Chapter 18.
Determination of Protein Structures

"Several different techniques are used to study the structure of protein molecules"

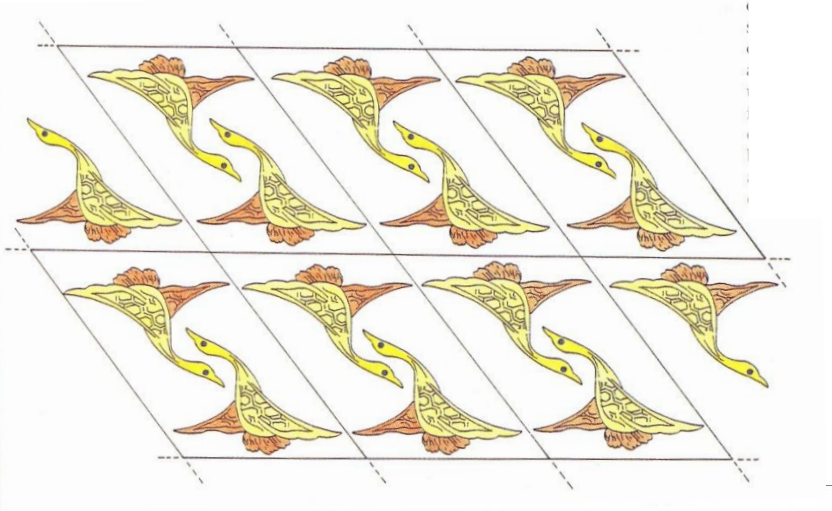
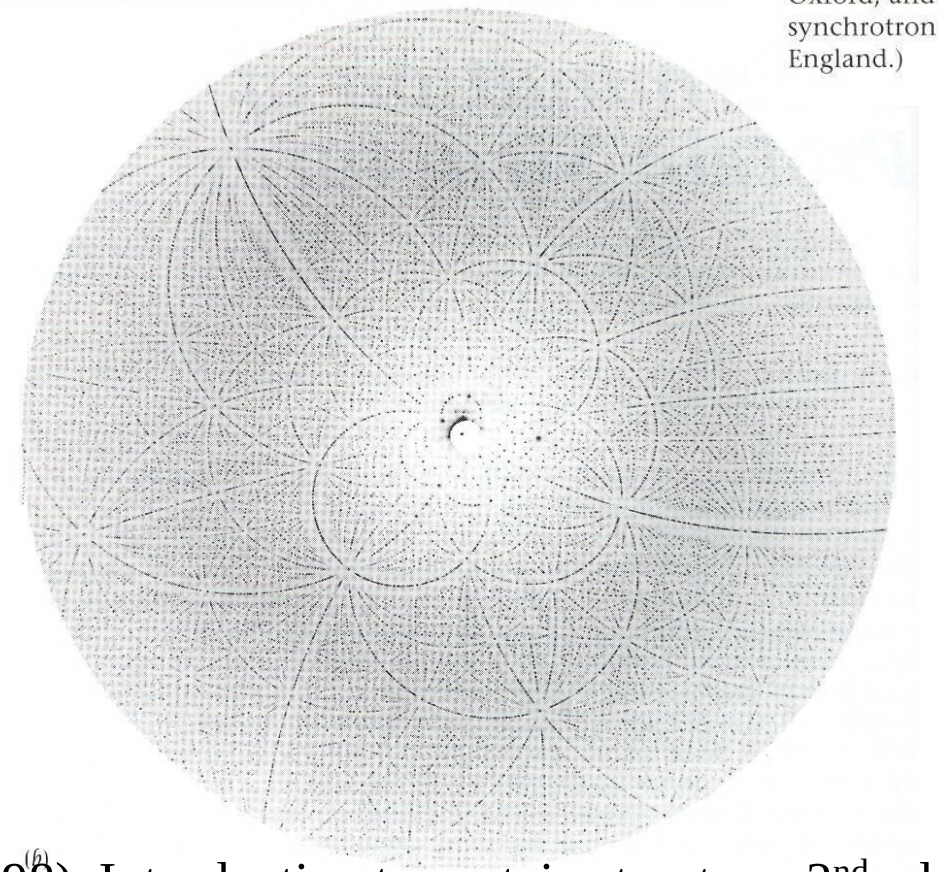


Figure 18.2 Well-ordered protein crystals diffract x-rays and produce diffraction patterns that can be recorded on film. The crystal shown in (a) is of the enzyme RuBisCo from spinach and the photograph in (b) is a recording (Laue photograph) of the diffraction pattern of a similar crystal of the same enzyme. The diffraction pattern was obtained using polychromatic radiation from a synchrotron source in the wavelength region 0.5 to 2.0 Å. More than 100,000 diffracted beams have been recorded on this film during an exposure of the crystal to x-rays for less than one second. (The Laue photograph was recorded by Janos Hajdu, Oxford, and Inger Andersson, Uppsala, at the synchrotron radiation source in Daresbury, England.)

(*a*)



(*b*)

Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.373-374.

Protein crystals are difficult to grow

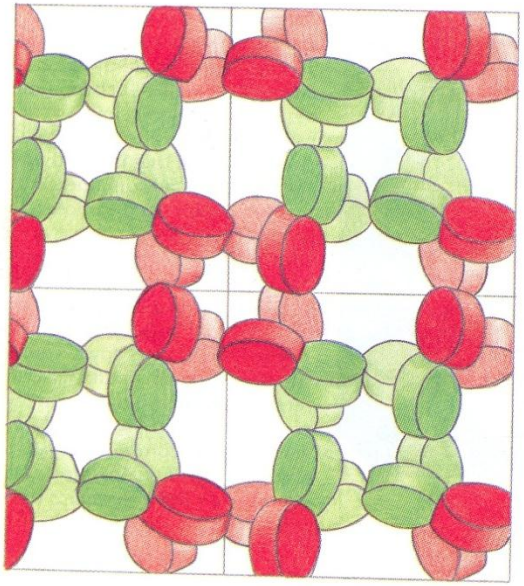
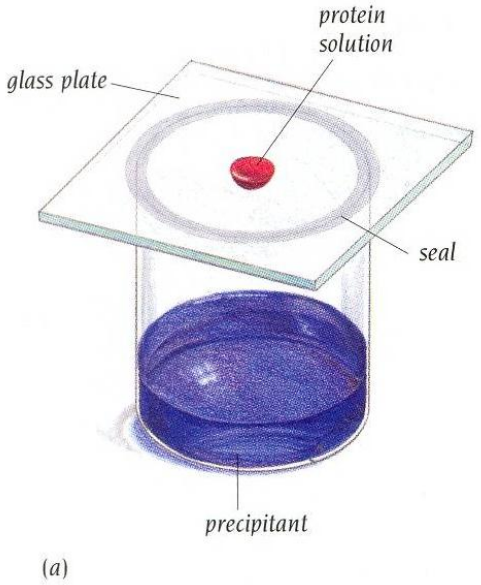


Figure 18.3 Protein crystals contain large channels and holes filled with solvent molecules, as shown in this diagram of the molecular packing in crystals of the enzyme glycolate oxidase. The subunits (colored disks) form octamers of molecular weight around 300 kDa, with a hole in the middle of each of about 15 Å diameter. Between the molecules there are channels (white) of around 70 Å diameter through the crystal. (Courtesy of Ylva Lindqvist, who determined the structure of this enzyme to 2.0 Å resolution in the laboratory of Carl Branden, Uppsala.)



