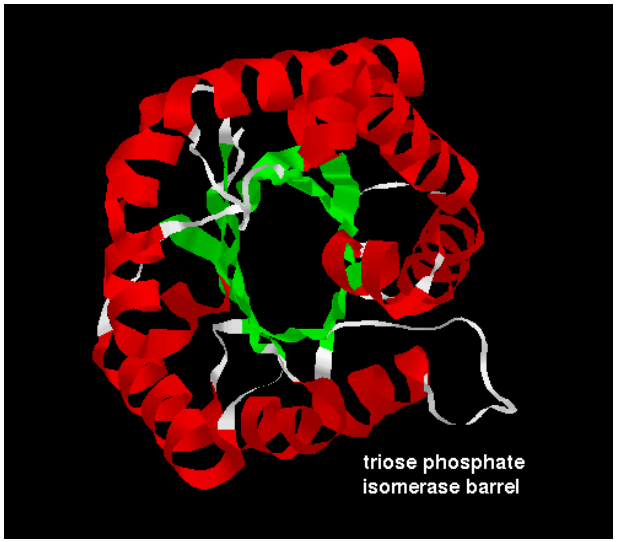
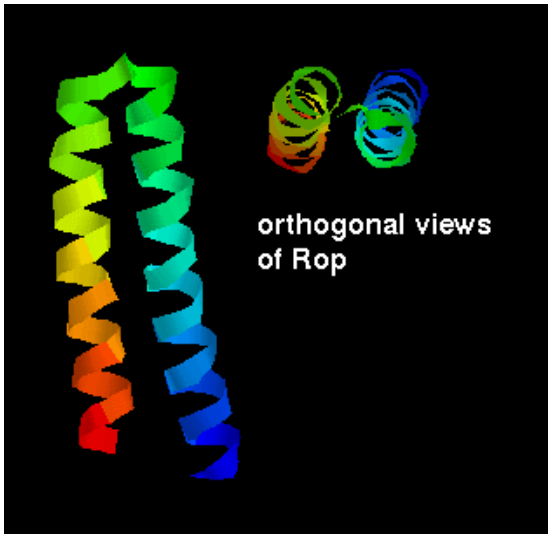
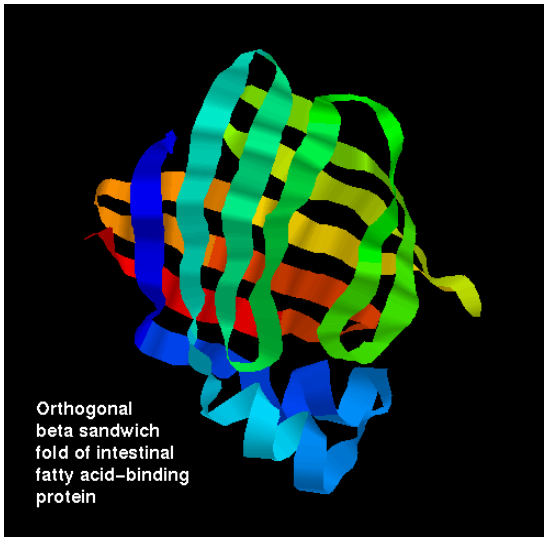
3D Structure
Determination & Data







David Wishart

Rm. 3-41 Athabasca Hall

david.wishart@ualberta.ca

Outline for Next 3 Weeks

- Introduction to protein structure
- Methods for visualizing structure
- Homology modeling
- Protein structure prediction
- Proteomics – part 1
- Proteomics – part 2

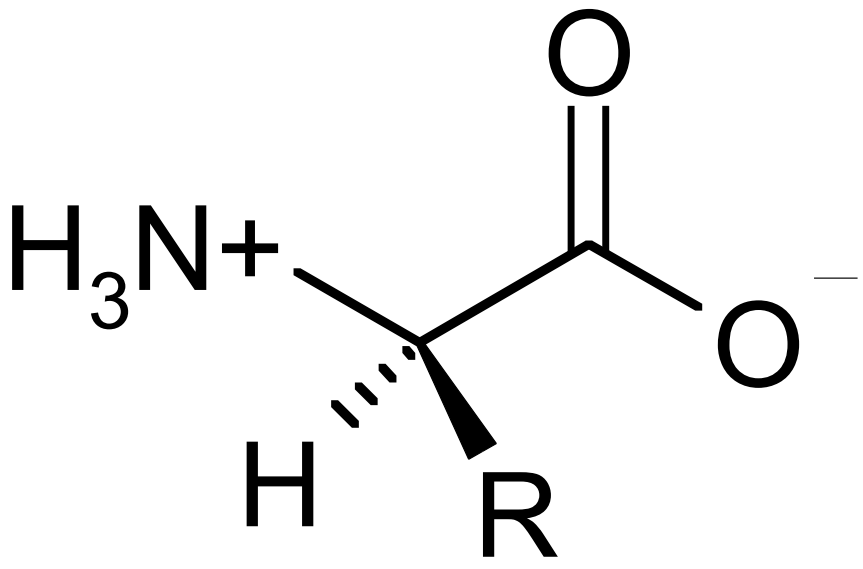
Objectives*

- Review of amino acid and polypeptide structure
- Gain and understanding of protein structure determination methods (NMR and X-ray)
- Learn about the Protein Data Bank
- Learn some of the approaches to Rendering, Modelling and Viewing 3D Structures

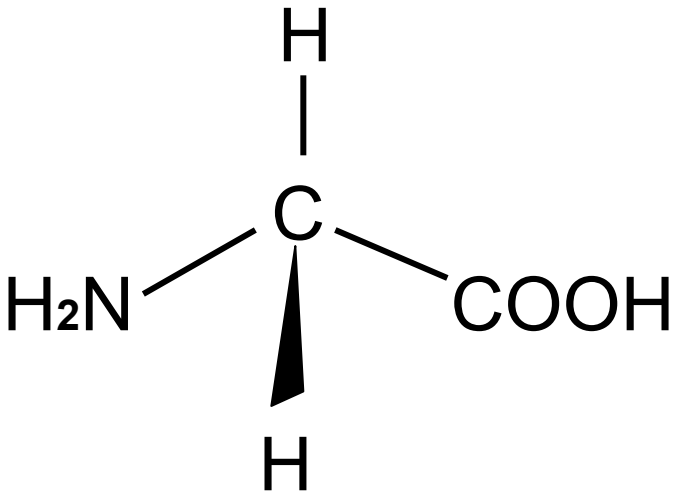
Much Ado About Structure*

- Structure ⟷ Function
- Structure ⟷ Mechanism
- Structure ⟷ Origins/Evolution
- Structure-based Drug Design
- Solving the Protein Folding Problem

