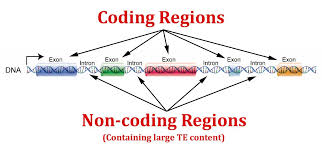
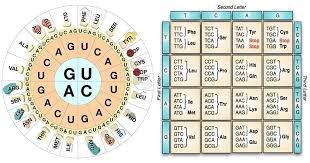
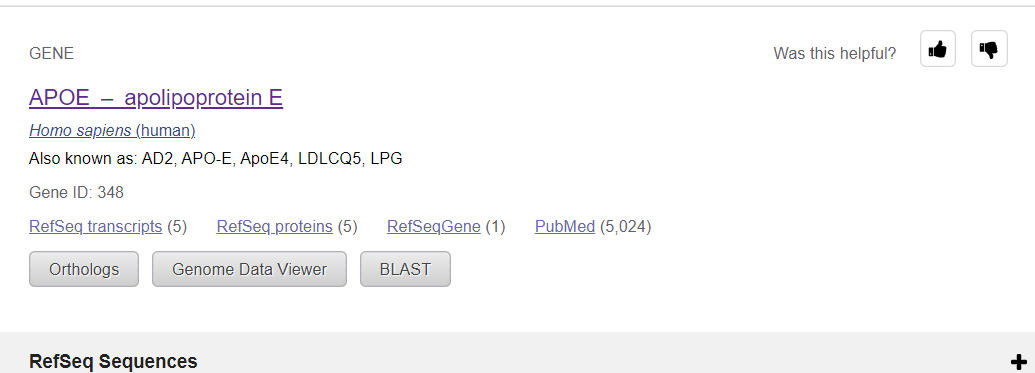
## Practical

Coding Region   
 comprising of exons and introns  
  
  
  
Genetic code –amino acids – proteins  
triplets of nucleotide – is formed in particular format only

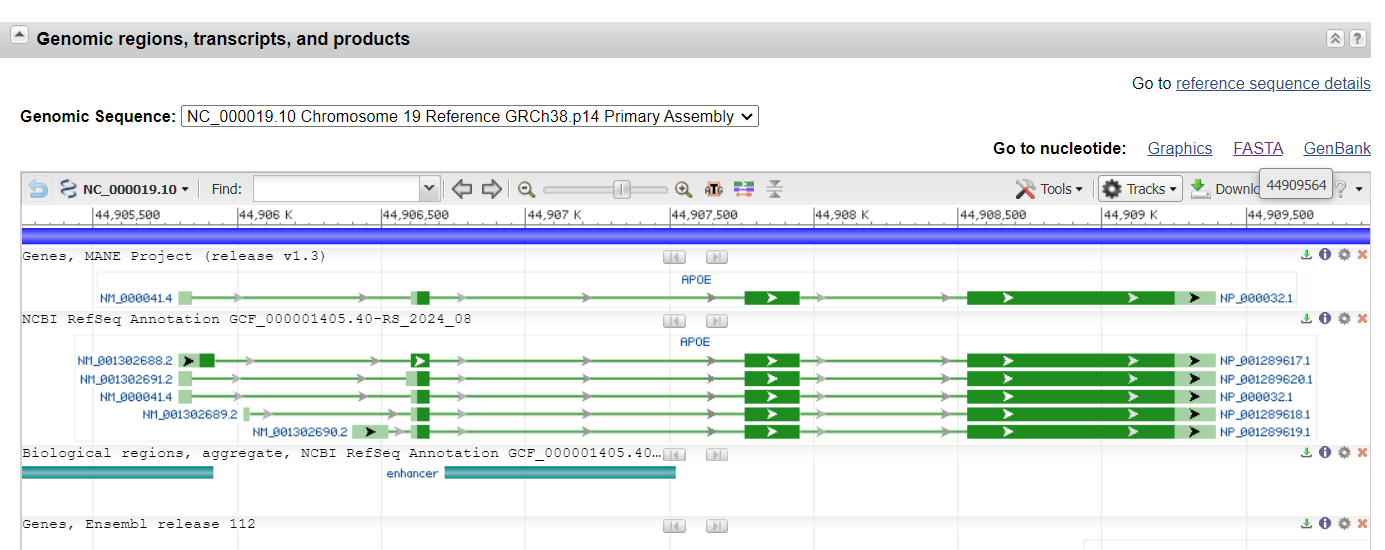
Start codon – dns seq starts

Stop codon – dna seq stops

ATGC – lang. of biology (DNA)

