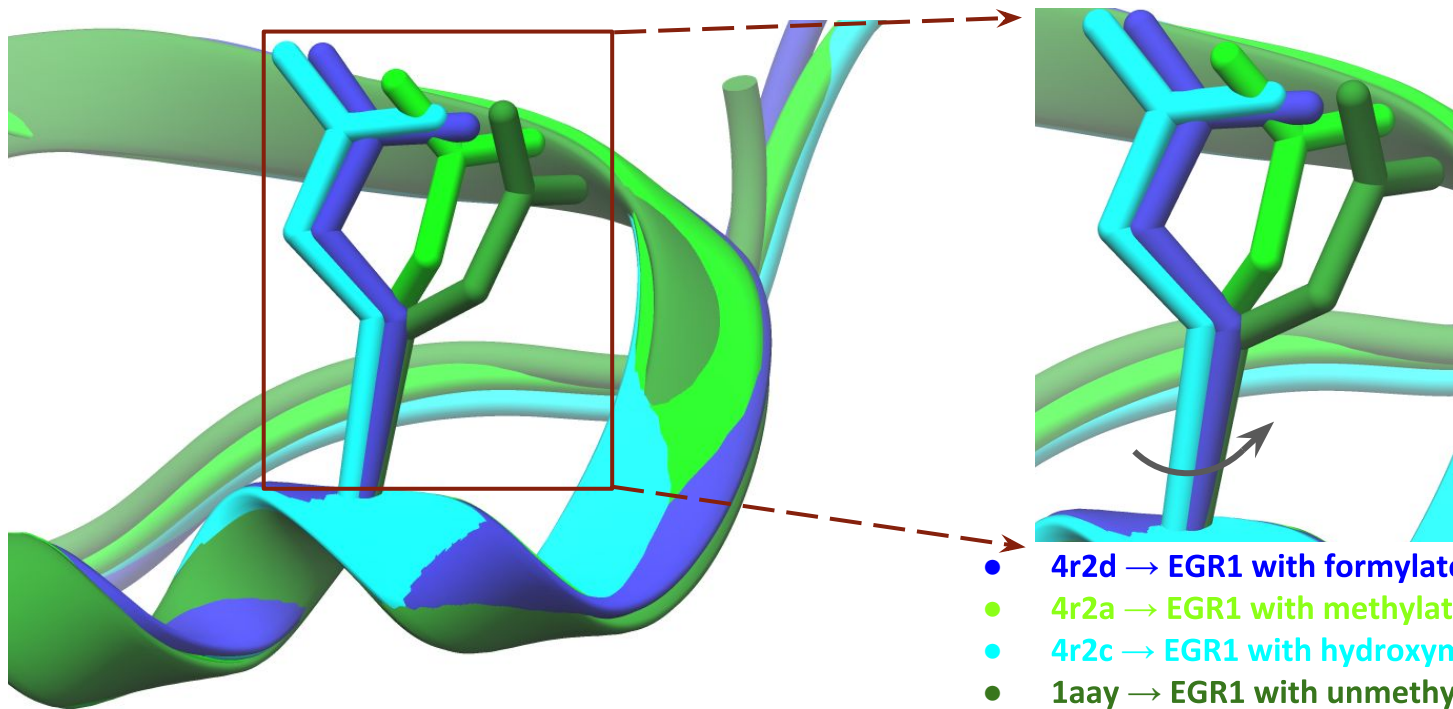
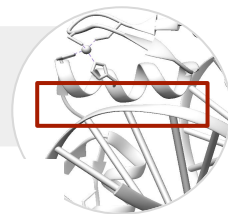


EGR1/Zif268

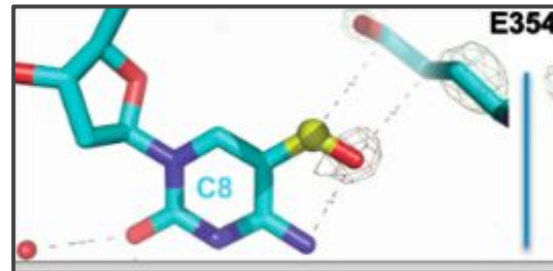
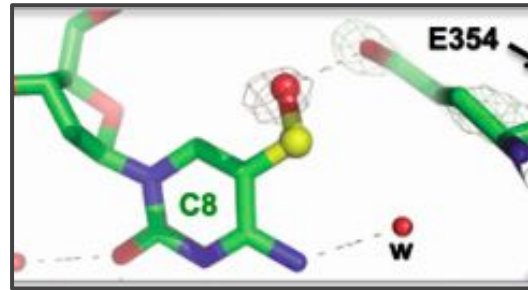
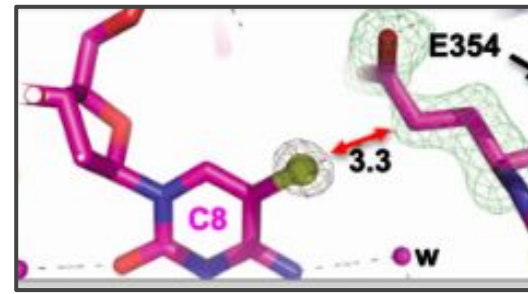
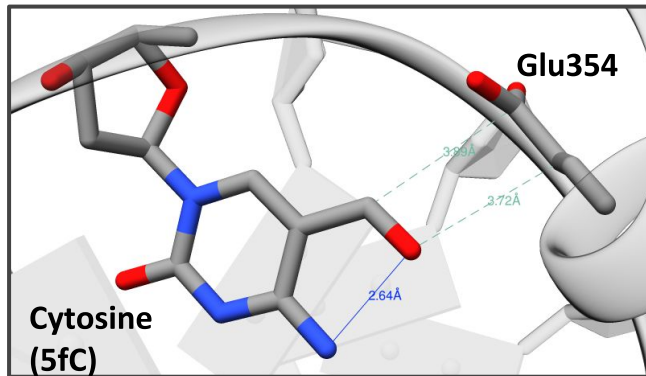
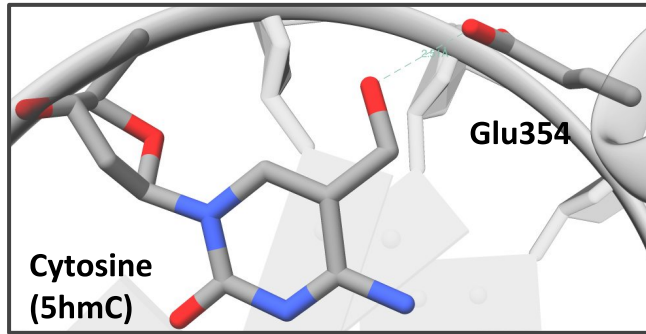
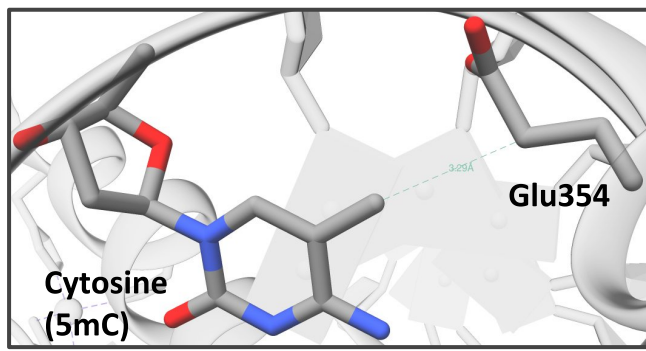
IMPORTANT RESIDUES AND INTERACTIONS

