



Exploring Biomolecular Structures with NCBI's iCn3D

Alexa M. Salsbury, Ph.D.

<https://bit.ly/3Xui4qr>

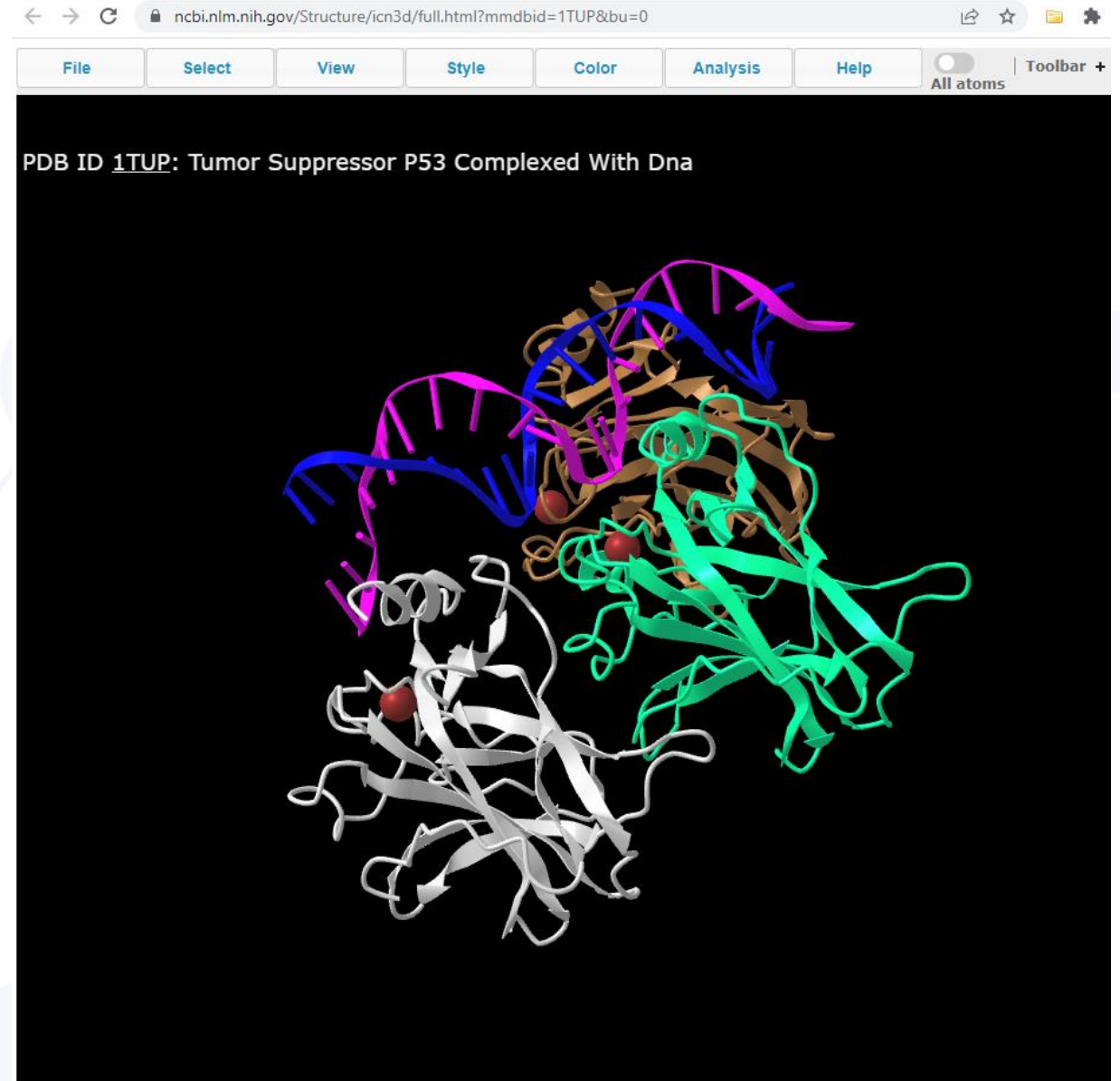


National Library of Medicine
National Center for Biotechnology Information



Overview

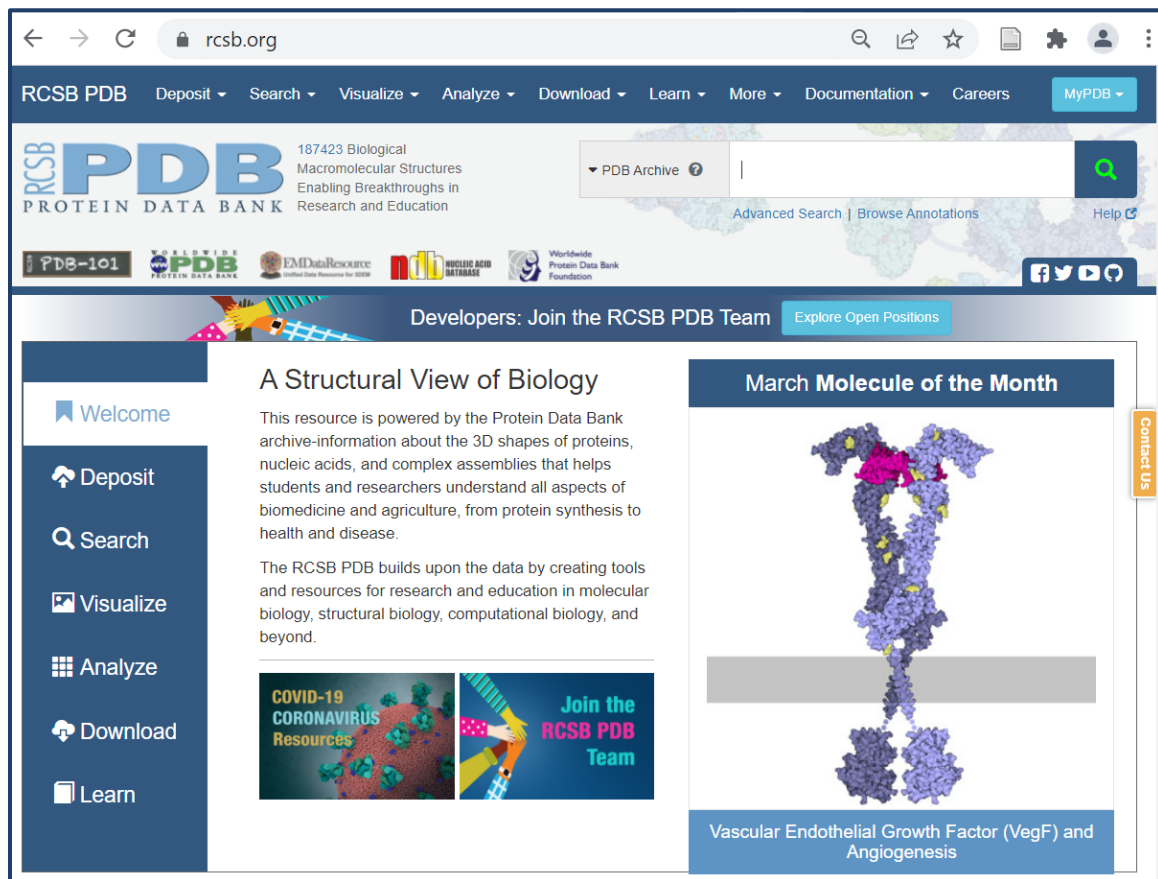
- Background
- iCn3D Fundamentals (Selection, Coloring, Style, and Sharing)
- Group Work
 - [Example 1](#): TP53 Mutation Analysis
 - [Example 2](#): TP53 from Structure to Function
 - [Example 3](#): Compare Crystal and AlphaFold TP53 Structures
- Group Discussion & Event wrap-up



Experimental techniques

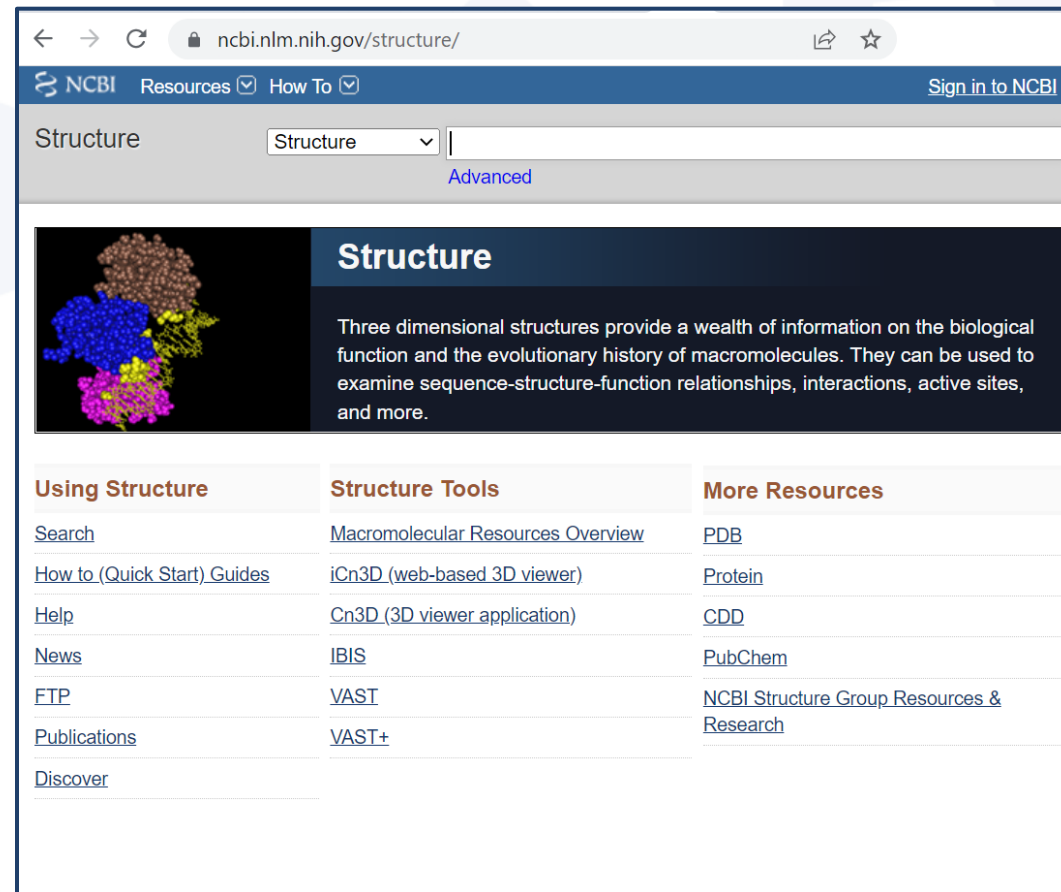
| | Advantages | Disadvantages |
|------------------------------|---|---|
| X-ray crystallography | <ul style="list-style-type: none">• Well developed• High resolution• Broad molecular weight range | <ul style="list-style-type: none">• Difficult sample prep• Static crystalline state |
| NMR | <ul style="list-style-type: none">• High resolution• 3D structure in solution• Good for dynamic study | <ul style="list-style-type: none">• Difficult sample prep• High sample purity needed |
| Cryo-EM | <ul style="list-style-type: none">• Simple sample prep• Structure in native state• Small sample size needed | <ul style="list-style-type: none">• Lower resolution• Works best for samples with high molecular weight• Equipment can be expensive, but costs are decreasing |

Where do I find experimentally determined structures?



