

Structural Bioinformatics (Part 1)

Presenter
Jose Lugo-Martinez
Lab Meeting

March 1, 2011





Outline

- What is structural bioinformatics?
- Fundamentals of protein structure (review)
- Experimental methods for structure determination
 - X-Ray Crystallography
 - NMR



What is Structural Bioinformatics?

- “Structural bioinformatics is the subdiscipline of bioinformatics that focuses on the representation, storage, retrieval, analysis, and display of structural information at the atomic and subcellular spatial scales” (Gu and Bourne, 2009)
- **Goals:**
 - Create methods for information manipulation of biological macromolecules and applying these methods to solve problems in biology
 - Create methods that elucidate new knowledge



Structural Bioinformatics Promise

- Creating an infrastructure for building up structural models from component parts
- Gaining the ability to understand the design principles of proteins so that new functionalities can be created
- Learning how to design drugs efficiently based on structural knowledge of their target
- Catalyzing the development of simulation models that can give insight into function based on structural simulations



Structural Bioinformatics Challenges

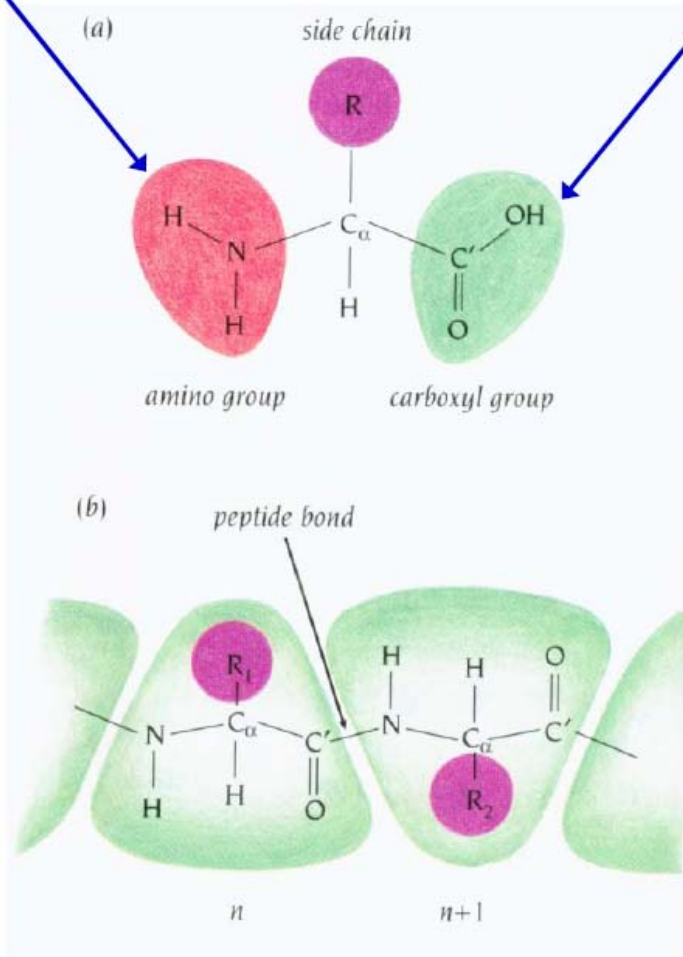
- Methods to support structural biology and structural genomics
 - target selection
 - tracking experimental crystallization trials
 - analysis of experimental data
 - storing molecular structures in databases
- Understanding structural basis of biological phenomena
 - visualization
 - characterization and classification
 - prediction
 - simulation
- Integrating structural data with other data types



