

# **CHAPTER 4**

---

# **THE THREE- DIMENSIONAL STRUCTURE OF PROTEINS**

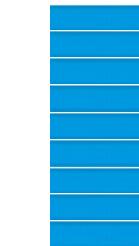
# Globular protein structures are very compact

Human serum albumin, 585 residues

---

**$\beta$  Conformation**

**$2,000 \times 5 \text{ \AA}$**



---

**$\alpha$  Helix**

**$900 \times 11 \text{ \AA}$**

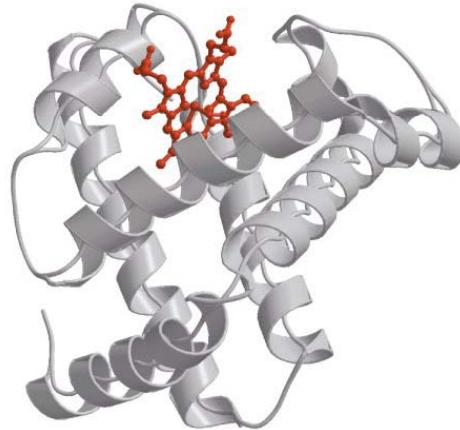


**Native globular form**

**$100 \times 60 \text{ \AA}$**

**The polypeptide  
chains must be very  
compactly folded**

# Tertiary structure of sperm whale myoglobin

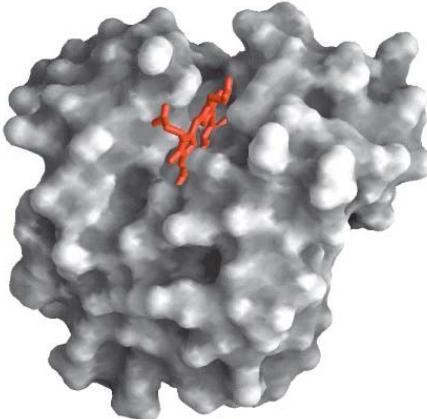


Ribbon  
(Cartoon)  
representation

## Secondary structure

$\alpha$ -helices: coiled ribbons or thick tubes  
 $\beta$ -strands: arrows  
Lines or thin tubes: random coils.

Originally conceived by  
Jane S. Richardson in  
1980s

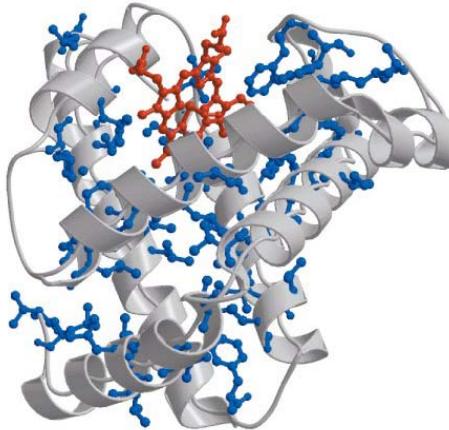


Surface contour  
image

Visualizing  
pockets where  
other molecules  
might bind

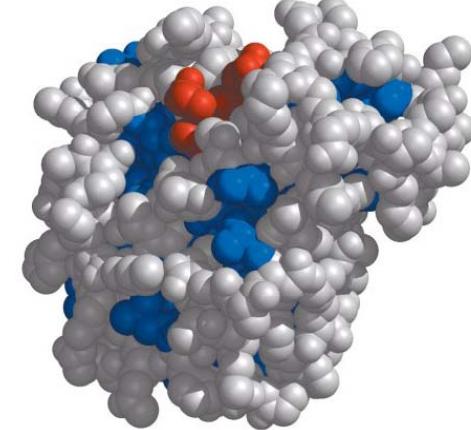


**Myoglobin: 153 a.a., 16.7 kDa**  
Revealed by John Kendrew et al. in 1950s



Ribbon including  
side chain of  
hydrophobic  
residues

Visualizing the  
cueing direction  
of the individual  
side chain



Space-filling  
model

All amino acid  
including side  
chains are  
represented by  
their van der  
Waals radii

Most hydrophobic  
residues are buried  
inside

# For their studies of the structures of globular proteins



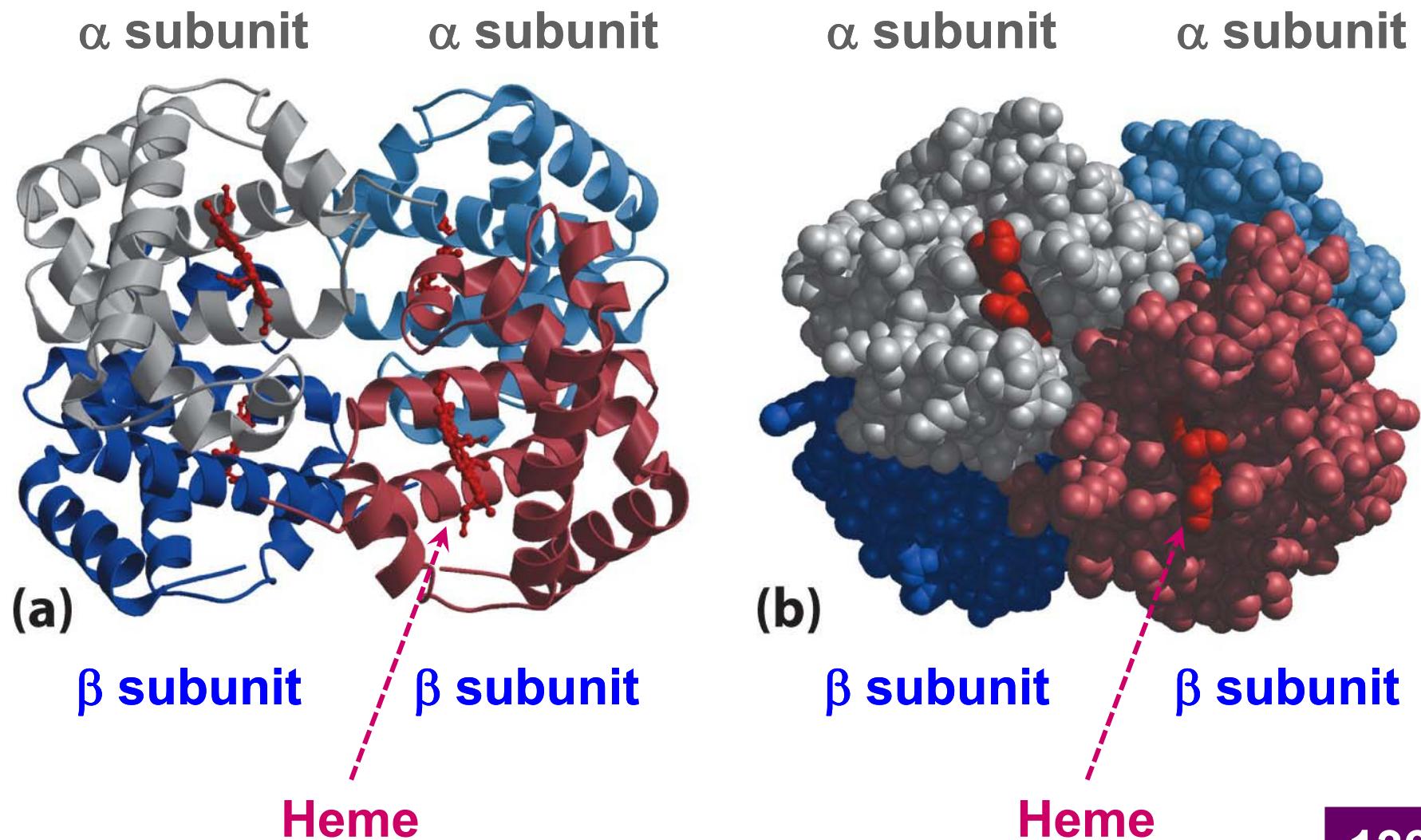
1962

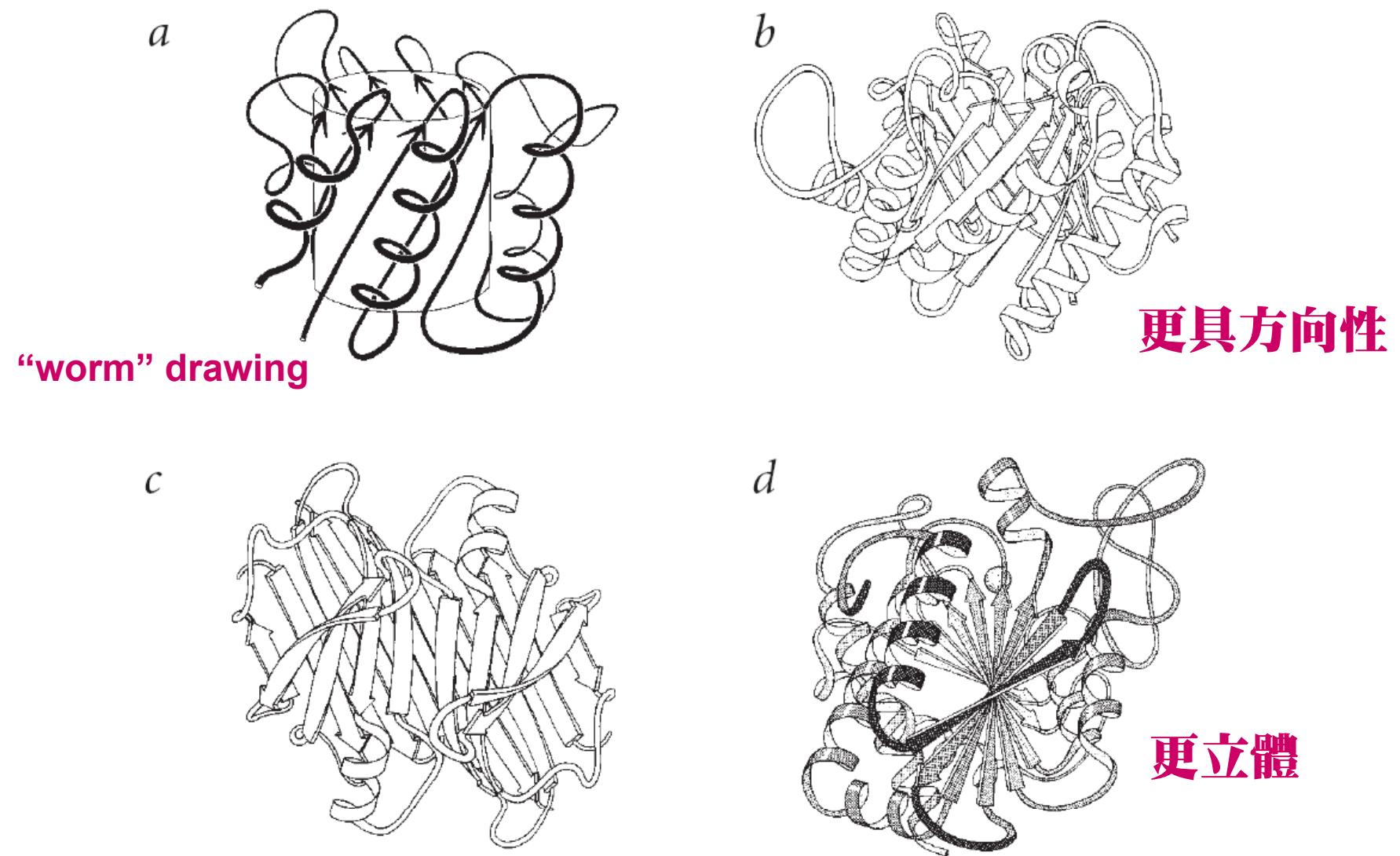


Max Perutz, 1914-2002 (left)   John/kendrew, 1917-1997 (right)

