Today's plan -

1) Charges on amino acids & peptides.

2) Continue with protein structure.

K_a and pK_a describe how completely a weak acid dissociates.

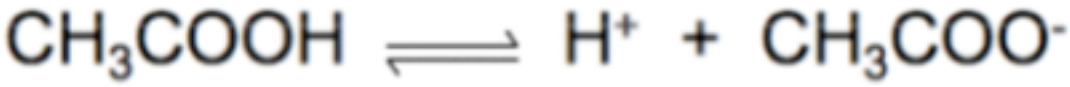
$$HA \rightleftharpoons H^+ + A^-$$

$$K_a = \frac{[H^+]\ [A^-]}{[HA]}$$

$$pK_a = -\log_{10} K_a$$

The pK_a of a weak acid is the pH at which $[HA] = [A^-]$

Example: acetic acid has a pKa of 4.7



So, in a solution of acetic acid at pH 4.7,

CH_3COOH and CH_3COO^- are present in equal amounts.

The Henderson-Hasselbalch equation describes how much of a weak acid is ionized at a particular pH:

$$pH = pKa + \log \frac{[\text{conjugate base}]}{[\text{acid}]}$$

The Henderson-Hasselbalch equation says: A change of one pH unit changes the ratio of acid to conjugate base by a factor of ten.

pH	Ratio $[CH_3COOH]$ / $[CH_3COO^-]$
3.7	[10] / [1]
4.7	[1] / [1]
5.7	[1] / [10]

Some pK_a values that every biochemist should know:

carboxyl group: pKa typically about 4

amine: pKa typically about 10

Draw the amino acid alanine (Ala).

Describe the charges on alanine at pH 7, 2 and 12.

pK_a values for some amino acid side chains:

Asp & Glu	pK_a is about 4
Lysine	pK_a is about 10.5
Arginine	pK_a is about 12
Tyrosine -OH	pK_a is about 10
Cysteine -SH	pK_a is about 8.3
Histidine ring	pK_a is about 6

Describe the charges on a tripeptide with sequence:

Ala-Lys-Cys at pH = 7

At what pH would this tripeptide have a charge of zero?

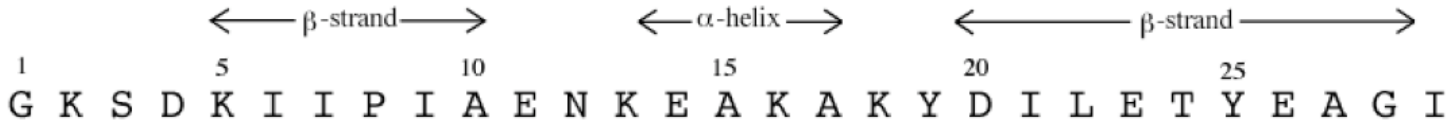
(this is the "isoelectric point" of the peptide)

Protein structure (continued)

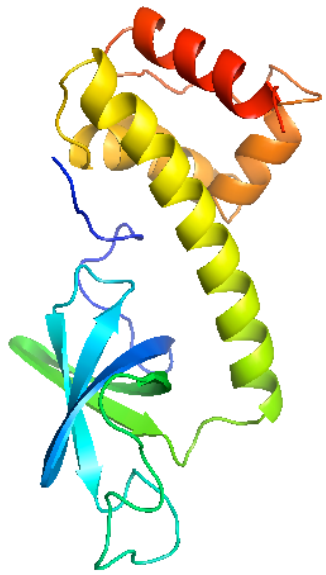
Primary, secondary, tertiary structure of proteins.

Primary structure is just the a.a. sequence.

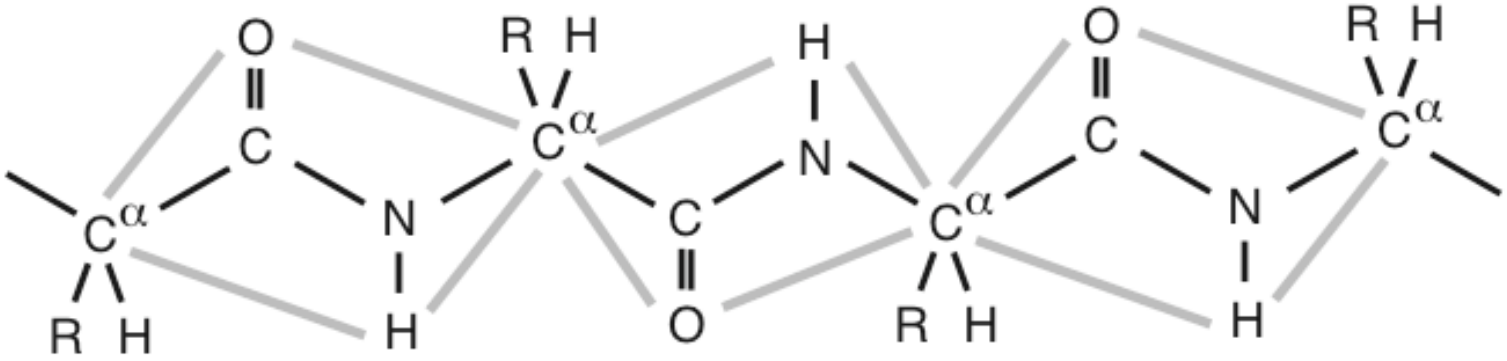
Secondary structure describes which parts of the protein are helices, beta strands, turns.



Tertiary structure describes 3-D fold.



Planar units within peptides are relatively rigid due to partial double bond character of C - N bond.

