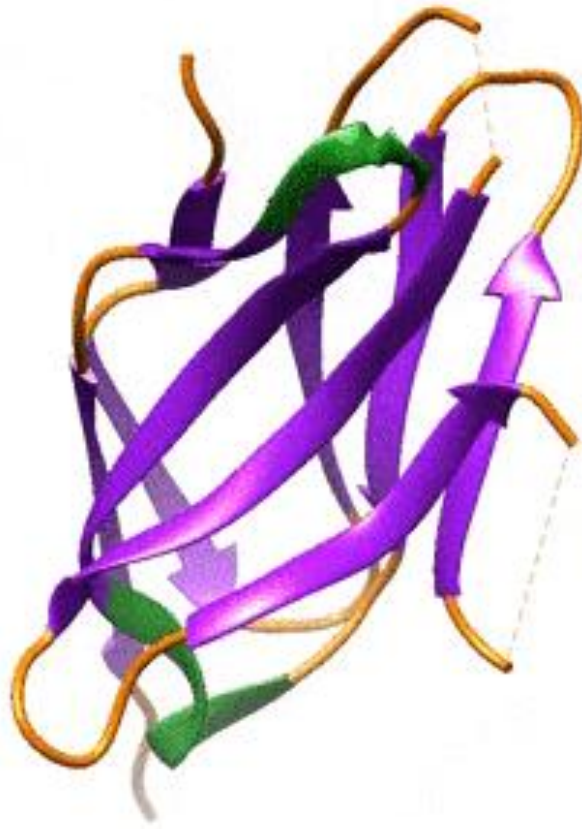


# **BUILDING PD-1 PROTEIN STRUCTURE FOR DIFFERENT MODELS USING HUMAN PD-1 STRUCTURE AS REFERENCE AND CHECKING THE STRUCTURAL AND SEQUENCE LEVEL SIMILARITIES BETWEEN THEM**

---



# Motivation

---

- Cancer is one of the leading causes of deaths in the world. It is a disease in which abnormal cells divide uncontrollably and destroy body tissue.
- The main cancer treatment include cancer surgery, radiotherapy, chemotherapy and hormone therapy.
- One of the emerging treatments for cancer is immune checkpoint inhibitors which has been approved for cancers like small cell carcinoma of lung, malign melanoma, renal cell carcinoma etc.
- Currently approved immune checkpoint inhibitors target the molecules CTLA4, PD1 and PDL1.

# Programmed Cell Death Protein 1: Functioning

---

- PD-1 is a protein found on T cells (a type of immune cell) that helps keep the body's immune responses in check.
- PD-1 extends from the surface of T-cells, and interacts with two similar ligand proteins, PD-L1 and PD-L2, that are found on the surface of regulatory T cells which control the immune system's response to self and foreign particles.
- Cancers such as melanomas evade the immune system by expressing PD-L1 on their surface, allowing them to trick the immune system by downregulating T-cells.
- Immune checkpoint inhibitors target these cancer cells by blocking the interaction of PD-1 and PD-L1, restoring T-cell function.

# Programmed Cell Death Protein 1: Properties

---

- Also called CD279
- Immunoglobulin superfamily, expressed on T-cells and pro B-cells
- Coded by PD1 gene which is expressed on chromosome 2 of human
- Composed of 288 amino acids
- Useful in prevention of autoimmune diseases
- Can prevent the immune system from attacking cancer cells
- Another reason for studying PD1: Drugs targeting PD1 along with other negative immune checkpoint receptors may augment immune response and/or facilitate HIV eradication.

# MODELS

MODEL	COMMON NAME	ACCESSION NUMBER
<i>Macaca mulatta</i>	Rhesus Monkey	NP_001107830.1
<i>Rattus norvegicus</i>	Brown Rat	NP_001100397.1
<i>Bos taurus</i>	Domesticated Cattle	BAX73992.1
<i>Sus scrofa</i>	Pig	NP_001191308

Is the result of the phylogenetic analysis correlating with the RMSD values for the above mentioned models?