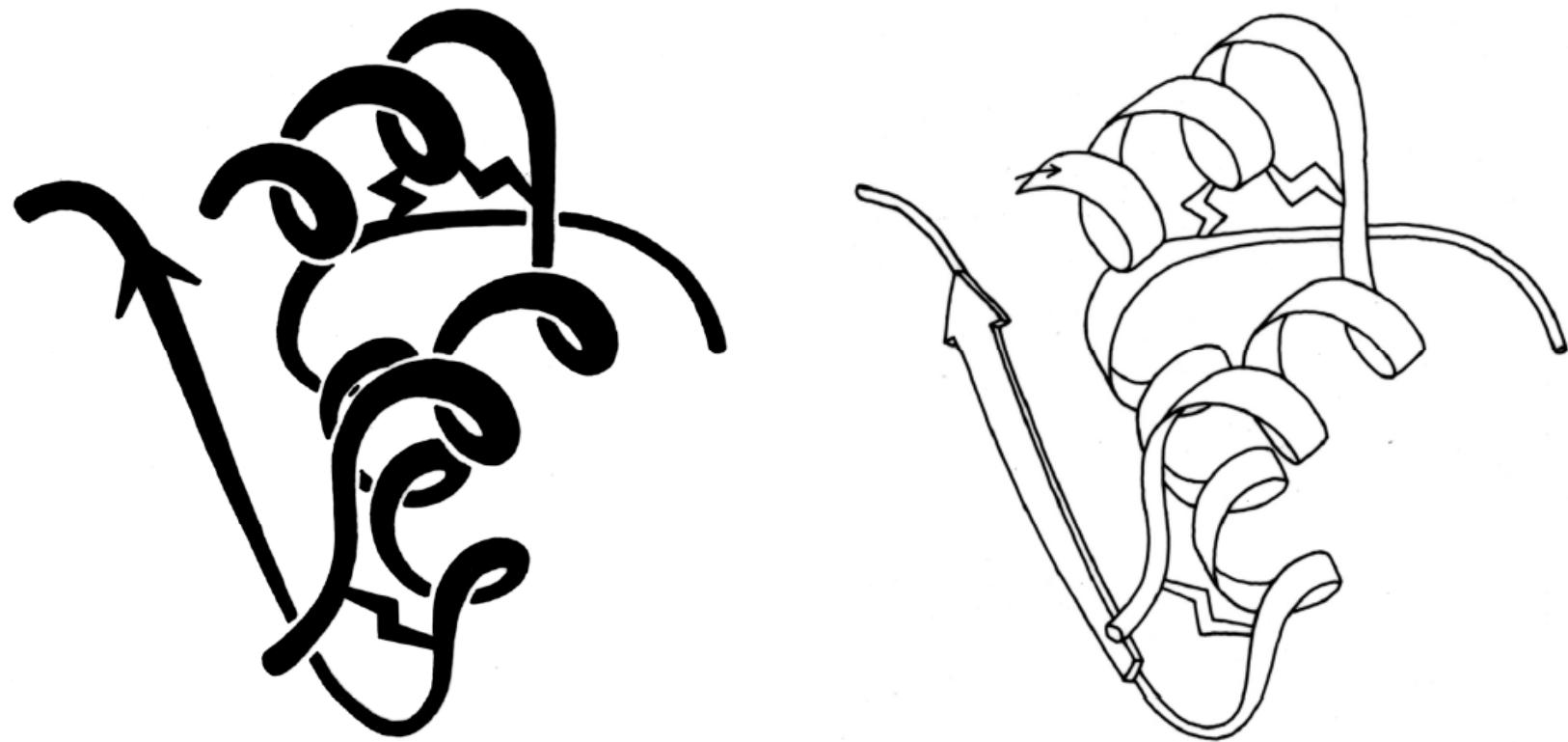
**Perutz and Kendrew won the Nobel Prize in 1962 for solving the structures of hemoglobin (Perutz) and myoglobin (Kendrew). This is the same year that Watson, Crick, and Wilkins won for the structure of DNA. Recall that Watson & Crick were working in the Perutz lab at the time of their discovery and Crick was actually working on the structure of hemoglobin as part of his Ph.D. thesis**

# Quaternary structure of deoxyhemoglobin





**Fig. 1** Some early schematic drawings of protein structures. **a**, Triose phosphate isomerase 'worm' drawing<sup>5</sup>. **b**, Triose phosphate isomerase ribbon drawing<sup>6</sup>. **c**, Prealbumin dimer<sup>6</sup>. **d**, Carboxypeptidase A<sup>6</sup>.



*Insulin*

*FIG. 52. A schematic backbone drawing of insulin, a small structure which is dependent on its disulfides for stability.*



FIG. 8. Sketch of  $\beta$ -strand arrows superimposed on the  $\alpha$ -carbon backbone.

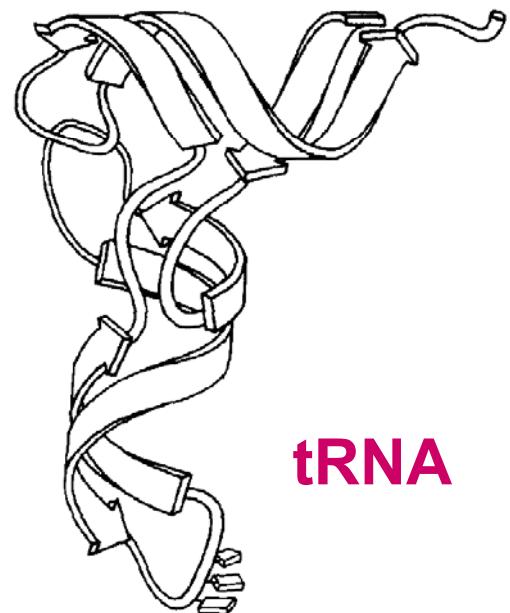


FIG. 2. Yeast tRNA<sup>Phe</sup>: schematic drawing of a nucleic acid structure.

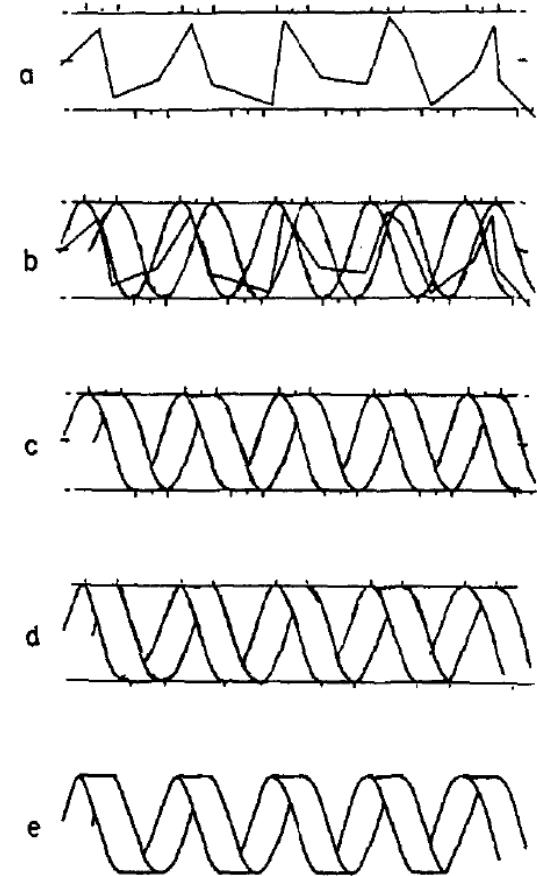


FIG. 5. Successive steps in drawing a spiral ribbon  $\alpha$ -helix.

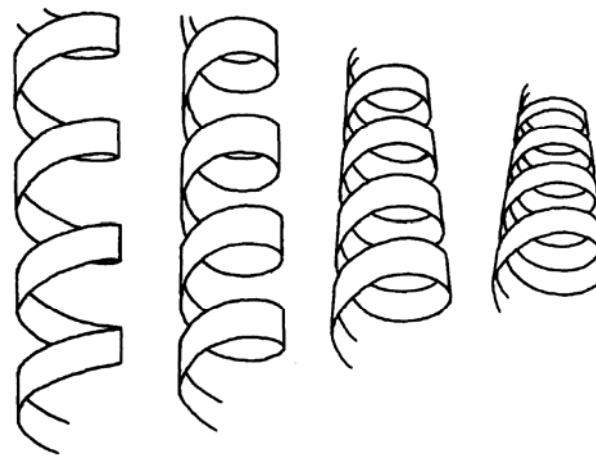
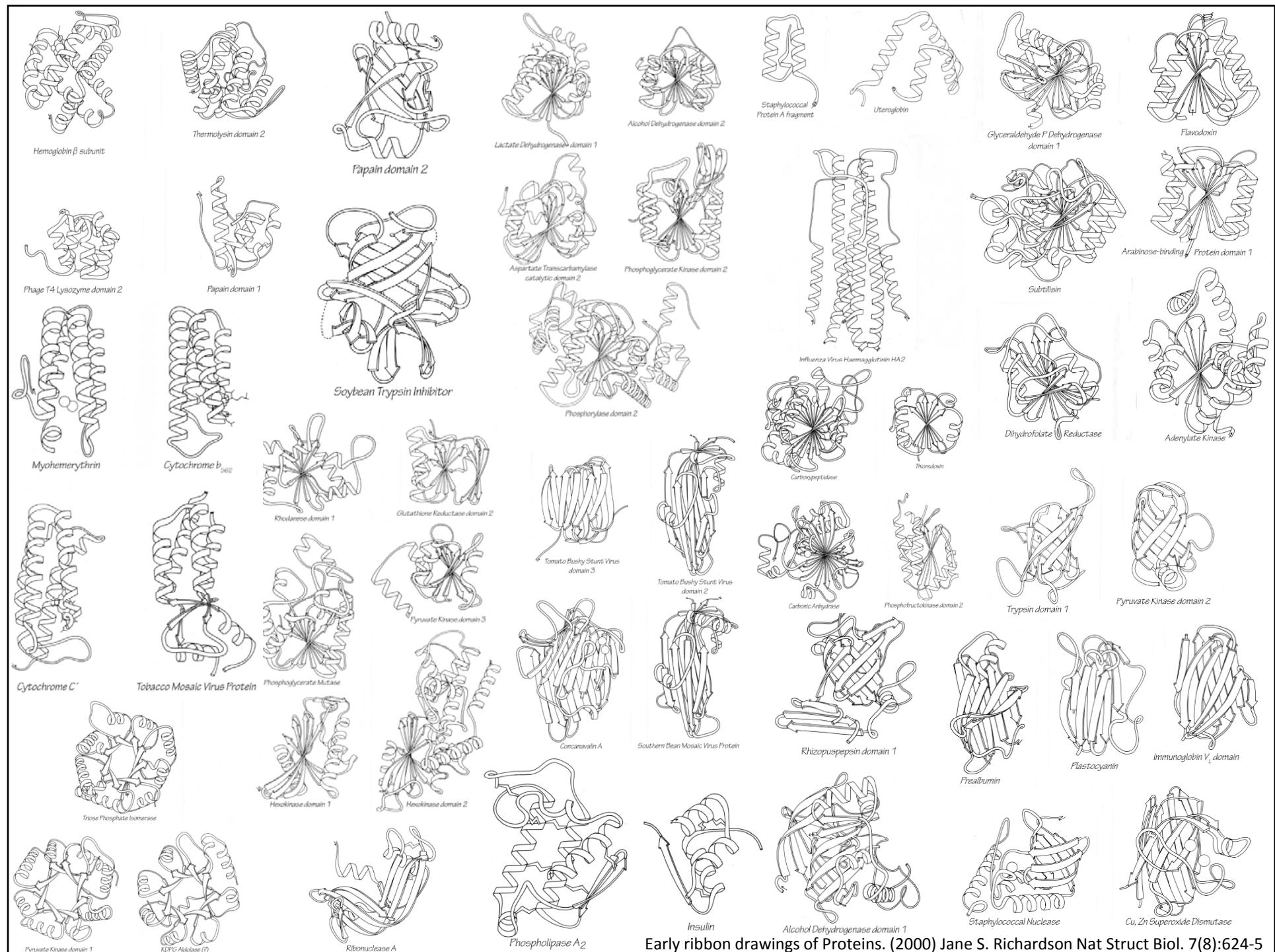
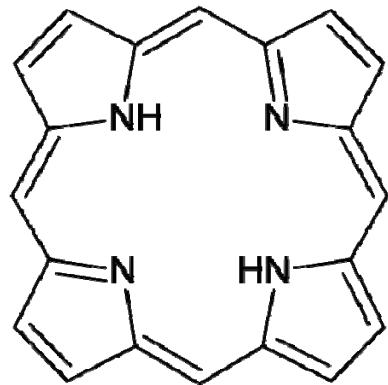


FIG. 7. A collection of  $\alpha$ -helices at varying angles to the page.



Early ribbon drawings of Proteins. (2000) Jane S. Richardson Nat Struct Biol. 7(8):624-5

# The heme group of heme proteins

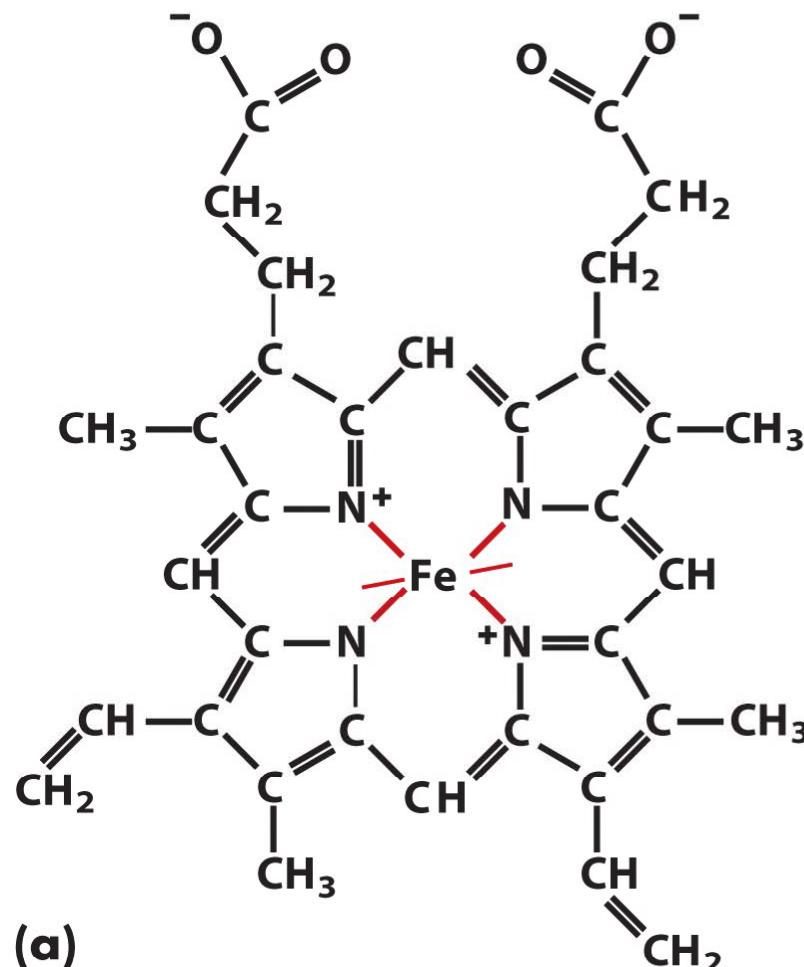


Simplest porphyrin

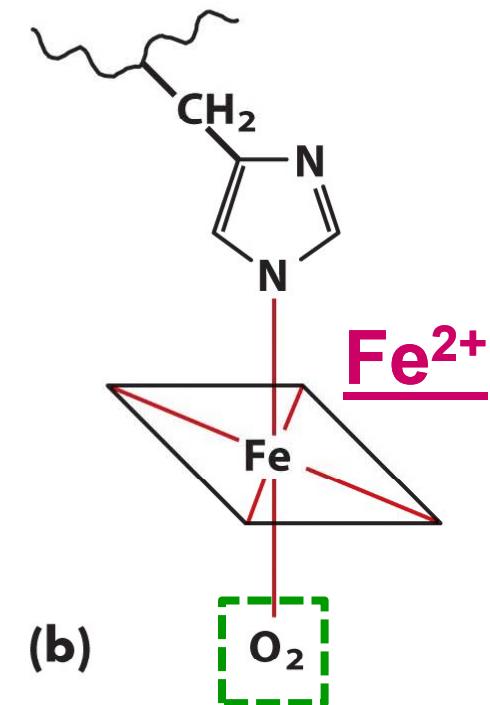
Porphyrin coordinated to iron: heme Fe

Porphyrin coordinated to magnesium: chlorophyll Mg

The one-carbon-shorter analogue corrole coordinated to a cobalt: vitamin B<sub>12</sub> Co



