

# Components of Protein Structures

BME 110: Computational Biology Tools

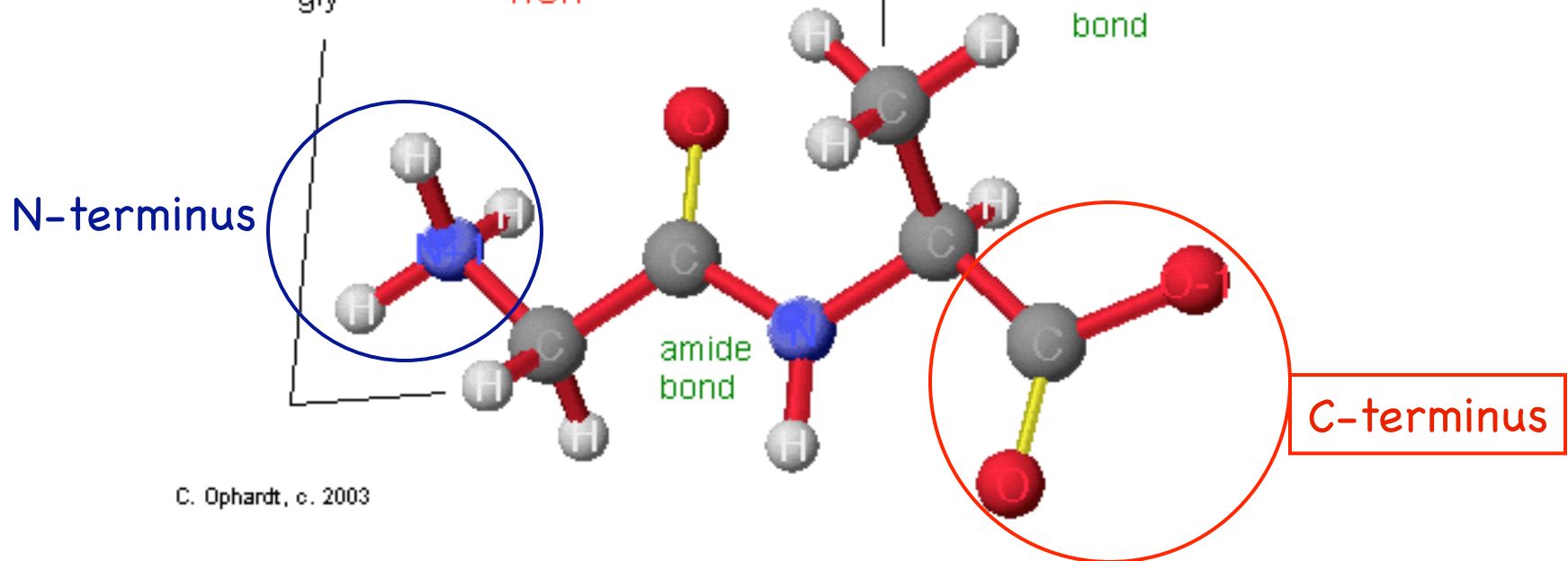
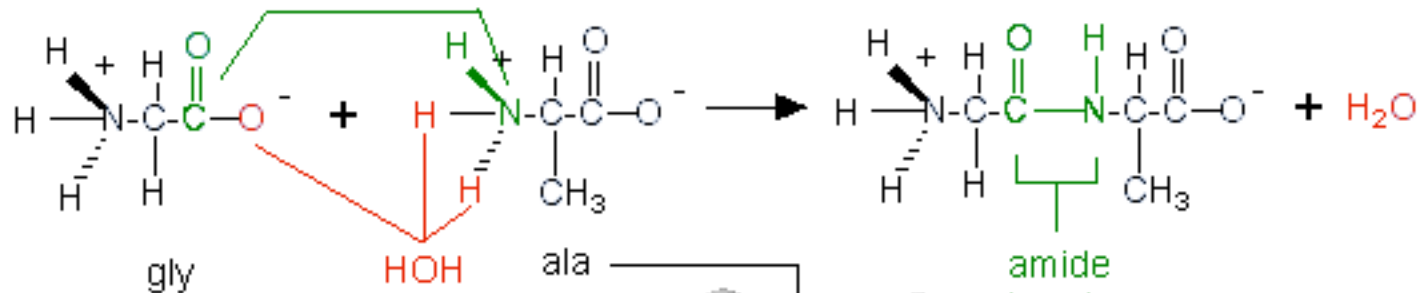
# Amino acids -- properties and symbols