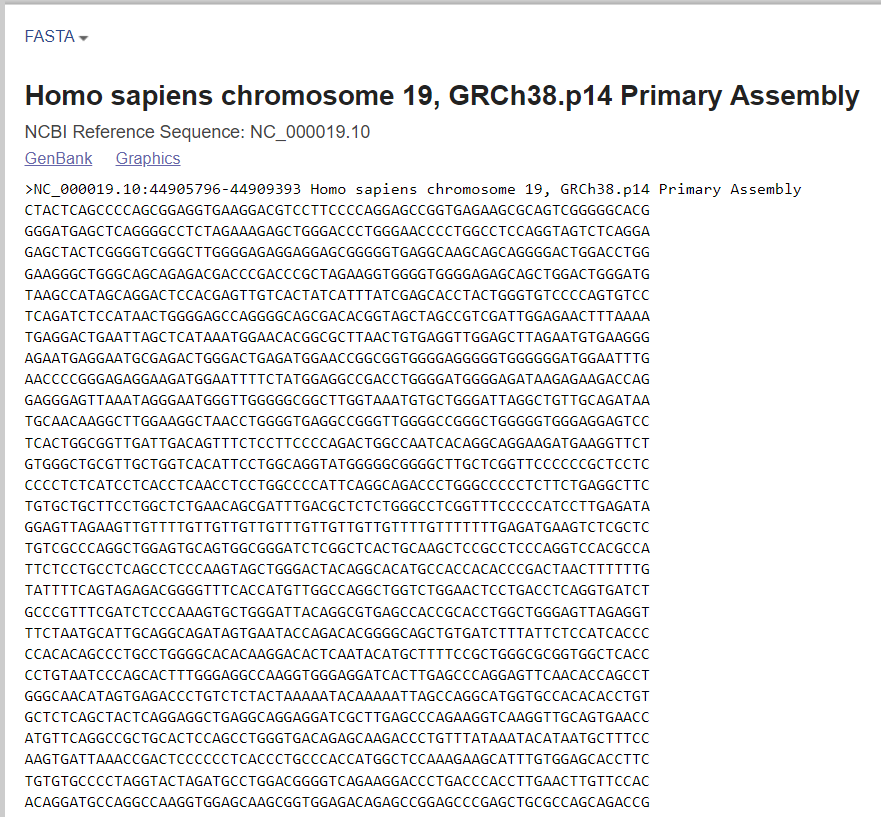
Gene (database) APOE – first link 

Select fasta :



Sequence starts from – CTA…

Header stores information about sequence – ‘>’ represents sequence



In 23 pairs of chromosome, this gene is located on position 44905796 – 44909393

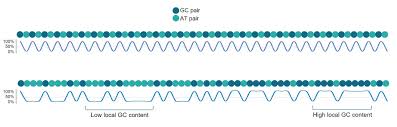
Percent of g and c in dna,rna – is called gc content

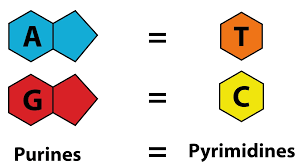
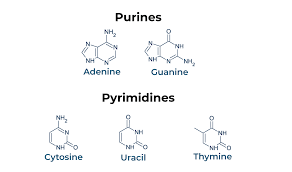
GC content – stability of DNA – More GC content- more stability of DNA – due to stronger hydrogen bind between g and c bases

DNA with higher gc content has higher melting temp, which is imp for laboratory techniques like PCR

Gene expression – GC content can influence gene expression - higher GC in promoter regions often associate with higher expression

