

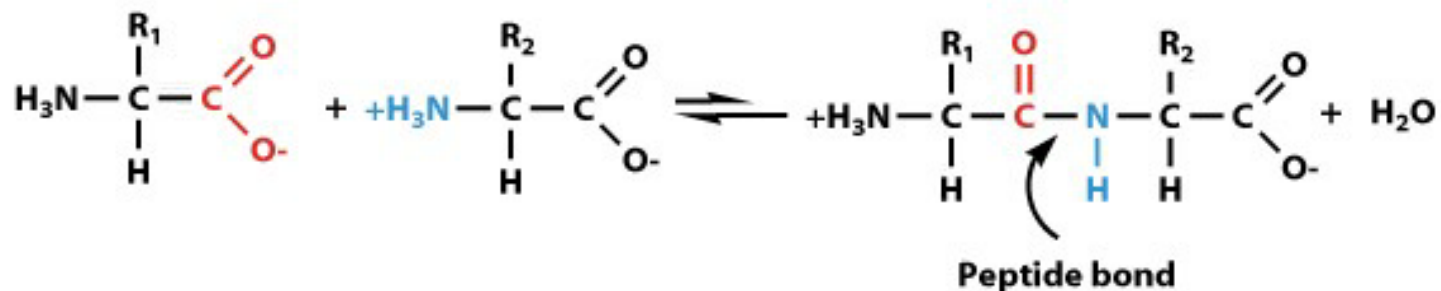
# From amino acids to proteins

Peptide bond

Energetics of folding

Backbone conformation

# The peptide bond



- Forward reaction is a dehydration, backward reaction is an hydrolysis
- Note energetics
- Two amino acids combine to form a dipeptide, three to form a tripeptide, several to form an oligopeptide, many to form a protein

# Protein sequence

- >sp|P54209|ATC1\_DUNBI CATION-TRANSPORTING ATPASE CA1 (EC 3.6.1.-) - Dunaliella bioculata.
- MVSHASSGRPSSRDTGMVYLGLGMQDAYSSSEVQEVAAFYHVDLDRGLSDRDVQQARIKYGRNQMEAE  
QSTPLWKLILKQFDDLLVKILLGAAIVDFIIAISEGESIQSGLIEPMVILLILVANATVGVVTERNAEKAIEQLKS  
YEADDATVLRNGQLQLIPSADIVPGDIVELAVGNKVPADTRVSHIYTTSLKIDQSLTGESQAVEKHTEVVH  
NEQAVYQDKLNMLFSGTLVVAGRARGIVVGTGSNTAIGKIRDAMGVEEDVVTPLKAKLDEFGALLSKVIA  
GICVLVWVWNINRFNDPALGGWFQGAIHYFKIAVALAVAAIPEGLPAVVTTCLALGTRKMARHNAIVRTL  
SVETLGCTTVICSDKTGTLTTNQMSVIKVAAVQSSSSQLAEFDVTGTTFSPEGMVLGPGGVVLRQPADTP  
CLAHAAQCAALCNDSSQVFVAQKTGTLQRIGESTEIALRVFAEKIGLPSSIRPDRPISRSQFGTNNFWQEDV  
ERLALLEFSRDRKMMSVLVKGSDRQHNIWSKGAPEFVLRKCSHVLANNGEAVPLTDNMRQAILSDMQ  
AFGSRQALRCLALAFKSVPTTTTKLDYSDESGLTFIGLLGMHDPPECRSALSTCHNAGIKVIMVTGDNK  
GTAEAVARQVGALSPSTALAGSDDEDNLGISYTGREFEEMGALGQAAATRNLVVLSRVEPMHKLRLVEL  
LKAQGHVVAMTGDGVNDAPALLRADIGIAMGSGTAVAKHAADMVLGDDNFATIVFAVAEGRVIFNNTKQF  
IRYMISSNIGEVAIFLAALLGLPEVLTPVQLLWVNLVTDGLPATALGFNRADKDMMARGPRRVDDPIVNG  
WLFLRYLIIGMYVGIVTVYGFIWWYISFPEGGNMTWSQLTHFQACASQPGGAKDCEVFHSHKPTTISMSV  
LVVVEMFNALNNLSEDSSLLRIPPWDNKWLVGAIATSMALHFGILYTGASAMFGVTGLSFAEWTMVIKLSA  
PVILVDEIMKAWSRRRQRHPASSRGGPVSLMEIQVPLTSSSRDEAALKLK
- FASTA format for presentation of amino acid sequence of a protein
- Different proteins with different sequence, related proteins with related sequence
- No obvious pattern

