

Secondary structure and amino acid properties

Magnus Andersson

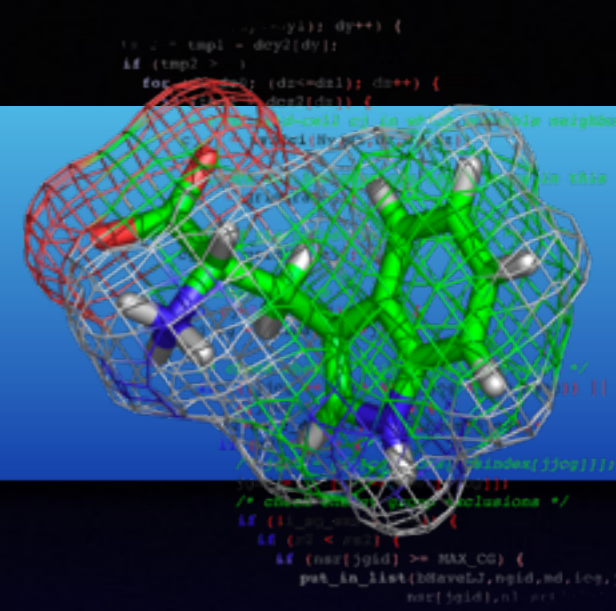
magnus.andersson@scilifelab.se

Theoretical & Computational Biophysics

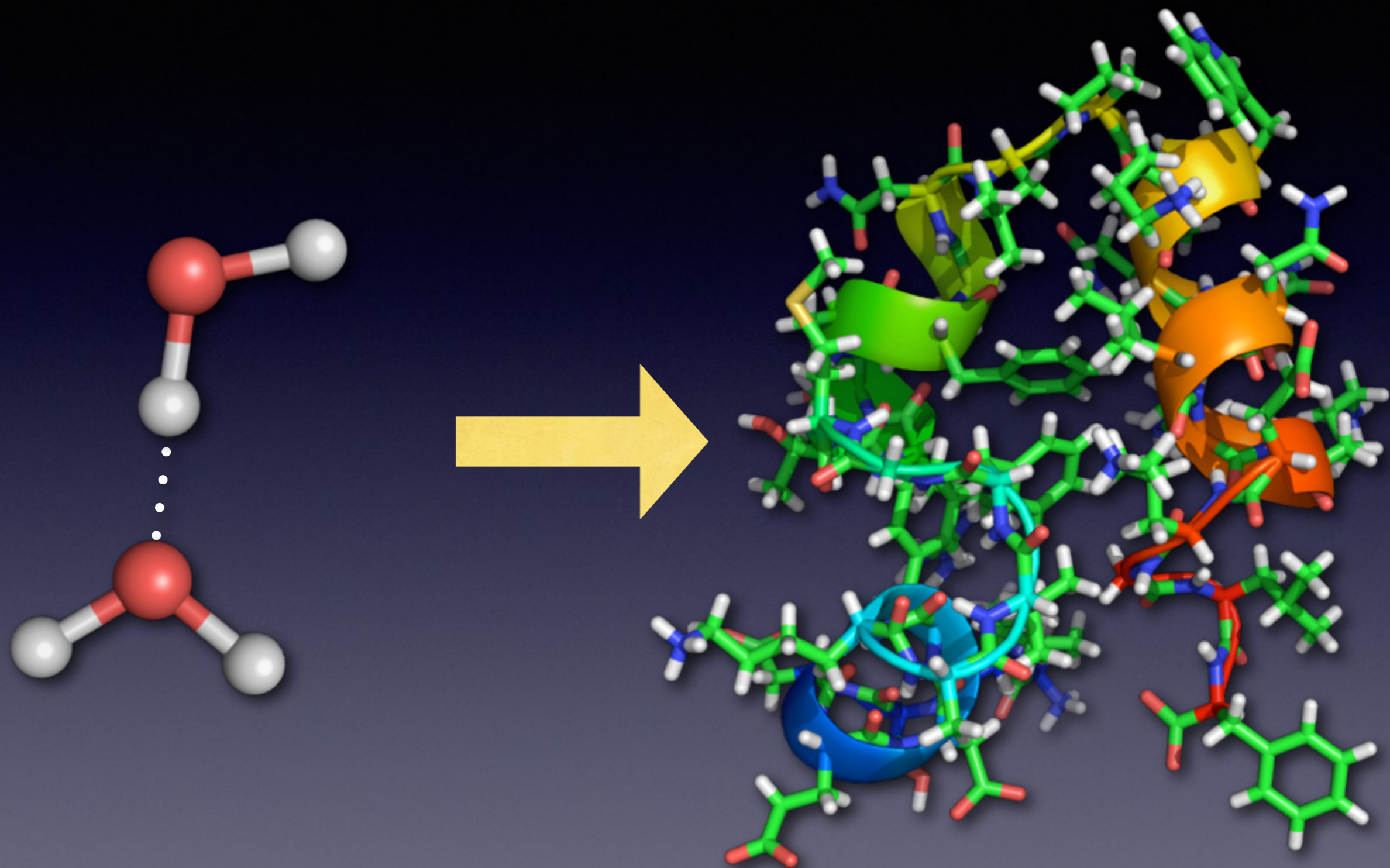
SciLifeLab



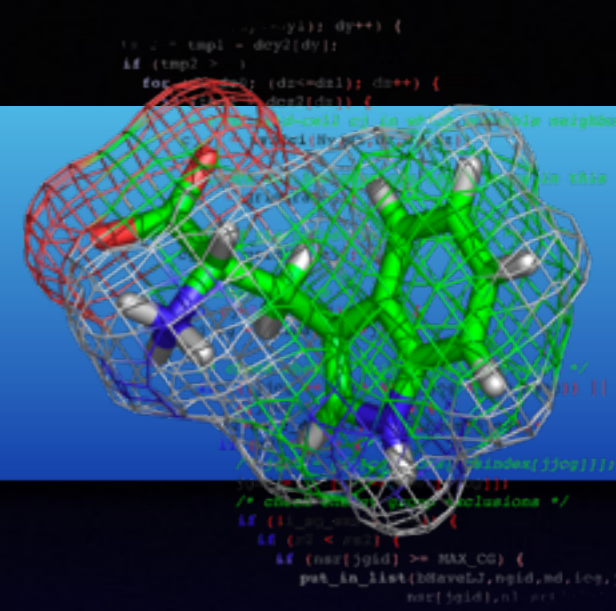
Recap



- **Hydrophobic effect**
- **Solubility & Partitioning**
- **Electrostatics is very strong**
 - **Special screening effects**
 - **Molecules reorient to maintain interactions**
 - **Leads to entropic effects**
- **Protein folding is largely determined by hydrophobicity, and entropy is critical**
- **The Molten Globule**

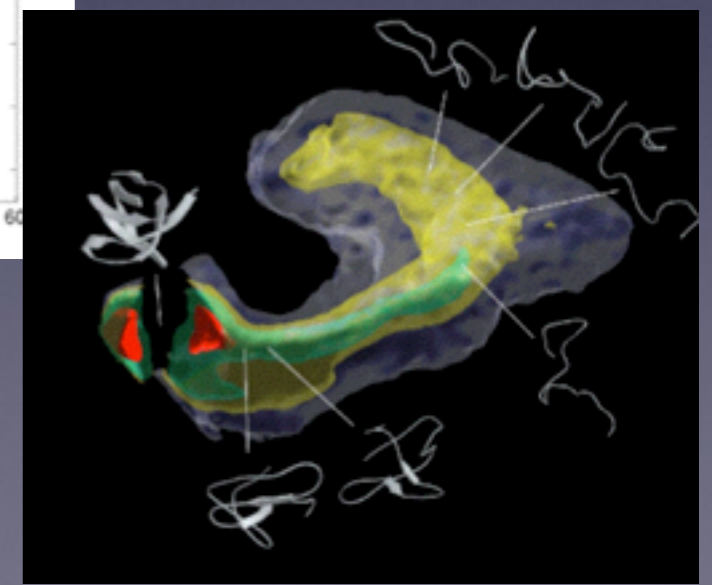
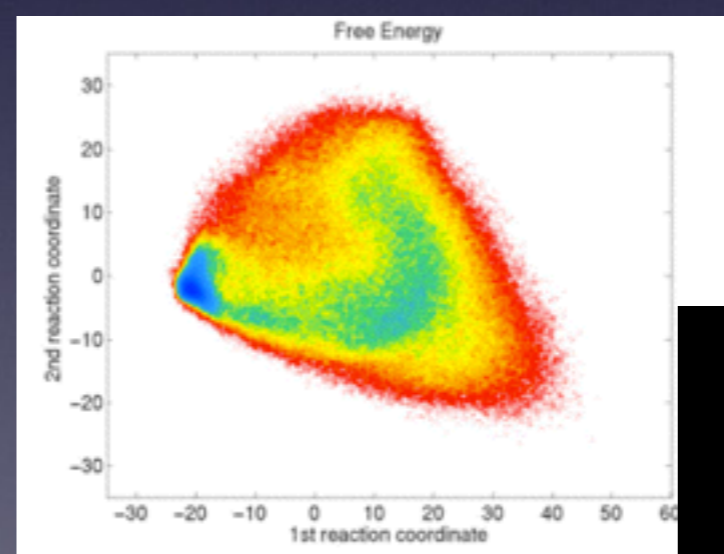
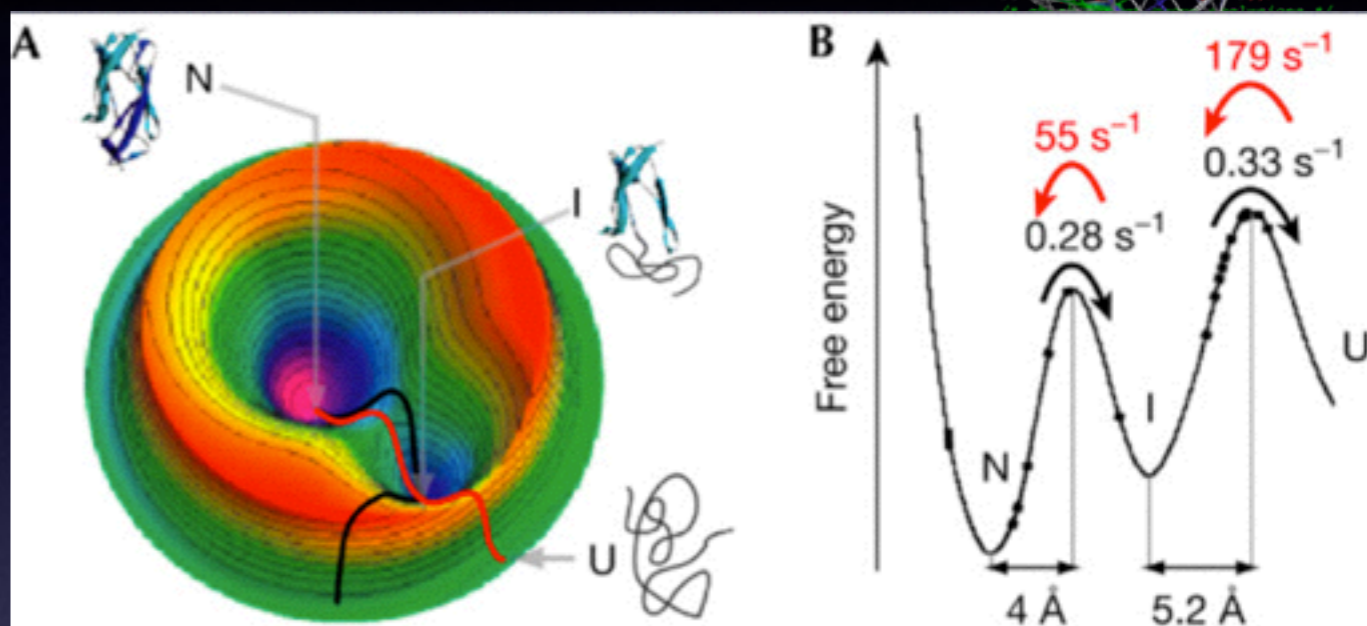
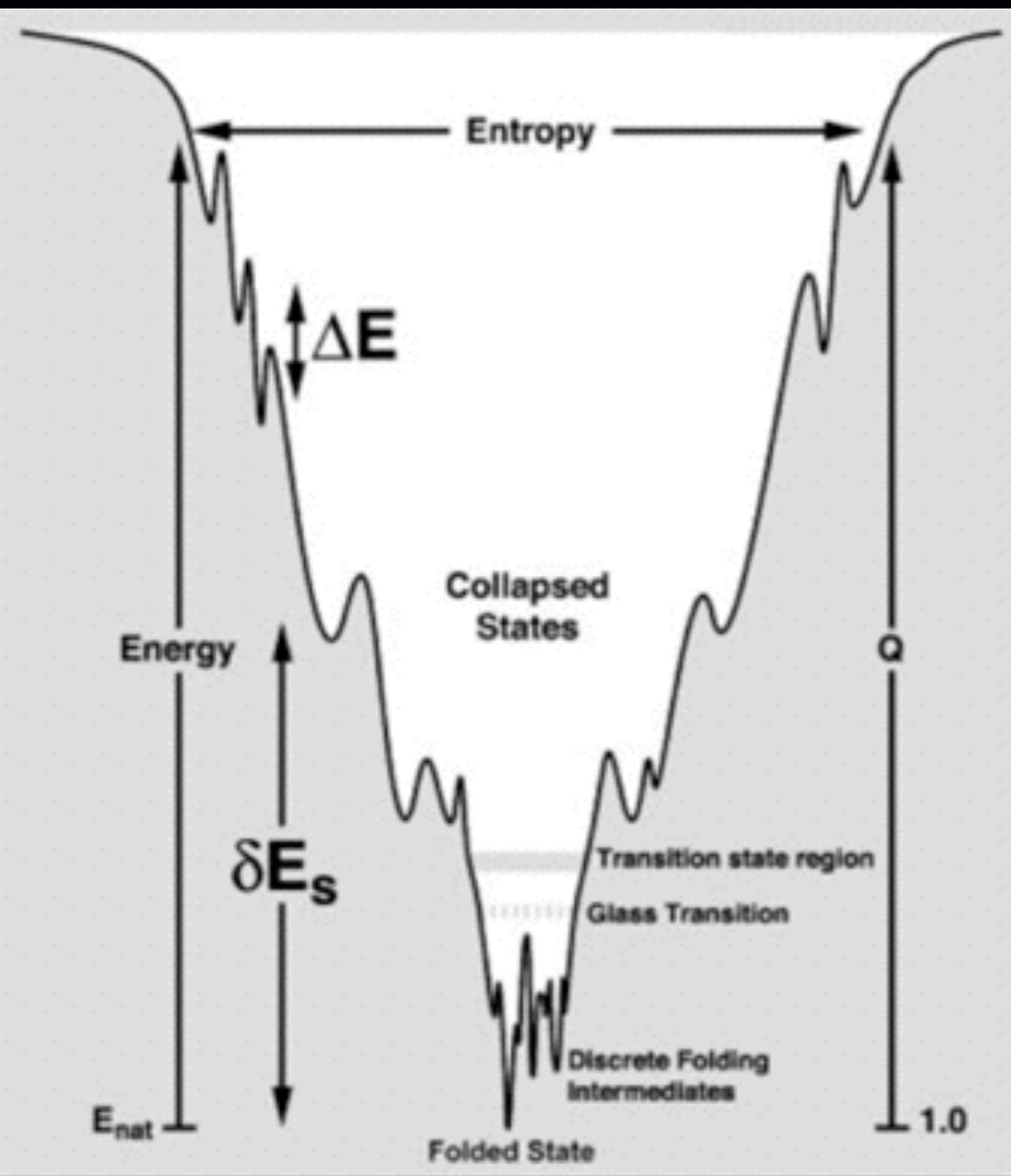
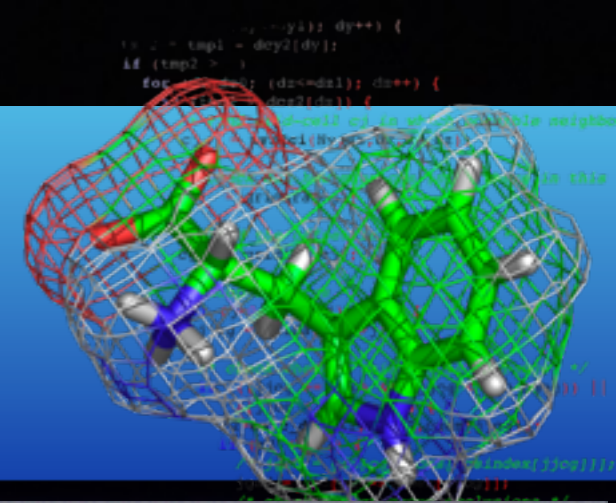


