

## Basic concept of motif identification 2.

How do we represent the position specific preference ?

BID_MOUSE	I	A	R	H	L	A	Q	I	G	D	E	M
BAD_MOUSE	Y	G	R	E	L	R	R	M	S	D	E	F
BAK_MOUSE	V	G	R	Q	L	A	L	I	G	D	D	I
BAXB_HUMAN	L	S	E	C	L	K	R	I	G	D	E	L
BimS	I	A	Q	E	L	R	R	I	G	D	E	F
HRK_HUMAN	T	A	A	R	L	K	A	L	G	D	E	L
Eg1-1	I	G	S	K	L	A	A	M	C	D	D	F

Statistical  
representation

G: 5 -> 71%

S: 1 -> 14 %

C: 1 -> 14 %

# Protein motif /domain

- Structural unit
- Functional unit
- Signature of protein family

**How are they defined?**

# Secondary structure prediction

Predict the likelihood of **amino acid x** to be in each of the three (four) types of secondary structure configuration

- **Helix**
- **Sheet**
- **Turn**
- **Coil**

Coiled-coil is two helices tangled together

(Slides before 2015/2016)

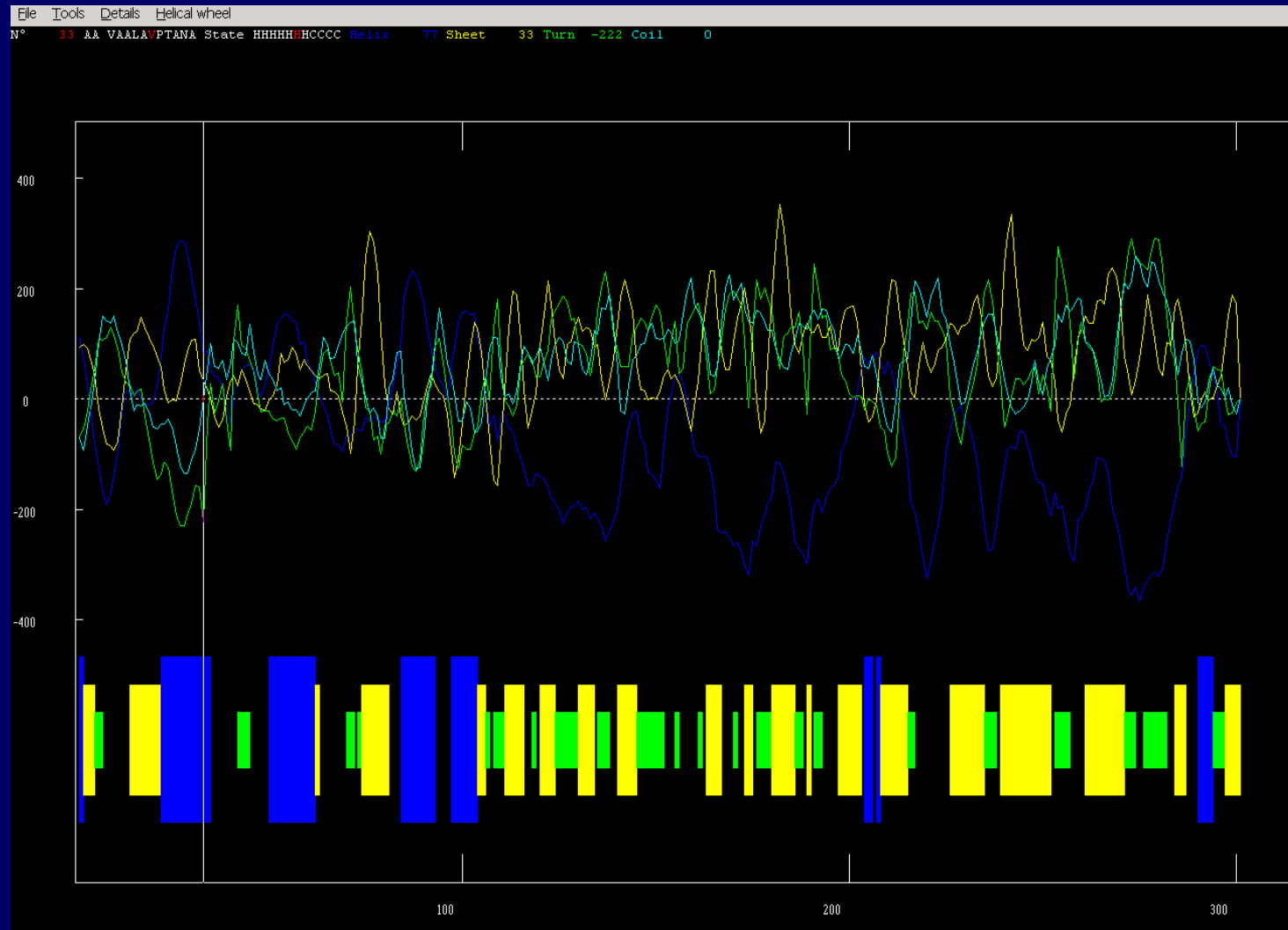
# **Secondary structure prediction**

## **- different strategies and algorithms**

- **Chou-Fasman / Garnier Method**
  - based on AA composition
- **Nearest Neighbor / Levin Method**
  - based on sequence similarity
- **Neural Network / PHD**
- **SOPM, DPM, DSC, etc.**

**(Slides before 2015/2016)**

# Results are given at single amino acid level



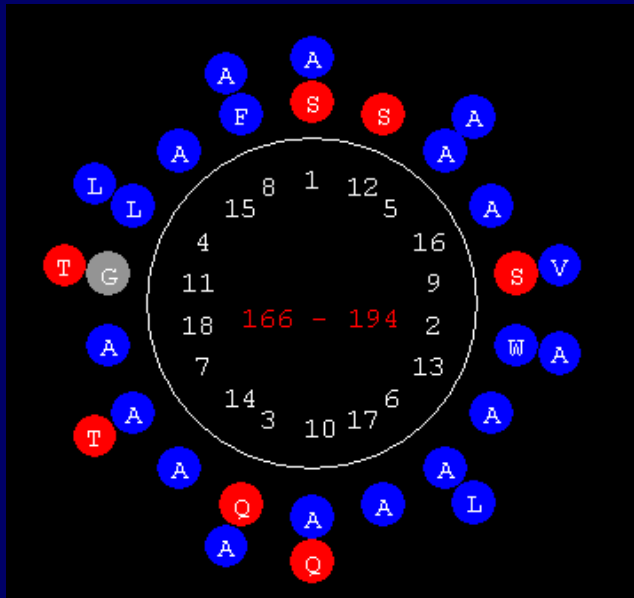
(Slides before 2015/2016)