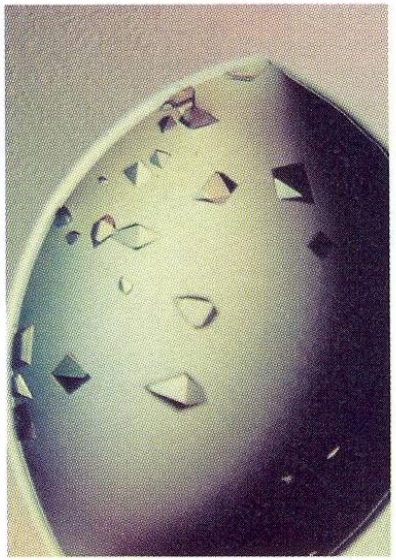


Figure 18.4 The hanging-drop method of protein crystallization. (a) About 10 µl of a 10 mg/ml protein solution in a buffer with added precipitant—such as ammonium sulfate, at a concentration below that at which it causes the protein to precipitate—is put on a thin glass plate that is sealed upside down on the top of a small container. In the container there is about 1 ml of concentrated precipitant solution. Equilibrium between the drop and the container is slowly reached through vapor diffusion, the precipitant concentration in the drop is increased by loss of water to the reservoir, and once the saturation point is reached the protein slowly comes out of solution. If other conditions such as pH and temperature are right, protein crystals will occur in the drop. (b) Crystals of recombinant enzyme RuBisCo from *Anacystis nidulans* formed by the hanging-drop method. (Courtesy of Janet Newman, Uppsala, who produced these crystals.)

Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.375-376.

X-ray sources

- Rotating anode generator: CuKα (1.5418 Å)
  Ni filter / graphite monochromator / Franck double focusing mirror / Osmic multilayer mirror

- Synchrotron: ~0.5-2.0 Å
  - High brilliance ➔ Tiny crystals (~50 μm) / Fast
  - Energy tunability ➔ MAD
  - White X-rays ➔ Laue diffraction method
    ➔ time-resolved crystallography
  - Pulsed beams: Multi vs single bunch

  - Bending magnet
  - Insertion devices
    Wiggler
    Undulator
- Free e- X-ray laser ➔ single-molecule diffraction

Third-Generation Large-scale SR Facilities

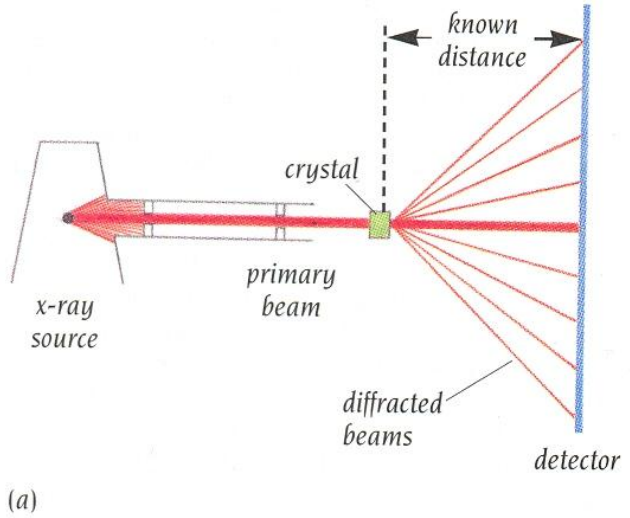
- > 5 GeV electron energy
- Capable of delivering X-rays from undulators

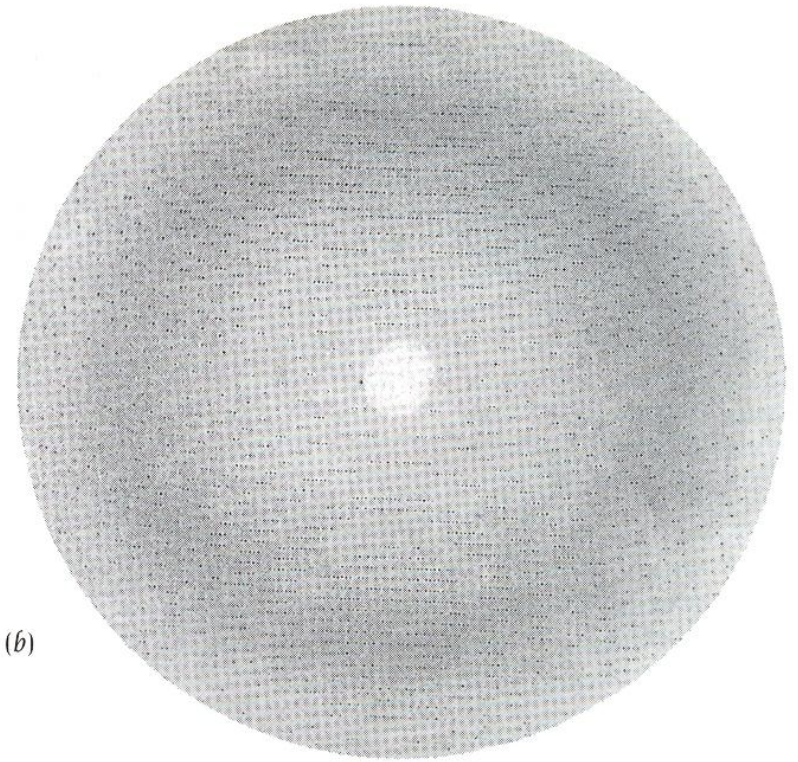
• SPring-8	Harima	8 GeV	-1997
• APS	Argonne	7 GeV	-1996
• ESRF	Grenoble	6 GeV	-1994

http://www.spring8.or.jp/e/general_info/overview/sr.html

- X-ray sources are either monochromatic or polychromatic

- X-ray data are recorded either on image plates or by electronic detectors





Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.377.