Outline today

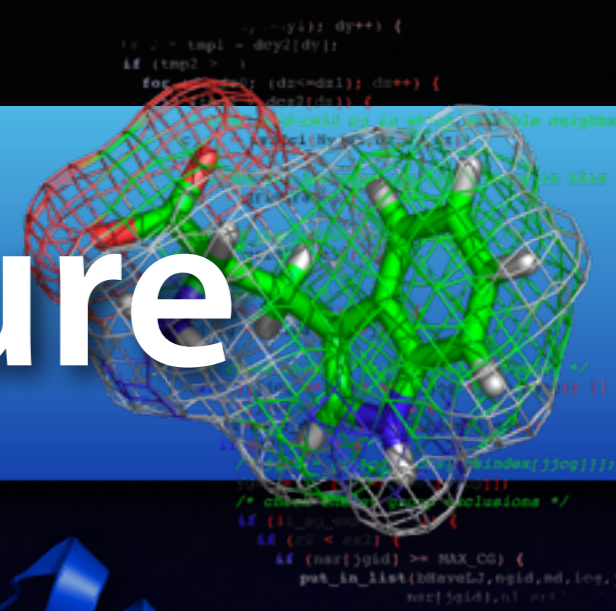


- **Back to the polypeptide chains**
- **Secondary structures & turns**
 - **Geometry, topology**
 - **Stabilization**
- **Amino acid properties & titration**

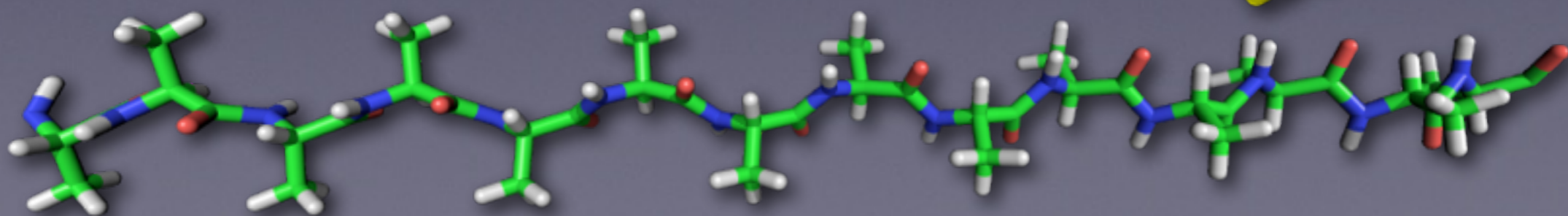
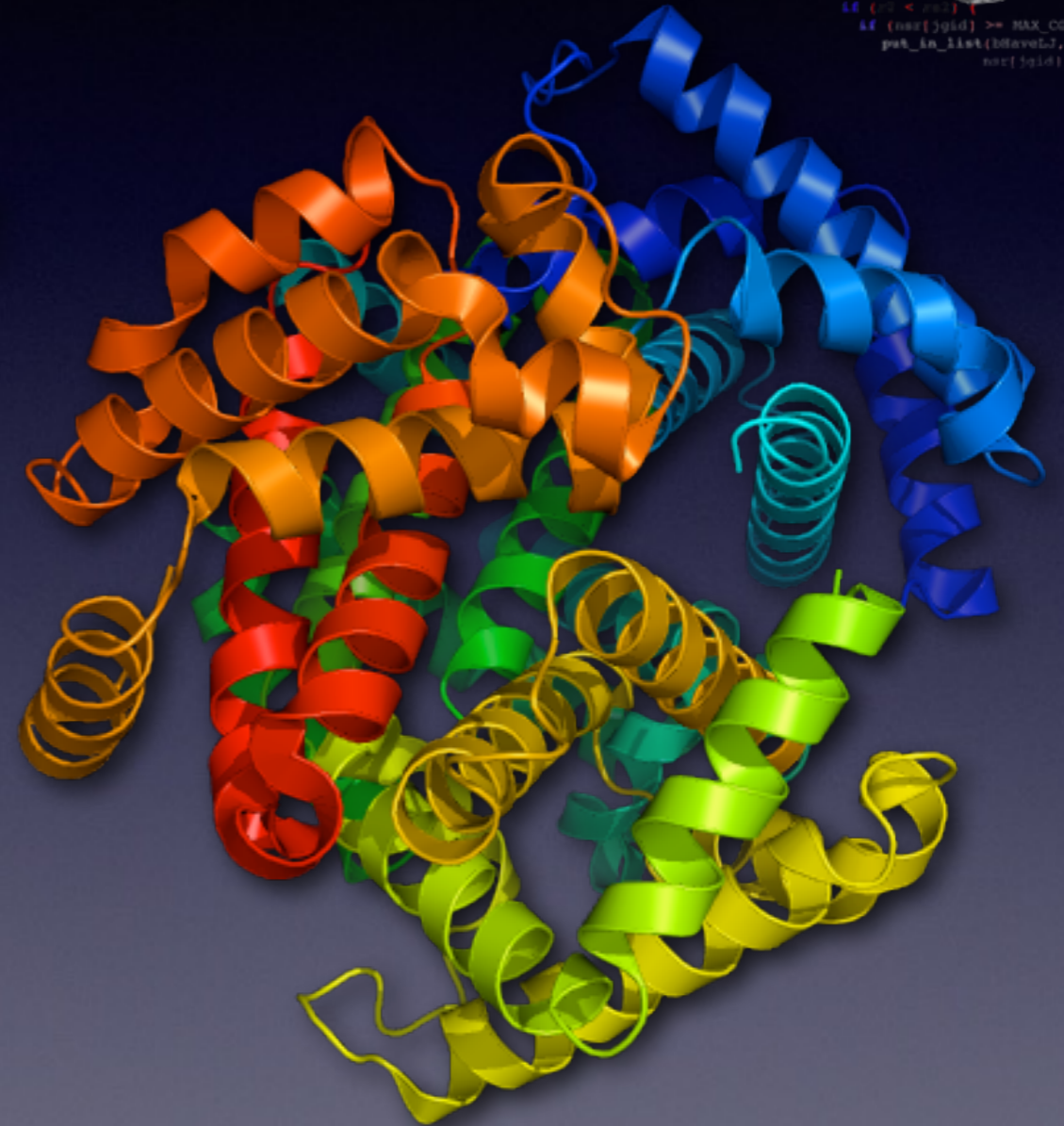
Energy landscapes



Secondary Structure



- Think in terms of ΔG now!
- What happens during folding?
- Why are interactions important?

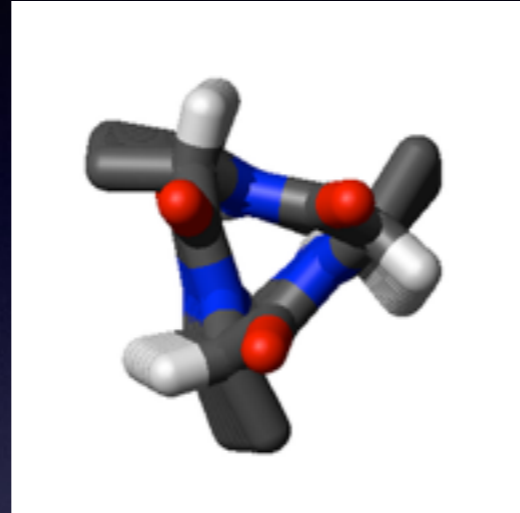
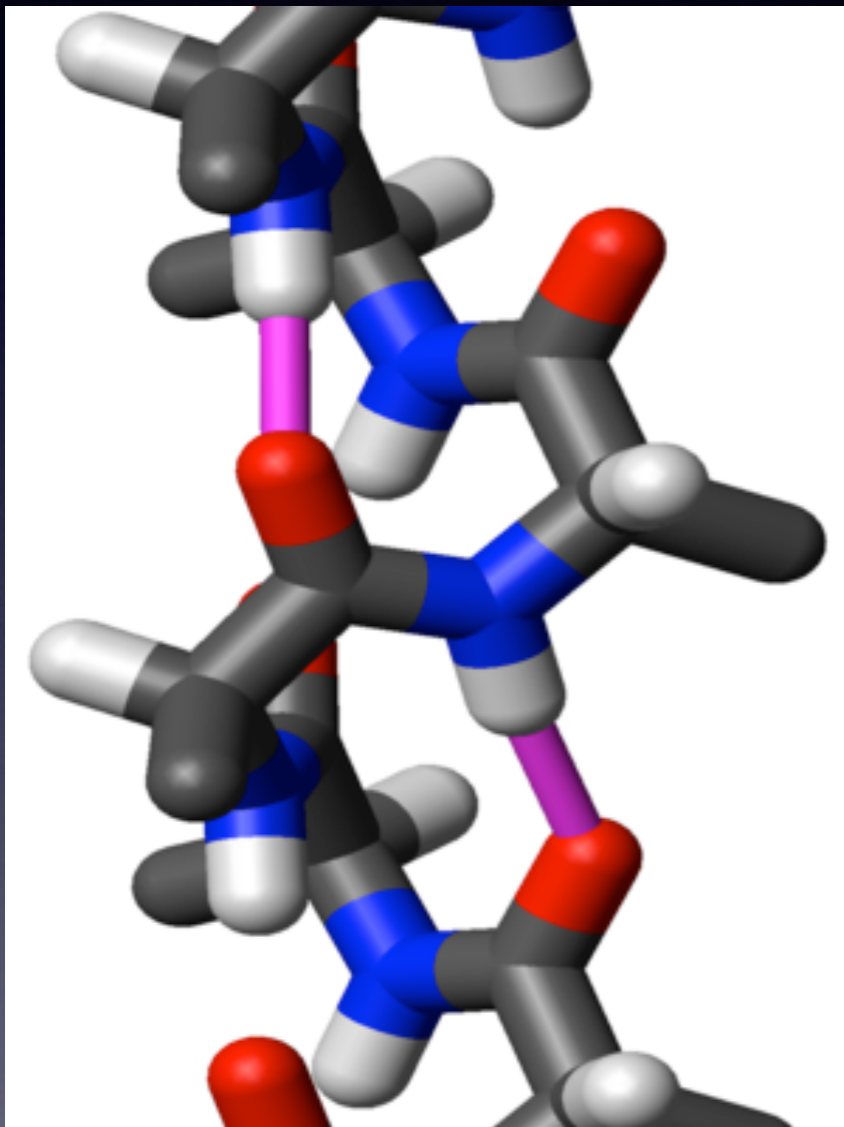
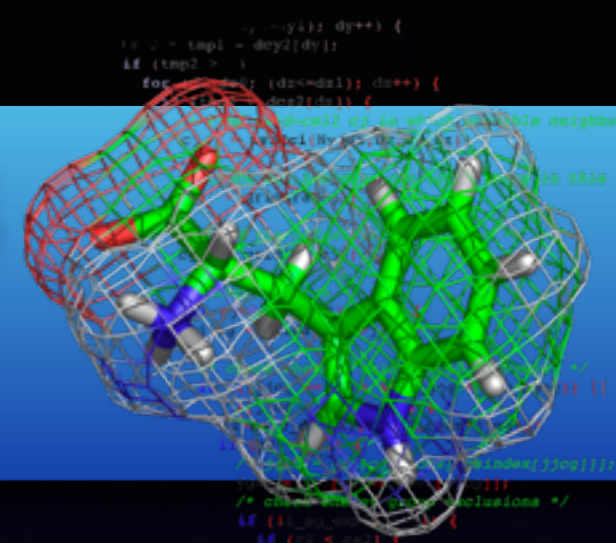


Alpha helices

- Hydrogen bonds: i to $i+4$
 - 0-4, 1-5, 2-6
- First hydrogen bond “locks” residues 1,2,3 in place
- Second stabilizes 2,3,4 (etc.)