# Helix wheel to discern helix subtype

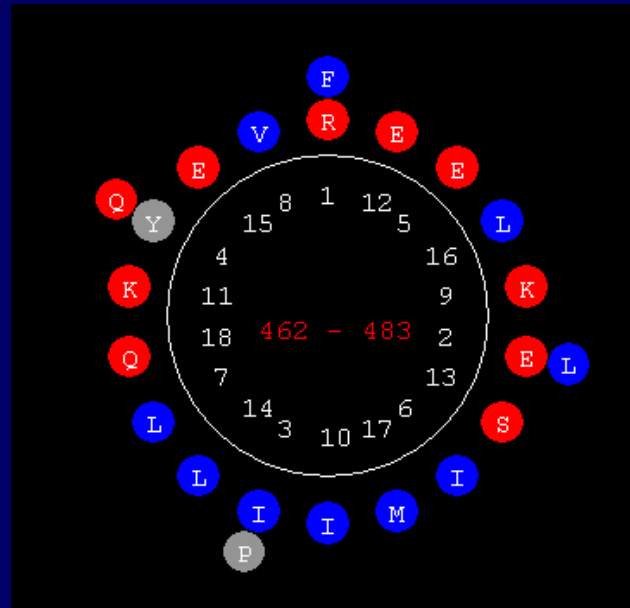
● Hydrophilic

● Hydrophobic

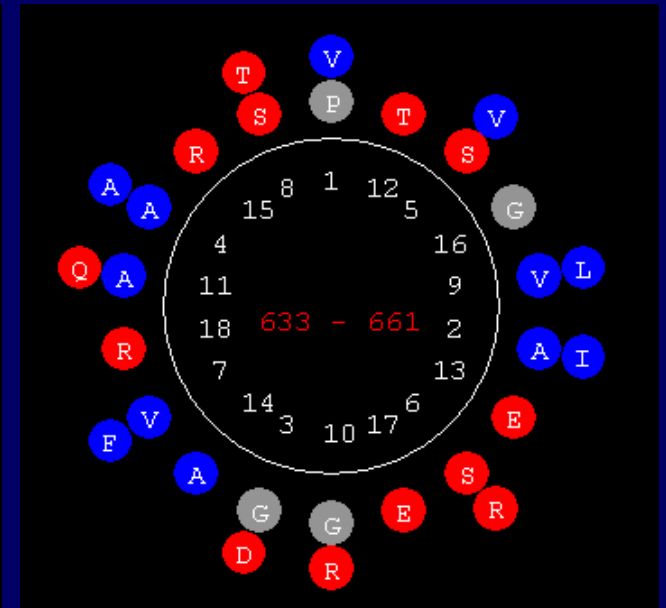
● Others



Hydrophobic



Amphipathic

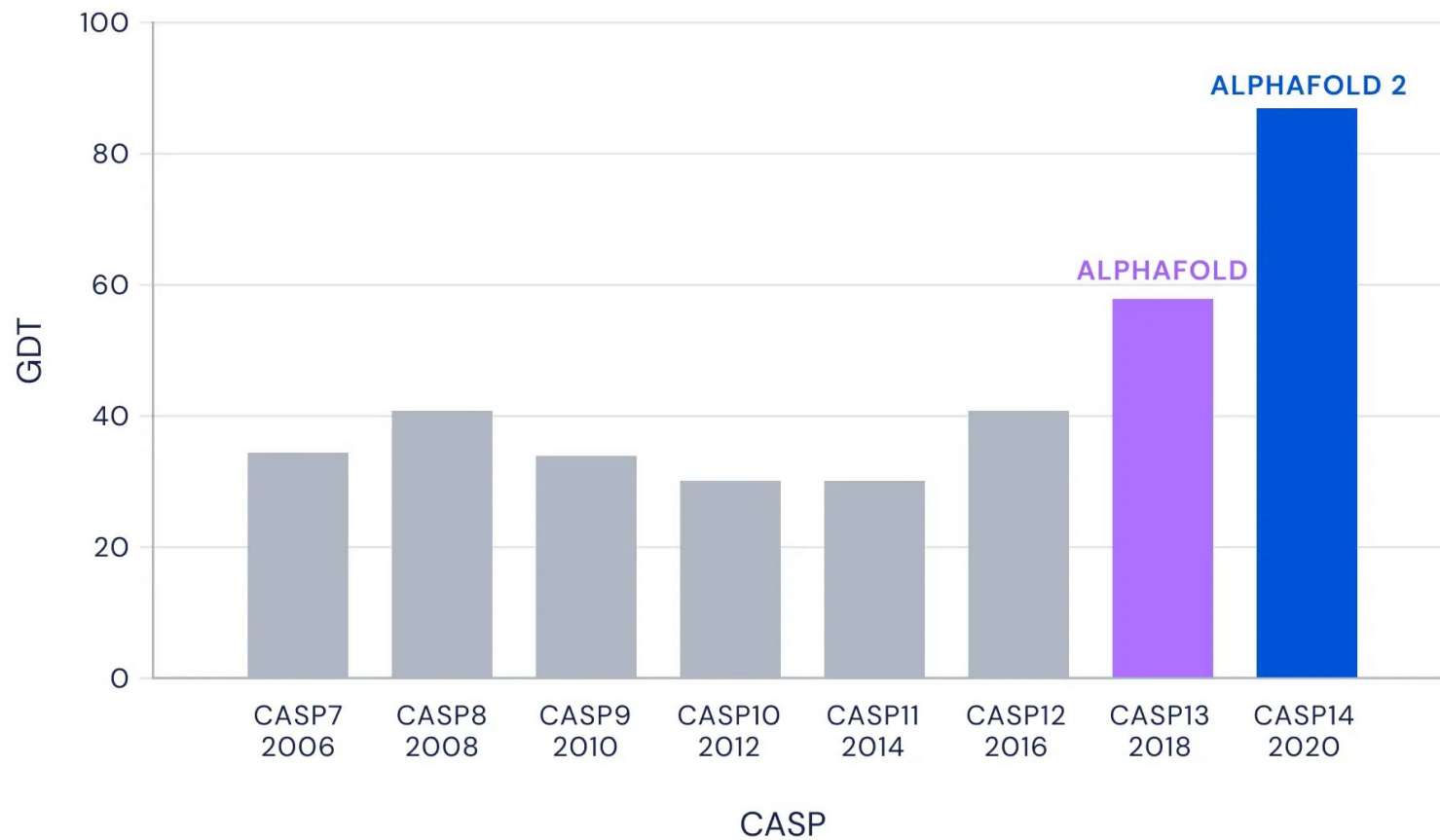


Hydrophilic

(Slides before 2015/2016)