The rules for diffraction are given by Bragg's law

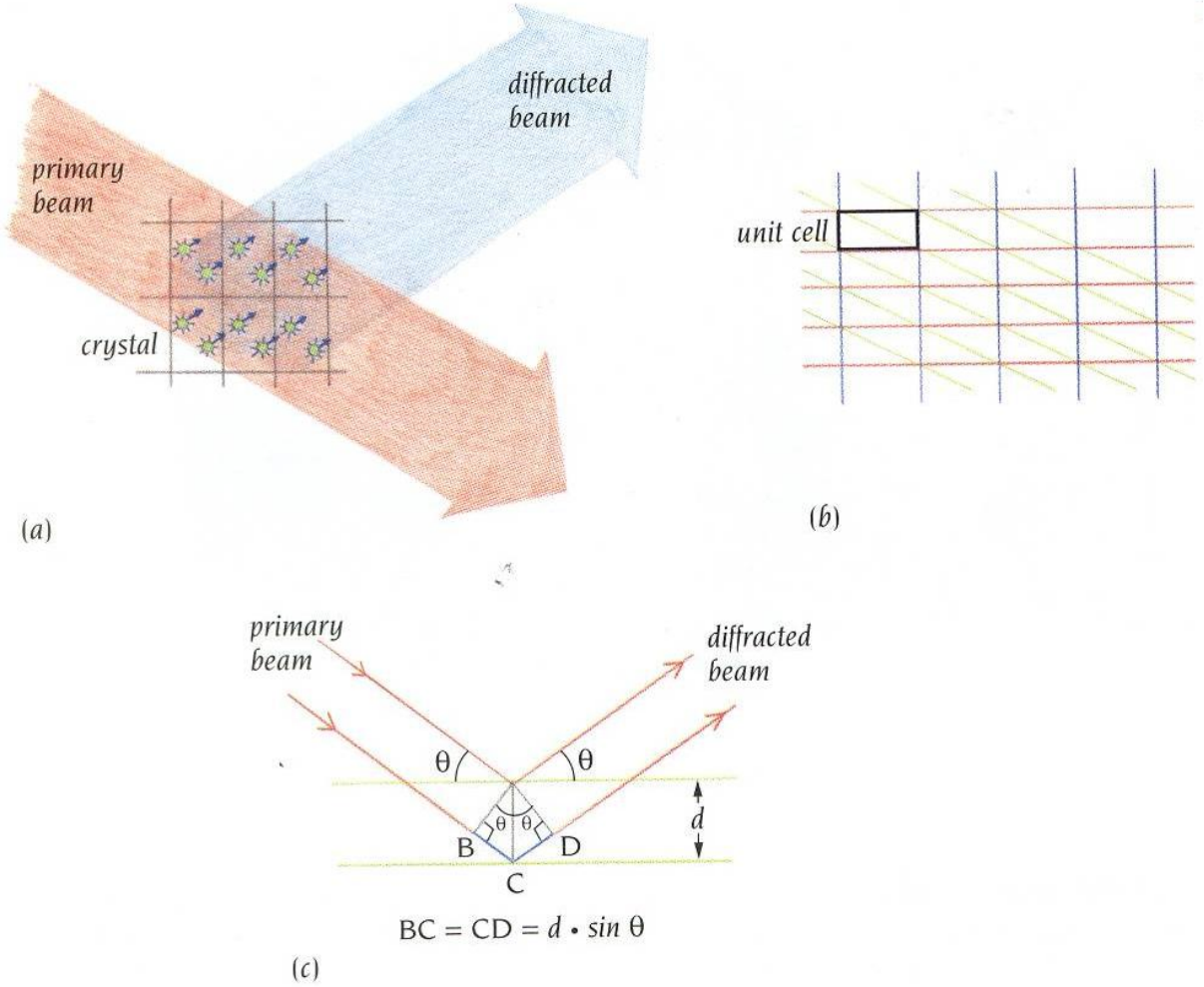


Figure 18.6 Diffraction of x-rays by a crystal. (a) When a beam of x-rays (red) shines on a crystal all atoms (green) in the crystal scatter x-rays in all directions. Most of these scattered x-rays cancel out, but in certain directions (blue arrow) they reinforce each other and add up to a diffracted beam. (b) Different sets of parallel planes can be arranged through the crystal so that each corner of all unit cells is on one of the planes of the set. The diagram shows in two dimensions three simple sets of parallel lines: red, blue, and green. A similar effect is seen when driving past a plantation of regularly spaced trees. One sees the trees arranged in different sets of parallel rows. (c) X-ray diffraction can be regarded as reflection of the primary beam from sets of parallel planes in the crystal. Two such planes are shown (green), separated by a distance d. The primary beam strikes the planes at an angle θ and the reflected beam leaves at the same angle, the reflection angle. X-rays (red) that are reflected from the lower plane have traveled farther than those from the upper plane by a distance BC + CD, which is equal to $2d \cdot \sin\theta$. Reflection can only occur when this distance is equal to the wavelength λ of the x-ray beam and Bragg's law—$2d \cdot \sin\theta = \lambda$—gives the conditions for diffraction. To determine the size of the unit cell, the crystal is oriented in the beam so that reflection is obtained from the specific set of planes in which any two adjacent planes are separated by the length of one of the unit cell axes. This distance, d, is then equal to $\lambda/(2 \cdot \sin\theta)$. The wavelength, λ, of the beam is known since we use monochromatic radiation. The reflection angle, θ, can be calculated from the position of the diffracted spot on the film, using the relation derived in Figure 18.7, where the crystal to film distance can be easily measured. The crystal is then reoriented, and the procedure is repeated for the other two axes of the unit cell.

Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.378.