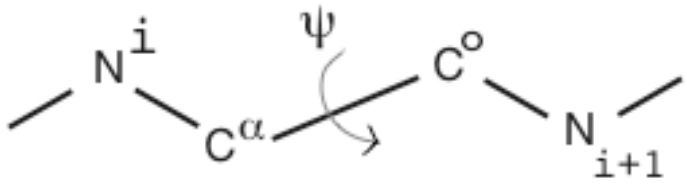


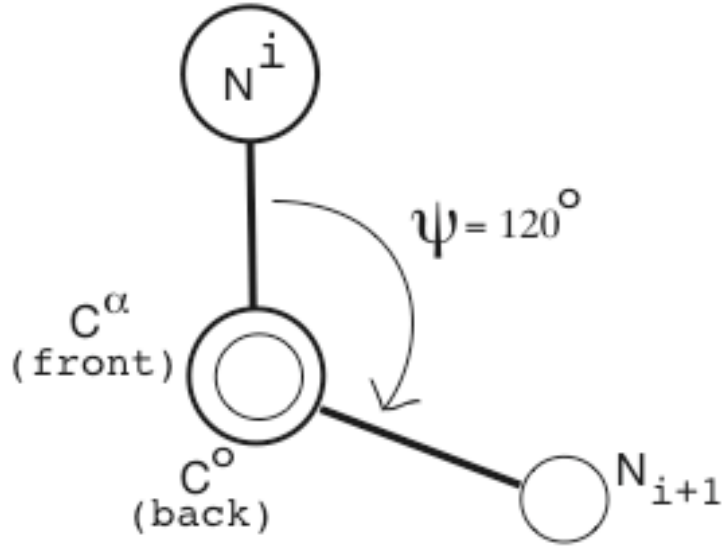
The peptide backbone can rotate around 2 torsion angles.

Torsion angles : psi ψ (rotation around C^{α} - C^{o} bond)
phi ϕ (rotation around N - C^{α} bond)

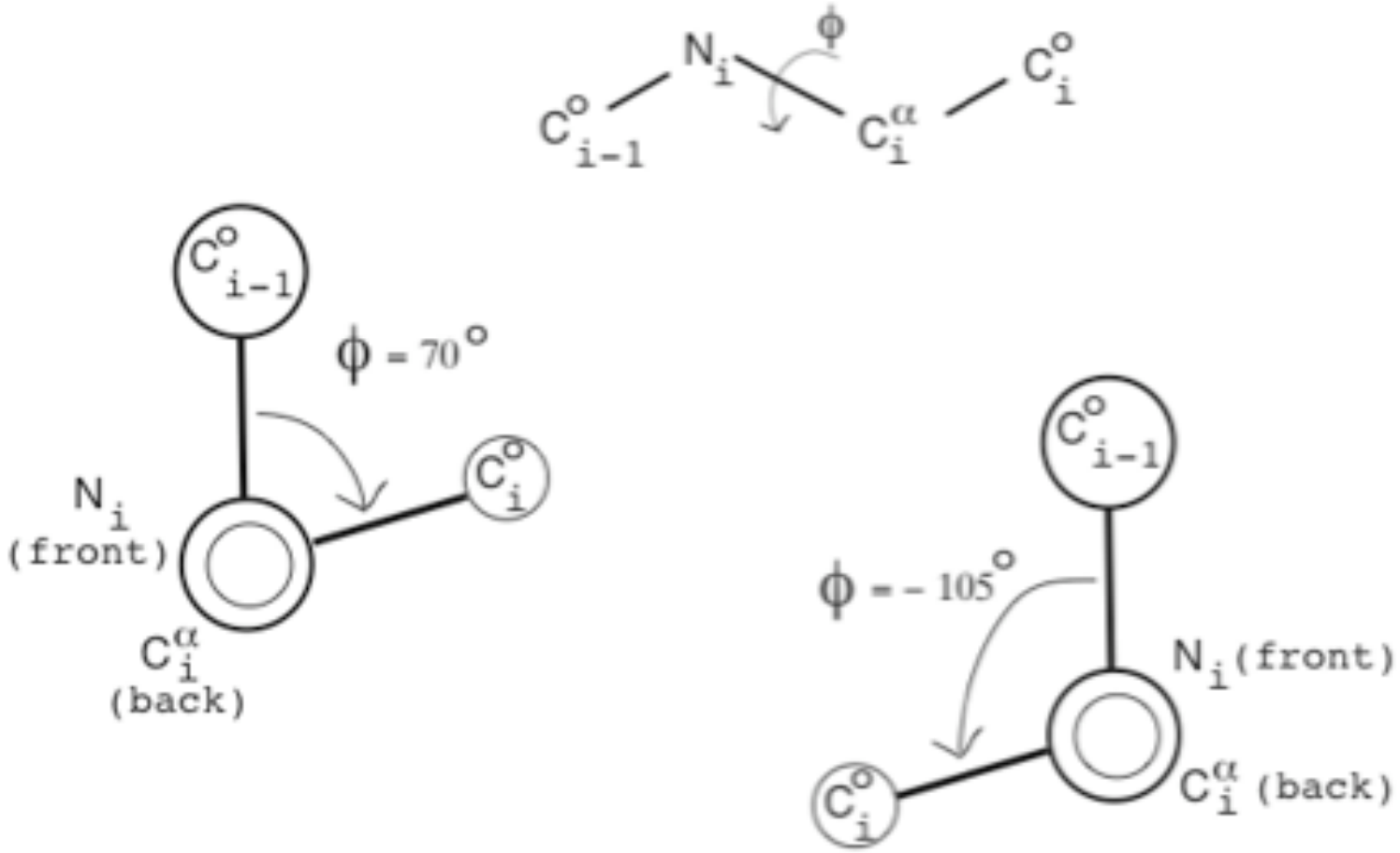
Describing geometry of peptide bonds using torsion angles:

4 atoms are needed to define a torsion angle.





phi


C^{o}_{i-1}
N_i
ϕ
C^{α}_{i}
C^{o}_{i}
C^{o}_{i-1}
$\phi = 70^{o}$
N_i (front)
C^{o}_{i}
C^{α}_{i} (back)
C^{o}_{i-1}
$\phi = -105^{o}$
N_i (front)
C^{o}_{i}
C^{α}_{i} (back)