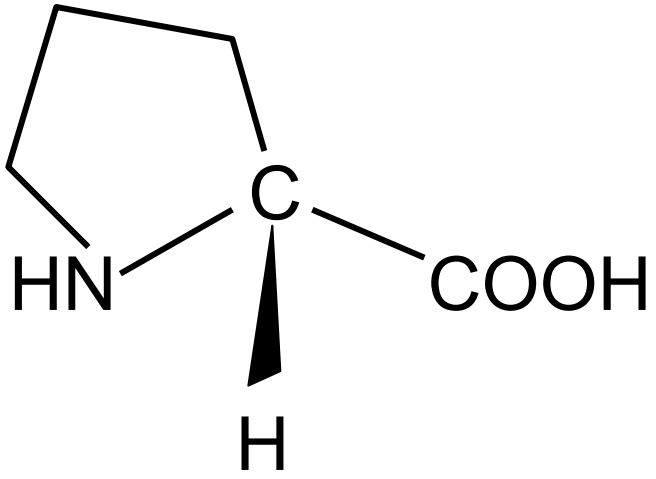
Amino Acids



Glycine and Proline*

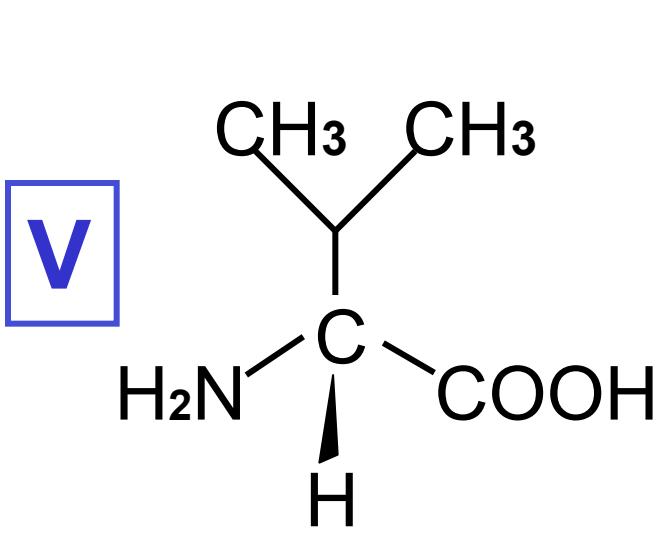


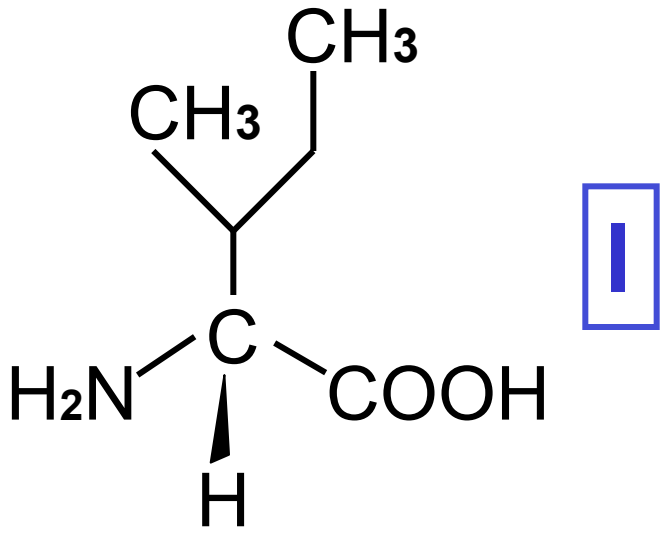
G

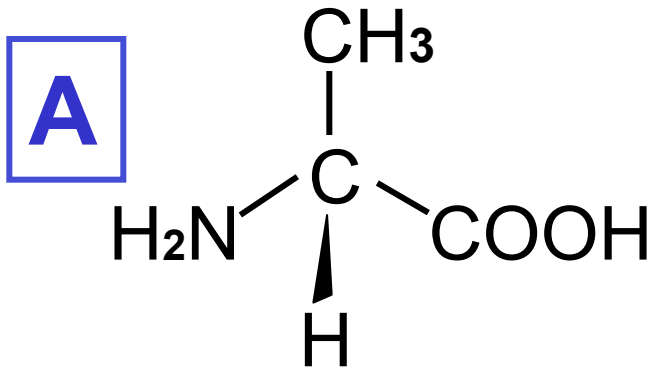


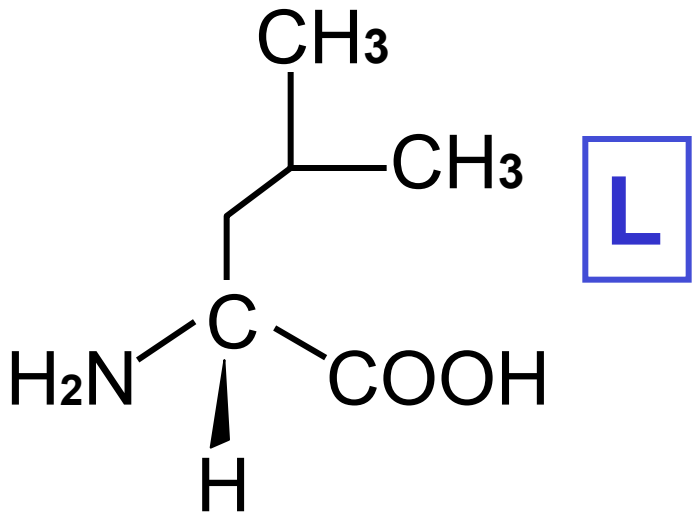
P

Aliphatic Amino Acids*

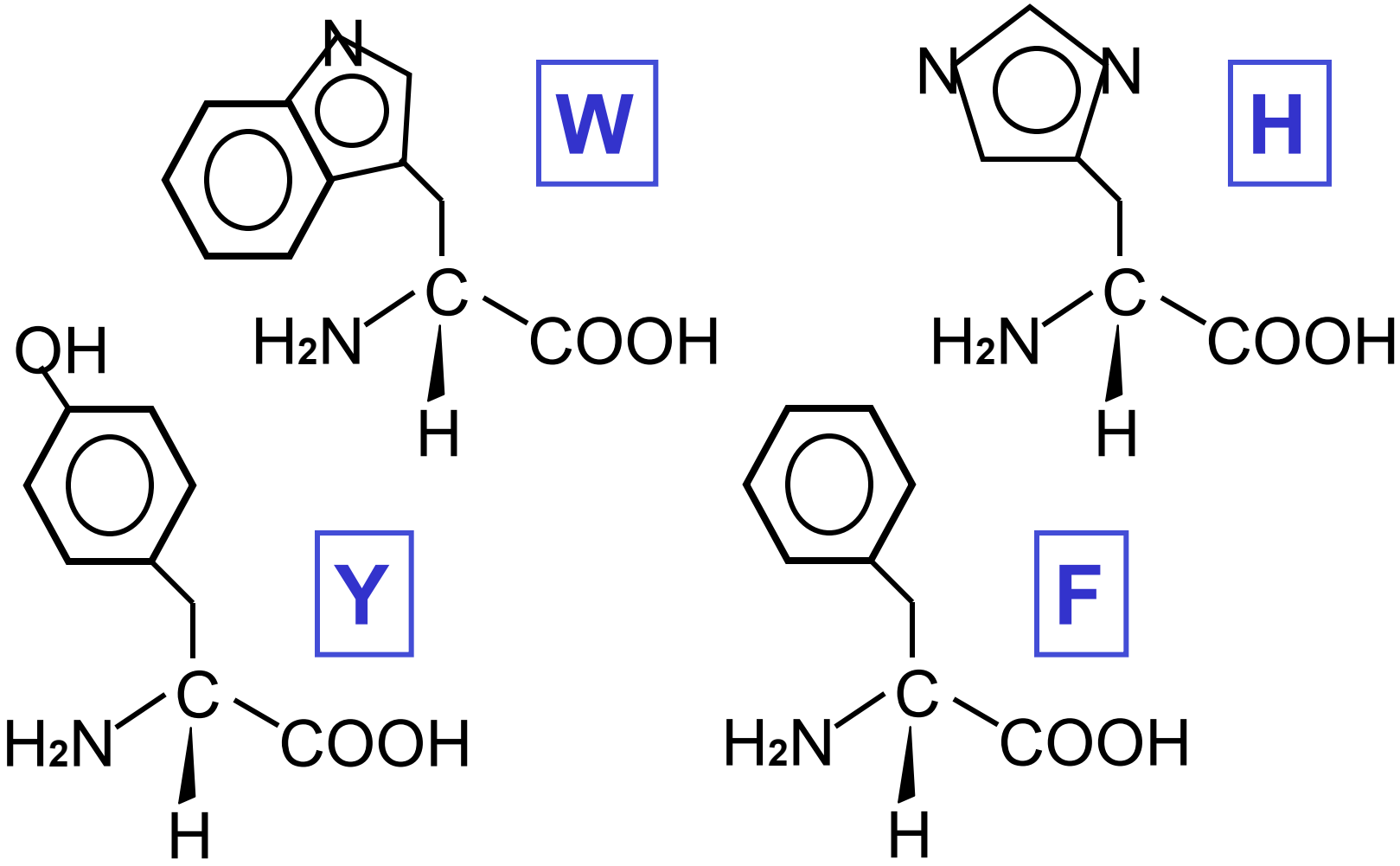




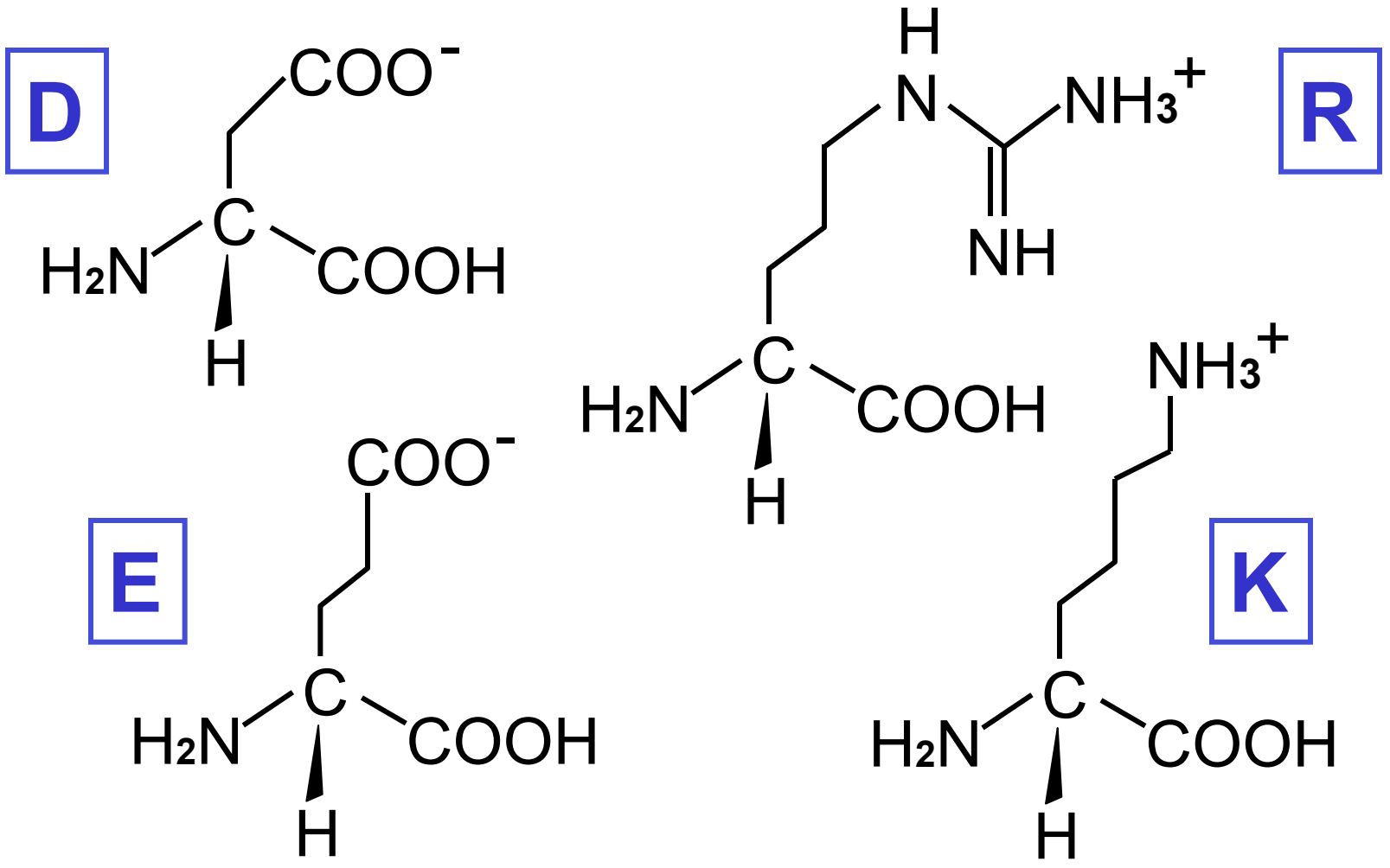




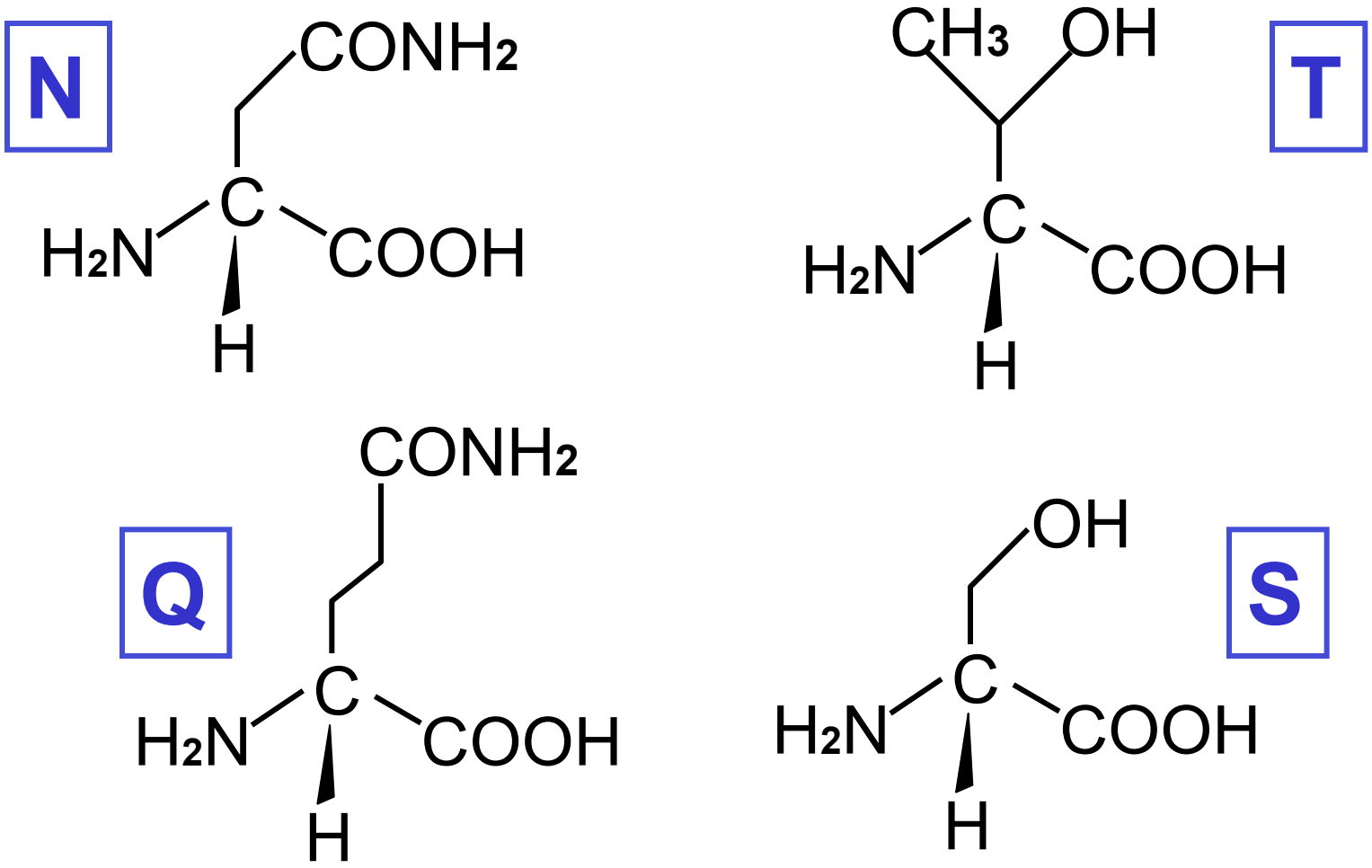
Aromatic Amino Acids*



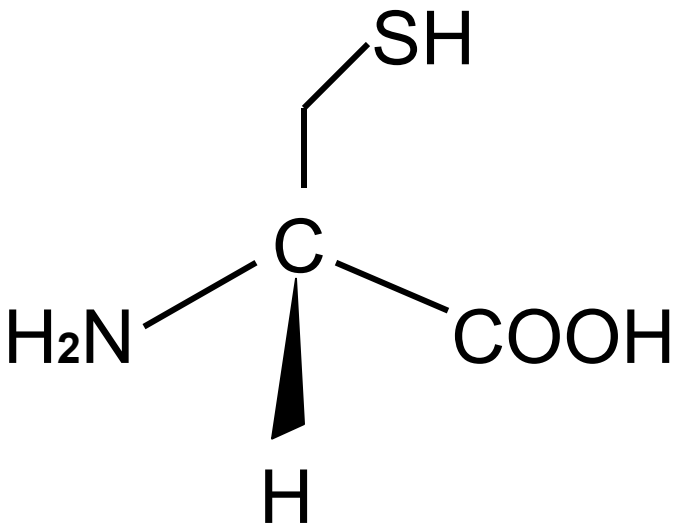
Charged Amino Acids*



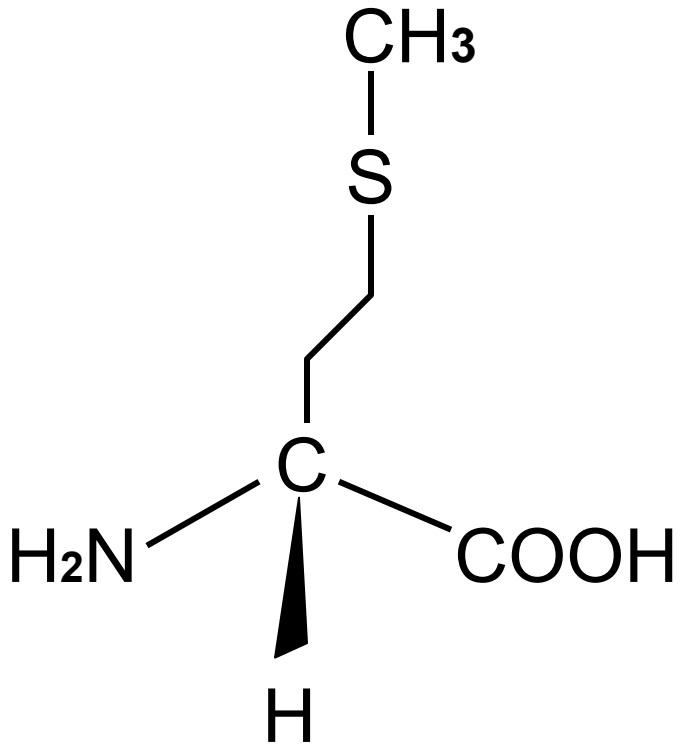
Polar Amino Acids*



Sulfo-Amino Acids*

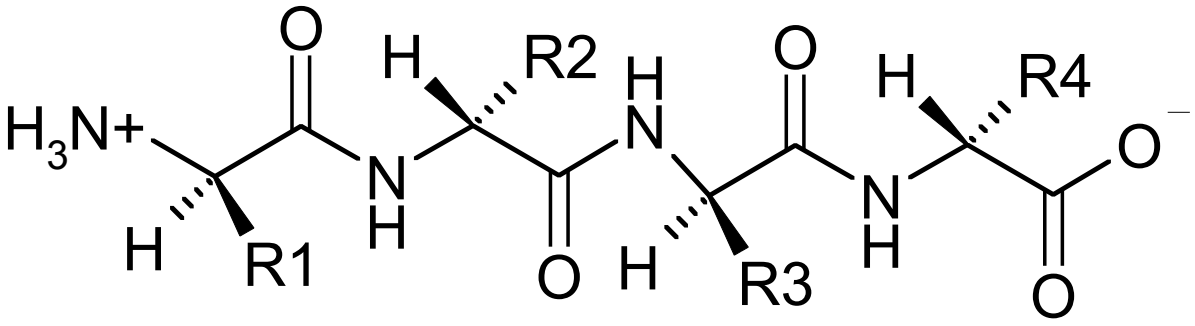


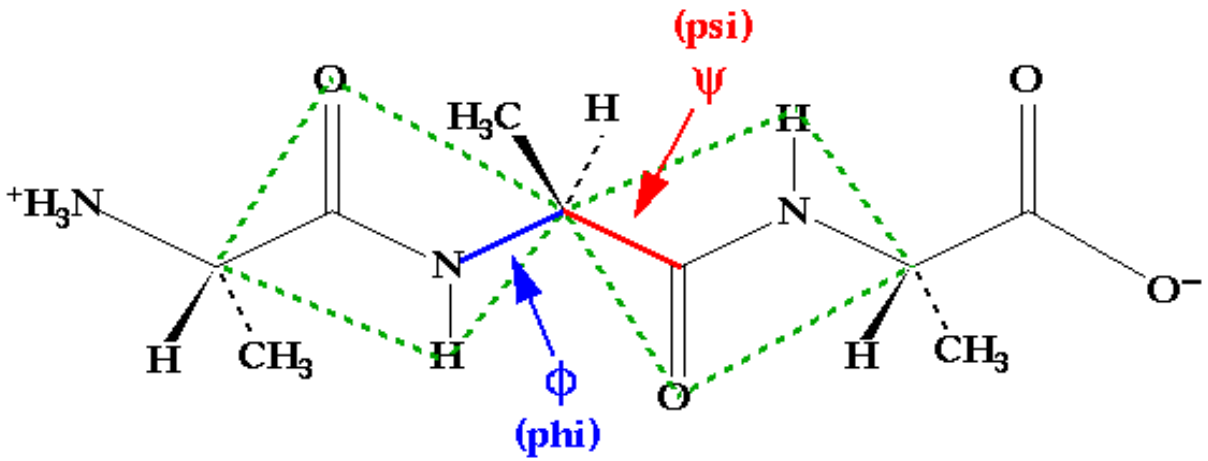
C



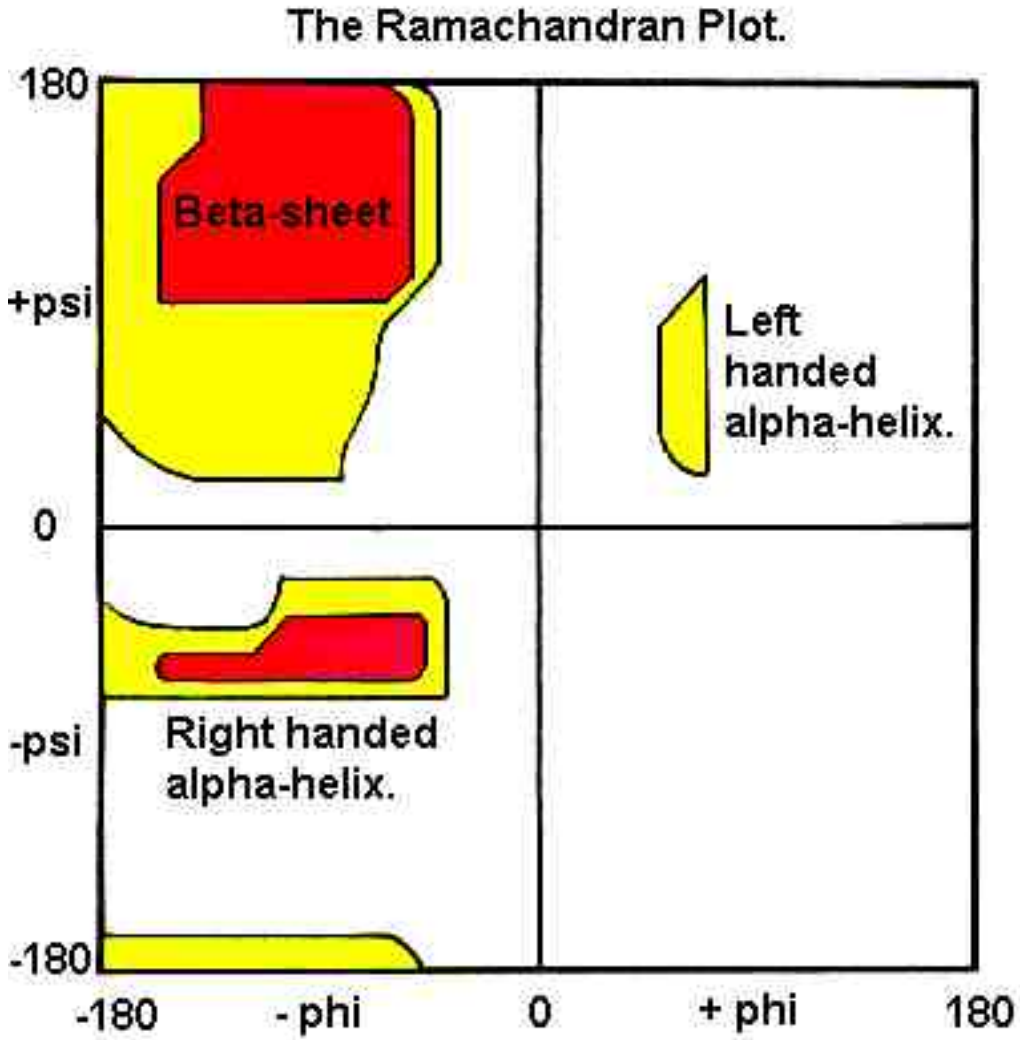
M

Polypeptides



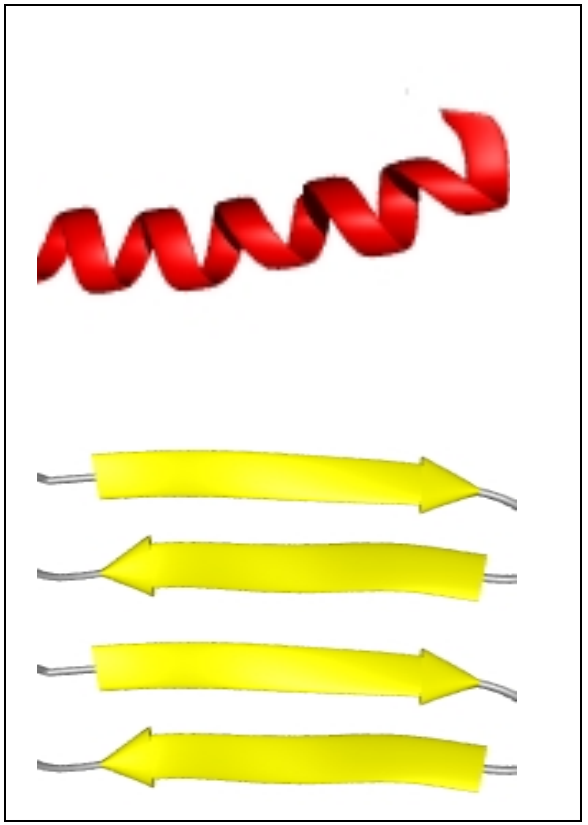


Ramachandran Plot*

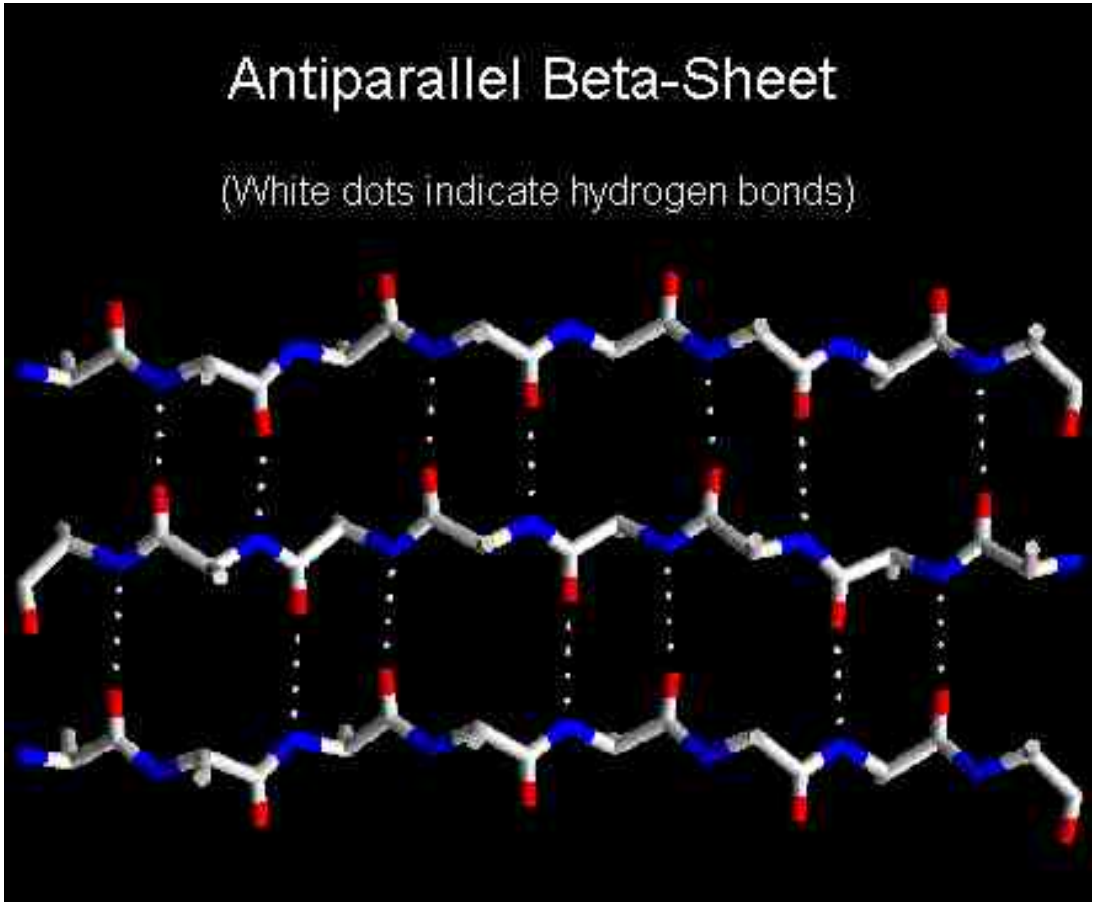


Secondary Structure*

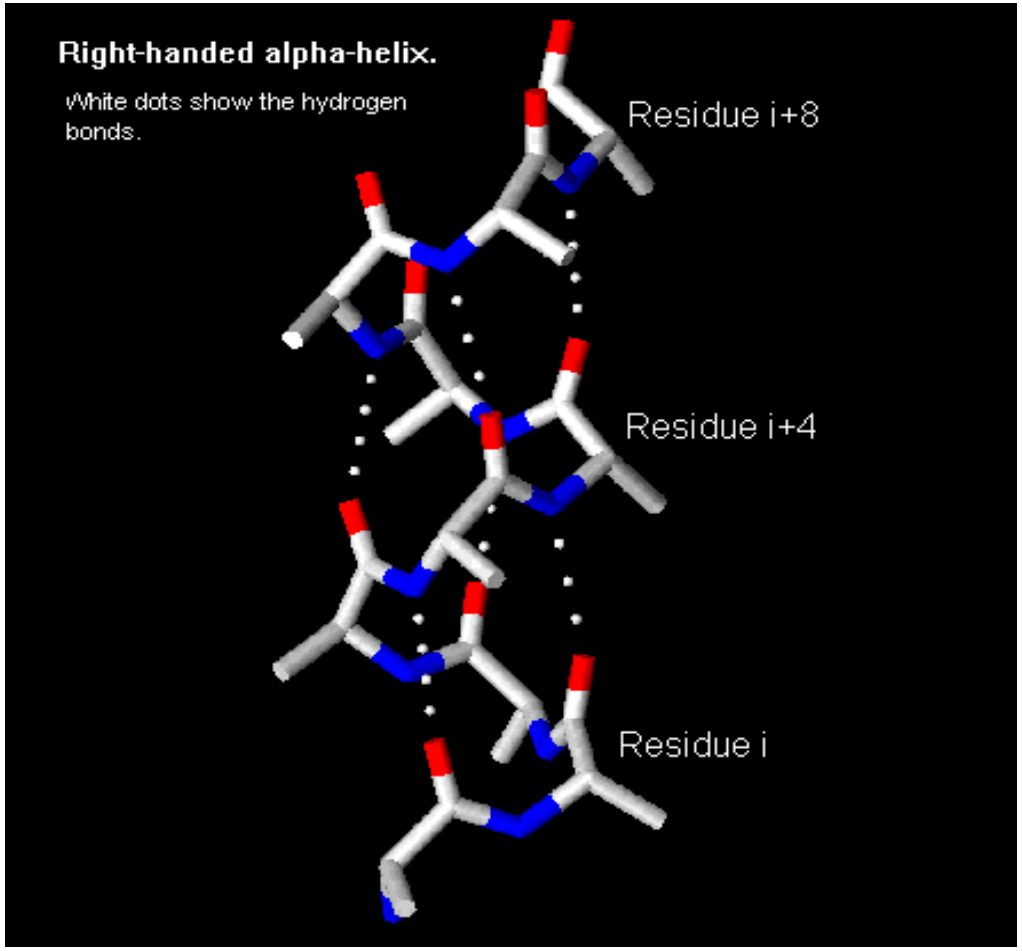
Phi & Psi angles for Regular Secondary Structure Conformations		
Structure	Phi (Φ)	Psi(Ψ)
Antiparallel β-sheet	-139	+135
Parallel β-Sheet	-119	+113
Right-handed α-helix	- 64	- 40
3_{10} helix	-49	-26
π helix	-57	-70
Polyproline I	-83	+158
Polyproline II	-78	+149
Polyglycine II	-80	+150



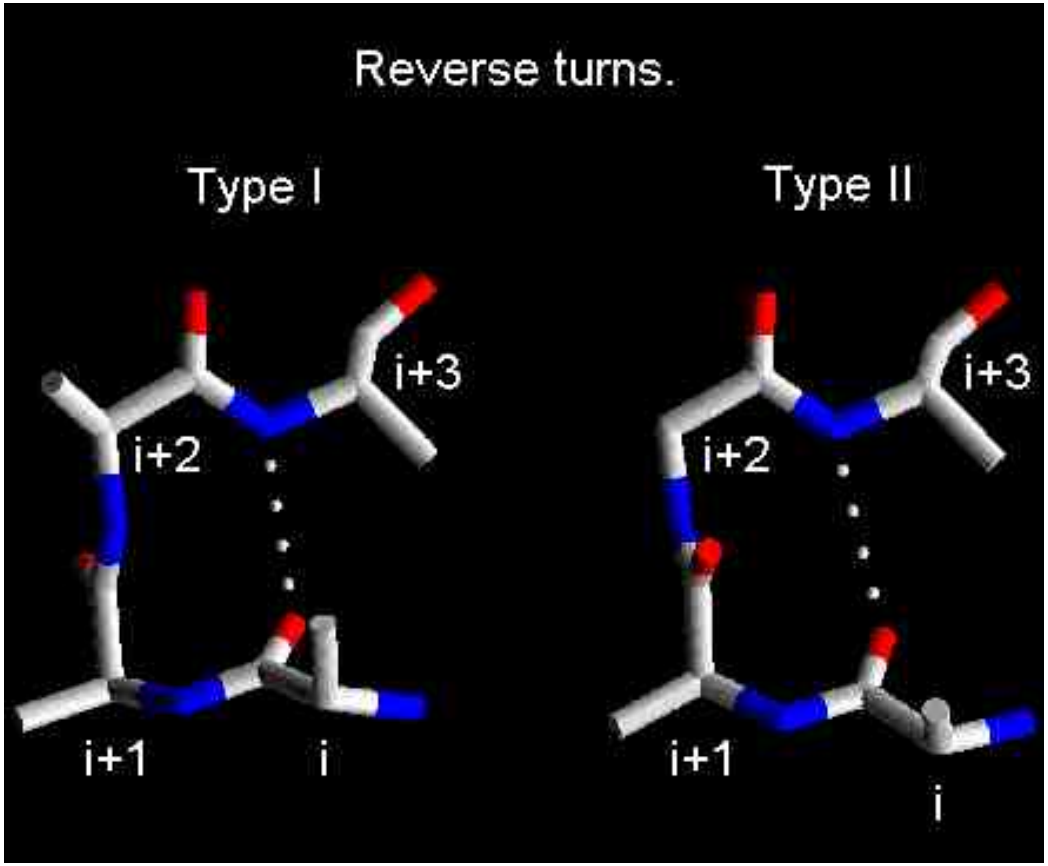
Beta Sheet


Antiparallel Beta-Sheet
(White dots indicate hydrogen bonds)

Alpha Helix

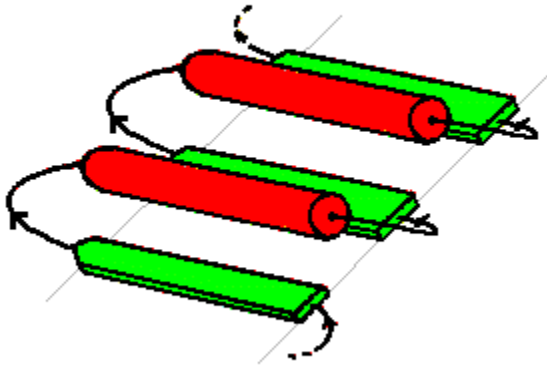

Right-handed alpha-helix.
White dots show the hydrogen bonds.
Residue i+8
Residue i+4
Residue i