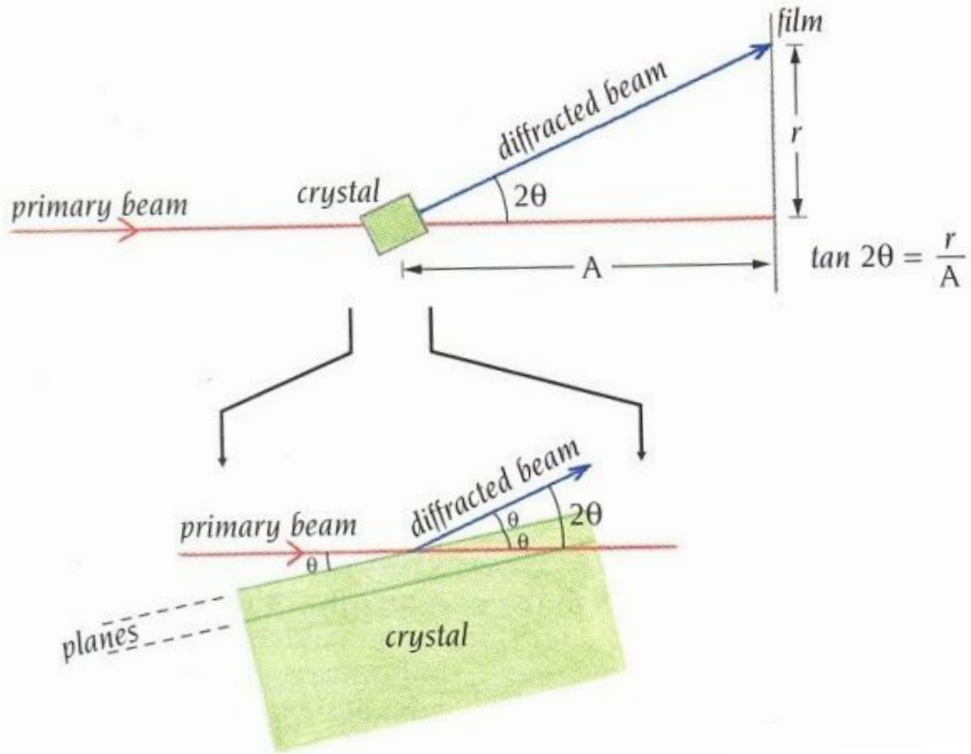
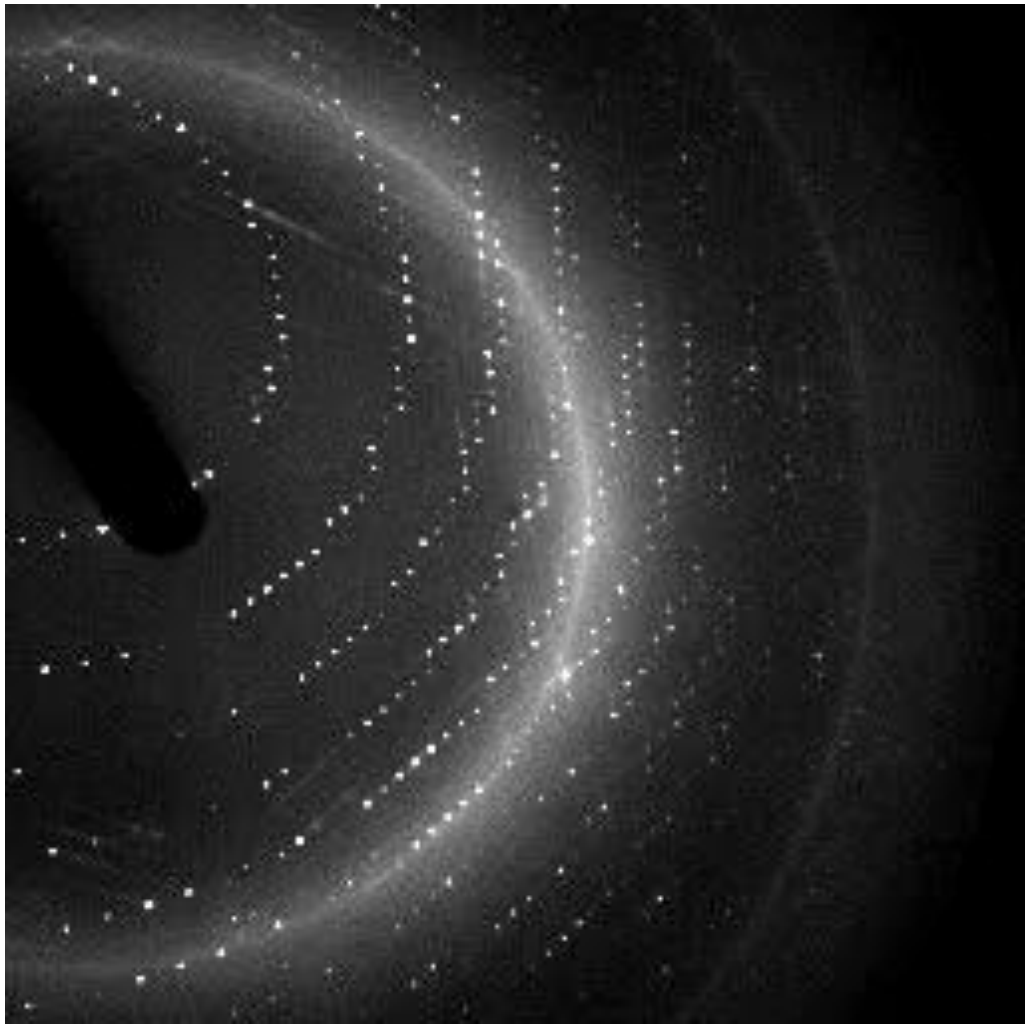


Figure 18.7 The reflection angle, θ, for a diffracted beam can be calculated from the distance (*r*) between the diffracted spot on a film and the position where the primary beam hits the film. From the geometry shown in the diagram the tangent of the angle $2\theta = r/A$. *A* is the distance between crystal and film that can be measured on the experimental equipment, while *r* can be measured on the film. Hence θ can be calculated. The angle between the primary beam and the diffracted beam is 2θ as can be seen on the enlarged insert at the bottom. It shows that this angle is equal to the angle between the primary beam and the reflecting plane plus the reflection angle, both of which are equal to θ.

Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.379.

An example of time-resolved crystallography

- Myoglobin



This movie is a compilation of actual x-ray diffraction data from a crystal of GCN4-N16A peptide in $P3_121$. Each frame is a 1-degree oscillation. This movie useful for showing people how Bragg planes intersecting the Ewald sphere make circles of spots, and how you can see that the lattice is hexagonal, and that you can tell if and when the dataset is complete.