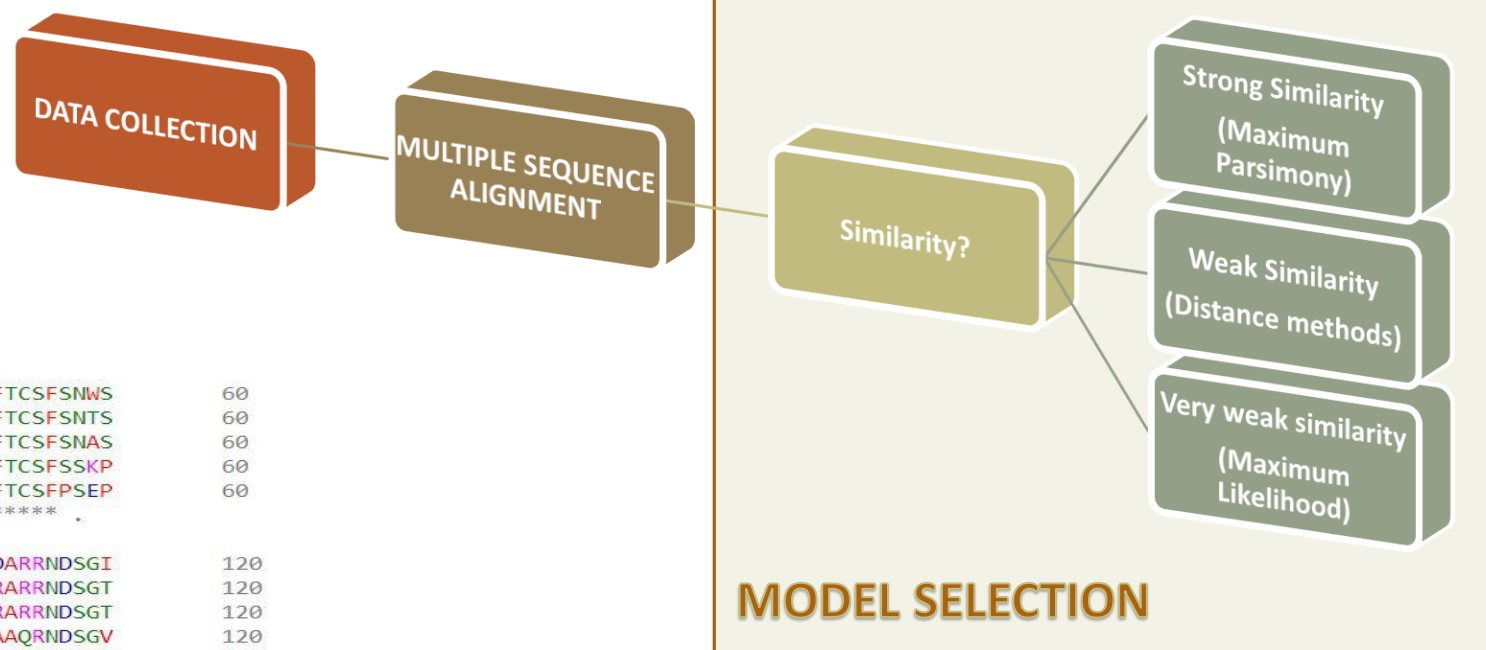
# SEQUENCE LEVEL ANALYSIS

## Tools Used

MegaX

Clustal Omega

# PHYLOGENETIC ANALYSIS



NP_001100397.1	MWVQQVPWSFTWAVLQLSWQSGWLLLEVLNKPWRPLTFSPTWLTVSEGANATFTCSFSNWS	60
NP_005009.2	MQIPQAPWPVWAVLQLGWRPGWFLDSPDRPWNPPTFSPALLVTEGDNATFTCSFSNTS	60
NP_001107830.1	MQIPQAPWPVWAVLQLGWRPGWFLDSPDRPWNPPTFSPALLVTEGDNATFTCSFSNAS	60
BAX73992.1	MGTPRALWPLWAVLQLGCWPGLLEASSRPWSALTFSPRLVPEGANATFTCSFSSKP	60
NP_001191308.1	MGTPRALWVWVVLQLRWPGWLLDAPSRPGPLTSPAQLTVPEGANATFTCSFPSEP	60
	* : . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . *	
NP_001100397.1	EDLKLNWYRLSPSNQTEKQAACNGYSQPVDRARFQIVQLPNGHDFHMINLDARRNDSGI	120
NP_005009.2	ESFVLNWYRMSPSNQTDKLAAFPEDRSQPGQDCRFRTQLPNGRDFHMSVVRARRNDSGT	120
NP_001107830.1	ESFVLNWYRMSPSNQTDKLAAFPEDRSQPGQDCRFRTQLPNGRDFHMSVVRARRNDSGT	120
BAX73992.1	ERFVLNWYRKSPSNQMDKLAAFPEDRSQPSRDRFRVTPLPDGQQFNMSIVAAQRNDSGV	120
NP_001191308.1	KHFILNWYRLSPSNQTDKLAAFSEDSQPGDRPRFHVTLPLNGRDFHMSVVATRRNDSGT	120
	: : * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . *	
NP_001100397.1	YLCGAISLPPKAQIKESPGAELVVTERILETPTRYPRSPKPEGQFQGLVIVIMSVLVGI	180
NP_005009.2	YLCGAISLAPKAQIKESLRAELRVTERRAEVPTAHPSPSPRPAGQFQTLVVGVVGGLLGS	180
NP_001107830.1	YLCGAISLAPKAQIKESLRAELRVTERRAEVPTAHPSPSPRPAGQFQALVVGVVGGLLGS	180
BAX73992.1	YFCGAIYLPPTQINESHSAELMVTEAVLEPPTPEPPSPQPRPEGQMQLVIGVTSVLLGV	180
NP_001191308.1	YFCGAIYLPPTQINESHQAKLTVTERVLELPTHEPSCPPRPEGHLEGQVLVITSVLLGL	180
	* . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . *	
NP_001100397.1	PVLLLLAWALAAFCSTGMSEAREAGRKEDPPKEAHAAAPVPSVAYEELDFQGREKTPEPA	240
NP_005009.2	LV--LLVWVLAVICSRRAAGTIGARRTGQPLKEDPSAVPVFSVDYGELDFQWREKTPEPP	238
NP_001107830.1	LV--LLVWVLAVICSRRAAGTIEARRTGQPLKEDPSAVPVFSVDYGELDFQWREKTPEPP	238
BAX73992.1	LLLPLLIWVLAADVLRATRGGCARRSQDQPPKEGCPSPAVTVDYGELDFQWREKTPEPA	240
NP_001191308.1	LLLLLLAWSLAFFLWAPRGDRAHRTENQPRKEGASSGLVFTVDYGELDFQWREKTPVPS	240
	: * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . *	
NP_001100397.1	-PCV--HTEYATIVFTEGLDASAIGRGSGADGPQGPRPPRHEDGHCSWPL	287
NP_005009.2	VPCVPEQTEYATIVFSPSGMGTSSPARRGSGADGPRSAQPLRPEDGHCSWPL	288
NP_001107830.1	APCVPEQTEYATIVFSPSGMGTSSPARRGSGADGPRSPRLRPEDGHCSWPL	288
BAX73992.1	APCVPEQTEYATIVF-----GRRASADSPQGPWPLRTEDGHCSWPL	282
NP_001191308.1	AACVSEQTEYATIVFPERPG--SPGRRASADSPQGPWPQRTEDGHCSWPL	288
	* . * . * . * . * . * . * . * . * . * . * . * . * . * . * . * . *	