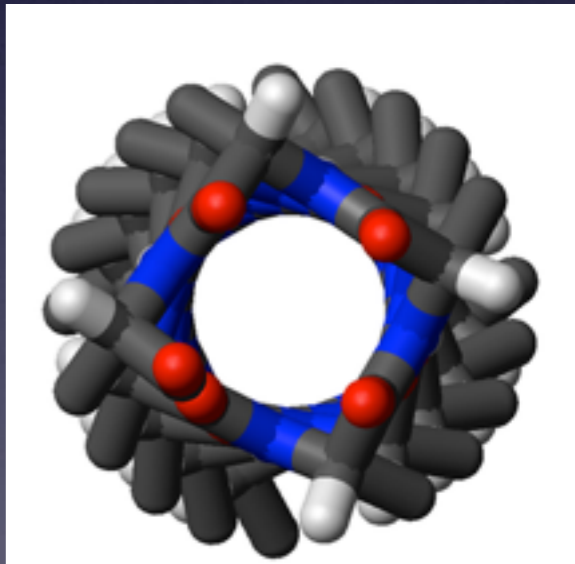
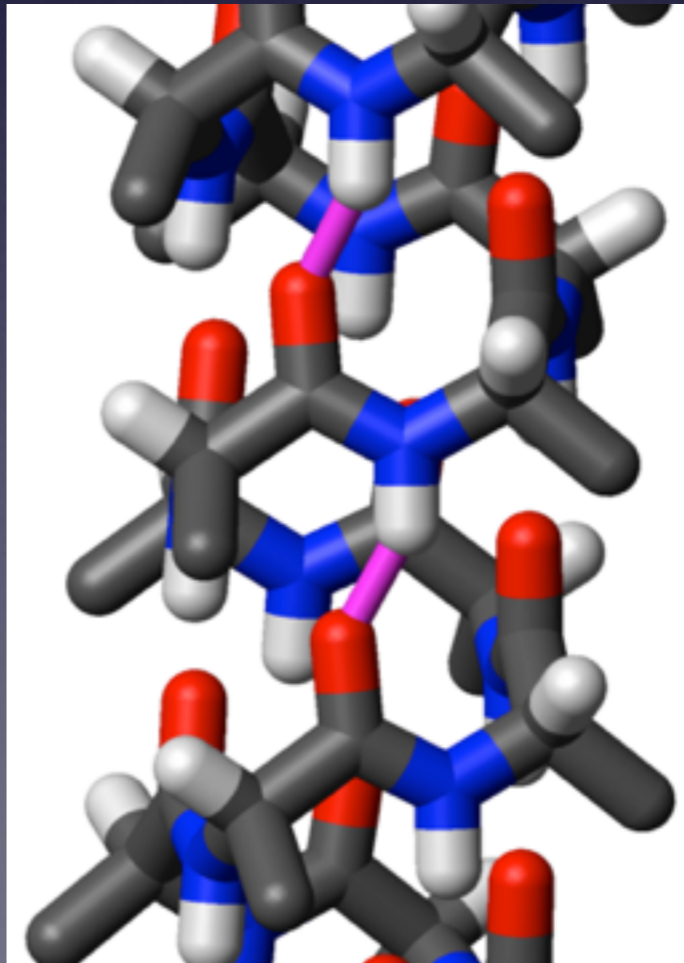


Why are alternative helices less common?



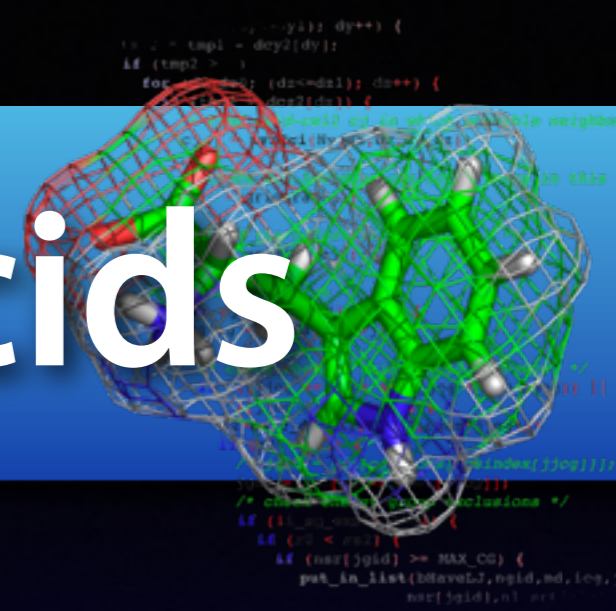
3_{10} helix

...and why are there only 3-4 different helix structures?



