I690/B680 Structural Bioinformatics Spring 2006

Protein Structure Determination by NMR Spectroscopy

Suggested Reading

(1) Van Holde, Johnson, Ho. "Principles of Physical Biochemistry", 2nd Ed., Prentice Hall, 2006, Chapter 12

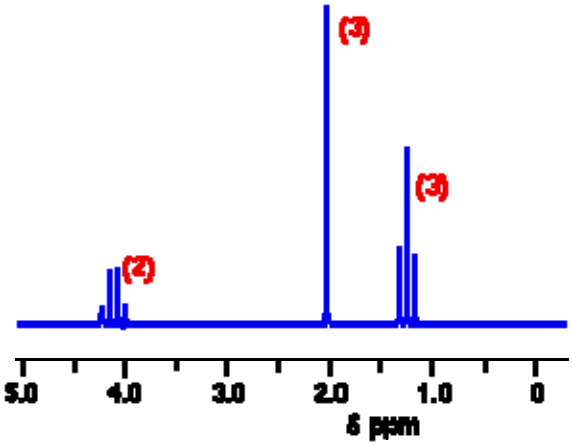
1. Basic Principles of NMR

- NMR is a routine technique used to determine the chemical structures of small organic molecules

- When an atomic nucleus with non-zero spin is placed in a magnetic field, its energy of interaction with the magnetic field becomes quantized. For nuclei with spin (I) of ½ [for example, ^{1}H, ^{13}C, ^{15}N, and ^{31}P] there are two possible energy levels.
- Transitions between these energy levels can be measured spectroscopically. They are in the radiofrequency range

- At equilibrium, there is an excess population in the lower energy level.
- Upon irradiation at radiofrequency, the energy level populations are equalized (saturated).
- 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. ^{1}H 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. ^{1}H NMR spectrum.
- Different ^{1}H 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. Thus the position of a resonance line in the spectrum is referred to as chemical shift, and is measured in ppm.
- Nuclei connected by a small number of chemical bonds influence each other resulting in "splitting" of the energy levels and the spectral lines. This splitting, measured in Hz, is referred to as spin-spin coupling or J-coupling.
- The ^{1}H NMR spectrum of ethanol demonstrates these various features.



- The spectra of proteins are much more complicated. Consequently, in order to separate the signals from each other, it is necessary to collect multidimensionial spectra (2D, 3D, 4D). We will briefly discuss what these are, but we will not go through any details.

2. How is NMR Used in Biochemistry?

- NMR can provide information about
  - Macromolecular structure
  - Macromolecular dynamics
  - Interactions of macromolecules with each other or with small ligands
  - Chemical equilibria (stabilities)
  - Protonation states
  - Rates of chemical processes or conformational equilibria
- The advantages of NMR over other methods:
  - It offers atomic level resolution (unlike other spectroscopic methods)
  - It is done in solution – no crystallization needed
  - It is very information-rich (see above)
- The disadvantages of NMR over other methods:
  - It is very insensitive – a typical sample is ~0.5 mL of 1 mM protein (10 mg of a 20 kDa protein)
  - Requires that the protein be highly soluble (>100 μM)
  - Size limit – <30 kDa for routine applications, possibly >50 kDa with sophisticated methods on a good day

3. The Steps in NMR Structure Determination of Proteins

1) Express the protein and incorporate isotope labels (^{15}N, ^{13}C, ^{2}H)
2) Optimize sample conditions
3) Assign the ^{1}H, ^{15}N, and ^{13}C resonances using multidimensional spectra
4) Collect spectra to obtain geometrical restraints
5) Do restrained MD using the geometrical restraints to obtain a family of structures

6) Analyze structures and iterate through steps 3-5 to correct errors, obtain additional restraints, and improve resolution of structural ensemble

- We will focus our discussion on the types of structural restraints that can be obtained from NMR spectra

4. Chemical Shifts

- The chemical shift of a nucleus is dependent on the chemical nature of the adjacent groups in the structure.
- Therefore the chemical shift contains information about the 3D structural environment of the nucleus.
- The theory relating chemical shift to 3D structure is not as well-developed as we would like so chemical shifts are not generally used as restraints.
- However, **C_α, C_β, and H_α chemical shifts** are quite sensitive to the **secondary structure** environment so can be used to **constrain the backbone torsion angles** in structure calculations