CLUSTAL W(1.60) multiple sequence alignment

1aay	-RPYACPVESCDRRFSRSD	E	LRHIRIHTGQKPFQCRICMRNFSRSDHLTTTHIRTHTGEKPFACDICGRKFARSDERKRHTKIHLR-
4r2dA	ERPYACPVESCDRRFSRSD	E	LRHIRIHTGQKPFQCRICMRNFSRSDHLTTTHIRTHTGEKPFACDICGRKFARSDERKRHTKIHLRQ
4r2aA	ERPYACPVESCDRRFSRSD	E	LRHIRIHTGQKPFQCRICMRNFSRSDHLTTTHIRTHTGEKPFACDICGRKFARSDERKRHTKIHLR-
4r2cA	ERPYACPVESCDRRFSRSD	E	LRHIRIHTGQKPFQCRICMRNFSRSDHLTTTHIRTHTGEKPFACDICGRKFARSDERKRHTKIHLR-



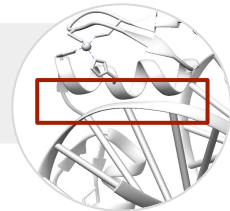
- 4r2d → EGR1 with formylated DNA
- 4r2a → EGR1 with methylated DNA
- 4r2c → EGR1 with hydroxymethylated DNA
- 1aay → EGR1 with unmethylated DNA



Hashimoto H, Olanrewaju OY, Zheng Y, Wilson GG, Zhang X, Cheng X. Wilms tumor protein recognizes 5-carboxylcytosine within a specific DNA sequence. *Genes Dev.* 2014; 28(20): 2304–2313.