Reverse Turn

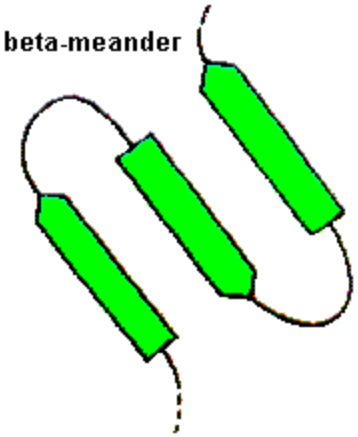


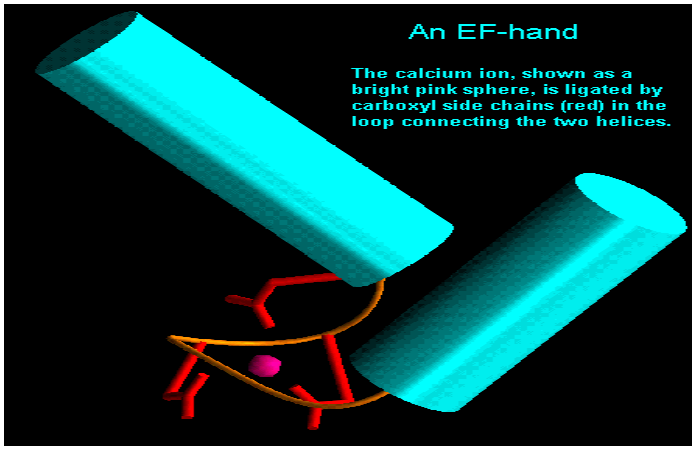
Supersecondary Structure*

he Rossman fold

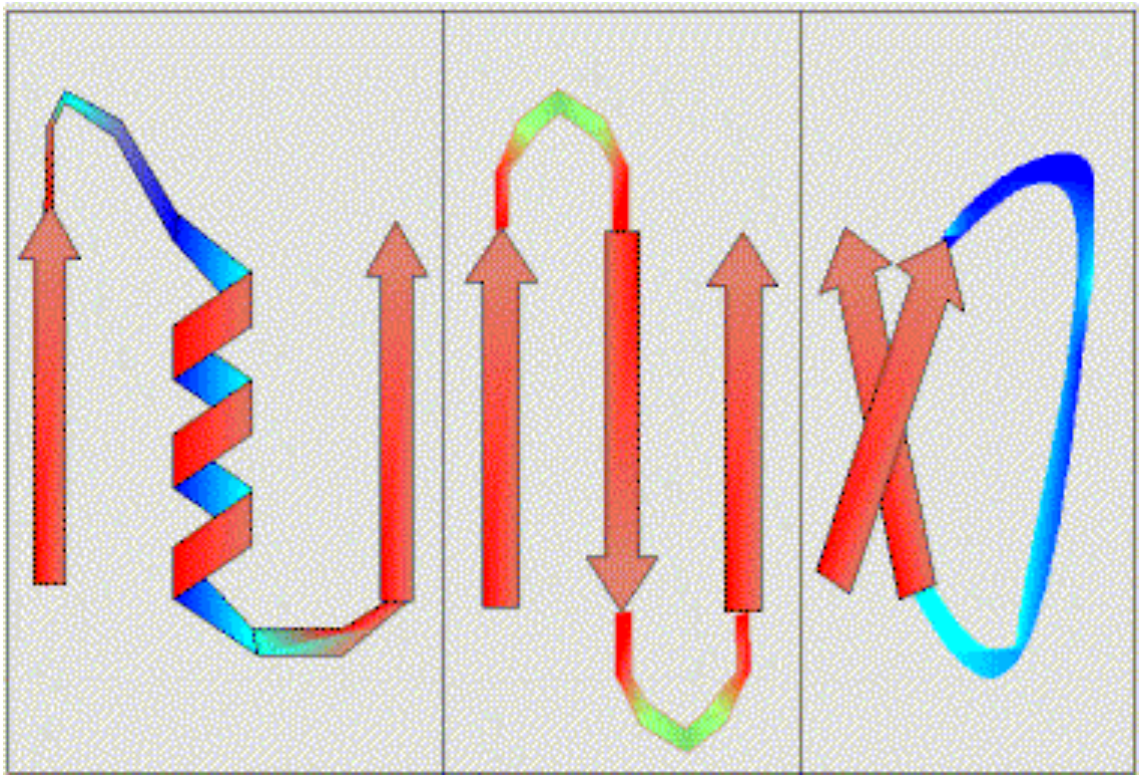


A beta-meander

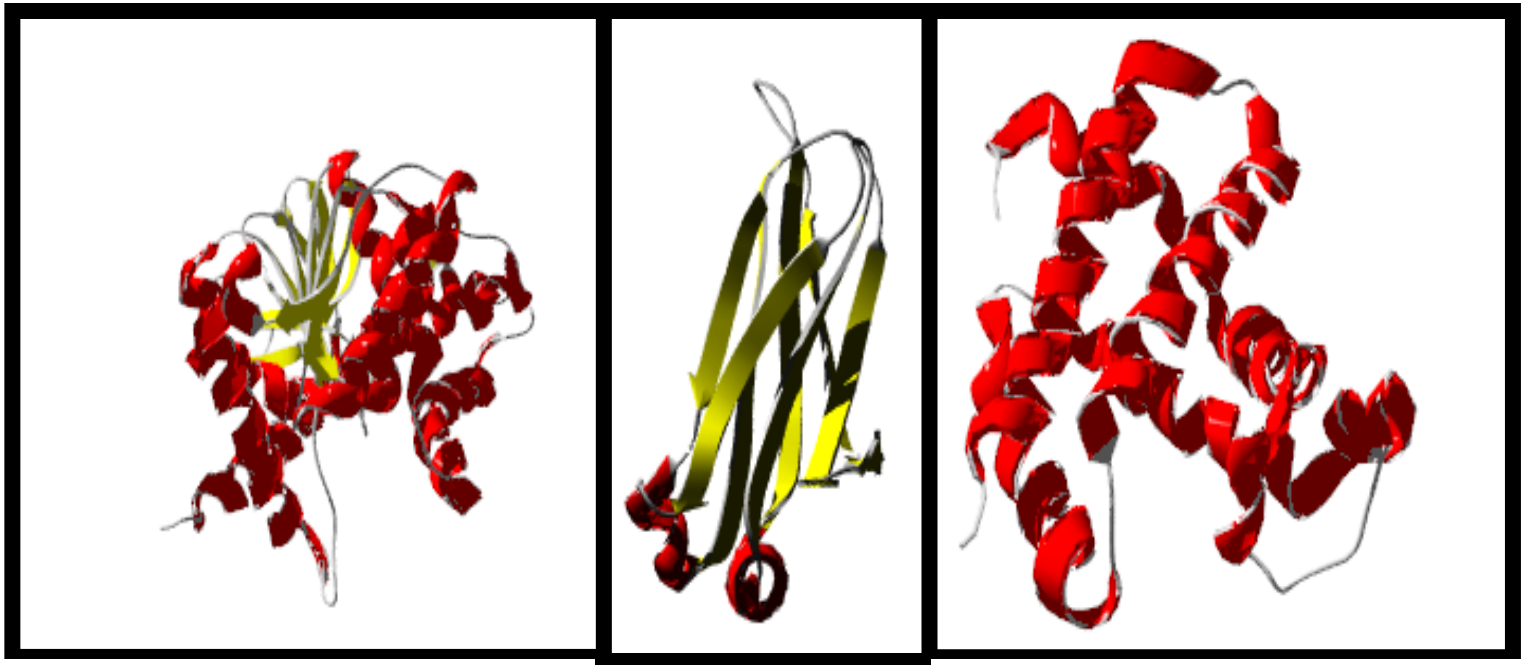




Supersecondary Structure



Tertiary Structure*



Lactate Dehydrogenase: Mixed α / β

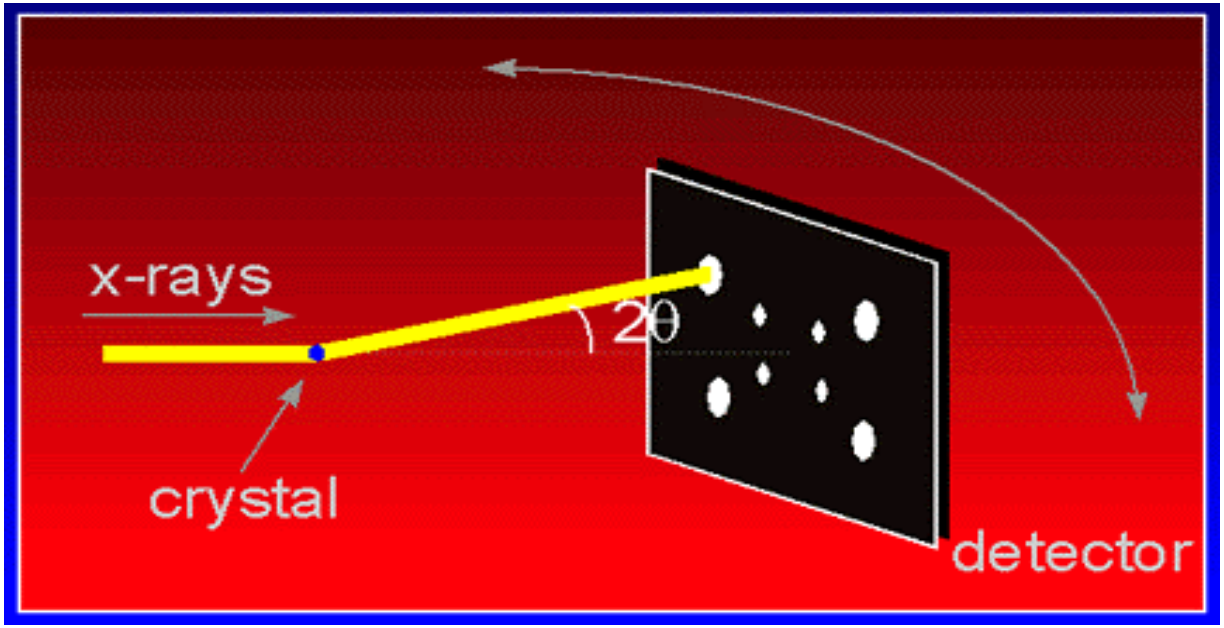
Immunoglobulin Fold: β

Hemoglobin B Chain: α

Solving Protein Structures*

- Only 2 kinds of techniques allow one to get atomic resolution pictures of macromolecules
- X-ray Crystallography (first applied in 1961 - Kendrew & Perutz)
- NMR Spectroscopy (first applied in 1983 - Ernst & Wuthrich)

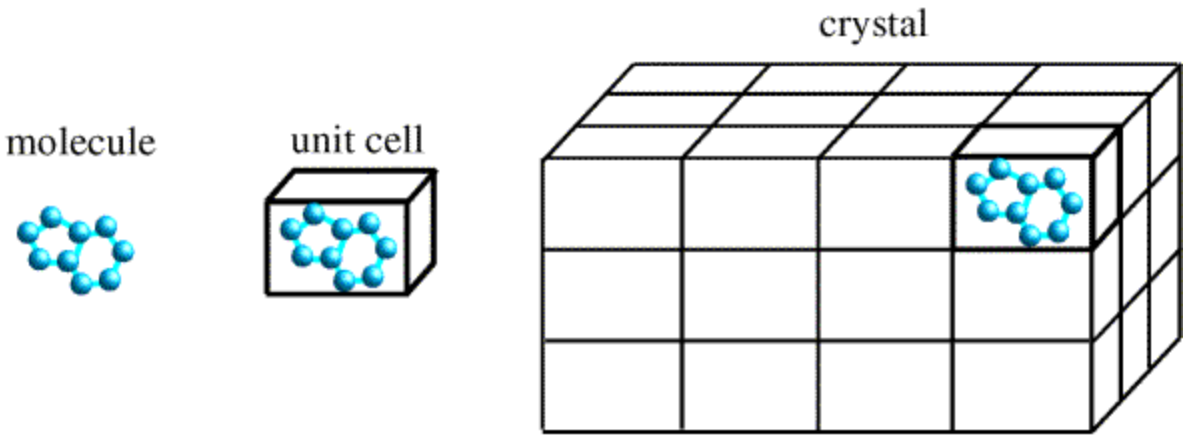
X-ray Crystallography



X-ray Crystallography*

- Crystallization
- Diffraction Apparatus
- Diffraction Principles
- Conversion of Diffraction Data to Electron Density
- Resolution
- Chain Tracing

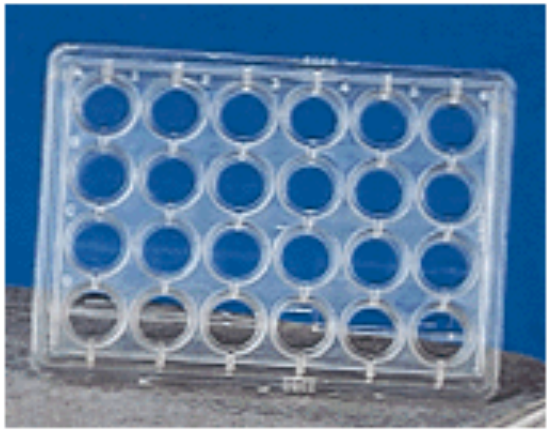
Crystallization

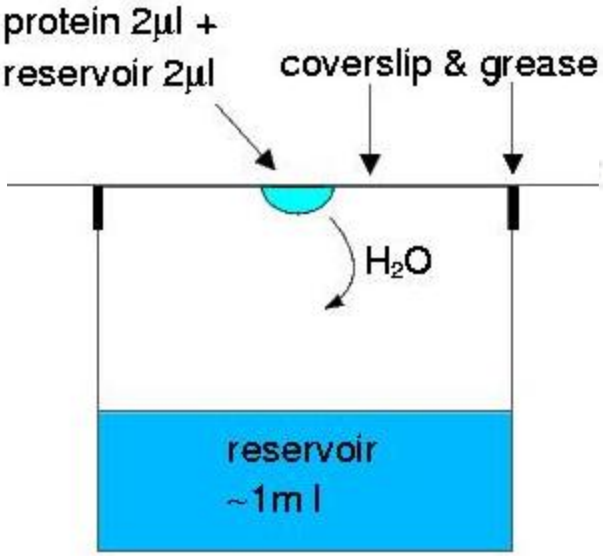


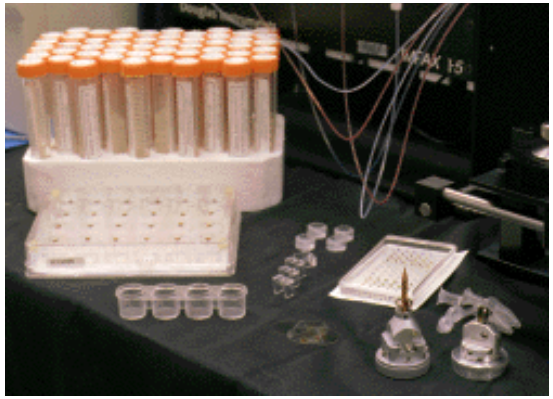
Protein Crystal

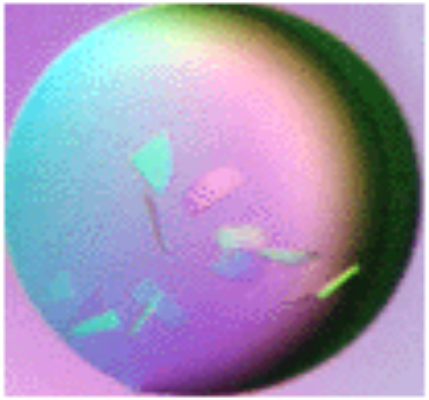
Protein crystallization proceeds in two steps. The first step screens precipitants and conditions that lead to protein precipitation. The second step tries to refine the precipitation steps so that the protein crystallizes instead of precipitating