Protein

base

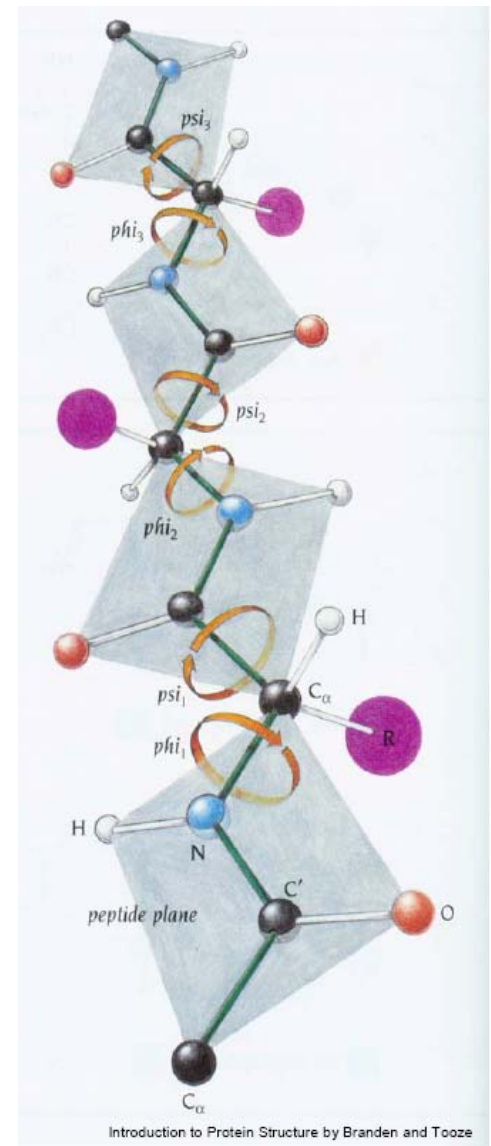
acid



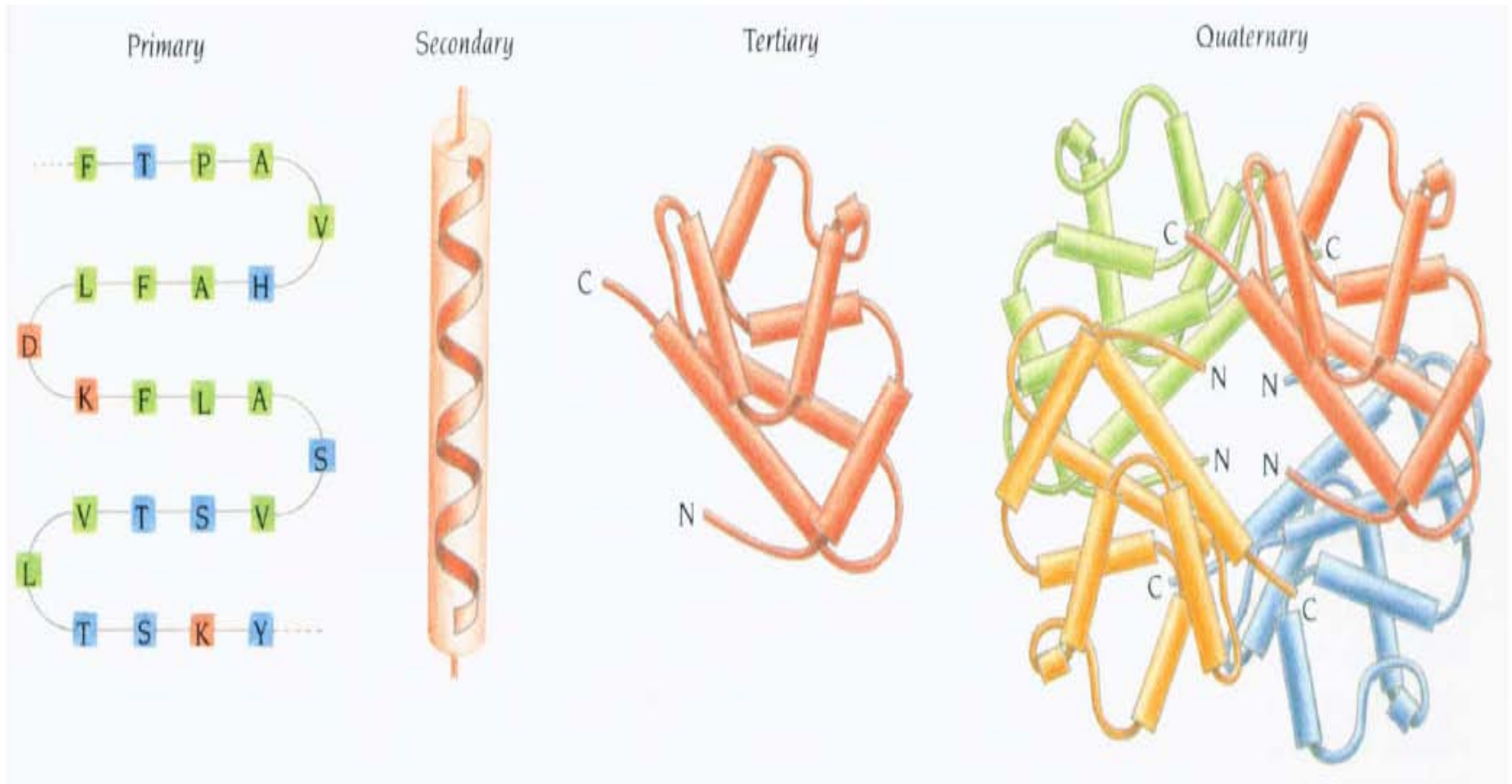
a) amino acid

b) peptide bond

Introduction to Protein Structure by Branden and Tooze



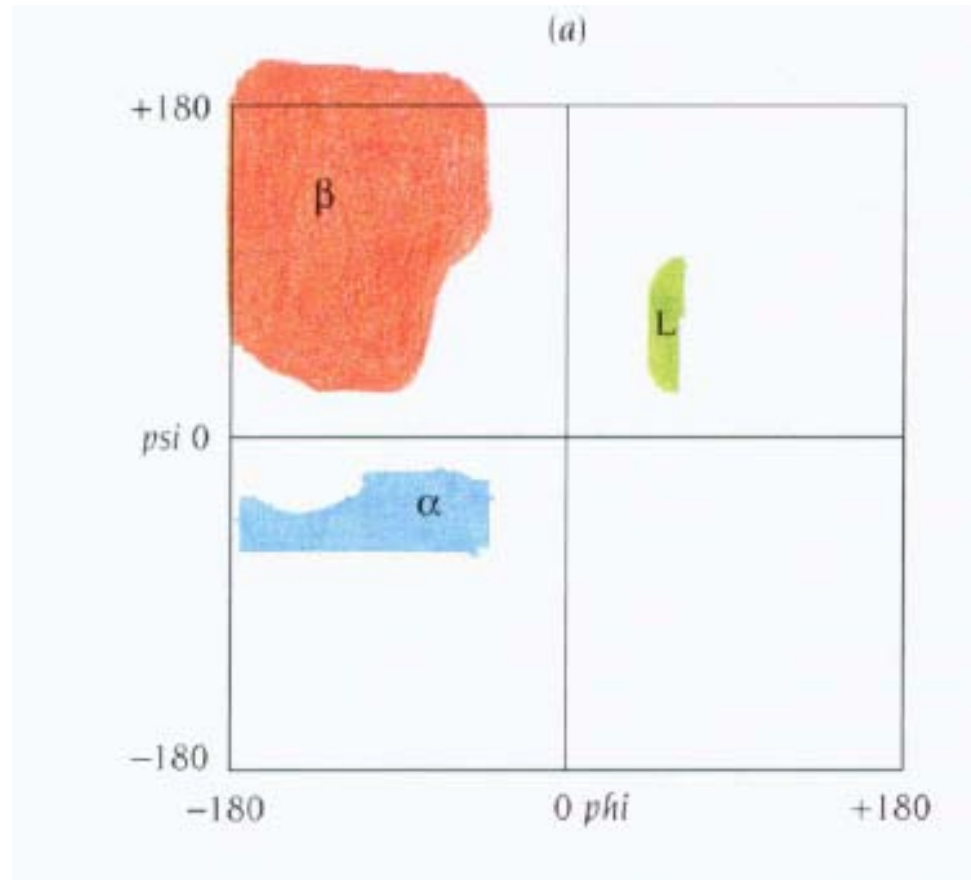
Levels of Protein Structure





INDIANA UNIVERSITY

Ramachandran Plot



Introduction to Protein Structure by Branden and Tooze



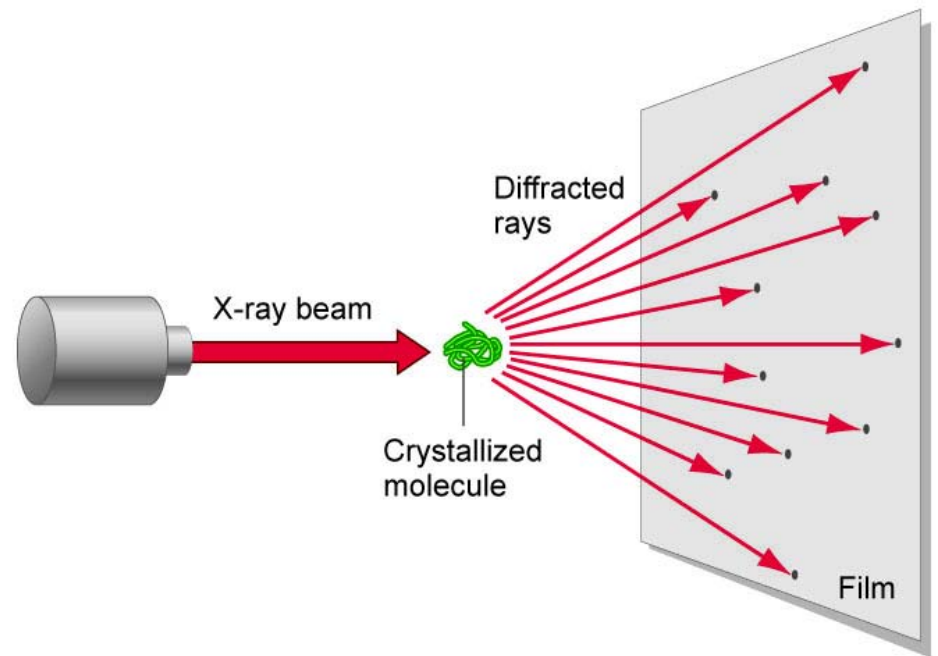
X-ray Crystallography