EGR1/Zif268

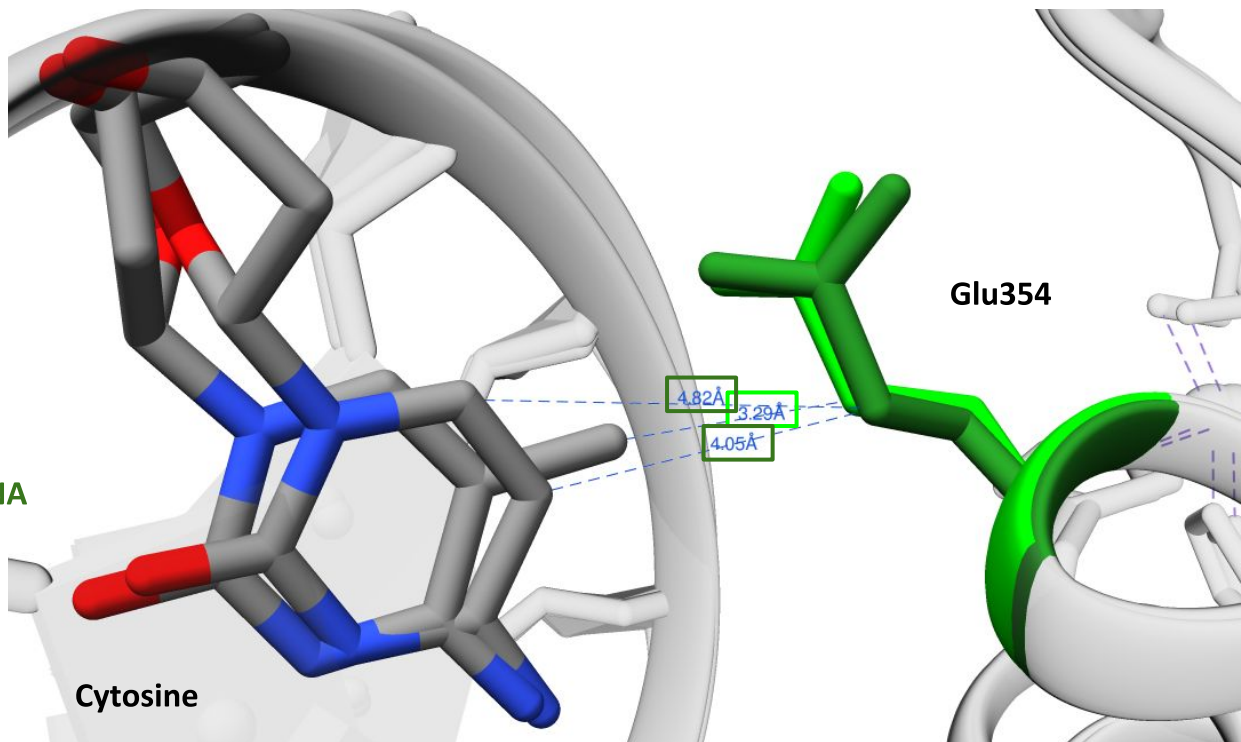
IMPORTANT RESIDUES AND INTERACTIONS



Van der Waals contacts

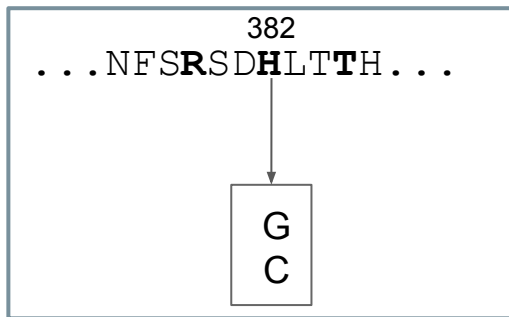
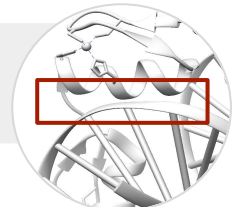
Glu354 - 5CM9	3.29 Å
Glu354 - C9	4.82 Å
Glu354 - C9	4.05 Å

- 4r2a → EGR1 with methylated DNA
- 1aay → EGR1 with unmethylated DNA



EGR1/Zif268

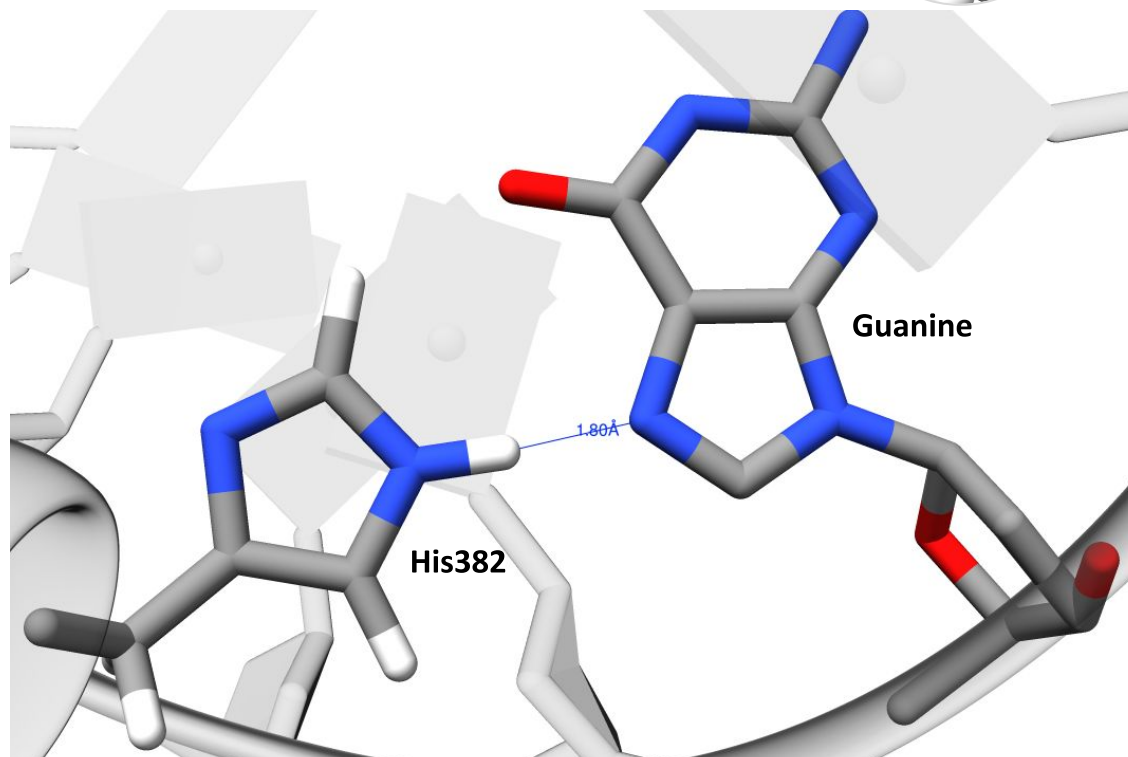
IMPORTANT RESIDUES AND INTERACTIONS



Hydrogen bonds

His382 (H) - DG6 (O)

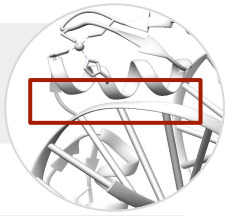
1.8 Å



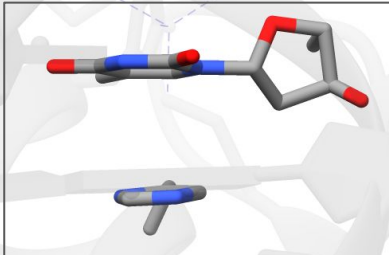
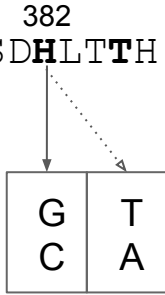
Zinc finger 2

EGR1/Zif268

IMPORTANT RESIDUES AND INTERACTIONS



...NFS**R**SD**H**LT**T**H...