(a)

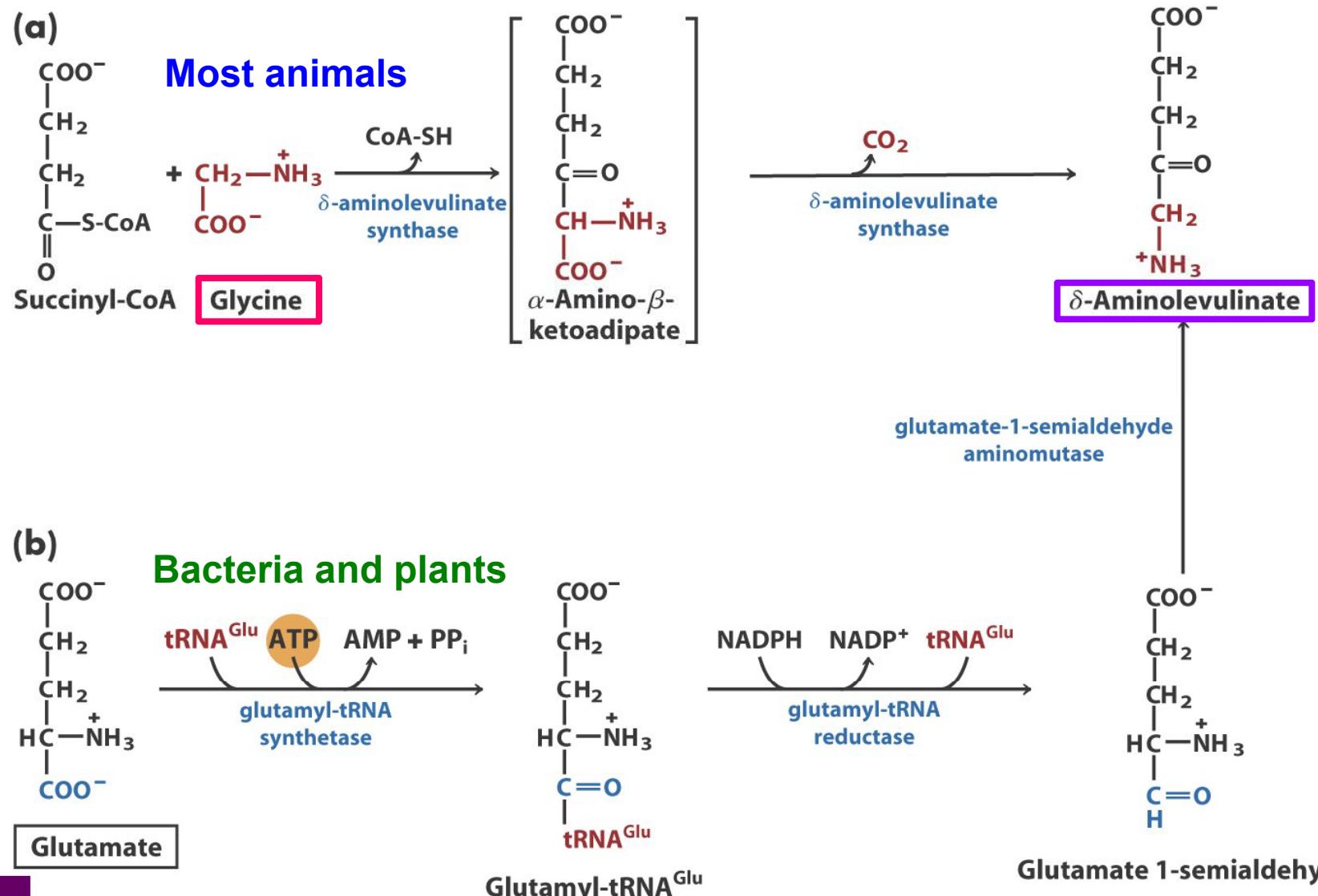


(b)

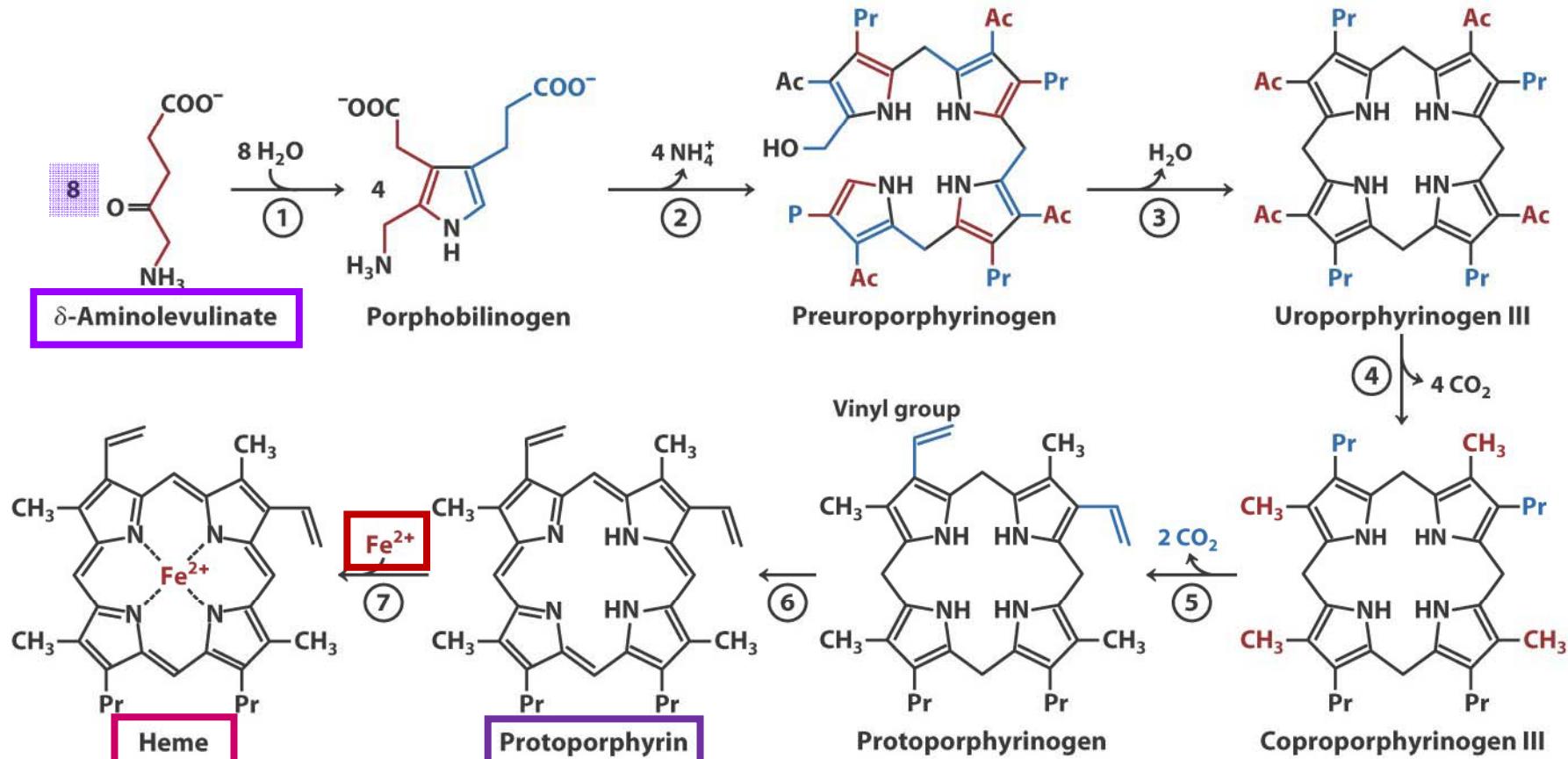
Myoglobin  
Hemoglobin  
Cytochrome

**Histidine**

# Specialized biomolecules derived from amino acids



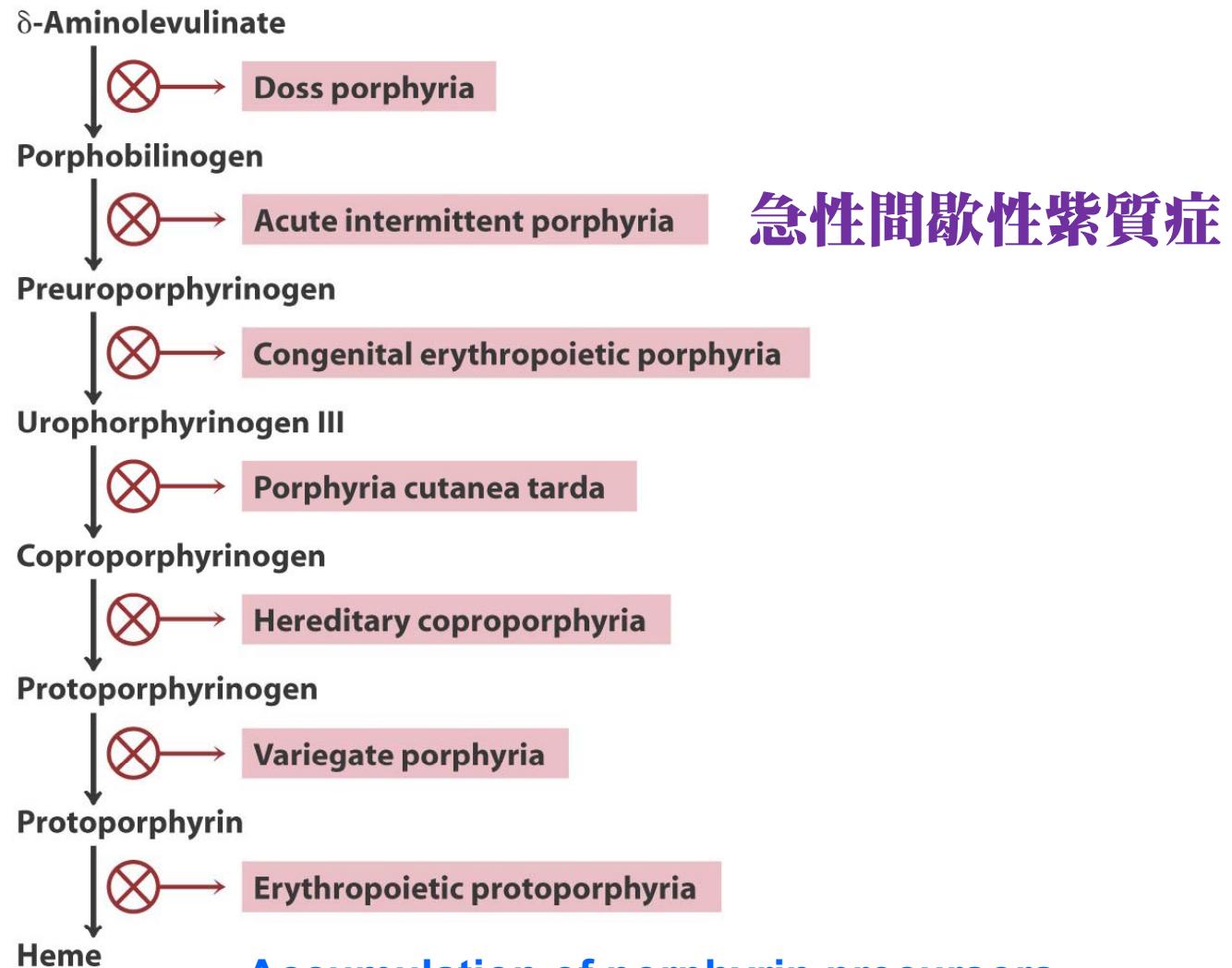
# Biosynthesis of heme from $\delta$ -aminolevulinate



- ① porphobilinogen synthase
- ② uroporphyrinogen synthase
- ③ uroporphyrinogen III cosynthase
- ④ uroporphyrinogen decarboxylase

- ⑤ coproporphyrinogen oxidase
- ⑥ protoporphyrinogen oxidase
- ⑦ ferrochelatase

# Porphyrias, defects in enzymes of the biosynthetic pathway from glycine to porphyrins



P. 875,  
BOX 22-2

Accumulation of porphyrin precursors  
Insufficient heme is synthesized

**TABLE 4-3** Approximate Amounts of  $\alpha$  Helix and  $\beta$  Conformation in Some Single-Chain Proteins

Protein (total residues)	Residues (%)*)	
	$\alpha$ Helix	$\beta$ Conformation
Chymotrypsin (247)	14	45
Ribonuclease (124)	26	35
Carboxypeptidase (307)	38	17
Cytochrome c (104)	39	0
Lysozyme (129)	40	12
Myoglobin (153)	78	0

Source: Data from Cantor, C.R. & Schimmel, P.R. (1980) *Biophysical Chemistry, Part I: The Conformation of Biological Macromolecules*, p. 100, W. H. Freeman and Company, New York.