Diffraction pattern movie
http://ucxray.berkeley.edu/~jamesh/movies/

Phase determination is the major crystallographic problem

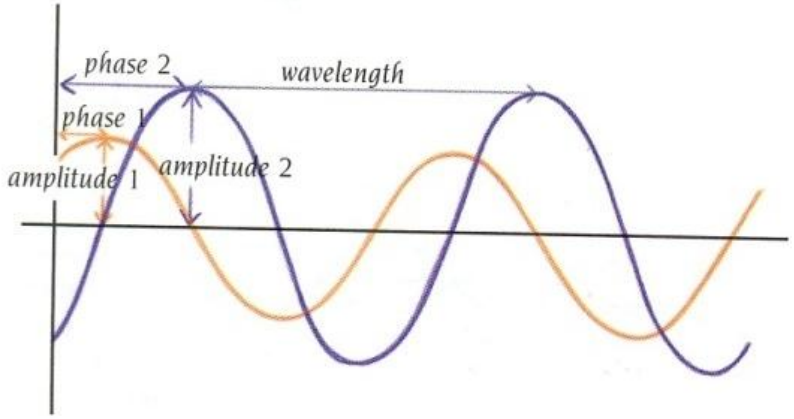
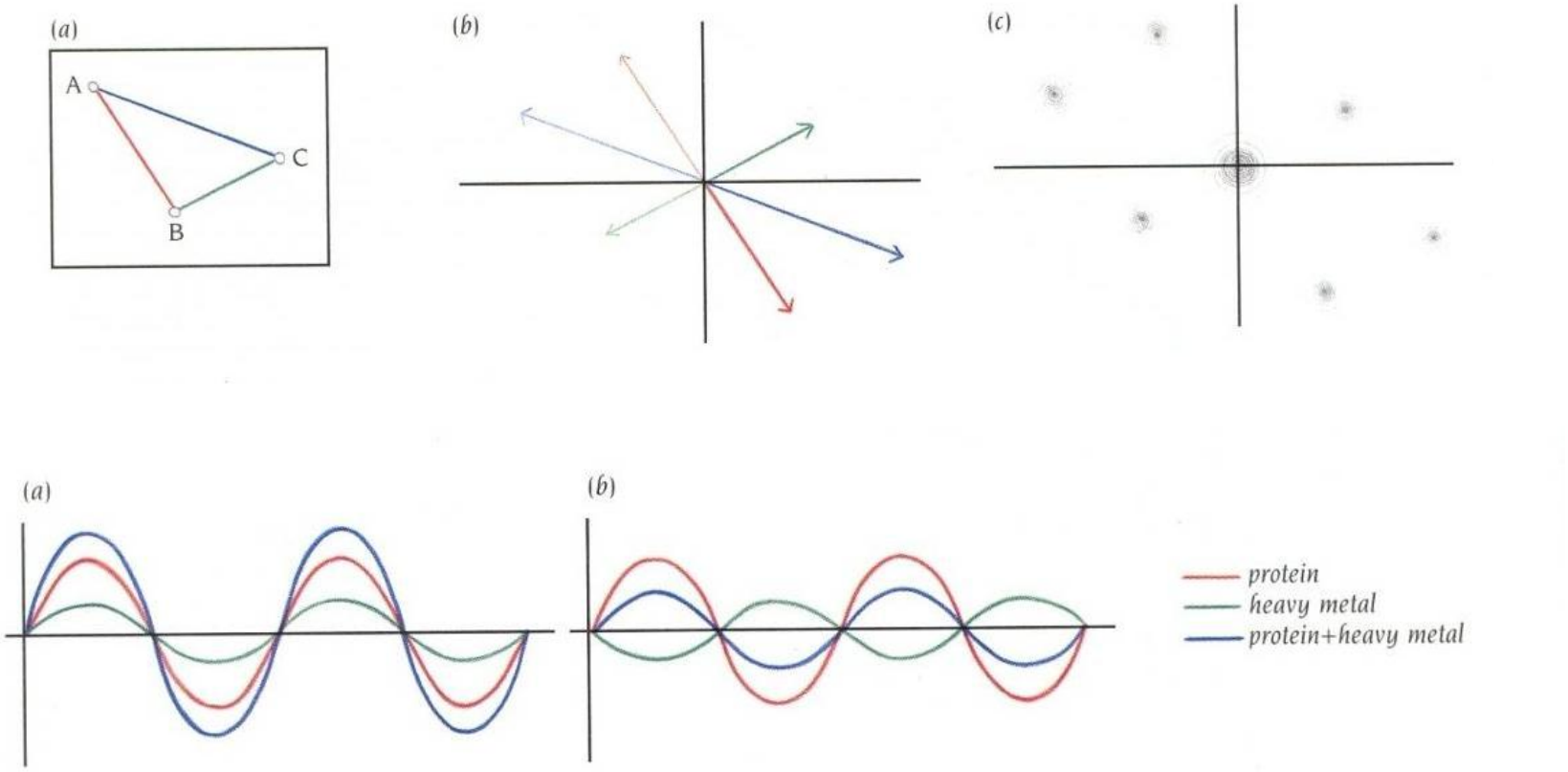


Figure 18.8 Two diffracted beams (purple and orange), each of which is defined by three properties: amplitude, which is a measure of the strength of the beam and which is proportional to the intensity of the recorded spot; phase, which is related to its interference, positive or negative, with other beams; and wavelength, which is set by the x-ray source for monochromatic radiation.

Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.379.



Phase information can also be obtained by Multiwavelength Anomalous Diffraction experiments

Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.380.

Why electron density?

- What we see as the result of a crystallographic experiment is the distribution of electrons in the molecule, i.e., an electron density map. However, since the electrons are mostly tightly localized around the nuclei, the electron density map gives us a pretty good picture of the molecule.

- X-rays interact with matter through its fluctuating electric field, which accelerates charged particles. You can think of the electrons fluctuating in position and, through their accelerations, emitting electromagnetic radiation in turn. Because electrons have a much higher charge to mass ratio than atomic nuclei or even protons, they are much more efficient in this process. Intensity of scattered radiation is proportional to the square of the charge/mass ratio, and the proton is about 2000 times as massive as the electron.