# Percent Identity Matrix - created by Clustal2.1

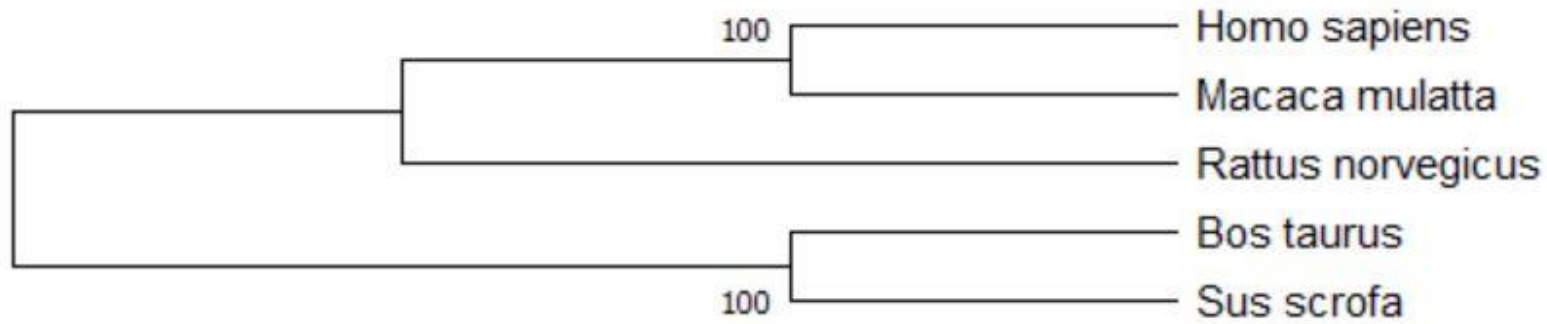
#

#

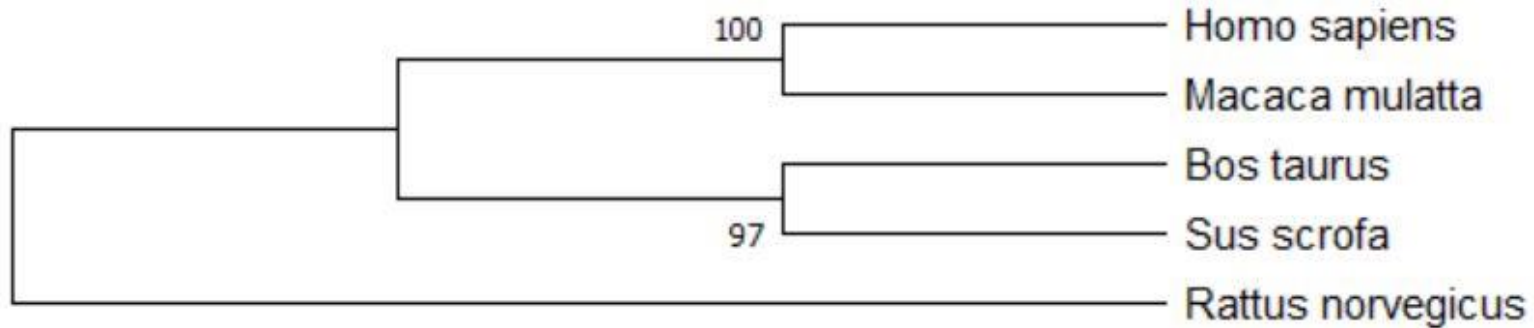
1:	NP_001100397.1	100.00	61.75	63.86	58.42	57.89
2:	NP_005009.2	61.75	100.00	96.18	66.43	64.34
3:	NP_001107830.1	63.86	96.18	100.00	67.14	64.69
4:	BAX73992.1	58.42	66.43	67.14	100.00	73.05
5:	NP_001191308.1	57.89	64.34	64.69	73.05	100.00