Amino acid				Amino acid			
Alanine	A	Ala	Neutral Non-polar	Methionine	M	Met	Neutral Non-polar
Cysteine	C	Cys	Neutral Slightly Polar	Asparagine	N	Asn	Neutral Polar
Aspartate	D	Asp	Acidic Polar	Proline	P	Pro	Neutral Non-polar
Glutamate	E	Glu	Acidic Polar	Glutamine	Q	Gln	Neutral Polar
Phenylalanine	F	Phe	Neutral Non-polar	Arganine	R	Arg	Basic Polar
Glycine	G	Gly	Neutral Non-polar	Serine	S	Ser	Neutral Polar
Histidine	H	His	Basic Polar	Threonine	T	Thr	Neutral Polar
Isoleucine	I	Ile	Neutral Non-polar	Valine	V	Val	Neutral Non-polar
Lysine	K	Lys	Basic Polar	Tryptophan	W	Trp	Neutral Slightly polar
Leucine	L	Leu	Neutral Non-polar	Tyrosine	Y	Tyr	Neutral Polar

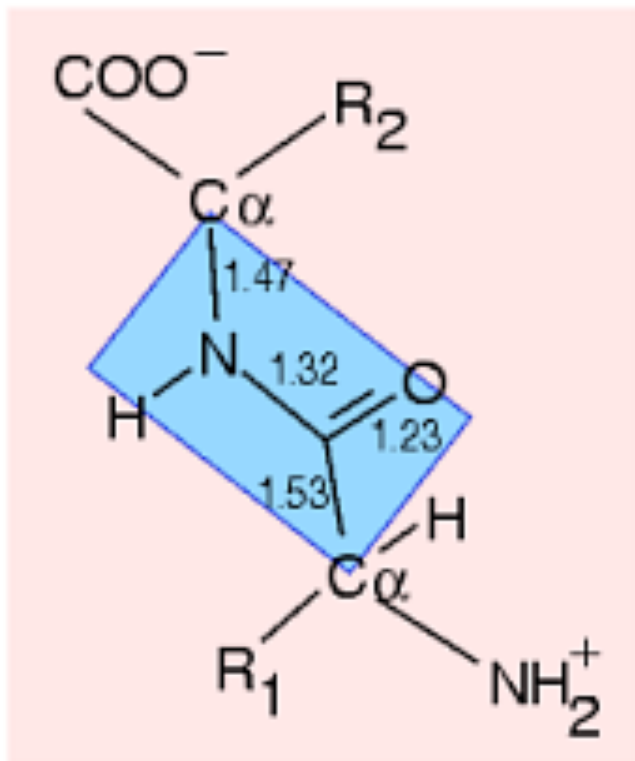


# Peptides and the peptide bond

## Peptide or Amide Synthesis



# peptide bond distances



$|x-H| \sim 1.05 \text{ \AA}$   
 $|N-C\alpha| \sim 1.45 \text{ \AA}$   
 $|N-C| \sim 1.37 \text{ \AA}$   
 $|C-O| \sim 1.23 \text{ \AA}$   
 $|C-C\alpha| \sim 1.49 \text{ \AA}$