The screenshot shows the RCSB PDB website. The header includes navigation links: Deposit, Search, Visualize, Analyze, Download, Learn, More, Documentation, Careers, and MyPDB. The main content area features a search bar with a dropdown menu for 'PDB Archive' and a search button. Below the search bar, there are several featured sections: 'A Structural View of Biology' with a description of the PDB archive, 'March Molecule of the Month' featuring a 3D model of a protein structure, and a 'Join the RCSB PDB Team' button. The footer includes social media links and a 'Contact Us' button.

RCSB Protein Data Bank



The screenshot shows the NCBI Structure Database website. The header includes navigation links: NCBI, Resources, How To, and Sign in to NCBI. The main content area features a search bar with a dropdown menu for 'Structure' and a search button. Below the search bar, there are several featured sections: 'Using Structure' with links to Search, How to (Quick Start) Guides, Help, News, FTP, Publications, and Discover; 'Structure Tools' with links to Macromolecular Resources Overview, iCn3D (web-based 3D viewer), Cn3D (3D viewer application), IBIS, VAST, and VAST+; and 'More Resources' with links to PDB, Protein, CDD, PubChem, and NCBI Structure Group Resources & Research.

NCBI Structure Database

Protein Data Bank ([PDB](https://www.rcsb.org/))

- New Structures are deposited daily

Each structure contains:

- 3D atomic coordinates
- Mandatory Metadata
 - Author Information
 - Primary citation
 - Experimental Data
 - Polymer sequence(s)- proteins, DNA, RNA
 - Small Chemical component structures- ligands, inhibitors, etc.

6LU7

The crystal structure of COVID-19 main protease in complex with an inhibitor N3

DOI: [10.2210/pdb6LU7/pdb](https://doi.org/10.2210/pdb6LU7/pdb)

Classification: [VIRAL PROTEIN](#)

Organism(s): [Severe acute respiratory syndrome coronavirus 2, synthetic construct](#)

Expression System: [Escherichia coli BL21\(DE3\)](#)

Mutation(s): No 

Deposited: 2020-01-26 **Released:** 2020-02-05

Deposition Author(s): [Liu, X.](#), [Zhang, B.](#), [Jin, Z.](#), [Yang, H.](#), [Rao, Z.](#)

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 2.16 Å

R-Value Free: 0.235

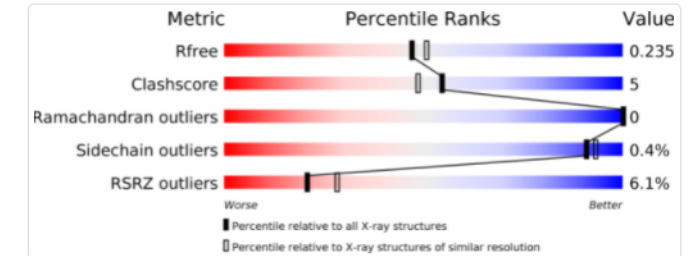
R-Value Work: 0.202

R-Value Observed: 0.204

wwPDB Validation

[3D Report](#)

[Full Report](#)



Literature

[Download Primary Citation](#)

Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors.

[Jin, Z.](#), [Du, X.](#), [Xu, Y.](#), [Deng, Y.](#), [Liu, M.](#), [Zhao, Y.](#), [Zhang, B.](#), [Li, X.](#), [Zhang, L.](#), [Peng, C.](#), [Duan, Y.](#), [Yu, J.](#), [Wang, L.](#), [Yang, K.](#), [Liu, F.](#), [Jiang, R.](#), [Yang, X.](#), [You, T.](#), [Liu, X.](#), [Yang, X.](#), [Bai, F.](#), [Liu, H.](#), [Liu, X.](#), [Guddat, L.W.](#), [Xu, W.](#), [Xiao, G.](#), [Qin, C.](#), [Shi, Z.](#), [Jiang, H.](#), [Rao, Z.](#), [Yang, H.](#)

(2020) *Nature* **582**: 289-293


PubMed: [32272481](#) [Search on PubMed](#)

DOI: [10.1038/s41586-020-2223-y](https://doi.org/10.1038/s41586-020-2223-y)

Primary Citation of Related Structures:

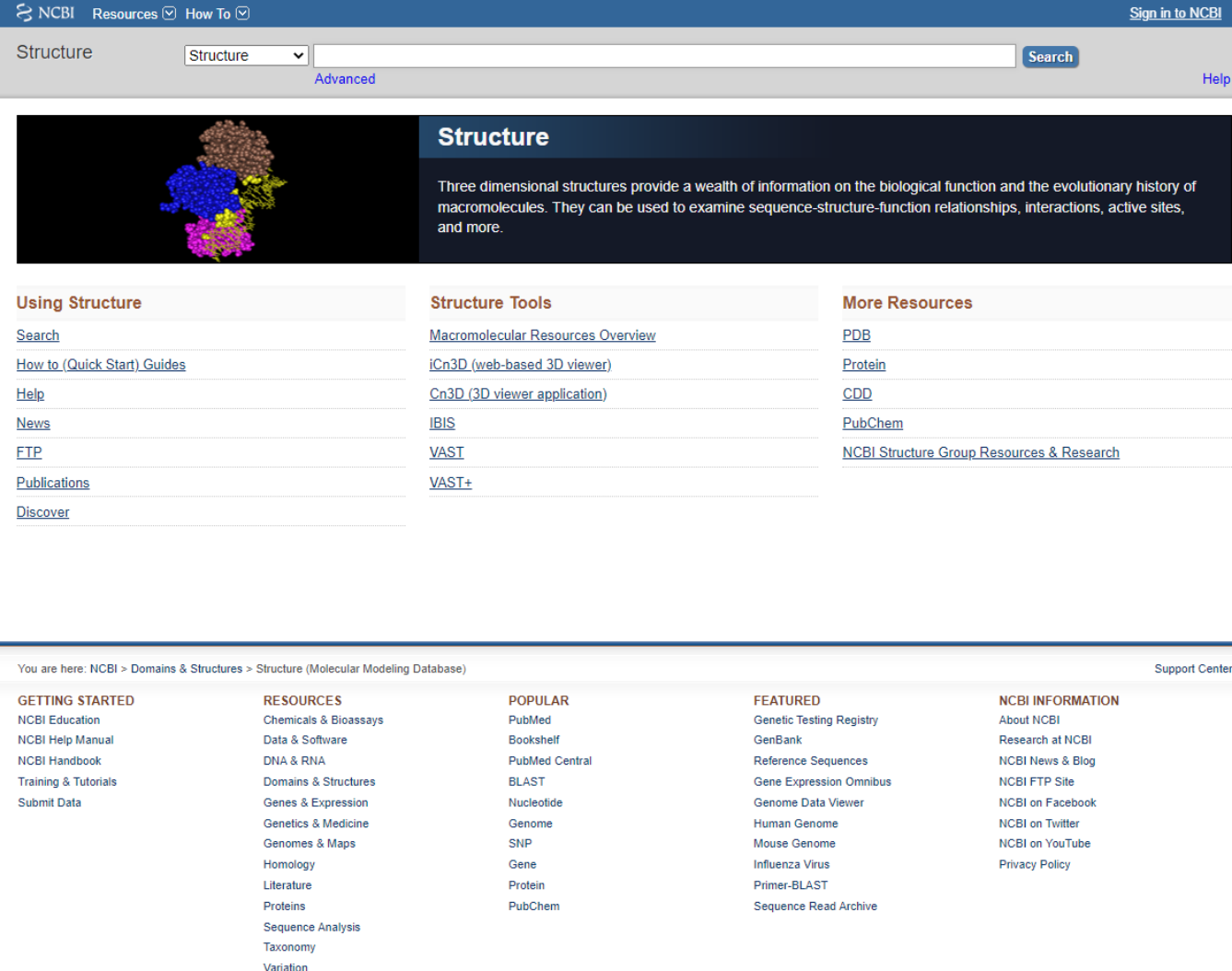
[7BQY](#), [6LU7](#)

PubMed Abstract:

A new coronavirus, known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is the aetiological agent responsible for the 2019-2020 viral pneumonia outbreak of coronavirus disease 2019 (COVID-19)¹⁻⁴. Currently, there are no targeted therapeutic agents for the treatment of this disease, and effective treatment options remain very limited ... 