\*Portions of the polypeptide chains that are not accounted for by  $\alpha$  helix or  $\beta$  conformation consist of bends and irregularly coiled or extended stretches. Segments of  $\alpha$  helix and  $\beta$  conformation sometimes deviate slightly from their normal dimensions and geometry.

# Supersecondary structures-Motif /Fold

A “fold” is a combination of secondary structures that form the core of a domain

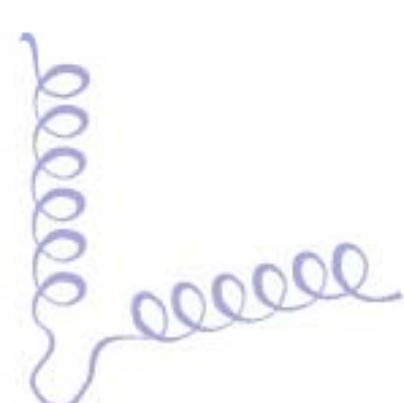
(a) **Helix-loop-helix** - two helices connected by a turn

(b) **Coiled-coil** - two amphipathic  $\alpha$  helices that interact in parallel through their hydrophobic edges

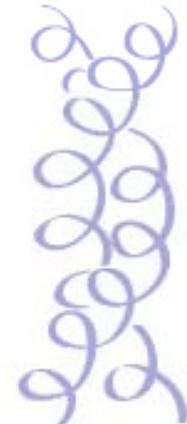
(c) **Helix bundle** - several  $\alpha$  helices that associate in an antiparallel manner to form a bundle

(d)  **$\beta\alpha\beta$  Unit** - two parallel  $\beta$  strands linked to an intervening  $\alpha$  helix by two loops

(a) Helix-loop-helix



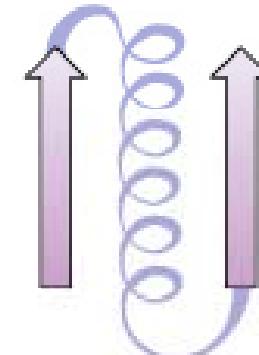
(b) Coiled coil



(c) Helix bundle

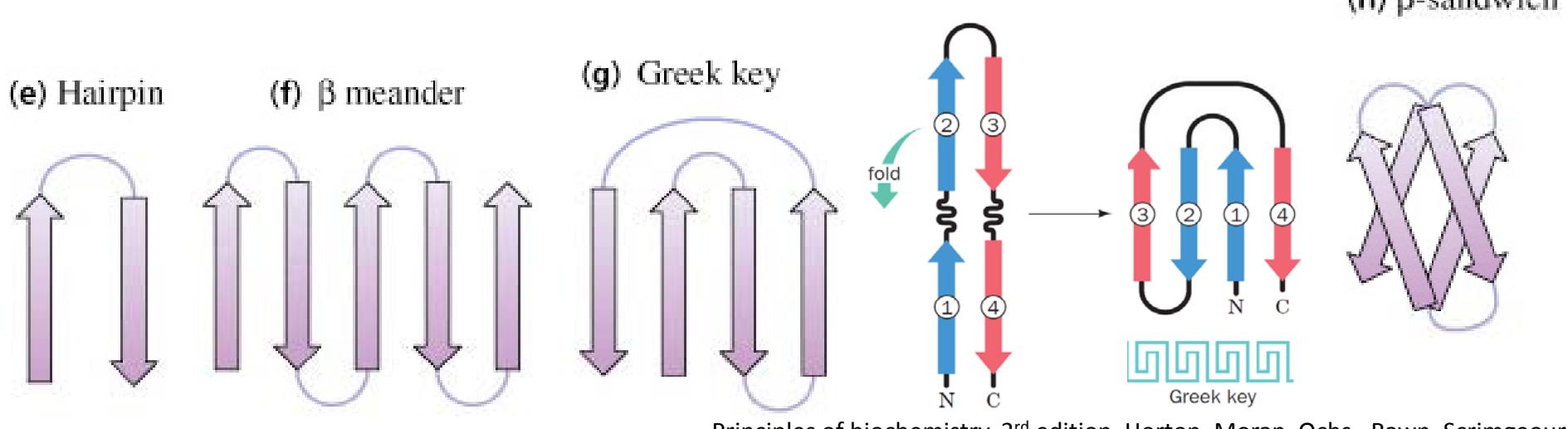


(d)  $\beta\alpha\beta$  unit



# Supersecondary structures-Motif /Fold

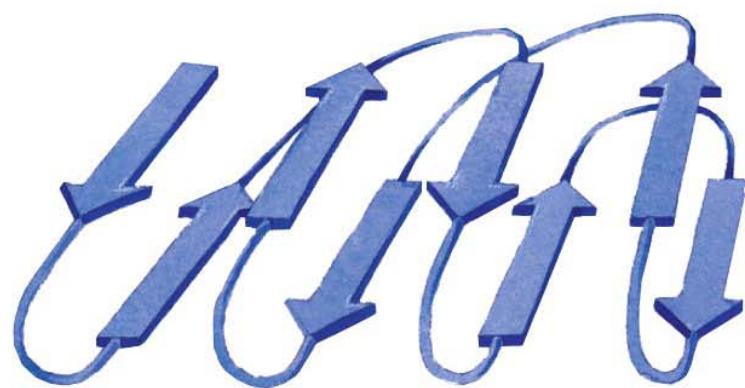
- (e) **Hairpin** - two adjacent antiparallel  $\beta$  strands connected by a  $\beta$  turn
- (f)  **$\beta$  Meander** - an antiparallel sheet composed of sequential  $\beta$  strands connected by loops or turns
- (g) **Greek key** - 4 antiparallel strands (strands 1,2 in the middle, 3 and 4 on the outer edges)
- (h)  **$\beta$  Sandwich** - stacked  $\beta$  strands or sheets



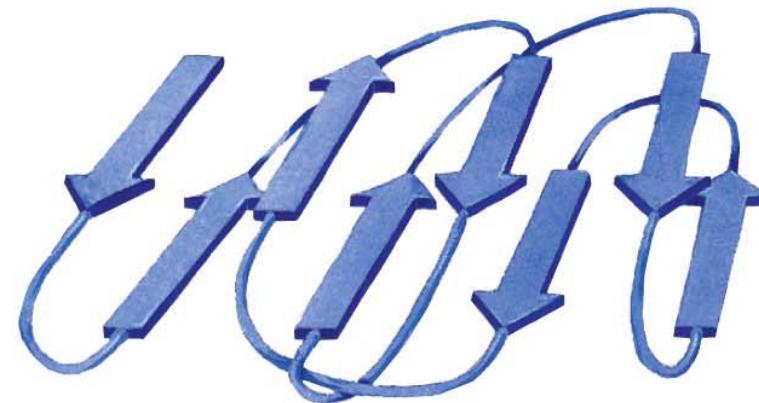
# Loops and turns

- **Loops and turns connect  $\alpha$  helices and  $\beta$  strands** and allow a peptide chain to fold back on itself to make a compact structure
- **Loops** - often contain hydrophilic residues and are found on protein surfaces
- **Turns** - loops containing 5 residues or less
- **$\beta$  Turns (reverse turns)** - connect different antiparallel  $\beta$  strands

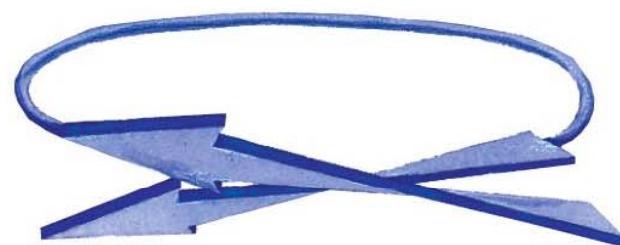
# $\beta$ strand folding patterns in proteins



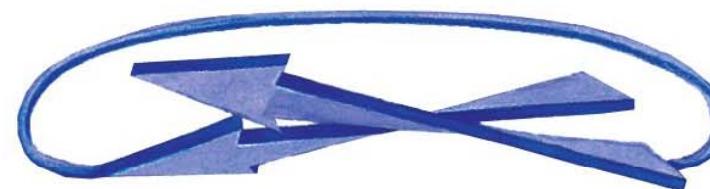
**Typical connections  
in an all- $\beta$  motif**



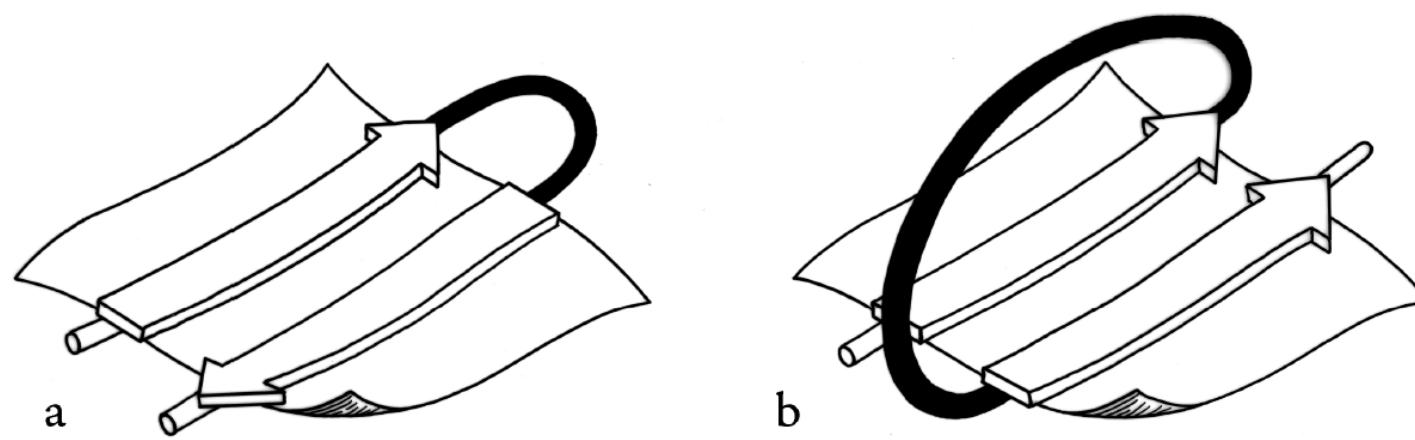
**Crossover connection  
(not observed)**



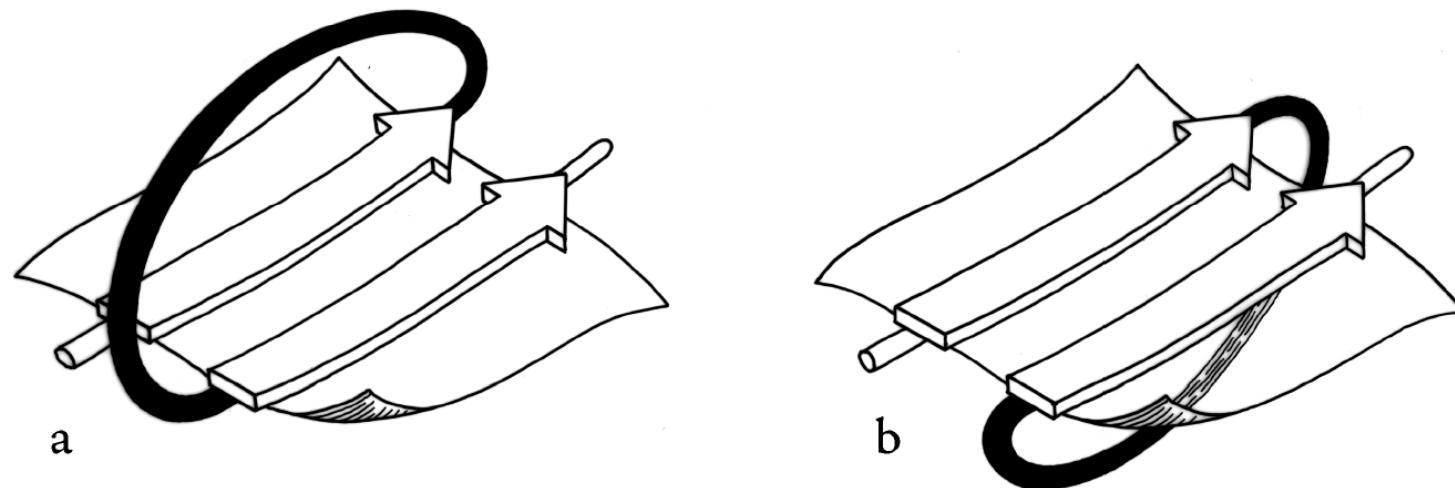
**Right-handed connection  
between  $\beta$  strands**