5. Spin-Spin (J-) Couplings

- Spin-spin couplings represent the influence of one nucleus on another mediated through the intervening chemical bonds (the electronic orbitals)
- Vicinal (3-bond) J-couplings are sensitive to the torsion angle (θ) of the middle bond, according to the Karplus relationship:
  $^3J = A \cos^2\theta + B \cos\theta + C$
- For particular types of couplings in a protein, the constants A, B, and C have been quantified empirically. For example, for H_N to H_α couplings:
  $^3J_{HNH\alpha} = 6.4 \cos^2\phi - 1.4 \cos\phi + 1.9$
  and typical values in various secondary structures are:

Structure Type	ϕ	$^3J_{HNH\alpha}$ (Hz)
α-helix	-57º	3.9
3_{10}-helix	-60º	4.2
Antiparallel β-sheet	-139º	8.9
Parallel β-sheet	-119º	9.7

- Based upon the above table, it is reasonable to constrain a residue to have a helical conformation if J<5 Hz and an extended conformation if J>7 Hz.
- If the residue is fluctuating between multiple conformations J will be time-averaged.

6. Dipolar Couplings

- In addition to interacting via bonds (J-coupling) any two nuclei influence each other through space via dipolar couplings.

- The dipolar coupling between two nuclei is sensitive to both the distance between the nuclei and the **angle between the permanent magnetic field and the internuclear vector**.

- If a molecule is static (as in a solid), with a specific orientation to the permanent magnetic field, certain internuclear dipolar couplings can be very strong, dramatically complicating the spectrum
- As a molecule tumbles freely in solution, each dipolar coupling averages to zero.
- However, **if a molecule is weakly aligned relative to the magnetic field**, the dipolar couplings are scaled down. The **residual dipolar couplings** are measurable and contain information about the average orientation of the internuclear vector relative to the permanent magnetic field.
- Weak alignment can be induced by adding to the solution some substance that has an intrinsic preference for alignment relative to the permanent magnetic field; the most commonly used are bicelles and phage particles. Occasional collisions of the macromolecule with the aligned material lead to an average weak alignment of the macromolecule.
- The residual dipolar couplings provide constraints about the angle between the internuclear vector and the alignment tensor (axis) of the macromolecule. Thus they provide **long-range angular information**.

7. <u>The Nuclear Overhauser Effect: Through-Space Distances</u>

- As noted above, as a molecule tumbles freely in solution, the dipolar coupling oscillates, averaging to zero.
- However, the oscillation of this coupling influences the rate at which nuclei relax from excited states back to the ground state.
- This "cross-relaxation" effect is called the nuclear Overhauser effect (NOE). It is easily measurable and is exquisitely sensitive to the internuclear distance (r).
  $NOE \propto 1/r^6$
- NOEs between protons are usually measurable only if $r < 5$ Å.
- Most of the restraints used in NMR structure calculations are typically NOE-derived distance restraints.
- NOEs may be classified as, for example:
  - Strong (< 2.5 Å)
  - Medium (< 3.8 Å)
  - Weak (< 5 Å)