NCBI's Structure Database

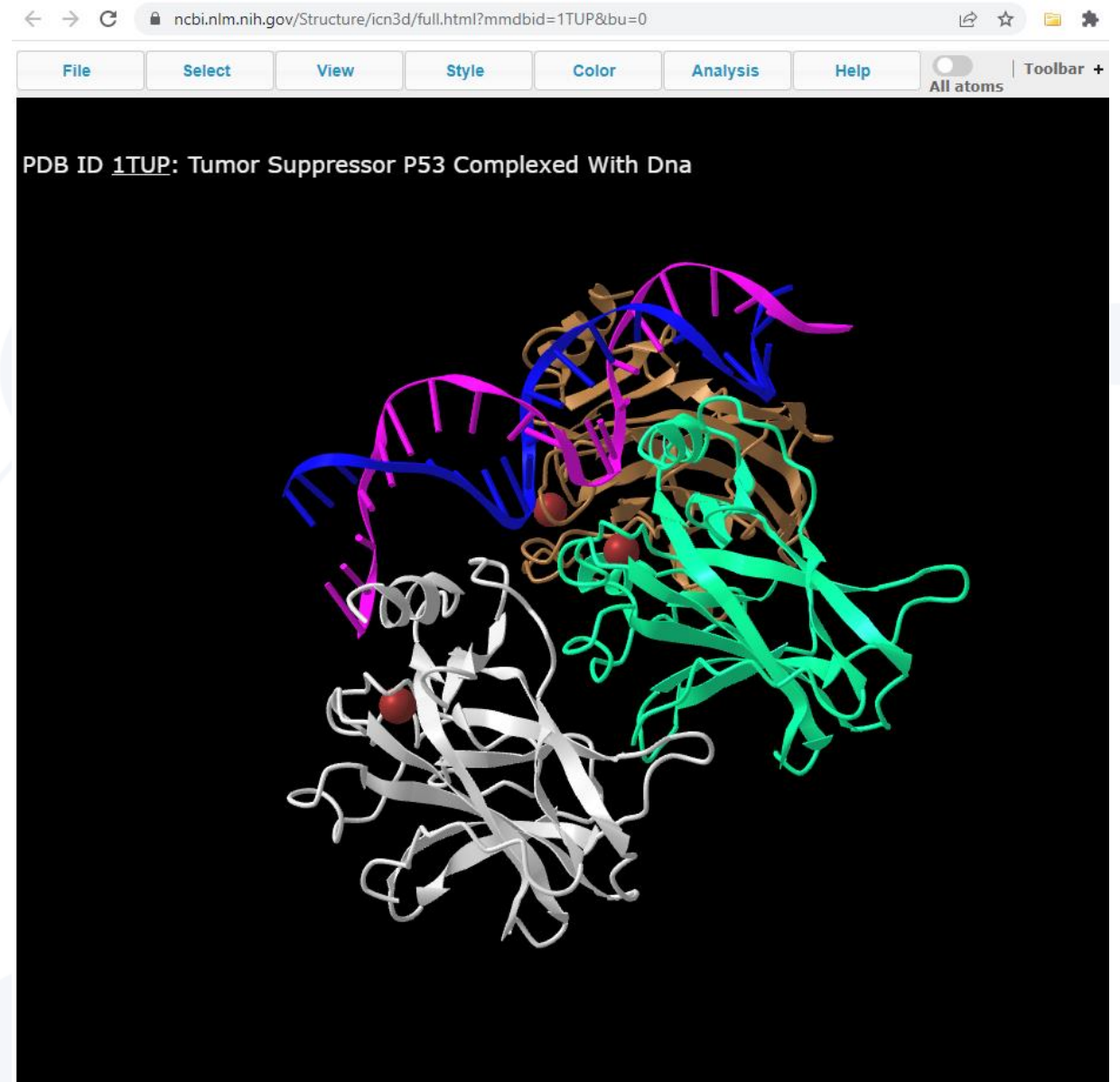
- Updated monthly
- Derived from PDB records
- Additional information added, including:
 - Explicit chemical graph information
 - Validation (secondary structure elements)
 - Includes taxonomy
- Connects 3D to associated literature, molecular data, chemical data, and other NCBI tools



The screenshot displays the NCBI Structure Database homepage. At the top, there is a navigation bar with links for 'NCBI', 'Resources', and 'How To', along with a 'Sign in to NCBI' link. Below this is a search bar with a dropdown menu set to 'Structure' and a 'Search' button. A 'Help' link is also present. The main content area features a large image of a protein structure on the left and a text box on the right stating: 'Three dimensional structures provide a wealth of information on the biological function and the evolutionary history of macromolecules. They can be used to examine sequence-structure-function relationships, interactions, active sites, and more.' Below this, there are three columns of links: 'Using Structure' (Search, How to (Quick Start) Guides, Help, News, FTP, Publications, Discover), 'Structure Tools' (Macromolecular Resources Overview, iCn3D (web-based 3D viewer), Cn3D (3D viewer application), IBIS, VAST, VAST+), and 'More Resources' (PDB, Protein, CDD, PubChem, NCBI Structure Group Resources & Research). At the bottom, there is a footer section with five columns of links: 'GETTING STARTED' (NCBI Education, NCBI Help Manual, NCBI Handbook, Training & Tutorials, Submit Data), 'RESOURCES' (Chemicals & Bioassays, Data & Software, DNA & RNA, Domains & Structures, Genes & Expression, Genetics & Medicine, Genomes & Maps, Homology, Literature, Proteins, Sequence Analysis, Taxonomy, Variation), 'POPULAR' (PubMed, Bookshelf, PubMed Central, BLAST, Nucleotide, Genome, SNP, Gene, Protein, PubChem), 'FEATURED' (Genetic Testing Registry, GenBank, Reference Sequences, Gene Expression Omnibus, Genome Data Viewer, Human Genome, Mouse Genome, Influenza Virus, Primer-BLAST, Sequence Read Archive), and 'NCBI INFORMATION' (About NCBI, Research at NCBI, NCBI News & Blog, NCBI FTP Site, NCBI on Facebook, NCBI on Twitter, NCBI on YouTube, Privacy Policy). A 'Support Center' link is located in the bottom right corner.

iCn3D

- Interactive, web-based 3D structure viewer
 - No installation needed!
- Users can
 - Visualize structure in 1D, 2D, and 3D
 - View sequence and structure alignments
 - Probe perturbations
 - And more!



3D Viewer Feature Comparison

| | Web- based | 1D Sequence | 2D Diagram | Annotation | Align | Share Link | Script | Jupyter Notebook | Virtual Reality | 3D Printing |
|----------------|---------------|----------------|---------------|------------|----------------|----------------|----------------|---------------------|--------------------|----------------|
| iCn3D | ✓ | ✓ | ✓ | ✓ | ✓ ^a | ✓ ^b | ✓ ^c | ✓ ^d | ✓ | ✓ |
| Mol* | ✓ | ✓ | Web | Web | | | | | | |
| Aquaria | ✓ | ✓ | | ✓ | | | | | ✓ | |
| Chimera | | ✓ | | ✓ | | | | | ✓ | ✓ |
| PyMol | | ✓ | | ✓ | | | ✓ | | | |
| Cn3D | | ✓ | Web | ✓ | ✓ | | | | | |

^a: iCn3D aligns structures (PDB or AlphaFold) based on structures or sequences.

^b: iCn3D sharable links could be a [short URL](#) or a URL containing the [address of an iCn3D PNG Image](#)

^c: iCn3D supports command-line analysis with either [Python scripts](#) or [Node.js scripts](#)

^d: iCn3D can also be [used in Jupyter Notebook](#)

iCn3D Features of Interest

- iCn3D aligns structures (PDB or AlphaFold) based on structures or sequences.
- iCn3D sharable links (<https://structure.ncbi.nlm.nih.gov/icn3d/share.html?XCxR6fSTmXHxR3o1A>)
- iCn3D supports command-line analysis with either [Python scripts](#) or [Node.js scripts](#)
- iCn3D can also be used in Jupyter Notebook (<https://pypi.org/project/icn3dpy>)
- 3D printing: structure.ncbi.nlm.nih.gov/icn3d/share.html?wt4TDqzhC2rhCYTD7
- Contact map: structure.ncbi.nlm.nih.gov/icn3d/share.html?rnMbe26tNsAjJLGK9
- Precalculated symmetry: structure.ncbi.nlm.nih.gov/icn3d/share.html?bGH1BfLsiGFhhTDn8
- Symmetry dynamically: structure.ncbi.nlm.nih.gov/icn3d/share.html?6NvhQ45XrnbuXyGe6
- Electron density map: structure.ncbi.nlm.nih.gov/icn3d/share.html?QpqNZ3k65ToYFvUB6
- EM map: structure.ncbi.nlm.nih.gov/icn3d/share.html?L4C4WYE85tYRiFeK7
- Transmembrane protein: structure.ncbi.nlm.nih.gov/icn3d/share.html?jMN16mJyR9STUx6E6
- Solvent Accessible Area: structure.ncbi.nlm.nih.gov/icn3d/share.html?xKSyfd1umbKstGh29

iCn3D Shortcuts

Rotate

- **Left mouse button** can be used to rotate the structure
- **Key L** - left
- **Key J** - right
- **Key I** - up
- **Key M** - down
- **Shift + Key L** - left 90°
- **Shift + Key J** - right 90°
- **Shift + Key I** - up 90°
- **Shift + Key M** - down 90°

Zoom

- **Middle mouse button** OR **Left Mouse + Shift** - can be used to zoom
- **Key Z** - zoom in
- **Key X** - zoom out

Translate

- **Right mouse button** OR **Left Mouse + Ctrl** - can be used to translate the structure to a different location within the 3D window
- **Keyboard arrows**

Select

- **Alt + Click (PC)** or **Option + Click (Mac)**- can be used to select atom/residue/strand , hold **Ctrl + Click** to add another

A faint, light blue molecular structure graphic is visible in the background, consisting of several interconnected circles (nodes) of varying sizes, representing atoms and bonds in a chemical network.

iCn3D Fundamentals Demo

<https://bit.ly/3Xui4qr>

Group Exercises

- We will now split into groups to work on one of the three examples:

- [Example 1](#): TP53 Mutation Analysis
- [Example 2](#): TP53 from Structure to Function
- [Example 3](#): Compare Crystal and AlphaFold TP53 Structures

Review the objectives on each of these pages and decide on an example that best suites your learning goals. Work with your group and NCBI experts on completing the steps and rendering an image(s) with iCn3D.

Group Discussion & Event Wrap-up

Example 1: TP53 Mutation Analysis

- **P53-DNA Binding:** Positively charged residues (like Lysine) in P53 interact with negatively charged DNA backbone.
- **Mutation Analysis:** Mutation K120A in TP53 disrupts interactions with DNA, potentially weakening binding.
- **ClinVar Feature:** Identified pathogenic mutations (K120) in P53 linked to Li-Fraumeni-like syndrome. Loss of interactions in mutants suggests K120 is critical for function.

Group Discussion & Event Wrap-up

Example 2: TP53 from Structure to Function

- **Structure:** TP53 is a high ordered structure, and this structure is important for biological function.
- **Interactions:** Charged and polar residues likely mediate DNA and protein binding.
- **DNA Binding:** Positively charged amino acids interact with negatively charged DNA backbone.
- **Zinc Binding:** Functional region overlaps with protein dimerization, suggesting zinc binding plays a role in dimerization

Group Discussion & Event Wrap-up

Example 3: Compare Crystal and AlphaFold TP53 Structures

- **AlphaFold Limitations:** Predicts monomers only (not dimers/tetramers crucial for TP53 function).
- **Structure Comparison:** pLDDT scores highlight regions of varying confidence (loops tend to be lower). Alignment metrics (RMSD, TM-score) indicate overall structural similarity.

