

Protein Structure Amino Acid Sequences & Attributes

BSC 5936: Introduction to BioInformatics

© Michael S. Chapman, FSU, 1994-03

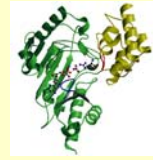
3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

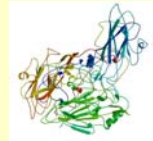
1

Tertiary Structure

- From 1-D information to...
 - Structure
 - Function
- DNA → Protein
- Same building blocks
 - 20 amino acids
 - Secondary structures
- Many different shapes & functions
 - Encoded by different sequences
 - Governed by forces between atoms



Arginine kinase, enzyme



Rhinovirus capsid proteins

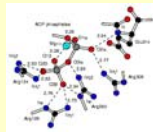
3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

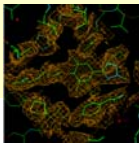
2

Interatomic Forces

- Tertiary structure
- Protein interactions
 - Substrates
 - Drugs
 - Assembly
 - Recognition
- Dynamics

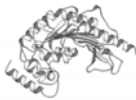


Arginine kinase, active site



Rhinovirus, drug-binding

Arginine kinase, substrate-induced conformational change



3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

Next few lectures

- Basics of Protein 3-D structure
 - Simulation of function (Dr. Zhou)
- Transition from 1-D information → structure / function
 - How does sequence → unique properties?
 - Can we predict structure / function from genomic sequence?

3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

4

Covalent Structure of Proteins

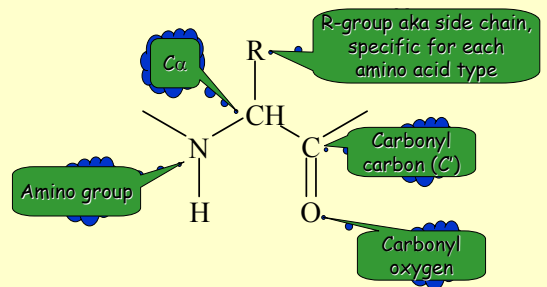
3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

5

Amino acid nomenclature

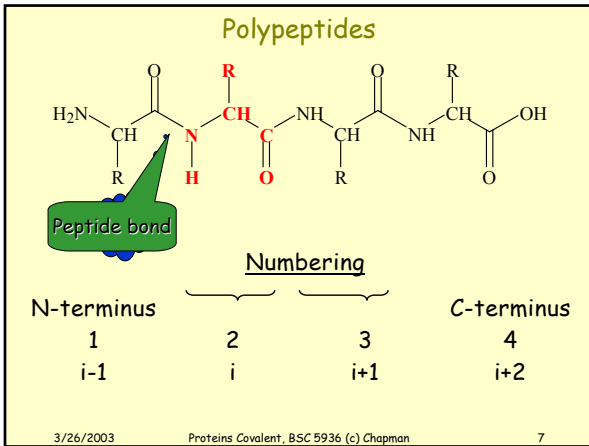
- What do we call the pieces?



3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

6



Geometry - bond lengths (covalent)

- X-H ~ 1 Å
- all others ~ 1.5 Å (approximately),
 - e.g. C-O, C-N.
 - don't worry about the differences
 - Stretching and compression of bonds is negligible.
- Bond angles are also...
 - Predictable, based on atoms
 - Don't change too much
 - ... not a high priority for us...

3/26/2003 Proteins Covalent, BSC 5936 (c) Chapman 8

Geometry - bond angles

- Flexibility small: ~ 2.5°.
- Geometry predictable according to valence.
- Where do you find examples of sp^3 , sp^2 hybridization?
 - sp^2 for 3 bonds; sp^3 for 4 bonds...

Atom	Valence	Hybridization	Coordination	Bond angle
Nitrogen	3	sp^2	Trigonal-planar	120°
Carbon - C_α	4	sp^3	tetrahedral	109°
Carbon carbonyl	4	sp^2	Trigonal-planar	120°

3/26/2003 Proteins Covalent, BSC 5936 (c) Chapman 9

Torsion Angles are the key (Dihedral angles)

- ~ "Free" rotation about single covalent bonds
- ~ Only conformational variables
 - Determinant of structure
 - As bond angles, lengths are ~ constant.

3/26/2003 Proteins Covalent, BSC 5936 (c) Chapman 10

Variation in Torsion Angles

- Determined by van der Waals forces
- Staggered conformers (often 3)
 - E.g. 60°, 180°, 300° (-60°).
- Variation within each conformer:
 - Deviations from "ideal" ± ~20°
- allows protein
 - to be "distorted" from "perfect" structure
 - to move dynamically

3/26/2003 Proteins Covalent, BSC 5936 (c) Chapman 11

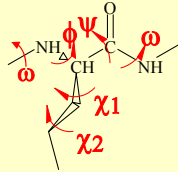
Enumerating torsion angles

- Defined by 4 atoms surrounding the bond
- Conventions on which atoms to choose, directions of the angles etc..

3/26/2003 Proteins Covalent, BSC 5936 (c) Chapman 12

Names of Protein Torsion Angles

- Side chains:
 - χ_1, χ_2 for bonds farther down the side chain.
- Backbone:
 - ω, ϕ, ψ , in αλπηβετιχλ order for bonds before N, C_α & C'.