**Must traverse sharper angles**  
**Left-handed connection  
between  $\beta$  strands  
(very rare)**



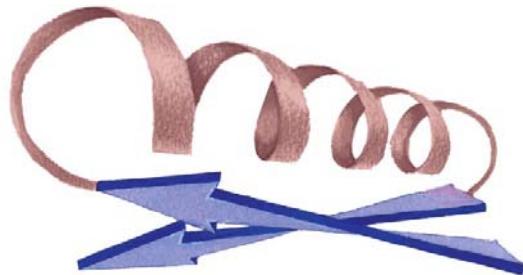
*FIG. 24. The two major sorts of connection between  $\beta$  strands: (a) a “hairpin,” or same-end, connection (this example is type +1, to a nearest-neighbor strand); (b) a “crossover,” or opposite-end, connection (this one is type +1x).*



*FIG. 26. (a) A right-handed +1x crossover connection; (b) a left-handed +1x crossover connection.*

## **subtilisin and glucose-phosphate isomerase**

# Constructing large motifs from small ones

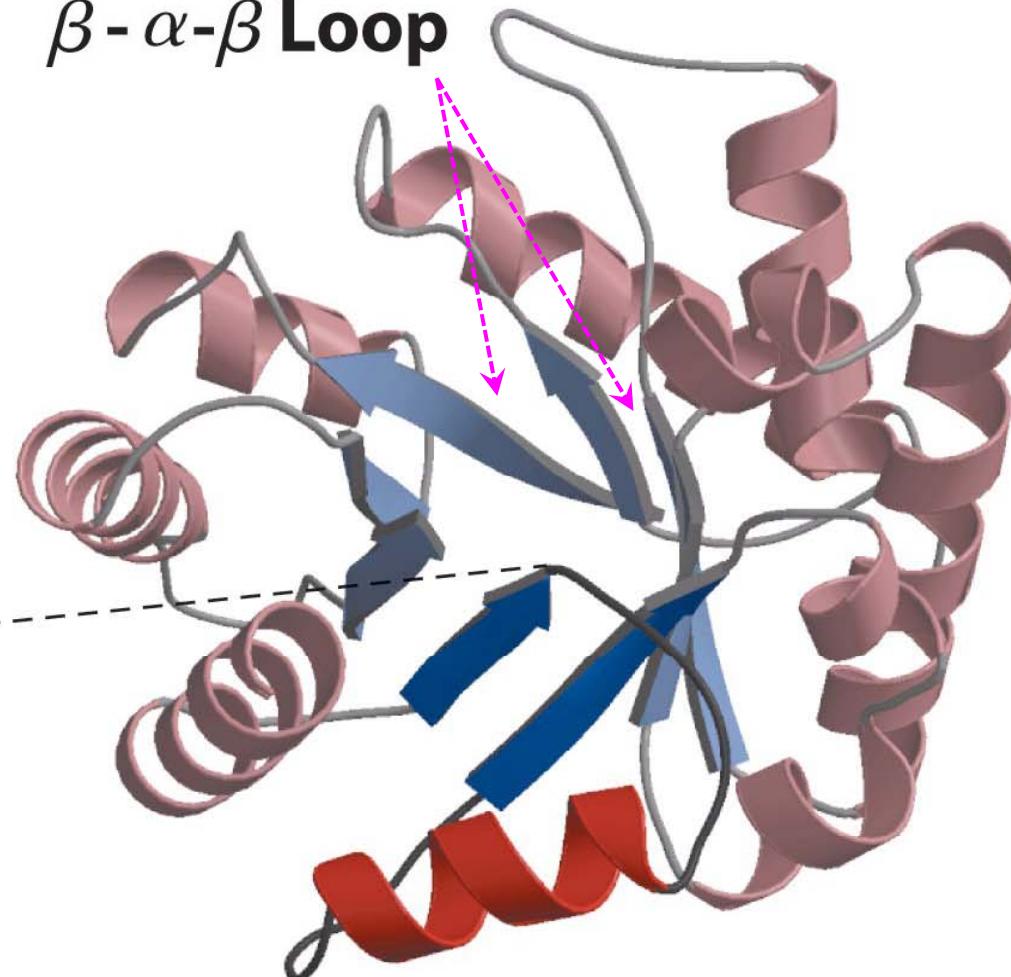


$\beta$ - $\alpha$ - $\beta$  Loop  
parallel



$\beta$ - $\alpha$ - $\beta$  Loop

$\beta$ - $\alpha$ - $\beta$  Loop



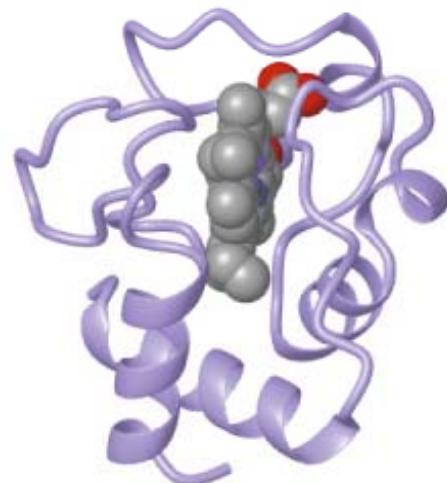
$\alpha/\beta$  Barrel

# Domain Structure and Function

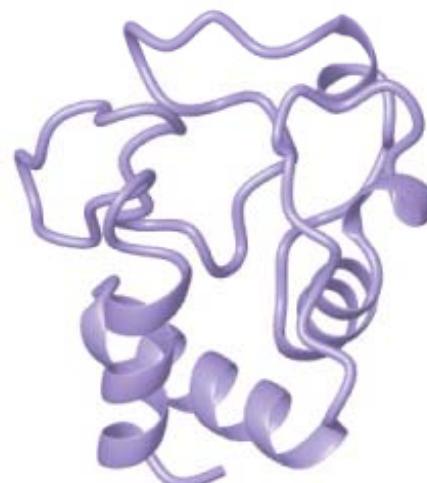
- **Independently folded**, compact units in proteins
- **Domain size: ~25 to ~300 amino acid residues**
- Domains are **connected to each other by loops, bound by weak interactions between side chains**
- **A domain may still retain its three-dimensional structure even when separated**
- Domains illustrate the **evolutionary conservation** of protein structure
- **A single domain** may have a **particular function** (e.g. binding small molecules, catalyzing a single reaction)
- **Interfaces between two separate domains** provide crevices, grooves, and pockets for binding or catalytic sites
- In **multifunctional enzymes**, each catalytic activity can be on one of several domains

# Conservation of cytochrome c structure

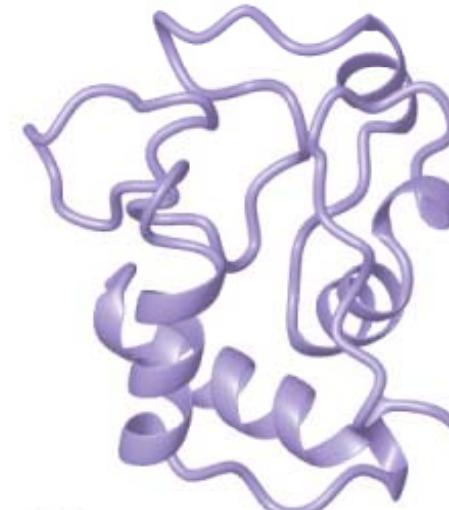
(a)



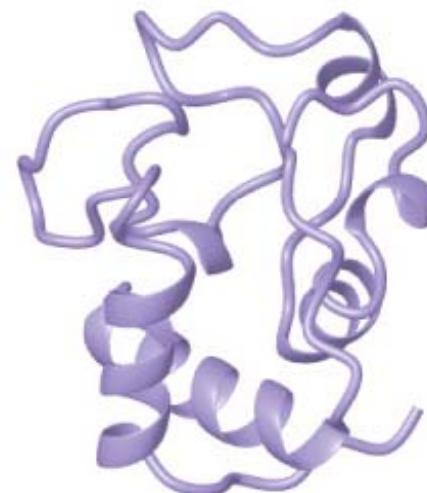
(b)



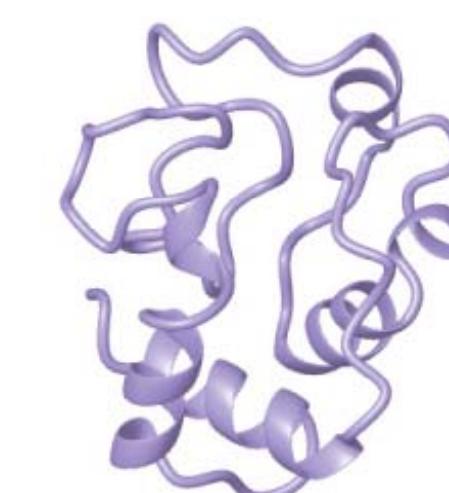
(c)



(d)



(e)

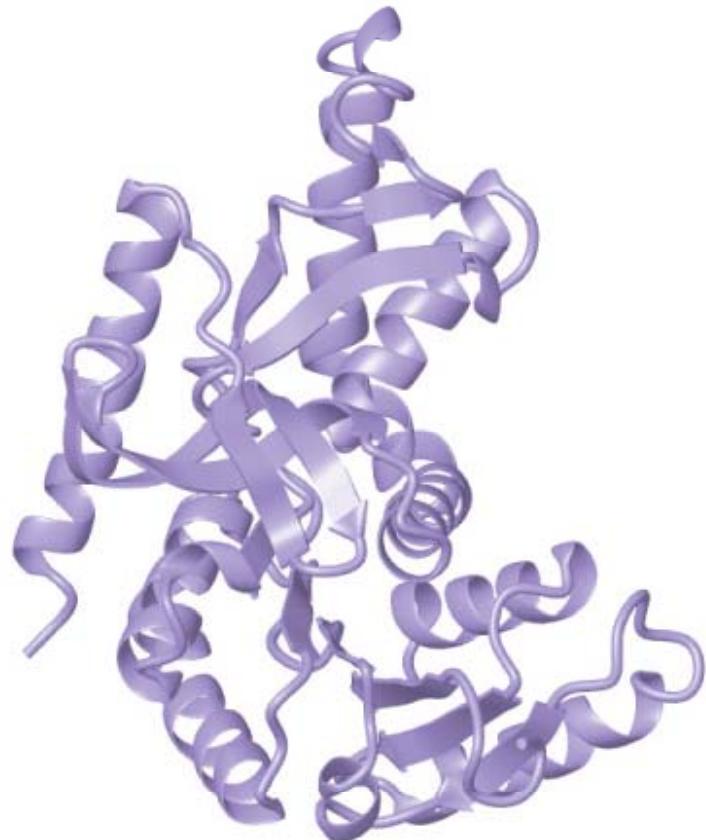


- (a) Tuna (+heme)
- (b) Tuna
- (c) Rice
- (d) Yeast
- (e) Bacteria

## Structural similarity of (a) lactate dehydrogenase, (b) malate dehydrogenase

*B. stereothermophilus*

(a)



*E. coli*

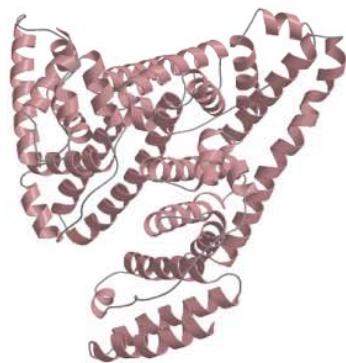
(b)



# Four categories of protein domains

- (1) All  $\alpha$  - domains consist almost entirely of  $\alpha$  helices and loops
- (2) All  $\beta$  - all domains contain only  $\beta$  sheets and non-repetitive structures that link the  $\beta$  strands
- (3) Mixed  $\alpha/\beta$  - contain supersecondary structures such as the  $\alpha\beta\alpha$  motif, where regions of  $\alpha$  helix and  $\beta$  strand alternate
- (4)  $\alpha + \beta$  - domains consist of local clusters of  $\alpha$  helices and  $\beta$  sheets in separate, contiguous regions of the polypeptide chain

## All $\alpha$



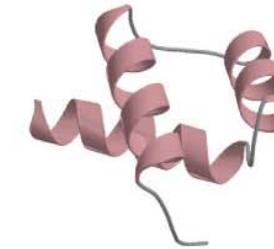
1AO6  
Serum albumin  
Serum albumin  
Serum albumin  
Serum albumin  
Human (*Homo sapiens*)



1BCF  
Ferritin-like  
Ferritin-like  
Ferritin  
Bacterioferritin (cytochrome  $b_1$ )  
*Escherichia coli*



1GAI  
 $\alpha/\alpha$  toroid  
Glycosyltransferases of the superhelical fold  
Glucoamylase  
Glucoamylase  
*Aspergillus awamori*, variant x100