Group Discussion & Event Wrap-up

- In your chosen example, what did you find most challenging to understand?
- Were there any specific features of iCn3D you struggled with?
- Based on these exercises, what specific questions do you still have about iCn3D?

Continue learning about iCn3D

Tutorials and help documents are available [here](#):

The screenshot displays the iCn3D web interface, which is part of the U.S. National Library of Medicine and NCBI National Center for Biotechnology Information. The interface features a top navigation bar with the NIH logo and the text "U.S. National Library of Medicine" and "NCBI National Center for Biotechnology Information". Below this, the "iCn3D" logo is prominently displayed, followed by the text "AlphaFold-related gallery with live examples".

A central menu is visible, listing options: "About iCn3D", "Live Gallery", "Tutorial >" (highlighted), "Search Structure", "Citing iCn3D", "Source Code >", "Develop >", and "Help Doc". To the right of the menu, a sub-menu lists "Use iCn3D", "iCn3D Videos", "URL Parameters", and "Commands".

The main content area is divided into two panels, each showing a protein structure visualization and its corresponding "Sequences and Annotations" panel.

Left Panel: The protein structure is labeled "UniProt ID A0A044R7Z7: ALPHAFOLD MONOMER V2". The structure is shown in a ribbon representation with various colors. The "Sequences and Annotations" panel for this protein includes a "Summary" tab and a "Details" tab. Under "Annotations", the "Conserved Domains" checkbox is checked, and the "3D Domains" checkbox is also checked. The "Proteins" section lists the protein "A0A044R7Z7_A" with a length of 479 residues. It shows domain annotations: "domain: PG3" (176 Res), "3D domain 1 of A0..." (188 Res), "3D domain 2 of A0..." (250 Res), and "3D domain 3 of A0..." (42 Res).

Right Panel: The protein structure is labeled "AlphaFold UniProt ID Q08426: ALPHAFOLD MONOMER V2". The structure is shown in a ribbon representation with various colors. The "Sequences and Annotations" panel for this protein includes a "Summary" tab and a "Details" tab. Under "Annotations", the "Conserved Domains" checkbox is checked, and the "ClinVar" checkbox is also checked. The "Proteins" section lists the protein "Q08426_A" with a length of 52 residues. It shows domain annotations: "domain: fcdJ" (52 Res) and "3D domain 1 of Q08426_A" (52 Res). A legend indicates the confidence of the AlphaFold prediction: "Very high (pLDDT > 90)", "Confident (90 > pLDDT > 70)", "Low (70 > pLDDT > 50)", and "Very low (pLDDT < 50)".

Below the left panel, a caption reads: "AlphaFold structures with conserved domain and 3D domain annotations (Uniprot ID A0A044R7Z7)". Below the right panel, a caption reads: "AlphaFold structures with SNP and ClinVar annotations (Uniprot ID Q08426)".

Continue learning about NCBI Resources

- Join us for workshops, webinars, or codeathons!

[NCBI Insights Blog](#)

- Follow us on social media:



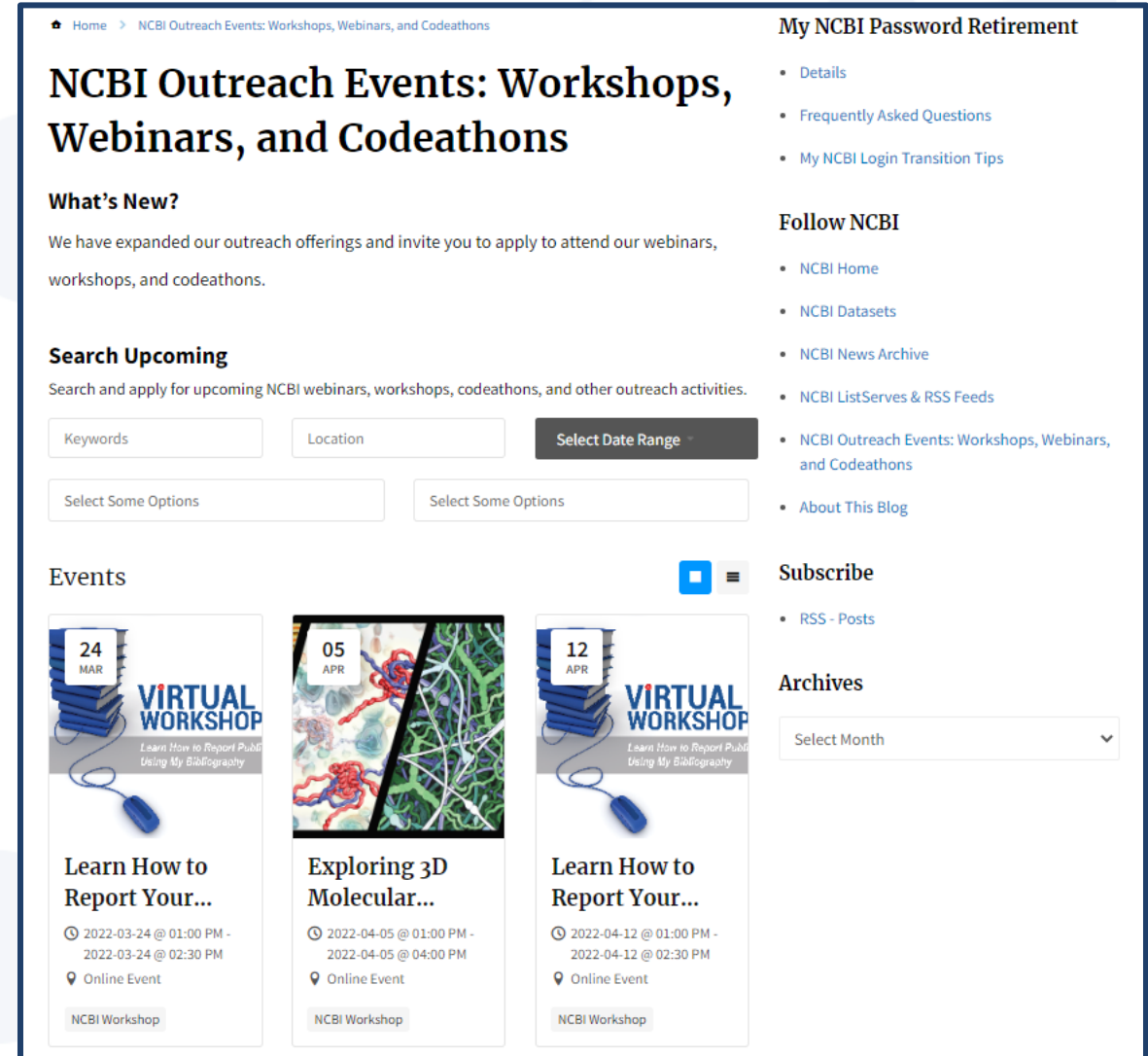
[X](#)



[LinkedIn](#)



[Facebook](#)



Home > NCBI Outreach Events: Workshops, Webinars, and Codeathons

NCBI Outreach Events: Workshops, Webinars, and Codeathons

What's New?
We have expanded our outreach offerings and invite you to apply to attend our webinars, workshops, and codeathons.

Search Upcoming
Search and apply for upcoming NCBI webinars, workshops, codeathons, and other outreach activities.

Keywords Location Select Date Range

Select Some Options Select Some Options

Events

- 24 MAR**
VIRTUAL WORKSHOP
Learn How to Report Your...
2022-03-24 @ 01:00 PM - 2022-03-24 @ 02:30 PM
Online Event
NCBI Workshop
- 05 APR**
Exploring 3D Molecular...
2022-04-05 @ 01:00 PM - 2022-04-05 @ 04:00 PM
Online Event
NCBI Workshop
- 12 APR**
VIRTUAL WORKSHOP
Learn How to Report Your...
2022-04-12 @ 01:00 PM - 2022-04-12 @ 02:30 PM
Online Event
NCBI Workshop

My NCBI Password Retirement

- Details
- Frequently Asked Questions
- My NCBI Login Transition Tips

Follow NCBI

- NCBI Home
- NCBI Datasets
- NCBI News Archive
- NCBI ListServes & RSS Feeds
- NCBI Outreach Events: Workshops, Webinars, and Codeathons
- About This Blog

Subscribe

- RSS - Posts

Archives

Select Month



Exploring Biomolecular Structures with NCBI's iCn3D Supplemental Learning Materials

Alexa M. Salsbury, Ph.D.