# Protein sequences fold

- Such sequences may fold to an organized structure
  - Not necessarily a simple linear collapse
    - [http://www.youtube.com/watch?v=swEc\\_sUVz5I](http://www.youtube.com/watch?v=swEc_sUVz5I)
  - Folding occurs by formation of intermediate domains
    - <http://www.youtube.com/watch?v=AlfvWESPyZY&feature=related>
- What structure actually forms depends on sequence
- Driving force for collapse = hydrophobicity
- Controlling force for specific structure = stronger interactions, or loss thereof

# Getting along with your neighbors

Bond	Energy	Example
Covalent bonds	100-600 KJ/mole	-S-S-
Ionic bonds	15 KJ/mole	COO <sup>-</sup> - NH <sub>3</sub> <sup>+</sup>
Hydrogen bonds	5 KJ/mole	-O-H ···· O=C-
Van der Waals	4 KJ/mole	-C-H ... H-C-

- Amino acids capable of all of these interactions
- Covalent bonds: too strong to break or rearrange, so all folding depends on weak interactions
- BUT – the energies of weak interactions enter in, not as positive contributors, but as costs...
- ...incurred in transition from unfolded to folded forms

# Collapse to a folded structure

## Energetics

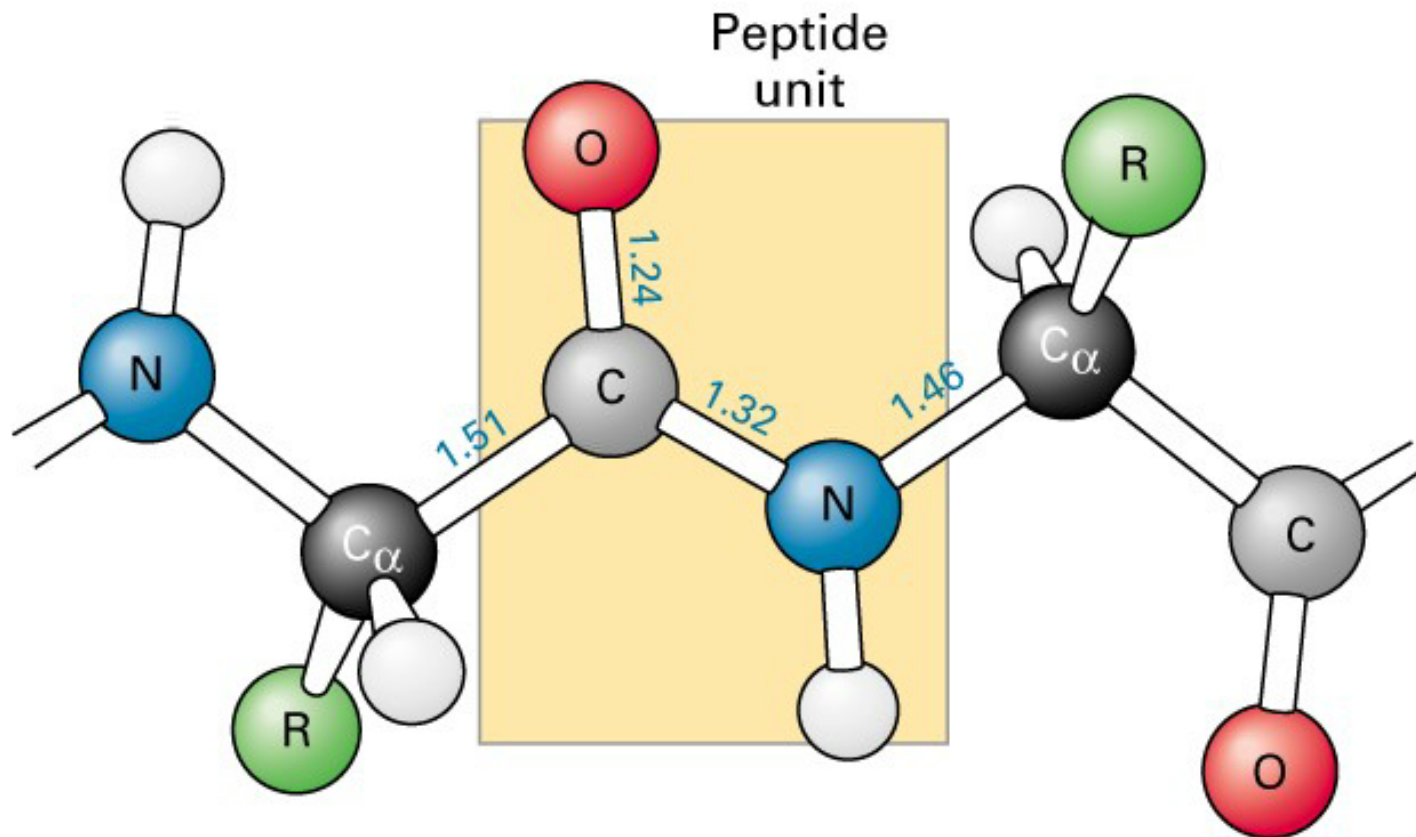
- Van der Waals interactions always available
- VdW sufficient to replace hydrogen bonds
- In water, hydrocarbons replace hydrogen bonds between water molecules with van der Waals interactions – a small, but cumulative cost.
- Clustering hydrocarbons together retains their van der Waals interactions, and regenerates water H-bonds