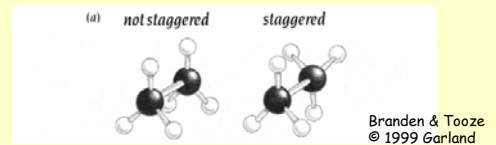
3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

13

Torsion Angle Values - Principles

- What values are most favored?
 - Those corresponding to staggered angles



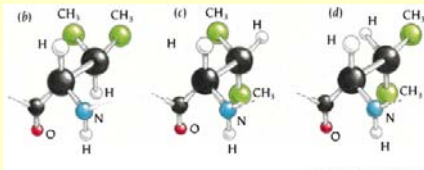
- ϕ, ψ are perhaps the most important
- χ are perhaps easier to understand
 - We will start with χ .

3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

14

χ_1 - Valine - a Simple Example



- 3 staggered rotamers, possible & actually seen
 - Not equally probable
- Rotamers c & d have CH₃ between (large) backbone atoms
- Rotamer b is preferred, because CH₃ next to H_α.
 - (all other things being equal)

3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

15

Concept of a Rotamer

- Side chains have several $\chi_1, \chi_2, \chi_3 \dots$
- Each χ approximates a staggered conformation
- 3 or 6 discrete options for each χ
- Allowed rotamers =
 - Product of these options
 - Minus those with steric conflicts
 - (Atoms that bump into one another)
 - 6 to 12 for each type of side chain
- 3 or 4 of these are favored rotamers
 - E.g. His (χ_1, χ_2): (180°, 90°), (60°, 90°), (60°, -90°).
 - Rotamer libraries

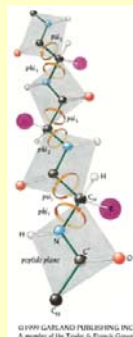
3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

16

Backbone Torsion Angles - Importance

- Describe overall fold
 - Almost completely
- Because...
 - Bond lengths ~fixed
 - Bond angles ~fixed
 - Only torsion angles variable
 - Only 2 variables / amino acid
 - (for the backbone)



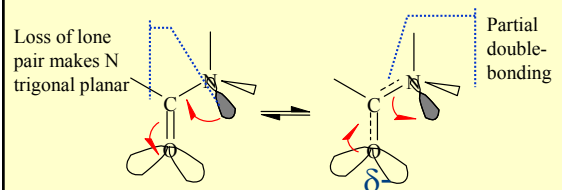
3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

17

3 Torsion Angles: ω, ϕ, ψ , but ω is fixed

- Consider a peptide bond...
- Lone pair electrons → partial double C—N bond
 - restricts rotation to $\pm 6^\circ$ about planar
 - Usually trans ($\sim 180^\circ$)



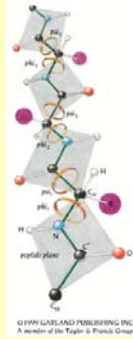
3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

18

ϕ, ψ Describe the overall fold

- If ϕ, ψ were listed for every amino acid
 - Topology defined
- Ramachandran considered what combinations of ϕ, ψ were favorable for each amino acid.
 - Peptides: $(Ala)_n, (Gly)_n, \dots$
 - Only van der Waals forces
 - Plotted areas where E favorable
 - As a 2-D function of ϕ, ψ .



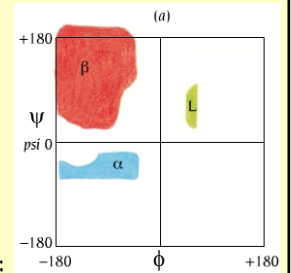
3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

19

Ramachandran Plot

- Shaded regions are favorable combinations of ϕ, ψ for an individual amino acid.
 - Backbone staggered
 - Large groups well separated
- If same values repeated:
 - Regular (secondary) structure
 - Naming of energy minima



Branden & Tooze
© 1999 Garland

Torsion angles on
kinemage 1

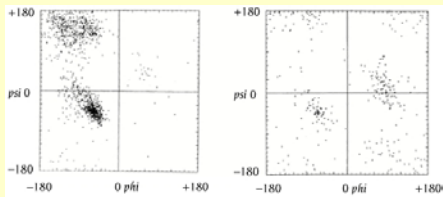
3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

20

Peptide Conformation

- Observed ϕ, ψ values for each amino acid of a protein always fall near the calculated energy minima
 - Well, nearly always
 - Non-glycine
 - Glycine



3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

21

Levinthal's paradox & dictionaries

- How many backbone conformations of a 300 residue protein are possible? Assume:
 1. Only ϕ, ψ important.
 2. ϕ, ψ need only be given $\pm 15^\circ$
 - i.e. sampled every 30° .
 3. Consider only minima of Ramachandran plot.
- Still **Approx. 10^{30} conformations!**
- Levinthal paradox(es):
 - How is the right conformation found?
 - Why are there only ~5,000 protein folds?

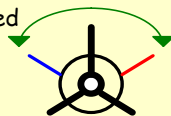
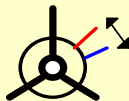
3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

22

Torsion angles and Dynamics

- Deformation within a minimum:
 - $\pm 15^\circ$ is easy; rapid vibrations.
- Jumping between minima:
 - Possible if does not cause steric overlap of other atoms.
 - Slow, as activation energy barrier passing through eclipsed conformation..



3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

23

Reading.

- Further reading:
 - Branden, C.-I. & Tooze: Introduction to Protein Structure, Chapter 1

3/26/2003

Proteins Covalent, BSC 5936 (c) Chapman

24