# Phylogenetic trees based on distance based models



Neighbor Joining



UPGMA



# STRUCTURE LEVEL ANALYSIS

## TOOLS USED

NCBI Blast

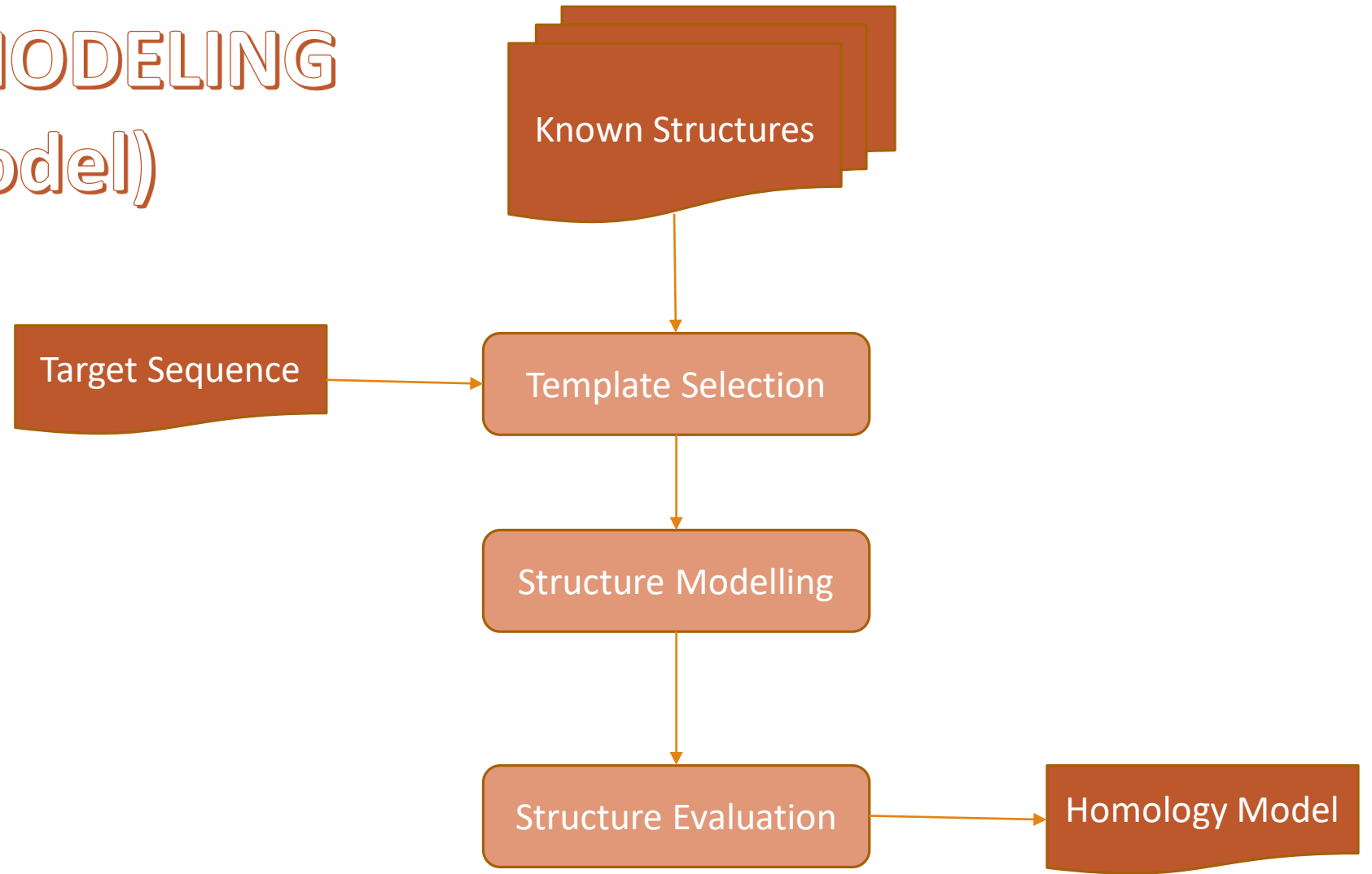
Swiss Model

Modeller

Chimera

ProCheck

# HOMOLOGY MODELING (Swiss Model)



# HOMOLOGY MODELING(Modeller)

## DATA COLLECTION

- Download the target protein sequence
- Run Blast
- Select the template with the lowest resolution
- Download the aligned sequence and the structure of the selected template

## DATA PREPARATION

- Multiple sequence alignment
- Keep only the part of the template with which the sequence got aligned
- Remove the residues

## MODELLING

- Define the number of models
- Select the model which gives the least DOPE value

## LOOP REFINEMENT

- Refine a region of an existing coordinate
- Define the number of models
- Select the model which gives the least DOPE value