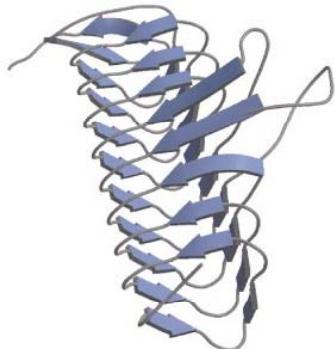


1ENH  
DNA-binding 3-helical bundle  
Homeodomain-like  
Homeodomain  
engrailed Homeodomain  
*Drosophila melanogaster*

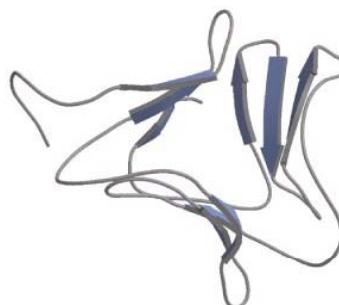
## All $\beta$



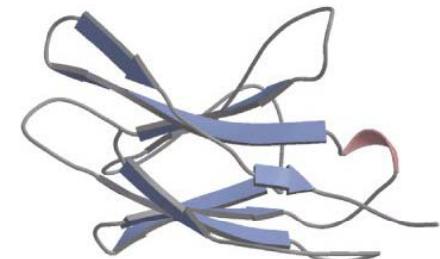
1HOE  
 $\alpha$ -Amylase inhibitor  
 $\alpha$ -Amylase inhibitor  
 $\alpha$ -Amylase inhibitor  
HOE-467A  
*Streptomyces tendae* 4158



1LXA  
Single-stranded left-handed  $\beta$  helix  
Trimeric LpxA-like enzymes  
UDP N-acetylglucosamine acyltransferase  
UDP N-acetylglucosamine acyltransferase  
*Escherichia coli*

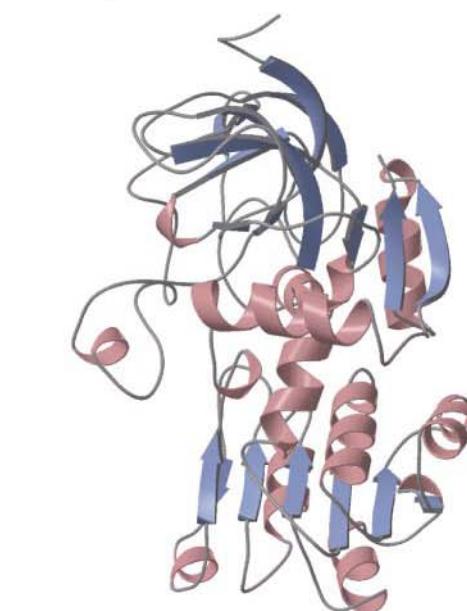


1JPC  
 $\beta$ -Prism II  
 $\alpha$ -D-Mannose-specific plant lectins  
 $\alpha$ -D-Mannose-specific plant lectins  
Lectin (agglutinin)  
Snowdrop (*Galanthus nivalis*)

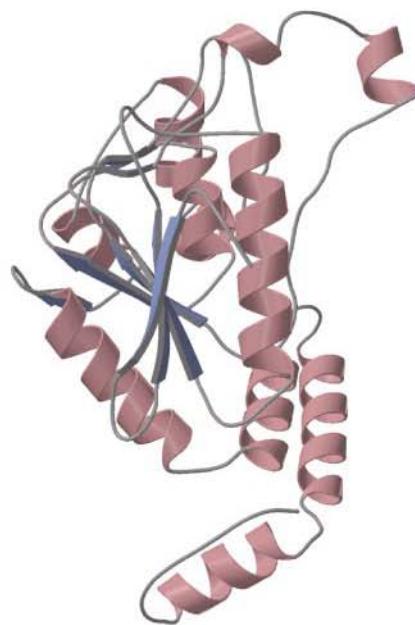


1CD8  
Immunoglobulin-like  $\beta$  sandwich  
Immunoglobulin  
Antibody variable domain-like  
CD8  
Human (*Homo sapiens*)

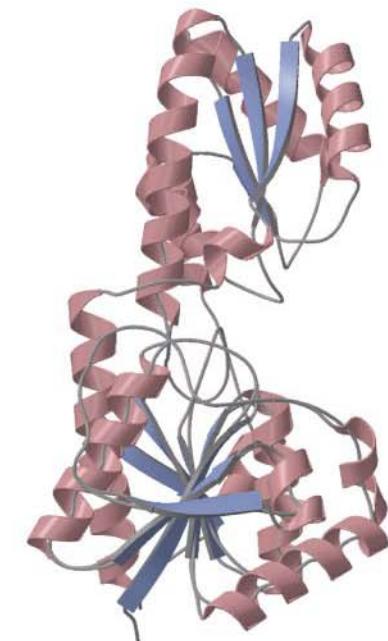
$\alpha/\beta$



1DEH  
NAD(P)-binding Rossmann-fold domains  
NAD(P)-binding Rossmann-fold domains  
Alcohol/glucose dehydrogenases,  
carboxyl-terminal domain  
Alcohol dehydrogenase  
Human (*Homo sapiens*)



1DUB  
Crotonase-like  
Crotonase-like  
Crotonase-like  
Enoyl-CoA hydratase  
Rat (*Rattus norvegicus*)



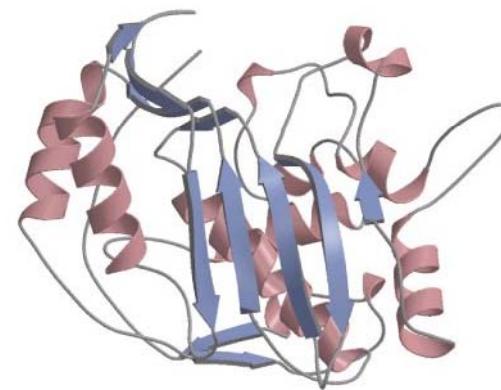
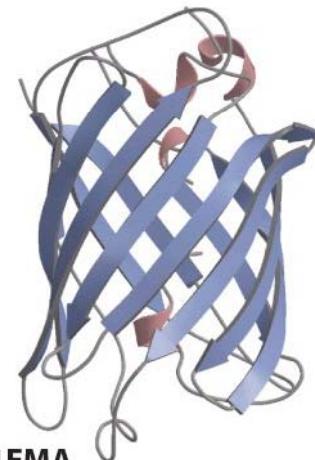
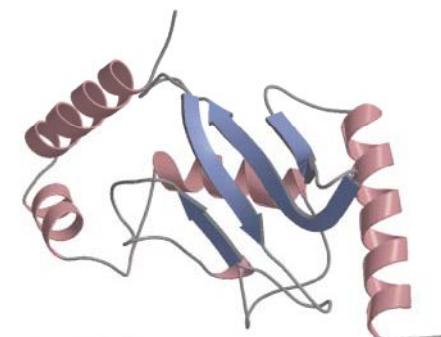
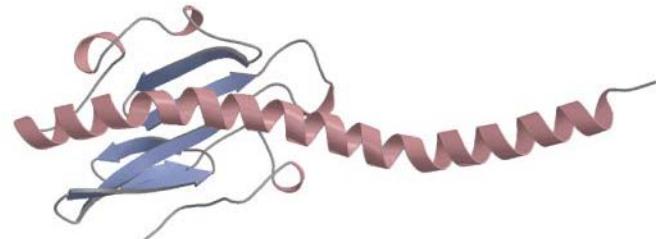
1PFK  
Phosphofructokinase  
Phosphofructokinase  
Phosphofructokinase  
Phosphofructokinase  
*Escherichia coli*

## Protein Data Bank (PDB)

The PDB is an archive of experimentally determined three-dimensional structures of biological macromolecules. Each structure is assigned a unique four-character identifier called the PDB ID. The first character must be a digit (1~9) and the PDB ID does not necessarily have a relationship to the macromolecule's name.

## $\alpha$ and $\beta$ 分開於不同區域

$\alpha + \beta$



$\beta$  barrel

PDB identifier  
Fold  
Superfamily  
Family  
Protein  
Species

Contact Us | Print

PDB ID or Text

Search Advanced Search

MyPDB

Login to your Account  
Register a New Account

Home

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

Deposition

All Deposit Services  
Electron Microscopy  
X-ray | NMR  
Validation Server  
BioSync Beamline  
Related Tools

Search

Advanced Search  
Latest Release  
New Structure Papers  
Sequence Search  
Chemical Components  
Unreleased Entries  
Browne Database  
Histograms

Tools

Contact Us | Print

PDB ID or Text Advanced Search

MyPDB

Login to your Account  
Register a New Account

Home

News & Publications  
Usage/Reference Policies

Deposition Policies  
Website FAQ

Deposition FAQ  
Contact Us  
About Us  
Careers  
External Links  
Sitemap  
New Website Features

Deposition

All Deposit Services  
Electron Microscopy  
X-ray | NMR  
Validation Server  
BioSync Beamline  
Related Tools

Search

Advanced Search  
Latest Release  
New Structure Papers  
Sequence Search  
Chemical Components  
Unreleased Entries  
Browne Database  
Histograms

## A Resource for Studying Biological Macromolecules

The PDB archive contains information about experimentally-determined structures of proteins, nucleic acids, and complex assemblies. As a member of the [wwPDB](#), the RCSB PDB curates and annotates PDB data according to agreed upon standards.

The RCSB PDB also provides a variety of tools and resources. Users can perform simple and advanced searches based on annotations relating to sequence, structure and function. These molecules are visualized, downloaded, and analyzed by users who range from students to specialized scientists.

[Hide Welcome Message](#)

Featured Molecules ([MotM Category View](#) / [Previous Features: MotM](#) | [PSI](#))

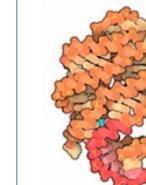
Structural View of Biology



Protein Synthesis



Molecule of the Month:  
**Riboswitches**



Why use two or more molecules when one will do? In our own cells, protein synthesis is controlled by thousands of regulatory proteins, which work together to decide when a particular protein will be made. Bacteria are masters of economy, however, and in some cases, they have figured out a way for messenger RNA to control itself, without the need for help by proteins.

[Full Article...](#)



Protein Structure Initiative Featured Molecule:  
**Alpha and beta barrels**

Analysis of the metabolic network of the bacterium *Thermatoga maritima* reveals the evolution of enzyme folding and function.

[Full Article](#) | [PSI Structural Biology Knowledgebase](#)

New Features

**PDBMobile: Search PDB & Browse Results**

Latest features released:

Website Release Archive:

RCSB PDB News

[Weekly](#) | [Quarterly](#) | [Yearly](#)

2010-10-12

Categorized View of Molecule of the Month



Explore new Molecule of the Month feature categories.

- Latest Website Release
- Poster Prize Awarded at ECM
- Analyze small molecule interactions in the PDB with Ligand Explorer
- Redesigned BioSync

wwPDB News

Statement on Retraction of PDB Entries

2010-10-08

Contact Us | Print

PDB ID or Text Advanced Search

Summary

Sequence

Derived Data

Seq. Similarity

3D Similarity

Literature

Bio & Chem.

Methods

Geometry

Links

Primary Citation

Structure and refinement of oxymyoglobin at 1.6 angstroms resolution

DOI:10.2210/pdb1mbo/pdb

Structure and refinement of oxymyoglobin at 1.6 A resolution.

Phillips, S.E.<sup>1</sup>

Journal: (1980) J.Mol.Biol. 142: 531-554

PubMed: 7463482

[Search Related Articles in PubMed](#)

PubMed Abstract:

No abstract available... [\[Read More & Search PubMed Abstracts\]](#)

1 Molecular Description

Classification: Oxygen Storage

Structure Weight: 17979.65

Molecule: MYOGLOBIN

Polymer: 1 Type: polypeptide(L)

Chains: A

Length: 153

Source

Polymer: 1

Scientific Name: Phycerote catodon

Taxonomy

Common Name: Sperm whale

1MBO

Display Files Download Files Print this Page Share this Page



More Images...

Biological assembly assigned by authors

[Deposition Summary](#)

THE CRYSTAL STRUCTURE OF HUMAN DEOXYHAEMOGLOBIN AT 1.74 ANGSTROMS RESOLUTION

DOI:10.2210/pdb2hhb/pdb

ENTRY ZHHB SUPERSEDES 1HHB

Primary Citation

The crystal structure of human deoxyhaemoglobin at 1.74 A resolution.

Fermi, G., Penz, M.F., Shanahan, B.J., Fourme, R.

Journal: (1981) J.Mol.Biol. 175: 159-174

PubMed: 7028803

[Search Related Articles in PubMed](#)

PubMed Abstract:

The structure of human deoxyhaemoglobin was refined at 1.74 A resolution using data collected on film at room temperature from a synchrotron X-ray source. The crystallographic R-factor is 16.0%. The estimated error in atomic positions is 0.1 A overall, 0.14...

[\[Read More & Search PubMed Abstracts\]](#)

# Related Citations in PDB Entry (REMARK 1)

Stereocchemistry of Iron in Deoxyhaemoglobin

Fermi, G., Penz, M.F., Shanahan, B.J., Duda, P.J., Seissler, J.L., Hahn, J.E.

(1982) *Nature* **295**: 525

Fermi, G., Penz, M.F.