π helix

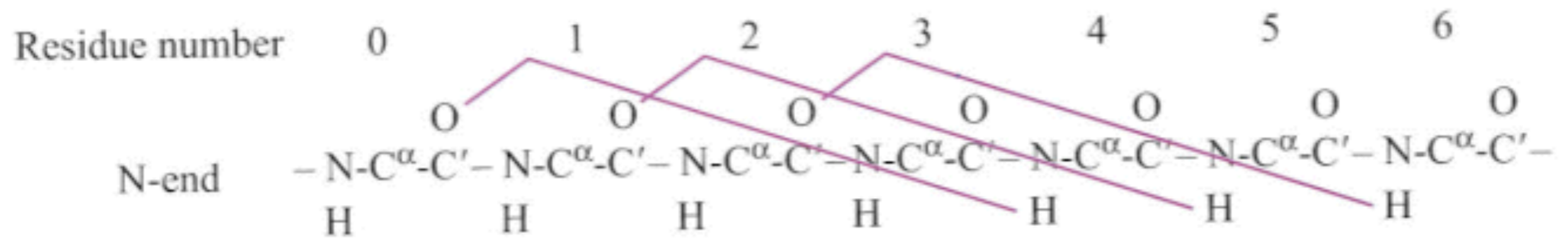
titratable amino acids



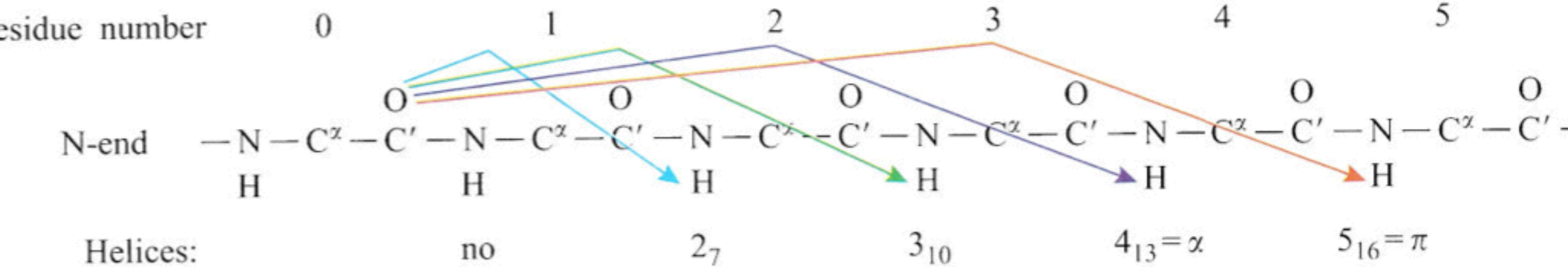
50.0us



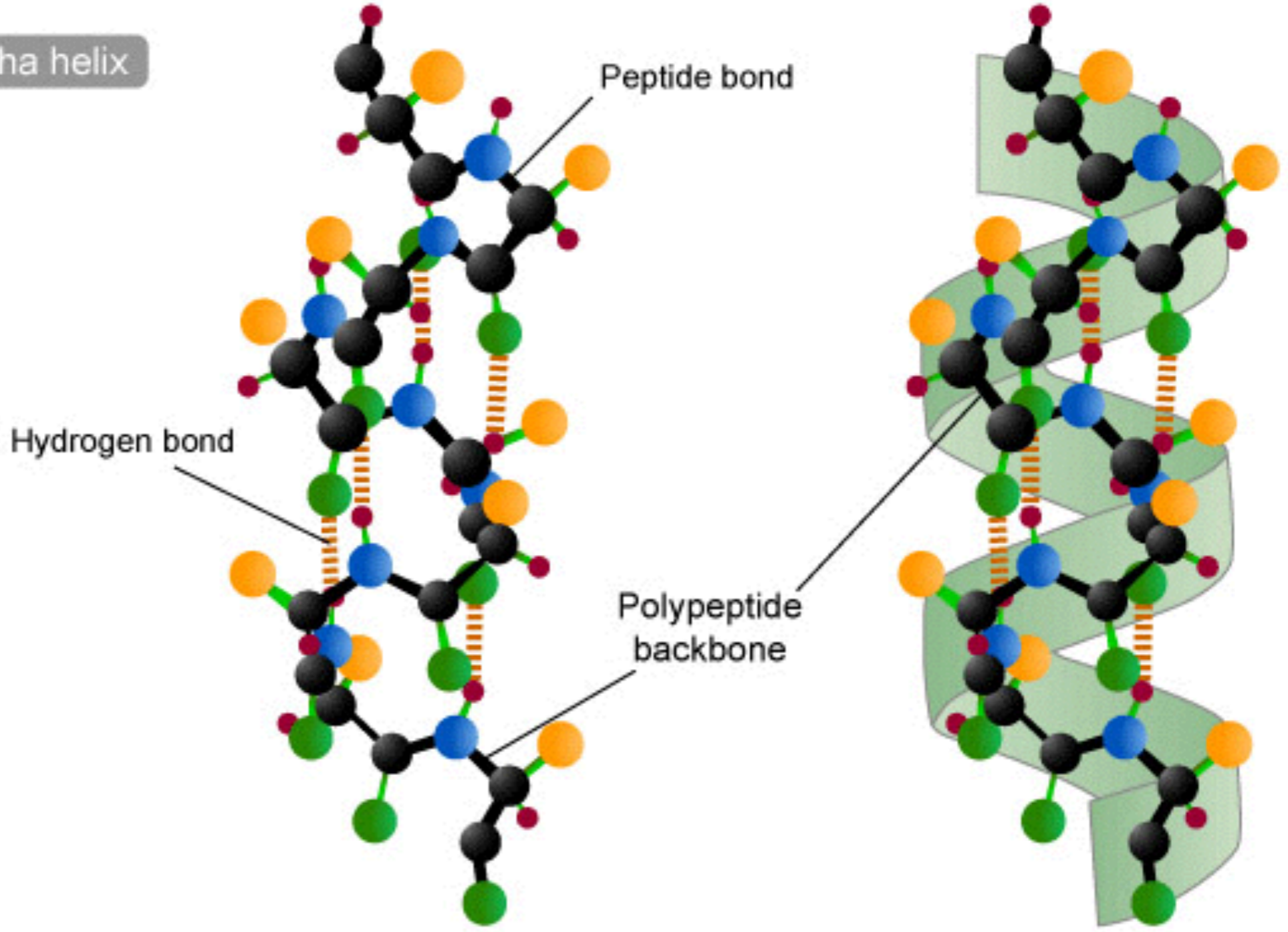
Alpha



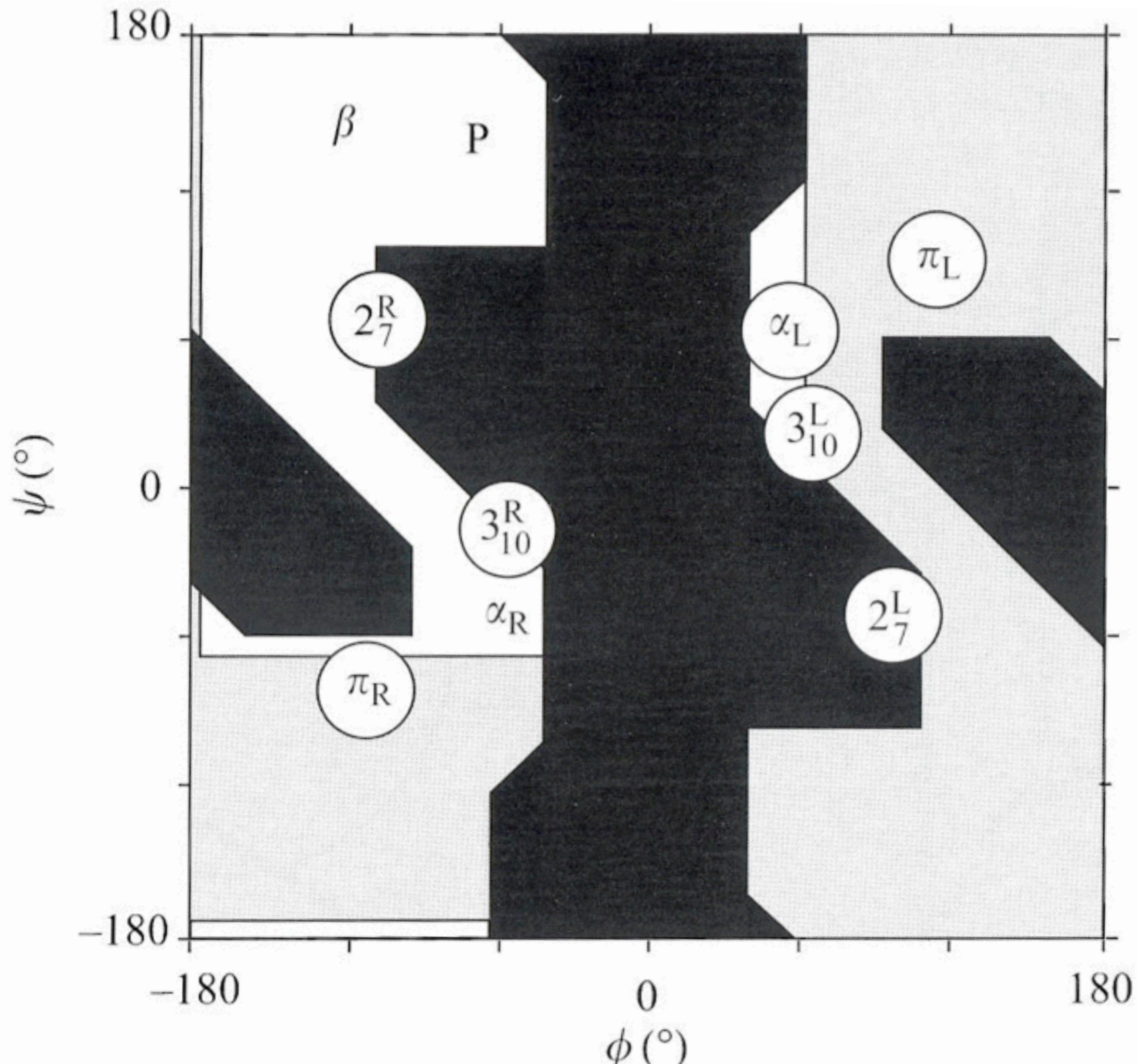
Other helices



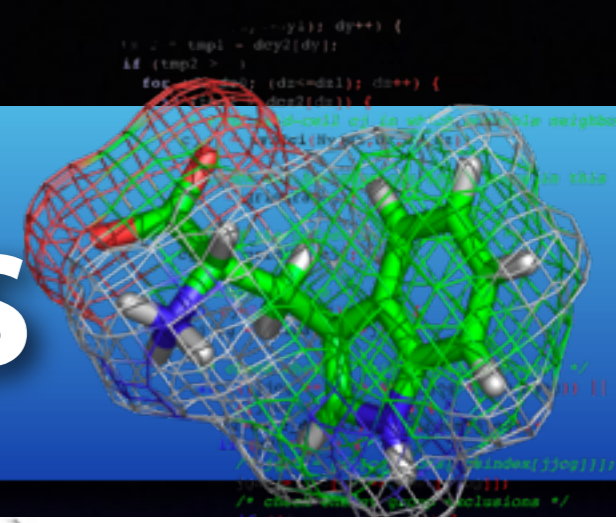
Alpha helix



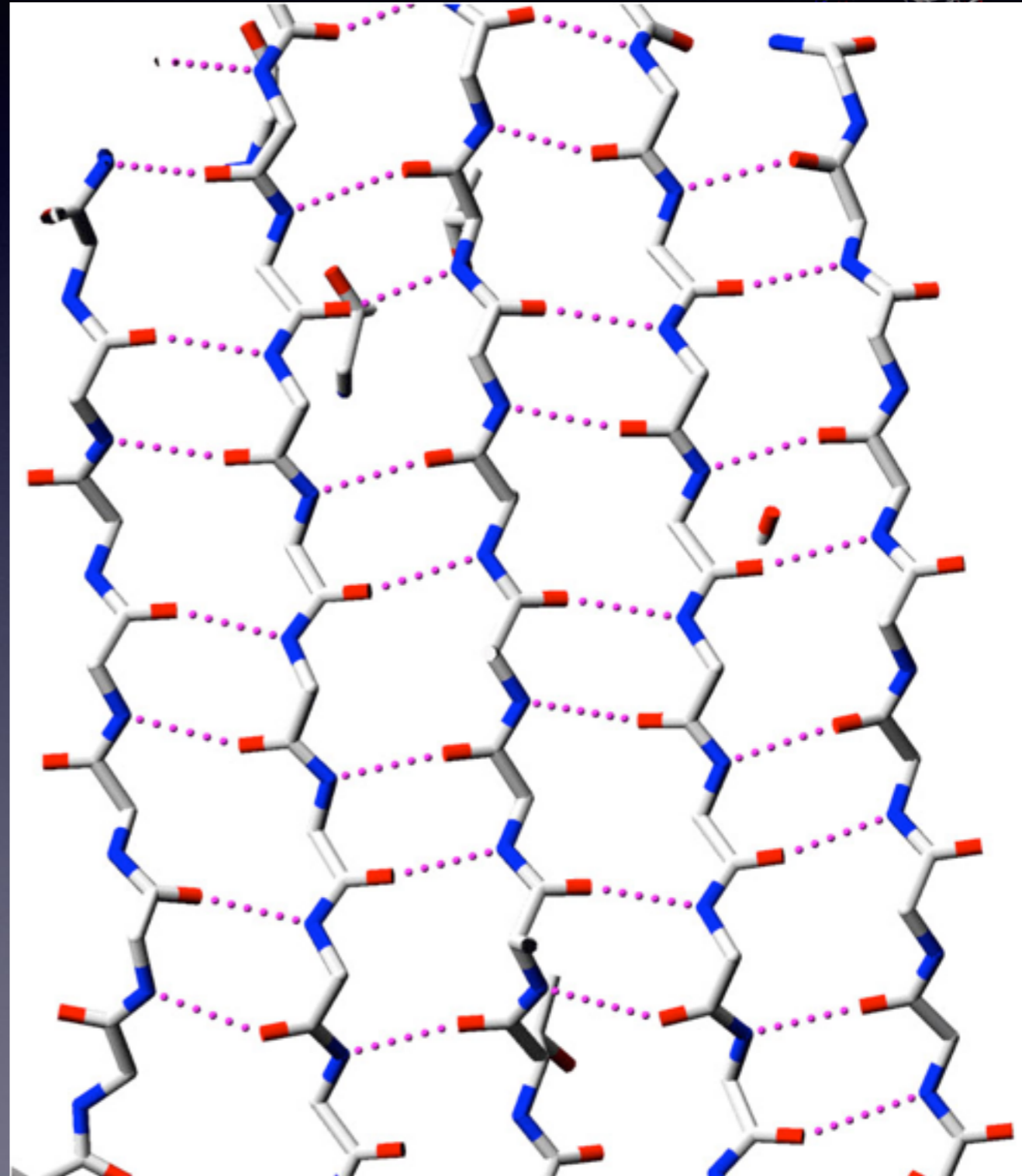
- Carbon
- Nitrogen
- R-group
- Oxygen
- Hydrogen