π - Stacking interaction

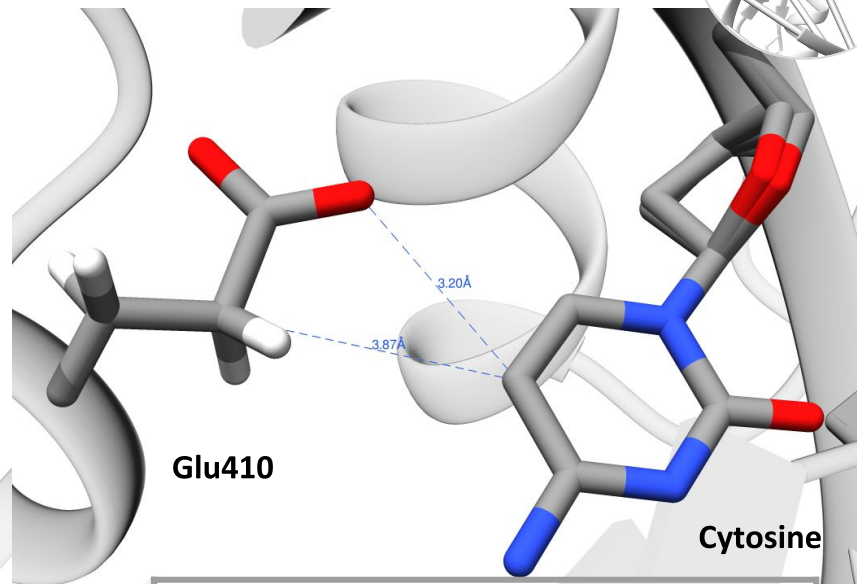
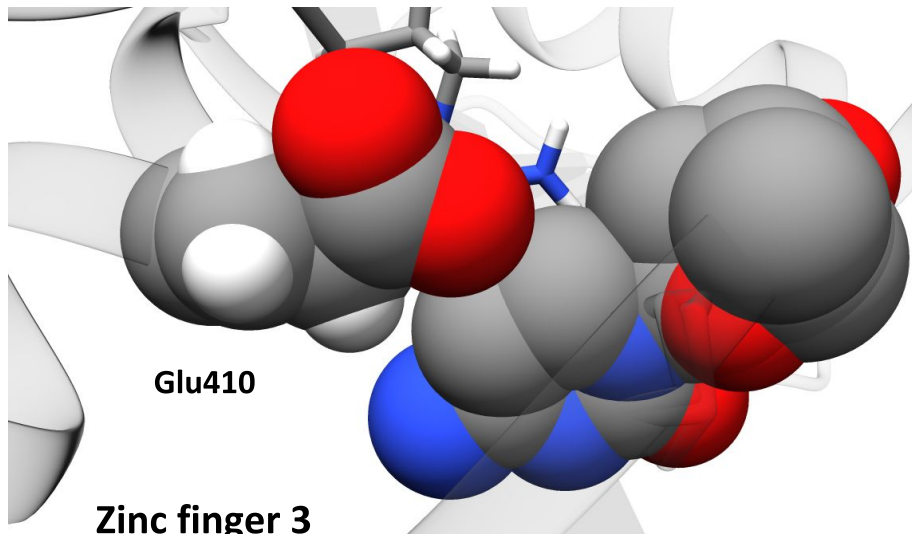
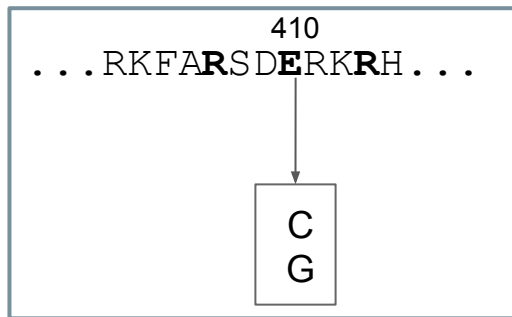
Thymidine

His382

Zinc finger 2

EGR1/Zif268

IMPORTANT RESIDUES AND INTERACTIONS



C-H...O type Hydrogen bond

Glu410 (O) - DC3 (C)