- For small-molecules - highly automated
 - Cambridge Structural Database (CCSD)
 - Growth rate: 15,000 new structures/year
- For macromolecules – manual intervention
 - Protein Data Bank (PDB)
 - Growth rate: 1,500 new structures/year
 - **Goal:**
 - High-throughput structure determination
 - How?
 - Automate!! (reduce/eliminate human intervention)



X-ray Crystallography in a Nutshell

- Steps
 - Express/purify the protein
 - Crystallize the protein
 - Collect diffraction data
 - Determine the space group, unit cell dimensions, and number/symmetry of molecules per unit cell
 - Solve the “phase problem”
 - Calculate an electron density map
 - Build a model of the molecule to fit the electron density map
 - Refine the model



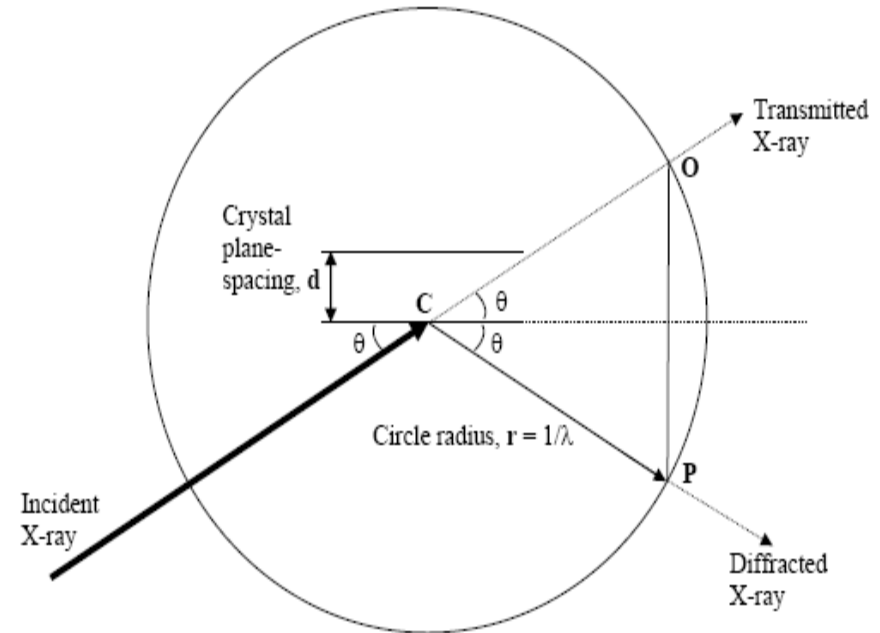
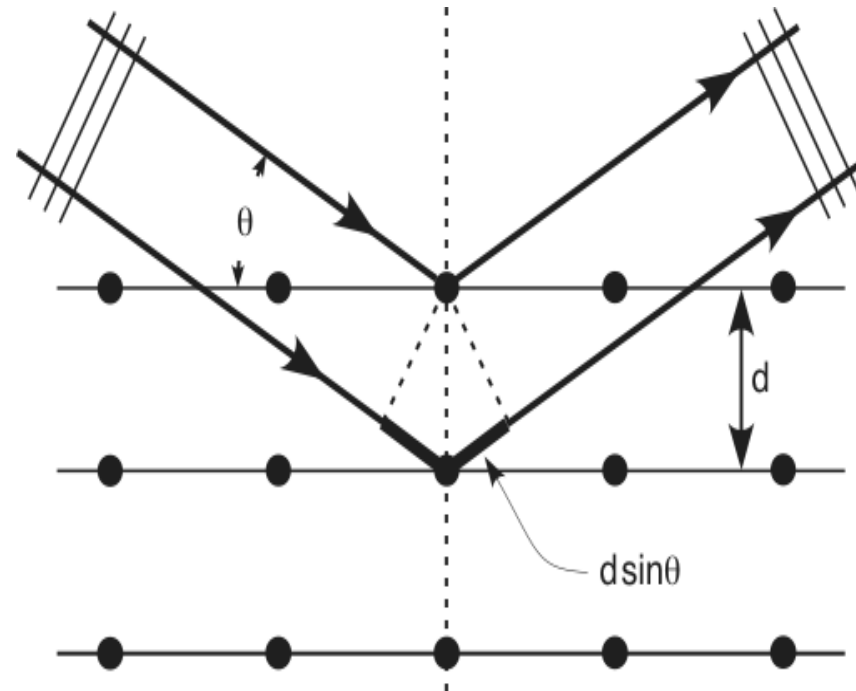
~85-90% of PDB entries