8. <u>Protection from Hydrogen-Exchange: Evidence for Hydrogen-Bonding</u>

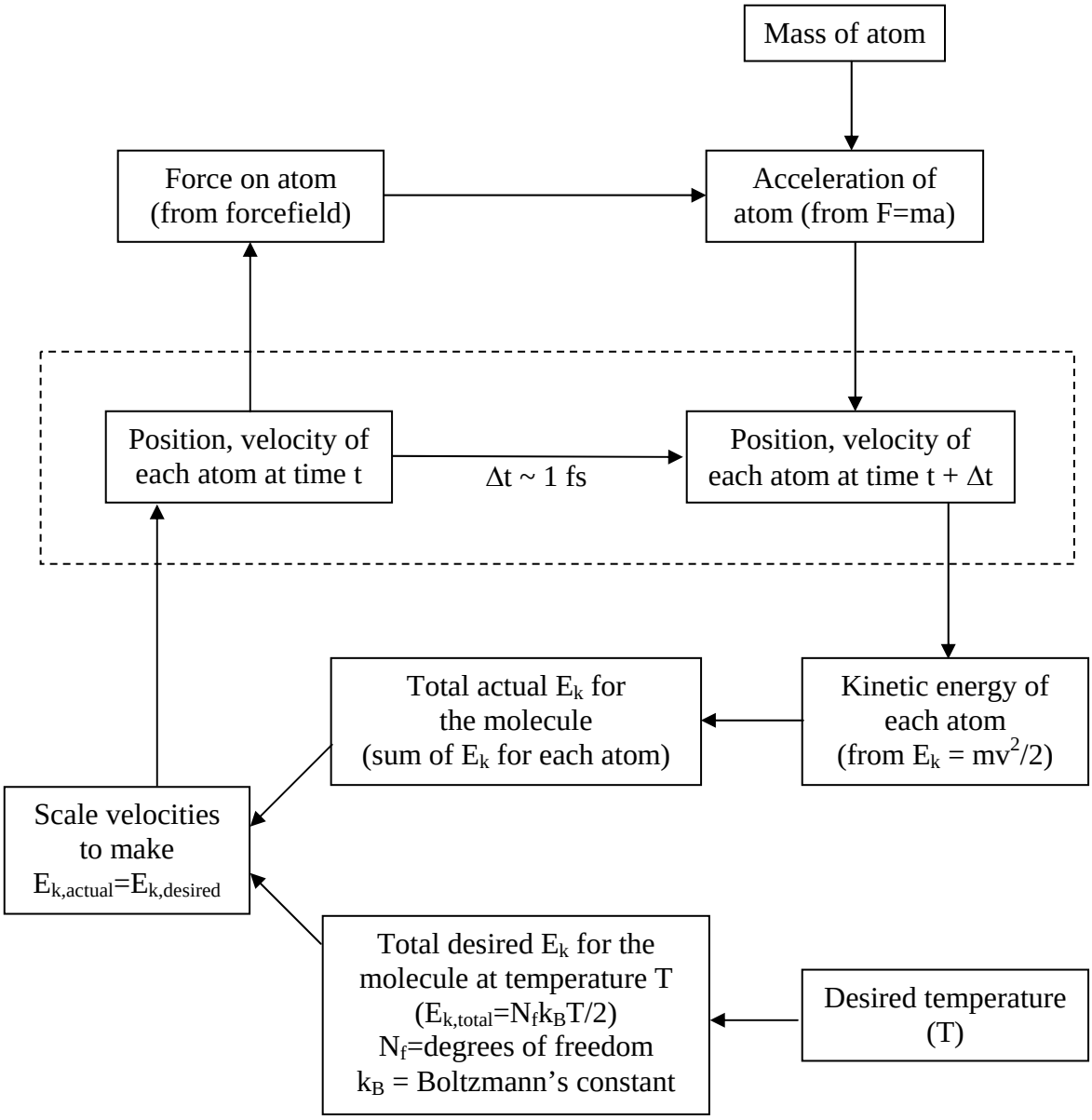
- When a protein is expressed and purified in protic solvents (water-based buffers) the backbone amides are all protonated (i.e. NH)
- When the protein is then quickly transferred to D_2O-based buffer, the NH groups begin to exchange to ND.
- Under typical NMR conditions, exchange occurs within seconds for exposed NH groups.

- However, exchange may be as slow as hours or days if the NH groups are protected in hydrogen bonds or buried within the protein structure.
- Typically, an H-D exchange data set is collected and the amide protons remaining after a few hours are identified. In early stages of structure calculations, it becomes apparent what the hydrogen-bonding partners of these groups are. The hydrogen bonds are then input as restraints for subsequent rounds of calculations.

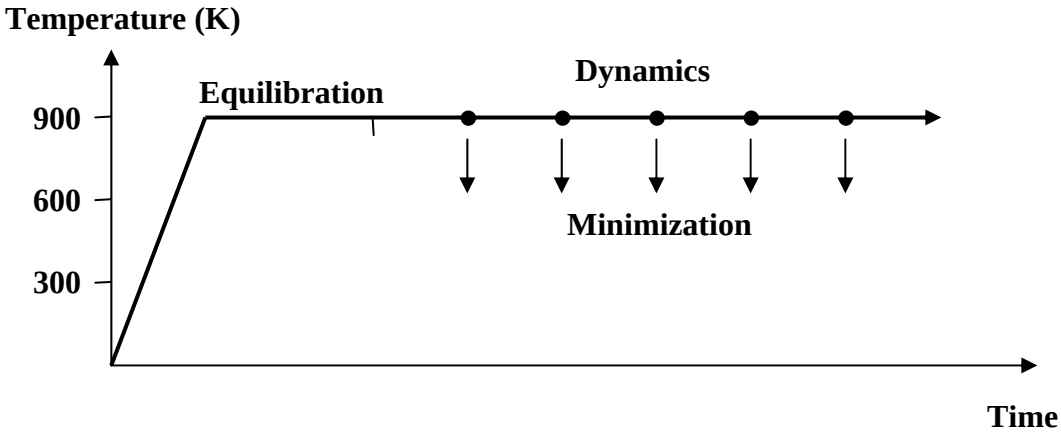
9. Structure Calculations

- 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.

Molecular Dynamics Simulation



(A) MD and EM



(B) Simulated Annealing