

TFEB: An important transcription factor

重要转录因子**TFEB**的生物信息学分析

Group members:

陈杰, 高凯瑜, 聂超, 张婧

Outline:

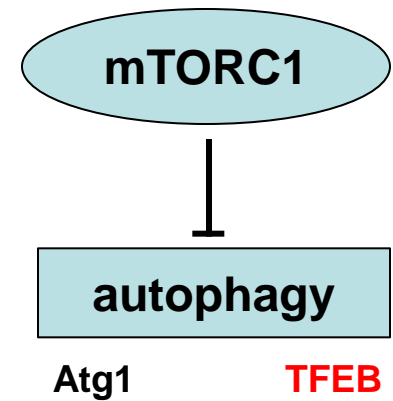
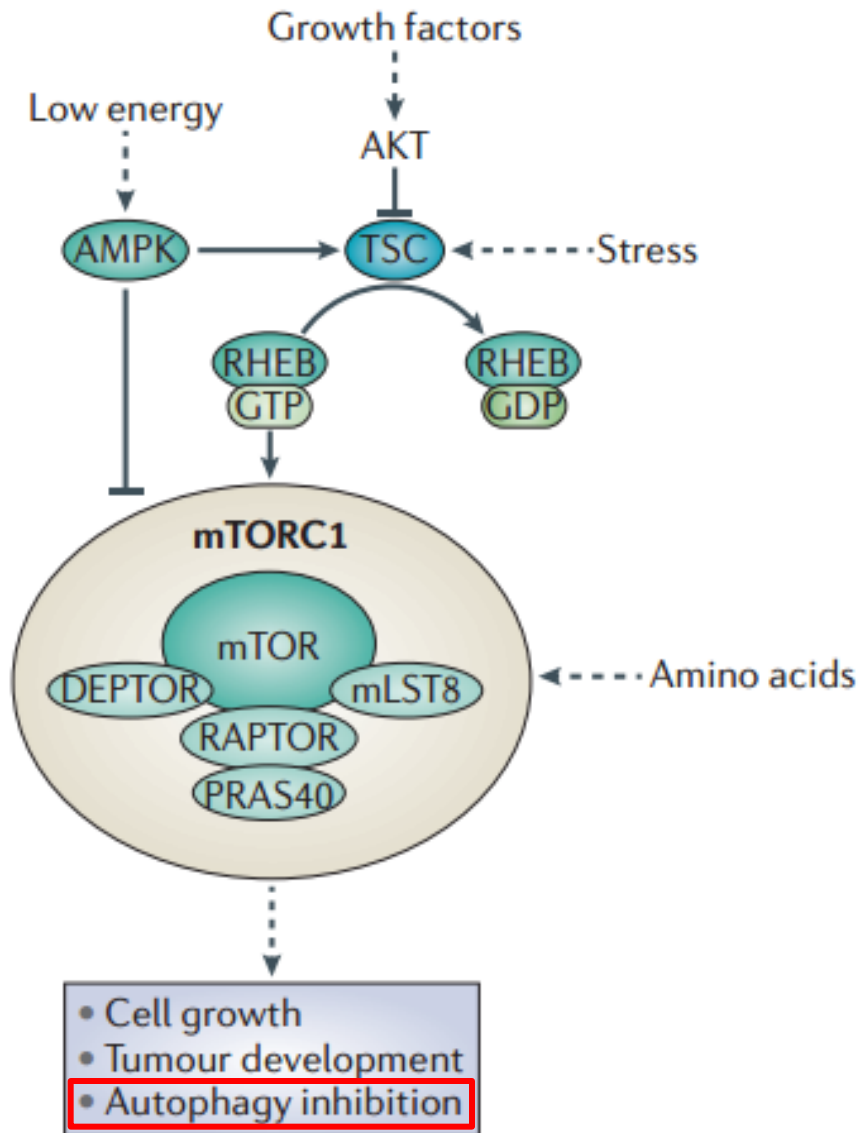
Background