# Ending up at a specific structure

## Avoiding energy costs

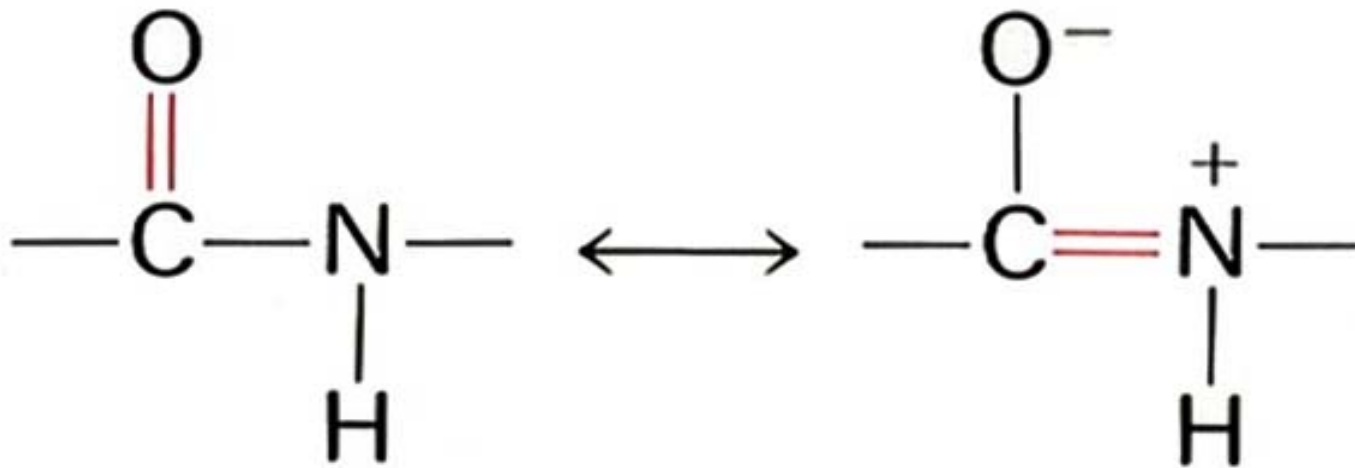
- Ionic interactions/hydrogen bonds:  
abundantly available in unfolded state  
(with water and dissolved salts)
  - Little gain from forming those interactions in folded state
  - Very high cost if folded state does not permit these interactions

# The Peptide Bond



Six atoms, (C<sub>α</sub><sub>1</sub>, C, O, N, H, C<sub>α</sub><sub>2</sub>) all lie in a plane.