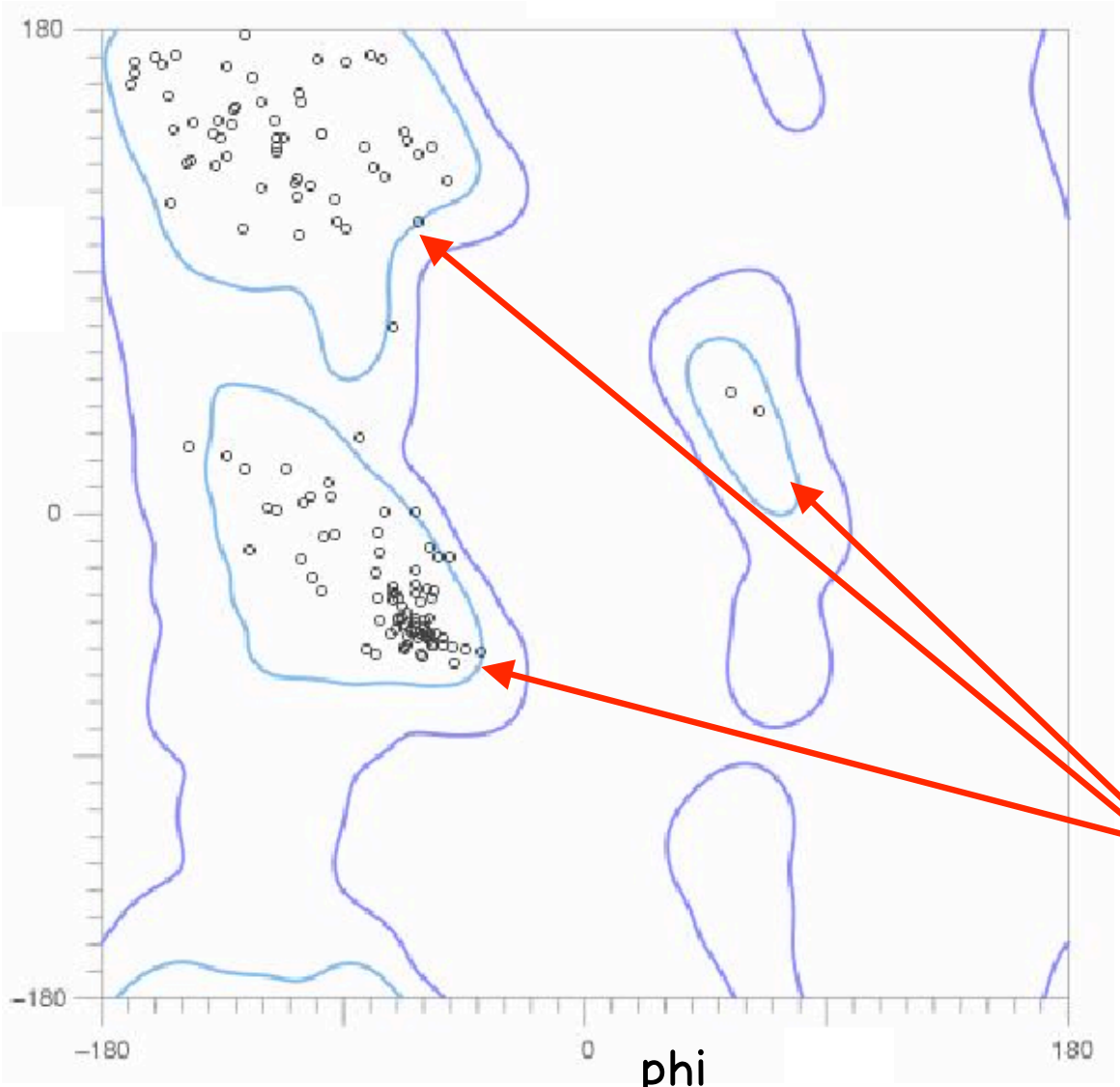
Some values of phi and psi are very common in proteins, and other values do not occur at all.

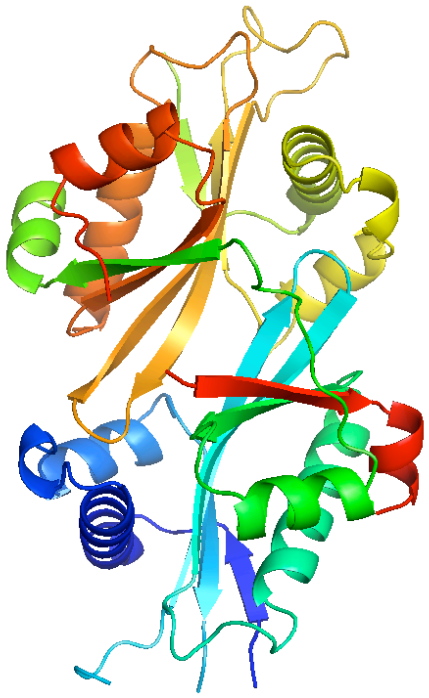
Some very common values of phi and psi are:

phi ≅ -60 degrees
psi ≅ -60 degrees
} alpha helix

phi ≅ -140 degrees
psi ≅ +120 degrees
} beta strand

Ramachandran plot for spermine-spermidine acetyl transferase (SSAT)