3.20 Å

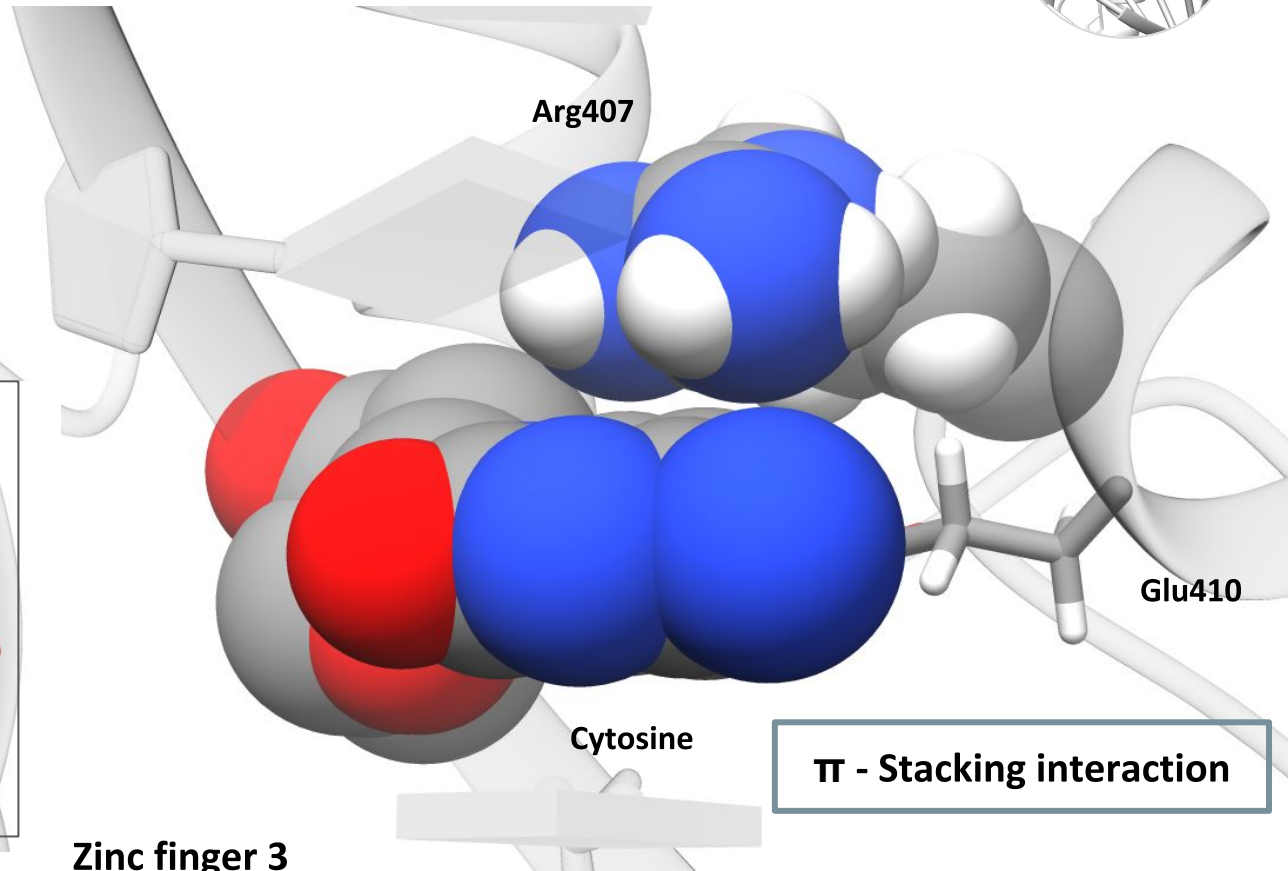
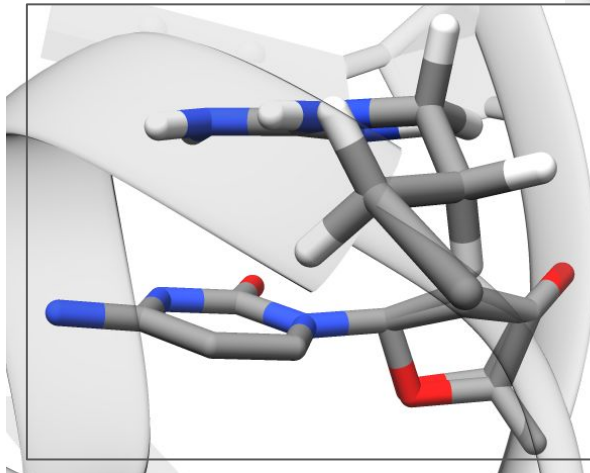
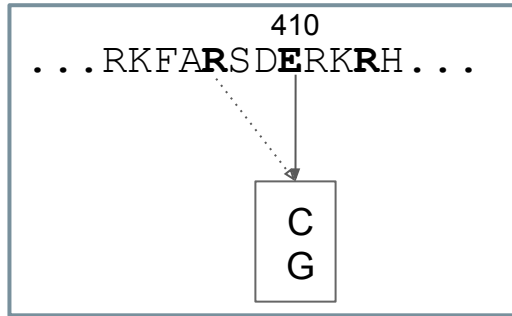
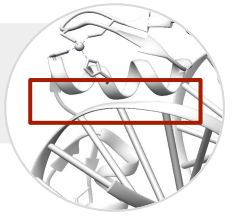
Van der Waals contacts

Glu410 (C) - DC3 (C)

3.87 Å

EGR1/Zif268

IMPORTANT RESIDUES AND INTERACTIONS





CONCLUSIONS

CONCLUSIONS

Transcription factors are important in the regulation of gene expression in order to control the cell viability

c-Fos/c-Jun

- ✓ Within the **same family**, both sequence and structure have been conserved.
- ✓ Comparing with **other species**, there is **conservation** in the two main domains. The only residue involved in a motif that has mutated is a Cys to a Ser.
- ✓ In the **bZIP** family, **leucines** allow the **dimerization** and consensus motif binds to a specific region of the **DNA** (AP-1 site).

Egr1

- ✓ Within the **same family**, the **DNA** binding residues are **not** conserved whereas those involved in the **stabilization** of the fold **are**.
- ✓ During **evolution**, important residues have been **conserved**.
- ✓ In the **C2H2** family, the hydrophobic core holds together the antiparallel **beta-sheets** and **alpha-helices**.
- ✓ Depending on the **state** of the **DNA**, a glutamic acid changes its **conformation**.