Every gene has 2 copies in our body



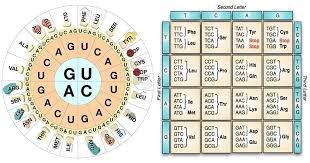


A and g – purines

C , t, uracils – pyrimides

Purine : pyrimides = AT/GC

Coding req – 1 amino acid req 3 nucleotide



Data frame has been formed using a codon (triplet) this is for amino acid

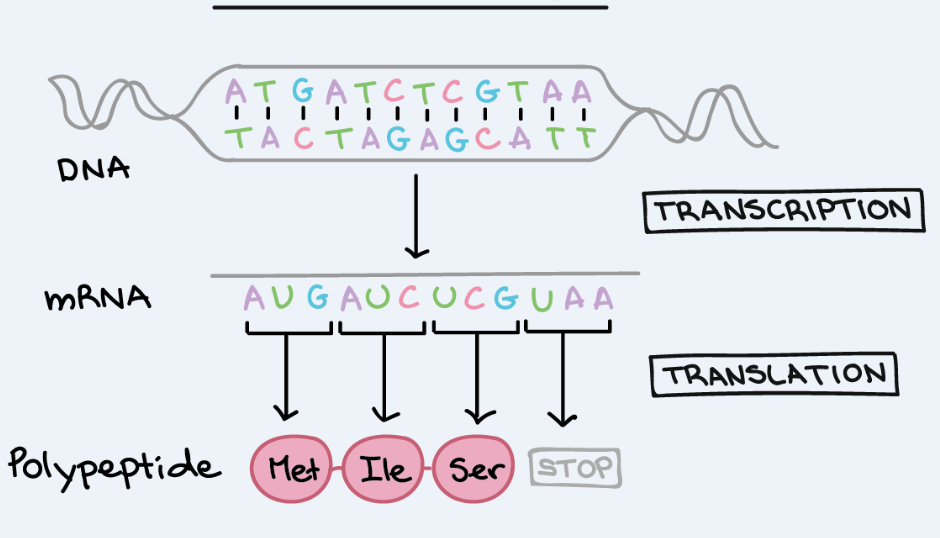
* Start codon

The point at which protein translation begins. The most common start codon is AUG. In eukaryotes, the start codon codes for methionine. In prokaryotes, it codes for a modified methionine.

* Stop codon

The point at which protein translation ends. The typical stop codons are UAG, UGA, and UAA.

https://ibiologia.com/codon/



Dna -> rna

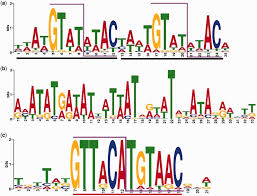
Nucleotide -> Codon -> amino acid -> polypeptide