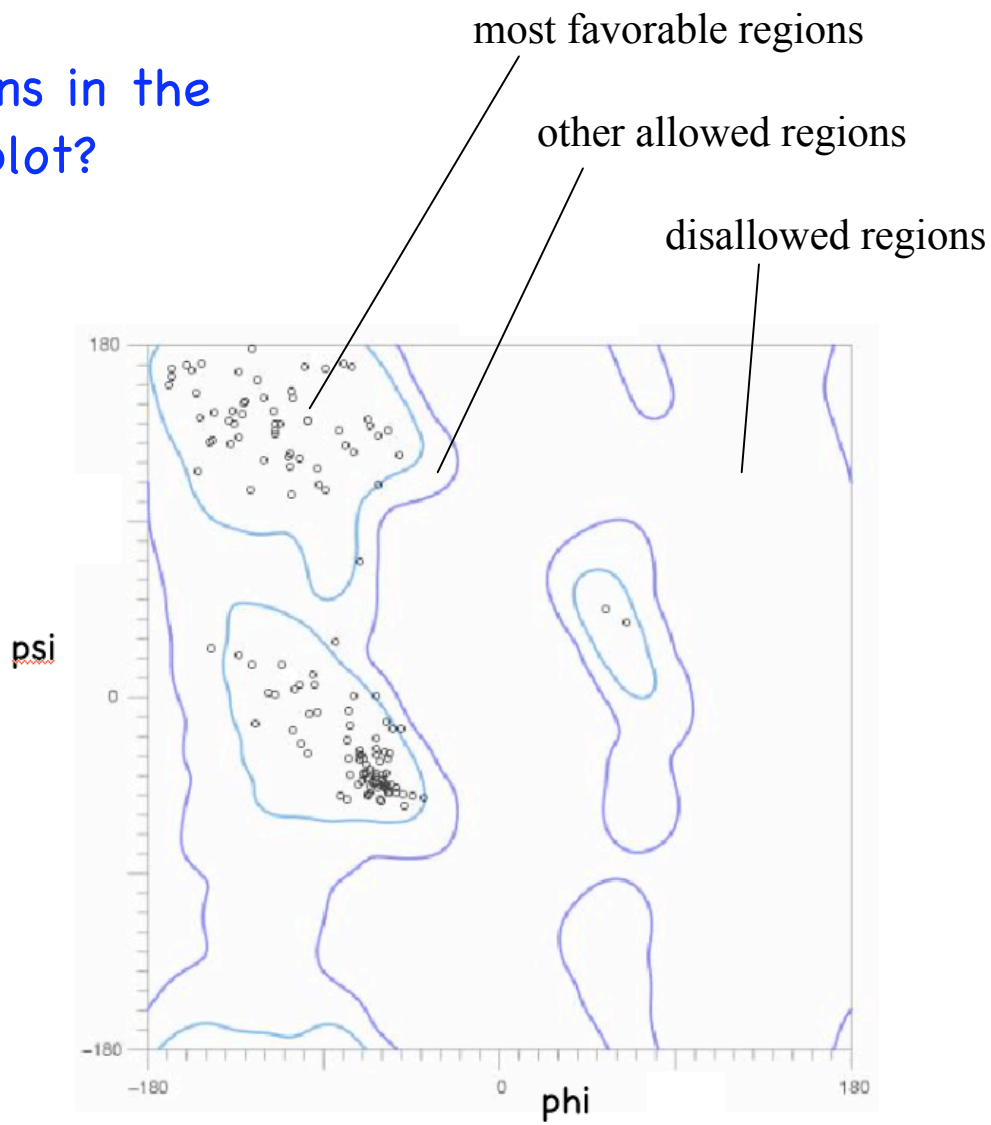




Blue circles indicate the most energetically favorable regions.

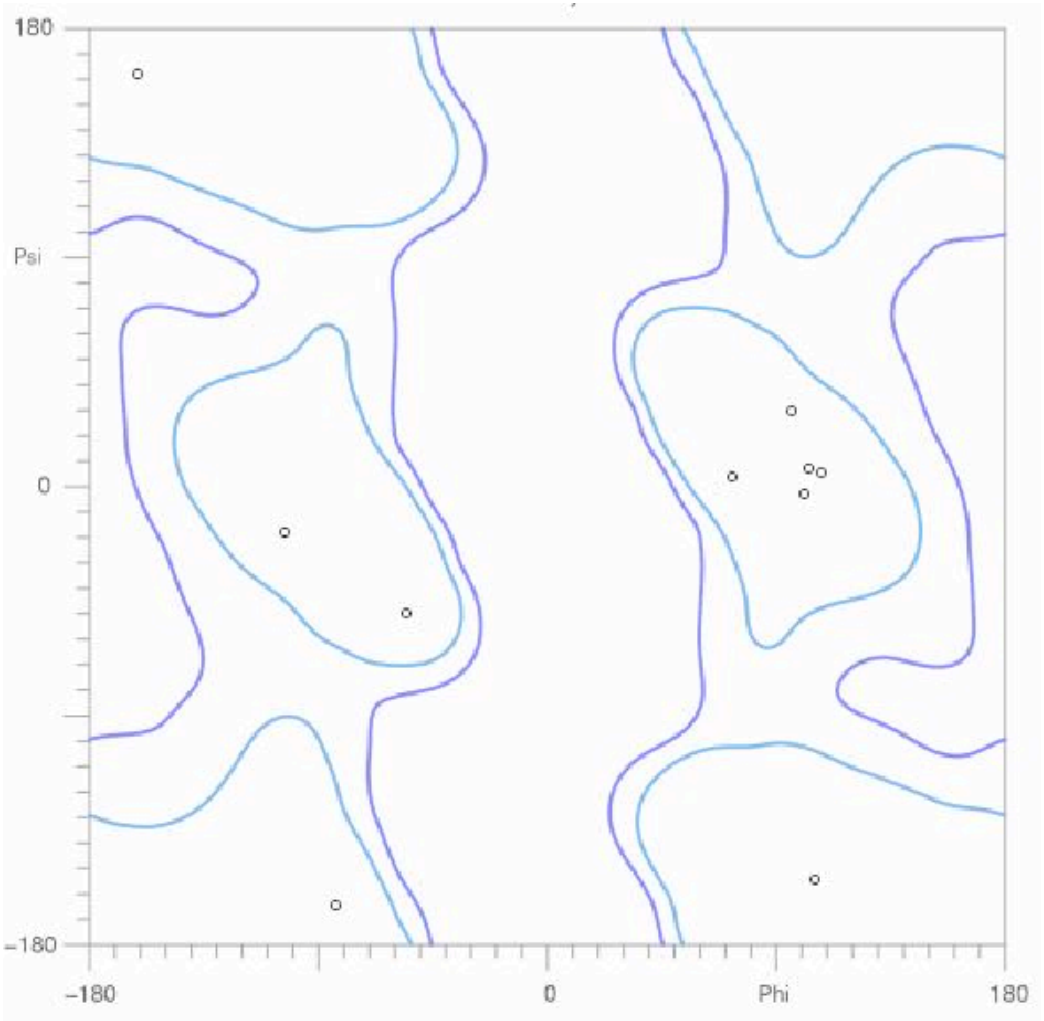
Plot for all a.a. types except Gly and Pro.

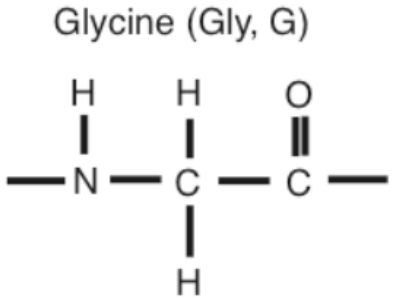
Why are there disallowed regions in the Ramachandran plot?



Disallowed regions are due to steric interactions among side chains.