PEM QUESTIONS

1. c-Jun/c-Fos TFs...
 - a) Are a heterodimer formed by two alpha-helices
 - b) Bind to AP-1 site of the DNA
 - c) Both are correct
 - d) Form 'knobs into holes' packing
 - e) **All of them are correct**
2. Which of these domains are conserved within evolution?
 - a) Basic DNA binding domain
 - b) Leucine zipper domain
 - c) **Both are correct**
 - d) Catalytic domain
 - e) All of them
3. AP-1 TFs
 - a) Are involved in cell proliferation, survival, apoptosis, transformation and oncogenesis
 - b) Are from Jun, Fos and ATF family
 - c) Almost all of them need to form dimers to bind to the DNA
 - d) Can bind to TRE and CRE sites
 - e) **All of them are correct**
4. The hydrophobic core in c-Jun/c-Fos interaction...
 - a) Is found in the external surface of the heterodimer
 - b) **Is formed by the hydrophobic interactions between amino acids in positions a and d**
 - c) Is formed by hydrogen bonds between leucines
 - d) Is formed by salt bridges between leucines
 - e) All of them are wrong
5. The main type/s of interactions that have been found between c-Jun/c-Fos and the DNA are:
 - a) Hydrogen bonds
 - b) Van der Waals
 - c) **both of above**
 - d) Salt bridges
 - e) All of them

PEM QUESTIONS

6. Which is the secondary structure of a zinc finger?

- a) An alpha helix
- b) An antiparallel beta sheet
- c) **Both of above are correct**
- d) Two alpha helices and one antiparallel beta sheet
- e) All of them are correct

7. Which residue/s are implicated in the zinc ion recognition?

- a) 1 cysteine and 2 histidines
- b) **2 histidines and 2 cysteines**
- c) Phenylalanine
- d) Cysteine and arginine
- e) None of them is correct

8. The binding domain recognize the...

- a) **Major groove of the DNA**
- b) Minor groove of the DNA
- c) Both of above are correct
- d) An specific region of minor groove of the DNA
- e) All of them are correct

9. Which of the following statements are correct for Egr1:

- a) Belongs to the C2H2 zinc fingers family
- b) Has 3 zinc fingers
- c) Fits into the major groove
- d) Each zinc finger interacts with a zinc ion
- e) **All of them are correct**

10. Which of them are superclasses of transcription factors:

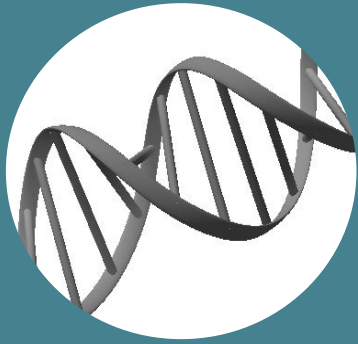
- a) Basic domains
- b) Helix turn Helix
- c) Zinc coordinating DNA binding
- d) Beta-scaffold factors with minor groove contacts
- e) **All of them are correct**

REFERENCES

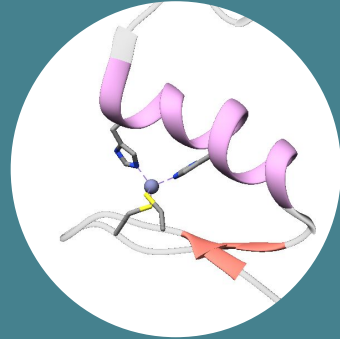
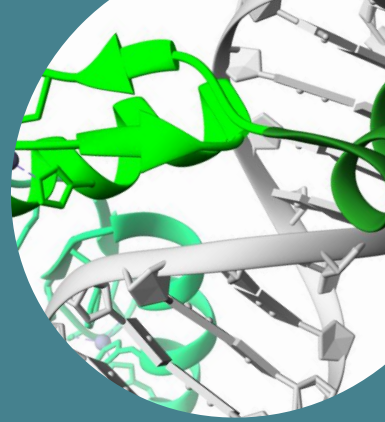
- [1] Andreucci JJ, Grant D, Cox DM, Tomc LK, Prywes R, Goldhamer DJ, Rodrigues N, Bédard PA, et al. Composition and function of AP-1 transcription complexes during muscle cell differentiation. *J Biol Chem*. 2002; 277 (19): 16426-32.
- [2] Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer L. The many faces of the helix-turn-helix domain: Transcription regulation and beyond. *FEMS. Microbiol Rev*. 2005;29(2):231-262.
- [3] Burmeister SS, Fernald RD. Evolutionary Conservation of the Egr-1 Immediate-Early Gene Response in a Teleost. *J Comp Neurol*. 2005;481(2):220-32.
- [4] Cruzalegui, F. c-Jun functions as a calcium-regulated transcriptional activator in the absence of JNK/SAPK1 activation. *EMBO J*. 1999; 18(5), pp.1335-1344.
- [5] Elrod-Erickson M, Rould MA, Nekludova L, Pabo CO. Structural impact of complete CpG methylation within target DNA on specific complex formation of the inducible transcription factor Egr-1. *Structure*. 1996;4(10):1171-80.
- [6] EMBL-EBI I. Helix-turn-helix motif (IPR000047) < InterPro EMBL-EBI [Internet]. Ebi.ac.uk. 2019 [cited 25 February 2019]. Available from: <https://www.ebi.ac.uk/interpro/entry/IPR000047>
- [7] Eyries M, Collins T, Khachigian L. Modulation of Growth Factor Gene Expression in Vascular Cells by Oxidative Stress. *Endothelium*. 2004;11(2):133-139.
- [8] Ferguson HA, Goodrich JA. Expression and purification of recombinant human c-Fos/c-Jun that is highly active in DNA binding and transcriptional activation in vitro. *Nucleic Acid Res*. 2001; 29(20): 1-6.
- [9] Glover JN, Harrison SC. Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature*. 1995; 373(6511): 257-61.
- [10] Gustems M, Woellmer A, Rothbauer U, Eck SH, Wieland T, Lutter D, Hammerschmidt W. c-Jun/c-Fos heterodimers regulate cellular genes via a newly identified class of methylated DNA sequence motifs. *Nucleic Acid Res*. 2014; 42(5): 3059-72.
- [11] Hashimoto H, Olanrewaju OY, Zheng Y, Wilson GG, Zhang X, Cheng X. Wilms tumor protein recognizes 5-carboxylcytosine within a specific DNA sequence. *Genes Dev*. 2014; 28(20): 2304–2313.
- [12] Latchman DS. Eukaryotic transcription factors. 4th. ed. London: Elsevier, 2004.