## The reason for the planarity



The peptide bond is rigid and planar because

Atoms involved in double bonds distribute the bonds in a plane

The double bond character of the C puts C, O, and N in a plane

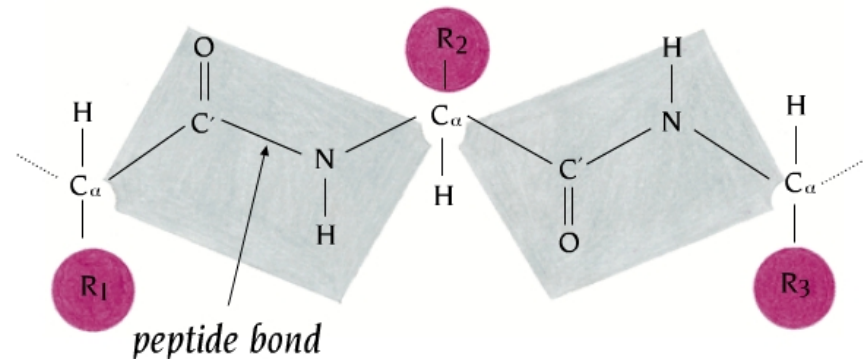
The double bond character of the N puts C, N, H, and Ca2 in (the same) plane

# Details

- Barrier to Rotation about C-N bond is 60 kJ/mole
- Peptide bond is planar with C $\alpha$  groups typically trans to the peptide bond (better accommodate R groups)
- C $\alpha$  groups are still sp<sup>3</sup> hybridized, still have 109.5 bond angles and can rotate around the single bond linking them to the atoms of the peptide bond
- The latter rotations allow side chains to move in three dimensions
- Not all conformations are possible

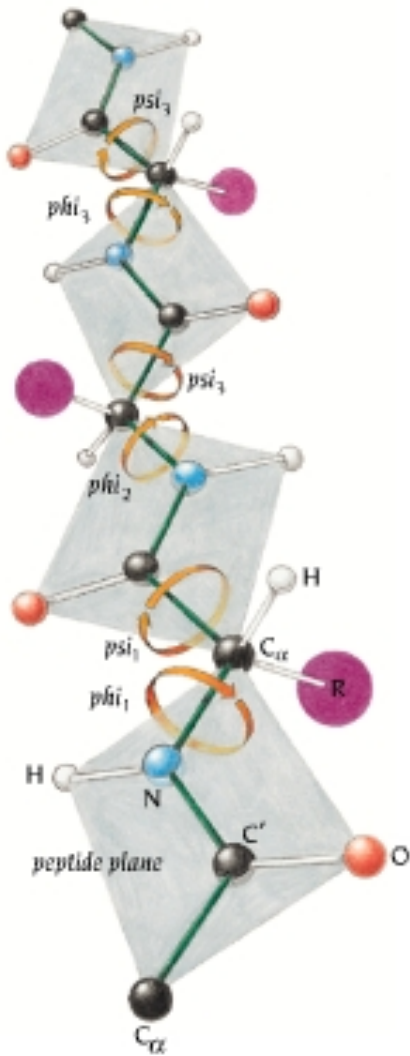
# A different view of the protein backbone

- Succession of units extending from one C $\alpha$  carbon to the next
- Each unit consists of atoms in a single plane
- Each unit can rotate around the bond connecting it to the C $\alpha$  carbon
- The two rotational angles at each carbon are, in principle, independent.
- In fact, they are not independent