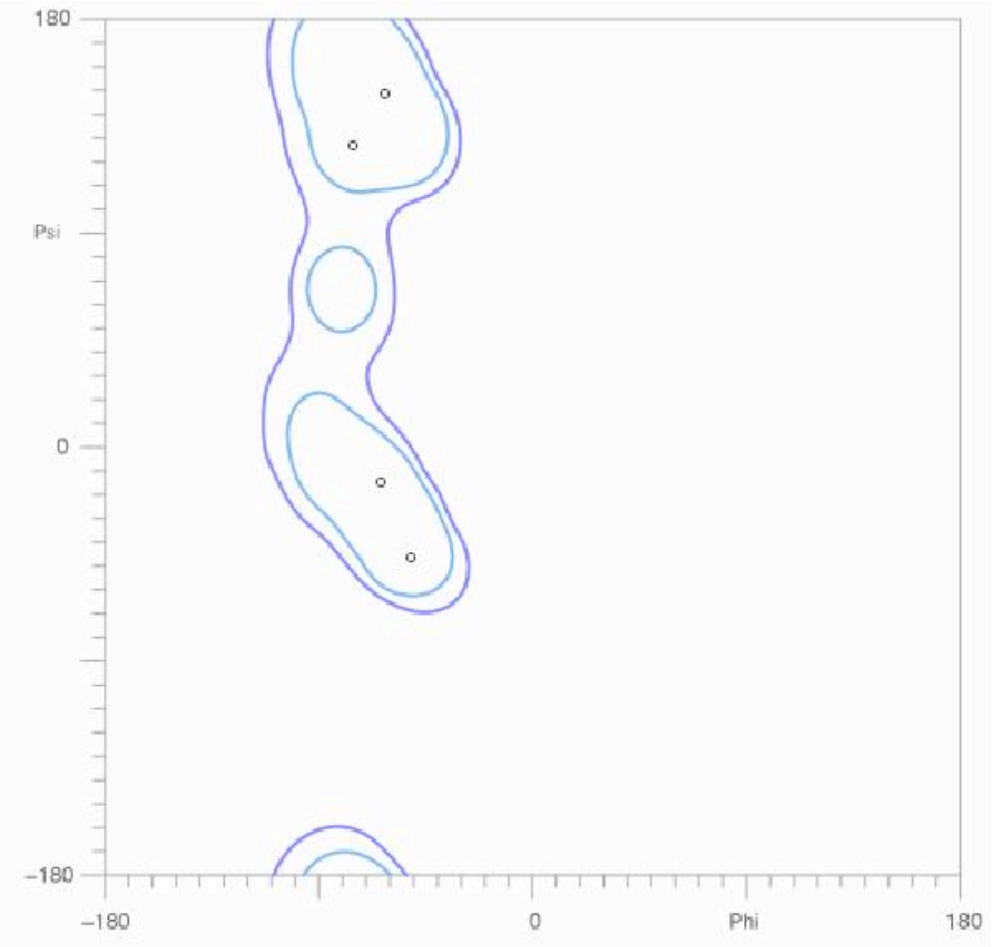
Ramachandran plot for glycines in SSAT.

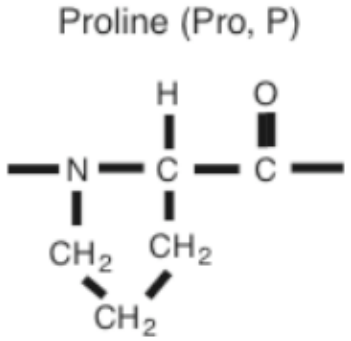




Allowed regions are large for glycine, compared to other a.a.

Ramachandran plot for prolines in SSAT.





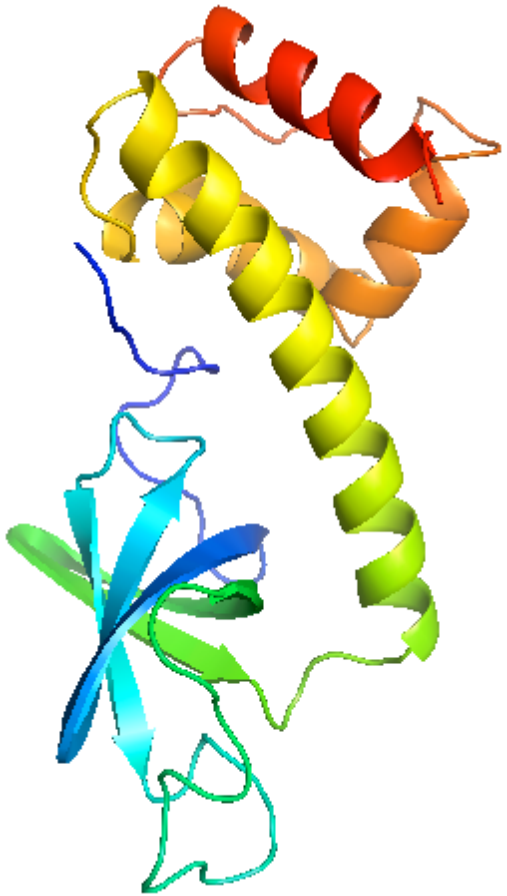
(notice allowed regions are small for proline, due to its cyclic side chain)

Proteins are flexible.

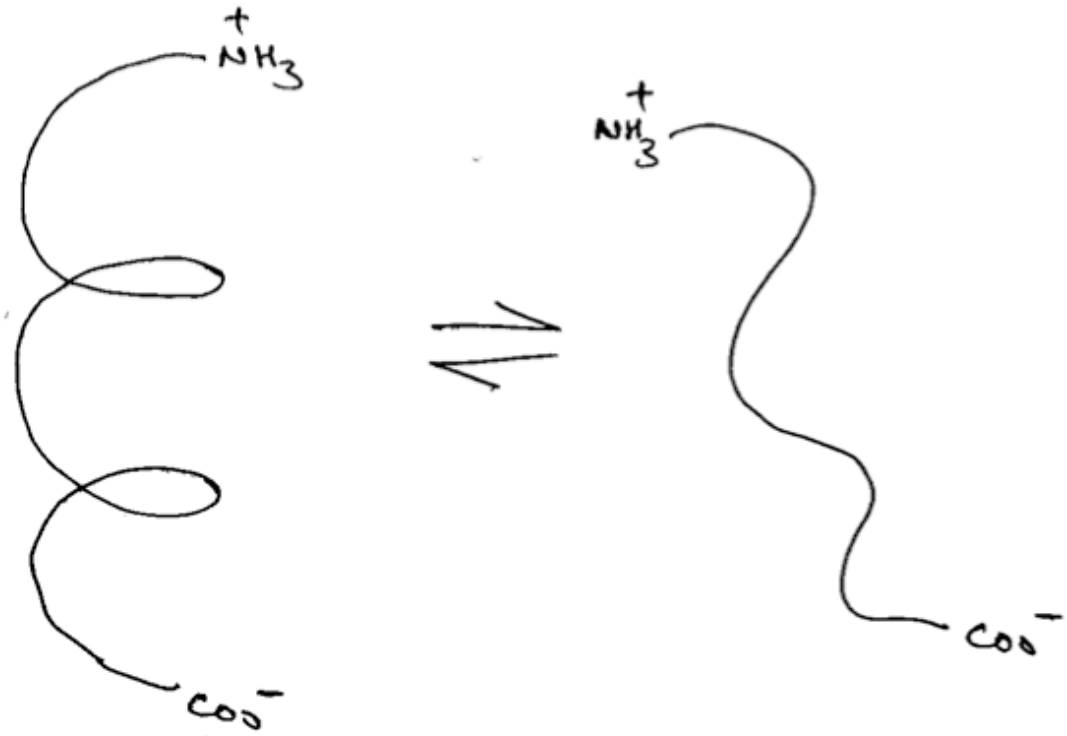
They exist in a rapid equilibrium between a larger number of conformational states.

