

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)

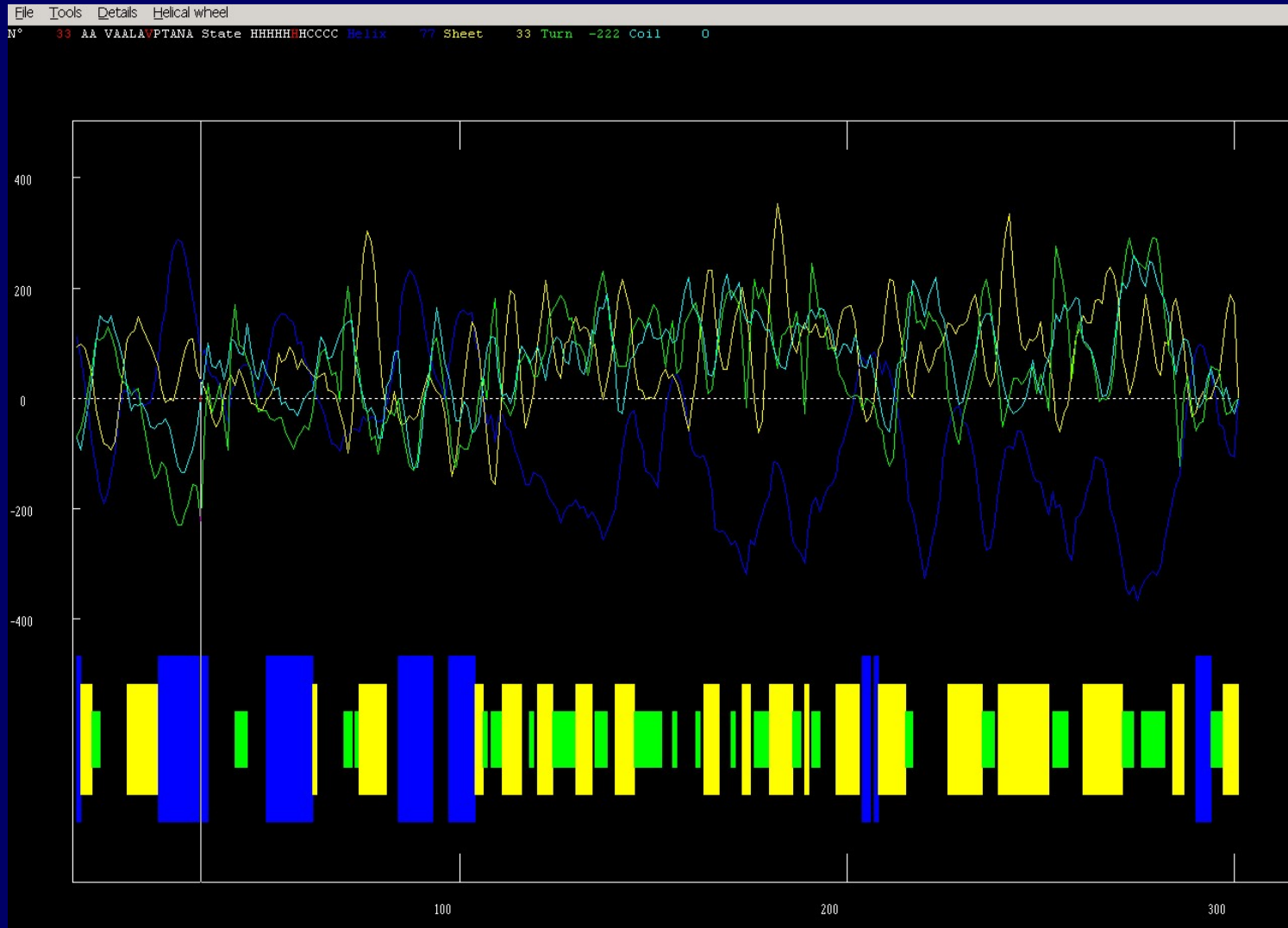
Secondary structure prediction

- 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



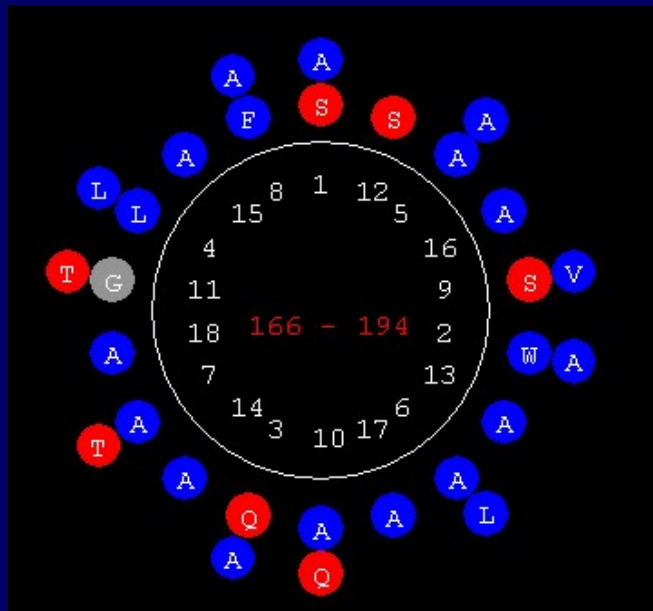
(Slides before 2015/2016)