Seq of coding reg from start to end of ebola

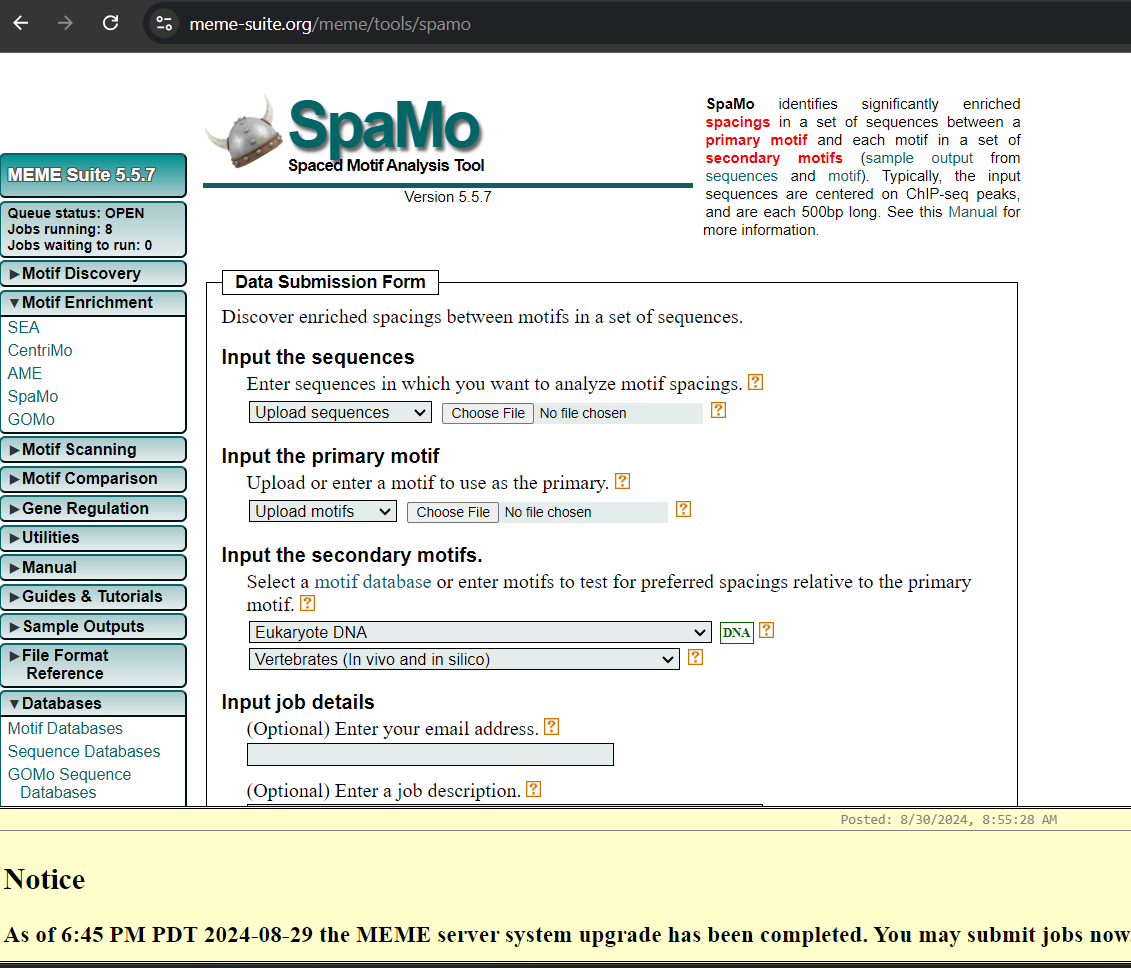
Day 2

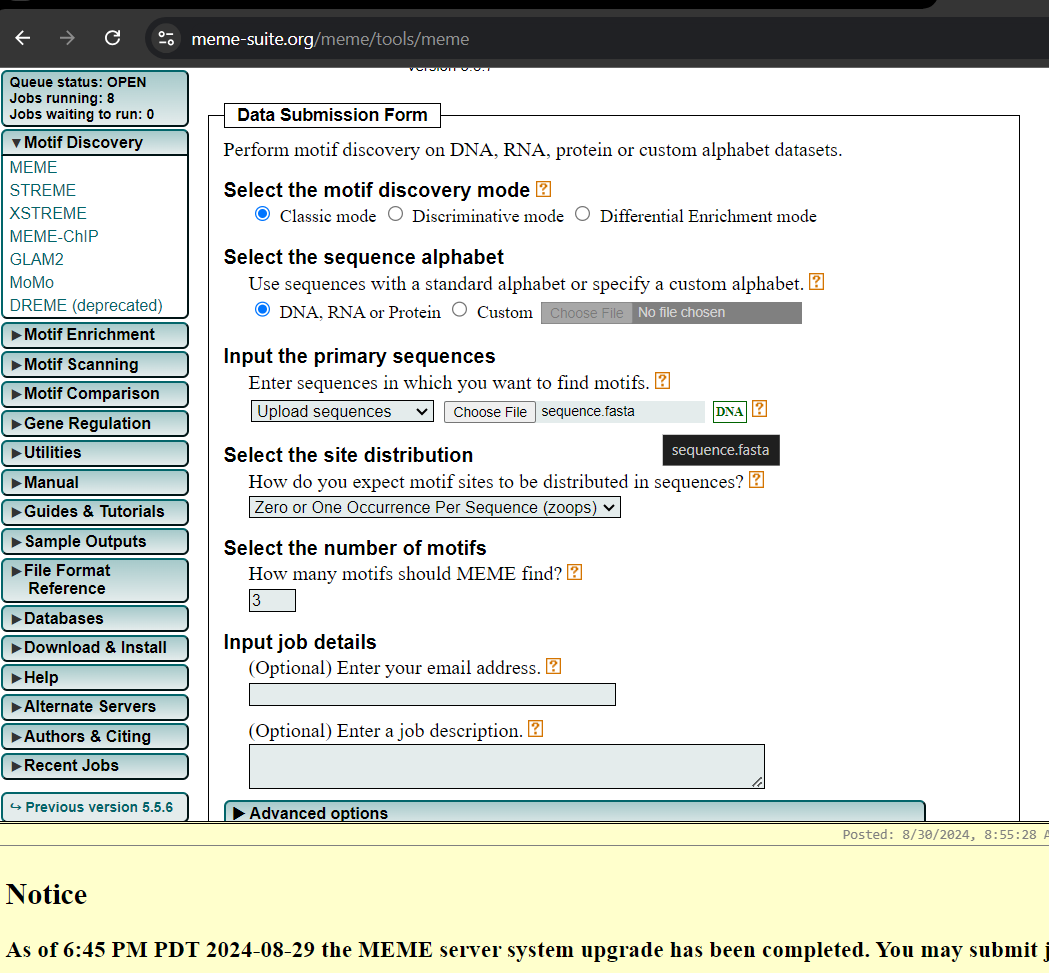
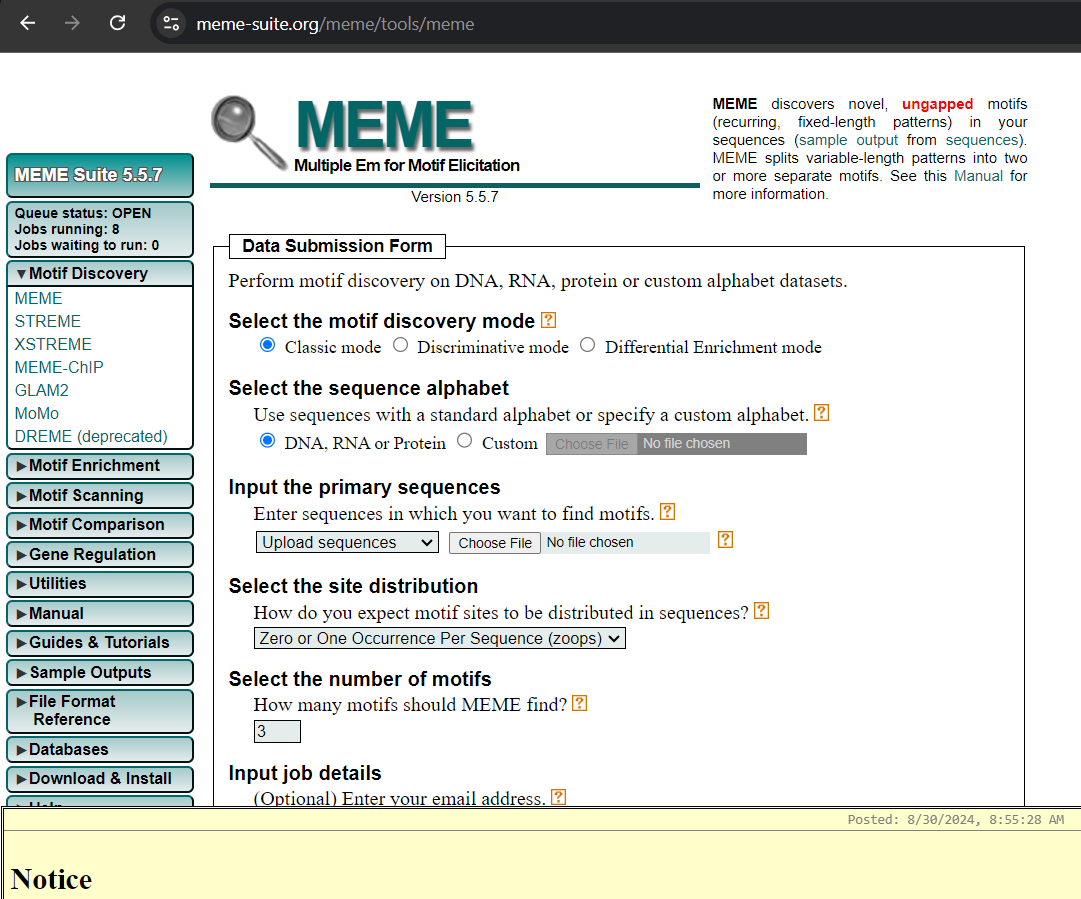
Out of 64 combinations of triplets of ATGC – 61 are amino acids