The population of each state is governed by the Gibbs free energy of the state.

The "native" conformation of a protein is the conformation with the lowest free energy.



Think of a helix within a protein as being flexible, in equilibrium between a folded and unfolded form.



What stabilizes a helix, relative to an unfolded state?

1) Consider entropy.

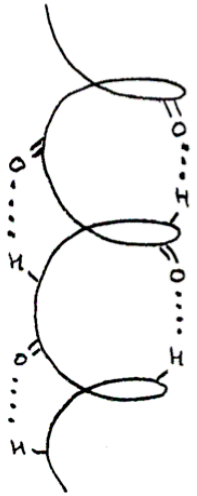
Helix: phi & psi torsion angles have only one allowed value ($\approx -60°$) per a.a.

unfolded: phi & psi have 3 allowed values for each amino acid.

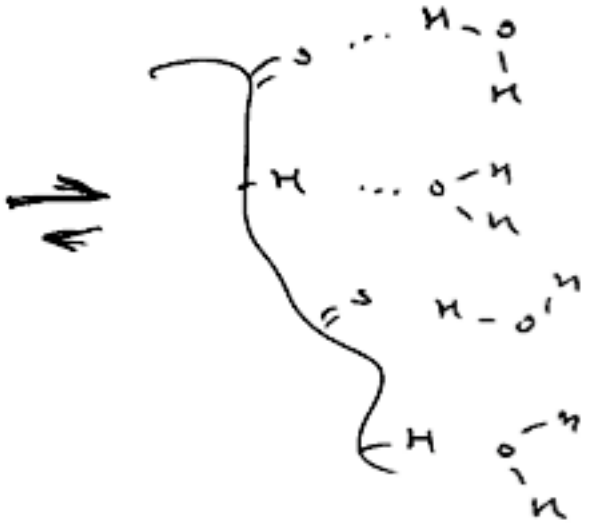
∴ entropy favors unfolded form.

2) Hydrogen bonds

Hydrogen bonds stabilize both the folded and unfolded state of a helix.



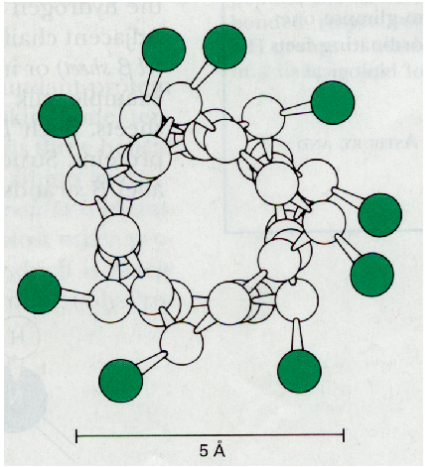
Intra-molecular H-bonds stabilize the folded helix.



H-bonds with water stabilize the unfolded form.

3) Van der Waals interactions between side chains are important for stabilizing a helix.

What does the potential energy curve for the van der Waals interaction looks like?

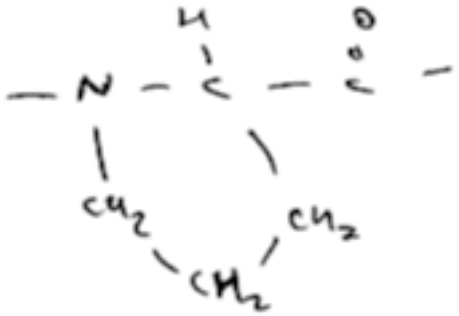


Some a.a. types have side chains that pack particularly well in helices (and are very common in helices):

Leu, Ala, Lys, Arg, Glu, Gln

Some a.a. types are relatively uncommon in helices:

Proline is not often found in a helix.



Proline has more restriction on phi and psi than other a.a. (due to cyclic side chain).

Also, proline has no amide (N-H) proton, so it can't hydrogen bond with another a.a. in a helix.

From the van Holde text book:

TABLE 4.2 PROPENSITIES OF AMINO ACIDS TO FORM α-HELICES ($\langle P_\alpha \rangle$) AND β-SHEETS ($\langle P_\beta \rangle$)