Crystallization*



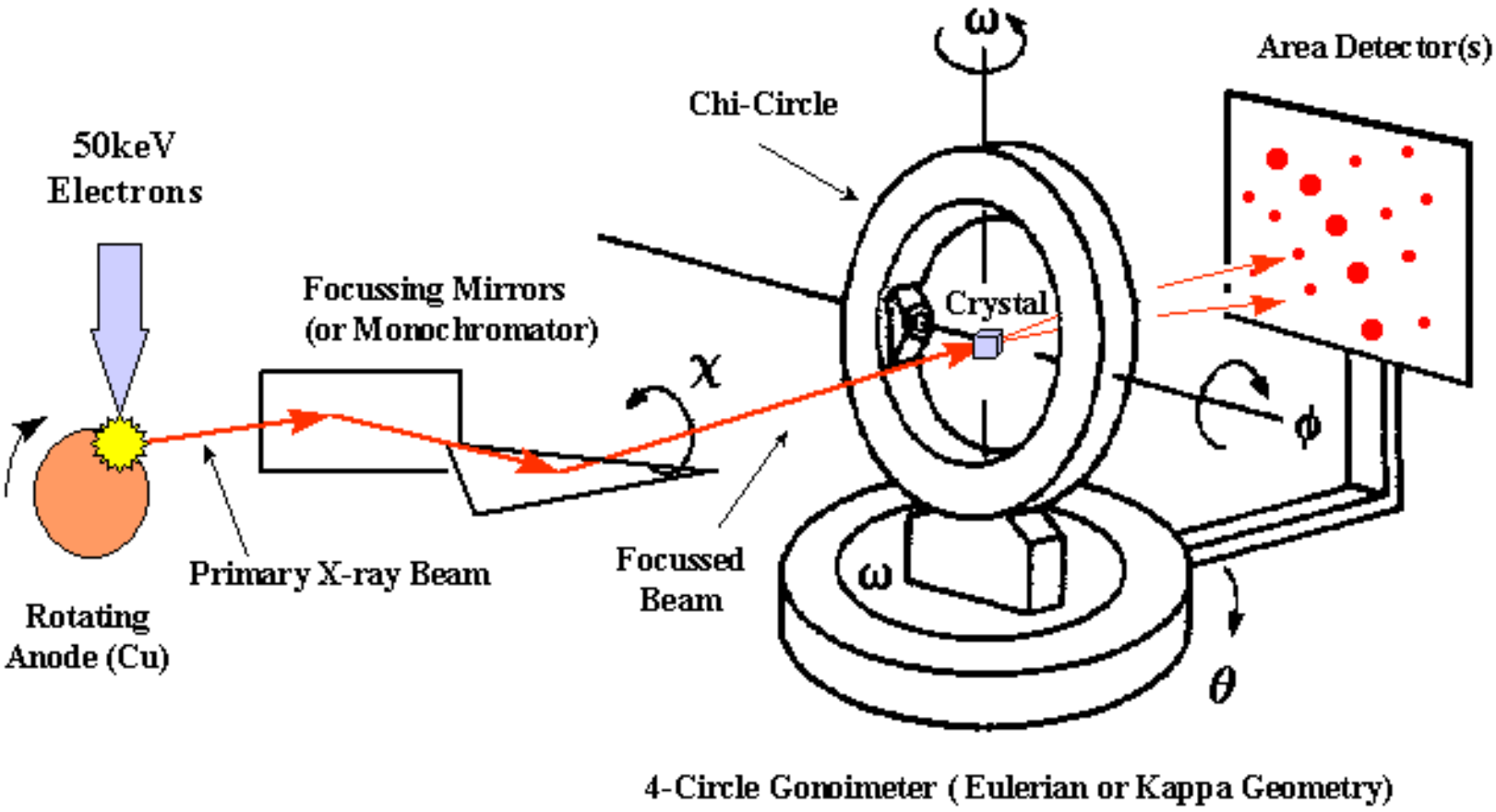




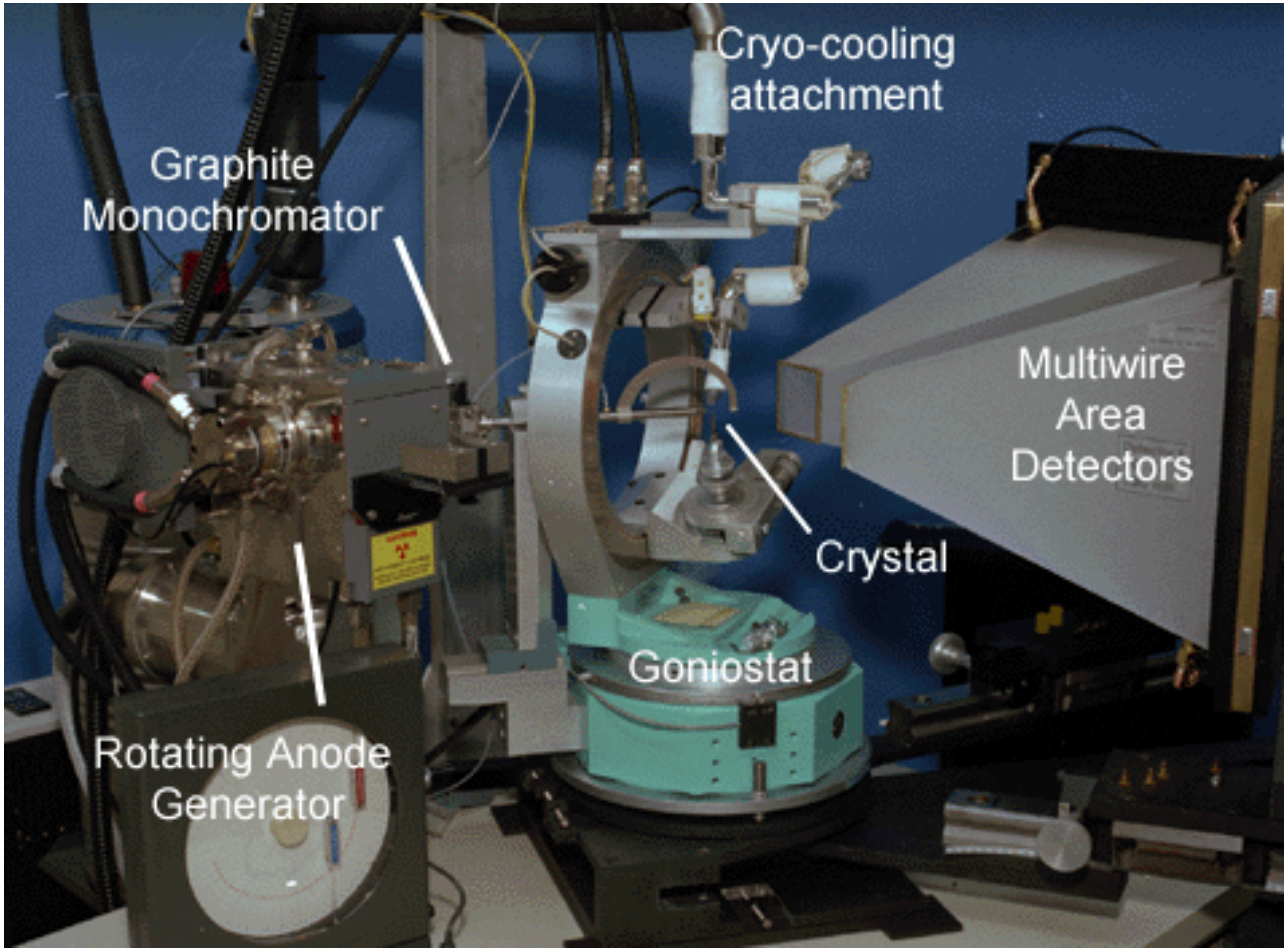


Hanging Drop Experiment for Cyrstallization

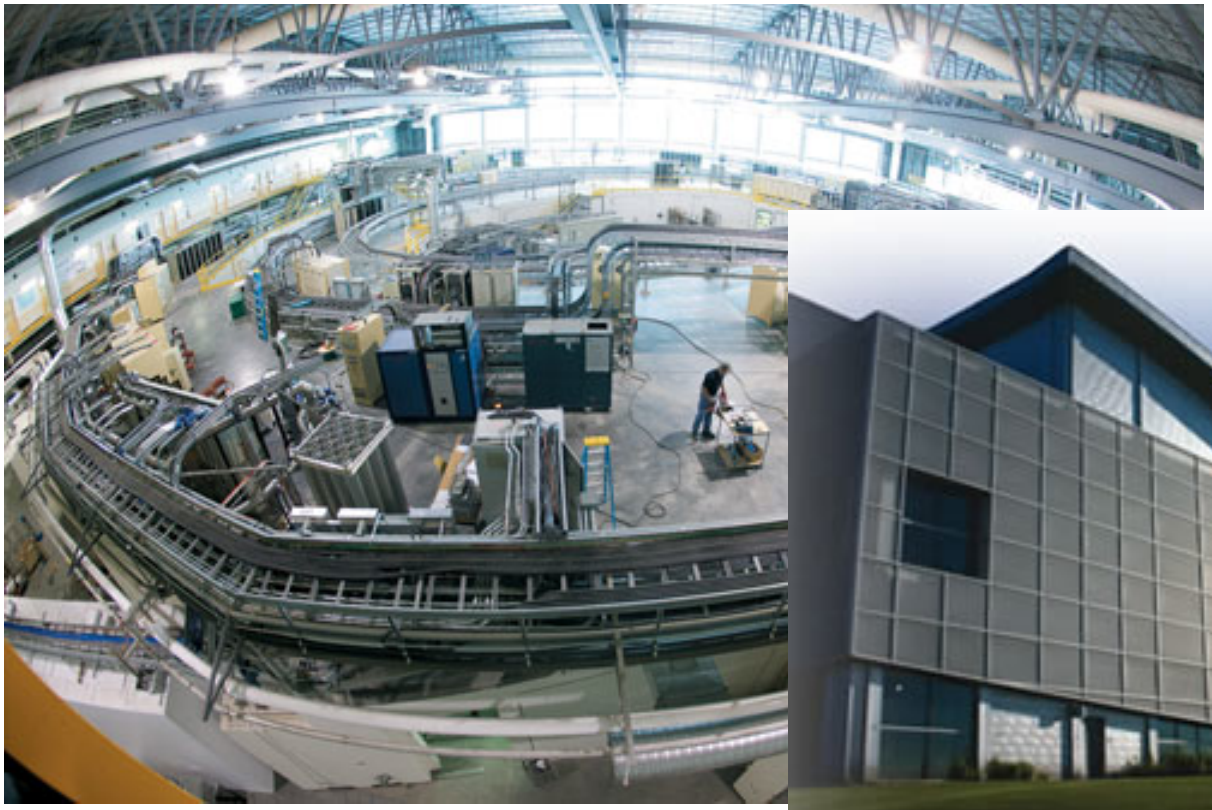
Diffraction Apparatus*



Diffraction Apparatus

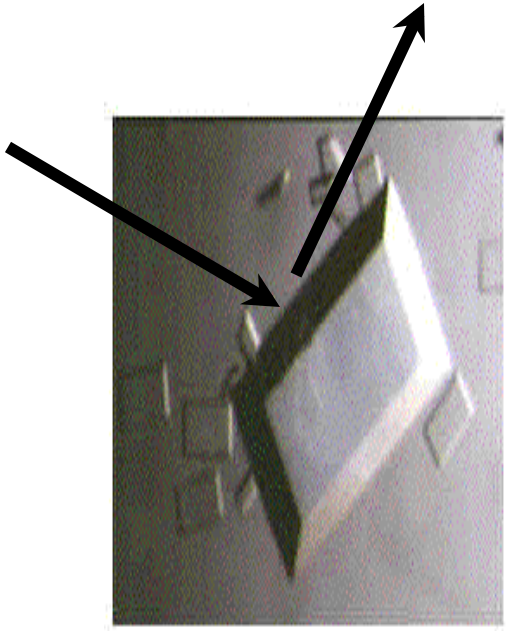


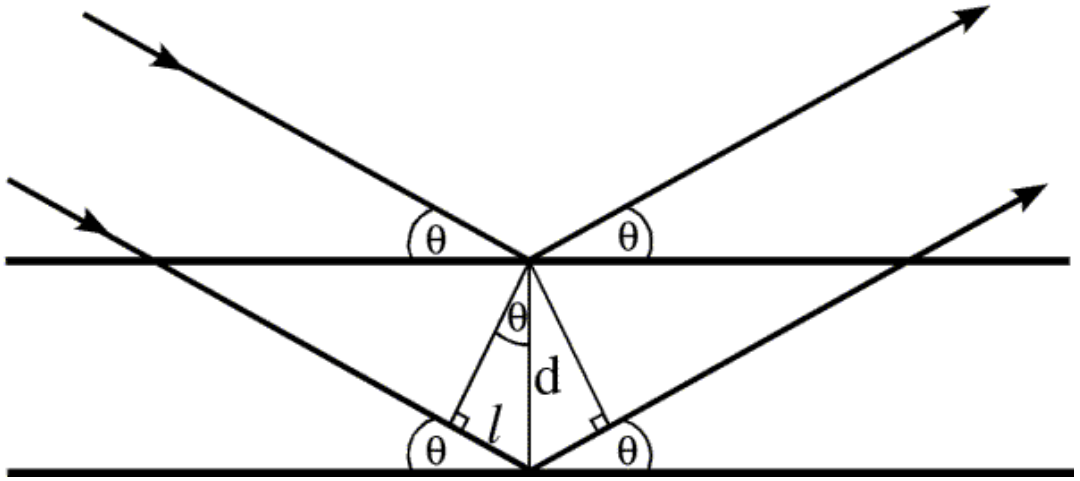
A Bigger Diffraction Apparatus



Synchrotron Light Source

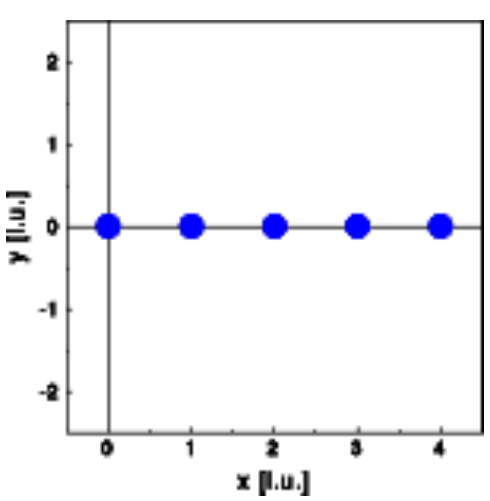
Diffraction Principles*



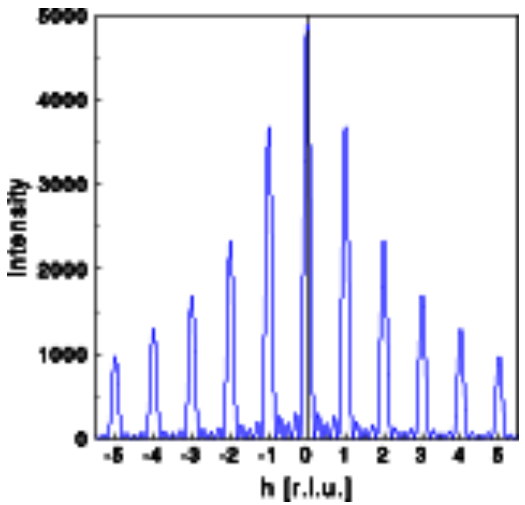


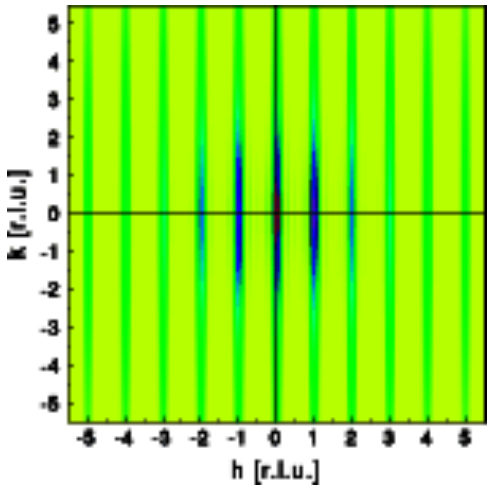
$$n\lambda = 2d\sin\theta$$

Diffraction Principles

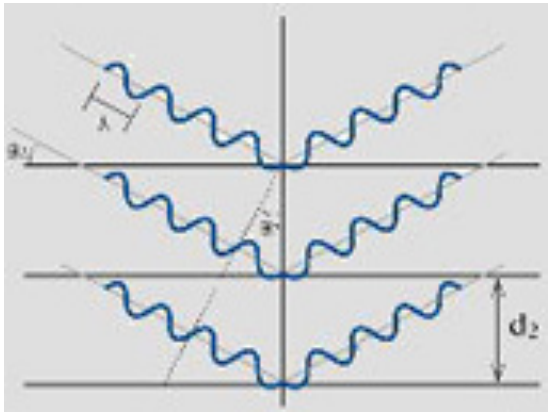


A string of atoms

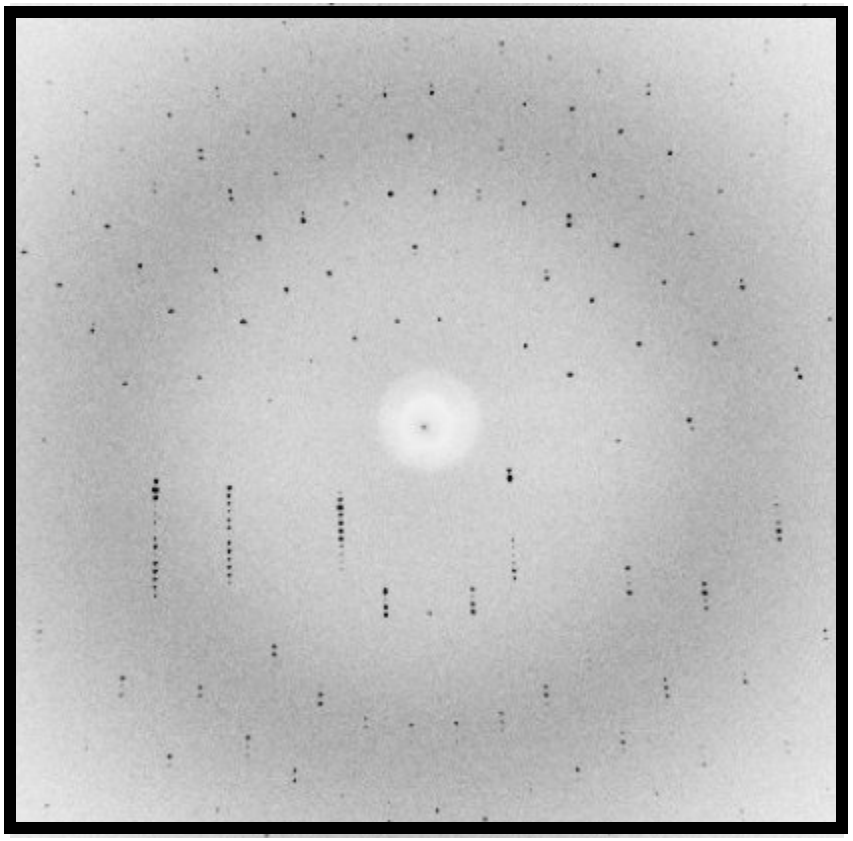




Corresponding Diffraction Pattern

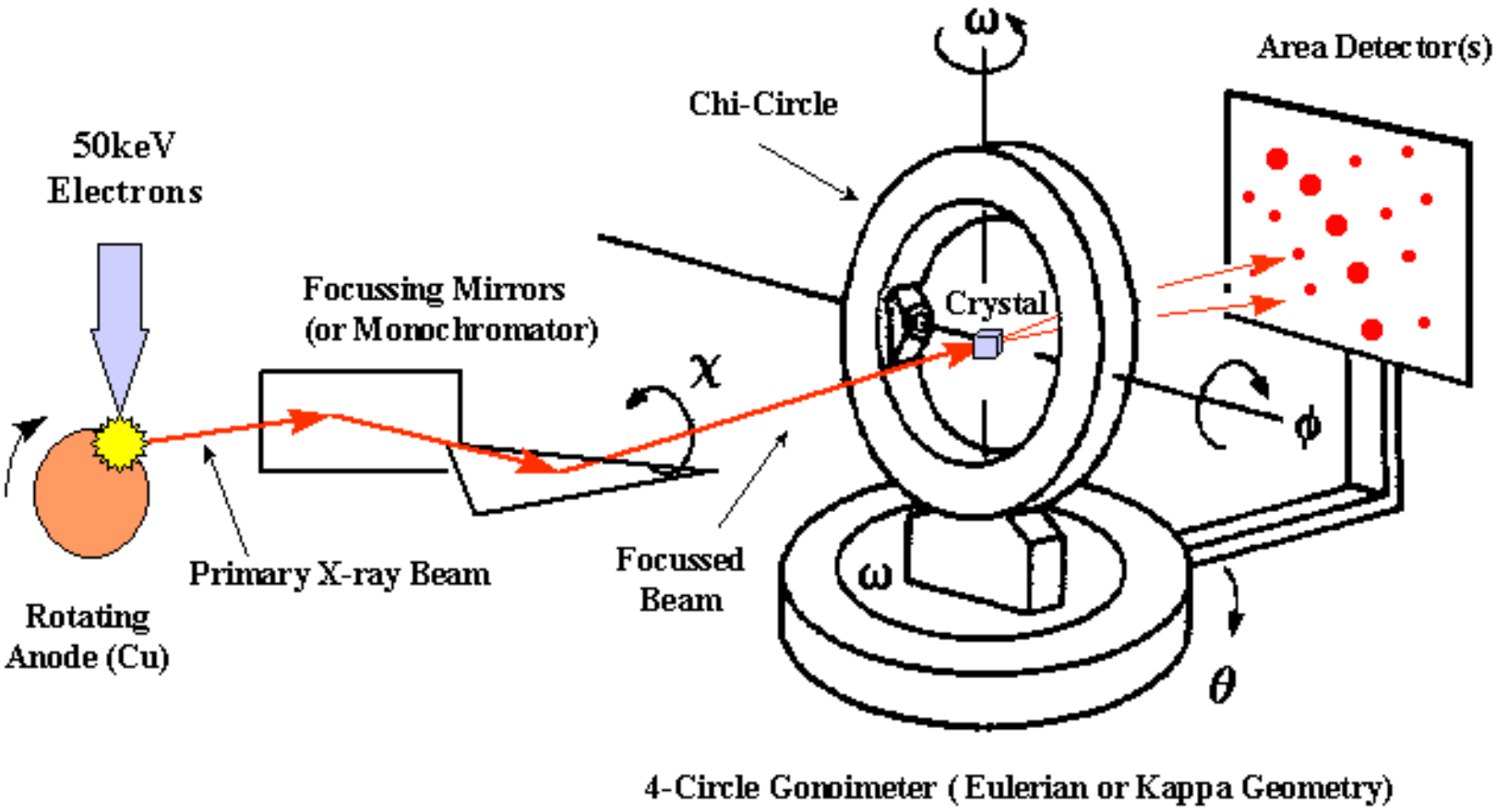


Protein Crystal Diffraction

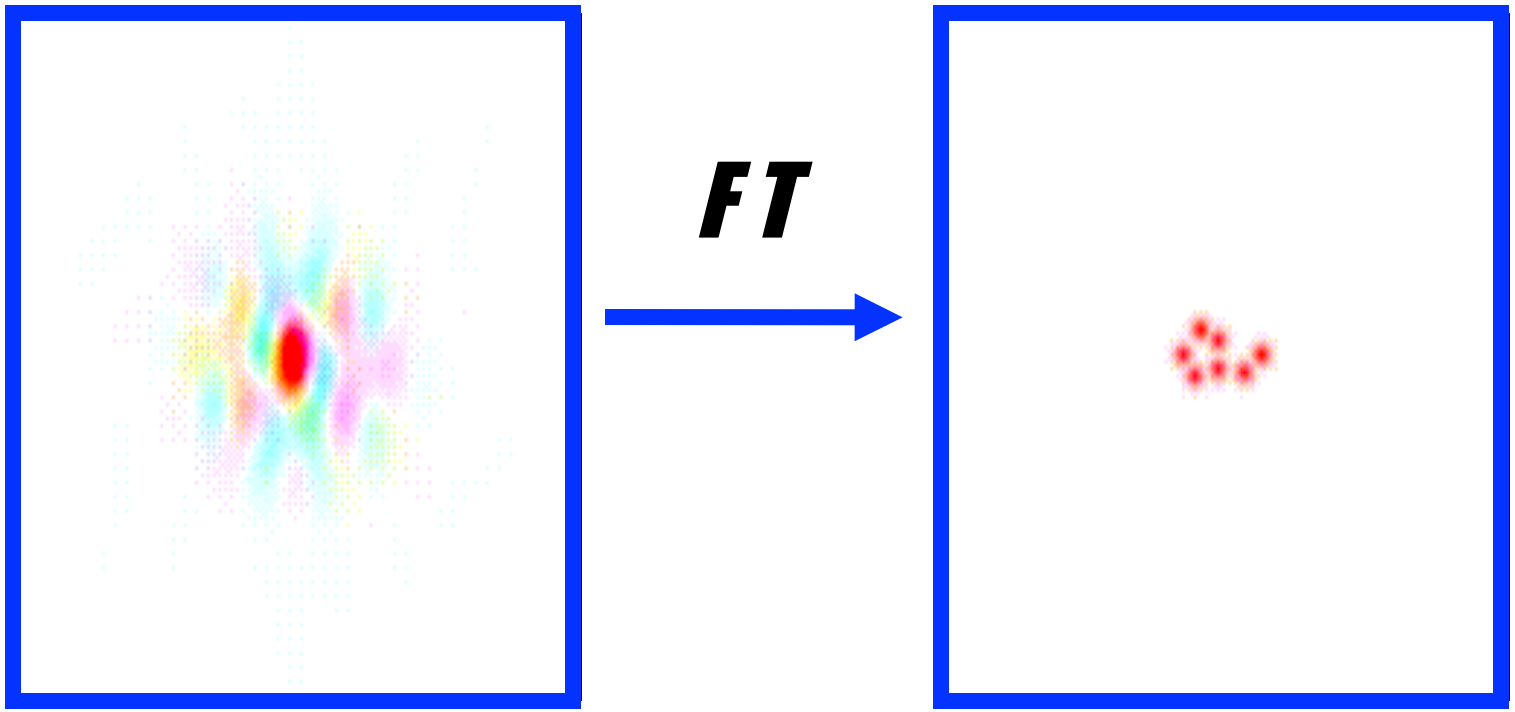


Diffraction Pattern

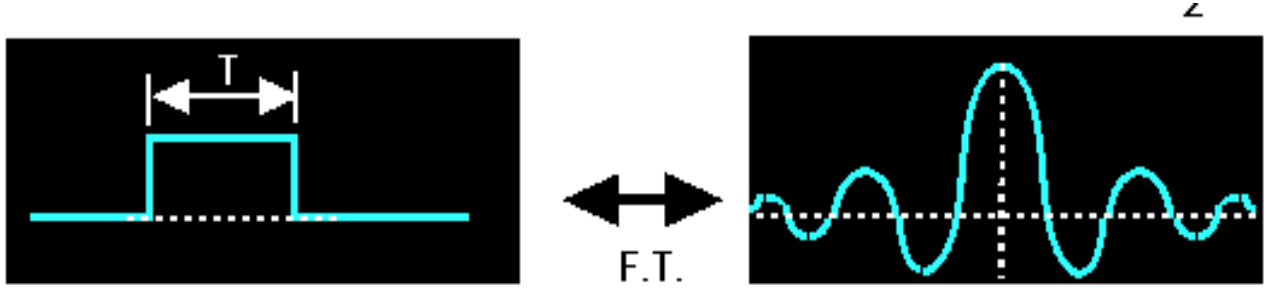
Diffraction Apparatus

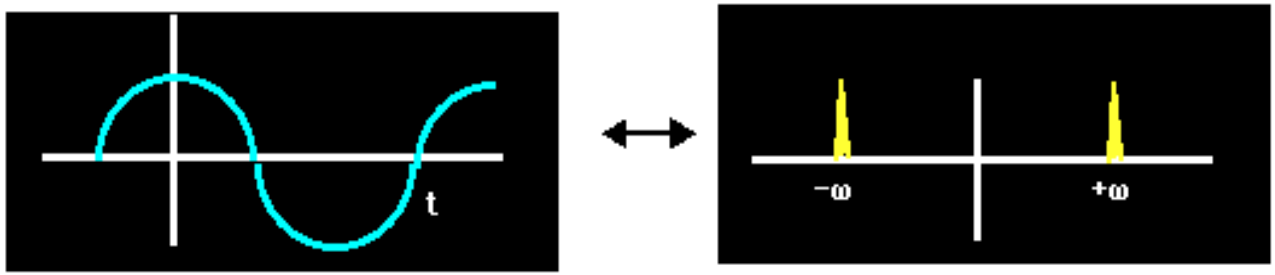


Converting Diffraction Data to Electron Density



Fourier Transformation*

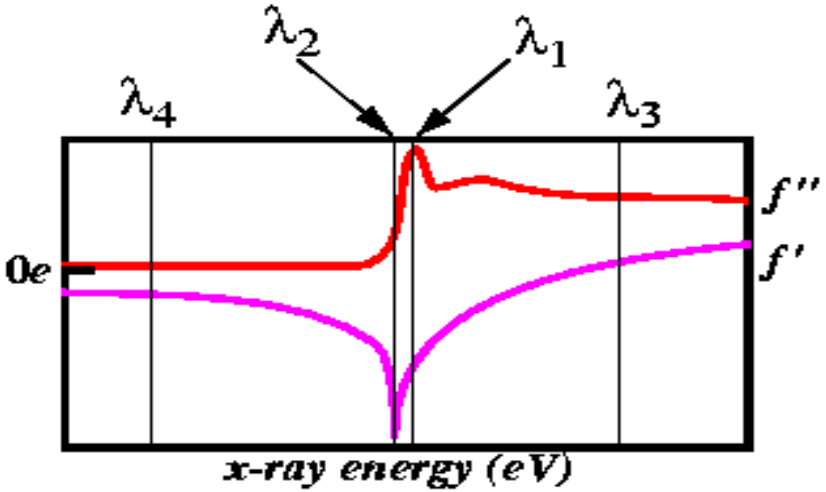


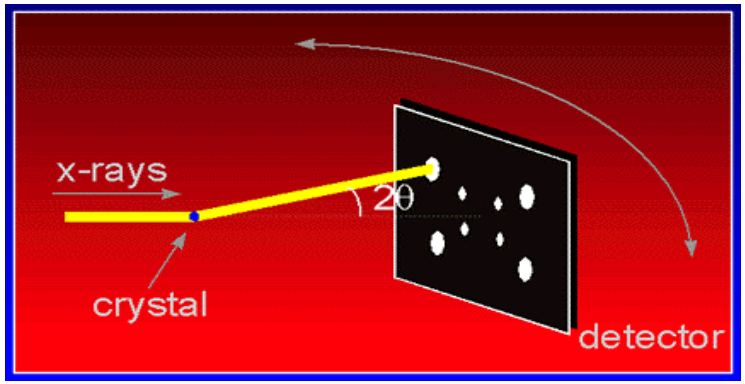


$$F(x,y,z) = \int f(hkl) e^{\iota(xyz)(hkl)} d(hkl)$$

Converts from units of inverse space to cartesian coordinates

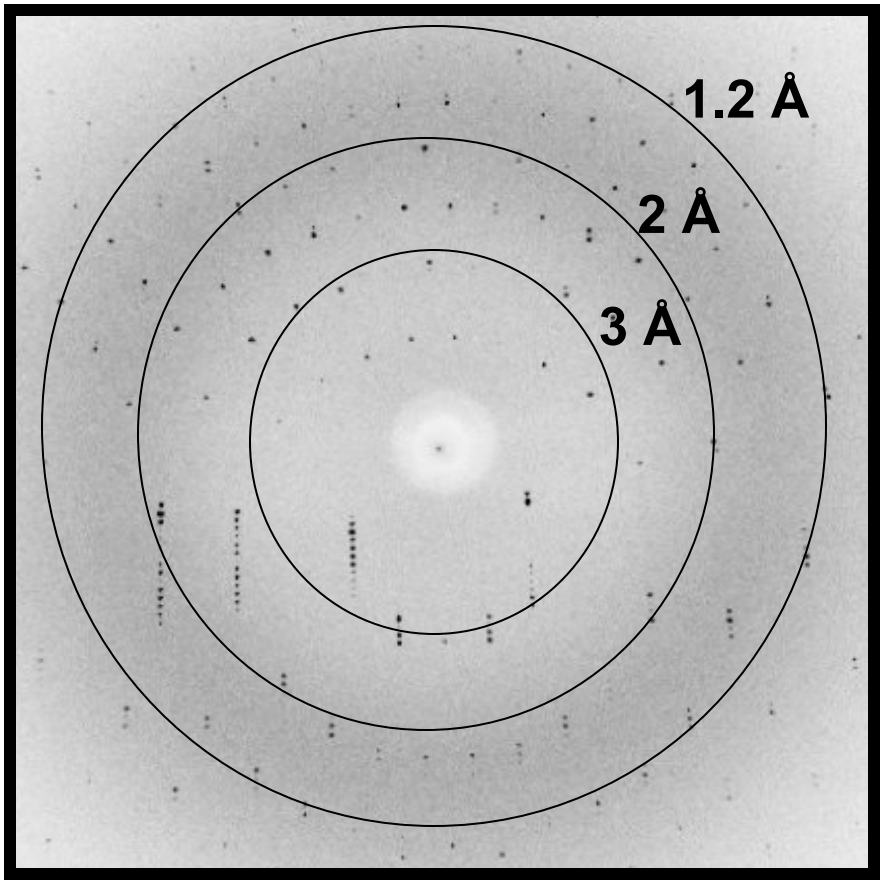
MAD & X-ray Crystallography*

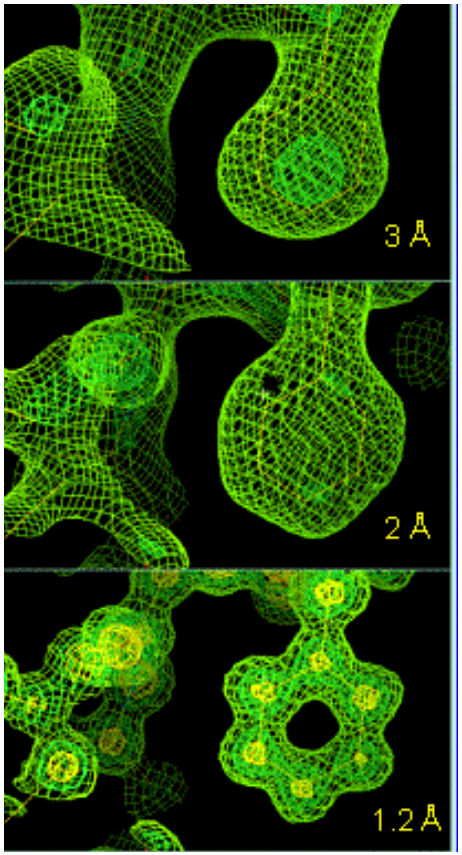




- MAD (Multiwavelength Anomalous Dispersion
- Requires synchrotron beam lines
- Requires protein with multiple scattering centres (selenomethionine labeled)
- Allows rapid phasing
- Proteins can now be “solved” in just 1-2 days

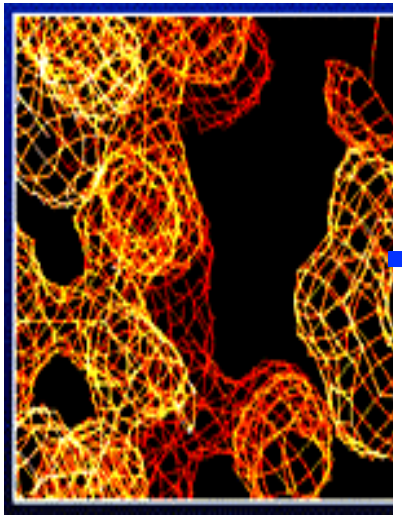
Resolution*



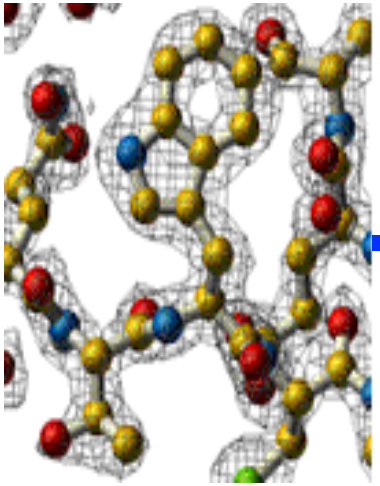


Resolution describes the ability of an imaging system to resolve detail in the object that is being imaged.

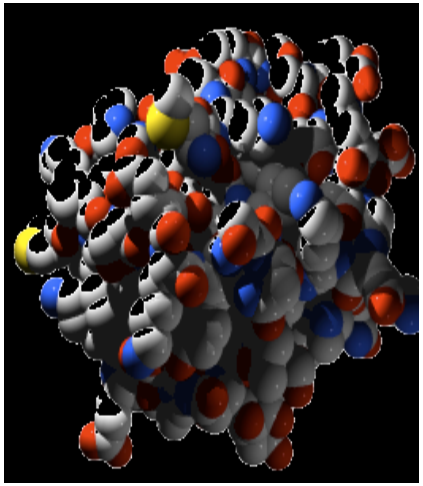
Chain Tracing



Electron Density



Chain Trace