from Pauling, L. 1951

# primary structure -- 1TIM

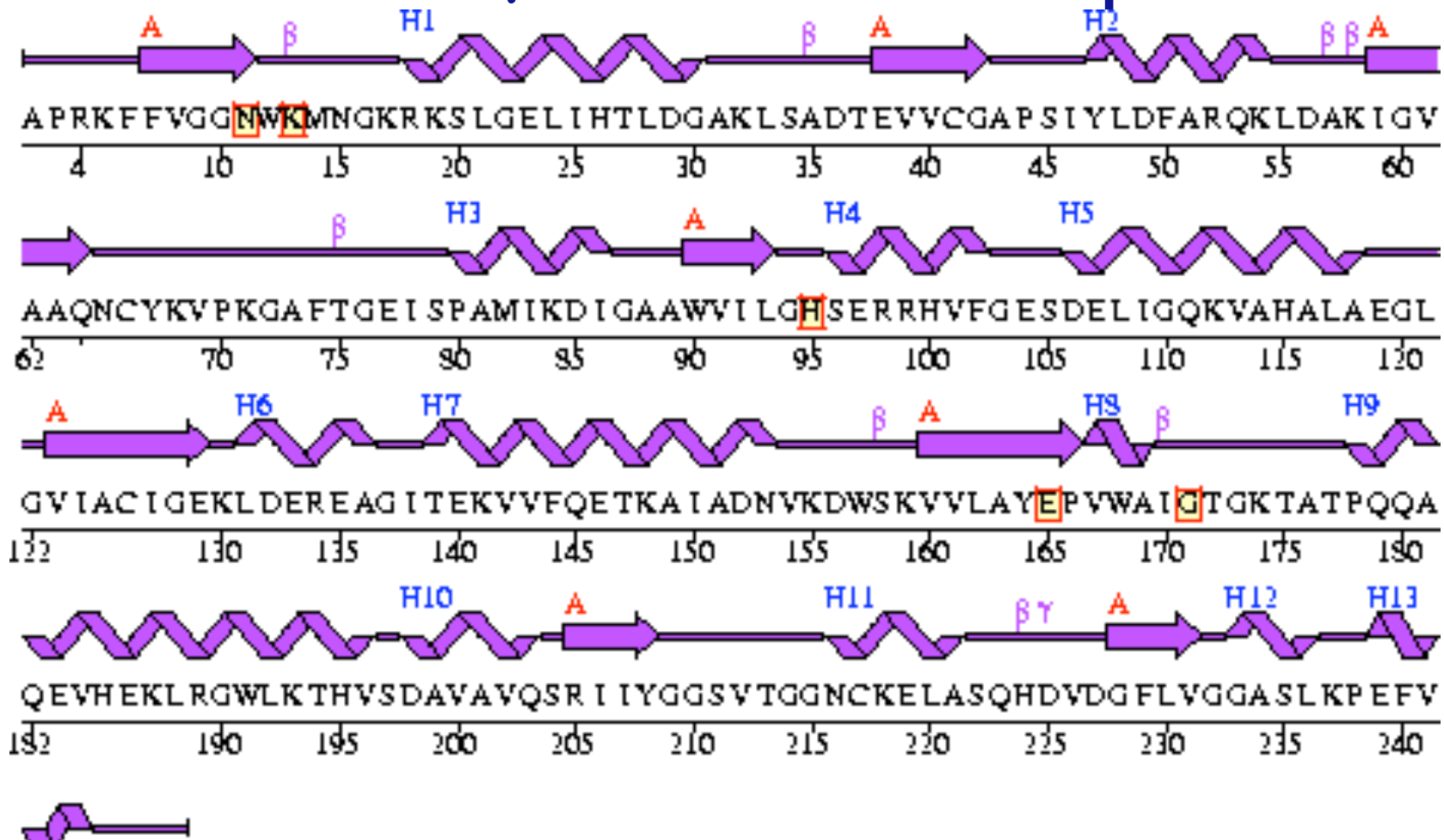
- primary -- sequence

>1TIM:A|PDBID|CHAIN|SEQUENCE

APRKFFVGGNWKMNGKRKSLGELIHTLDGAKLSADTEVVCGAPSIYLDFARQKLDK  
IGVAAQNCYKVPKGAFTEGEISPAMIKDIGAAWVILGHSERRHVFGESDELIGQKVAH  
ALAEGLGVIACIGEKLDEREAGITEKVVFQETKAIADNVKDWSKVVLAYEPVWAIGT  
GKTATPQQAQEVHEKLRGWLKTHVSDAVAVQSRIIYGGSVTGGNCKELASQHDVDGF  
LVGGASLKPEFVDIINAKH