Sample Preparation and Crystallization

- Involves screening a wide range of solution conditions to find those conducive to crystal formation
- Common parameters “tweaked” during trials
 - protein concentration
 - buffers and ions
 - ionic strength
 - pH
 - temperature
 - precipitants

Diffraction from Crystals



If an X-ray impinges on a plane of atoms in the crystal at an appropriate angle, it will be diffracted, interfere constructively with parallel beams, and result in a detectable diffracted X-ray (a reflection).



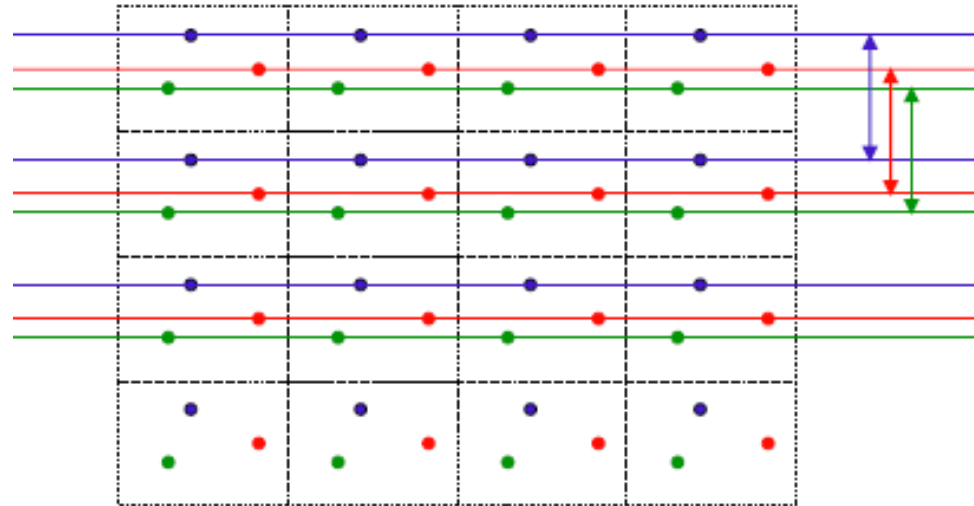
Collection of Diffraction Data

1. Shine an X-ray at a particular angle to the crystal. The X-ray will hit the different planes at different angles and some will diffract constructively, giving reflections at certain angles from the source beam.
2. Now rotate the crystal around an axis perpendicular to the source beam. Repeat 1

Note: Reflections positions do not provide any information about the structures of the molecules