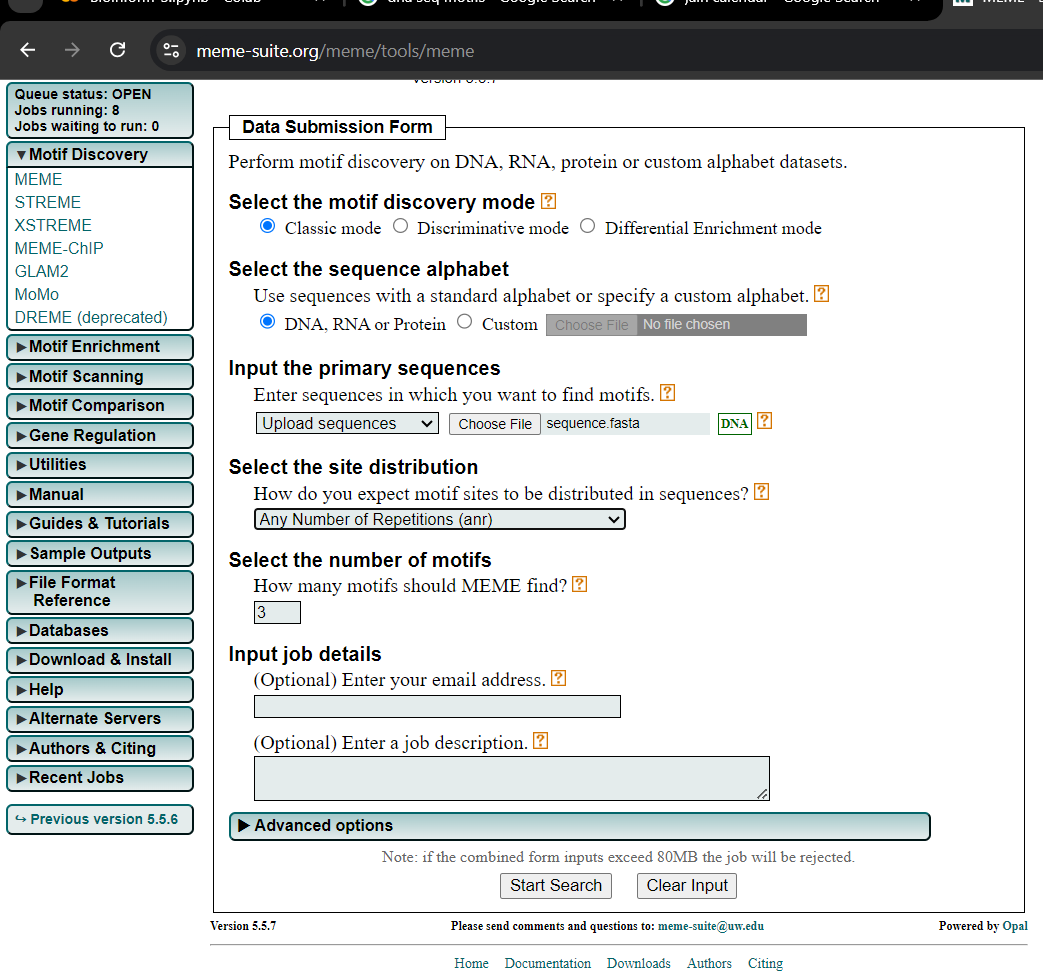
Dna seq motifs – repetitive reoccurring pattern in dna – in stretch on nucleotide