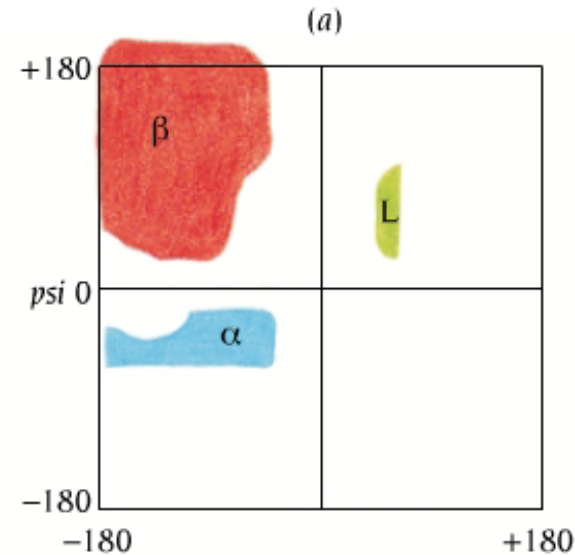
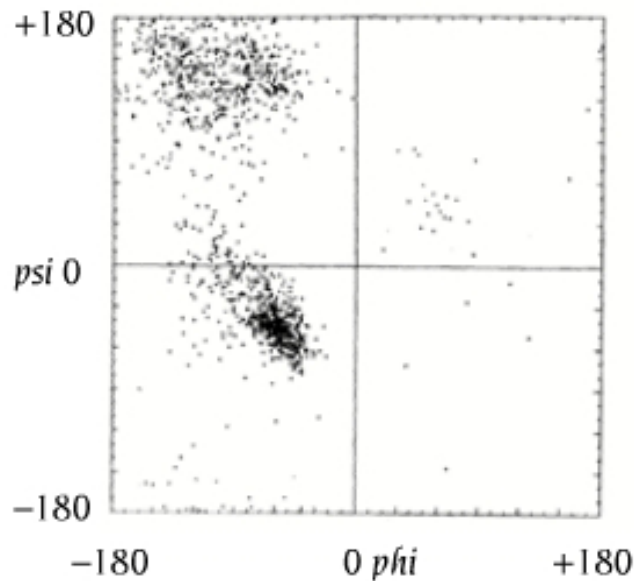
©1999 GARLAND PUBLISHING INC.  
A member of the Taylor & Francis Group

# Phi and Psi angles



- Note in this diagram which is N-terminal direction, which is C-terminal
- Phi angle is with the phormer amino acid (n-1)
- Psi angle is with the Psubsequent amino acid (n+1)
- For each C $\alpha$  carbon, can measure and plot phi, psi
- = Ramachandran plot

## Measured, theoretical phi, psi angles



- Relatively restricted range of possibilities because of steric interactions between side chain and backbone atoms

# And those are sweet spots...

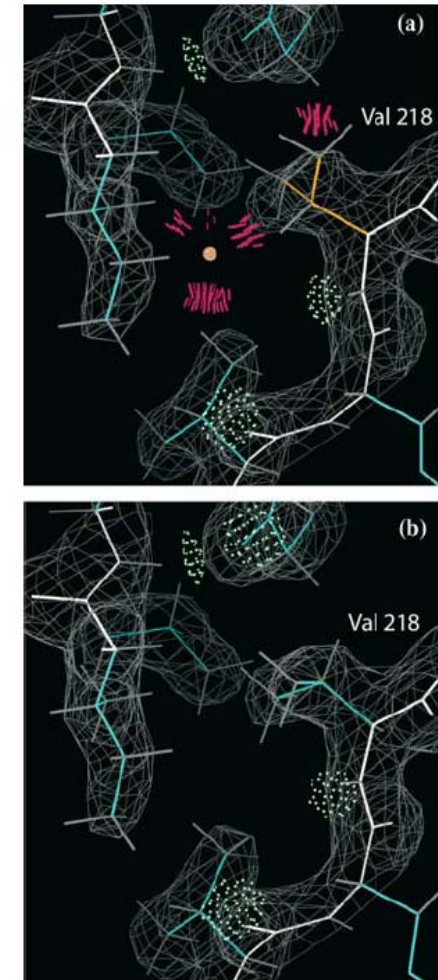
*Journal of Structural and Functional Genomics* 6: 1–11, 2005.  
DOI 10.1007/s10969-005-3138-4

© Springer 2005

## A test of enhancing model accuracy in high-throughput crystallography

W. Bryan Arendall III<sup>1</sup>, Wolfram Tempel<sup>2</sup>, Jane S. Richardson<sup>1</sup>, Weihong Zhou<sup>2</sup>, Shuren Wang<sup>1,4</sup>, Ian W. Davis<sup>1</sup>, Zhi-Jie Liu<sup>2</sup>, John P. Rose<sup>2</sup>, W. Michael Carson<sup>3</sup>, Ming Luo<sup>3</sup>, David C. Richardson<sup>1,\*</sup> & Bi-Cheng Wang<sup>2,\*</sup>

- Look in detail at amino acid configurations
  - Amine/Amide orientations in Asn/Gln/His
  - Similar considerations for vals (see fig ->)
  - Ramachandran angles
- Resulting adjustments
  - 99% of amino acid side chains in relaxed (trans) configuration
  - >98% Ramachandran angles favorable



2. An example of using these methods to refit valine sidechain and an incorrect 'water' (centre)