Building a model involves subjective interpretation of the data

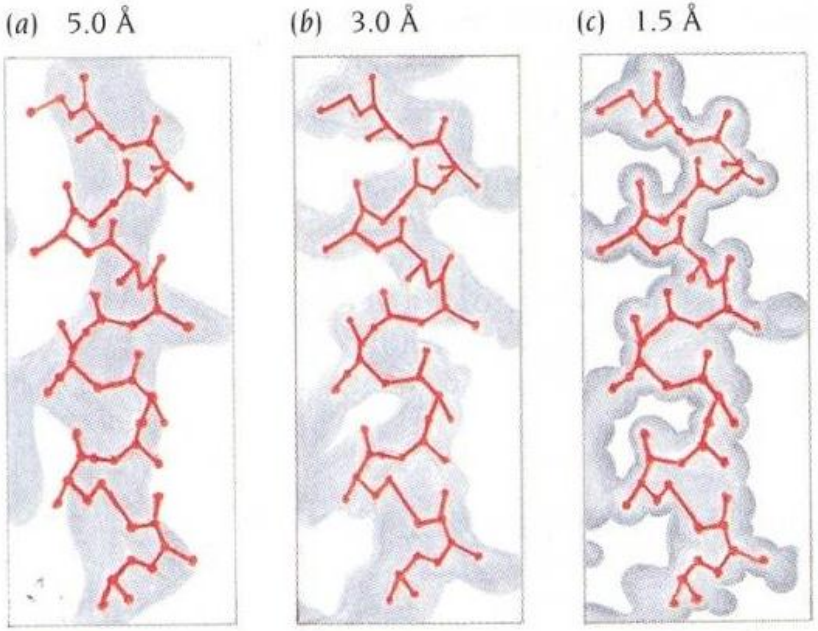
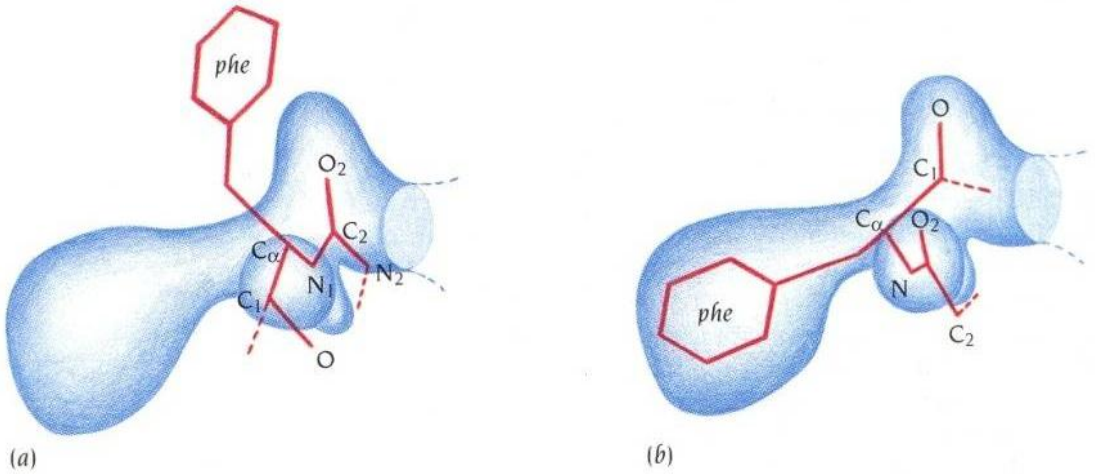
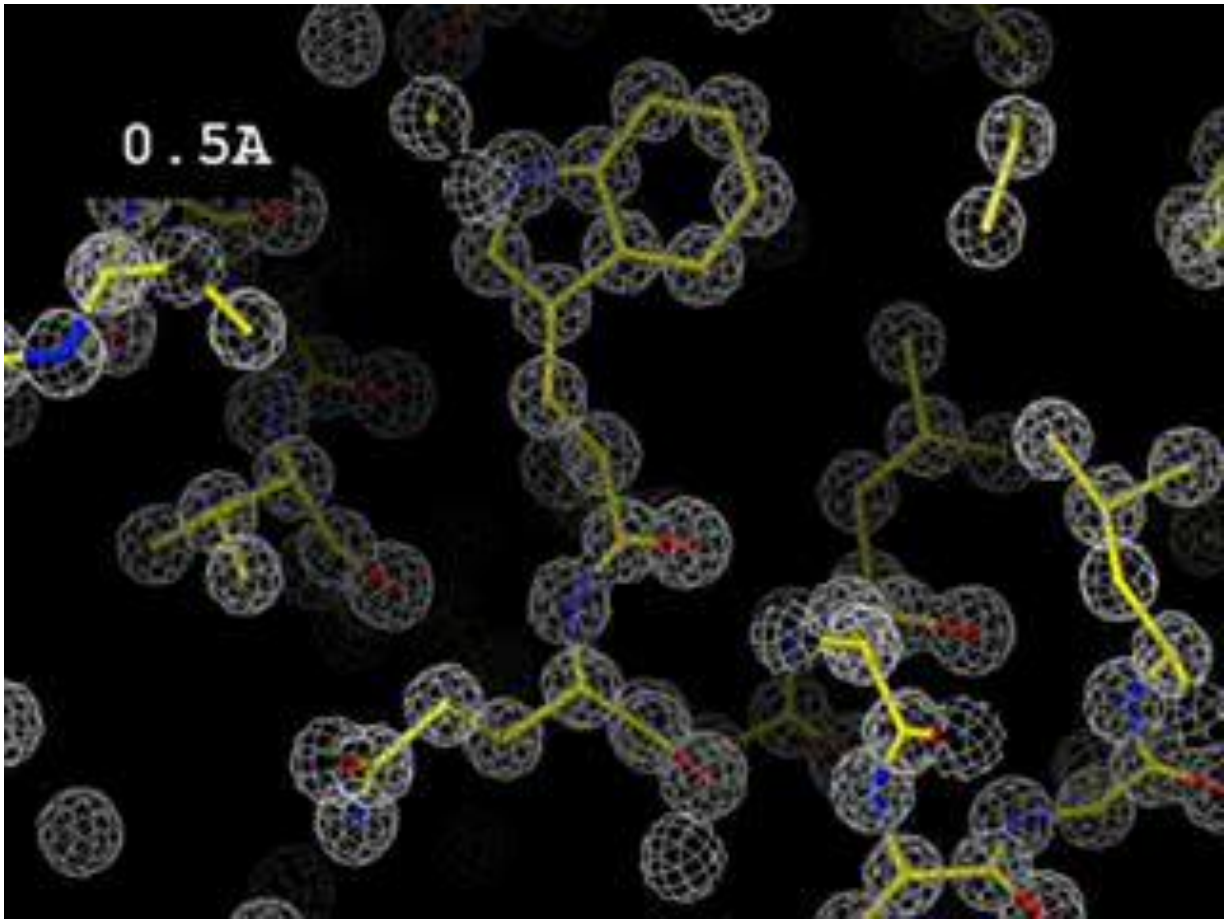


Figure 18.11 Electron-density maps at different resolution show more detail at higher resolution. (a) At low resolution (5.0 Å) individual groups of atoms are not resolved, and only the rodlike feature of an α helix can be deduced. (b) At medium resolution (3.0 Å) the path of the polypeptide chain can be traced, and (c) at high resolution (1.5 Å) individual atoms start to become resolved. Relevant parts of the protein chain (red) are superimposed on the electron densities (gray). The diagrams show one α helix from a small protein, myohemerythrin. [Adapted from W.A. Hendrickson in *Protein Engineering* (eds. D.L. Oxender and C.F. Fox.), p. 11. New York: Liss, 1987.]



Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.382.



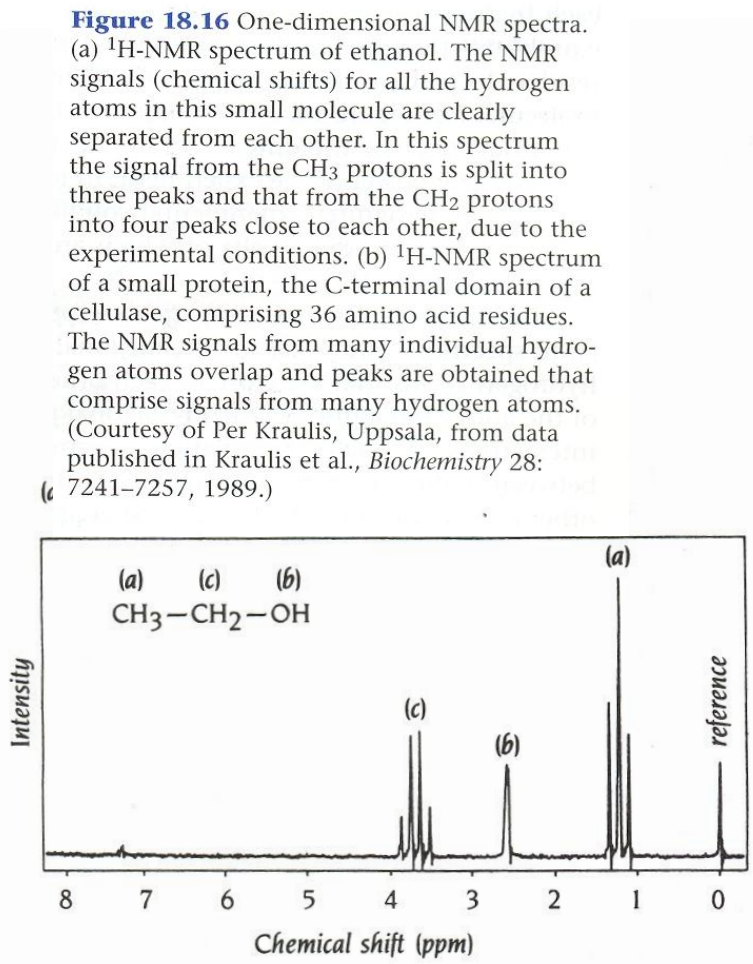
This movie displays a calculated electron density map, contoured at 1 sigma, as the resolution limit is adjusted slowly from 0.5A to 6A. The phases are perfect, and so are the amplitudes (R-factor = 0.0%) for all the resolutions displayed. Note that, even for a perfect map, you expect side chains to poke out of density at 3.5A.