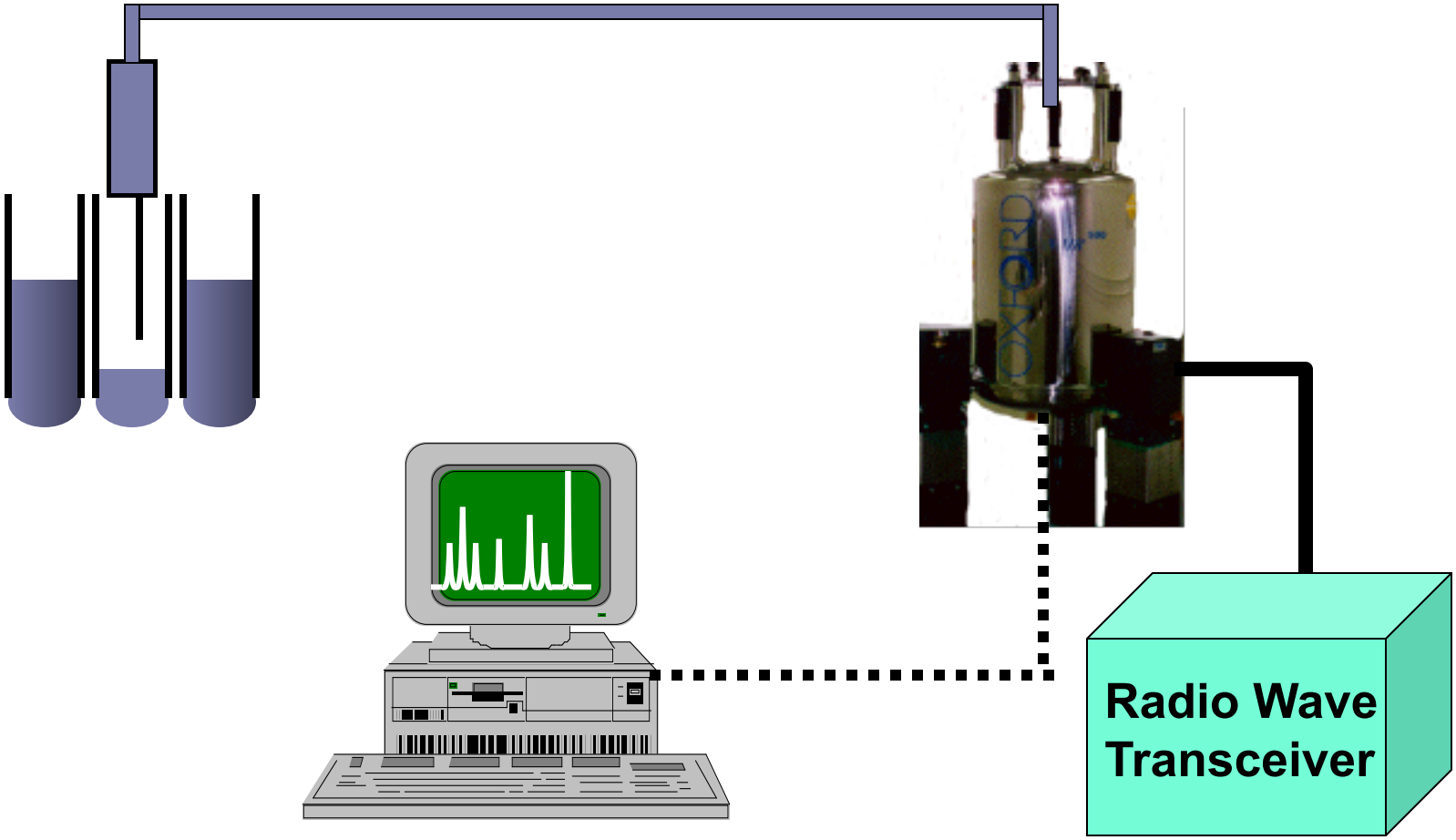
Final Model

The Final Result

```
ORIGX2      0.000000  1.000000  0.000000        0.00000                      2TRX 147
ORIGX3      0.000000  0.000000  1.000000        0.00000                      2TRX 148
SCALE1      0.011173  0.000000  0.004858        0.00000                      2TRX 149
SCALE2      0.000000  0.019585  0.000000        0.00000                      2TRX 150
SCALE3      0.000000  0.000000  0.018039        0.00000                      2TRX 151
ATOM      1  N   SER A   1      21.389  25.406  -4.628  1.00 23.22           2TRX 152
ATOM      2  CA  SER A   1      21.628  26.691  -3.983  1.00 24.42           2TRX 153
ATOM      3  C   SER A   1      20.937  26.944  -2.679  1.00 24.21           2TRX 154
ATOM      4  O   SER A   1      21.072  28.079  -2.093  1.00 24.97           2TRX 155
ATOM      5  CB  SER A   1      21.117  27.770  -5.002  1.00 28.27           2TRX 156
ATOM      6  OG  SER A   1      22.276  27.925  -5.861  1.00 32.61           2TRX 157
ATOM      7  N   ASP A   2      20.173  26.028  -2.163  1.00 21.39           2TRX 158
ATOM      8  CA  ASP A   2      19.395  26.125  -0.949  1.00 21.57           2TRX 159
ATOM      9  C   ASP A   2      20.264  26.214   0.297  1.00 20.89           2TRX 160
ATOM     10  O   ASP A   2      19.760  26.575   1.371  1.00 21.49           2TRX 161
ATOM     11  CB  ASP A   2      18.439  24.914  -0.856  1.00 22.14           2TRX 162
```

http://www.ruppweb.org/Xray/101index.html

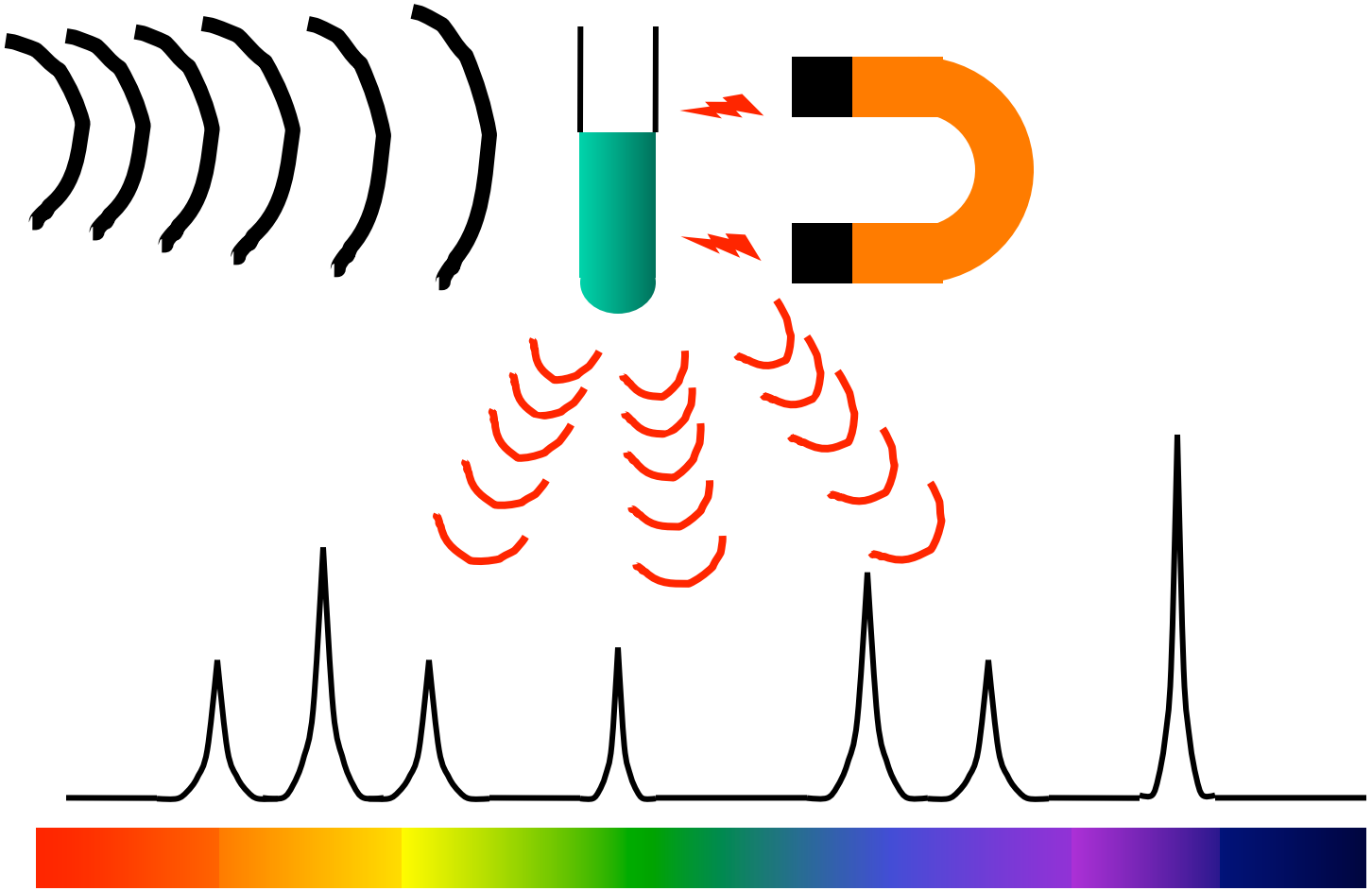
NMR Spectroscopy



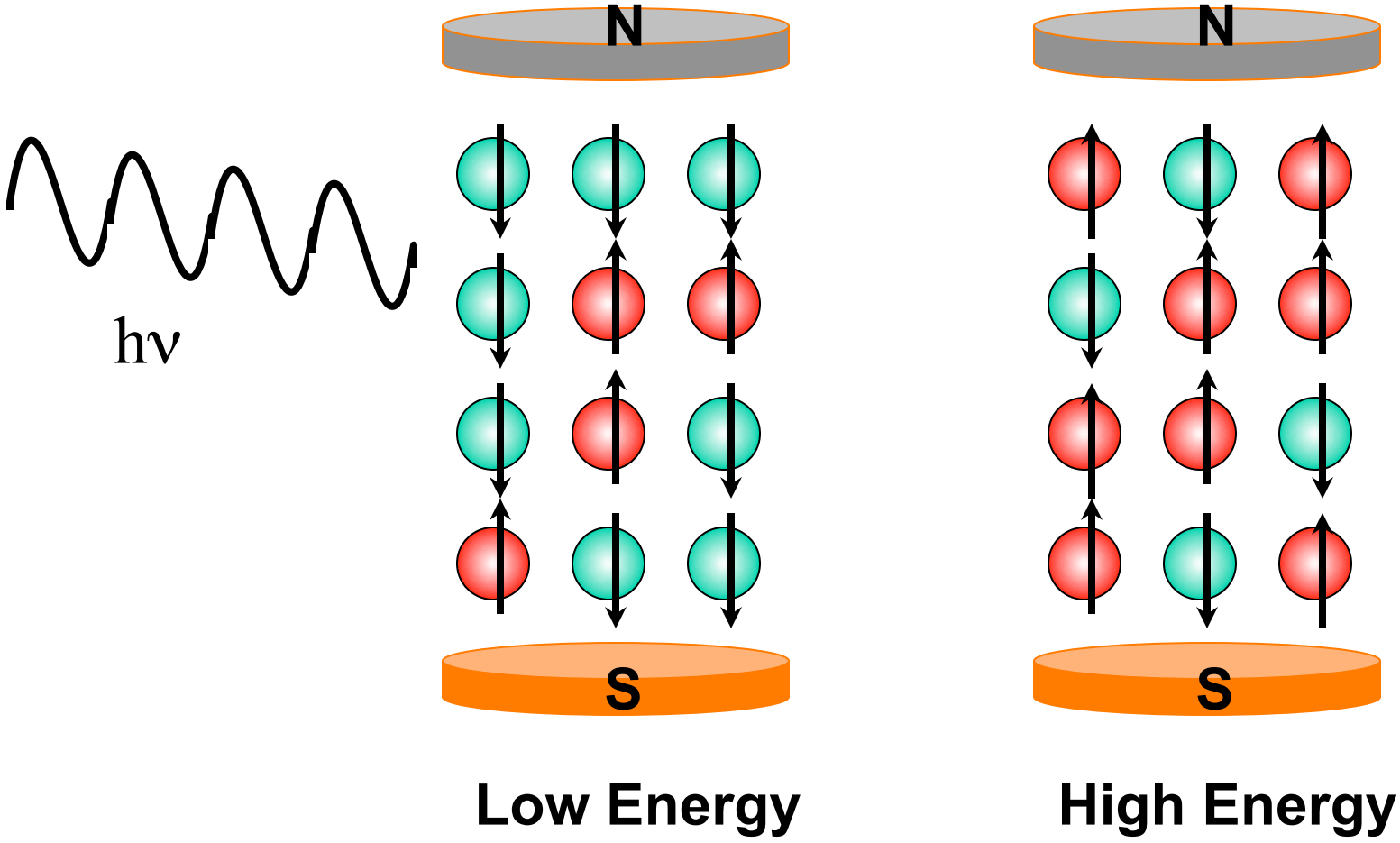
Principles of NMR*

- Measures nuclear magnetism or changes in nuclear magnetism in a molecule
- NMR spectroscopy measures the absorption of light (radio waves) due to changes in nuclear spin orientation
- NMR only occurs when a sample is in a strong magnetic field
- Different nuclei absorb at different energies (frequencies)