# AlphaFold Revolution

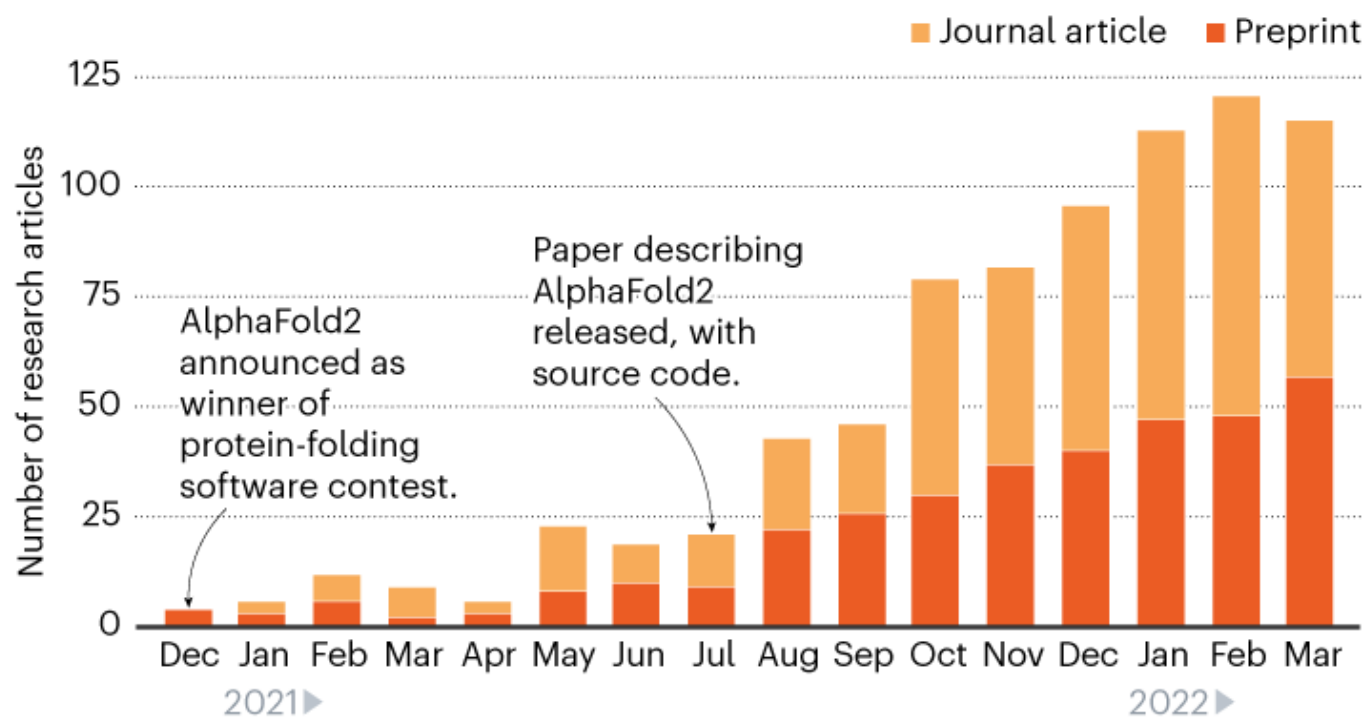
Median Free-Modelling Accuracy



# AlphaFold Revolution

## ALPHAFOLD MANIA

The number of research papers and preprints citing the AlphaFold2 AI software has shot up since its source code was released in July 2021\*.



\*Nature analysis using Dimensions database; removing duplicate preprints and papers/R. Van Noorden, E. Callaway.



**Practice –Running AlphFold2 in  
HiperGator**

# **New frontiers of protein structure analysis**

- **Effect of mutation(s) on protein structure.**
- **Complexes:**
  - **Protein : Protein complexes**
  - **Protein :non-peptide macromolecule complexes**
  - **...**
- **Intrinsically disordered protein regions (IDRs).**

# View and manipulate protein structures

Many options:

- ❖ Jalview
- ❖ PyMol
- ❖ UCSF ChimeraX
- ❖ ...

**Observe – Manipulating and viewing  
protein structures in PyMol.**