(1981) *Hemoglobin and Myoglobin, Atlas of Molecular Structures in Biology* 2:

ZHHB

Display Files Download Files Print this Page Share this Page

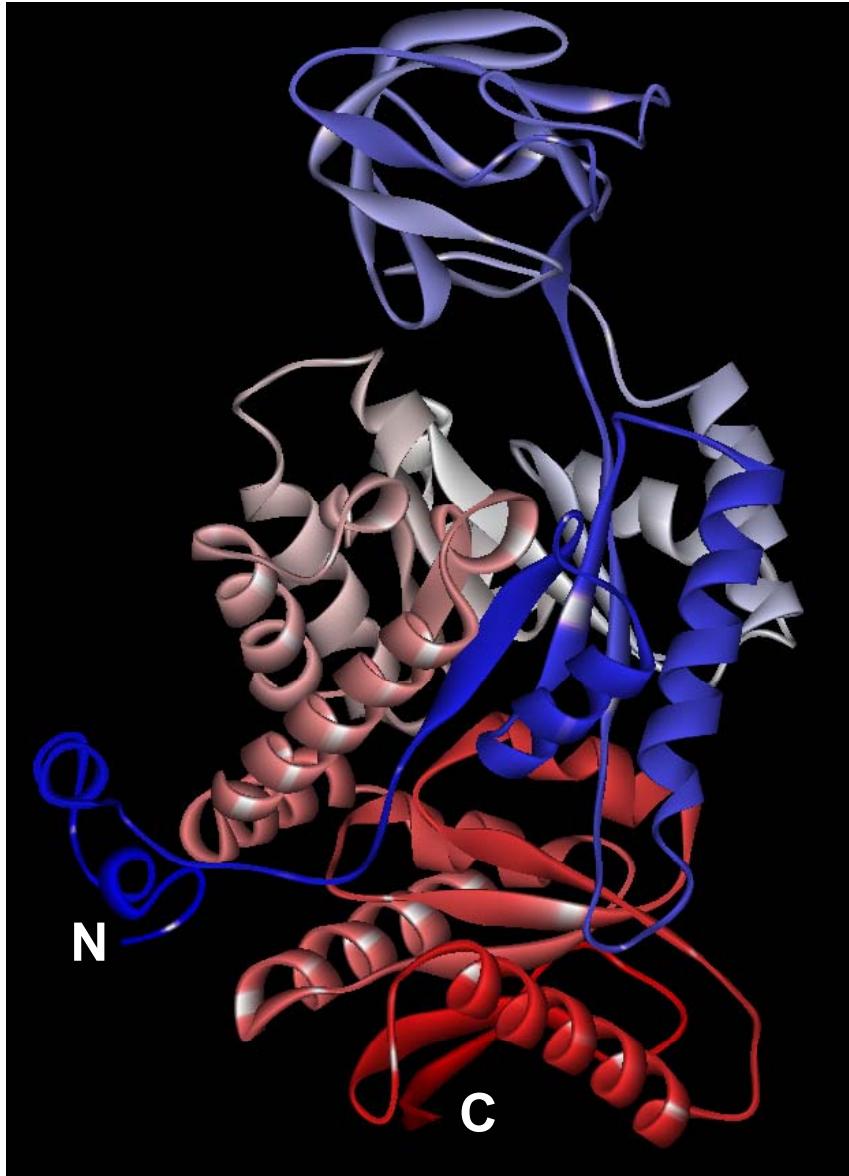


More Images...

Biological assembly assigned by authors and generated by PISA (software)

[Deposition Summary](#)

# Main polypeptide chain of pyruvate kinase folds into three distinct domains

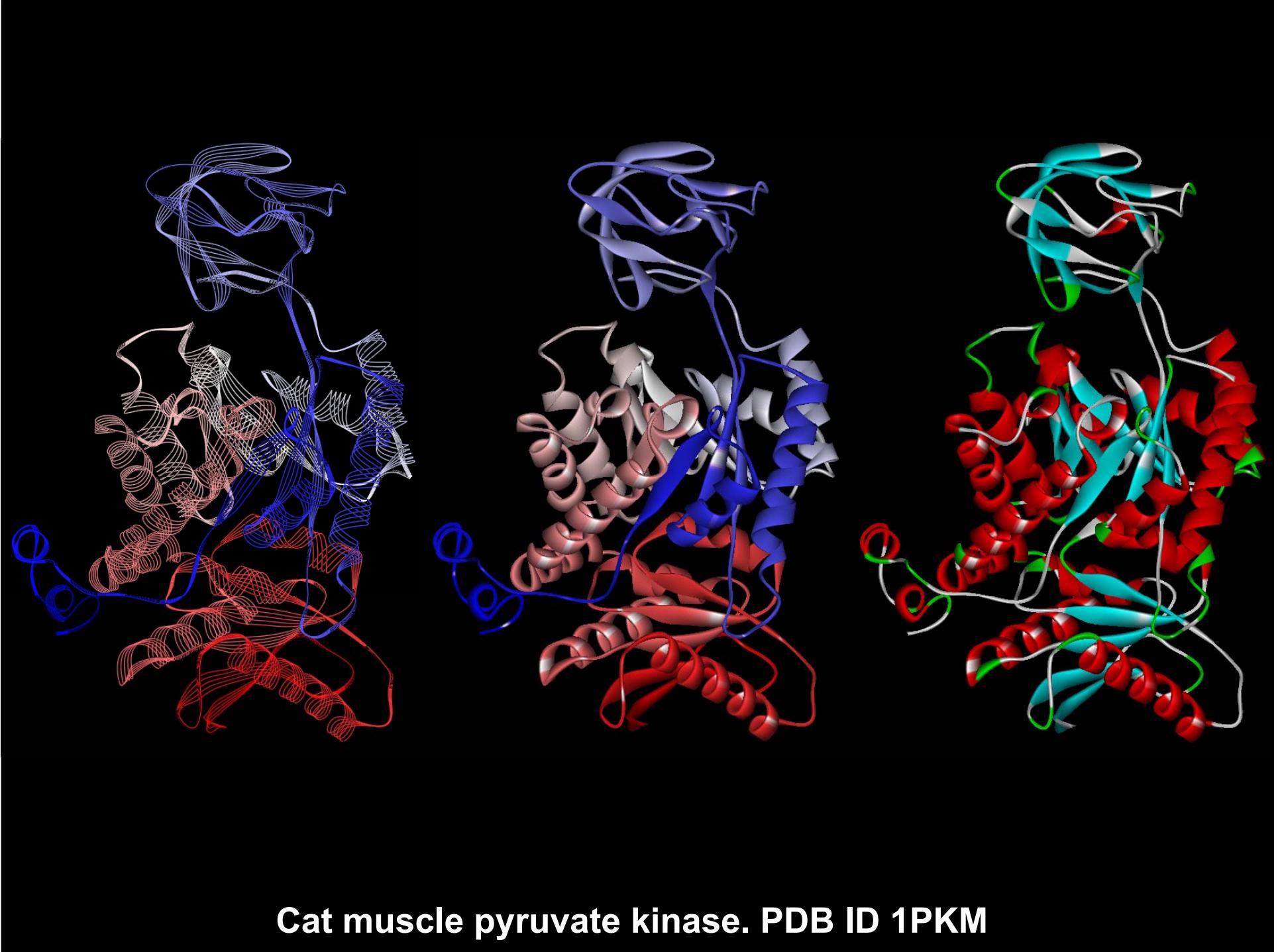


**Top:** a b-barrel domain, residues 116 to 219.

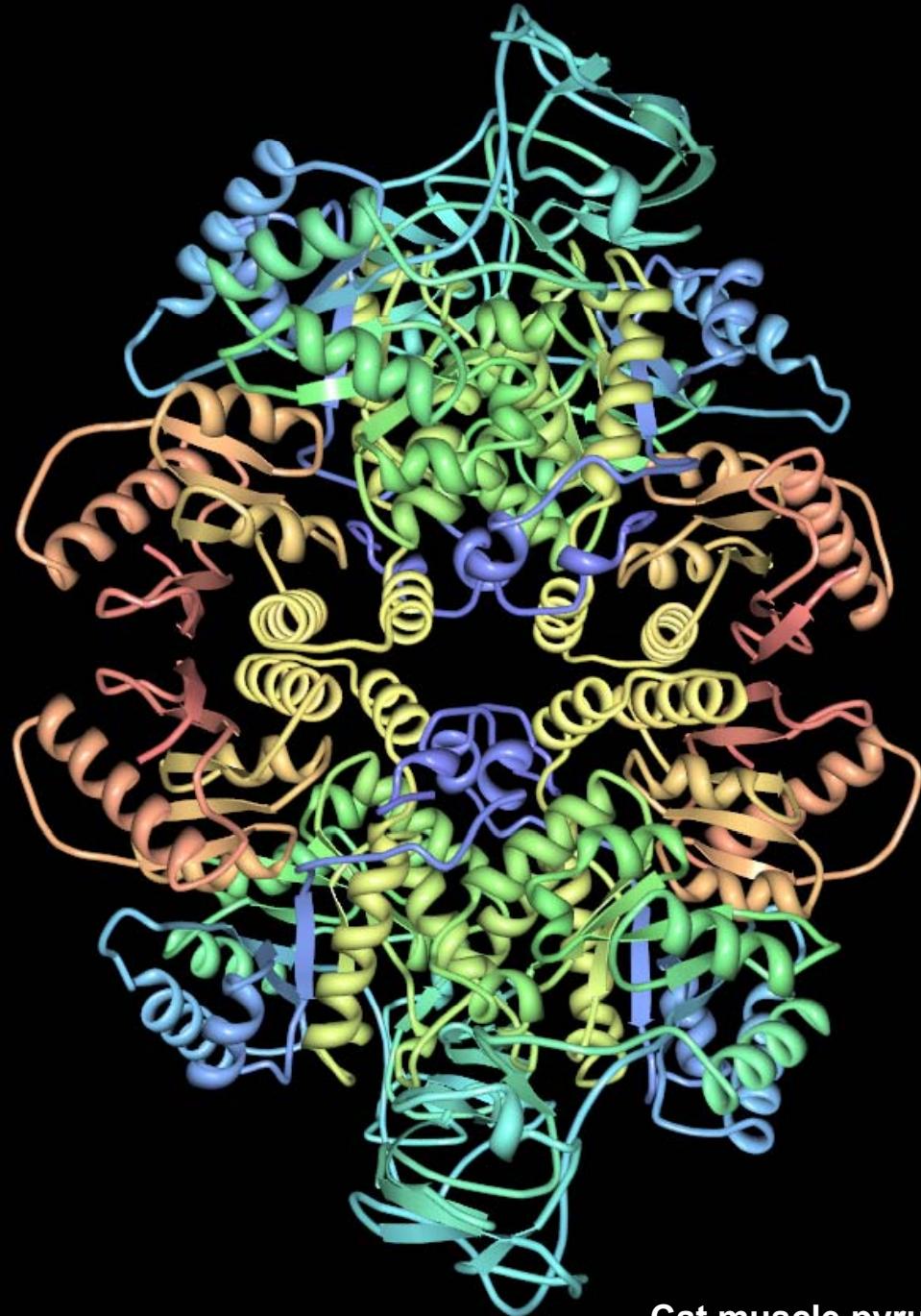
**Middle:** comprises 1 to 115 and 220 to 388.

The construction of a domain from two separate parts of the chain, rather than a contiguous sequence, is unusual, and occurs in about 25% of multidomain proteins.