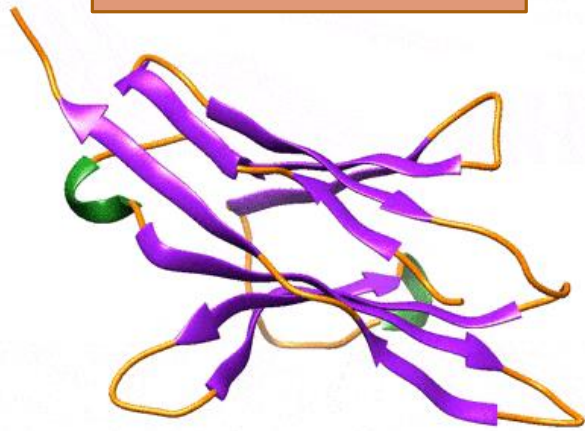
## MODEL EVALUATION

- RMSD Score
- Ramachandran plot

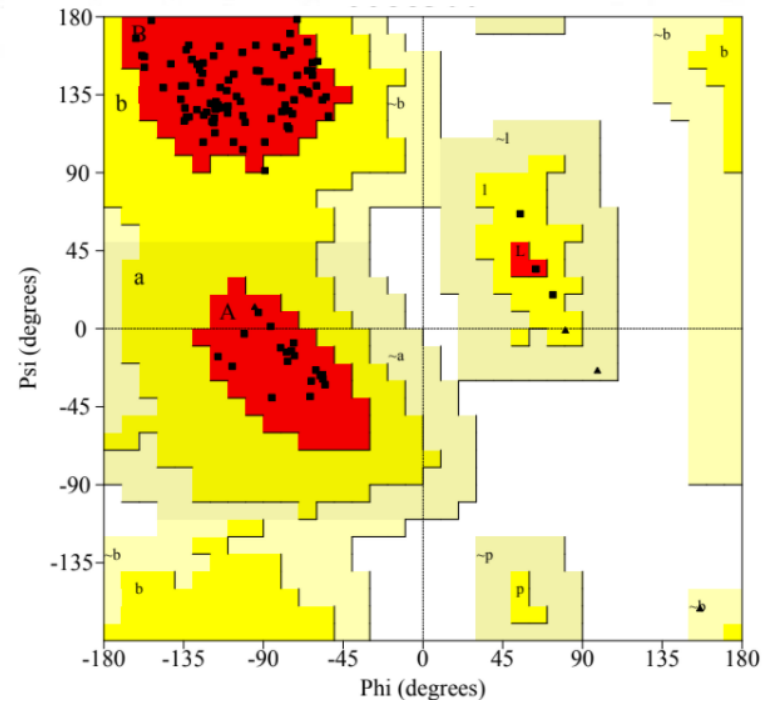
# TEMPLATES

MODELS	PDB ID	
	Swiss Model	Modeller
<i>Macca mulatta</i>	6K0Y	6J14
<i>Rattus norvegicus</i>	3BP5	5WT9
<i>Bos taurus</i>	6K0Y	5WT9
<i>Sus scrofa</i>	6K0Y	5WT9

Swiss Model

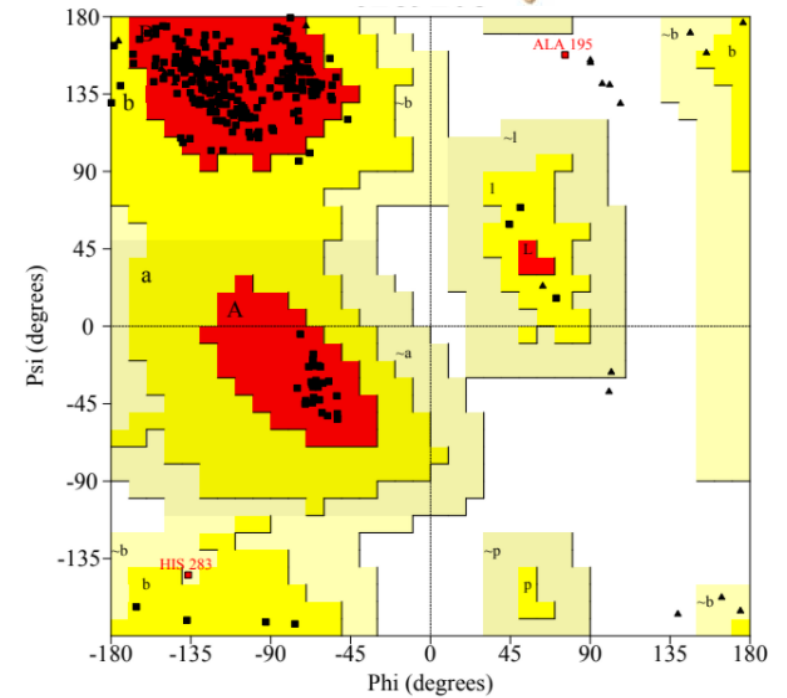


*Macaca mulatta*  
(Rhesus Monkey)