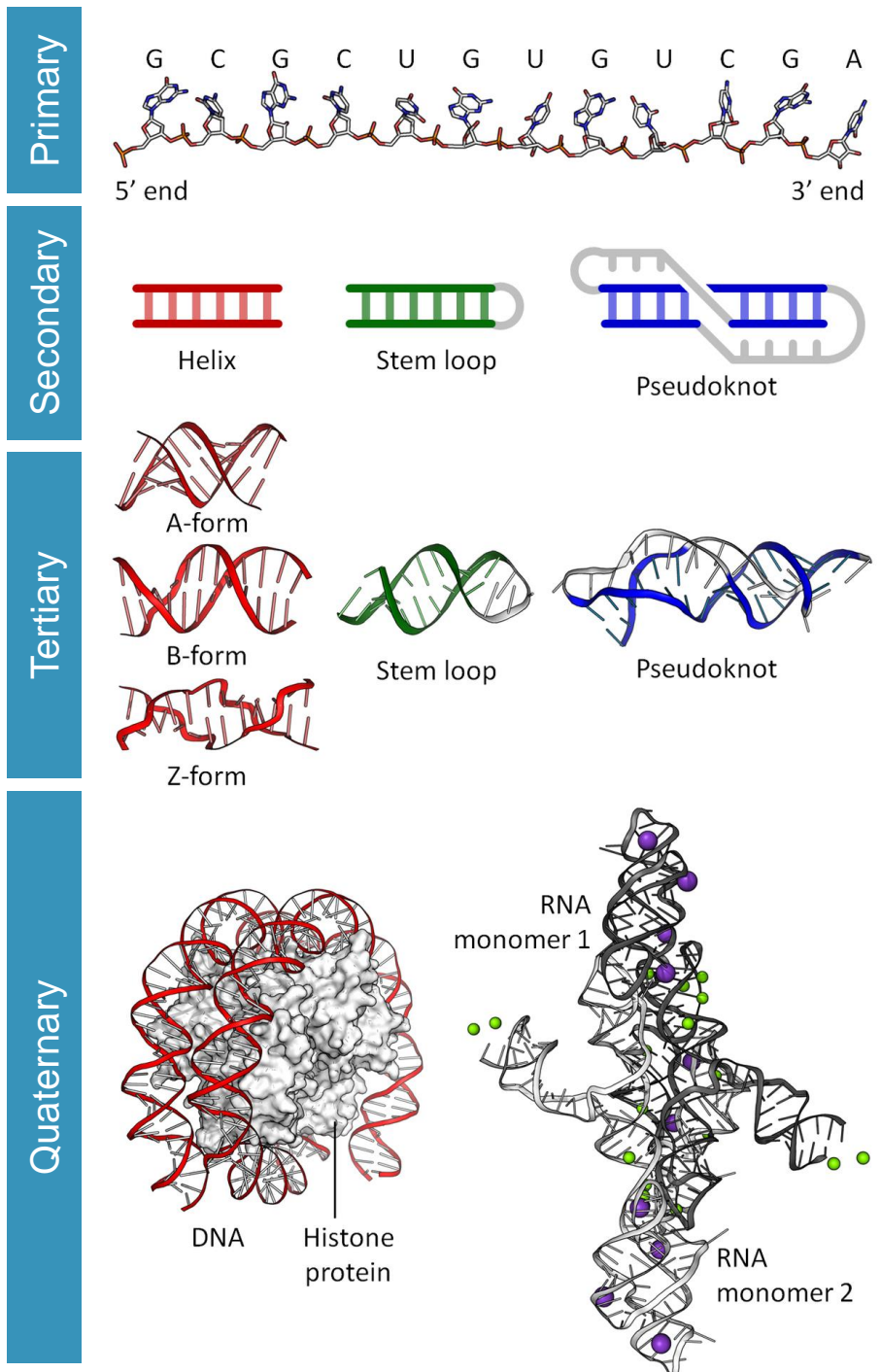


National Library of Medicine
National Center for Biotechnology Information



Nucleic Acid Structure

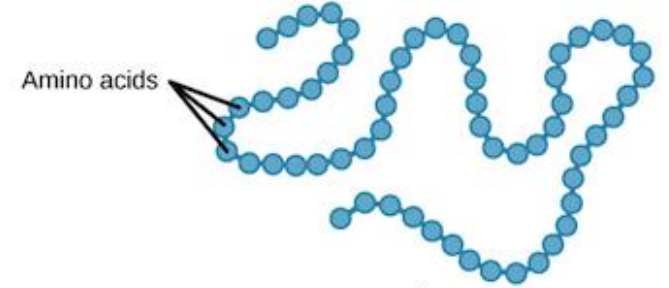
- **Primary**- sequence of nucleotides
- **Secondary**- base pairing interactions between polymers (DNA) or within a single polymer (RNA)
- **Tertiary**- 3D folding pattern
- **Quaternary**- interactions of nucleic acids with other molecules (DNA, RNA, or Protein)



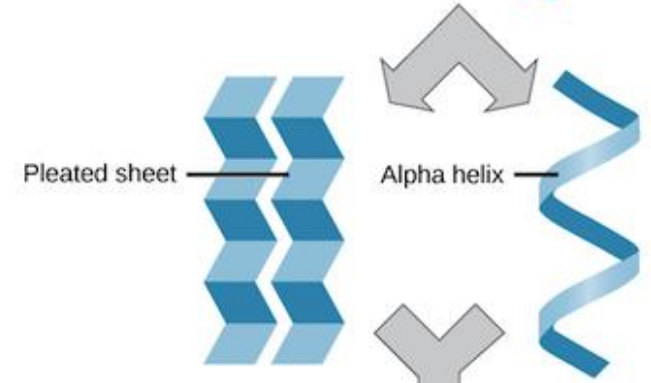
Protein Structure

- **Primary**- sequence of amino acids
- **Secondary**- hydrogen bonding of the peptide backbone that causes amino acids to fold into a repeating pattern
- **Tertiary**- 3D folding pattern of a protein due to side chain interactions
- **Quaternary**- protein consisting of more than one polypeptide

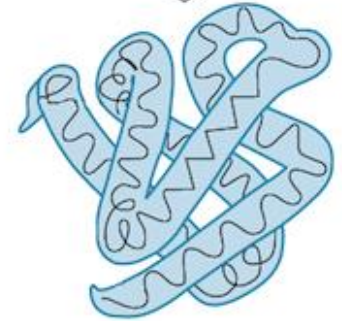
Primary



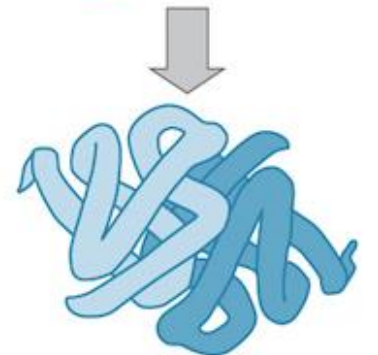
Secondary



Tertiary



Quaternary



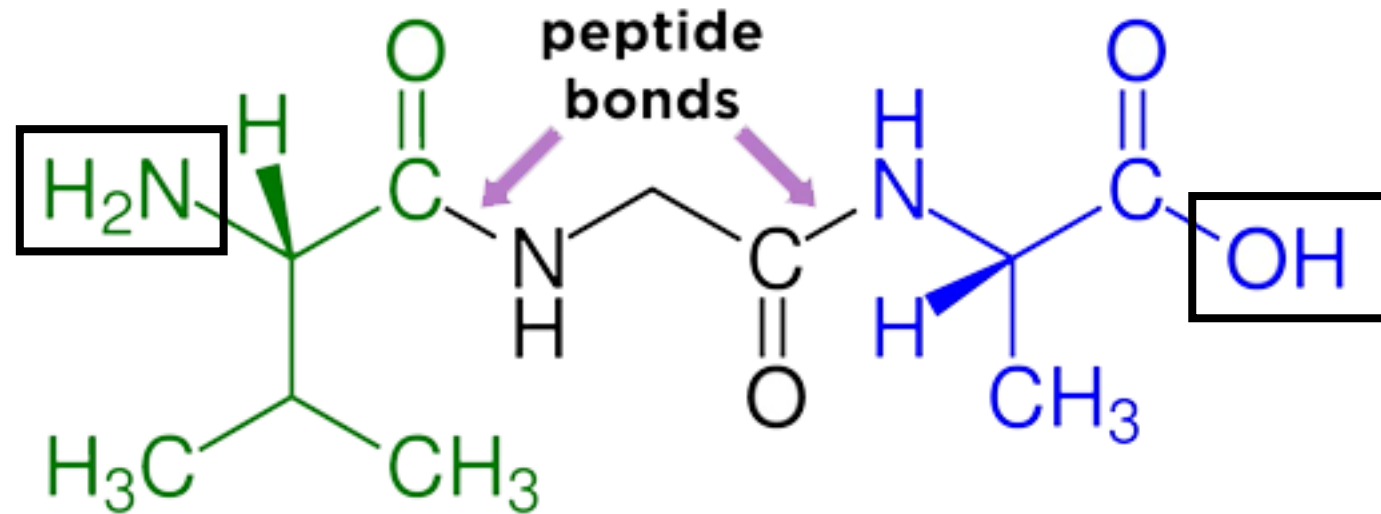
Terminology

N-terminus

(ends in amino group)

C-terminus

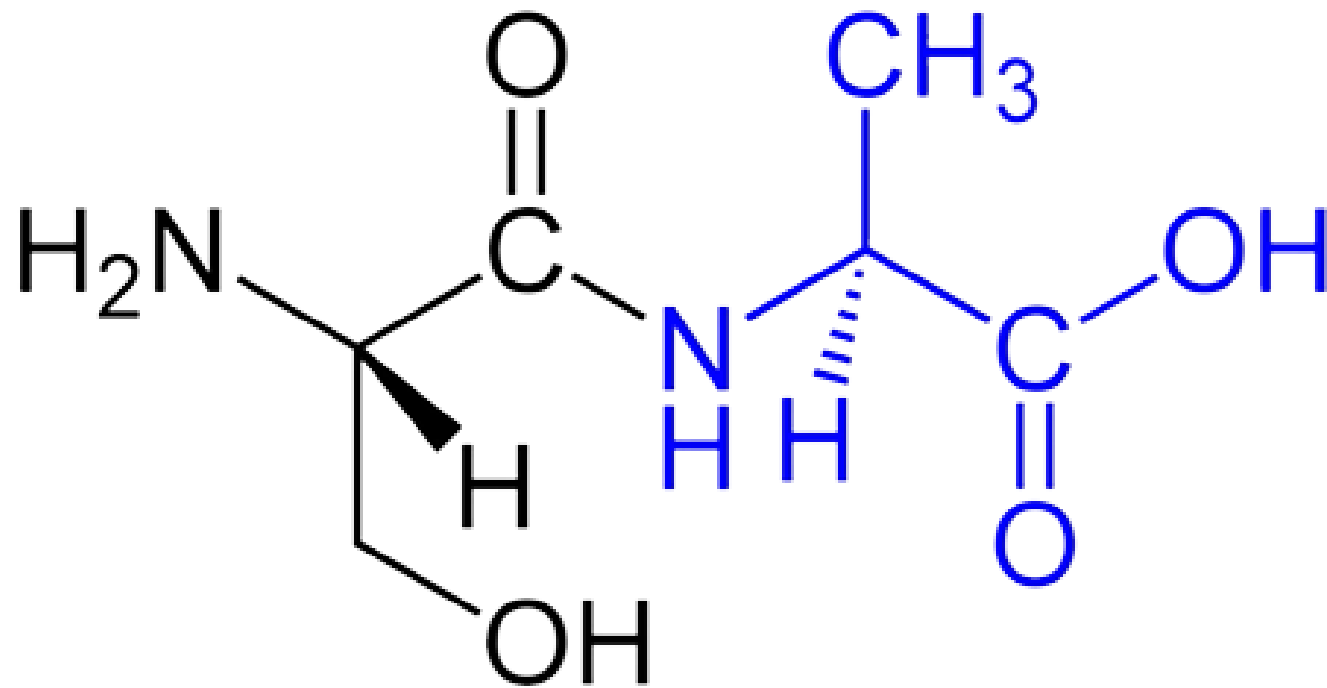
(ends in carboxyl group)