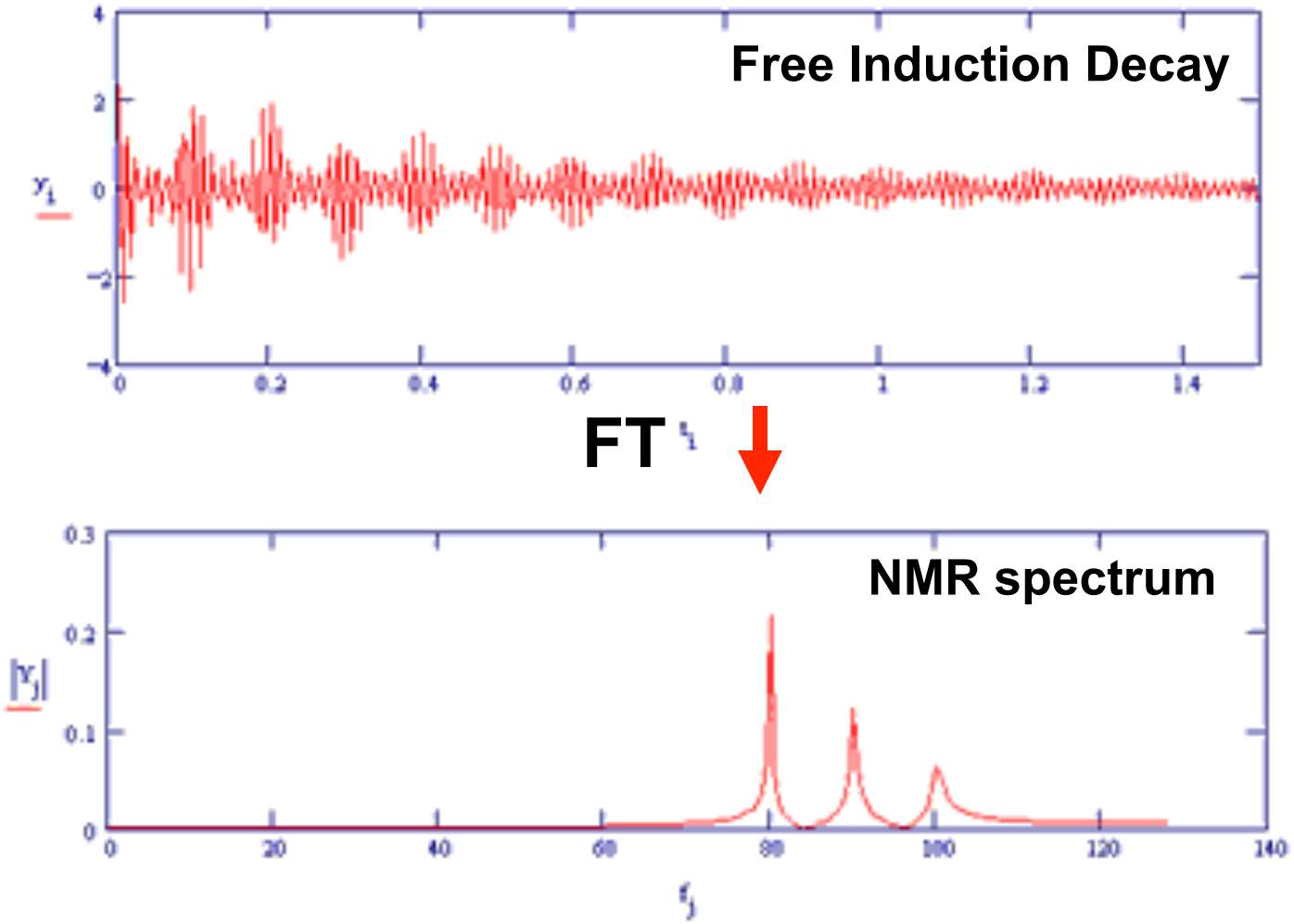
Principles of NMR*



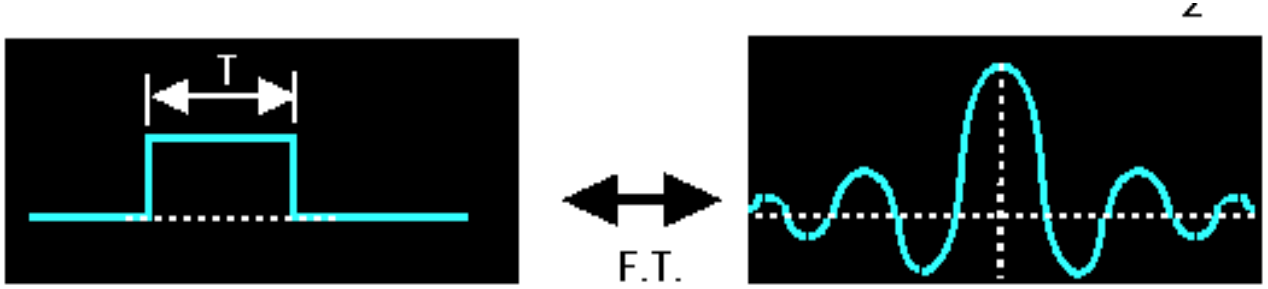
Principles of NMR*

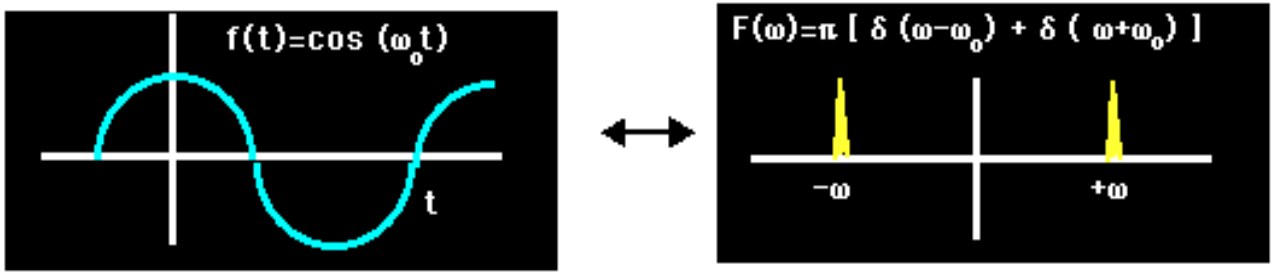


FT NMR



Fourier Transformation*



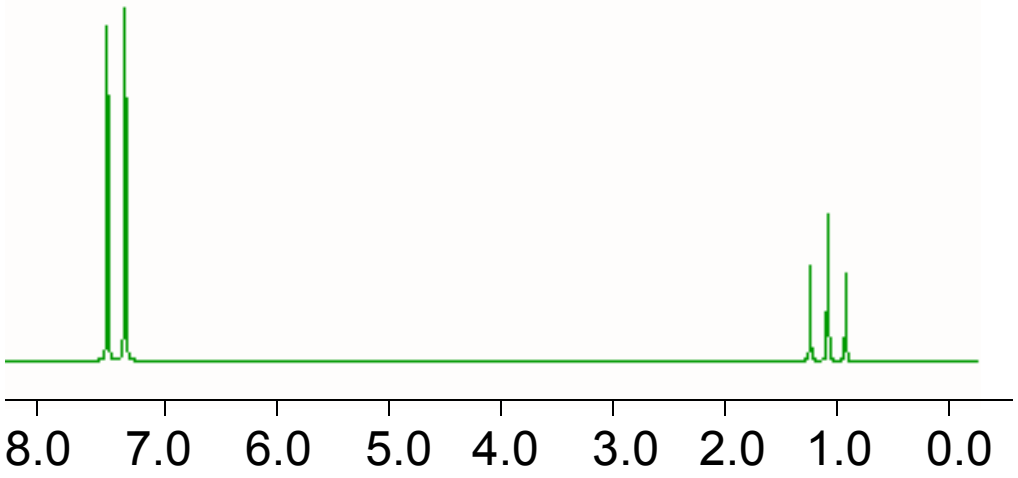


$$F(\omega) = \int f(t) e^{\iota\omega t} dt$$

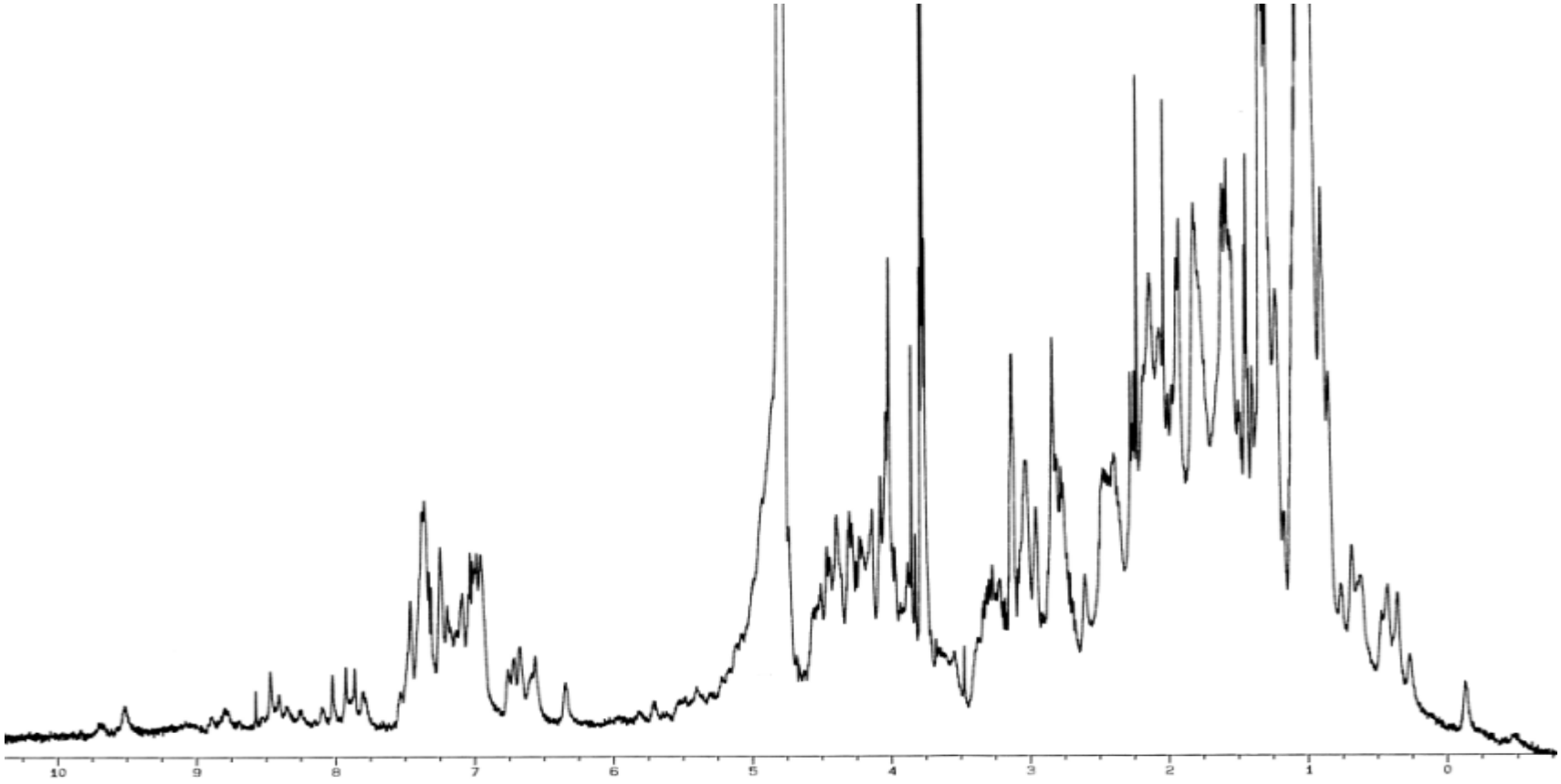
Converts from units of time to units of frequency

^{1}H NMR Spectra Exhibit...*

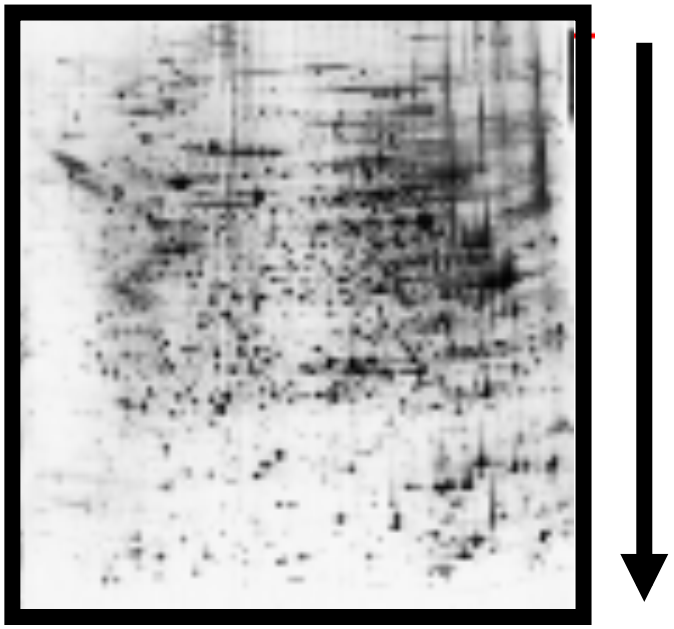
- **Chemical Shifts (peaks at different frequencies or ppm values)**
- **Splitting Patterns (from spin coupling)**
- **Different Peak Intensities (# ^{1}H)**



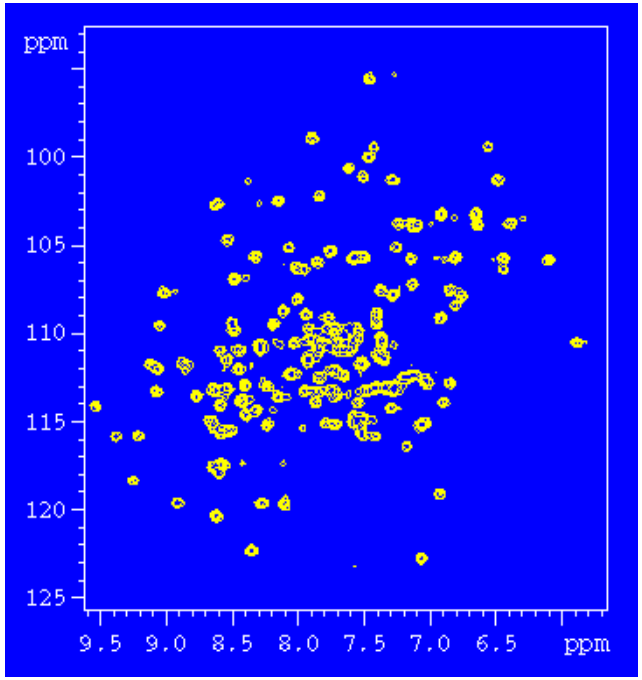
Protein NMR Spectrum



2D Gels & 2D NMR

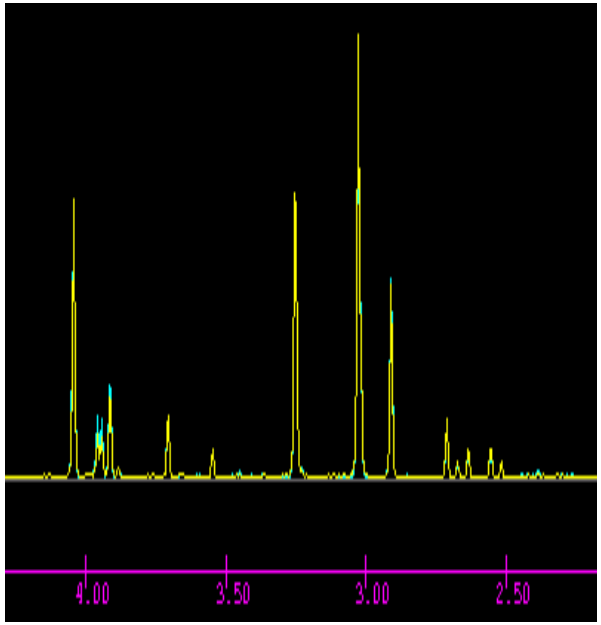


SDS
PAGE



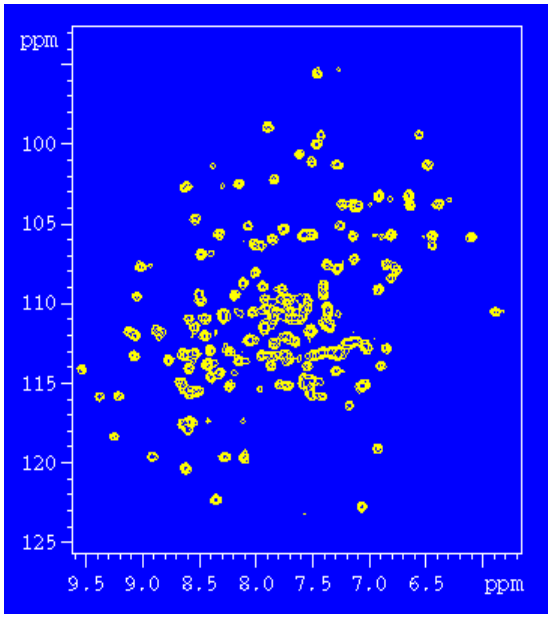
Multidimensional NMR*

1D



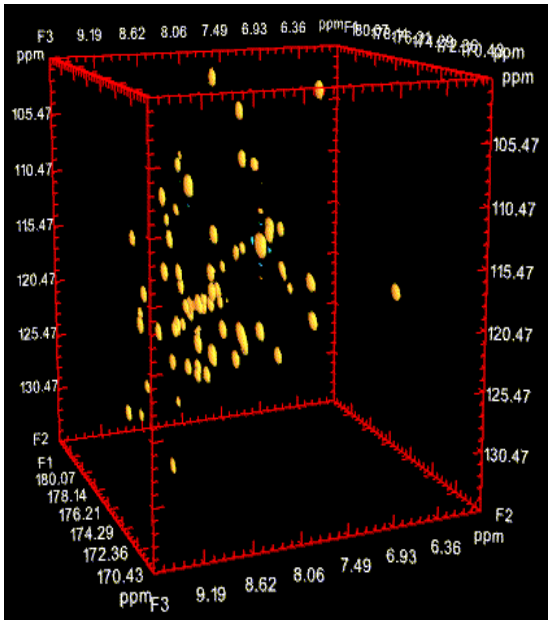
MW ~ 500

2D



MW ~ 10,000

3D

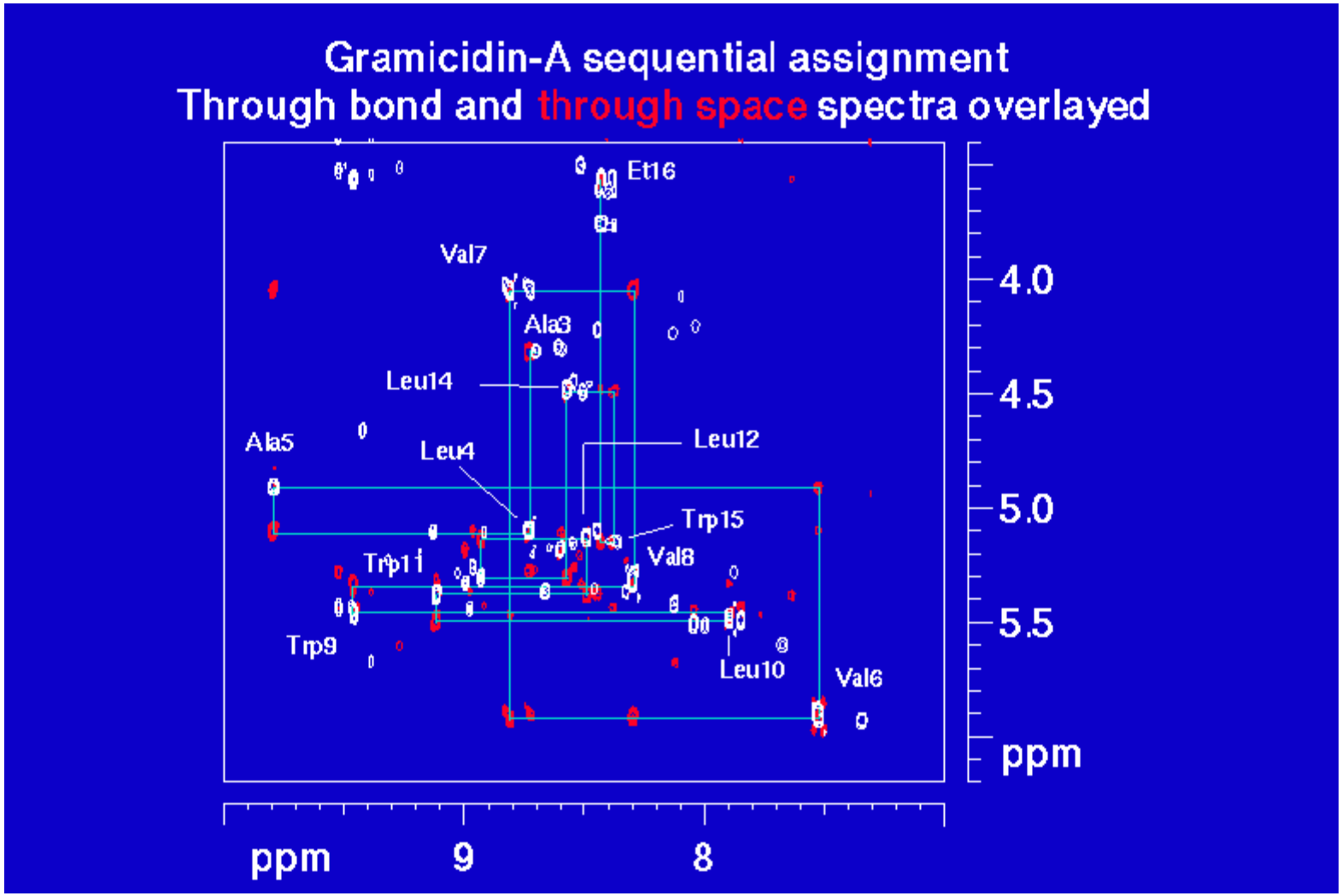


MW ~ 30,000

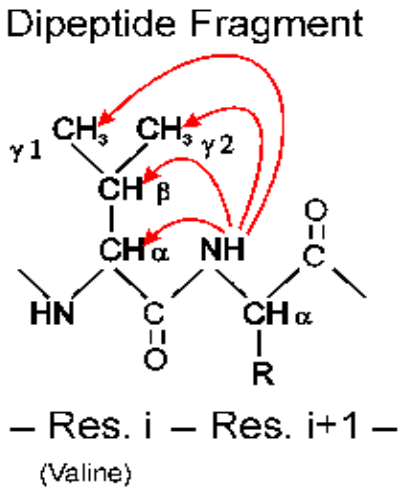
The NMR Process*

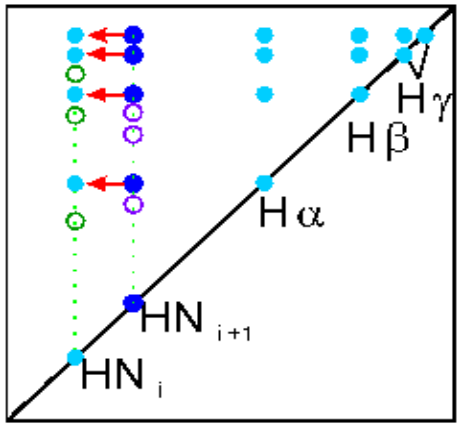
- Obtain protein sequence
- Collect TOCSY & NOESY data
- Use chemical shift tables and known sequence to assign TOCSY spectrum
- Use TOCSY to assign NOESY spectrum
- Obtain inter and intra-residue distance information from NOESY data
- Feed data to computer to solve structure