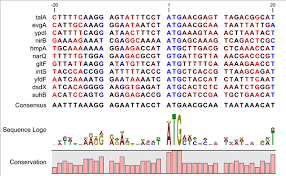
MEME tool –







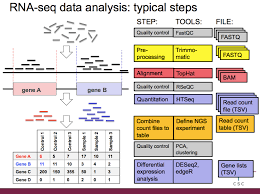
Motifs – conserved signatures



Write motifs code

RNA – seq (basically mRNA)

rna seq data process biological sample ->mrna transcript -> rna fragments -> dna fragments



Cancer detection – mutations finding

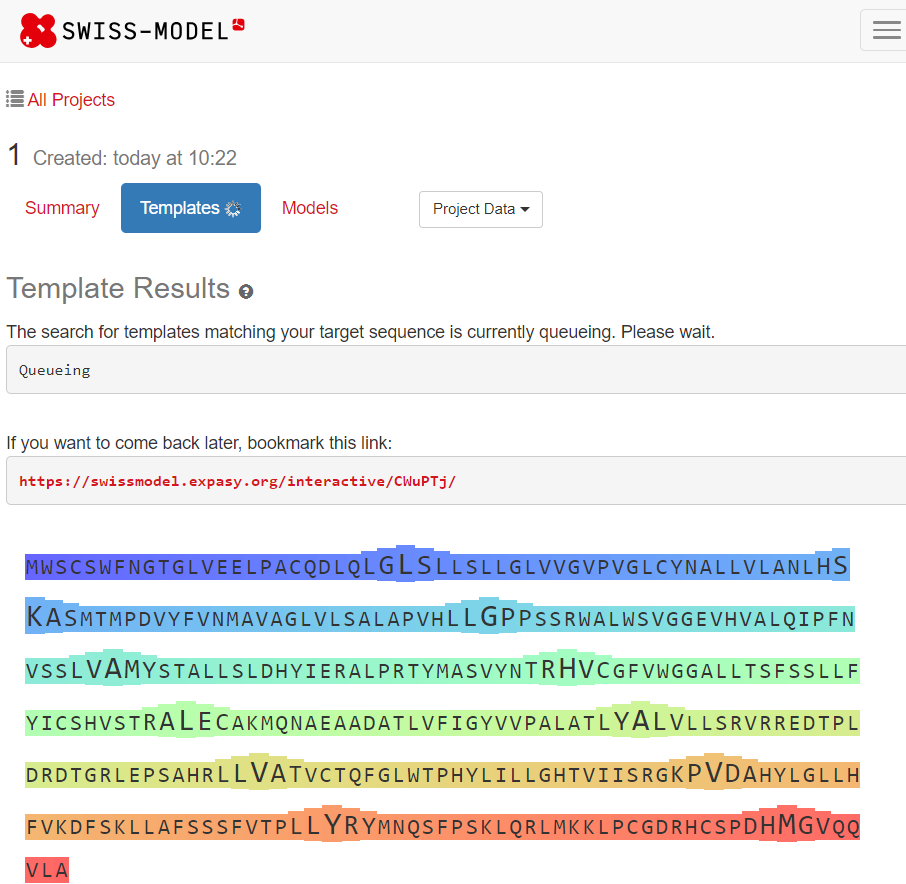
Sampling – sequencing – align reference in ncbi genome – check similarity in both patient’s seq and ref and find mutations