Alignment and construction of gene tree

physicochemical properties

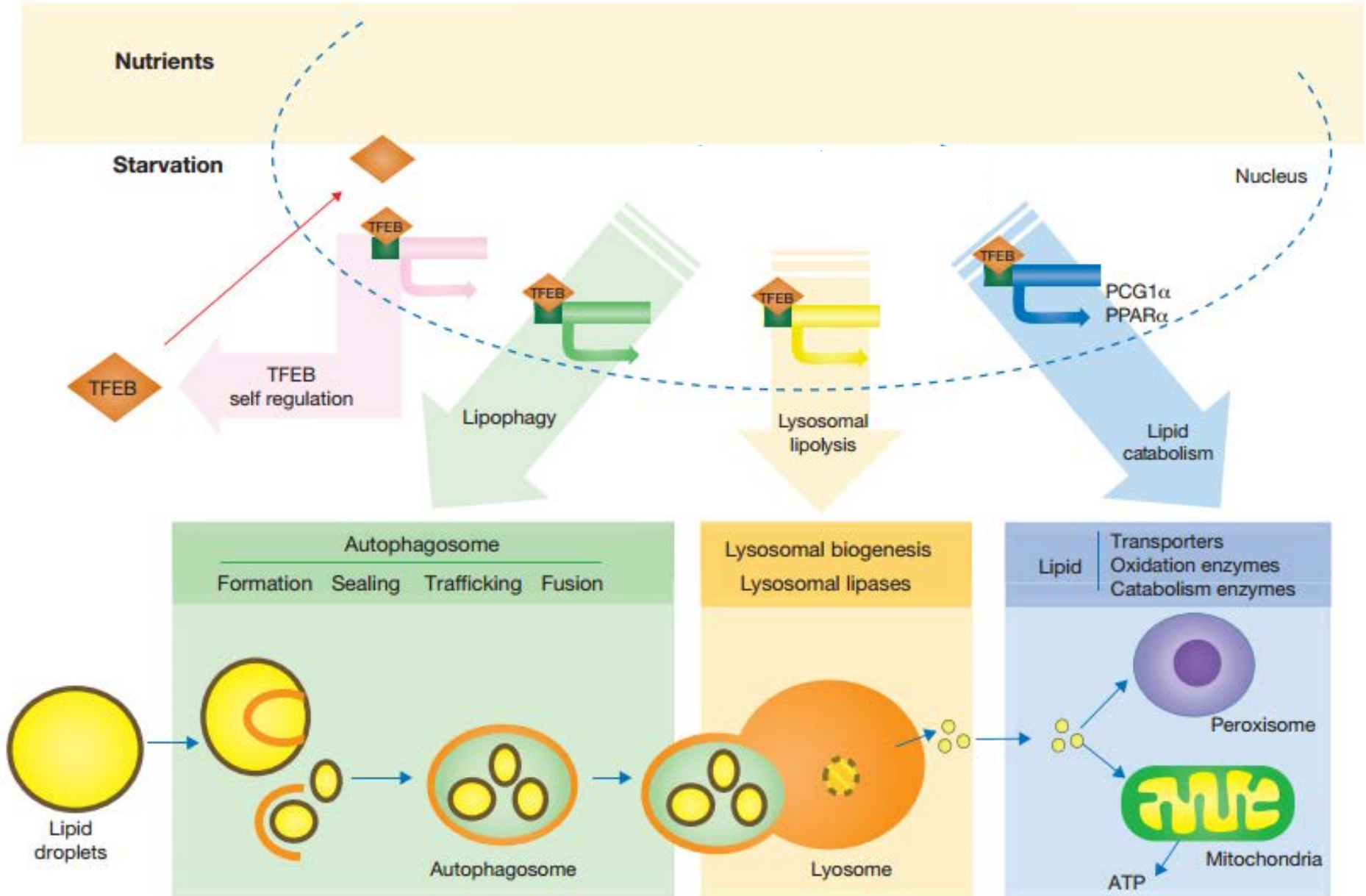
Homologous modeling

1. Background



1. Background

Time course of TFEB activating of target genes



Outline:

Background

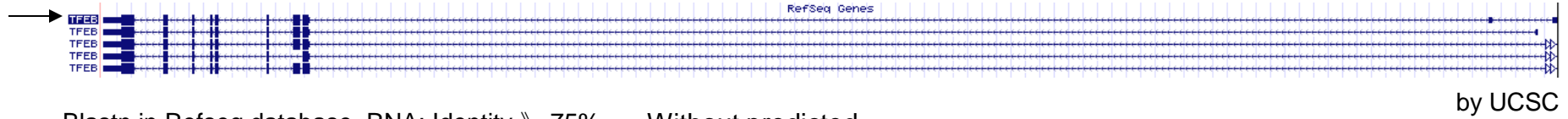
Alignment and construction of gene tree

physicochemical properties

Homologous modeling

2.Alignment and construction of gene tree

mRNA:



Species	Refseq Accession	Identity
Homo sapiens	NM_007162.2	100%
Macaca mulatta	NM_001266320.1	98%
Papio anubis	NM_001168746.1	98%
Oryctolagus cuniculus	NM_001171339.1	91%
Bos taurus	NM_001205666.1	88%
Gallus gallus	NM_001030922.1	81%
Mus musculus	NM_011549.3	80%
Rattus norvegicus	NM_001025707.1	77%



Macaca mulatta



Papio anubis



Oryctolagus cuniculus



Bos taurus

2.Alignment and construction of gene tree

Protein:

NCBI Reference Sequence: NP_009093.1 (Homo)