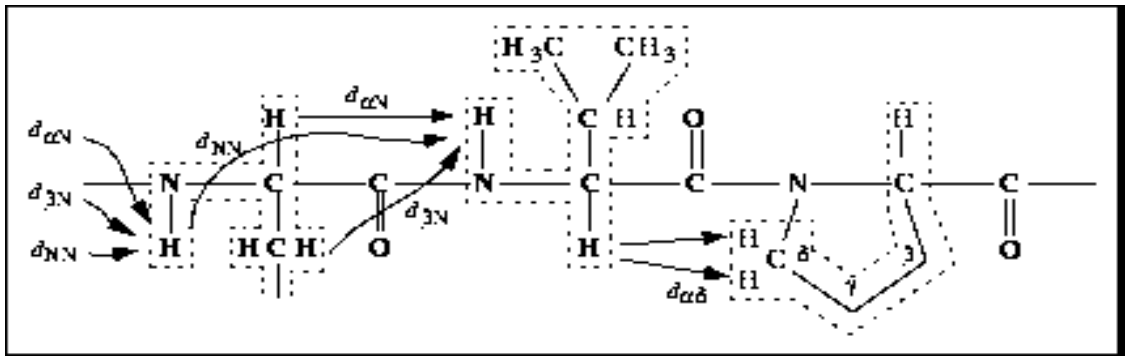
Assigning Chemical Shifts



Measuring NOEs*





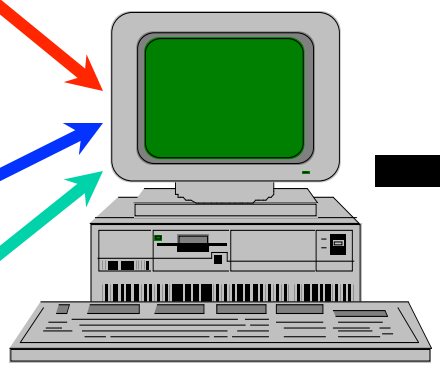


NMR Spectroscopy*

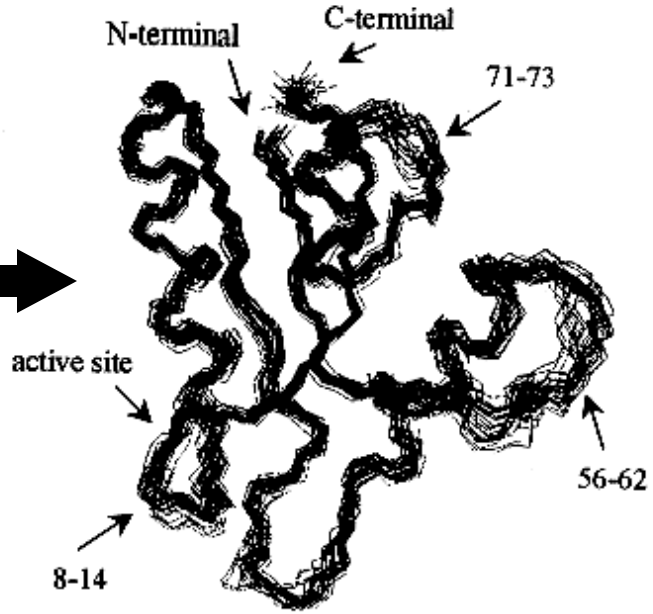
Chemical Shift Assignments

NOE Intensities

J-Couplings



Distance Geometry Simulated Annealing



The Final Result

```
ORIGX2      0.000000  1.000000  0.000000        0.00000                      2TRX 147
ORIGX3      0.000000  0.000000  1.000000        0.00000                      2TRX 148
SCALE1      0.011173  0.000000  0.004858        0.00000                      2TRX 149
SCALE2      0.000000  0.019585  0.000000        0.00000                      2TRX 150
SCALE3      0.000000  0.000000  0.018039        0.00000                      2TRX 151
ATOM      1  N   SER A   1      21.389  25.406  -4.628  1.00 23.22           2TRX 152
ATOM      2  CA  SER A   1      21.628  26.691  -3.983  1.00 24.42           2TRX 153
ATOM      3  C   SER A   1      20.937  26.944  -2.679  1.00 24.21           2TRX 154
ATOM      4  O   SER A   1      21.072  28.079  -2.093  1.00 24.97           2TRX 155
ATOM      5  CB  SER A   1      21.117  27.770  -5.002  1.00 28.27           2TRX 156
ATOM      6  OG  SER A   1      22.276  27.925  -5.861  1.00 32.61           2TRX 157
ATOM      7  N   ASP A   2      20.173  26.028  -2.163  1.00 21.39           2TRX 158
ATOM      8  CA  ASP A   2      19.395  26.125  -0.949  1.00 21.57           2TRX 159
ATOM      9  C   ASP A   2      20.264  26.214   0.297  1.00 20.89           2TRX 160
ATOM     10  O   ASP A   2      19.760  26.575   1.371  1.00 21.49           2TRX 161
ATOM     11  CB  ASP A   2      18.439  24.914  -0.856  1.00 22.14           2TRX 162
```

X-ray Versus NMR*

X-ray

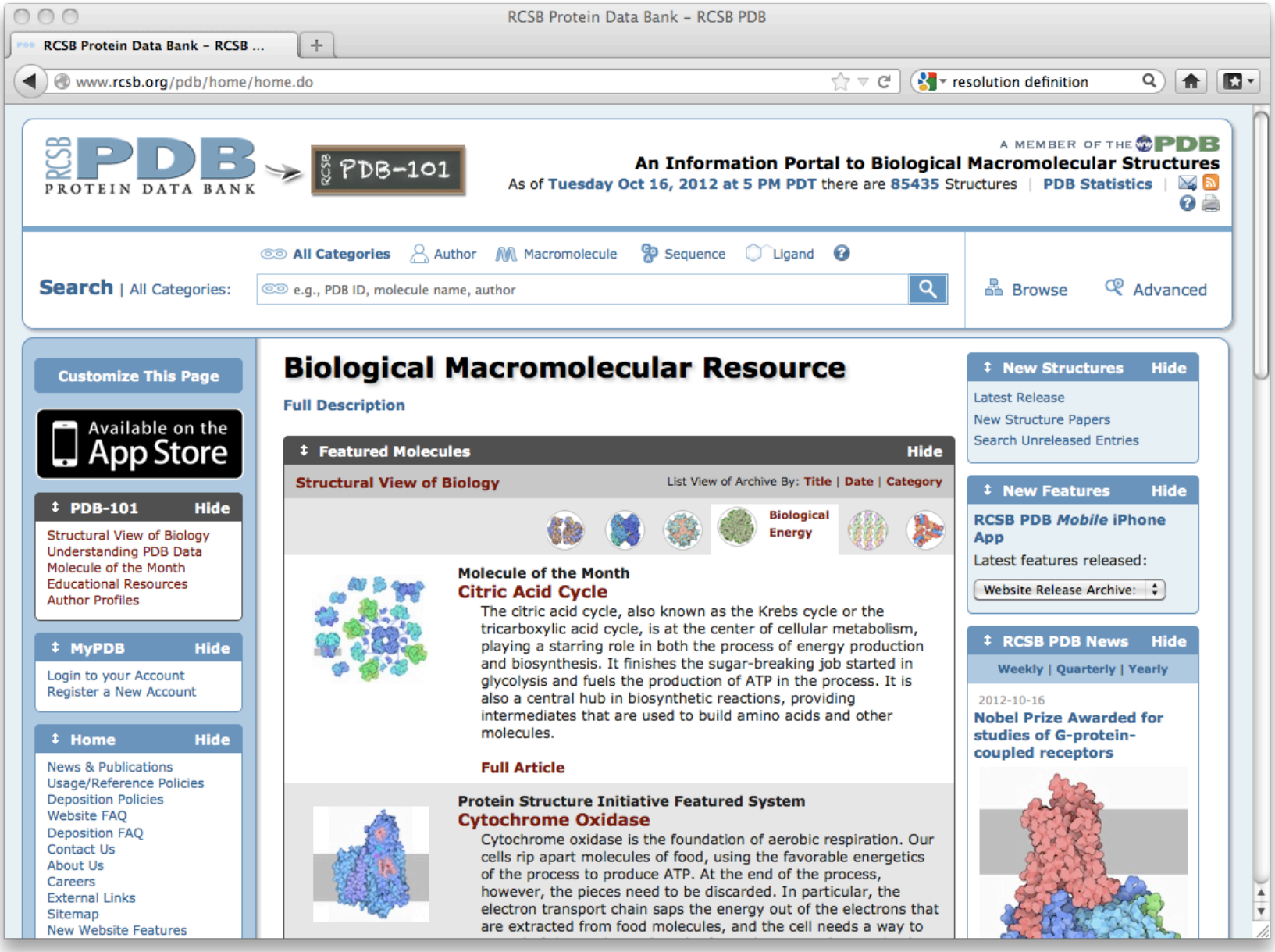
- Producing enough protein for trials
- Crystallization time and effort
- Crystal quality, stability and size control
- Finding isomorphous derivatives
- Chain tracing & checking

NMR

- Producing enough labeled protein for collection
- Sample “conditioning”
- Size of protein
- Assignment process is slow and error prone
- Measuring NOE’s is slow and error prone

The PDB*

- PDB - Protein Data Bank
- Established in 1971 at Brookhaven National Lab (7 structures)
- Primary archive for macromolecular structures (proteins, nucleic acids, carbohydrates)
- Moved from BNL to RCSB (Research Collaboratory for Structural Bioinformatics) in 1998

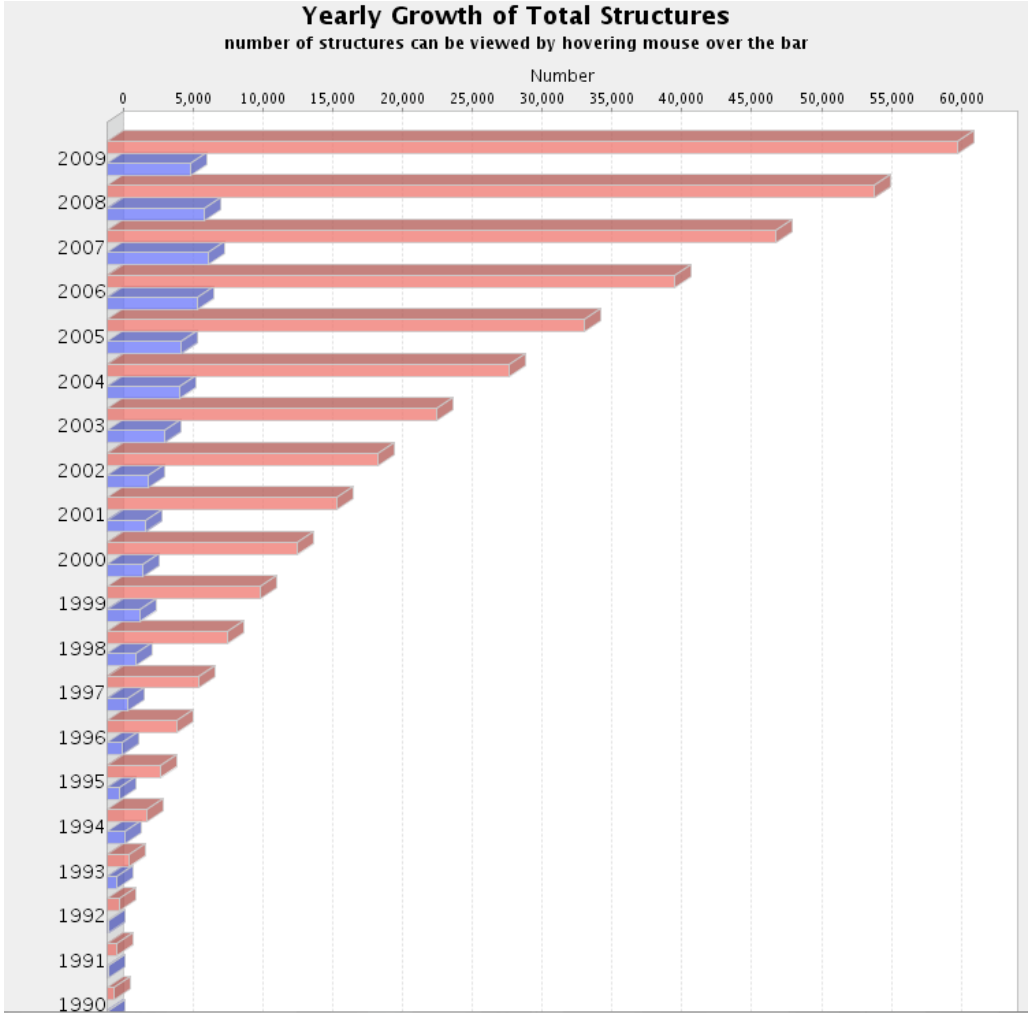


http://www.rcsb.org/pdb/

The PDB

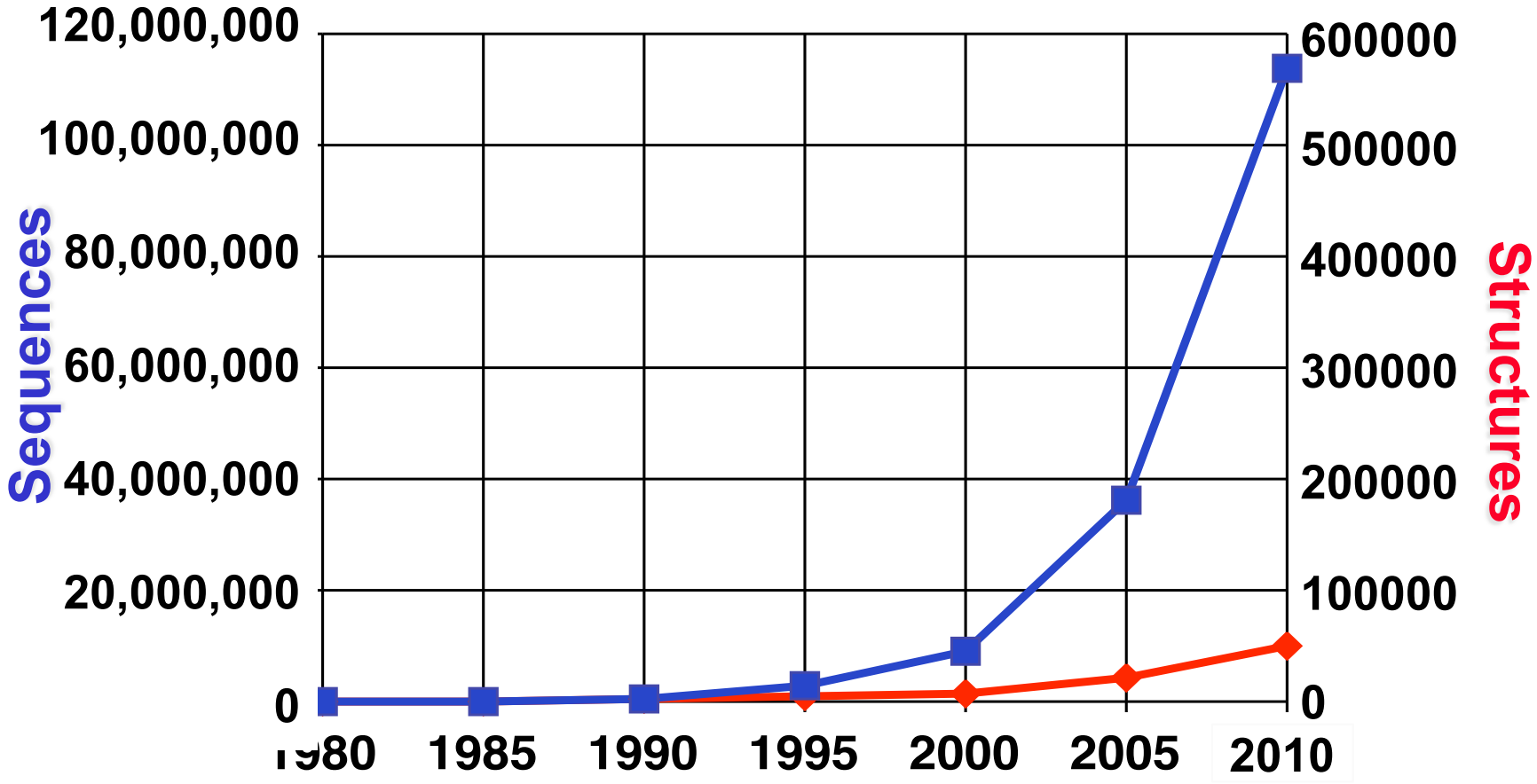
- Contains coordinate data (primarily) from X-ray, NMR and modelling
- Contains files in 2 formats
  - PDB format
  - mmCIF (macrmolecular Crystallographic Information File Format)
- Contains >80,000 entries
- Currently growing exponentially

PDB Growth



- Growth has been exponential for the past 10 years
- Approximately 8000 new structures being added each year

Structural Proteomics: The Motivation



PDB Composition (2009)

	Proteins	Protein/DNA complexes	DNA/RNA
X-ray	51000 (85%)	1636 (88%)	1602 (67%)
NMR	7500 (14%)	182 (11%)	824 (32%)
Models	321 (1%)	24 (1%)	8 (1%)

PDB File Format*

```
HEADER    ELECTRON TRANSPORT                      19-MAR-90   2TRX       2TRXA   1
COMPND    THIOREDOXIN                                                    2TRXA   2
SOURCE    (ESCHERICHIA $COLI)                                            2TRX    4
AUTHOR    S.K.KATTI,D.M.LE*MASTER,H.EKLUND                               2TRX    5
REVDAT   2   15-JAN-93 2TRXA    1       HEADER COMPND                     2TRXA   3
REVDAT   1   15-OCT-91 2TRX     0                                         2TRX    6
JRNL        AUTH   S.K.KATTI,D.M.LE*MASTER,H.EKLUND                      2TRX    7
JRNL        TITL   CRYSTAL STRUCTURE OF THIOREDOXIN FROM ESCHERICHIA     2TRX    8
JRNL        TITL 2 $COLI AT 1.68 ANGSTROMS RESOLUTION                    2TRX    9
JRNL        REF    J.MOL.BIOL.                   V. 212   167 1990       2TRX   10
JRNL        REFN   ASTM JMOBAK  UK ISSN 0022-2836                  070   2TRX   11
REMARK   1
```

HEADER - PDB accession, date, function
CMPND - name of molecule or compound
SOURCE - origin or source of molecule (species)
REVDAT - revision dates
JRNL - primary reference (journal) describing structure
REMARK - a comment made by depositor

PDB File Format*

```
REMARK   6 CORRECTION. CORRECT CLASSIFICATION ON HEADER RECORD AND      2TRXA  5
REMARK   6  REMOVE E.C. CODE.  15-JAN-93.                              2TRXA  6
SEQRES   1 A  108  SER ASP LYS ILE ILE HIS LEU THR ASP ASP SER PHE ASP  2TRX  74
SEQRES   2 A  108  THR ASP VAL LEU LYS ALA ASP GLY ALA ILE LEU VAL ASP  2TRX  75
SEQRES   3 A  108  PHE TRP ALA GLU TRP CYS GLY PRO CYS LYS MET ILE ALA  2TRX  76
SEQRES   4 A  108  PRO ILE LEU ASP GLU ILE ALA ASP GLU TYR GLN GLY LYS  2TRX  77
SEQRES   5 A  108  LEU THR VAL ALA LYS LEU ASN ILE ASP GLN ASN PRO GLY  2TRX  78
SEQRES   6 A  108  THR ALA PRO LYS TYR GLY ILE ARG GLY ILE PRO THR LEU  2TRX  79
SEQRES   7 A  108  LEU LEU PHE LYS ASN GLY GLU VAL ALA ALA THR LYS VAL  2TRX  80
SEQRES   8 A  108  GLY ALA LEU SER LYS GLY GLN LEU LYS GLU PHE LEU ASP  2TRX  81
SEQRES   9 A  108  ALA ASN LEU ALA                                      2TRX  82
HET     CU    109       1     COPPER ++ ION                            2TRX 100
HET     CU    109       1     COPPER ++ ION                            2TRX 101
HET    MPD    601       8     2-METHYL-2,4-PENTANEDIOL                 2TRX 102
HET    MPD    602       8     2-METHYL-2,4-PENTANEDIOL                 2TRX 103
```

REMARK - a comment made by depositor
SEQRES - sequence of protein in 3 letter code
HET - names of heteroatoms

PDB File Format*

```
FORMUL   3   CU    2(CU1 ++)                                                2TRX 110
FORMUL   4  MPD    8(C6 H14 O2)                                             2TRX 111
FORMUL   5  HOH   *140(H2 O1)                                               2TRX 112
HELIX    1 A1A SER A   11  LEU A   17  1 DISORDERED IN MOLECULE B           2TRX 113
HELIX    2 A2A CYS A   32  TYR A   49  1 BENT BY 30 DEGREES AT RES 39       2TRX 114
HELIX    3 A3A ASN A   59  ASN A   63  1                                    2TRX 115
HELIX    4 31A THR A   66  TYR A   70  5 DISTORTED H-BONDING C-TERMINS      2TRX 116
HELIX    5 A4A SER A   95  LEU A  107  1                                    2TRX 117
SHEET    1 B1A 5 LYS A   3  THR A   8  0                                   2TRX 123
SHEET    2 B1A 5 LEU A  53  ASN A  59  1  O  VAL A  55   N  ILE A   5     2TRX 124
SHEET    3 B1A 5 GLY A  21  TRP A  28  1  N  TRP A  28   O  LEU A  58     2TRX 125
SHEET    4 B1A 5 PRO A  76  LYS A  82 -1  O  THR A  77   N  PHE A  27     2TRX 126
SHEET    5 B1A 5 VAL A  86  GLY A  92 -1  N  GLY A  92   O  LYS A  82     2TRX 127
SSBOND   1 CYS A   32    CYS A   35                                         2TRX 143
```

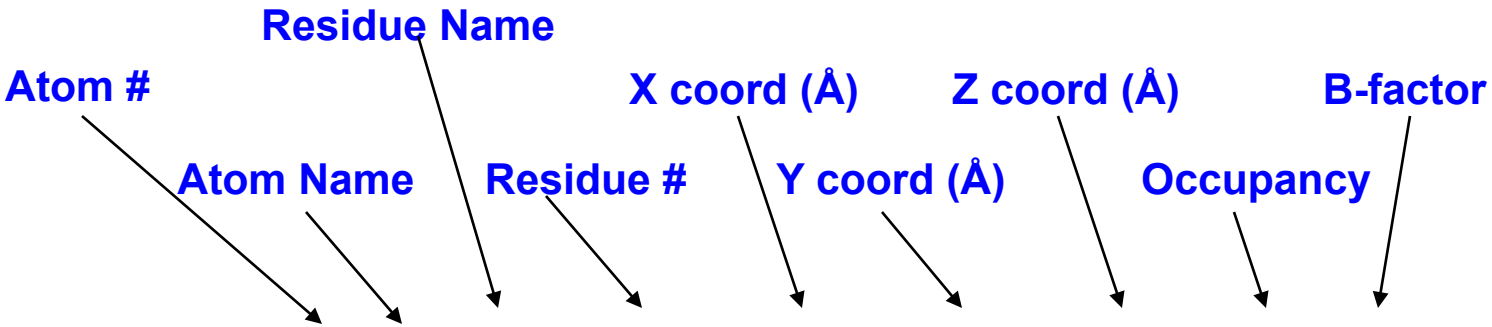
FORMUL - chemical formula of heteroatoms
HELIX - location of helices as identified by depositor
SHEET location of beta sheets as identified by depositor
SSBOND - location and exisitence of disulfide bond

PDB File Format*

```
ORIGX1      1.000000  0.000000  0.000000        0.00000                      2TRX 146
ORIGX2      0.000000  1.000000  0.000000        0.00000                      2TRX 147
ORIGX3      0.000000  0.000000  1.000000        0.00000                      2TRX 148
SCALE1      0.011173  0.000000  0.004858        0.00000                      2TRX 149
SCALE2      0.000000  0.019585  0.000000        0.00000                      2TRX 150
SCALE3      0.000000  0.000000  0.018039        0.00000                      2TRX 151
ATOM      1  N   SER A   1      21.389  25.406  -4.628  1.00 23.22           2TRX 152
ATOM      2  CA  SER A   1      21.628  26.691  -3.983  1.00 24.42           2TRX 153
ATOM      3  C   SER A   1      20.937  26.944  -2.679  1.00 24.21           2TRX 154
ATOM      4  O   SER A   1      21.072  28.079  -2.093  1.00 24.97           2TRX 155
ATOM      5  CB  SER A   1      21.117  27.770  -5.002  1.00 28.27           2TRX 156
ATOM      6  OG  SER A   1      22.276  27.925  -5.861  1.00 32.61           2TRX 157
ATOM      7  N   ASP A   2      20.173  26.028  -2.163  1.00 21.39           2TRX 158
ATOM      8  CA  ASP A   2      19.395  26.125  -0.949  1.00 21.57           2TRX 159
ATOM      9  C   ASP A   2      20.264  26.214   0.297  1.00 20.89           2TRX 160
ATOM     10  O   ASP A   2      19.760  26.575   1.371  1.00 21.49           2TRX 161
```