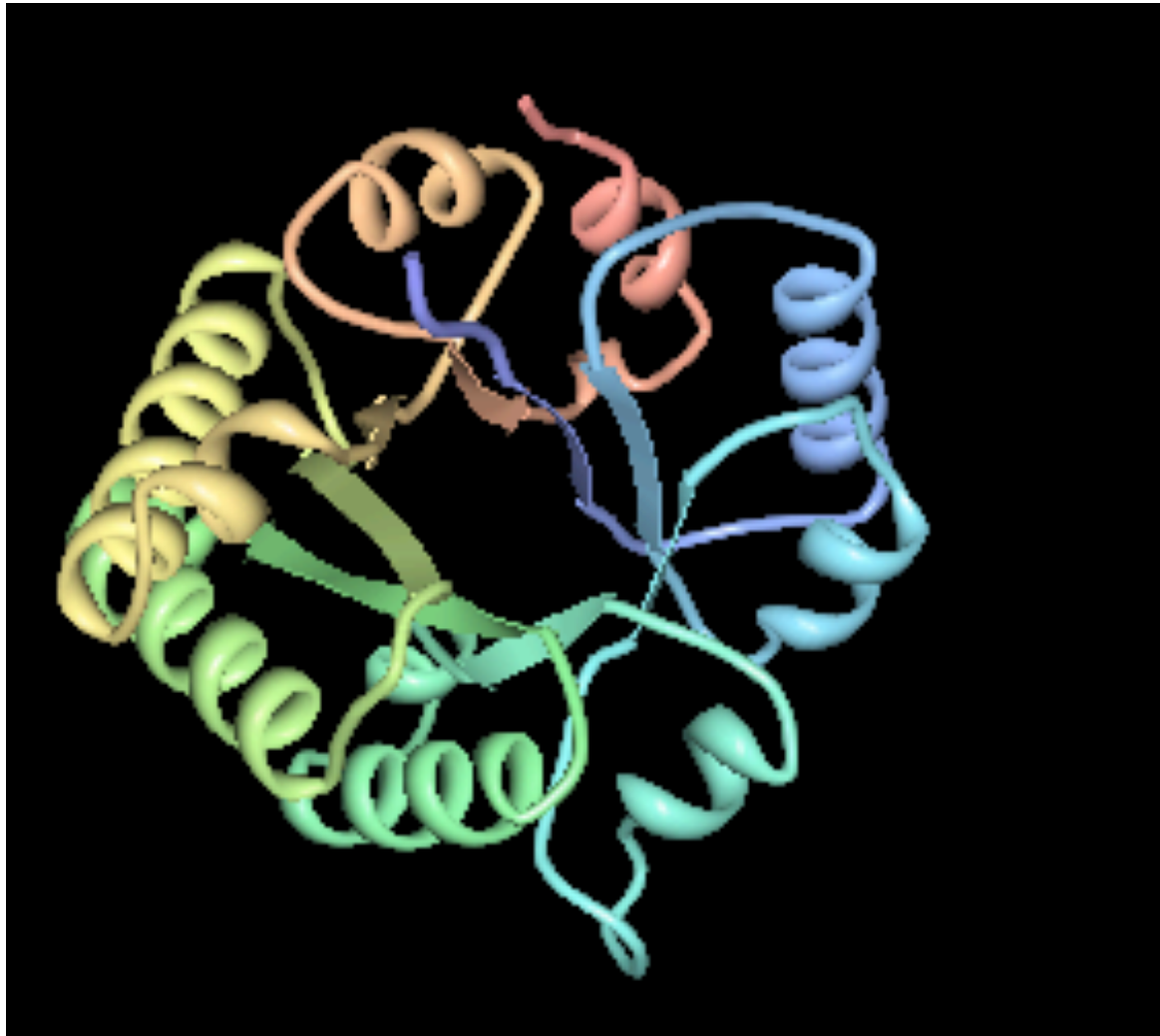
# secondary structure - 1TIM

## helix, strand or loop



<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum>

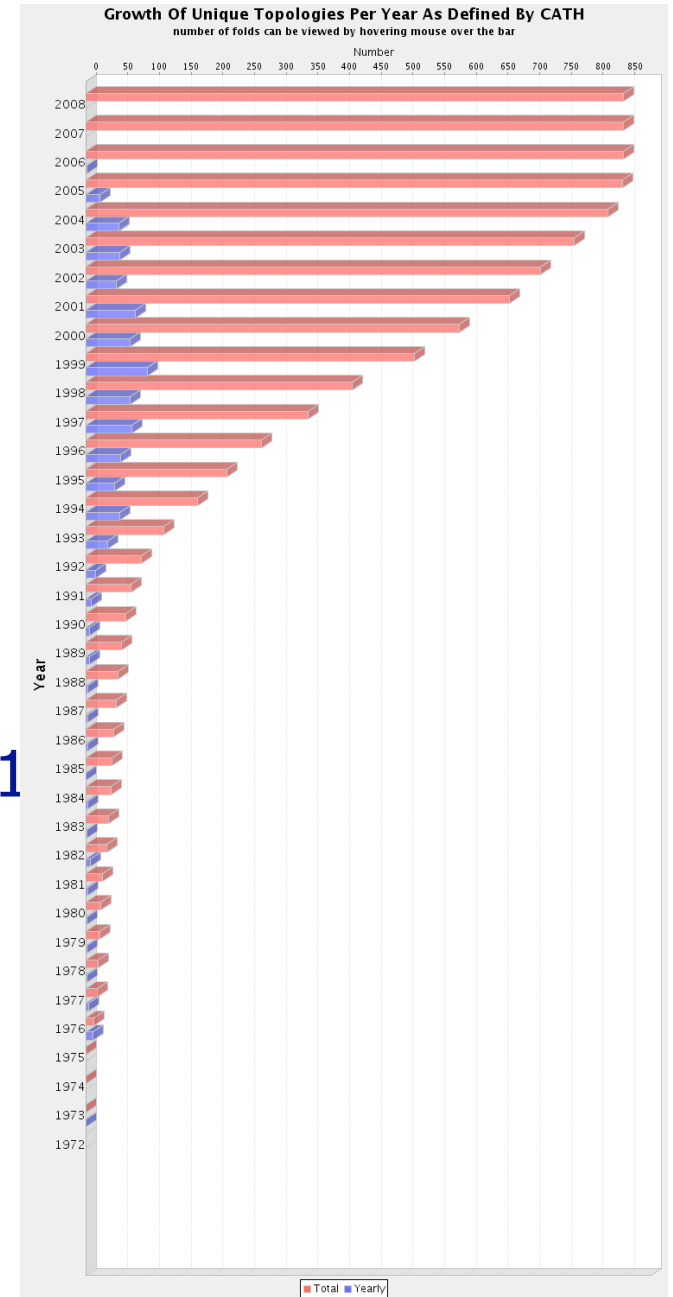
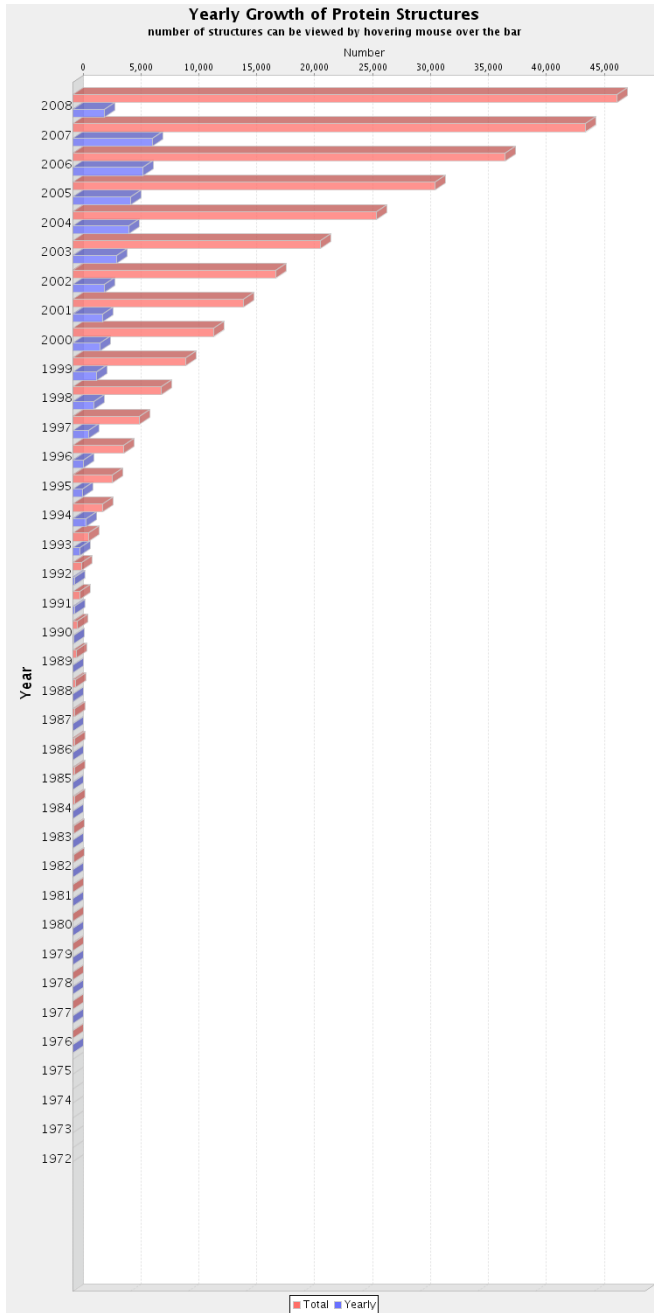
# tertiary structure -- 1TIM