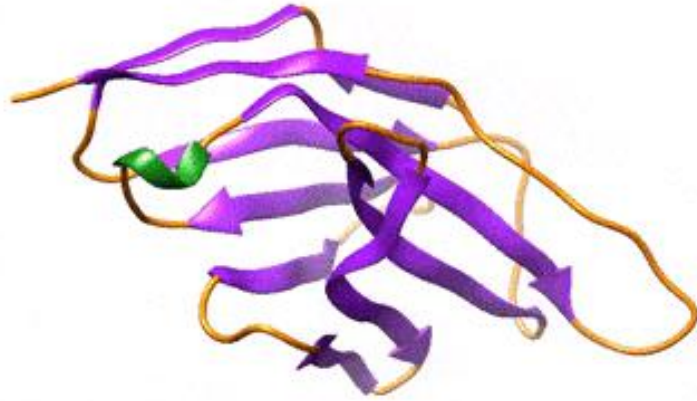
Residues in most favoured regions : 97%

Modeller

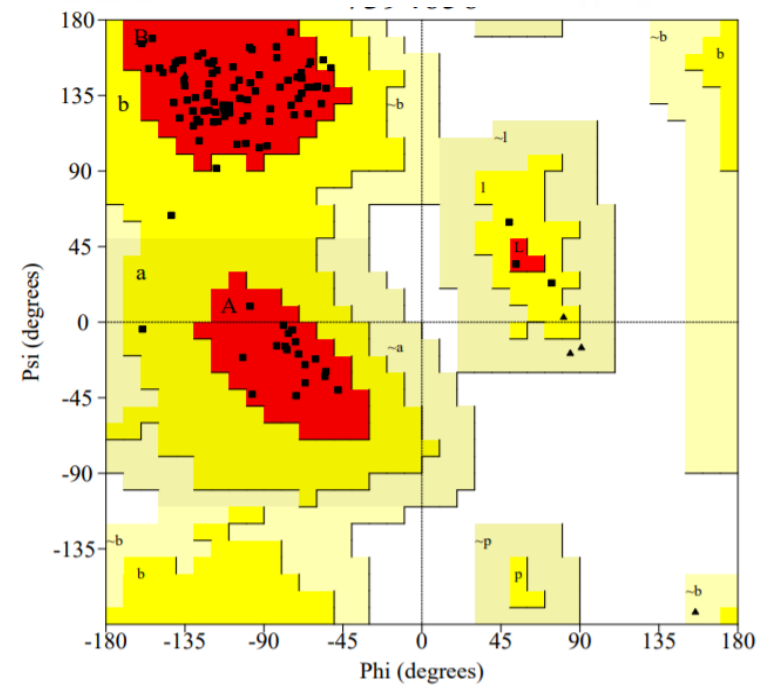


Residues in most favoured regions : 93.1%

Swiss Model

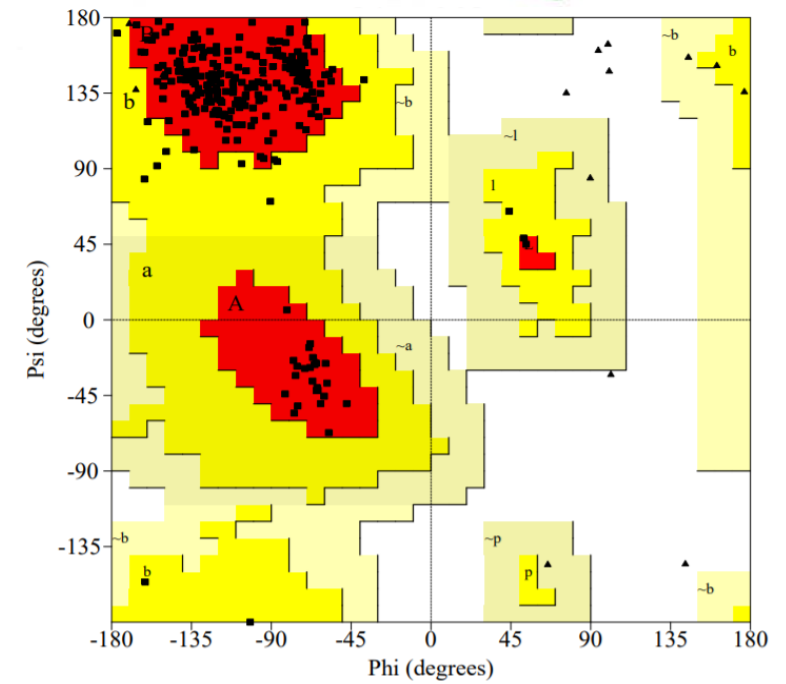
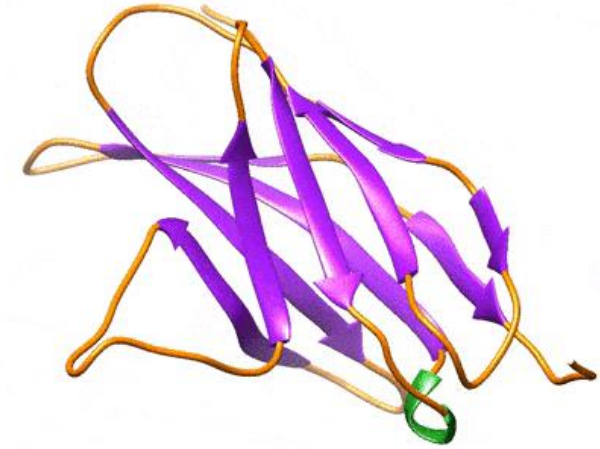


*Rattus norvegicus*  
(BROWN RAT)