ORIGXn - scaling factors to transform from orthogonal coords.
SCALEn - scaling factors to transform to fractional cryst. Coords.
ATOM - atomic coordinates of molecule

PDB File Format*




```
ATOM      1  N   SER A   1      21.389  25.406  -4.628  1.00 23.22      2TRX 152
ATOM      2  CA  SER A   1      21.628  26.691  -3.983  1.00 24.42      2TRX 153
ATOM      3  C   SER A   1      20.937  26.944  -2.679  1.00 24.21      2TRX 154
ATOM      4  O   SER A   1      21.072  28.079  -2.093  1.00 24.97      2TRX 155
ATOM      5  CB  SER A   1      21.117  27.770  -5.002  1.00 28.27      2TRX 156
ATOM      6  OG  SER A   1      22.276  27.925  -5.861  1.00 32.61      2TRX 157
ATOM      7  N   ASP A   2      20.173  26.028  -2.163  1.00 21.39      2TRX 158
ATOM      8  CA  ASP A   2      19.395  26.125  -0.949  1.00 21.57      2TRX 159
ATOM      9  C   ASP A   2      20.264  26.214   0.297  1.00 20.89      2TRX 160
ATOM     10  O   ASP A   2      19.760  26.575   1.371  1.00 21.49      2TRX 161
```

PDB File Format

- Spacing is critical (Fortran compatible)
- Often inconsistent (40+ years old)
- Watch for unusual residues (ACE, SME)
- Some files have 1 structure (X-ray), others have 2 structures (chain A and B in unit cell), others have >20 (NMR)
- Some have missing atoms, others have hydrogens, others don't

Structure File Conversion

Alchemy (t)	AMBER PREP (prep)	Ball and Stick (bs)
Biosym .CAR (car)	Boogie (boog)	Cacao Cartesian (caccrt)
Cambridge CADPAC (cadpac)	CHARMm (charmm)	Chem3D Cartesian 1 (c3d1)
Chem3D Cartesian 2 (c3d2)	CSD CSSR (cssr)	CSD FDAT (fdat)
CSD GSTAT (c)	Feature (feat)	Free Form Fractional (f)
GAMESS Output (famout)	Gaussian Z-Matrix (g)	Gaussian Output (gauout)
Hyperchem (hin)	MDL Isis (isis)	Mac Molecule(macmol)
Macromodel (k)	Micro World (micro)	MM2 Input (mi)
MM2 Ouput (mo)	MM3 (mm3)	MMADS (mmads)
MDL MOLfile (mdl)	MOLIN (molen)	Mopac Cartesian (ac)
Mopac Internal (ai)	Mopac Output (ao)	PC Model (pc)
PDB (p)	Quanta (quanta)	ShelX (shelx)
Spartan (spar)	Spartan Semi-Empirical (semi)	Spartan Mol. Mechanics (spmm
Sybyl Mol (mol)	Sybyl Mol2 (mol2)	Conjure (con)
Maccs 2d (maccs2)	Maccs 3d (maccs3)	UniChem XYZ (unixyz)
XYZ (x)	XED (xed)	

http://smog.com/chem/babel/

RCSB PDB

http://www.pdb.org/pdb/home/home.do

Most Visited | Getting Started | Latest Headlines

RCSB PDB

RCSB PDB PROTEIN DATA BANK → PDB-101

A MEMBER OF THE PDB

An Information Portal to Biological Macromolecular Structures

As of Tuesday Nov 01, 2011 at 5 PM PDT there are 76970 Structures | PDB Statistics

Type in thioredoxin

All Categories | Author | Macromolecule | Sequence | Ligand

Search | All Categories: e.g., PDB ID, molecule name, author

Browse | Advanced

Customize This Page

MyPDB Hide

Login to your Account
Register a New Account

Home Hide

News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us
About Us
Careers
External Links
Sitemap
New Website Features

Deposition Hide

All Deposit Services
Electron Microscopy
X-ray | NMR
Validation Server
BioSync Beamlines/Facilities
Related Tools

Tools Hide

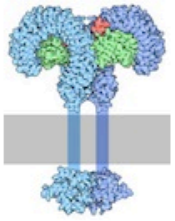
Biological Macromolecular Resource

Full Description

Featured Molecules Hide

Structural View of Biology — List View of Archive By: Title | Date | Category

Enzymes

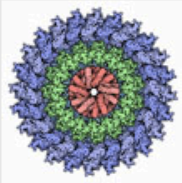


Molecule of the Month

Toll-like Receptors

The world is filled with bacteria and viruses, all eager to infect our cells. We have two lines of defense against this constant assault. Our first defense is the innate immune system, which stands guard against the most common attackers and mounts a quick defense when they are found. This innate system is found widely in animals, plants, and fungi, and for most, is the only line of defense.

Full Article



Protein Structure Initiative Featured System

The Perils of Protein Secretion

Salmonella bacteria attack cells by injecting deadly proteins. PSI researchers are revealing how these proteins work, and how the bacteria control their action.

Full Article | Archive | PSI Structural Biology Knowledgebase

New Structures Hide

Latest Release
New Structure Papers
Search Unreleased Entries

New Features Hide

PDB-101: RSS Feed

Latest features released:

Website Release Archive:

RCSB PDB News Hide

Weekly | Quarterly | Yearly

2011-11-01

New Target Registration Database

Done

RCSB PDB – Query Results

http://www.pdb.org/pdb/results/results.do?qrid=BF2C3165&tabtoshow=Current

Most Visited | Getting Started | Latest Headlines

RCSB PDB – Query Results

MyPDB Hide
- Login to your Account
- Register a New Account
- Query Results (538)
- Query History (1)

Home Hide
- News & Publications
- Usage/Reference Policies
- Deposition Policies
- Website FAQ
- Deposition FAQ
- Contact Us
- About Us
- Careers
- External Links
- Sitemap
- New Website Features

Deposition Hide
- All Deposit Services
- Electron Microscopy
- X-ray | NMR
- Validation Server
- BioSync Beamlines/Facilities
- Related Tools

Tools Hide
- Download Files
- Compare Structures
- File Formats
- Services: RESTful | SOAP
- Widgets

PDB-101 Hide
- Structural View of Biology
- Understanding PDB Data
- Molecule of the Month
- Educational Resources

Help Hide

538 Structure Hits | 13 Unreleased Structures | 248 Citations | 127 Ligand Hits | 3 Web Page Hits

Query Parameters: Query Details | Save Query to MyPDB

Text Search for: thioredoxin

Query Refinements Show

Organism | Taxonomy | Exp. Method | X-Ray Resolution | Release Date | Polymer Type | Enzyme Classification | SCOP Classification

Refine Query — Remove Similar: Select Percent Similarity

Display/Download: | Generate Reports: | Sort by: Revelance | Results per Page:

Displaying results 1 - 25 of 538 total | Page 1 of 22 | Jump to page: GO

Click here to learn more

2L5L — **Solution Structure of Thioredoxin from Bacteroides Vulgatus**

Authors: Harris, R., Foti, R., Seidel, R.D., Bonanno, J.B., Freeman, J., Bain, K.T., Sauder, J.M., Burley, S.K., Girvin, M.E., Almo, S.C., New York SGX Research Center for Structural Genomics,

Release Date: 2010-11-24 **Classification:** Transport Protein

Experiment: SOLUTION NMR

Compound: 1 Polymer [Display Full Polymer Details | Display for All Results]

Citation: PubMed ID is not available.

3HZ4 — **Crystal Structure of Thioredoxin from Methanosarcina mazei**

Authors: Syed Ibrahim, B., Burley, S.K., Swaminathan, S., New York SGX Research Center for Structural Genomics,

Release Date: 2009-06-30 **Classification:** Oxidoreductase

Experiment: X-RAY DIFFRACTION with resolution of 2.30 Å

Compound: 1 Polymer [Display Full Polymer Details | Display for All Results]

Done

Click on the tabs to learn more

MyPDB Hide
Login to your Account
Register a New Account
Query Results (538)
Query History (1)

Home Hide
News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us
About Us
Careers
External Links
Sitemap
New Website Features

Deposition Hide
All Deposit Services
Electron Microscopy
X-ray | NMR
Validation Server
BioSync Beamlines/Facilities
Related Tools

Tools Hide
Download Files
Compare Structures
File Formats
Services: RESTful | SOAP
Widgets

PDB-101 Hide
Structural View of Biology
Understanding PDB Data
Molecule of the Month
Educational Resources

Summary | Sequence | Annotations | Seq. Similarity | 3D Similarity | Literature | Biol. & Chem. | Methods | Geometry | Links

CRYSTAL STRUCTURE OF THIOREDOXIN FROM ESCHERICHIA COLI AT 1.68 ANGSTROMS RESOLUTION

2TRX

Display Files
Download Files
Share this Page

Annotation data related to this entry.

Domain Annotation: SCOP Classification (version 1.75) Hide

Domain Info	Class	Fold	Superfamily	Family	Domain	Species
d2trxa_	Alpha and beta proteins (a/b)	Thioredoxin fold	Thioredoxin-like	Thioltransferase	Thioredoxin	Escherichia coli [TaxId: 562]
d2trxb_	Alpha and beta proteins (a/b)	Thioredoxin fold	Thioredoxin-like	Thioltransferase	Thioredoxin	Escherichia coli [TaxId: 562]

Domain Annotation: CATH Classification (version 3.4.0) Hide

Domain	Class	Architecture	Topology	Homology
2trxA00	Alpha Beta	3-Layer(aba) Sandwich	Glutaredoxin	Glutaredoxin
2trxB00	Alpha Beta	3-Layer(aba) Sandwich	Glutaredoxin	Glutaredoxin

Protein Family Annotation: PFAM Classification Hide

Chain	PFAM Accession	PFAM ID	Description	Type	Clan ID
A	PF00085	Thioredoxin	Thioredoxin	Domain	
B	PF00085	Thioredoxin	Thioredoxin	Domain	

Gene Product Annotation: GO Terms Hide

Polymer	Molecular Function	Biological Process	Cellular Component
THIOREDOXIN (2TRX:A,B)	○ protein binding ○ protein disulfide	○ transport ○ electron transport chain	○ none

All Deposit Services
Electron Microscopy
X-ray | NMR
Validation Server
BioSync Beamlines/Facilities
Related Tools

Tools Hide
Download Files
Compare Structures
File Formats
Services: RESTful | SOAP
Widgets

PDB-101 Hide
Structural View of Biology
Understanding PDB Data
Molecule of the Month
Educational Resources

Help Hide
Launch Help System
Display Settings
Video Tutorials
Glossary of Terms
PDB*Mobile* FAQ

Chain A : THIOREDOXIN

FASTA | **Sequence & DSSP** | **Image**
Polymer 1
Length: 108 residues
Chain Type: polypeptide(L)
Reference: **UniProtKB P0AA25**

Display Parameters

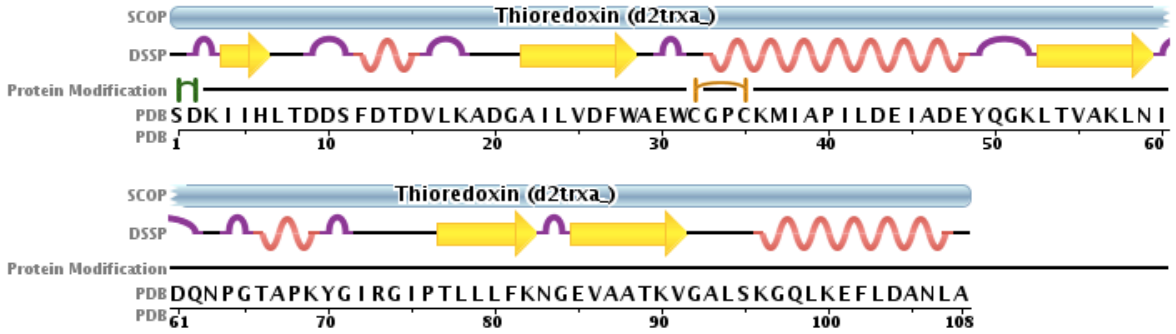
Identical chains: **B** | **show all chains**
Currently displayed: **SEQRES sequence.**
Display external (UniProtKB) sequence
Mouse over an annotation to see more details.

Annotations *Add Annotations* Select

Domain Assignment:**SCOP** [hide] [reference] — d2trxa_ Thioredoxin: 108 residues

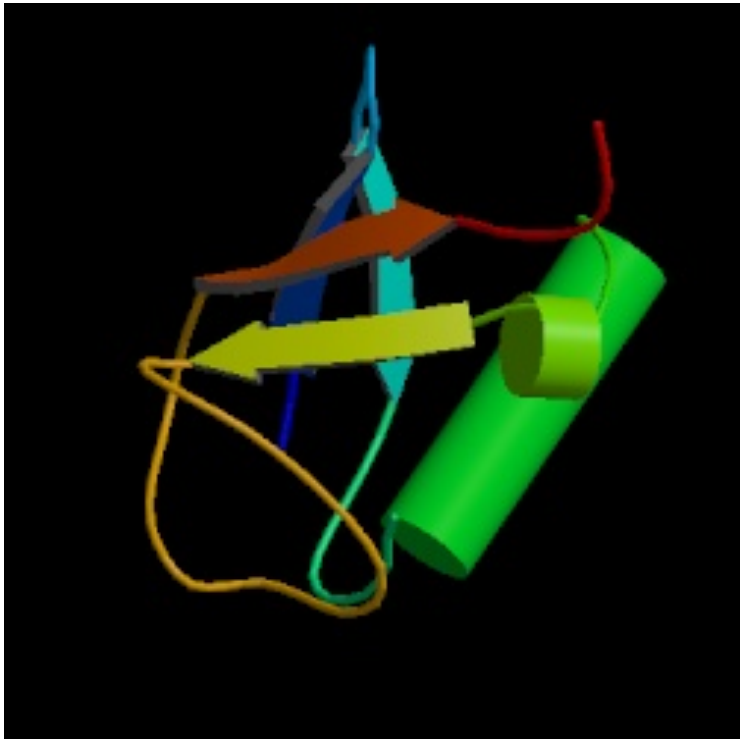
Secondary Structure:**DSSP** [hide] [reference] — 33% helical (4 helices; 36 residues); 27% beta sheet (5 strands; 30 residues)

Structural Feature:**Protein Modification** [hide] [reference] [reference] — 332 copper (II) ion (Metal coordination, CU) *PDB*:**CU**; 0018 L-cystine (redox-active center, disulfide bond) *RESID*:**AA0025**; *PSI-MOD*:**MOD:00034**

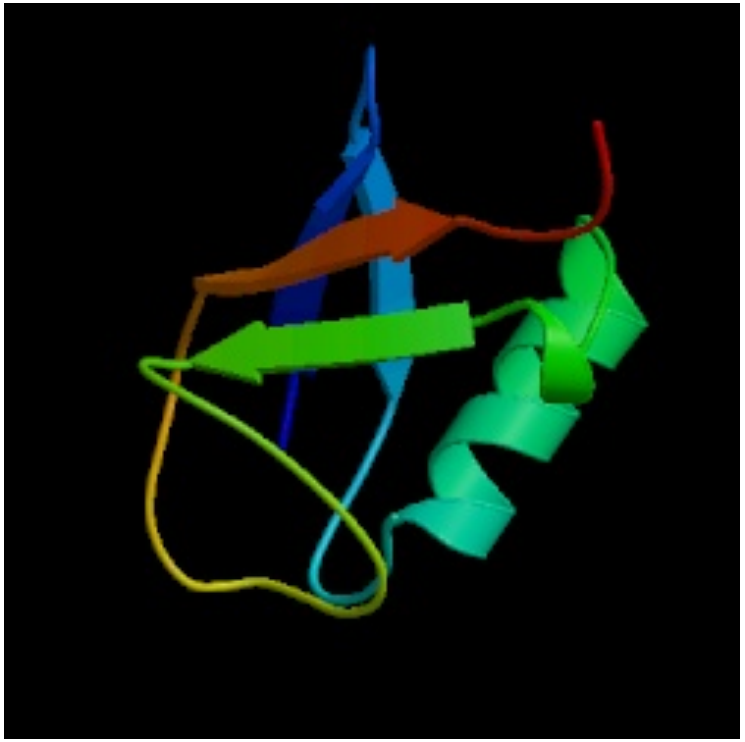


Protein Modification Legend
L-cystine (redox-active center, disulfide bond)
copper (II) ion (Metal coordination, CU)

Protein Rendering*



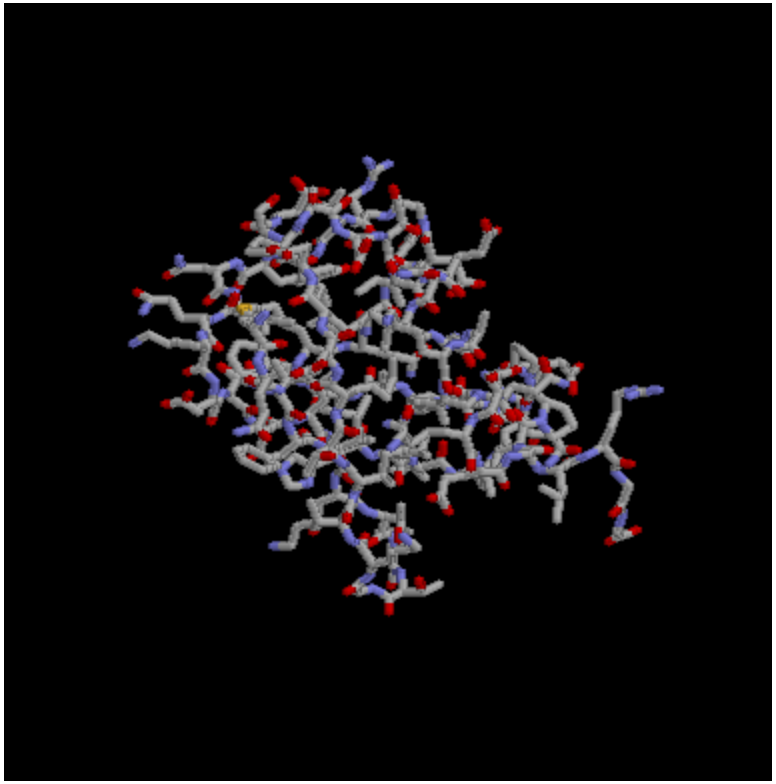
Cylinder



Ribbon (N-C gradient)

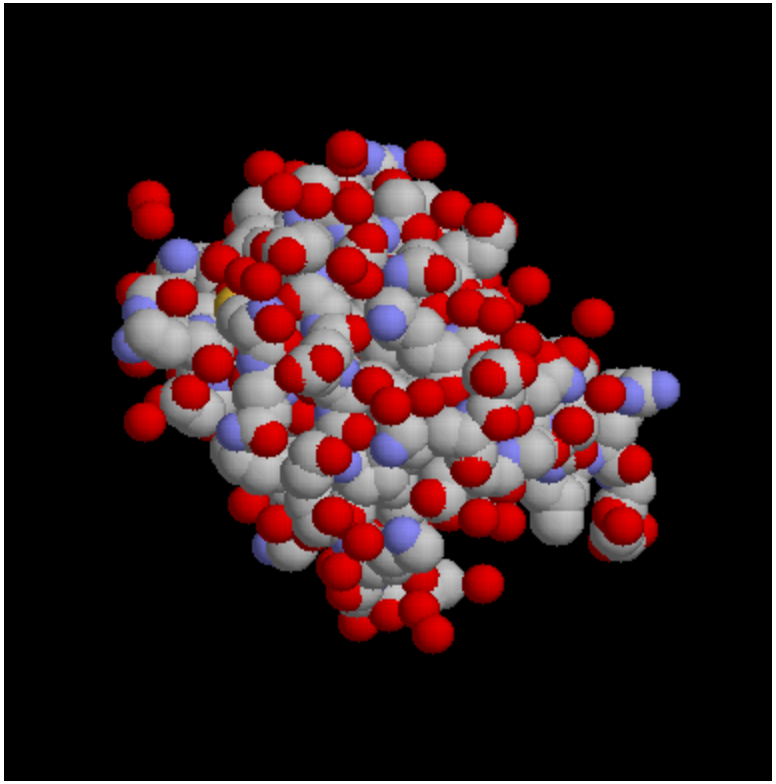
Protein Rendering*

Ribbon (2º structure)

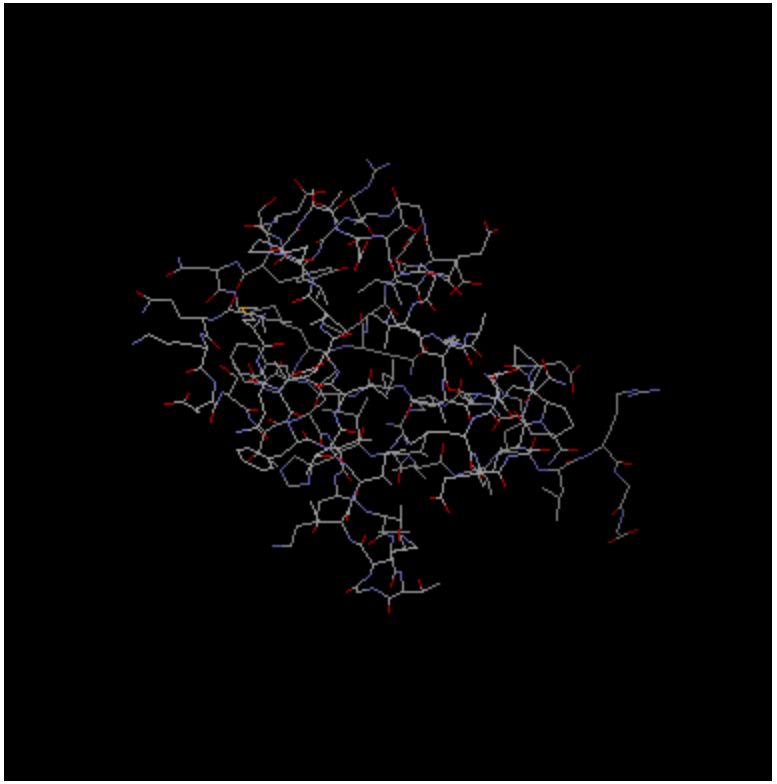


Stick

Protein Rendering*



Space Filling



Wire Frame (Vector)

Summary

- **3D structure of proteins is determined by side chains and sequence of constituent amino acids**
- **X-ray and NMR are the only methods that can yield high resolution 3D structures at atomic scale**
- **The PDB houses essentially all known atomic structures of proteins and DNA/RNA**
- ***Homework – spend 10-15 minutes browsing through the PDB and learn how to use it***