# Protein Data Bank

[www.pdb.org](http://www.pdb.org)

- as of 5/20/2008, there were 50961 stored structures
- with 1056 unique folds(SCOP)



# structures

<http://www.pdb.org/pdb/explore.do?structureId=1TIM>



type X-RAY DIFFRACTION

Resolution[Å]  
2.50

R-Value  
n/a

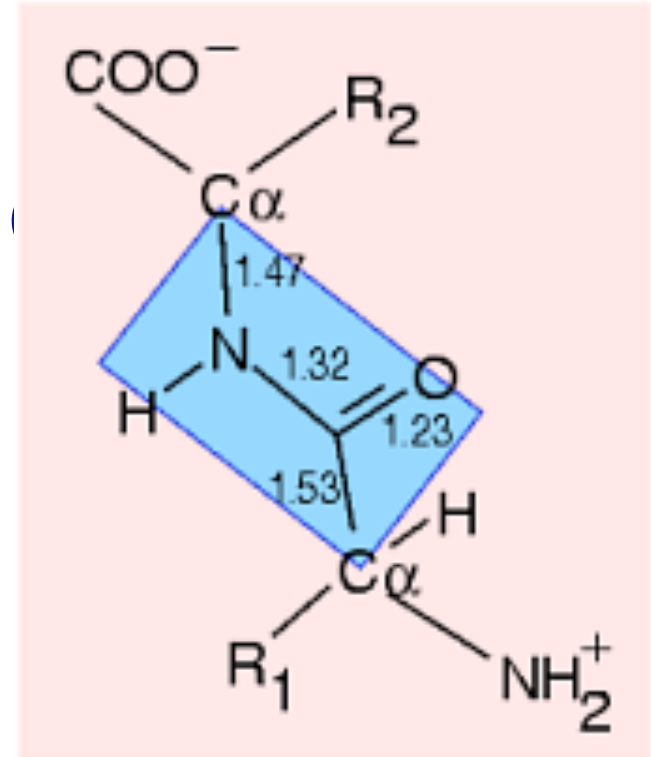
R-Free  
n/a

Space Group  
P 2<sub>1</sub> 2<sub>1</sub> 2<sub>1</sub>

- **Banner, D.W., Bloomer, A., Petsko, G.A., Phillips, D.C., Wilson, I.A.**  
Atomic coordinates for triose phosphate isomerase from chicken muscle.  
*Biochem.Biophys.Res.Commun.*  
v72 pp.146-155 , 1976

# PDB structure records (1TIM)

record	atom	residue	coordinates
ATOM	1 N	ALAA 1	43.240 11.990 -6.915
ATOM	2 CA	ALAA 1	43.888 10.862 -6.231
ATOM	3 C	ALAA 1	44.791 11.378 -5.094
ATOM	4 O	ALAA 1	44.633 10.992 -3.937
ATOM	5 CB	ALAA 1	44.722 10.051 -7.240
ATOM	6 N	PRO A 2	45.714 12.244 -5.497
ATOM	7 CA	PRO A 2	46.689 12.815 -4.561



$$\begin{aligned}
 |C\alpha_{ALA}, N_{ALA}| &= \sqrt{\left(\left|\vec{X}\right|\right)^2 + \left(\left|\vec{Y}\right|\right)^2 + \left(\left|\vec{Z}\right|\right)^2} \\
 &= \sqrt{(43.240 - 43.888)^2 + (11.990 - 10.862)^2 + (-6.915 + 6.231)^2} \\
 &\approx 1.4697
 \end{aligned}$$

# polyproline helices

- polyproline helices occur in nature
- examine 2J6O -- SH3 domain
- examine 2FBJ -- IGA.
  - action-preset-pretty
  - select pro, resn pro
  - show sticks
  - display sequence on

## in class exercise

# Biometrics of 1TIM using PYMOL

- Polar contacts are hydrogen bonds
  - select bb, (name ca,n,o,c)
  - find polar contacts
- The measurement wizard
- show sticks, resi xx
- Helices -
  - What is the Pitch of a few of these  $3.6_{13}$  helices?
  - How many residues are there per turn?
- Strands -
  - These are all parallel strands
  - What is the bonding pattern?