- X-rays reflected from a plane through one type of atom will be **in-phase** with X-rays reflected from a parallel plane through the **same type of atom**. Thus, **constructive interference** between these reflected rays and the possibility of detecting a reflection.
- But X-rays reflected from a plane through one type of atom will NOT be in phase with X-rays reflected from a parallel plane through a different type of atom
- The reflected X-ray at the Bragg angle (θ) will be a convolution of the reflected X-rays coming from each set of parallel planes (i.e. each type of atom) creating "**Phase Problem**"





Reflection Intensities and Phases

- Need to determine intensity and phase of X-ray by
 - **de-convolute it into its component waves**
 - **analyze each component wave to deduce the *electron density map***
- Intensities can be determined from intensities of the reflection on the X-ray detector



“Phase” Problem Approaches

- *Molecular Replacement*
 - if structure of a similar protein is known, use it to build a model of your protein, then compare model with observed diffraction intensities
- *Multiple Isomorphous Replacement*
 - a series of heavy atom derivatives of the crystals are prepared and diffraction data collected on the derivatives
- *Multi-wavelength Anomalous Dispersion (MAD)*
 - typically samples are prepared such that methionine residues are replaced by selenomethionine. Diffraction data are collected at several wavelengths (at a synchrotron) and the scattering of X-rays by the selenium atom is used to obtain phases



Electron Density Map, Solving the Structure, ...

- Generate electron density map by converting the intensities and phases of the reflected X-rays to the electron densities at each position in the unit cell
 - using Fourier transformation
- Use electron density map and known protein sequence to build (then iteratively refine) molecular models that are in best agreement with the experimental data
- Submit to PDB
 - validation criteria to assess global and local correctness

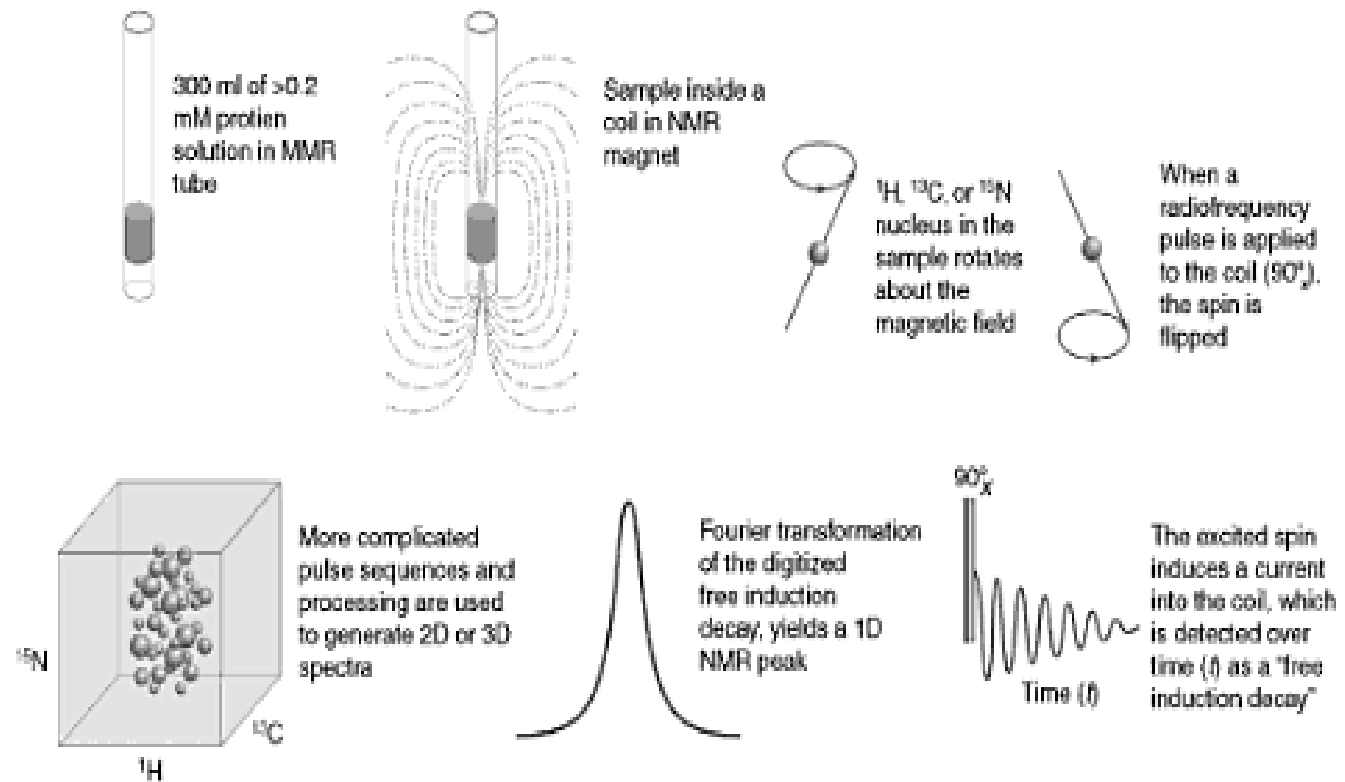
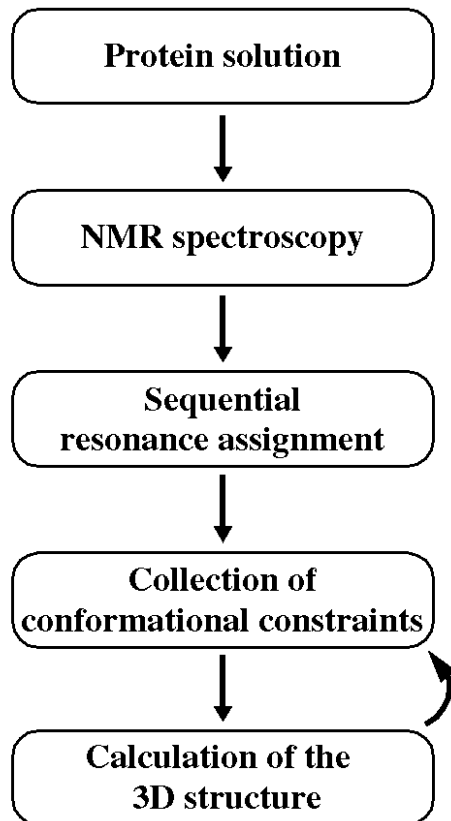