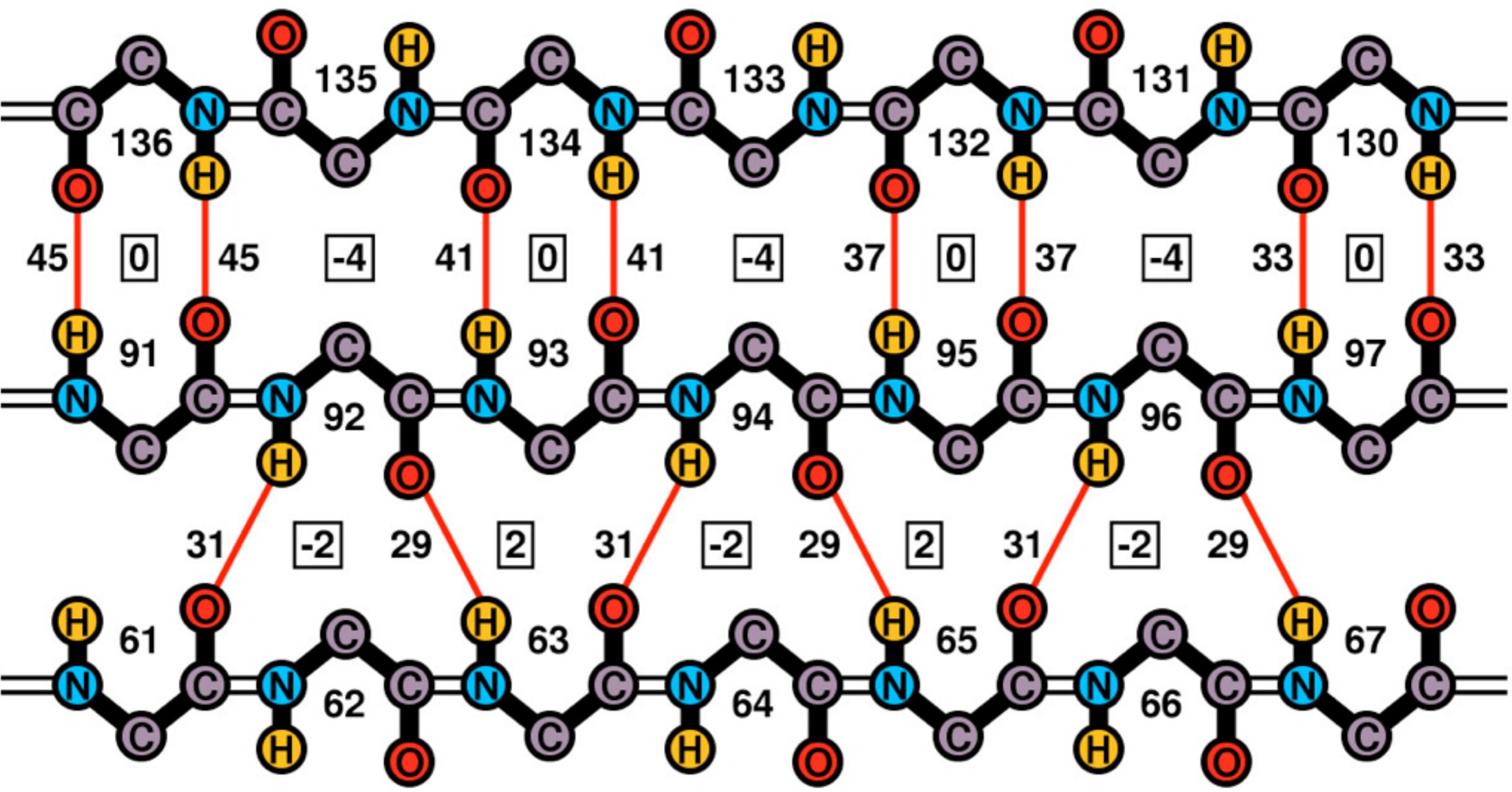


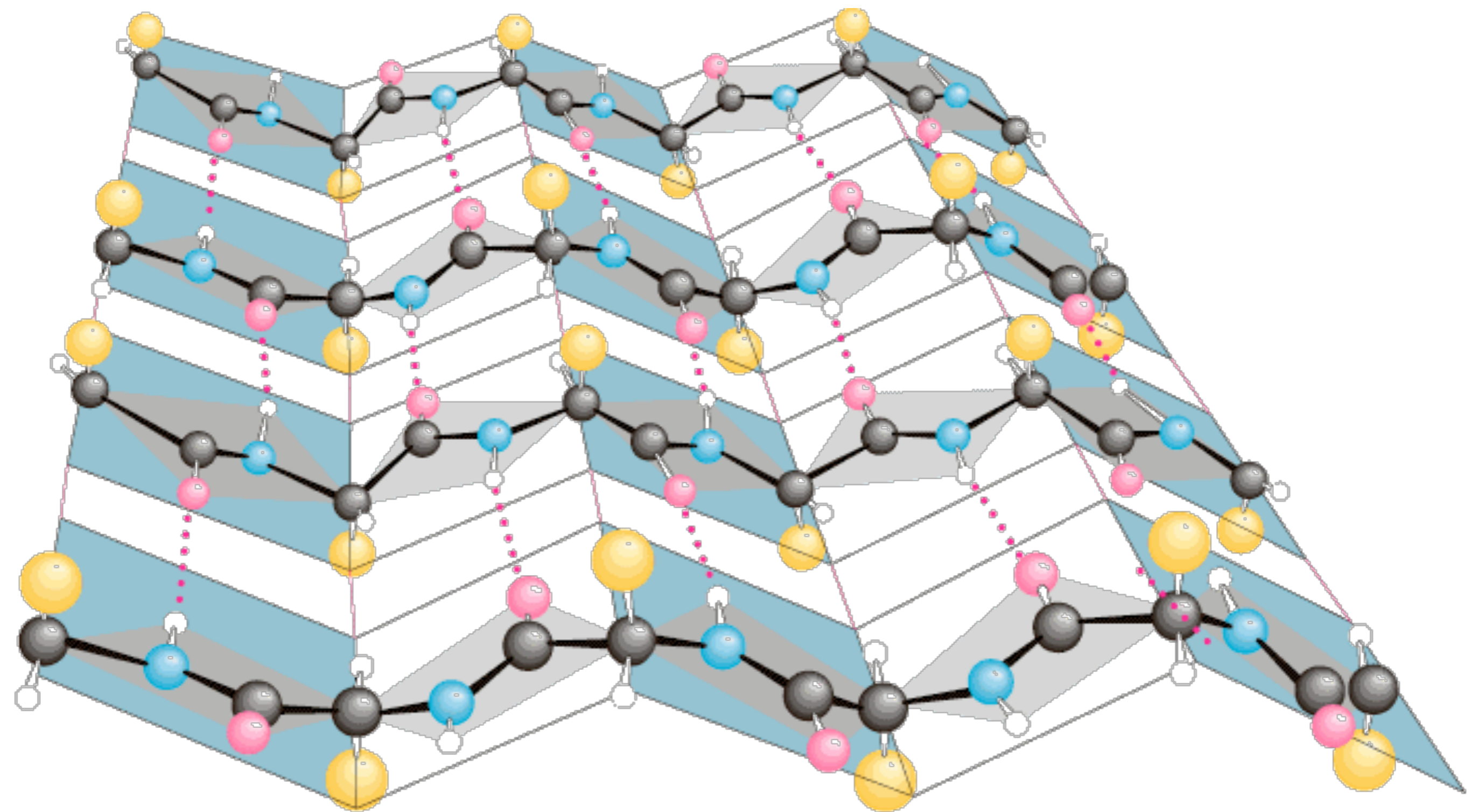
Beta strands & sheets



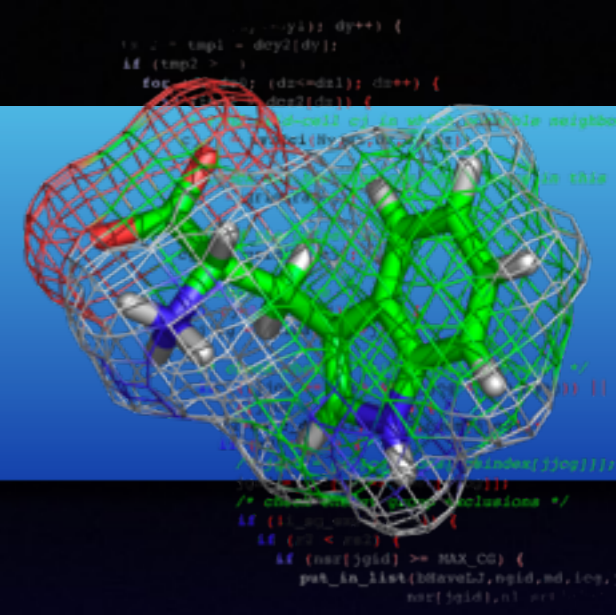
- How is this different from helices?
- Interaction patterns?
- Where are side chains pointing?
- Can you think of differences for the folding/formation?



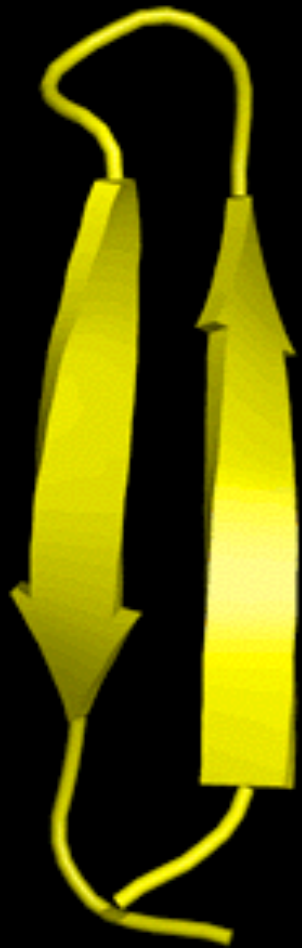




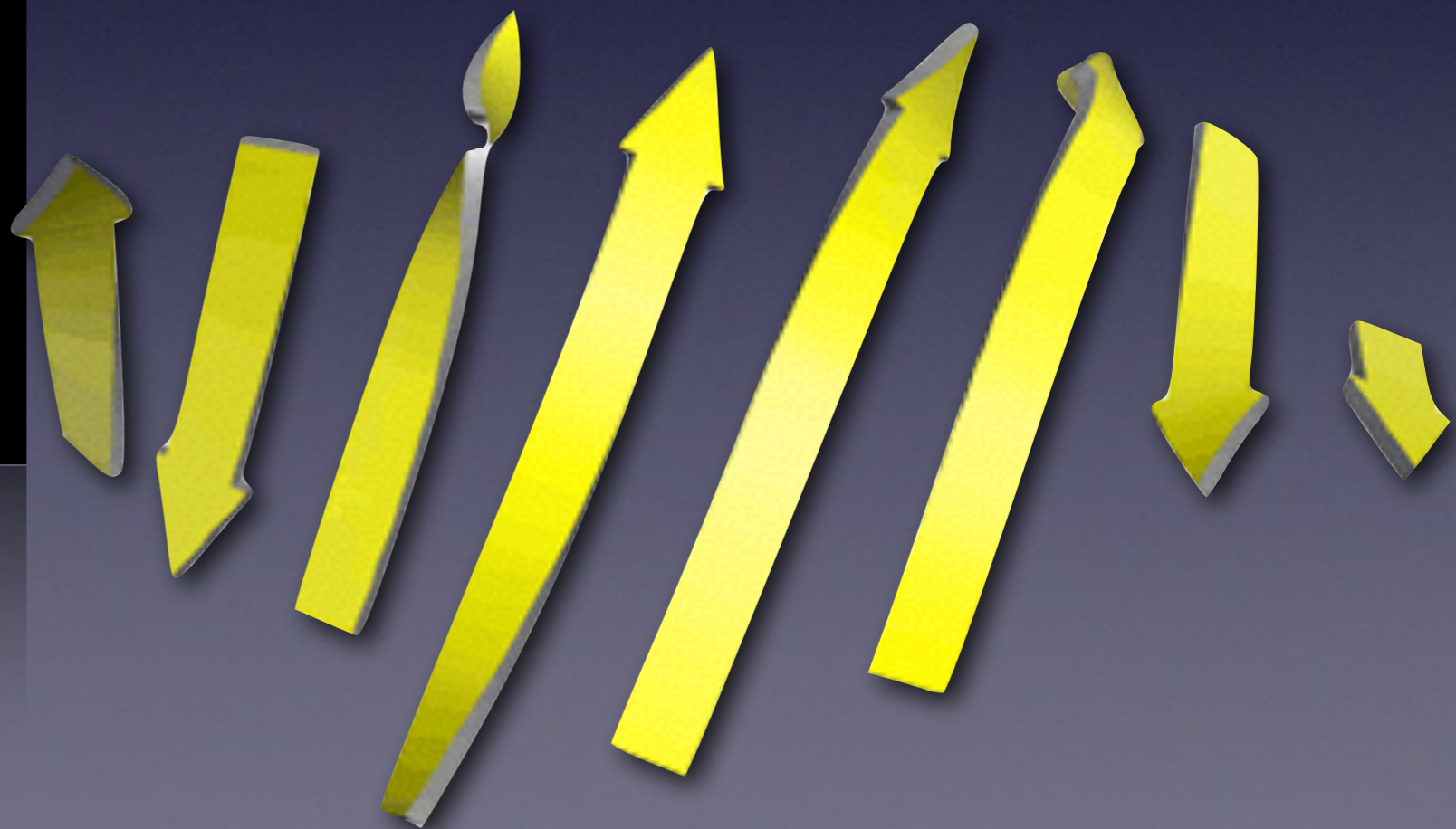
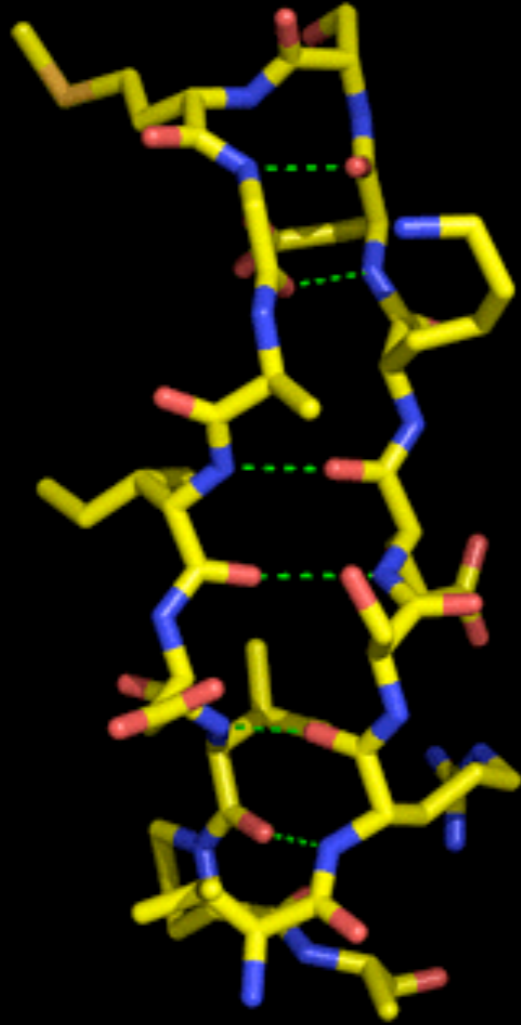
Beta twisting



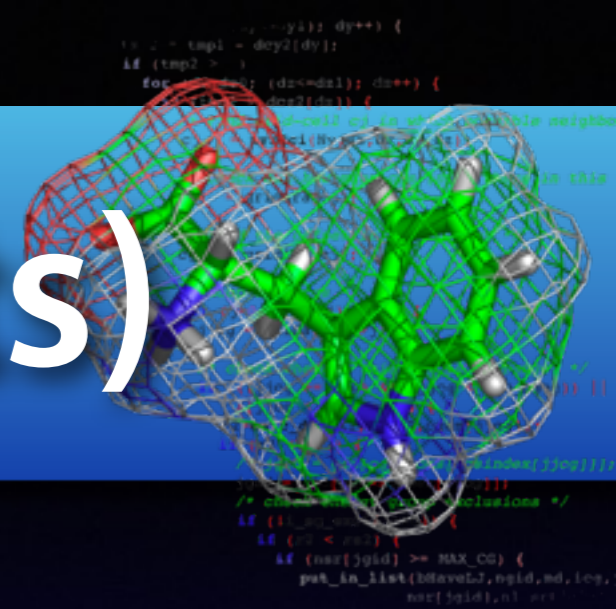
a.



b.

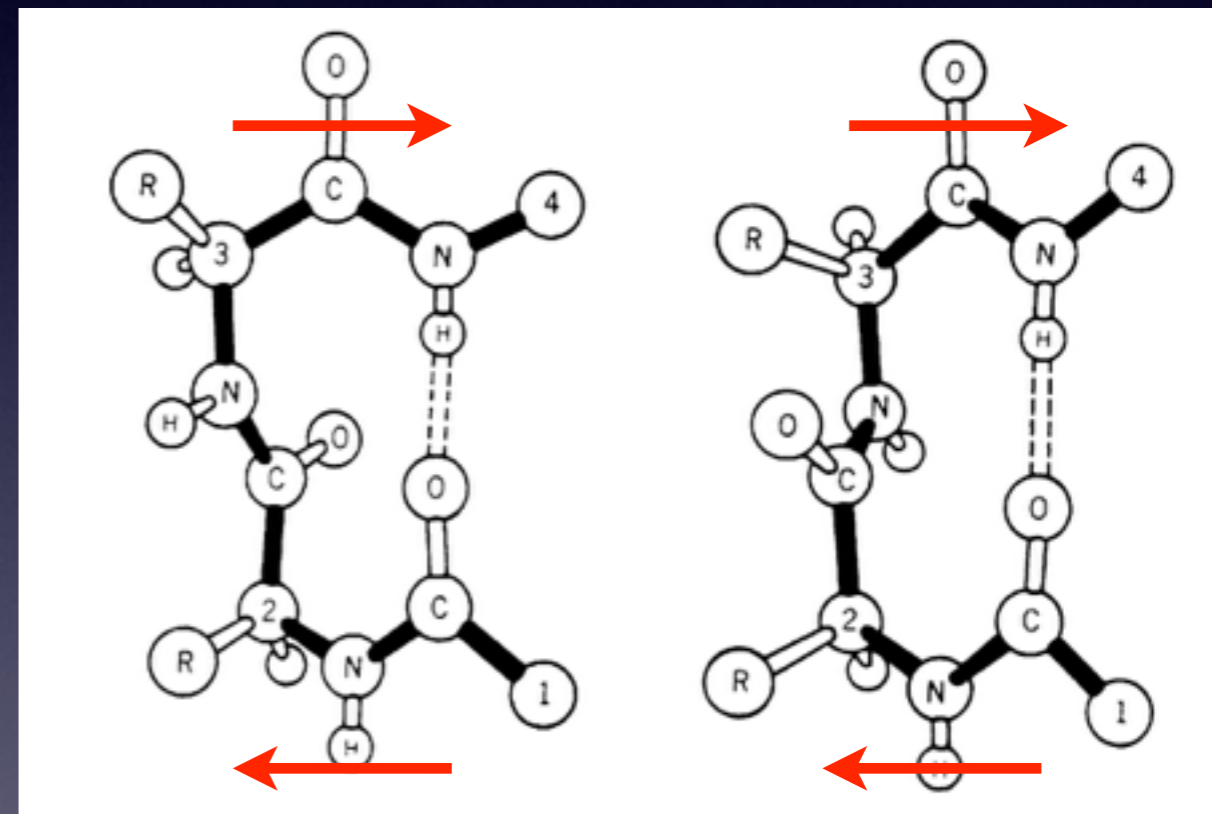


Tight turns (in sheets)