Helix wheel to discern helix subtype

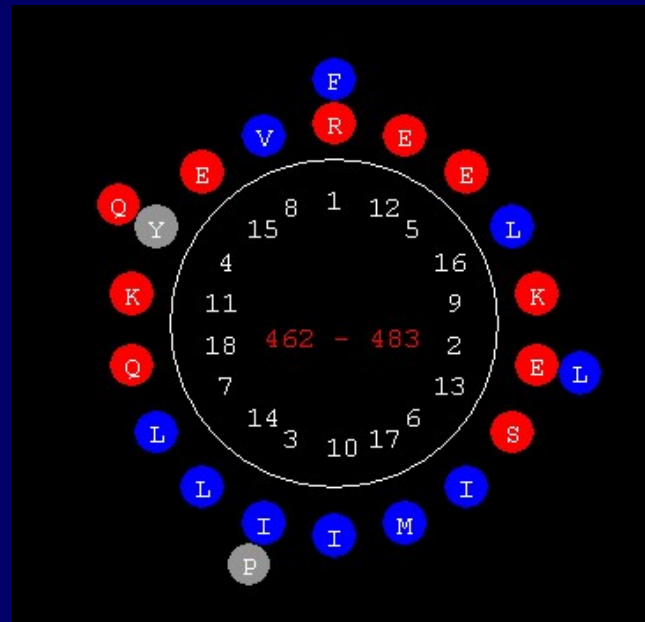
● Hydrophilic

● Hydrophobic

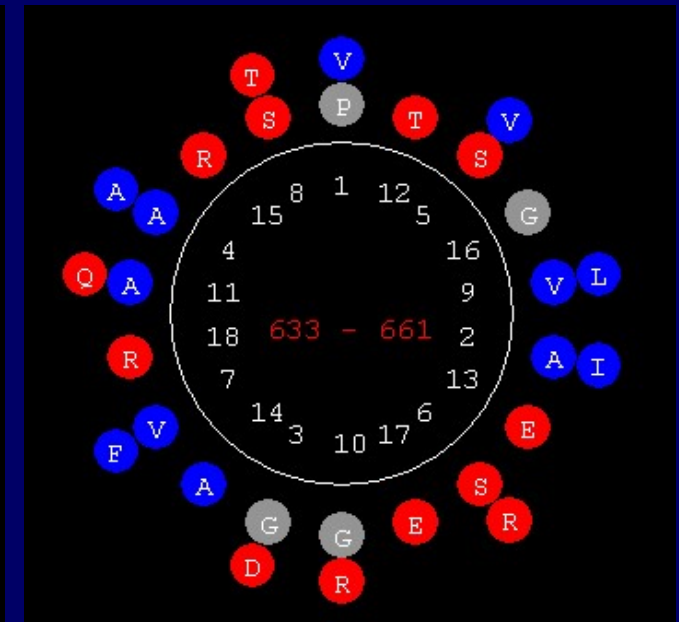