PyDESeq2 for python

Install pymol

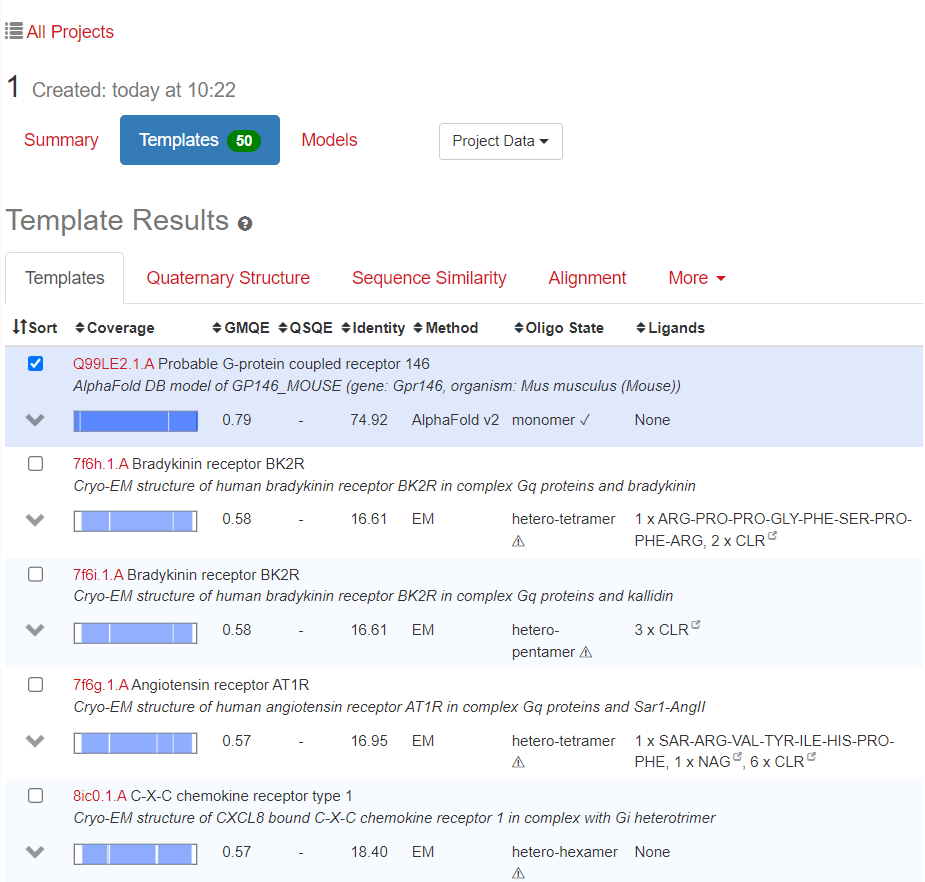
Molecular visualization tool

1. Get seq from xyz uniprot databaseds
2. Uniprot – search – copy seq – go to swiss model – start modelling– paste seq – give project name – give email (optional ) – search template button –