Venkatachalam, 1968 (models)
Simple steric repulsion

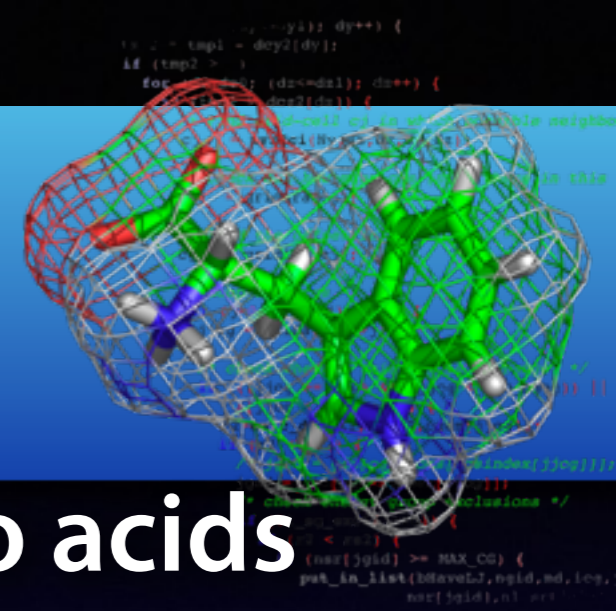
Type	$\varphi(i+1)$	$\psi(i+1)$	$\varphi(i+2)$	$\psi(i+2)$
I	-60	-30	-90	0
I'	60	30	90	0
II	-60	120	80	0
II'	60	-120	-80	0
IV	-61	10	-53	17
Vla1	-60	120	-90	0
Vla2	-120	120	-60	0
Vlb	-135	135	-75	160
VIII	-60	-30	-120	120



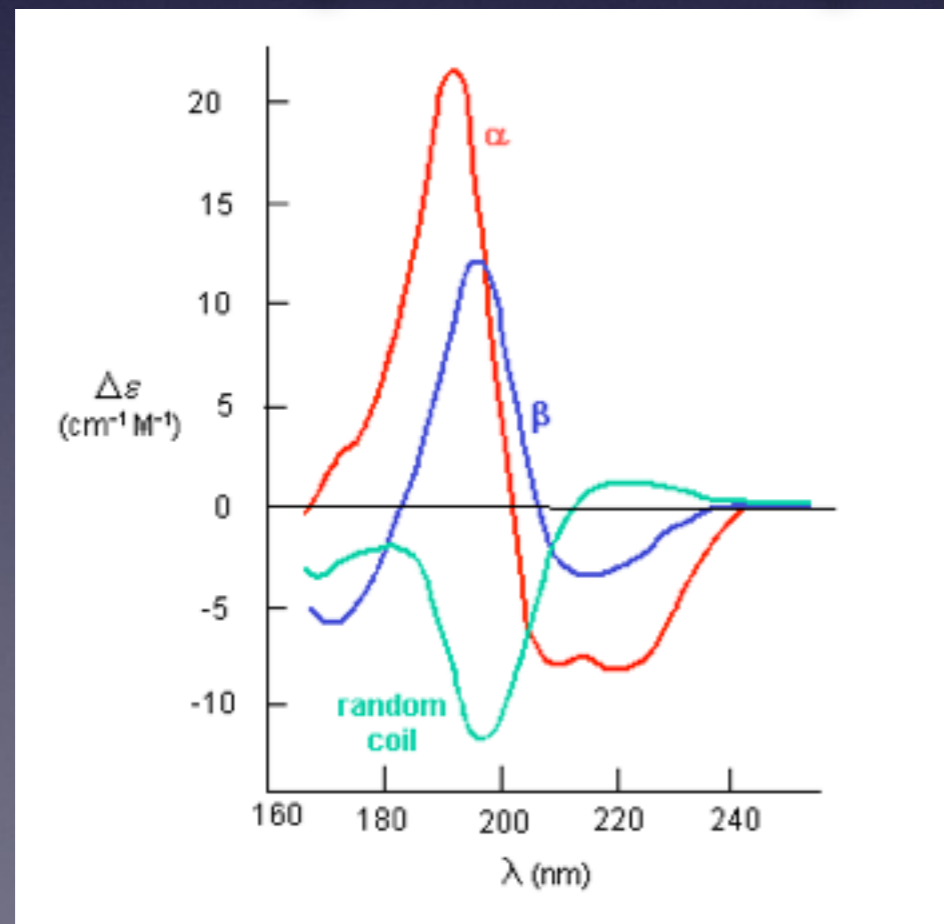
Type I

Type II

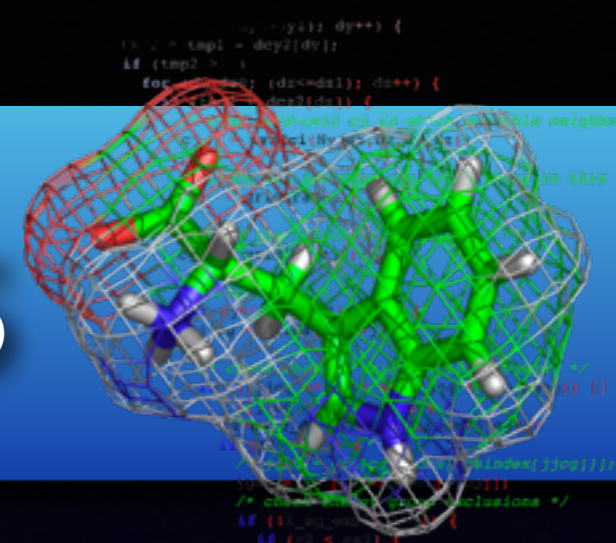
CD spectroscopy



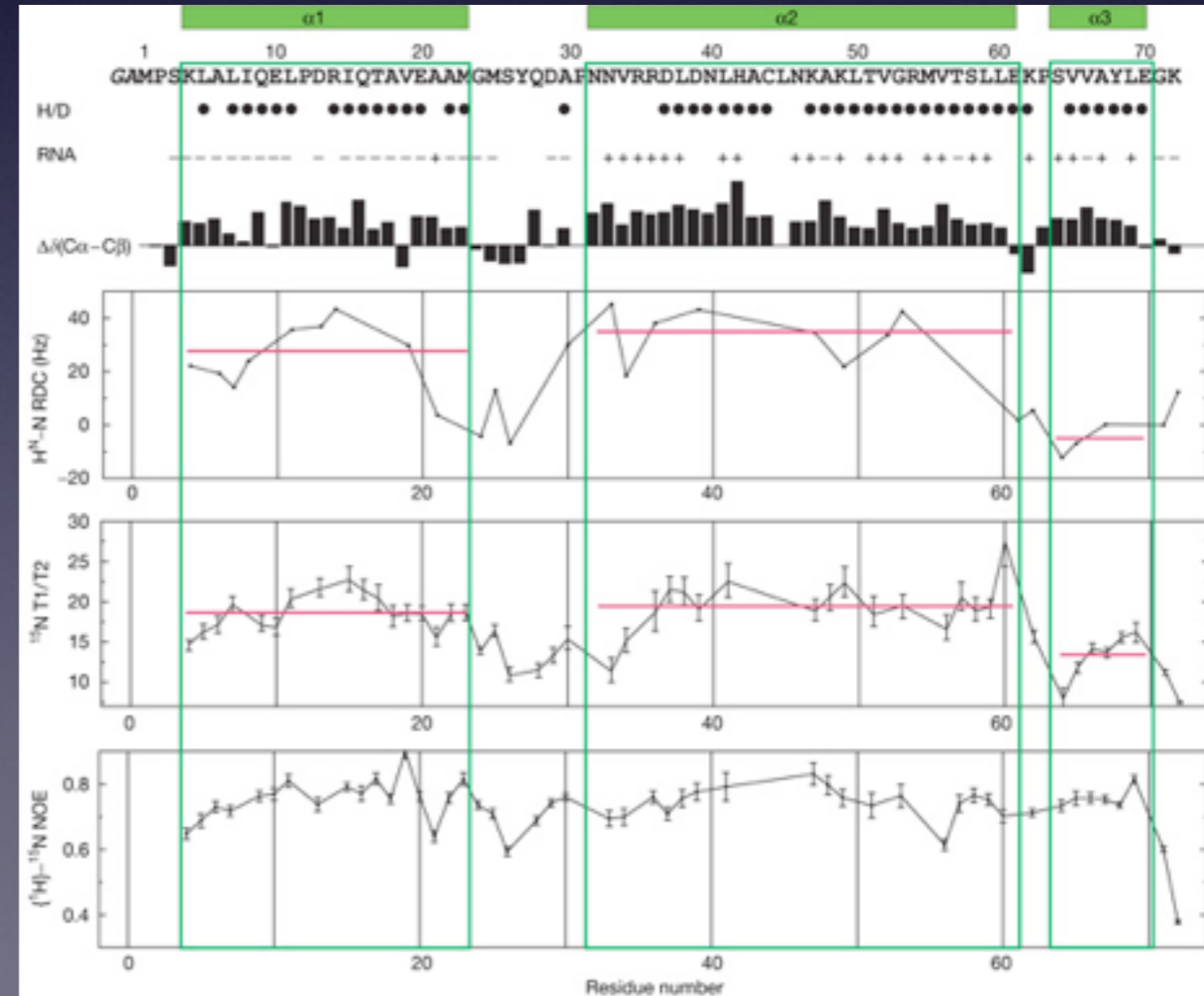
- Circular dichroism - chirality of amino acids will rotate polarized light
- Amount depends on the environment
- Cheap, fast, simple, no sequence resolution



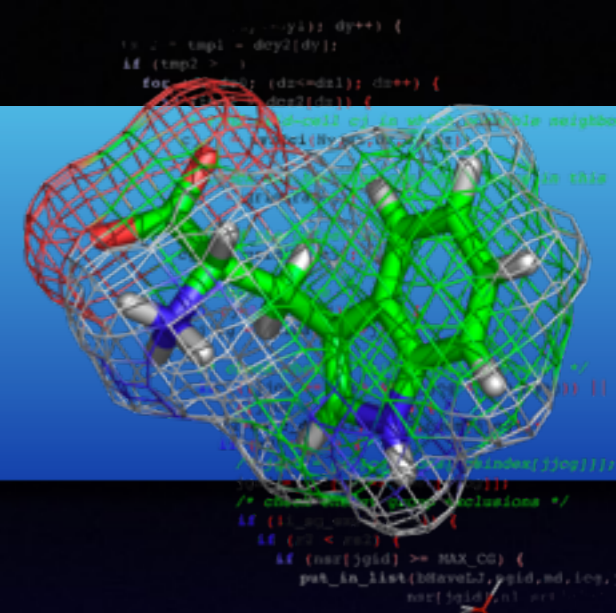
NMR chemical shifts



- Environment will shift frequency of nuclear spin resonance - 'chemical shifts'
- More complex than CD, but sequence resolved



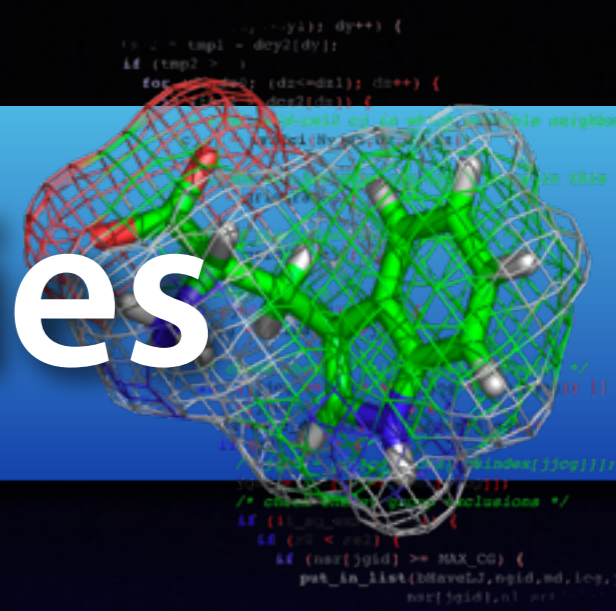
Helices vs. sheets



- **Helix**
 - Local h-bonds
 - Gradual (but fast) growth
 - Low initiation barrier
- **Sheets**
 - Non-local h-bonds
 - Collective interactions; all-or-nothing
 - High initiation barrier - very slow formation
- Next week: Phase/folding transitions!