REFERENCES

- [13] Pavletich NP, Pabo CO. Zinc Finger-DNA Recognition: Crystal Structure of a Zif268-DNA Complex at 2.1 Å. *Science*. 1991;252(5007):809-817.
- [14] Pulverer, B., Kyriakis, J., Avruch, J., Nikolakaki, E. and Woodgett, J. Phosphorylation of c-jun mediated by MAP kinases. *Nature*. 1991; 353(6345), pp.670-674.
- [15] Razin SV, Borunova VV, Maksimenko OG, Kantidze OL. Cys2His2 Zinc Finger Protein Family: Classification, Functions, and Major Members. *Biochemistry (Mosc)*. 2012;77(3):217-26.
- [16] Ryu T, Jung J, Lee S, Nam HJ, Hong SW, Yoo JW, Lee DK, Lee D. bZIPDB : A database of regulatory information for human bZIP transcription factors. *BMC Genomics*. 2007; 8: 130-136.
- [17] Sekiya T, Kato K, Kawaguchi A, Nagata K. Involvement of CTCF in transcription regulation of EGR1 at early G1 phase as an architecture factor. *Sci Rep*. 2019;9(1):329.
- [18] Yan S, Fujita T, Lu J, Okada K, Shan Zou Y, Mackman N et al. Egr-1, a master switch coordinating upregulation of divergent gene families underlying ischemic stress. *Nat Med*. 2000; 6(12):1355-1361.
- [19] Yura K, Tomoda S, Go M. Repeat of a helix–turn–helix module in DNA-binding proteins. *Protein Eng*. 1993;6(6):621-628.
- [20] Zandarashvili L, Esadze A, Vuzman D, Kemme C, Levy Y, Iwahara J. Balancing between affinity and speed in target DNA search by zinc-finger proteins via modulation of dynamic conformational ensemble. *Proc Natl Acad Sci*. 2015;112(37):5142-5149.
- [21] Zandarashvili L, White M, Esadze A, Iwahara J. Structural impact of complete CpG methylation within target DNA on specific complex formation of the inducible transcription factor Egr-1. *FEBS Lett*. 2015;589(15):1748-1753.
- [22] Li, L., Feng, Z. and Porter, A. JNK-dependent Phosphorylation of c-Jun on Serine 63 Mediates Nitric Oxide-induced Apoptosis of Neuroblastoma Cells. *J Biol Chem*. 2003; 279(6), pp.4058-4065.



**THANK YOU FOR
YOUR ATTENTION!**



Maddalen Arizkorreta, Marta Figueras, Gal·la Garcia, Eva Martínez, Mar Prieto

ANNEX

EGR1

EVOLUTION AMONG DIFFERENT SPECIES - SEQUENCE ALIGNMENT (HMM)