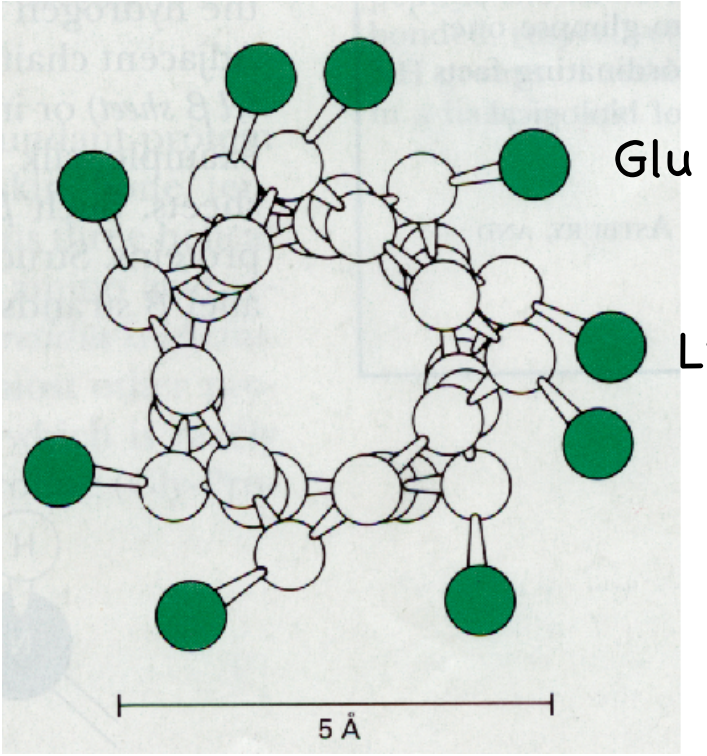
α-residues	$\langle P_\alpha \rangle$	β-Residues	$\langle P_\beta \rangle$
Glu	1.44 ± 0.06	Val	1.64 ± 0.07
Ala	1.39 ± 0.05	Ile	1.57 ± 0.08
Met	1.32 ± 0.11	Thr	1.33 ± 0.07
Leu	1.30 ± 0.05	Tyr	1.31 ± 0.09
Lys	1.21 ± 0.05	Trp	1.24 ± 0.14
His	1.12 ± 0.08	Phe	1.23 ± 0.09
Gln	1.12 ± 0.07	Leu	1.17 ± 0.06
Phe	1.11 ± 0.07	Cys	1.07 ± 0.12
Asp	1.06 ± 0.06	Met	1.01 ± 0.13
Trp	1.03 ± 0.10	Gln	1.00 ± 0.09
Arg	1.00 ± 0.07	Ser	0.94 ± 0.06
Ile	0.99 ± 0.06	Arg	0.94 ± 0.09
Val	0.97 ± 0.05	Gly	0.87 ± 0.05
Cys	0.95 ± 0.09	His	0.83 ± 0.09
Thr	0.78 ± 0.05	Ala	0.79 ± 0.05
Asn	0.78 ± 0.06	Lys	0.73 ± 0.06
Tyr	0.73 ± 0.06	Asp	0.66 ± 0.06
Ser	0.72 ± 0.04	Asn	0.66 ± 0.06
Gly	0.63 ± 0.04	Pro	0.62 ± 0.07
Pro	0.55 ± 0.05	Glu	0.51 ± 0.06

Listed are values compiled from the crystal structures of 64 proteins, and the assignments as former (*H* and *h*), indifferent (*I* and *i*) and breakers (*b* and *B*) for each type of structure.

From Chou, P. Y. in *Prediction of Protein Structure and the Principles of Protein Conformation* (1989), ed. G. D. Fasman, Plenum Press, New York, pp. 549–586.

4) Salt - bridges

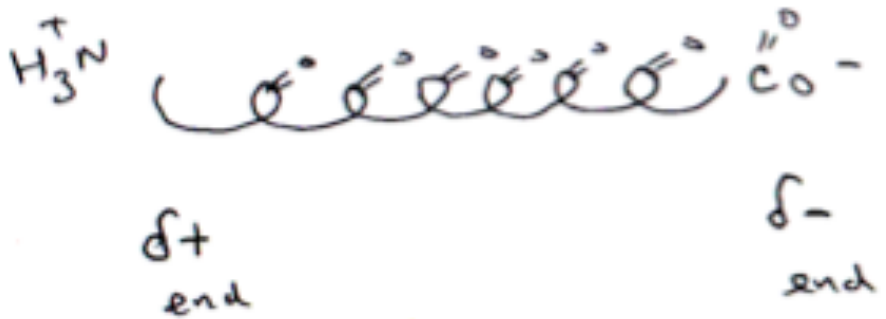
oppositely charged a.a. tend to be spaced by 3-4 aa. in primary sequence in very stable helices.



5) "helix dipole"

A separation of + & − charge is called an electric dipole.

Helices have a non-zero dipole moment; they definitely have a + & − end, even if no charged a.a. are present.



The stability of a helix with sequence:

EEEEEEEEEEAAAAAAAAAA

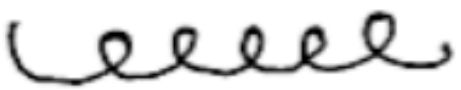
is very different from a helix with sequence:

AAAAAAAAAAEEEEEEEEEE

Melting points of these helices differ by about 30 °C.

Reason for difference in stability: Interaction of the negatively charged Glu (E) with the helix dipole.

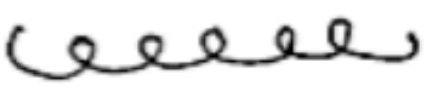
⊕ Glu_{10} Ala_{10} ⊖



more stable

↳ negative glutamic acids interact w/ helix dipole favorably.