Residues in most favoured regions : 94.9%

Modeller

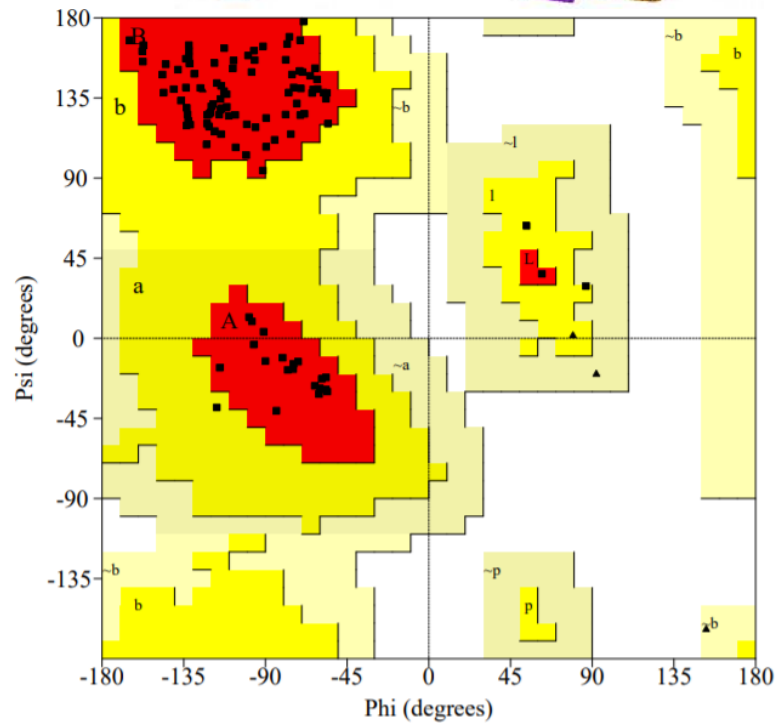


Residues in most favoured regions : 94.5%

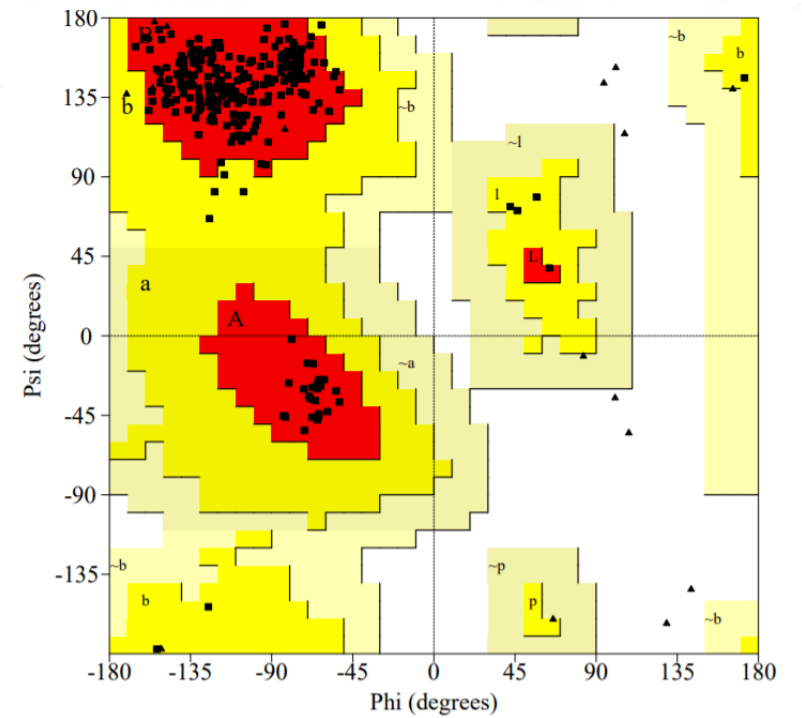
Swiss Model

Modeller

# *Bos taurus* (Cattle)



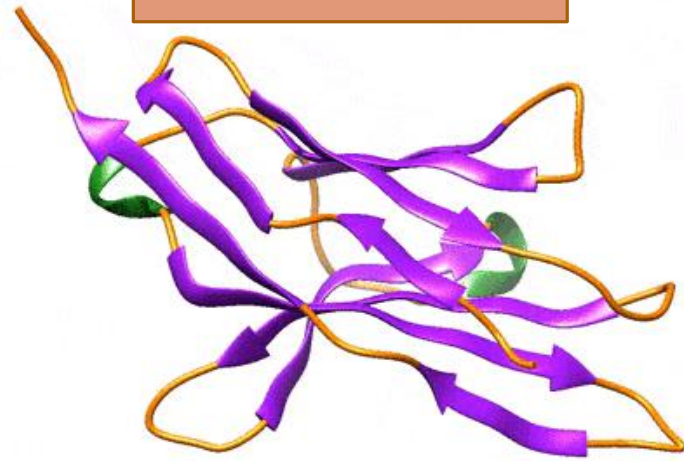
Residues in most favoured regions : 97%



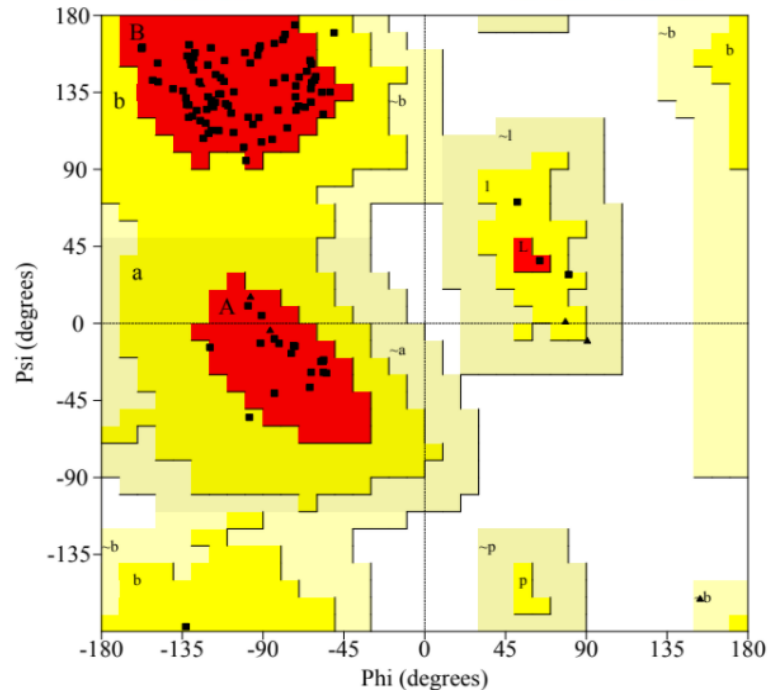
Residues in most favoured regions : 94.7%