Blastp in Refseq database_Protein: Identity » 75% Without predicted

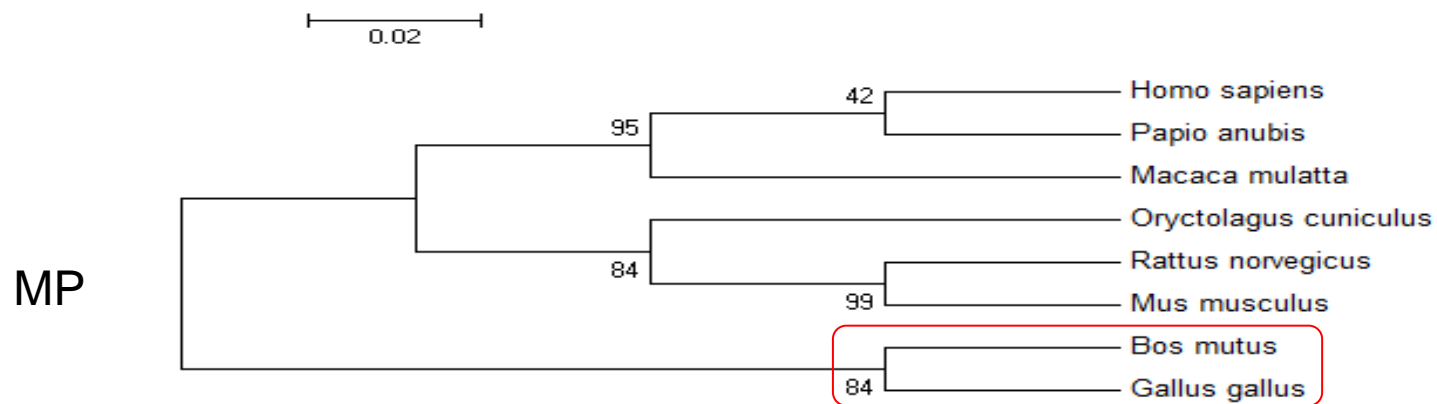
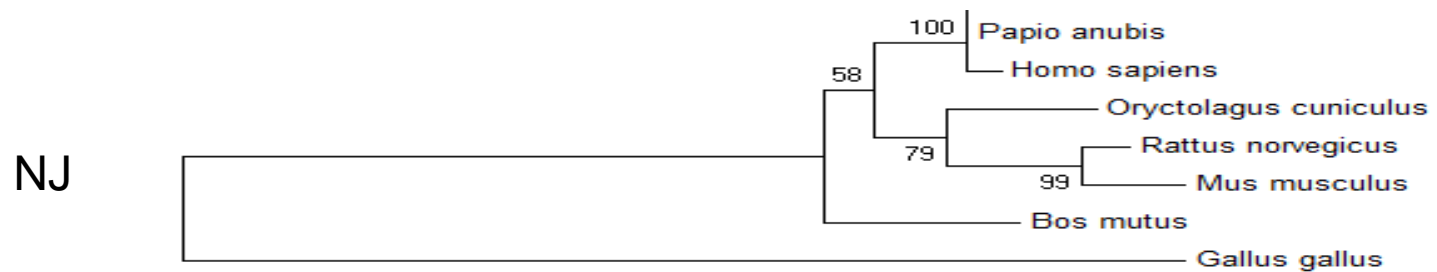
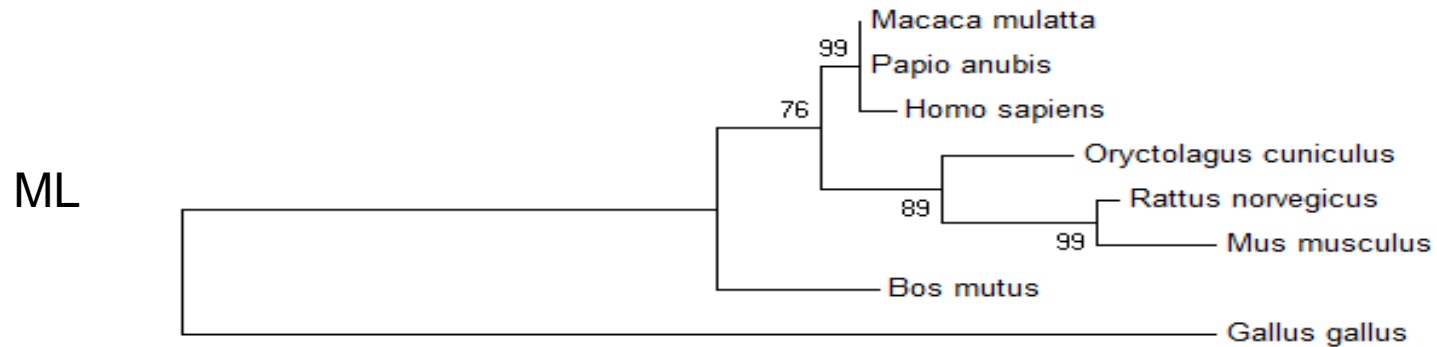
	Species	Refseq Accession	Identity
mammal	Homo sapiens	NP_009093.1	100%
	Macaca mulatta	NP_001253249.1	99%
	Papio anubis	NP_001162217.1	99%
	Oryctolagus cuniculus	NP_001164810.1	94%
	Rattus norvegicus	NP_001020878.1	94%
	Mus musculus	NP_035679.3	93%
	Bos taurus	NP_001192595.1	93%
	Gallus gallus	NP_001026093.1	76%