⊕ Ala_{10} Glu_{10} ⊖



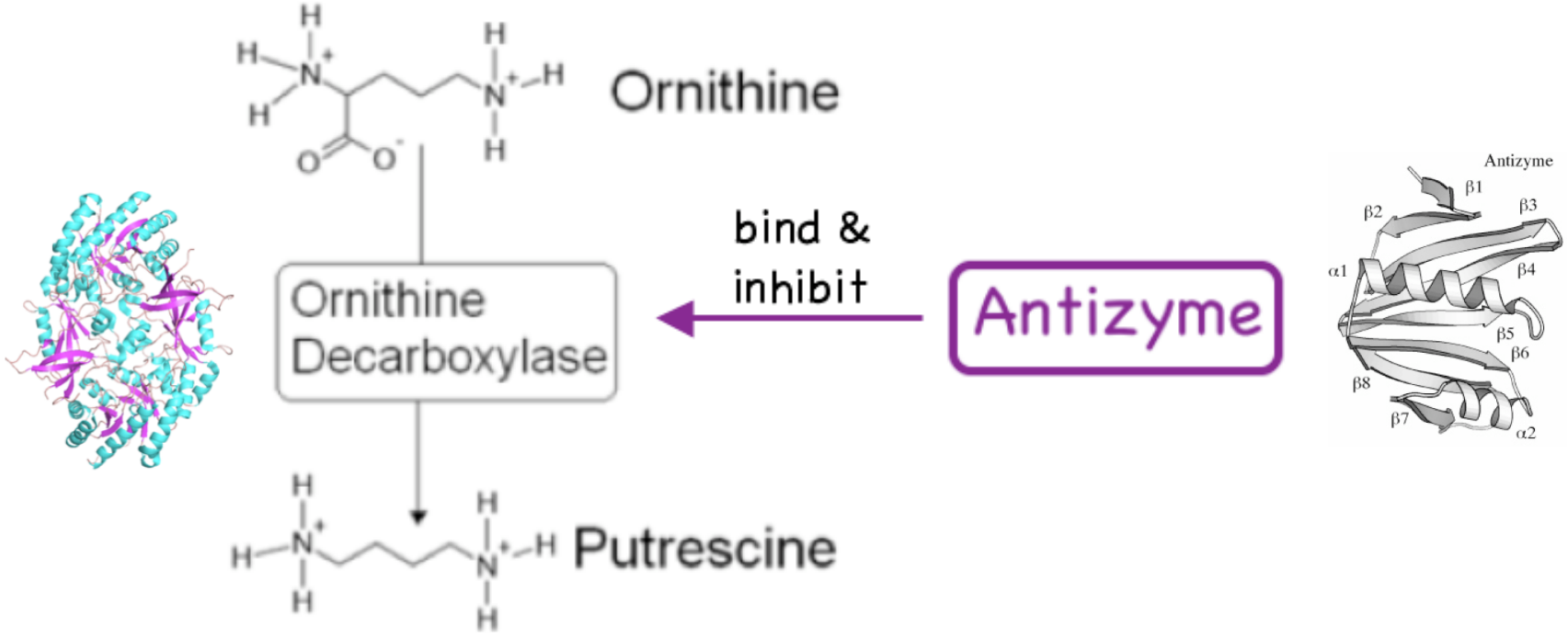
less stable

Last topic for today:

Once the structure of a protein is known (by x-ray crystallography or NMR), how do you know which part of the protein is important for function?

For example:

How could you figure out which amino acids in Antizyme are involved in binding and inhibiting the enzyme Ornithine Decarboxylase?



Answer:

Compare the amino acid sequences of antizyme from many different species.

The functionally important amino acids are probably the ones that are on the protein surface, and don't vary between species.

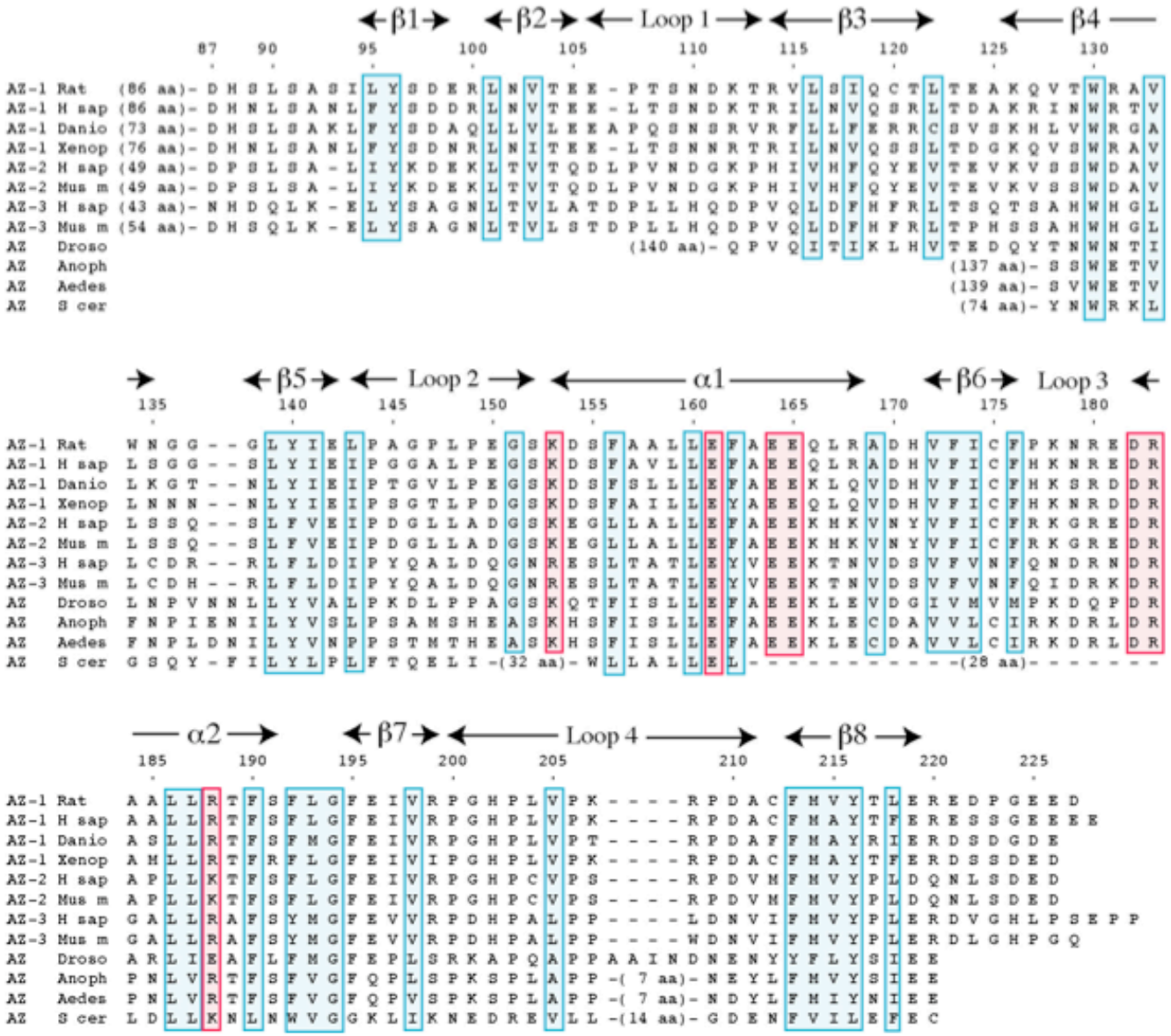
(a.a. that don't vary between species are called "conserved").

Why does this work?

Amino acids that are essential for protein function do not vary between species (because mutations result in a non-functional protein).

Interior are important for protein stability; surface a.a. are more likely to be important for protein function.

Comparison of a.a. sequences of antizyme from 12 differenct species.



Blue = conserved, inside protein ; Red = conserved, on surface of protein.

Antizyme contains several conserved glutamic acids (E161, E164, E165) on its surface. These are likely to be important in binding ornithine decarboxylase.