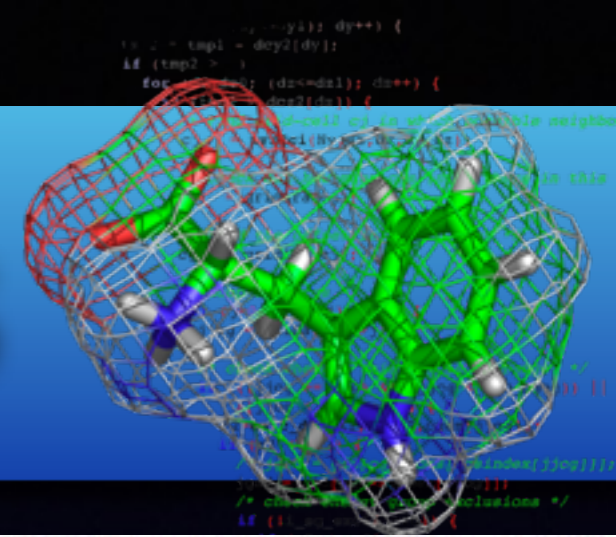


Amino acid properties

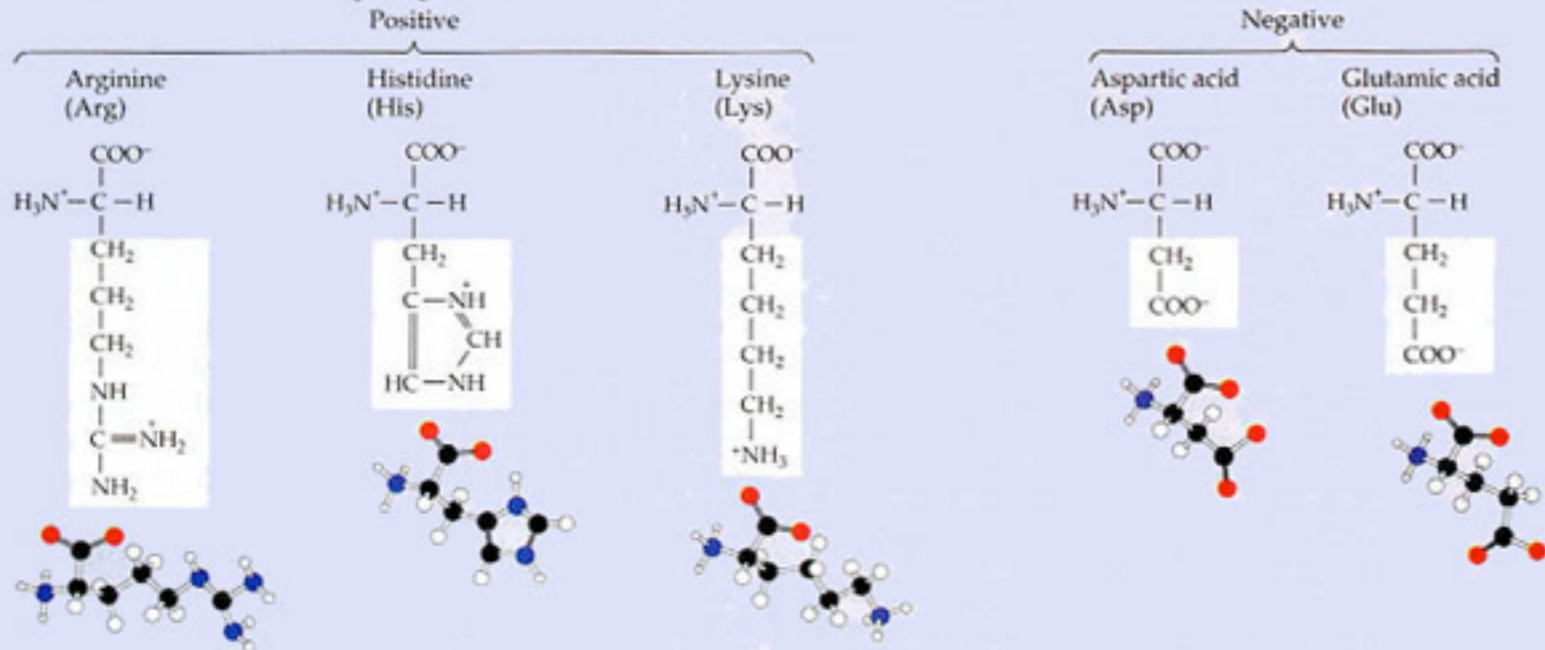


- All amino acids are not equal
 - Proline is very rare in alpha helices
 - Glycine is common in tight turns
 - Some residues common at helix ends
 - Differences inside/surface of proteins
- What is the cause of these differences, and can it be useful?

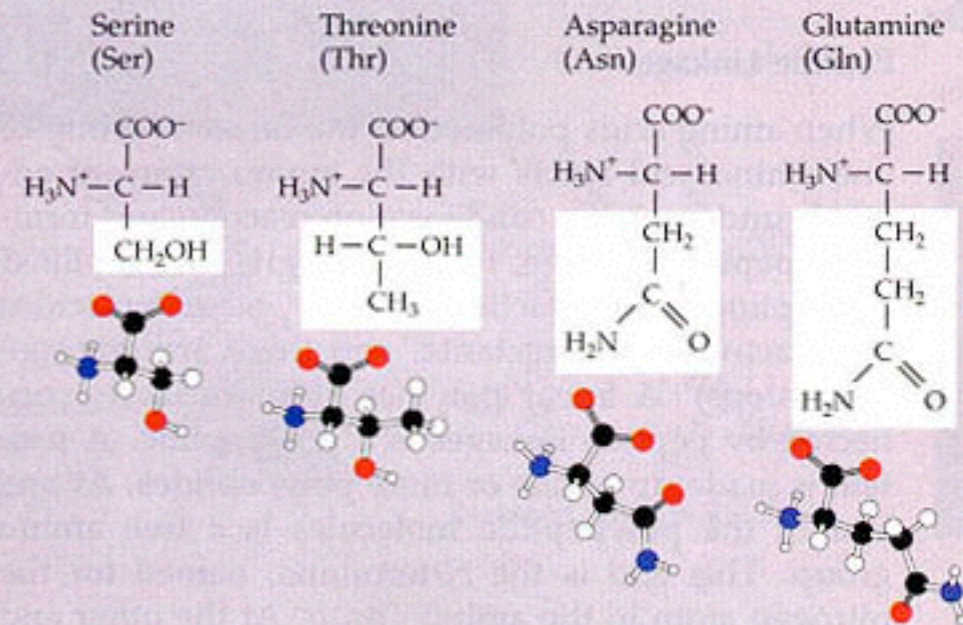
Natural amino acids



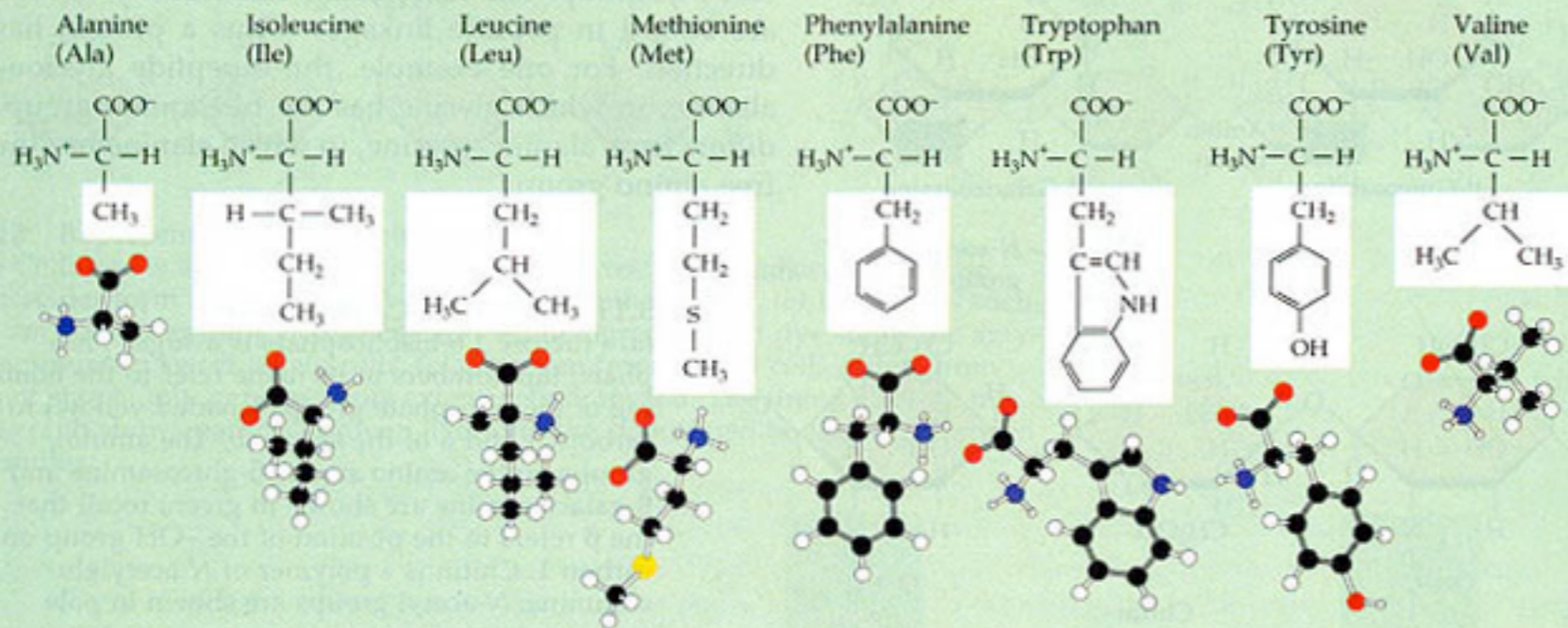
A. Amino acids with electrically charged side chains



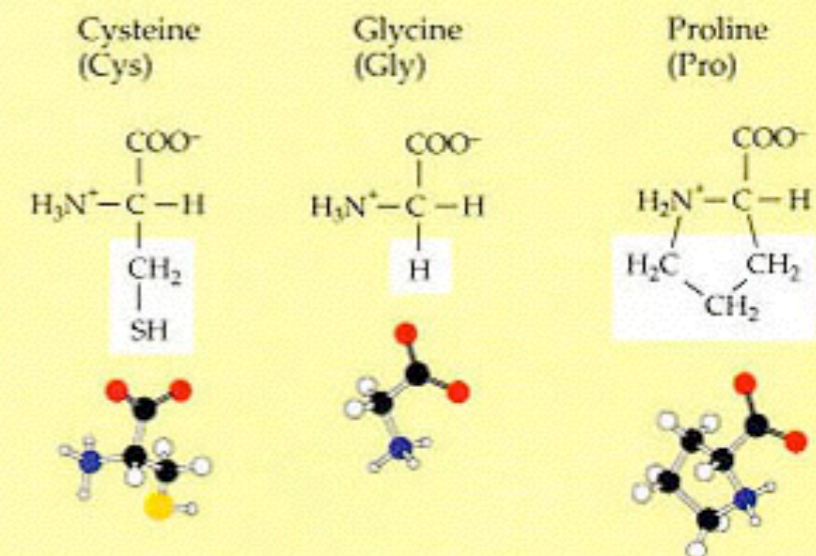
B. Amino acids with polar but uncharged side chains



D. Amino acids with hydrophobic side chains



C. Special cases

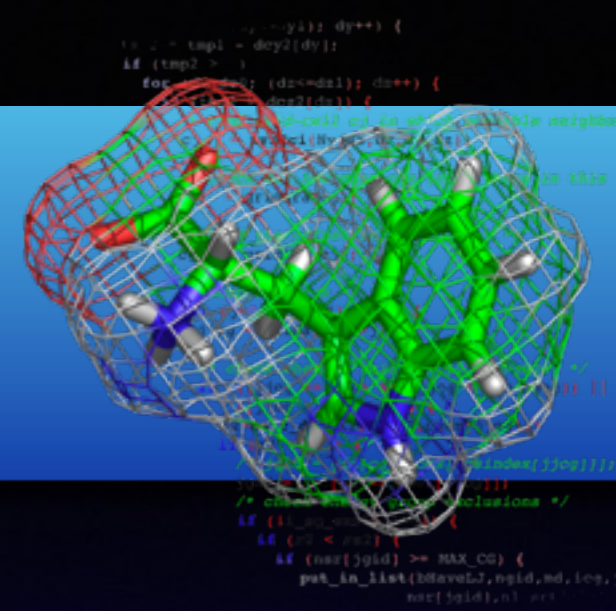


Name	3-letter code	1-letter code	Abundance	ΔG solvation
Glycine	GLY	G	6,89%	
Alanine	ALA	A	7,34%	1,94
Proline	PRO	P	5%	
Glutamic acid	GLU	E	6,22%	-79,12
Glutamine	GLN	Q	3,96%	-9,38
Aspartic acid	ASP	D	5,12%	-80,65
Asparagine	ASN	N	4,57%	-9,7
Serine	SER	S	7,38%	-5,06
Histidine	HIS	H	2,26%	-10.27/-64.13
Lysine	LYS	K	5,81%	-69,24
Arginine	ARG	R	5,2%	~ -60
Threonine	THR	T	5,85%	-4,88
Valine	VAL	V	6,48%	1,99
Isoleucine	ILE	I	5,76%	2,15
Leucine	LEU	L	9,36%	2,28
Metionine	MET	M	2,32%	-1,48
Phenylalanine	PHE	F	4,12%	-0,76
Tyrosine	TYR	Y	3,25%	-6,11
Cysteine	CYS	C	1,76%	-1,24
Tryptophan	TRP	W	1,34%	-5,88
GLU or GLN	GLX	Z (= E OR Q)		
ASP or ASN	ASX	B (= D OR N)		
Any amino acid	XXX	X		(kcal/mol)