valine-glycine-alanine

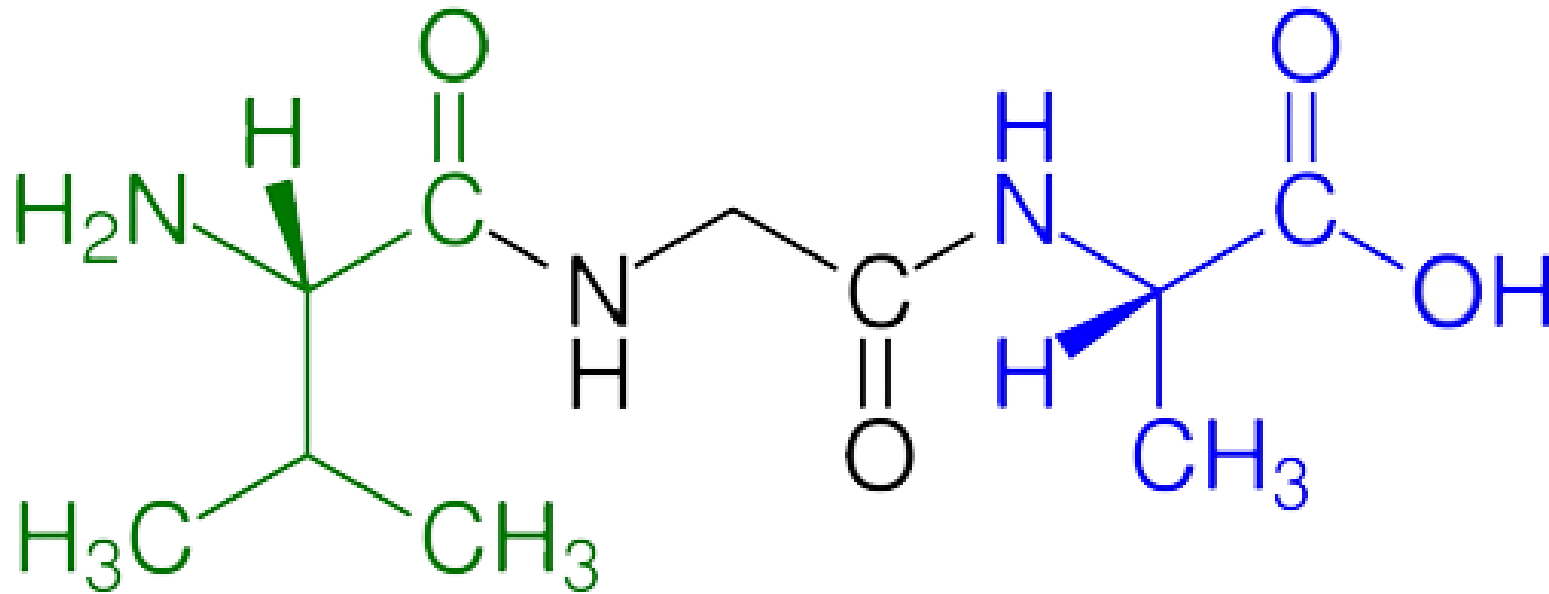
Terminology

dipeptide (2 amino acids)



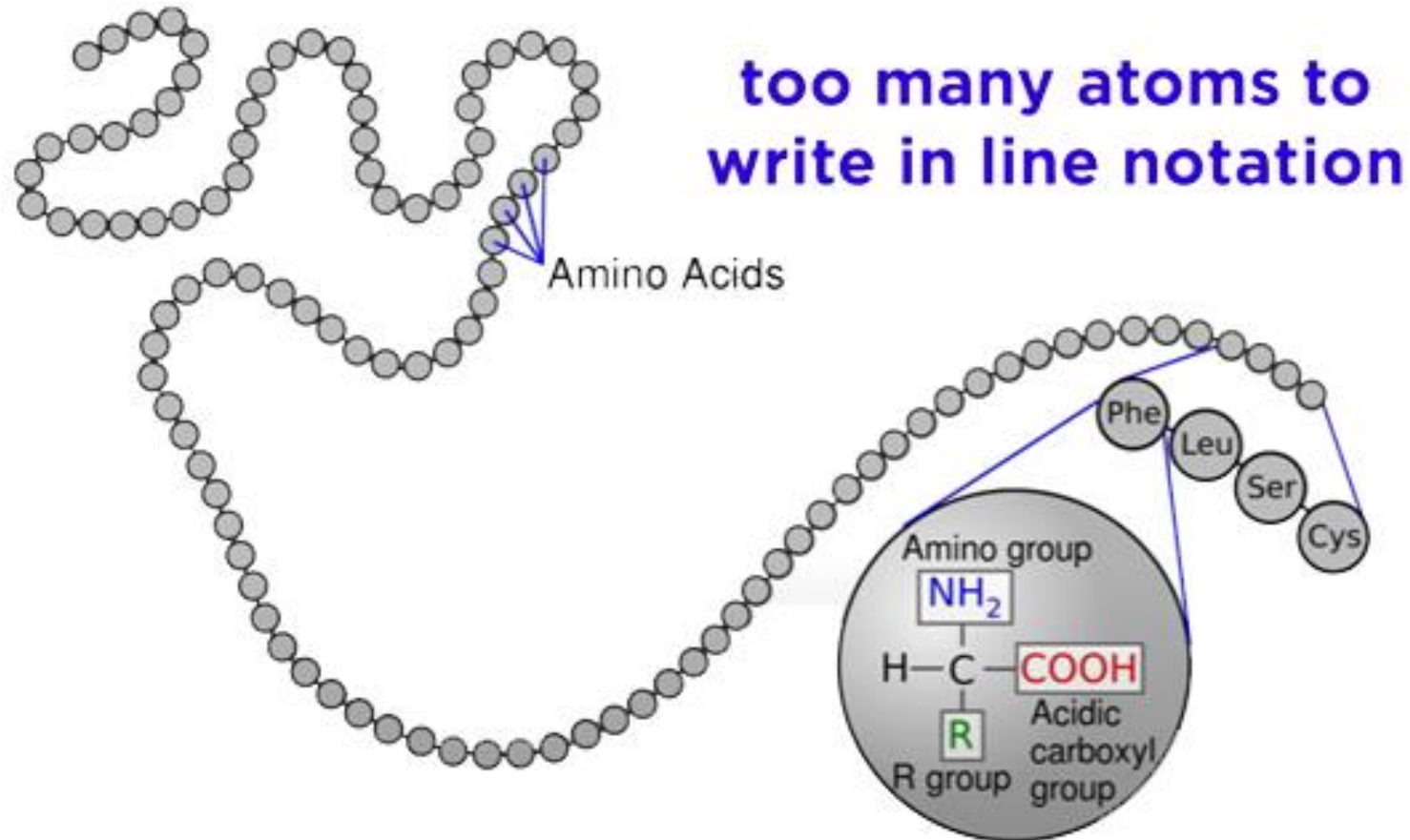
Terminology

oligopeptide (3-10 amino acids)



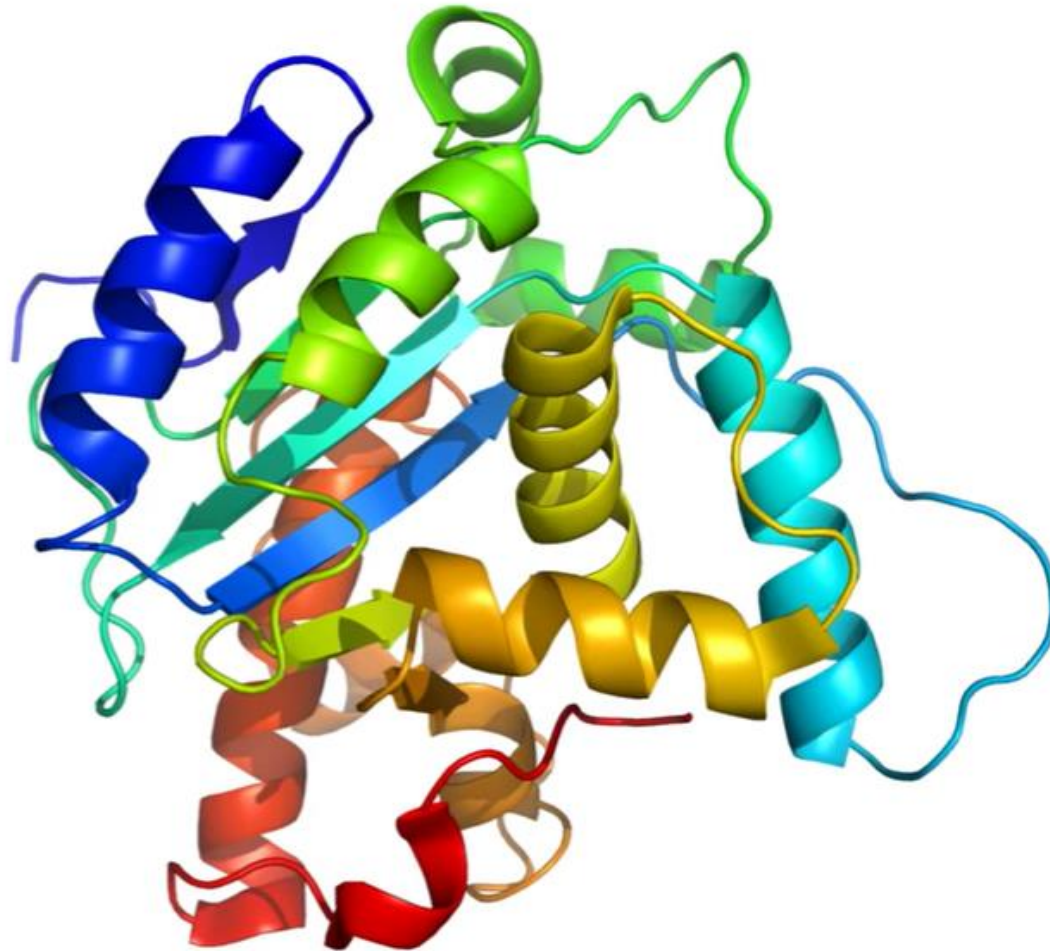
Terminology

polypeptide (>10 amino acids)