Swiss Model

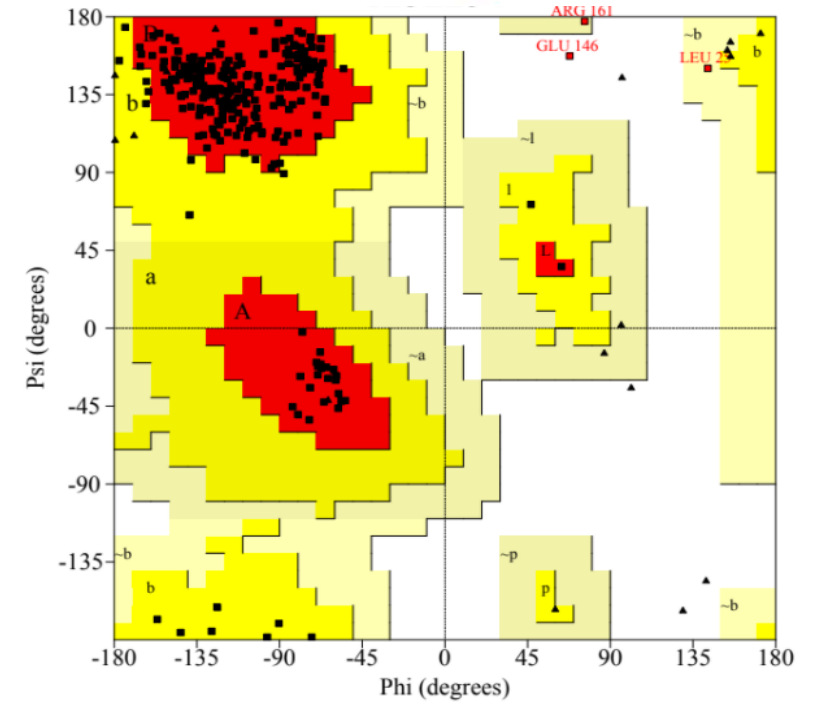
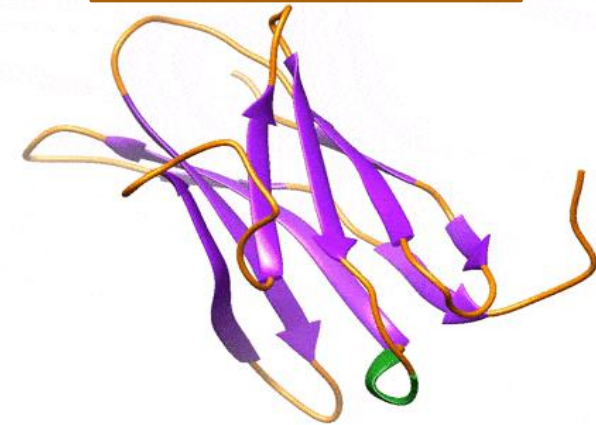


*Sus scrofa*  
(Pig)



Residues in most favoured regions : 96.9%

Modeller

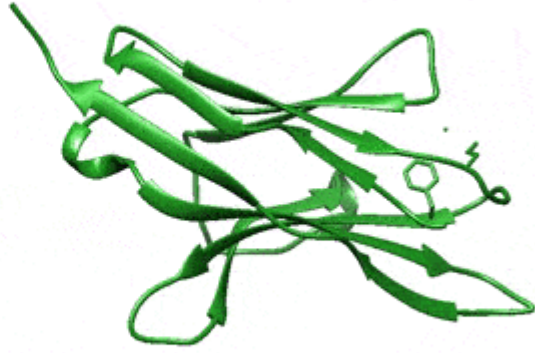


Residues in most favoured regions : 91%

RMSD : 0.073 °A



Macaca Mulatta

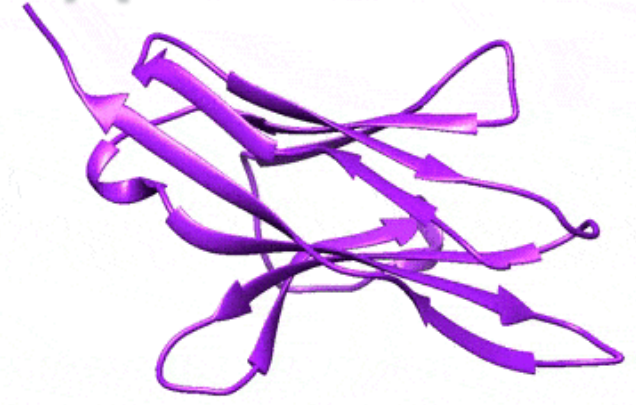


Homo Sapiens - 6K0Y

RMSD : 0.140 °A



Homo Sapiens - 6K0Y



Bos Taurus

RMSD : 0.142 °A



Sus Scrofa

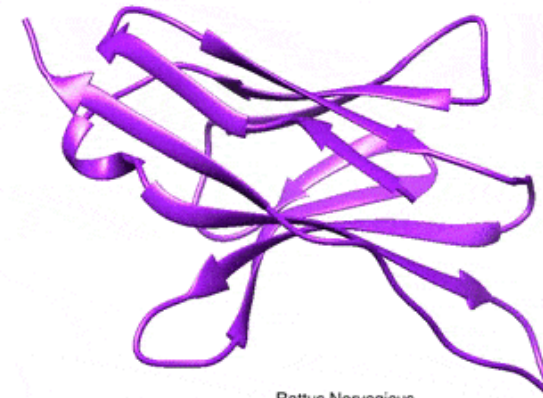


Homo Sapiens - 6K0Y

RMSD : 0.088 °A



Homo Sapiens - 3BP5



Rattus Norvegicus

# CONCLUSION

**Hypothesis :** Sequence analysis = Structure analysis

**Observations :**

- Phylogenetic analysis : *Macaca mulatta* is closest to *Homo sapiens*
- Homology Modeling : In terms of the RMSD, *Macaca mulatta* gives us the minimum score.

**Interpretation :**

Since, sequence and structural analysis give us the same results, we do not reject our hypothesis. Therefore, we conclude that *Macaca mulatta* will be the best of the four models for drug trials and studies related to PD1 protein of *Homo sapiens*.

Thank You!