● Others



Hydrophobic



Amphipathic

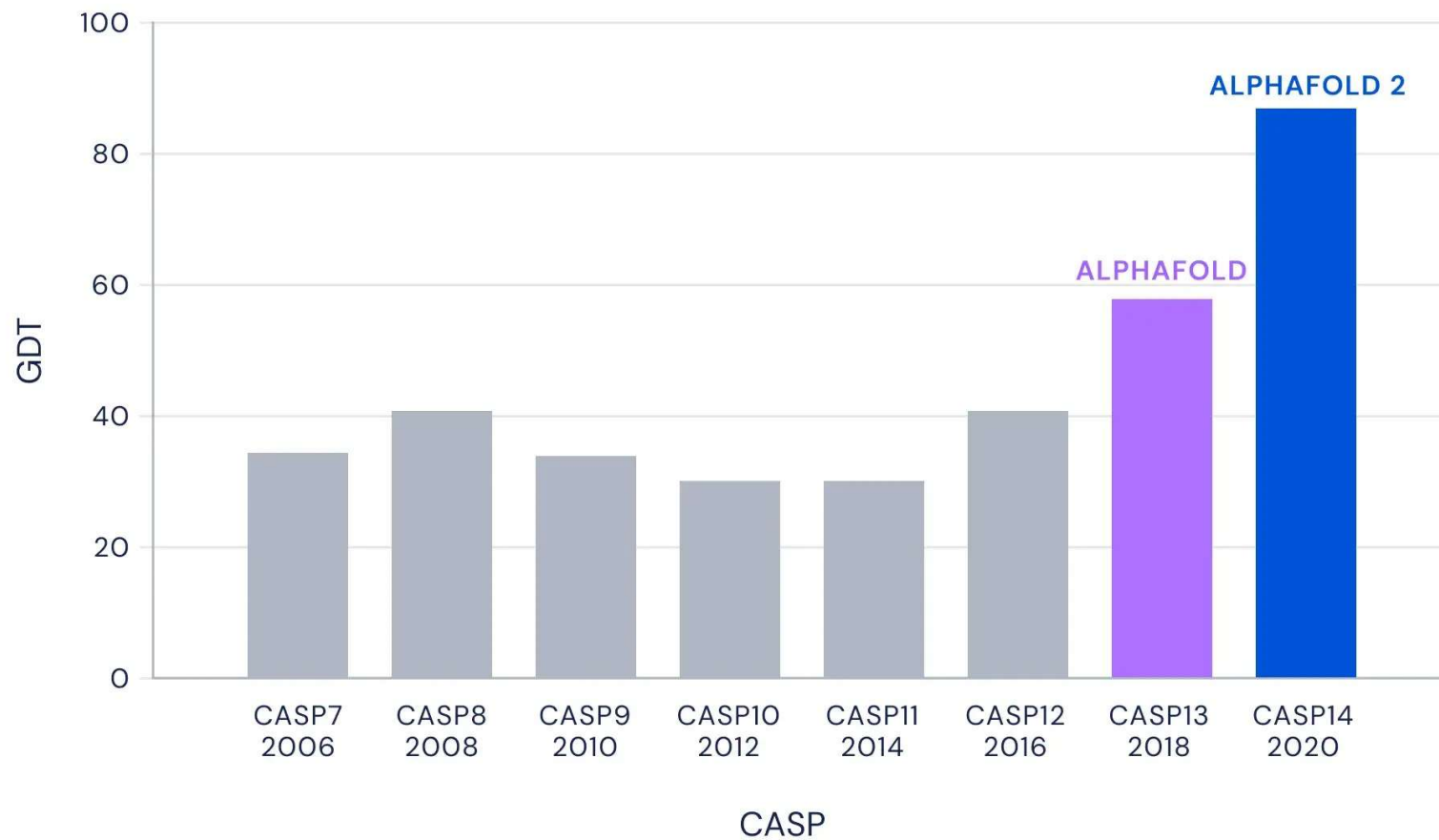


Hydrophilic

(Slides before 2015/2016)