**Bottom:** residues 389 to 530.

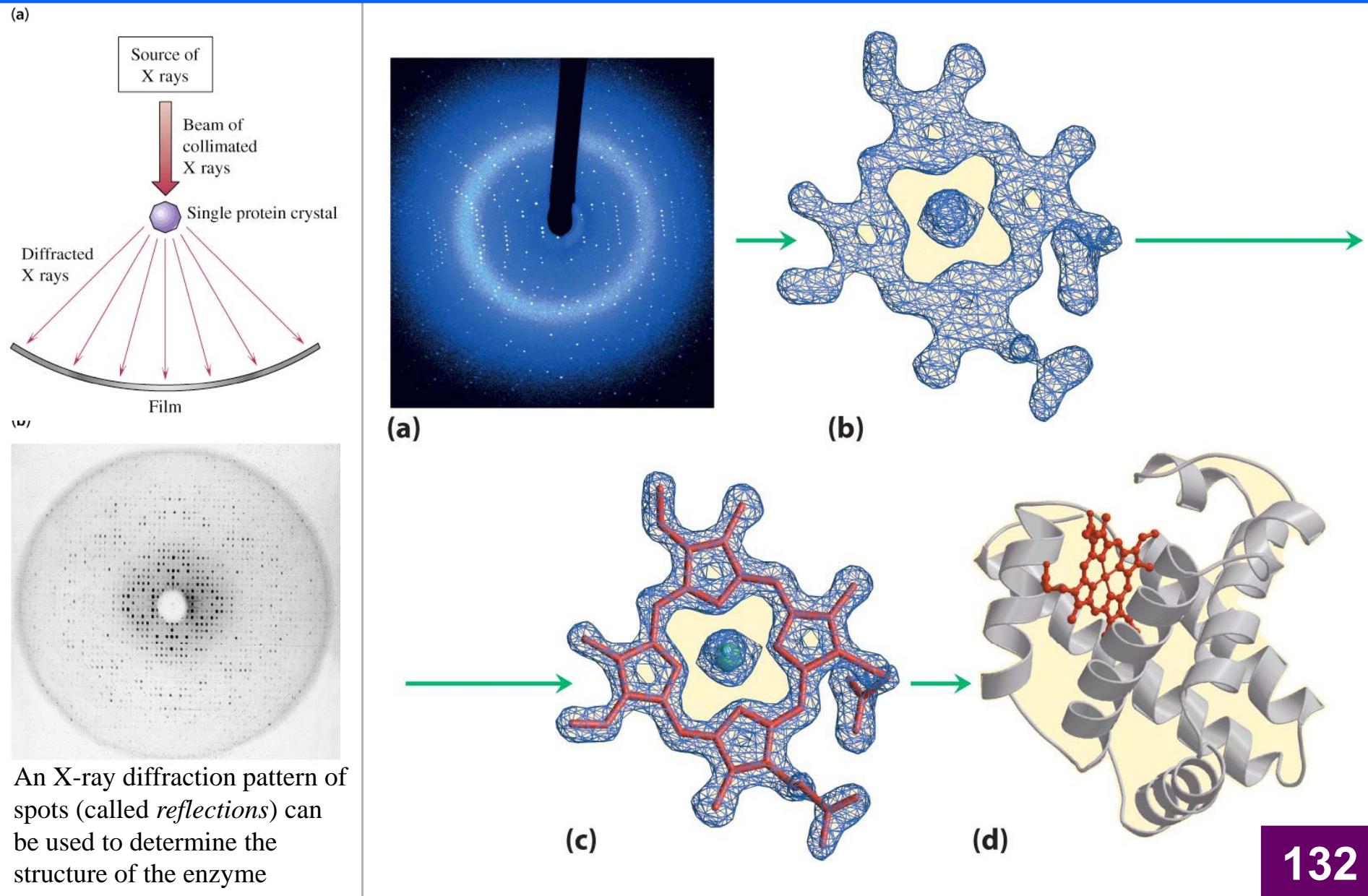


Cat muscle pyruvate kinase. PDB ID 1PKM



Cat muscle pyruvate kinase. PDB ID 1PKM

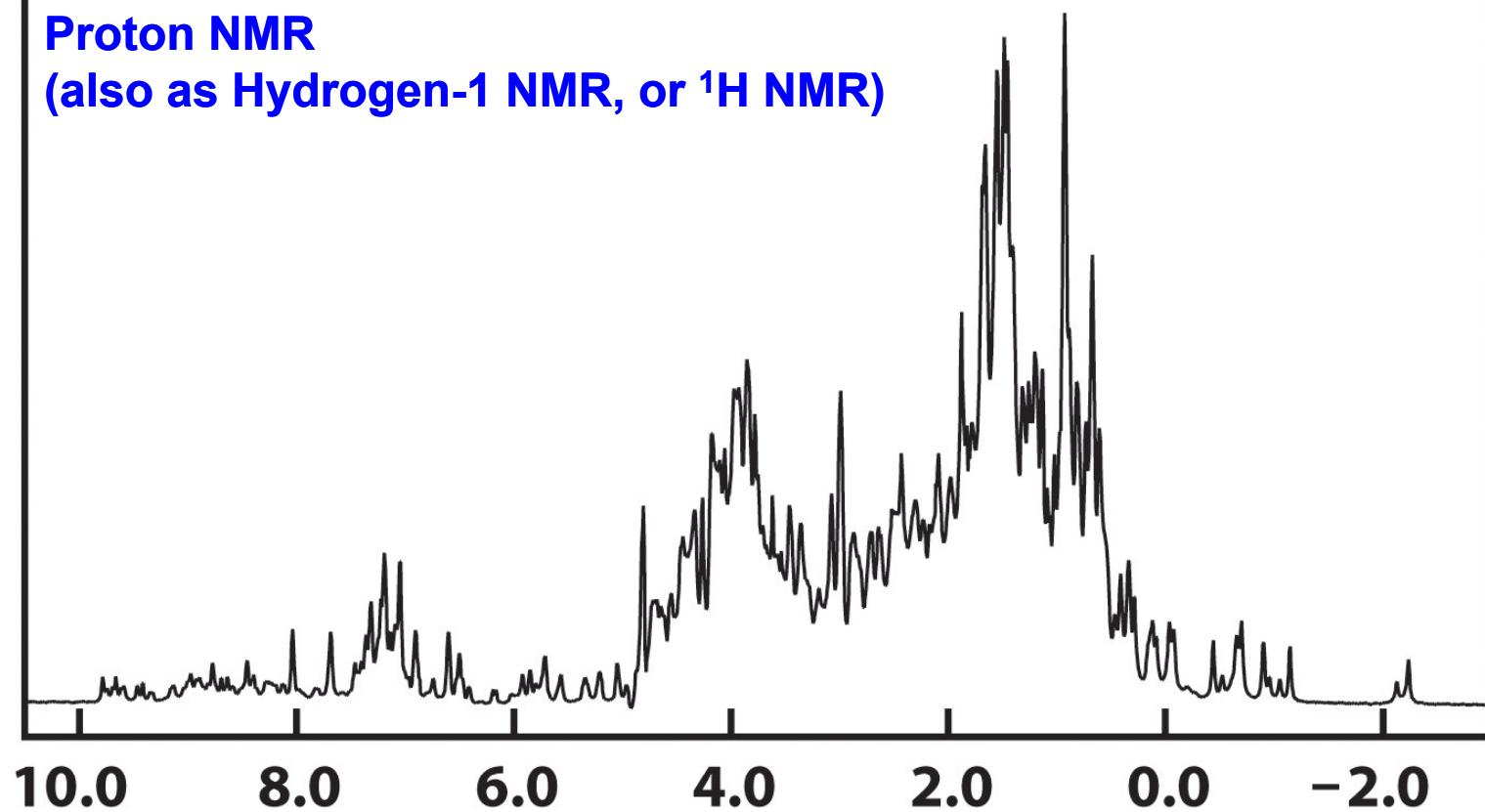
# Determining the three-dimensional structure of a protein by X-ray diffraction



# One-dimensional NMR spectrum

A globin from a marine blood worm

Proton NMR  
(also as Hydrogen-1 NMR, or  $^1\text{H}$  NMR)

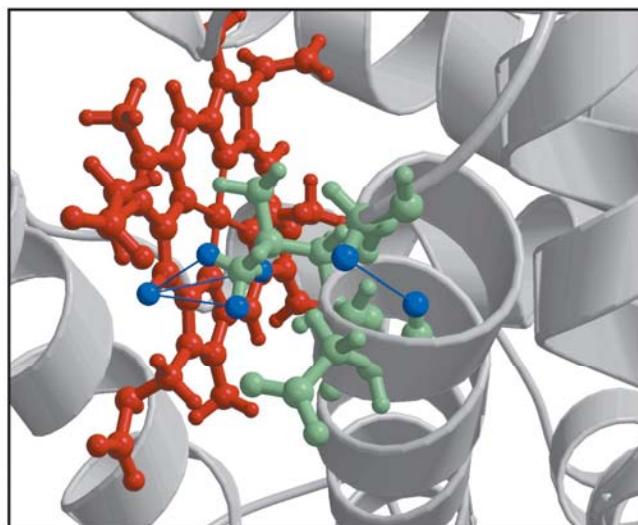
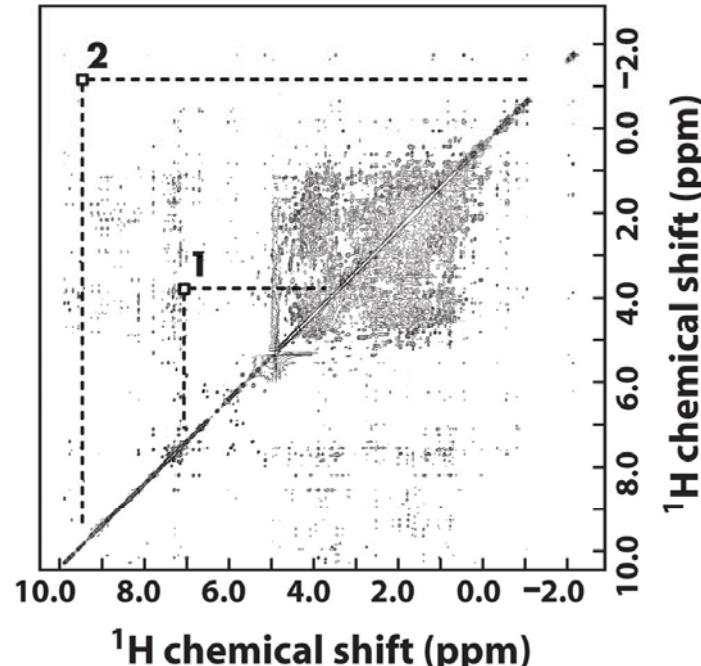


$^1\text{H}$  chemical shift (ppm)

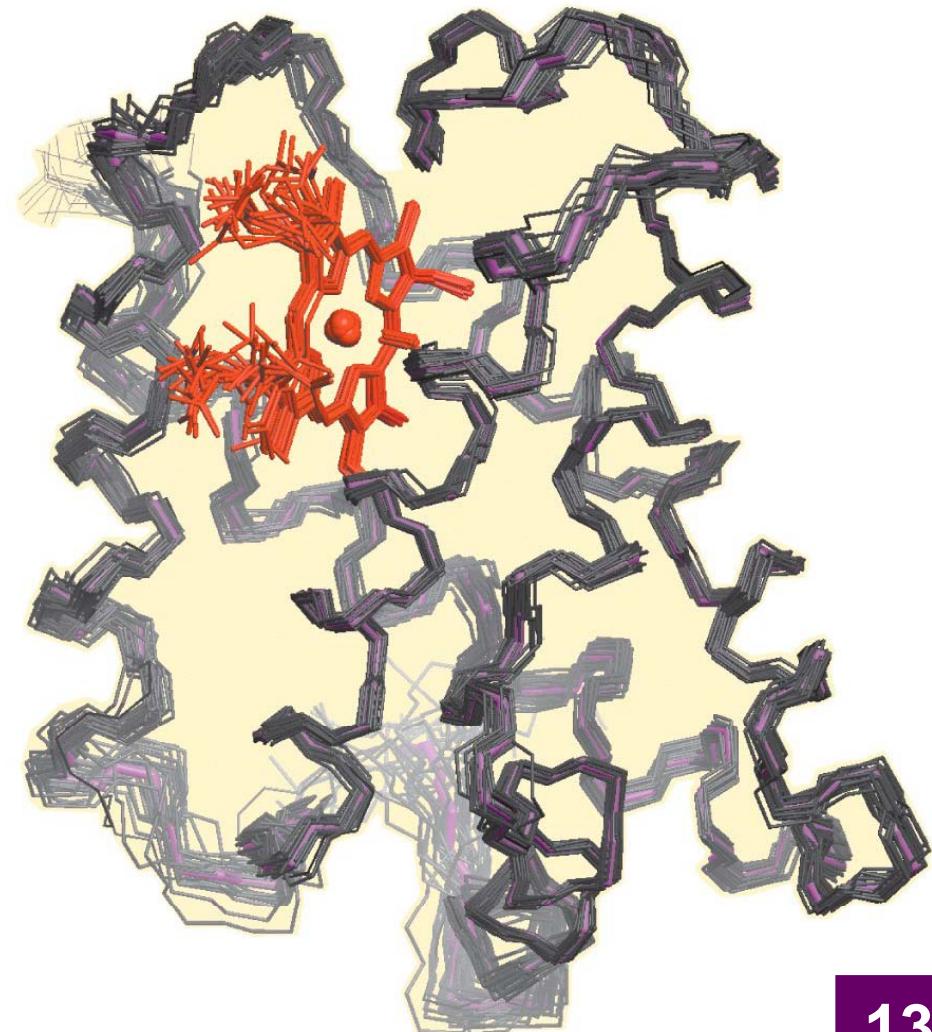
133

The most commonly studied nuclei are  $^1\text{H}$  (the most NMR-sensitive isotope after the radioactive  $^3\text{H}$ ) and  $^{13}\text{C}$ , although nuclei from isotopes of many other elements (e.g.  $^2\text{H}$ ,  $^{10}\text{B}$ ,  $^{11}\text{B}$ ,  $^{14}\text{N}$ ,  $^{15}\text{N}$ ,  $^{17}\text{O}$ ,  $^{19}\text{F}$ ,  $^{23}\text{Na}$ ,  $^{29}\text{Si}$ ,  $^{31}\text{P}$ ,  $^{35}\text{Cl}$ ,  $^{113}\text{Cd}$ ,  $^{129}\text{Xe}$ ,  $^{195}\text{Pt}$ ) are studied by high-field NMR spectroscopy as well.

# Two-dimensional NMR spectrum



Three-dimensional structure





John B. Fenn



Koichi Tanaka



Kurt Wüthrich

The Nobel Prize in Chemistry 2002 was awarded "for the development of methods for identification and structure analyses of biological macromolecules" with one half jointly to John B. Fenn and Koichi Tanaka "for their development of soft desorption ionisation methods for mass spectrometric analyses of biological macromolecules" and the other half to Kurt Wüthrich "for his development of nuclear magnetic resonance spectroscopy for determining the three-dimensional structure of biological macromolecules in solution".