Nuclear Magnetic Resonance

- Spectroscopy
 - the study of the interaction of electromagnetic radiation with matter
- NMR is a spectroscopic method
 - phenomenon that occurs when nuclei of certain atoms are immersed in a static magnetic field and exposed to a second oscillating magnetic field
 - some nuclei experience it, and others do not
 - study physical, chemical, and biological properties of matter (at atom-level)



NMR in a Nutshell



~10-15% of PDB entries



Why it works?

- When an atomic nucleus with non-zero spin is placed in a magnetic field, its energy of interaction with the magnetic field becomes quantized
- Transitions between these energy levels can be measured spectroscopically
- At equilibrium, there is an excess population in the lower energy level
- Upon irradiation at radiofrequency, the energy level populations are equalized
- The irradiating source is removed and the system gradually relaxes back to equilibrium, emitting radiowaves at the resonant frequency
- Different types of nuclei [e.g. ^1H versus ^{13}C] have very different frequencies so it is possible to collect a spectrum in which only one type of nucleus is observed, e.g. ^1H NMR spectrum.
- Different ^1H nuclei in the same molecule will be in different chemical environments, which shields them to varying extents from the permanent magnetic field. Consequently they resonate at slightly different frequencies.



Structure Calculations

- Once the many restraints above have been determined, one performs restrained molecular dynamics (rMD) calculations with the goal of identifying candidate structures that satisfy these experimental restraints.
- Typically one then goes through an iterative process of collecting or assigning additional restraints, correcting errors and repeating calculations until the family of structures obtained satisfies the experimental data adequately.



- **X-ray Crystallography**

- Pros
 - no size limitation exists
- Cons
 - suitable crystal must be found

- **NMR**

- Pros
 - atomic-level resolution (as opposed to other spectroscopic methods)
 - proteins molecules in the solution phase
 - lots of useful information
 - could be partially disordered or exist in various stable conformations
- Cons
 - small molecules (size limit 5-25 kDa)
 - requires highly soluble protein



What we covered today

- What is structural bioinformatics?
 - Chapter 1
- Fundamentals of protein structure
 - Chapter 2
- X-ray crystallography
 - Chapter 4
- Nuclear Magnetic Resonance
 - Chapter 5



INDIANA UNIVERSITY

Thank You!

QUESTIONS?