AlphaFold Revolution

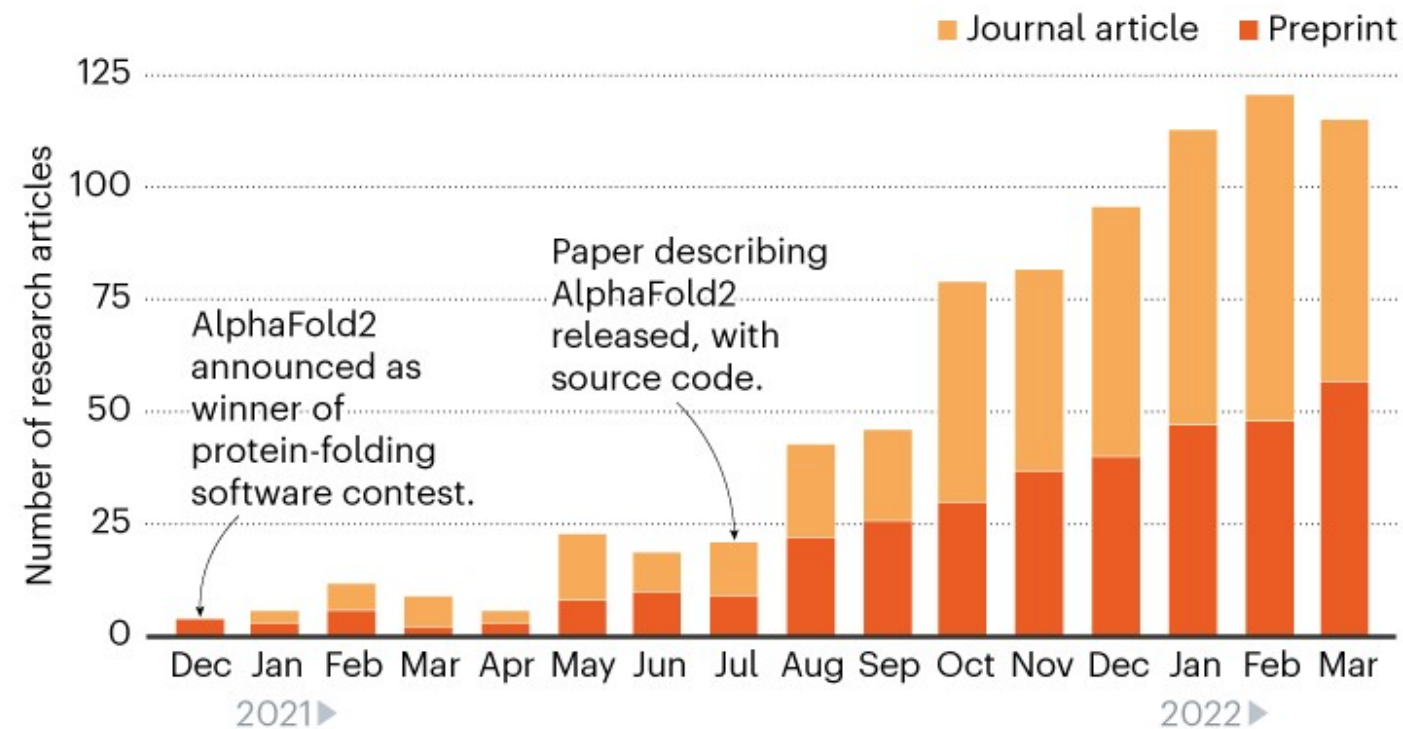
Median Free-Modelling Accuracy



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.

©nature

Transformer / Evoformer (AlphaFold2)

Attention Is All You Need

Ashish Vaswani*

Google Brain

avaswani@google.com

Noam Shazeer*

Google Brain

noam@google.com

Niki Parmar*

Google Research

nikip@google.com

Jakob Uszkoreit*

Google Research

usz@google.com

Llion Jones*

Google Research

llion@google.com

Aidan N. Gomez* †

University of Toronto

aidan@cs.toronto.edu

Łukasz Kaiser*

Google Brain

lukaszkaiser@google.com

Illia Polosukhin* ‡

illia.polosukhin@gmail.com

Practice –Running AlphFold2 in HiperGator

HPC resource allocation

Download SRA

```
#!/bin/sh
#SBATCH --job-name=Get_SRA
#SBATCH --mail-type=ALL
#SBATCH --mail-user=xxxx@ufl.edu
#SBATCH -t 12:00:00
#SBATCH --cpus-per-task=1
#SBATCH --mem-per-cpu=3gb
#SBATCH --qos=gms6014
#SBATCH --output=Get_SRA_%j.log
```

Alphafold

```
#!/bin/bash
#SBATCH --partition=gpu
#SBATCH --constraint=ai
#SBATCH --nodes=1
#SBATCH --ntasks=8
#SBATCH --gpus=1
#SBATCH --mem=48gb
#SBATCH --time=12:00:00
#SBATCH --mail-user=xxxxx@ufl.edu
#SBATCH --qos=gms6014
#SBATCH --output=alphafold_%j.log
```

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.**