CLUSTAL W(1.60) Multiple sequence alignment

```
Egr1_Mouse      - -maaaakaemqlnslqlisdpgfsfphsptndnypkleemllnsngapqflgaagtpgs
Egr1_Human      maaakaemqlnslqlisdpgfsfphsptndnypkleemllnsngapqflgaagtpgs
Egr1_Zebrafish  -----maaktelllpalqlisdplsfphsptndnykle
Egr1_Chicken    -----maakaemqlnslqlisdpgfvgfphsppamd
Egr1_Bovine      -maakaemqlnslqlisdpgfsfphsptndnypkleemllnsngapqflgaagtpgs
Egr1_Xenopus    -----navaktenlvslqlisdpgfsfphsptm
Egr1_Catfish    -----
```

```
Egr1_Mouse      ggnssstssggggggsgnssafnpgqepseqpyehltesfsdialnnekamvetsy
Egr1_Human      nssssssggggggggsgnssstfnpqadteqpyehltaesfpdislnneklvets
Egr1_Zebrafish  emlnsagatpflnatageavfsggeqefdhlagdtlseimekplsdqstytqrlp
Egr1_Chicken    hypkleemllnsgggqflapsagapesagfagaapeqsehfelaadtfeislnektm
Egr1_Bovine      ssgsgggggggggsgnssstfnpgaeaseqpyehltaesfpdislnneklvet
Egr1_Xenopus    dnypkleemllnagggpflgasvpdgsfnstvegaeqfhdltadafsemslssekav
Egr1_Catfish    -----
```

```
Egr1_Mouse      psqtrtlppitytgrfsllepapngntlwpeflfslvslvsmtnpptssssapass
Egr1_Human      ypsqtrtlppitytgrfsllepapngntlwpeflfslvslvsmtnppasssapass
Egr1_Zebrafish  psltytgrfsllepapngntlwpeflfslvslvsmtnppasssapass
Egr1_Chicken    psltytgrfsllepapngntlwpeflfslvslvsmtnppasssapass
Egr1_Bovine      sypsqtrtlppitytgrfsllepapngntlwpeflfslvslvsmtnppasssapass
Egr1_Xenopus    essyanqtrtlpsltytgrfsllepapngntlwpeflfslvslvsmtnppasssapass
Egr1_Catfish    -----lwaeplfnlsvglvgvnhststssatqtpssspastssssss
```

```
Egr1_Mouse      sssasqspplscavpsndsspiysaaptfptntdfpepgsqafpgsagtalqypypay
Egr1_Human      sasasqspplscavpsndsspiysaaptfptntdfpepgsqafpgsagtalqypypay
Egr1_Zebrafish  pssssstssaslscsvhsepnpiysaaptfptntdfpepgsqafpgsagtalqypypay
Egr1_Chicken    pssssstssaslscsvhsepnpiysaaptfptntdfpepgsqafpgsagtalqypypay
Egr1_Bovine      sssasqspplscavpsndsspiysaaptfptntdfpepgsqafpgsagtalqypypay
Egr1_Xenopus    sssssstssaslscsvhsepnpiysaaptfptntdfpepgsqafpgsagtalqypypay
Egr1_Catfish    sssssstssstsspslscsvhsepnpiysaaptfptntdfpepgsqafpgsagtalqypypay
```

```
Egr1_Mouse      patkggfvqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafatqsg
Egr1_Human      paakggfvqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafatqsg
Egr1_Zebrafish  tpsaktcnpsfsvmpmpdylftqqqseislvppdqkplqtqagqpaltplhtikafat
Egr1_Chicken    ptaktfnqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafatqsg
Egr1_Bovine      pgakggfvqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafatqsg
Egr1_Xenopus    ypvskttfvqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafatqsg
Egr1_Catfish    pppaypskacgasfvqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafat
```

```
Egr1_Mouse      sqdalkalnttyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Human      sqdalkalnttyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Zebrafish  qtsqsdqlksvyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Chicken    sqelkltlnttyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Bovine      sqdalkalnttyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Xenopus    sqdalkalnttyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Catfish    qtsqsdqlksvyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
```

```
Egr1_Mouse      patkggfvqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafatqsg
Egr1_Human      paakggfvqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafatqsg
Egr1_Zebrafish  tpsaktcnpsfsvmpmpdylftqqqseislvppdqkplqtqagqpaltplhtikafat
Egr1_Chicken    ptaktfnqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafatqsg
Egr1_Bovine      pgakggfvqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafatqsg
Egr1_Xenopus    ypvskttfvqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafatqsg
Egr1_Catfish    pppaypskacgasfvqvpmpdylfpqggdglgltpdqkpfqglennrtqqpsltplstikafat
```

```
Egr1_Mouse      sqdalkalnttyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Human      sqdalkalnttyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Zebrafish  qtsqsdqlksvyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Chicken    sqelkltlnttyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Bovine      sqdalkalnttyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Xenopus    sqdalkalnttyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
Egr1_Catfish    qtsqsdqlksvyqsqlikpsrmrkyknprsktppherpYACPVeSCDRRFSRDELTRHIR
```

```
Egr1_Mouse      IHTgqkpfqcrimrnfrsdlhtthirthtgekpfacdicgrkfarsderkrhtkhlr
Egr1_Human      IHTgqkpfqcrimrnfrsdlhtthirthtgekpfacdicgrkfarsderkrhtkhlr
Egr1_Zebrafish  IHTgqkpfqcrimrnfrsdlhtthirthtgekpfaceicgrkfarsderkrhtkhlr
Egr1_Chicken    IHTgqkpfqcrimrnfrsdlhtthirthtgekpfacdicgrkfarsderkrhtkhlr
Egr1_Bovine      IHTgqkpfqcrimrnfrsdlhtthirthtgekpfacdicgrkfarsderkrhtkhlr
Egr1_Xenopus    IHTgqkpfqcrimrnfrsdlhtthirthtgekpfacdicgrkfarsderkrhtkhlr
Egr1_Catfish    IHTgqkpfqcrimrnfrsdlhtthirthtgekpfacdicgrkfarsderkrhtkhlr
```

```
Egr1_Mouse      qdkkdkadksvvaspaasslssypspvatsypspattsfpsvpvtsyspsgsstypspahs
Egr1_Human      qdkkdkadksvvaspaasslssypspvatsypspattsfpsvpvtsyspsgsstypspahs
Egr1_Zebrafish  qdkkdkadksvvaspaasslssypspvatsypspattsfpsvpvtsyspsgsstypspahs
Egr1_Chicken    qdkkdkadksvvaspaasslssypspvatsypspattsfpsvpvtsyspsgsstypspahs
Egr1_Bovine      qdkkdkadksvvaspaasslssypspvatsypspattsfpsvpvtsyspsgsstypspahs
Egr1_Xenopus    qdkkdkadksvvaspaasslssypspvatsypspattsfpsvpvtsyspsgsstypspahs
Egr1_Catfish    qdkkdkadksvvaspaasslssypspvatsypspattsfpsvpvtsyspsgsstypspahs
```

```
Egr1_Mouse      gfpspsvattfasvpvpaftqvsssfpsagvssfststglsdmtatfsprtleic----
Egr1_Human      tpsvpvhsqfppspvatvtyssvpvpaqavsssfpsavtnsfasstglsdmtatfsprtleic
Egr1_Zebrafish  tpsvpvhsqfppspvatvtyssvpvpaqavsssfpsavtnsfasstglsdmtatfsprtleic
Egr1_Chicken    tpsvpvhsqfppspvatvtyssvpvpaqavsssfpsavtnsfasstglsdmtatfsprtleic
Egr1_Bovine      tpsvpvhsqfppspvatvtyssvpvpaqavsssfpsavtnsfasstglsdmtatfsprtleic
Egr1_Xenopus    tpsvpvhsqfppspvatvtyssvpvpaqavsssfpsavtnsfasstglsdmtatfsprtleic
Egr1_Catfish    tpsvpvhsqfppspvatvtyssvpvpaqavsssfpsavtnsfasstglsdmtatfsprtleic
```

```
Egr1_Mouse      ---
Egr1_Human      etc
Egr1_Zebrafish  ---
Egr1_Chicken    ---
Egr1_Bovine      ---
Egr1_Xenopus    ---
Egr1_Catfish    ---
```