Protein Identity > RNA Identity for mammalian but not chicken. More Reasonalbe

Alignment and construction of gene tree

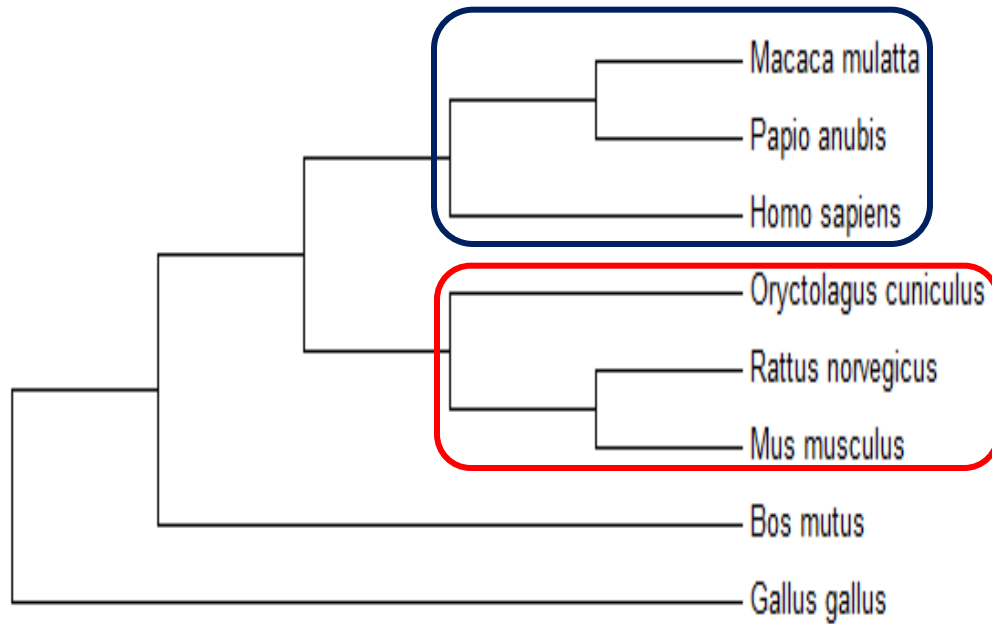
Phylogenetic tree

Common sense: Variation: $ML > NJ > MP$

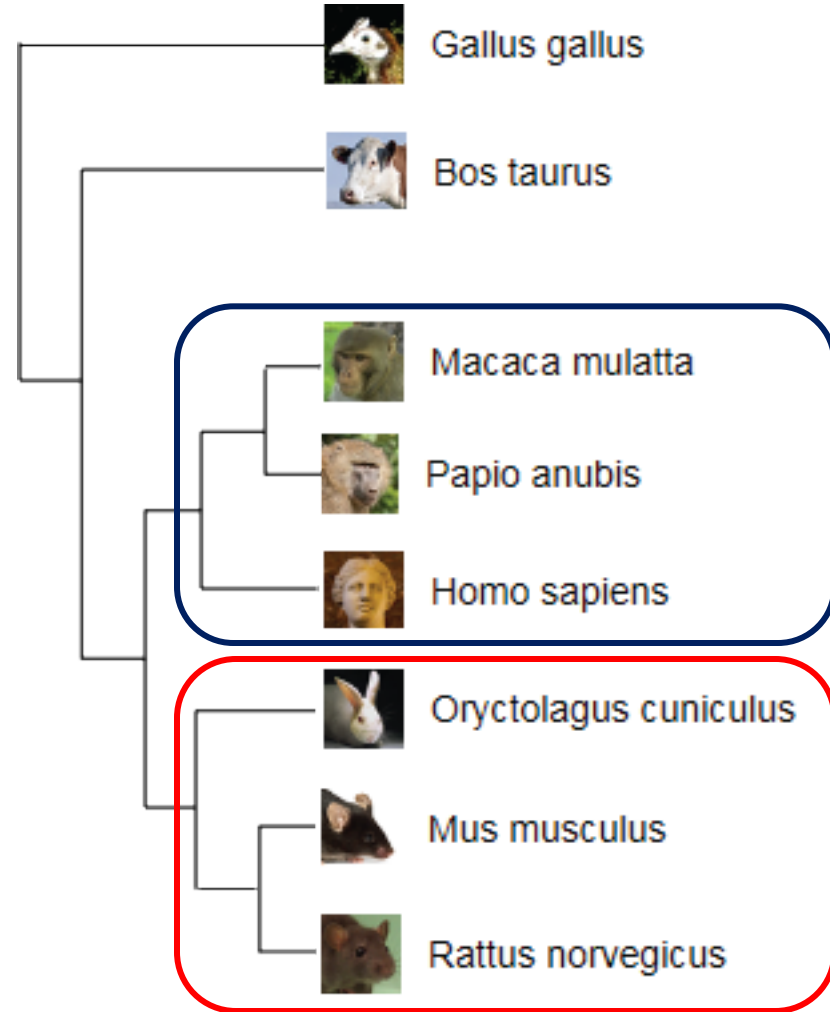


2. Alignment and construction of gene tree

Ensembl **species tree** vs **TFEB tree**



By MEGA



By Ensembl

Outline:

Background

Alignment and construction of gene tree

physicochemical properties

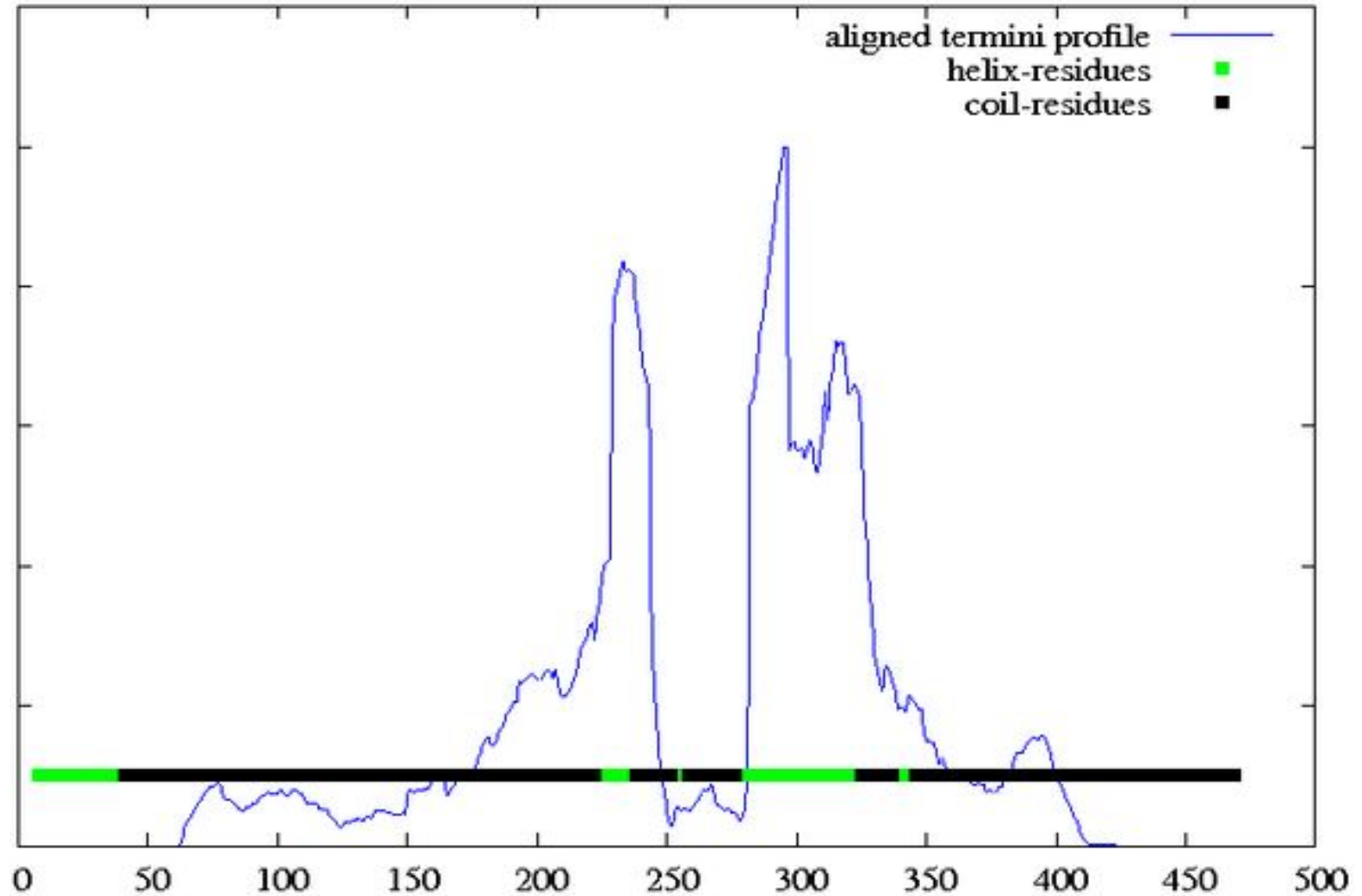
DNA binding domain predication

NLS predication

Homologous modeling

physicochemical properties

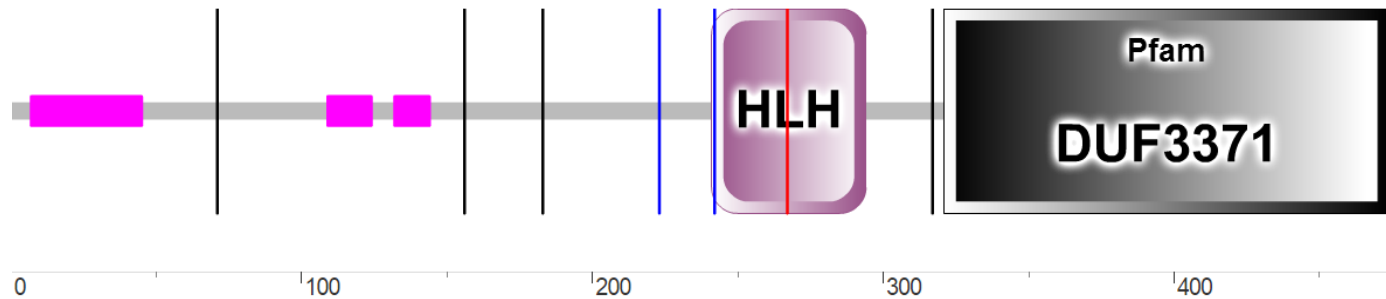
DNA binding domain predication



predicated by **Psipred**

physicochemical properties

DNA binding domain predication



Confidently predicted domains, repeats, motifs and features:




Name	Start ▲	End	E-value
low complexity	7	44	N/A
low complexity	109	123	N/A
low complexity	132	143	N/A
HLH	241	294	1.44e-15
Pfam:DUF3371	321	474	1.5e-38

http://smart.embl-heidelberg.de/smart/show_motifs.pl




physicochemical properties

DNA binding domain predication

Domains and Repeats

Feature key	Position(s)	Length	Description	Graphical view
Domain ⁱ	235 - 288	54	bHLH  PROSITE-ProRule annotation 	

Region

Feature key	Position(s)	Length	Description	Graphical view
Region ⁱ	156 - 165	10	Strong transcription activation domain  Sequence Analysis	
Region ⁱ	298 - 319	22	Leucine-zipper	

<http://www.uniprot.org/uniprot/P19484>

2.Alignment and construction of gene tree

Alignment:

Species	Refseq Accession	Identity
Homo sapiens	NP_009093.1	100%
Macaca mulatta	NP_001253249.1	99%
Papio anubis	NP_001162217.1	99%
Oryctolagus cuniculus	NP_001164810.1	94%
Rattus norvegicus	NP_001020878.1	94%
Mus musculus	NP_035679.3	93%
Bos taurus	NP_001192595.1	93%
Gallus gallus	NP_001026093.1	76%
Drosophila melanogaster	NP_001033808.1	57%
Caenorhabditis elegans	NP_500462.1	54%