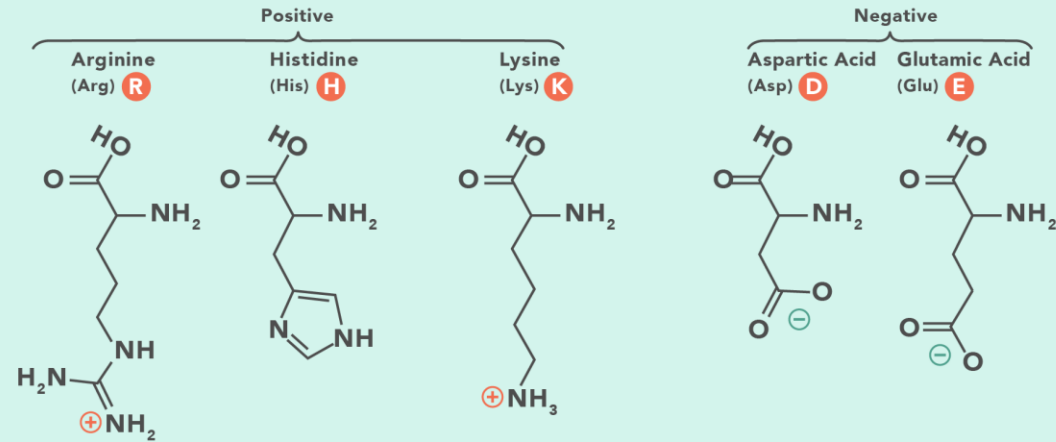


Terminology

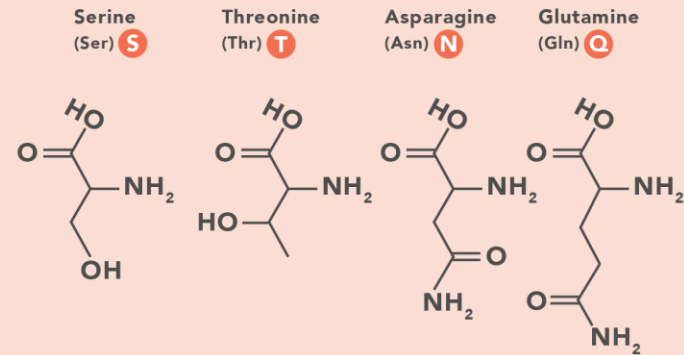
protein (generally 300-1000 amino acids)



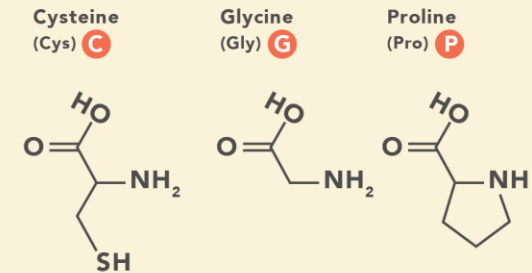
A. Amino Acids with Electrically Charged Side Chains



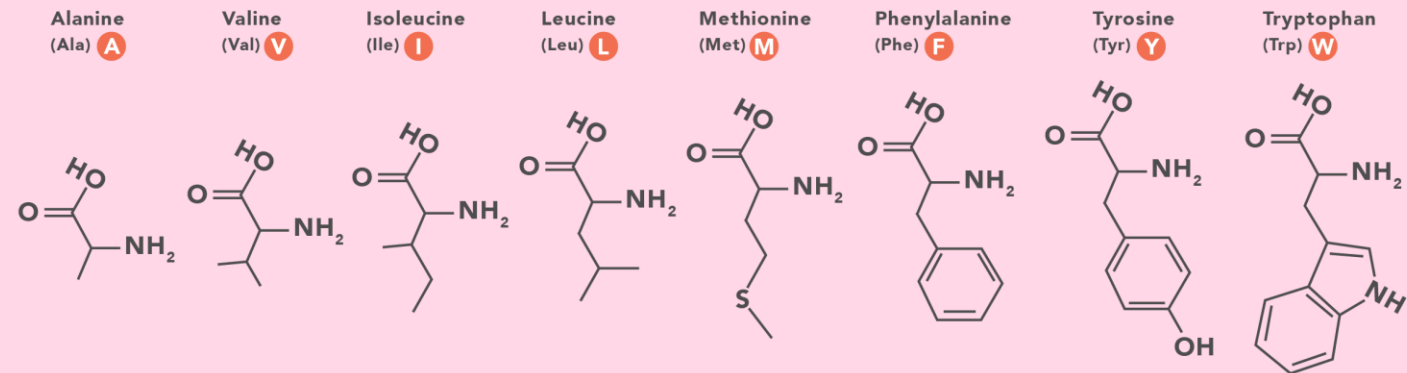
B. Amino Acids with Polar Uncharged Side Chains



C. Special Cases



D. Amino Acids with Hydrophobic Side Chains



Functional definition:

- Enzymes: Accelerate biochemical reactions
- Structural: Form biological structures
- Transport: Carry biochemically important substances
- Defense: Protect the body from foreign invaders

Structural definition:

- Globular: Complex folds, irregularly shaped tertiary structures
- Fibrous: Extended, simple folds -- generally structural proteins

Cellular localization definition:

- Membrane: In direct physical contact with a membrane; generally water insoluble.
- Soluble: Water soluble; can be anywhere in the cell

Experimental techniques



■ Single crystal X-ray diffraction (SC-XRD)

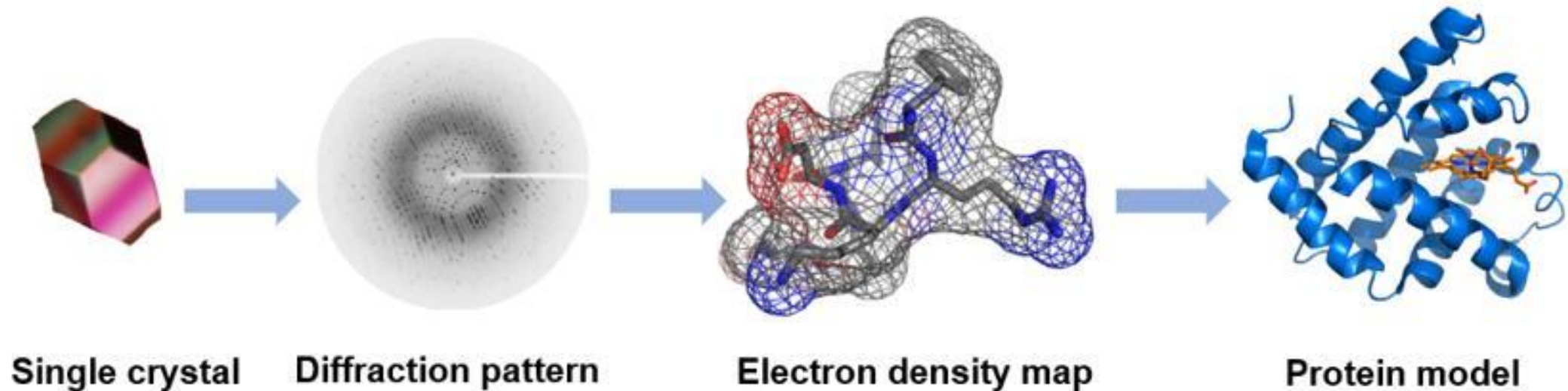
■ Nuclear magnetic resonance (NMR)

■ Cryo-electron microscopy (Cryo-EM)

Three main research techniques for structural biology.
According to the statistics of PDB (<https://www.rcsb.org/>)

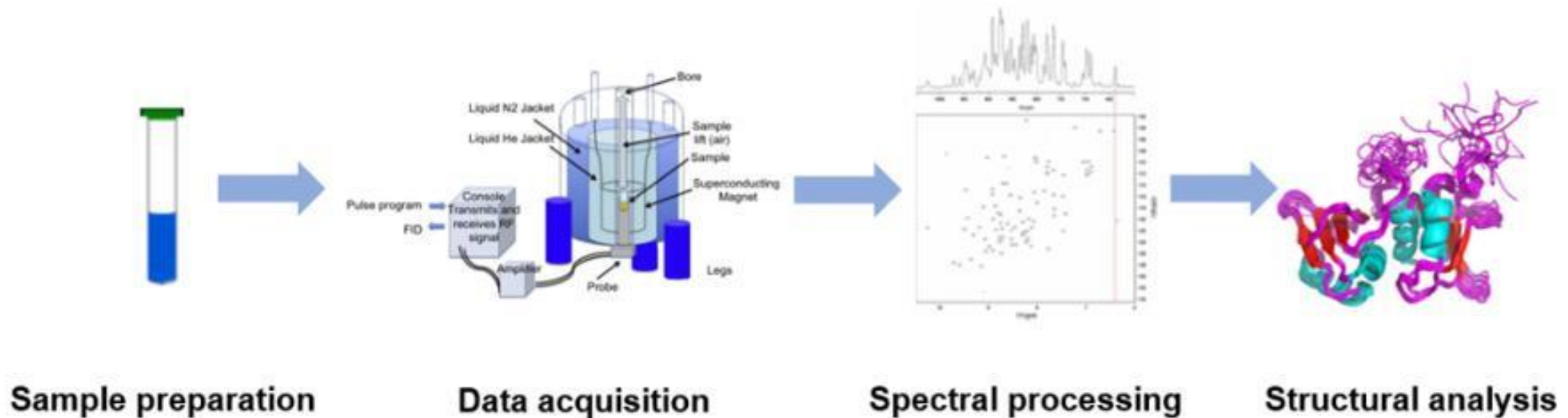
X-ray Crystallography

- Requires crystals, which can be hard to make
- Can handle very large proteins and complexes (e.g. ribosome)
- Provides a “flash picture” with little or no data about motions
- Can include packing artifacts from crystallization



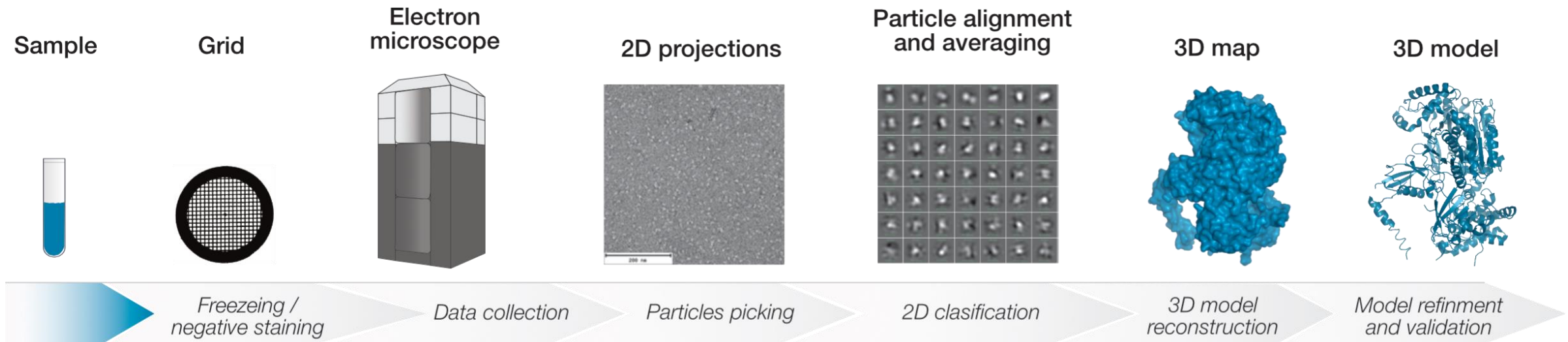
Nuclear Magnetic Resonance

- Requires highly concentrated, C13/N15-labeled protein solutions
- Limited to relatively small proteins (<30 kDa)
- Sensitive to molecular motions
- High protein concentrations may induce non-biological binding