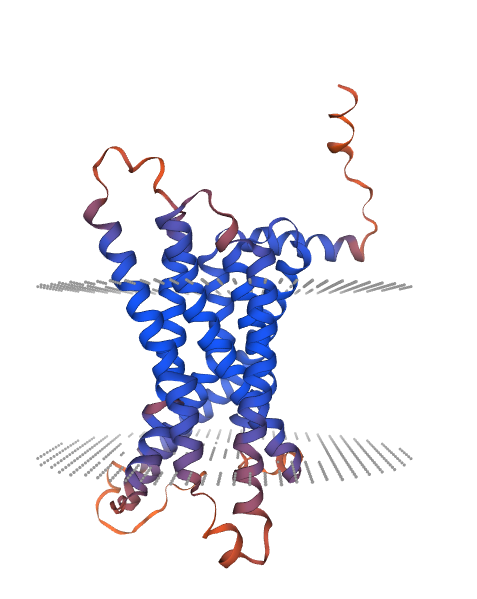
1. Generate all templates

* Select one which has higher identity number



* Select 74.9 one
* Select Build model

1. Predict 3d structure



Download it in PDB format and go to chimera

1. Do validation of 3d structure

Open image of pdb format in chimera

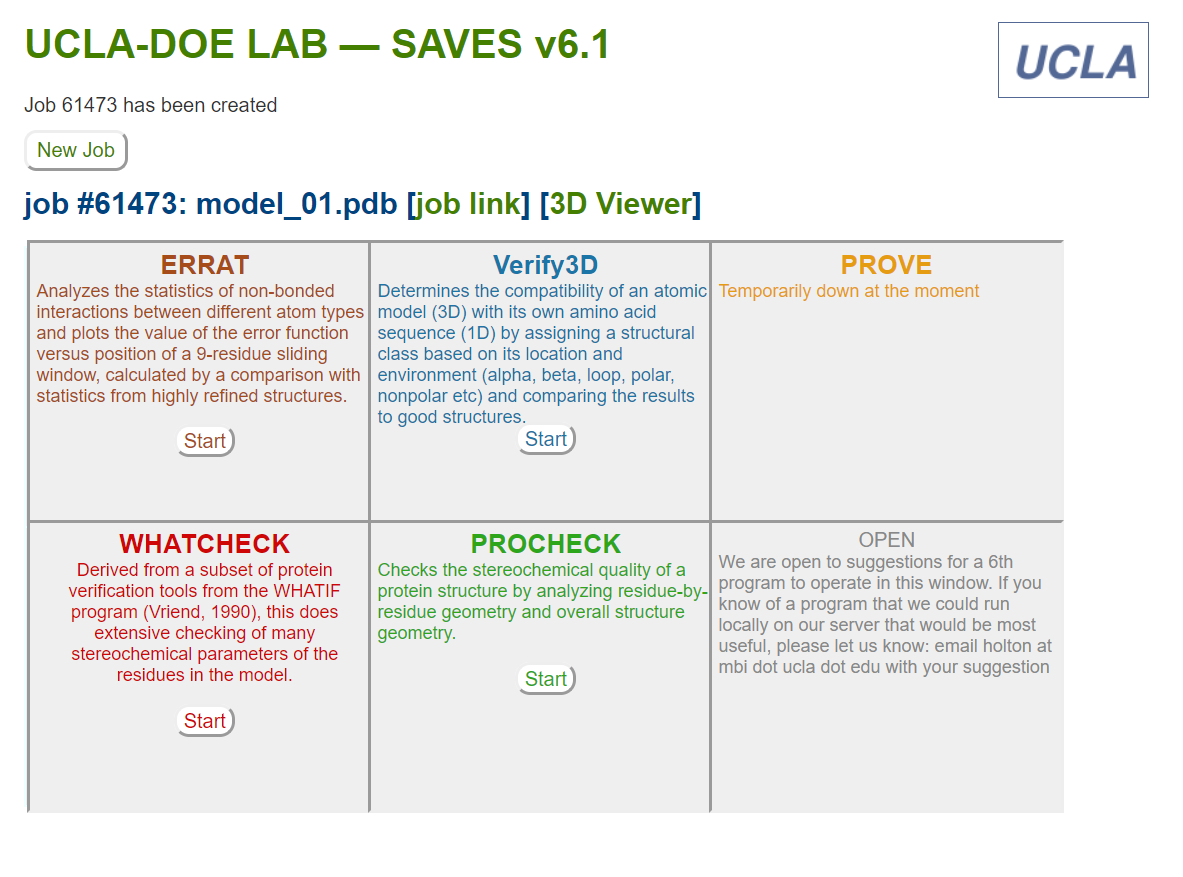
We can explore – color and all

If you go to actions->surface->show it shows different form of diagrams

Go to Saves v6.1 (website)– for validation of 3d structure of protein

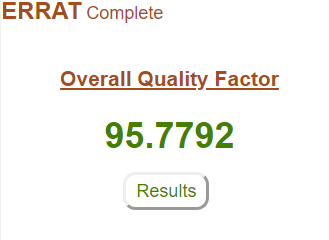
Upload pdb image here

Run program

You get these options   


Start all

Errat score close to 100 is a good score



Using ramchandran plot, we can say if protein structure is correct or not

It has 4 quarters- represents alpha helix and beta sheets

Docking – protein prep – ligand prep –

We take 2 molecules and we see how it binds, interacts , what is its energy

Pubchem – smile or pennicilia C drug – go to 3d conformer – file download in sdf convert to pdb or mol2 using chimera

Go to swiss dock for docking

Upload in swiss model submit ligand ‘submit model pdb image

Prepare target

Resize box or grid size

Show protein surface

Start docking – it will give job sort of bar will fill – it will take lot of time

It finds best possible cavity – it is computationally expensive

Go to hex docking server website

Go to hex server

Try haddock – website – good docking tool

Submit a new job -> Login -> academic creds -> go on

Threading server - > I-tasser -> give fasta seq as input

Go to plip ->

Hbplip

Pdbsum

[Vishwayogi@gmail.com](mailto:Vishwayogi@gmail.com)