2.Alignment and construction of gene tree

Species/Abbrv	Gr		*	***	**	*****	*****	*****	*****	*****	**		*	*****	*	***																																														
1. Homo sapiens		A	K	E	R	Q	K	K	D	N	H	N	L	I	E	R	R	R	R	F	I	N	D	R	I	K	E	L	G	M	L	I	P	K	A	N	D	-----	L	D	V	R	W	N	K	G	I	L	K	A	S	V	D	Y	I	R	R					
2. Macaca mulatta		A	K	E	R	Q	K	K	D	N	H	N	L	I	E	R	R	R	R	F	I	N	D	R	I	K	E	L	G	M	L	I	P	K	A	N	D	-----	L	D	V	R	W	N	K	G	I	L	K	A	S	V	D	Y	I	R	R					
3. Papio anubis		A	K	E	R	Q	K	K	D	N	H	N	L	I	E	R	R	R	R	F	I	N	D	R	I	K	E	L	G	M	L	I	P	K	A	N	D	-----	L	D	V	R	W	N	K	G	I	L	K	A	S	V	D	Y	I	R	R					
4. Oryctolagus cuniculus		A	K	E	R	Q	K	K	D	N	H	N	L	I	E	R	R	R	R	F	I	N	D	R	I	K	E	L	G	M	L	I	P	K	A	N	D	-----	L	D	V	R	W	N	K	G	I	L	K	A	S	V	D	Y	I	R	R					
5. Rattus norvegicus		A	K	E	R	Q	K	K	D	N	H	N	L	I	E	R	R	R	R	F	I	N	D	R	I	K	E	L	G	M	L	I	P	K	A	N	D	-----	L	D	V	R	W	N	K	G	I	L	K	A	S	V	D	Y	I	R	R					
6. Mus musculus		A	K	E	R	Q	K	K	D	N	H	N	L	I	E	R	R	R	R	F	I	N	D	R	I	K	E	L	G	M	L	I	P	K	A	N	D	-----	L	D	V	R	W	N	K	G	I	L	K	A	S	V	D	Y	I	R	R					
7. Bos mutus		A	K	E	R	Q	K	K	D	N	H	N	L	I	E	R	R	R	R	F	I	N	D	R	I	K	E	L	G	M	L	I	P	K	A	N	D	-----	L	D	V	R	W	N	K	G	I	L	K	A	S	V	D	Y	I	R	R					
8. Gallus gallus		A	K	E	R	Q	K	K	D	N	H	N	L	I	E	R	R	R	R	F	I	N	D	R	I	K	E	L	G	M	L	I	P	K	A	N	D	-----	L	D	V	R	W	N	K	G	I	L	K	A	S	V	D	Y	I	R	R					
9. Drosophila melanogaster		A	K	D	R	Q	K	K	D	N	H	N	M	I	E	R	R	R	R	F	I	N	D	R	I	K	E	L	G	T	L	L	P	K	G	S	D	A	F	Y	E	V	V	---	R	D	I	R	N	K	G	I	L	K	S	S	V	D	Y	I	K	C
10. Caenorhabditis elegans		Y	R	D	R	Q	K	K	D	I	H	N	M	I	E	R	R	R	R	F	I	N	D	R	I	K	E	L	G	M	L	P	K	N	T	S	E	D	M	K	L	N	K	G	I	L	K	A	S	C	L	Y	I	R	V							

By MEGA

Species/Abbrv	Gr		*		*	***		*																										
1. Homo sapiens		N	H	S	R	R	L	E	M	I	N	K	Q	L	W	L	R	I	Q	E	L	E	M	Q	A	R	V	H	G	L	P	I	T	S
2. Macaca mulatta		N	H	S	R	R	L	E	M	I	N	K	Q	L	W	L	R	I	Q	E	L	E	M	Q	A	R	V	H	G	L	P	I	T	S
3. Papio anubis		N	H	S	R	R	L	E	M	I	N	K	Q	L	W	L	R	I	Q	E	L	E	M	Q	A	R	V	H	G	L	P	I	T	S
4. Oryctolagus cuniculus		N	H	S	R	R	L	E	M	I	N	K	Q	L	W	L	R	I	Q	E	L	E	M	Q	A	R	V	H	G	L	P	I	T	S
5. Rattus norvegicus		N	H	S	R	R	L	E	M	I	N	K	Q	L	W	L	R	I	Q	E	L	E	M	Q	A	R	V	H	G	L	P	I	T	S
6. Mus musculus		N	H	S	R	R	L	E	M	I	N	K	Q	L	W	L	R	I	Q	E	L	E	M	Q	A	R	V	H	G	L	P	I	T	S
7. Bos mutus		N	H	S	R	R	L	E	M	I	N	K	Q	L	W	L	R	I	Q	E	L	E	M	Q	A	R	V	H	G	L	P	I	T	S
8. Gallus gallus		N	H	S	R	R	L	E	M	I	N	K	Q	L	L	L	R	I	Q	E	L	E	M	Q	A	R	V	H	G	L	P	I	T	S
9. Drosophila melanogaster		L	R	Q	R	Q	V	E	L	Q	N	R	K	L	M	S	R	I	K	E	L	E	M	Q	A	K	S	H	G	I	L	L	S	E
10. Caenorhabditis elegans		Q	Q	K	S	L	E	S	T	A	H	K	Y	A	D	R	V	K	E	L	E	E	M	L	A	R	G	V	V	B	E			

By MEGA

Uniprot:

Feature key	Position (s)	Length	Description
Domain ⁱ	235 – 288	54	bHLH
Region ⁱ	298 – 319	22	Leucine-zipper

Outline:

Background

Alignment and construction of gene tree

physicochemical properties

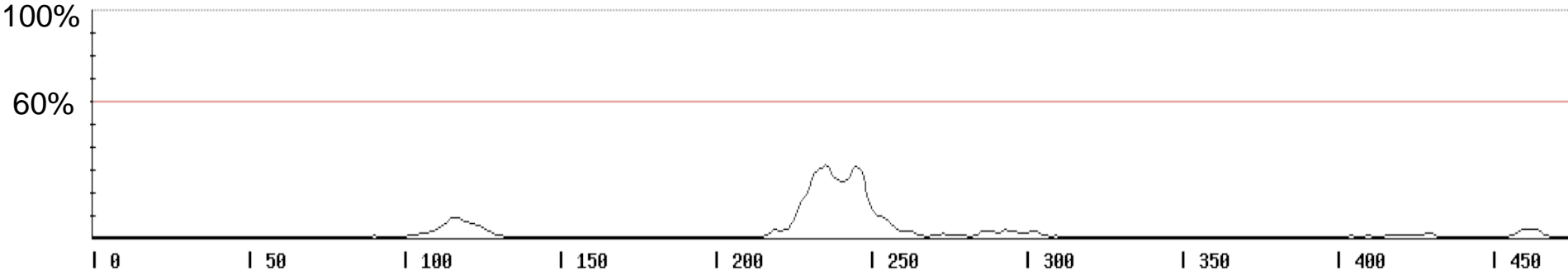
DNA binding domain predication

NLS predication

Homologous modeling

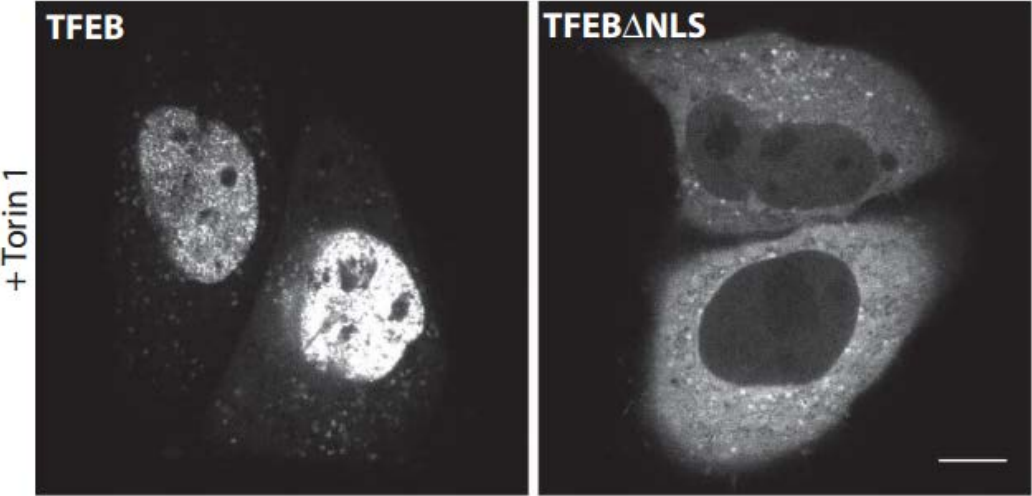
physicochemical properties

NLS predication



by NLStradamus

The true result from paper



Agnes Roczniak-Ferguson et.al. **Science**, 2012

Outline:

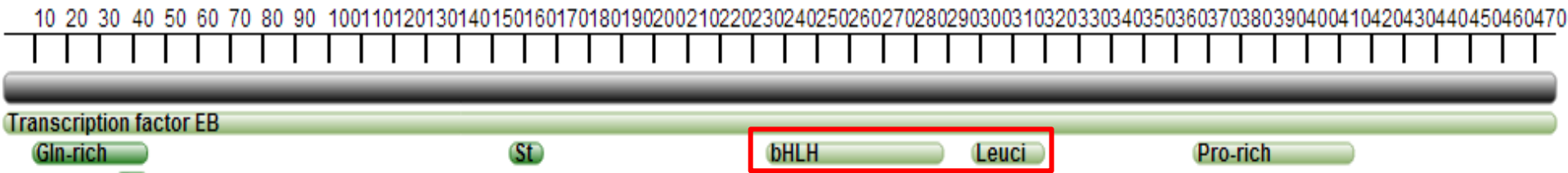
Background

Alignment and construction of gene tree

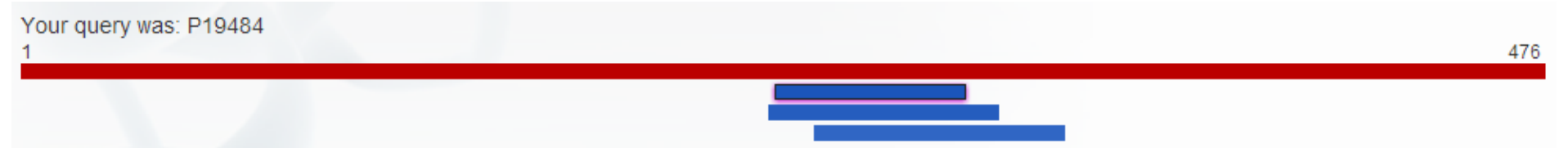
physicochemical properties

Homologous modeling

Homologous modeling



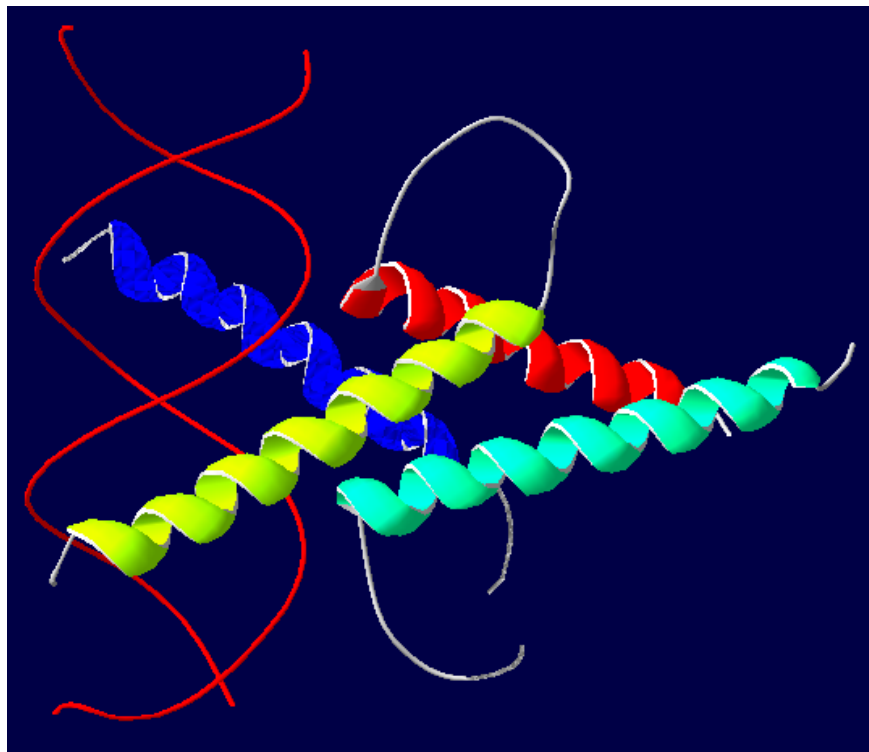
provide by **PDB**



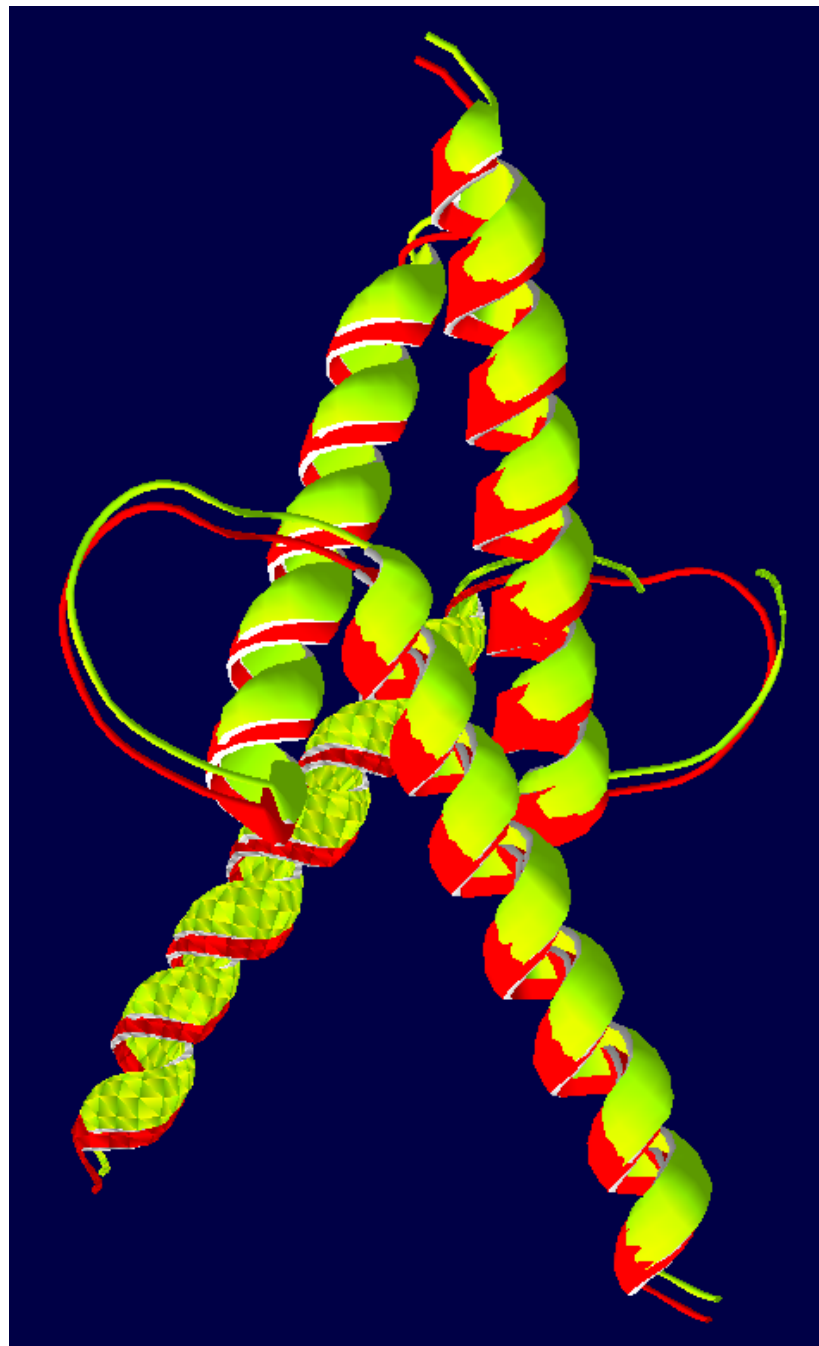
Predicated by **ProteinModelPortal**

Model	LEKKDNHNLI-ERRRRFNINDRIKELGMLIPKANDLDVRWNKGTILKASVDYIRRMQKDL QK
4atkB	--KKDNHNLIERRRRFNINDRIKELGTLIPKSNPDMRWKGTILKASVDYIRKLQREQ QR
id. residues	..KKDNHNLI..RRRRFNINDRIKELG.LIPK.ND.D.RWNKGTILKASVDYIR..Q... Q.

provided by **Swissmodel**



Template: 4atkB (provided by PDB)



TFEB fit with template

THX for your attention !