Cryo-electron microscopy

- Requires expensive equipment
- Only small amount of sample
- Rapid freezing sample allows sample to maintain a closer-to-native state
- Useful for biomolecules with high molecular weight



NCBI Structure Database Search Tips

Entrez is a molecular biology database system that provides access to a wealth of NCBI data

- More [Entrez Help](#) is available on the NCBI website

Finding structures with Entrez

```
"term1"[field1] AND/OR/NOT "term2"[field2] AND/OR/NOT ...
```

- Use field limits and Boolean operators
- Put phrases in quotes

NCBI Structure Database Search Examples

Useful Search Fields

Organism

Ex. "Homo sapiens"[orgn]

Experimental Method

Ex. "NMR"[exp]

Chemical Name

"zinc"[chemical name]

PDB Description

Ex. "Tumor Suppressor
p53"[title]

[Filter]

Ex. "Complex DNA"[filter]

[More Search Field Options](#)

```
term1[field1] AND/OR/NOT term2[field2] AND/OR/NOT ...
```

```
"Homo sapiens"[orgn] AND "X-ray  
diffraction"[exp]
```

Search results

Items: 1 to 20 of 47803

```
"Homo sapiens"[orgn] AND "X-ray  
diffraction"[exp] AND "zinc"[chemical name]
```

Search results

Items: 1 to 20 of 6092

```
"Homo sapiens"[orgn] AND "X-ray  
diffraction"[exp] AND "zinc"[chemical  
name] AND "Complex DNA"[filter]
```

Search results

Items: 1 to 20 of 288

```
1TUP
```

PDB File

| | | | |
|--------|--|--|-----------------|
| HEADER | ISOMERASE/DNA | 04-OCT-07 | 2RGR |
| TITLE | TOPOISOMERASE IIA BOUND TO G-SEGMENT DNA | | |
| COMPND | MOL_ID: 1; | | |
| COMPND | 2 MOLECULE: DNA TOPOISOMERASE 2; | | |
| COMPND | 3 CHAIN: A; | | |
| COMPND | 4 FRAGMENT: DNA BINDING AND CLEAVAGE DOMAIN (RESIDUES 419- | | |
| COMPND | 5 1177); | | |
| COMPND | 6 SYNONYM: DNA TOPOISOMERASE II; | | |
| COMPND | 7 EC: 5.99.1.3; | | |
| COMPND | 8 ENGINEERED: YES; | | |
| COMPND | 9 MOL_ID: 2; | | |
| COMPND | 10 MOLECULE: DNA; | | |
| COMPND | 11 CHAIN: C; | | |
| COMPND | 12 ENGINEERED: YES; | | |
| COMPND | 13 MOL_ID: 3; | | |
| COMPND | 14 MOLECULE: DNA; | | |
| COMPND | 15 CHAIN: D; | | |
| COMPND | 16 ENGINEERED: YES | | |
| SOURCE | MOL_ID: 1; | | |
| SOURCE | 2 ORGANISM_SCIENTIFIC: SACCHAROMYCES CEREVISIAE; | | |
| SOURCE | 3 ORGANISM_COMMON: BAKER'S YEAST; | | |
| SOURCE | 4 ORGANISM | | |
| SOURCE | 5 GENE: | | |
| SOURCE | 6 EXPRES | | |
| SOURCE | 7 EXPRES | | |
| SOURCE | 8 EXPRES | | |
| SOURCE | 9 EXPRES | | |
| SOURCE | 10 EXPRES | | |
| SOURCE | 11 EXPRES | | |
| SOURCE | 12 MOL_ID | | |
| SOURCE | 13 SYNTHETIC | | |
| SOURCE | 14 MOL_ID | | |
| SOURCE | 15 SYNTHETIC | | |
| REMARK | 2 | | |
| REMARK | 2 | RESOLUTION. | 3.00 ANGSTROMS. |
| REMARK | 3 | | |
| REMARK | 3 | REFINEMENT. | |
| REMARK | 3 | PROGRAM | : PHENIX |
| REMARK | 280 | | |
| REMARK | 280 | CRYSTAL | |
| REMARK | 280 | SOLVENT CONTENT, VS (%) | : 59.90 |
| REMARK | 280 | MATTHEWS COEFFICIENT, VM (ANGSTROMS**3/DA) | : 3.07 |
| REMARK | 280 | | |
| REMARK | 280 | CRYSTALLIZATION CONDITIONS: 12-20% PEG 1000, 100-250 MM MGCL2, | |
| REMARK | 280 | 100 MM SODIUM CACODYLATE, PH 7.0, VAPOR DIFFUSION, HANGING | |
| REMARK | 280 | DROP, TEMPERATURE 277K | |
| REMARK | 290 | | |

PDB File: Data

| | | | | | | | | | | | | |
|--------|----|-----|-----|---|-----|--------|--------|---------|------|-------|-----|---|
| ATOM | 46 | N | THR | A | 424 | 17.857 | 11.128 | -95.193 | 1.00 | 57.50 | N | |
| ANISOU | 46 | N | THR | A | 424 | 7252 | 7857 | 6740 | -72 | 842 | 107 | N |
| ATOM | 47 | CA | THR | A | 424 | 16.610 | 11.906 | -95.142 | 1.00 | 56.73 | C | |
| ANISOU | 47 | CA | THR | A | 424 | 7153 | 7780 | 6623 | -85 | 799 | 141 | C |
| ATOM | 48 | C | THR | A | 424 | 15.300 | 11.135 | -94.959 | 1.00 | 55.10 | C | |
| ATOM | 49 | O | THR | A | 424 | 6950 | 7583 | 6224 | -111 | 717 | 159 | O |
| ATOM | 50 | CB | THR | A | 424 | 16.451 | 12.75 | 15.406 | 1 | 60.37 | 18 | C |
| ATOM | 50 | CB | THR | A | 424 | 7628 | 8271 | 7271 | -81 | 860 | 205 | C |
| ATOM | 51 | OG1 | THR | A | 424 | 16.550 | 11.914 | -97.557 | 1.00 | 63.33 | O | |
| AN | 51 | OG1 | THR | A | 424 | 8049 | 8666 | 7346 | -104 | 846 | 163 | C |
| AT | 51 | CG2 | THR | A | 424 | 17.523 | 13.836 | -96.462 | 1.00 | 61.80 | C | |
| AN | 51 | CG2 | THR | A | 424 | 77 | 7271 | -81 | 860 | 205 | C | |

Name

Atom Number

Atom Name

Chain ID

Residue Name

Residue Number

X

Y

Z

Occupancy

Temperature Factor

Computational Structural Biology

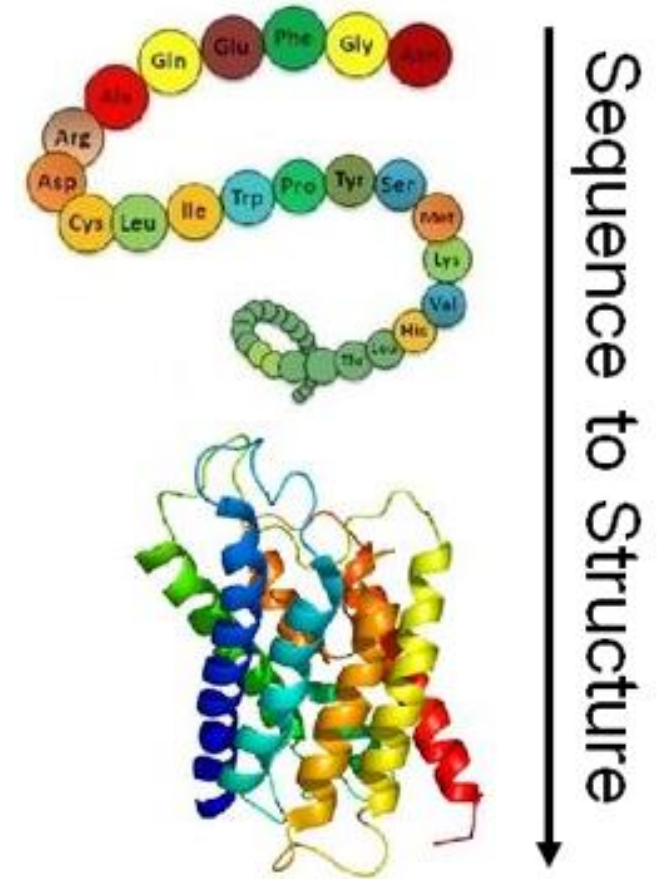
- Structure Prediction- inference of 3D structure from sequence data
- Molecular Docking- predicts the orientation of one molecule to another
- Molecular Dynamics Simulations- analyzes physical movements of atoms and molecules over time

Computational Structural Biology

- Structure Prediction- inference of 3D structure from sequence data
- Molecular Docking- predicts the orientation of one molecule to another
- Molecular Dynamics Simulations- analyzes physical movements of atoms and molecules over time
- Rely on experimental information from public databases
 - NCBI Databases and RCSB Protein Data Bank

Structure Prediction Methods

- Comparative Modeling
 - Prediction is based on amino acid sequence and structures of similar molecules available
- Fold recognition
 - Predicts folded structure by aligning a protein of **unknown** structure and a protein of **known structure** for low levels of sequence identity (<25%)
- Ab initio
 - Predicts the structure of proteins from the sequence and using molecular energy calculations (Schrodinger equation)



Homology Modeling vs *Ab initio* Prediction

| Ab initio Prediction | Comparative Modeling |
|--|---|
| Applicable to any sequence | Applicable to only those sequences with recognizable similarity to a template structure |
| Not very accurate ($>4\text{\AA}$ RMSD) | Fairly accurate ($<3\text{\AA}$ RMSD), similar to low resolution X-ray structure |
| Attempted for proteins of <100 residues | Not limited by size |
| Accuracy and applicability are limited by our understanding of the protein folding problem | Accuracy and applicability are limited by the number of known folds |