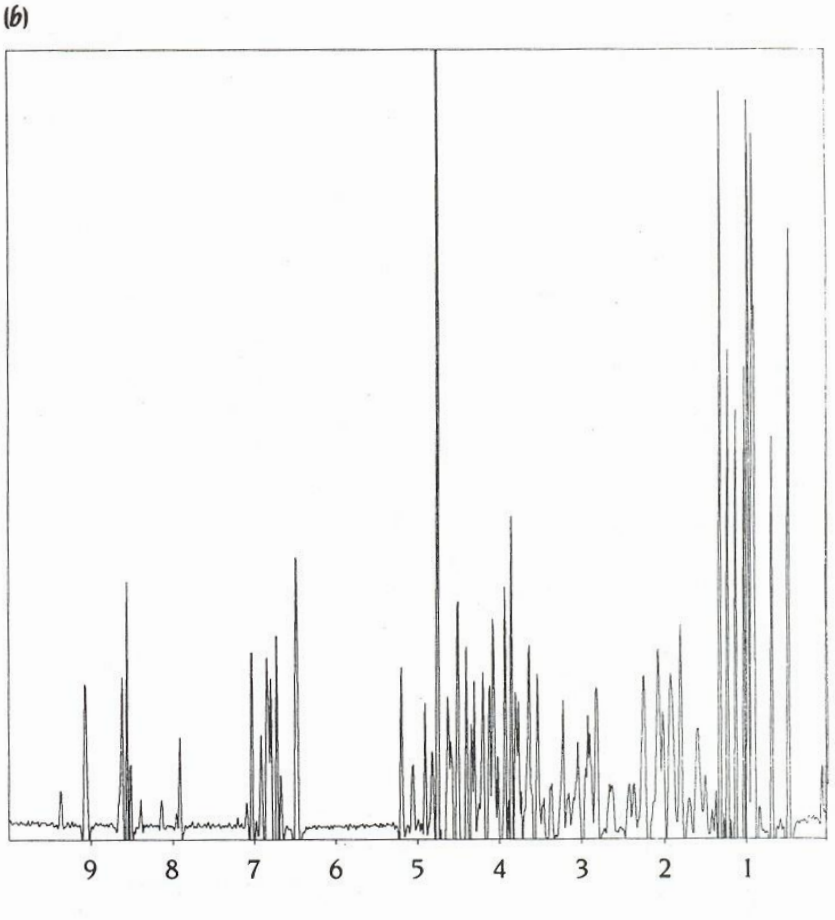
Resolution movie
http://ucxray.berkeley.edu/~jamesh/movies/

- Errors in the initial model are removed by refinement
- Recent technological advances have greatly influences protein crystallography

NMR methods use the magnetic properties of atomic nuclei

Figure 18.16 One-dimensional NMR spectra. (a) ^{1}H-NMR spectrum of ethanol. The NMR signals (chemical shifts) for all the hydrogen atoms in this small molecule are clearly separated from each other. In this spectrum the signal from the CH_3 protons is split into three peaks and that from the CH_2 protons into four peaks close to each other, due to the experimental conditions. (b) ^{1}H-NMR spectrum of a small protein, the C-terminal domain of a cellulase, comprising 36 amino acid residues. The NMR signals from many individual hydrogen atoms overlap and peaks are obtained that comprise signals from many hydrogen atoms. (Courtesy of Per Kraulis, Uppsala, from data published in Kraulis et al., *Biochemistry* 28: 7241–7257, 1989.)





Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.387.

Two dimensional NMR spectra of proteins are interpreted by the method of sequence assignment

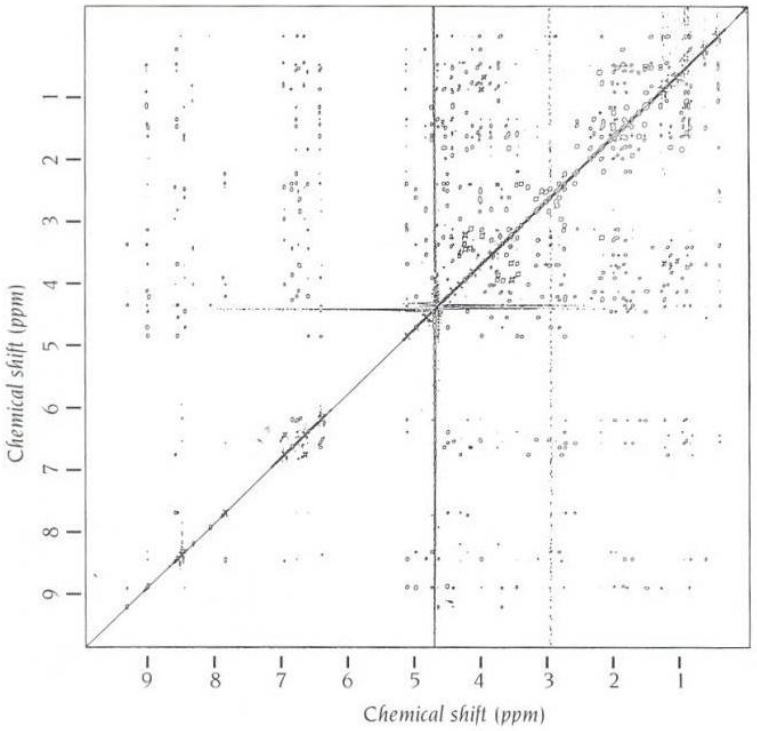


Figure 18.17 Two-dimensional NMR spectrum of the C-terminal domain of a cellulase. The peaks along the diagonal correspond to the spectrum shown in Figure 18.16b. The off-diagonal peaks in this NOE spectrum represent interactions between hydrogen atoms that are closer than 5 Å to each other in space. From such a spectrum one can obtain information on both the secondary and tertiary structures of the protein. (Courtesy of Per Kraulis, Uppsala.)

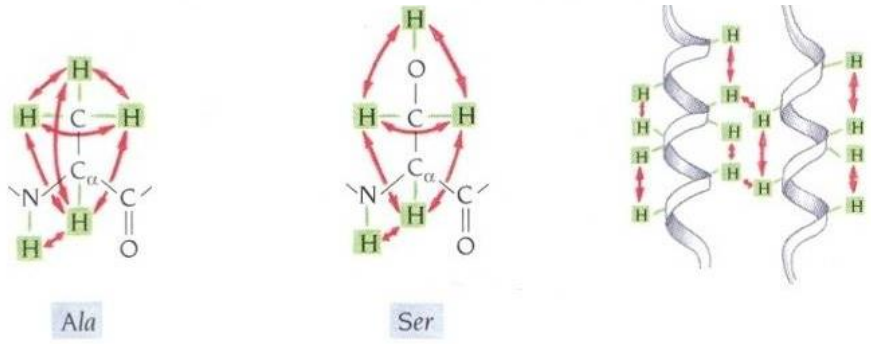


Figure 18.18 (a) COSY NMR experiments give signals that correspond to hydrogen atoms that are covalently connected through one or two other atoms. Since hydrogen atoms in two adjacent residues are covalently connected through at least three other atoms (for instance, HCα-C′-NH), all COSY signals reveal interactions within the same amino acid residue. These interactions are different for different types of side chains. The NMR signals therefore give a "fingerprint" of each amino acid. The diagram illustrates fingerprints (red) of residues Ala and Ser. (b) NOE NMR experiments give signals that correspond to hydrogen atoms that are close together in space (less than 5 Å), even though they may be far apart in the amino acid sequence. Both secondary and tertiary structures of small protein molecules can be derived from a collection of such signals, which define distance constraints between a number of hydrogen atoms along the polypeptide chain.

Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.388.

Nuclear Overhauser effect (NOE)

Definition: NOE is the change in intensity of the resonance of one nucleus, when the transitions of another nucleus (lying close to the first nucleus in space within 5 Å) are perturbed by irradiation.

Physical origin: NOE is a response of the total system to restore thermal equilibrium, when the transition of one nucleus is saturated with a weak R.F. field. When the population difference across a nucleus is changed (i.e., saturated) by irradiation, the system responds by changing the population differences across other low-lying nuclei as a means to compensate the change.

Distance constraints are used to derive possible structures of a protein molecule

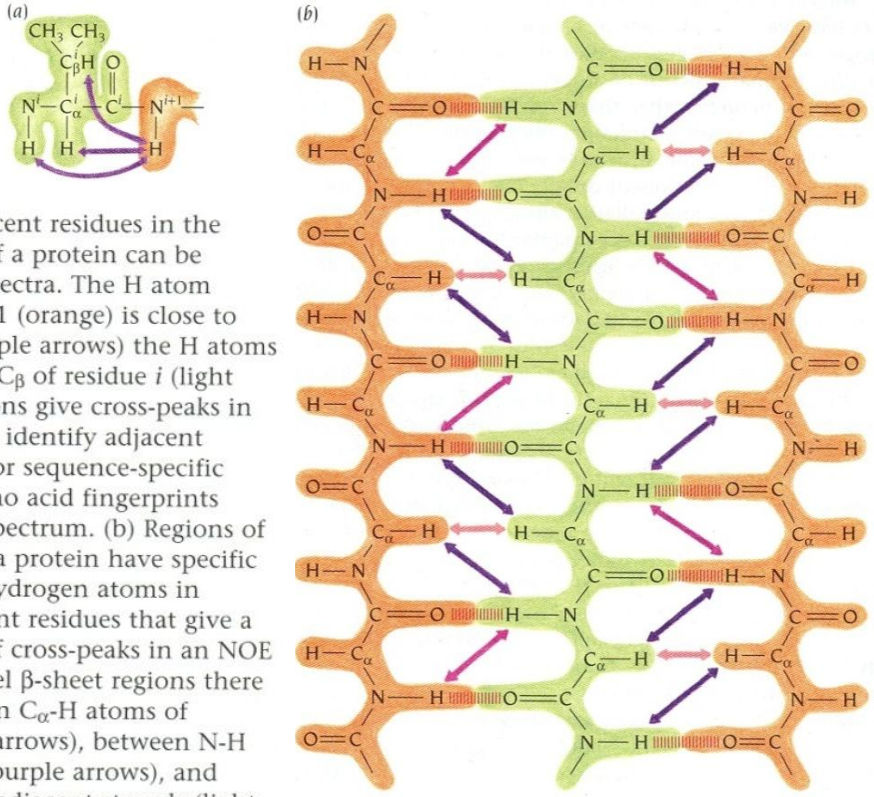


Figure 18.19 (a) Adjacent residues in the amino acid sequence of a protein can be identified from NOE spectra. The H atom attached to residue $i + 1$ (orange) is close to and interacts with (purple arrows) the H atoms attached to N, C_{α}, and C_{β} of residue i (light green). These interactions give cross-peaks in the NOE spectrum that identify adjacent residues and are used for sequence-specific assignment of the amino acid fingerprints derived from a COSY spectrum. (b) Regions of secondary structure in a protein have specific interactions between hydrogen atoms in sequentially nonadjacent residues that give a characteristic pattern of cross-peaks in an NOE spectrum. In antiparallel β-sheet regions there are interactions between C_{α}-H atoms of adjacent strands (pink arrows), between N-H and C_{α}-H atoms (dark purple arrows), and between N-H atoms of adjacent strands (light purple arrows). The corresponding pattern of cross-peaks in an NOE spectrum identifies the residues that form the antiparallel β sheet. Parallel β sheets and α helices are identified in a similar way.

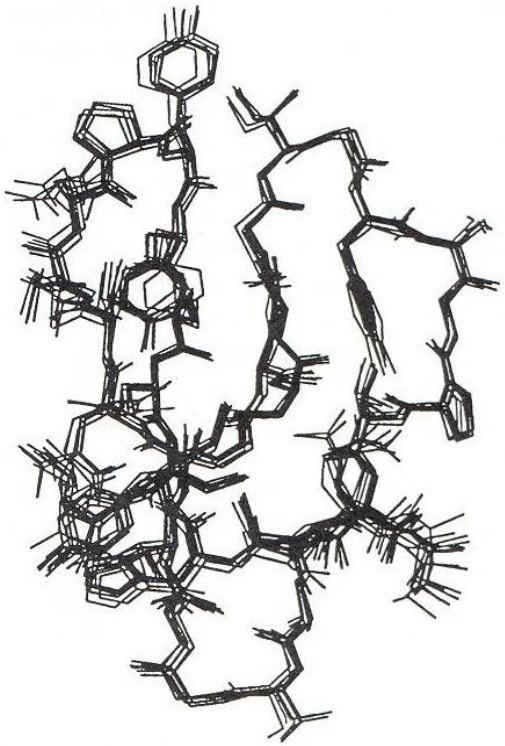


Figure 18.20 The two-dimensional NMR spectrum shown in Figure 18.17 was used to derive a number of distance constraints for different hydrogen atoms along the polypeptide chain of the C-terminal domain of a cellulase. The diagram shows 10 superimposed structures that all satisfy the distance constraints equally well. These structures are all quite similar since a large number of constraints were experimentally obtained. (Courtesy of P. Kraulis, Uppsala, from data published in P. Kraulis et al., *Biochemistry* 28: 7241–7257, 1989, by copyright permission of the American Chemical Society.)

Branden & Tooze (1998), Introduction to protein structure, 2nd ed., p.389-390.

Visualization of Protein Structures

Placement of protein and RNA structures into a 5 Å-resolution map of the 50S ribosomal subunit

Nenad Ban*†, Poul Nissen*, Jeffrey Hansen*†, Malcolm Capel‡, Peter B. Moore*§ & Thomas A. Steitz*†§

*Departments of * Molecular Biophysics & Biochemistry and § Chemistry, Yale University and the † Howard Hughes Medical Institute, New Haven, Connecticut 06520-8114, USA*

‡ Department of Biology, Brookhaven National Laboratory, Upton, New York 11973, USA

The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution

Nenad Ban *et al*. &Thomas A. Steitz
Science, Vol 289, Issue 5481 (2000) 905-920