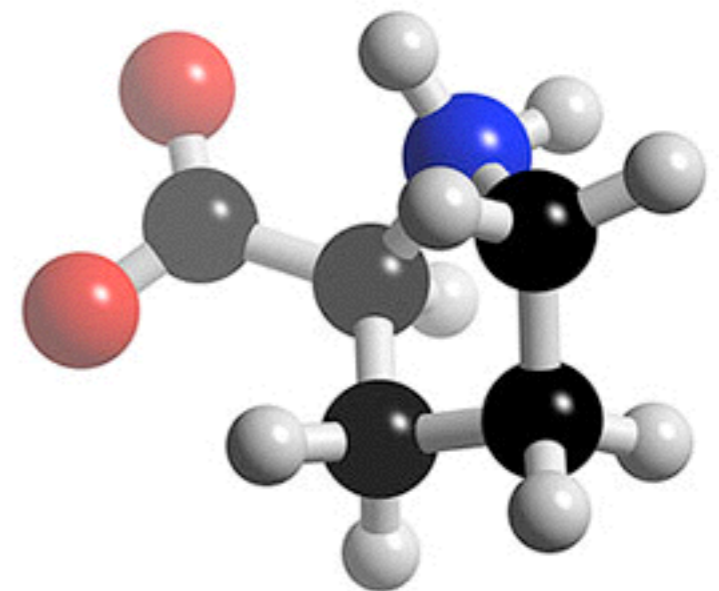
The Genetic Code

	U	C	A	G				
U	<p>UUU <u>Ph</u>enylalanine</p> <p>UUC <u>alanine</u></p> <p>UUG <u>Leu</u>cine</p> <p>UUA <u>Leu</u>cine</p>	<p>UCU <u>Ser</u>ine</p> <p>UCC <u>Ser</u>ine</p> <p>UCA <u>Ser</u>ine</p> <p>UCG <u>Ser</u>ine</p>	<p>UAU <u>Tyr</u>osine</p> <p>UAC <u>Tyr</u>osine</p> <p>UAA Stop</p> <p>UAG Stop</p>	<p>UGU <u>Cys</u>teine</p> <p>UGC <u>Cys</u>teine</p> <p>UGA Stop</p> <p>UGG <u>Tryp</u>tophan</p>	U	C	A	G
C	<p>CUU <u>Leu</u>cine</p> <p>CUC <u>Leu</u>cine</p> <p>CUA <u>Leu</u>cine</p> <p>CUG <u>Leu</u>cine</p>	<p>CCU <u>Pro</u>line</p> <p>CCC <u>Pro</u>line</p> <p>CCA <u>Pro</u>line</p> <p>CCG <u>Pro</u>line</p>	<p>CAU <u>His</u>tidine</p> <p>CAC <u>His</u>tidine</p> <p>CAA <u>Glut</u>amine</p> <p>CAG <u>Glut</u>amine</p>	<p>CGU <u>Arg</u>inine</p> <p>CGC <u>Arg</u>inine</p> <p>CGA <u>Arg</u>inine</p> <p>CGG <u>Arg</u>inine</p>	U	C	A	G
A	<p>AUU <u>Iso</u>leucine</p> <p>AUC <u>Iso</u>leucine</p> <p>AUA <u>Iso</u>leucine</p> <p>AUG <u>Met</u>hionine</p>	<p>ACU <u>Thr</u>eonine</p> <p>ACC <u>Thr</u>eonine</p> <p>ACA <u>Thr</u>eonine</p> <p>ACG <u>Thr</u>eonine</p>	<p>AAU <u>Asp</u>aragine</p> <p>AAC <u>Asp</u>aragine</p> <p>AAA <u>Lys</u>ine</p> <p>AAG <u>Lys</u>ine</p>	<p>AGU <u>Ser</u>ine</p> <p>AGC <u>Ser</u>ine</p> <p>AGA <u>Arg</u>inine</p> <p>AGG <u>Arg</u>inine</p>	U	C	A	G
G	<p>GUU <u>Val</u>ine</p> <p>GUC <u>Val</u>ine</p> <p>GUA <u>Val</u>ine</p> <p>GUG <u>Val</u>ine</p>	<p>GCU <u>Ala</u>nine</p> <p>GCC <u>Ala</u>nine</p> <p>GCA <u>Ala</u>nine</p> <p>GCG <u>Ala</u>nine</p>	<p>GAU <u>Asp</u>artic acid</p> <p>GAC <u>Asp</u>artic acid</p> <p>GAA <u>Glut</u>amic acid</p> <p>GAG <u>Glut</u>amic acid</p>	<p>GGU <u>Gly</u>cine</p> <p>GGC <u>Gly</u>cine</p> <p>GGA <u>Gly</u>cine</p> <p>GGG <u>Gly</u>cine</p>	U	C	A	G

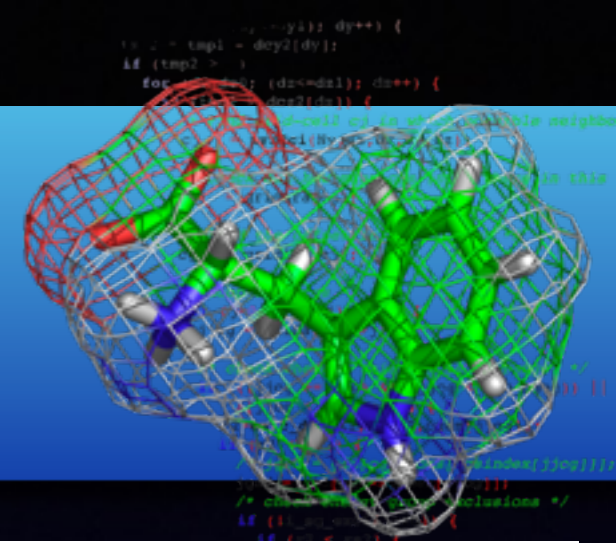
Proline



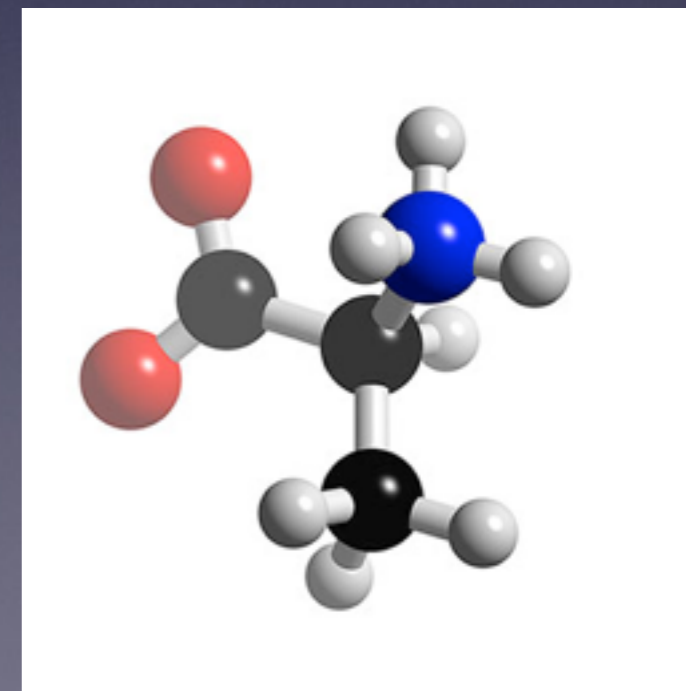
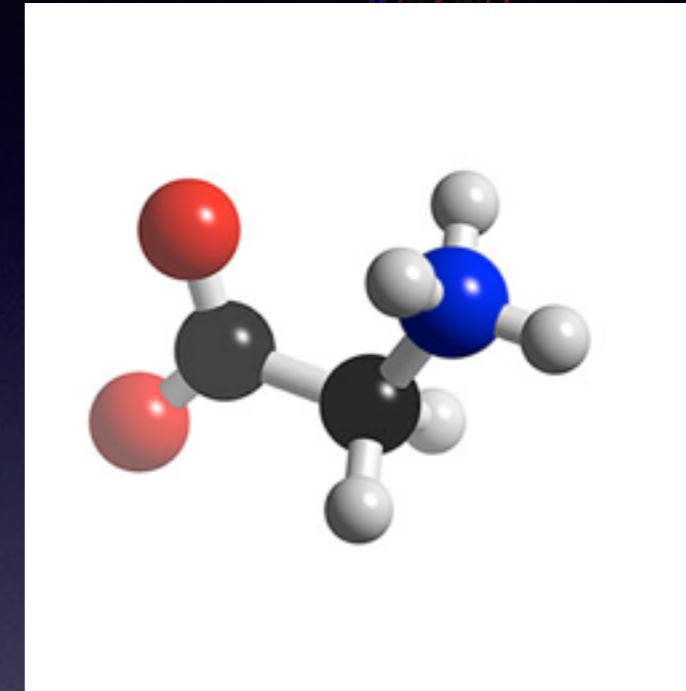
- **Proline:**
 - **Cannot form hydrogen bonds, bulky side-chain with two carbons connected to the backbone nitrogen atom**
 - **N-terminus of alpha helices**
 - **Turns**
 - **Normally not inside helices/sheets**



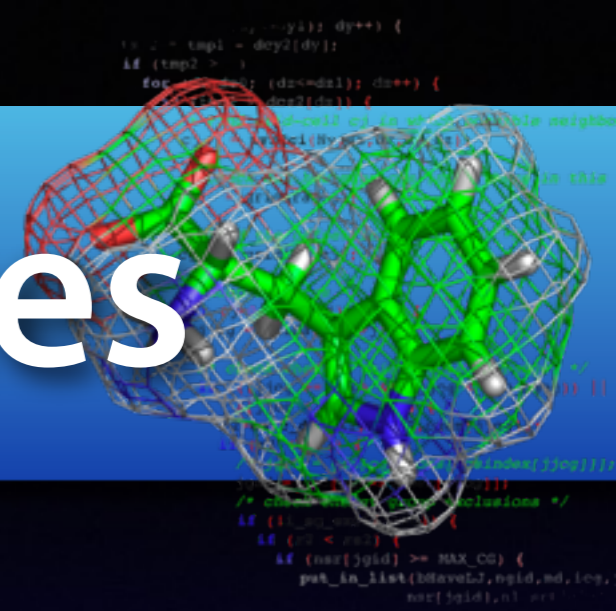
Glycine / Alanine



- Glycine
 - No side chain means no clashes
 - Flexible ramachandran map
 - Common in turns (flexible)
- Alanine
 - Methyl side chain
 - Slight helix preference, but sheet ok

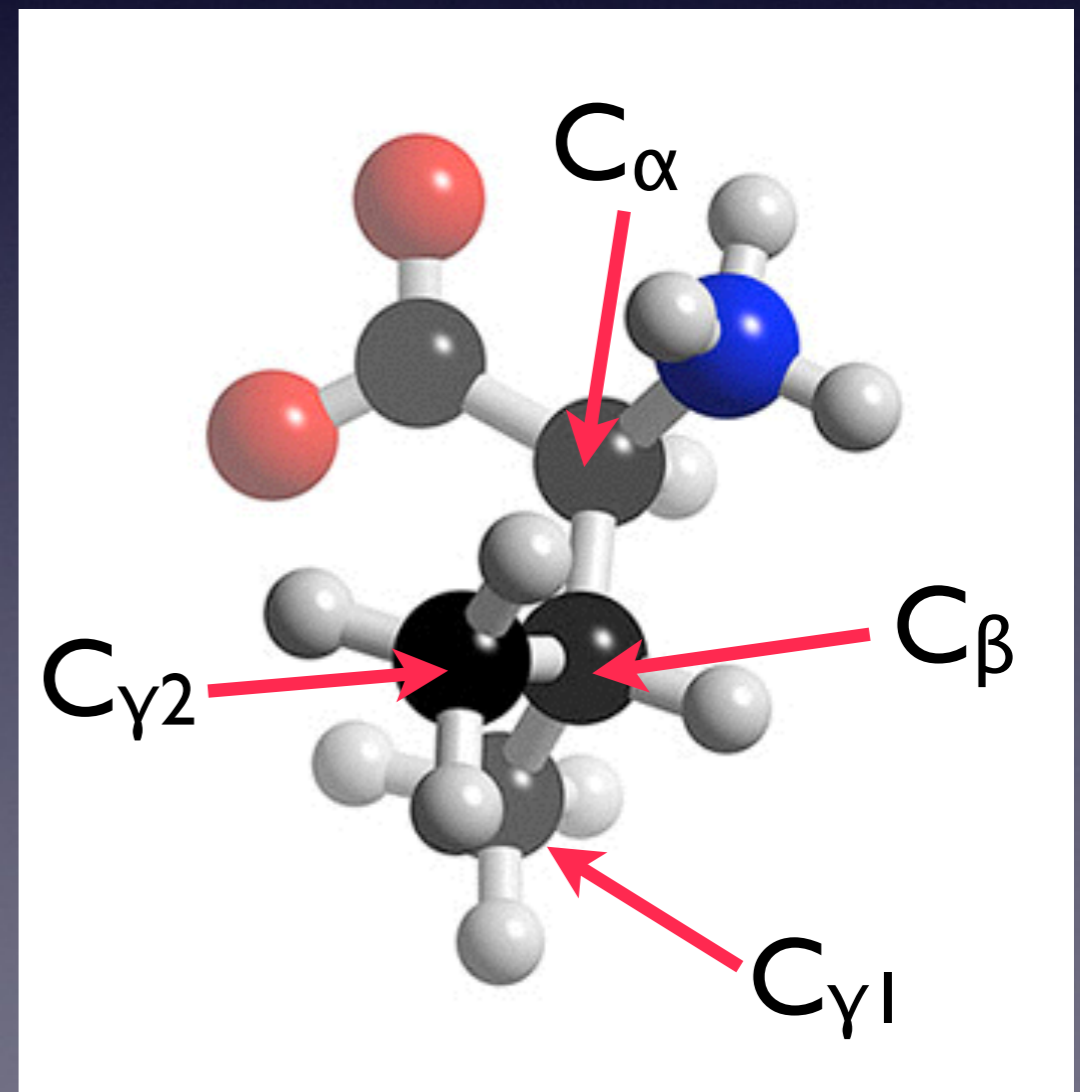


Hydrophobic residues

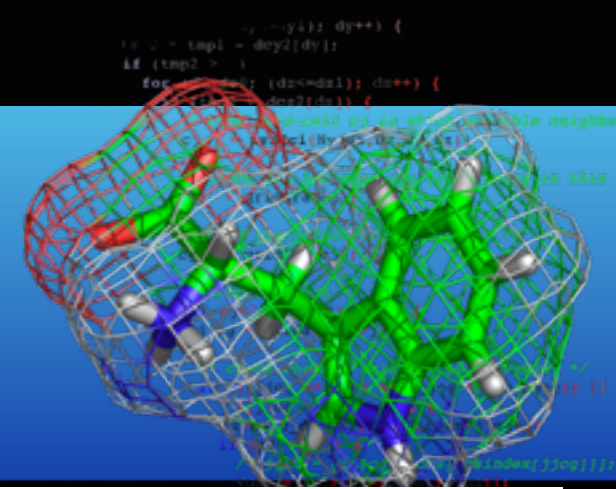


- Normally prefer beta sheets
- Side chains protrude on alternating sides
- More room for bulky side chains (often h-phobic)
- In particular residues with two γ carbons

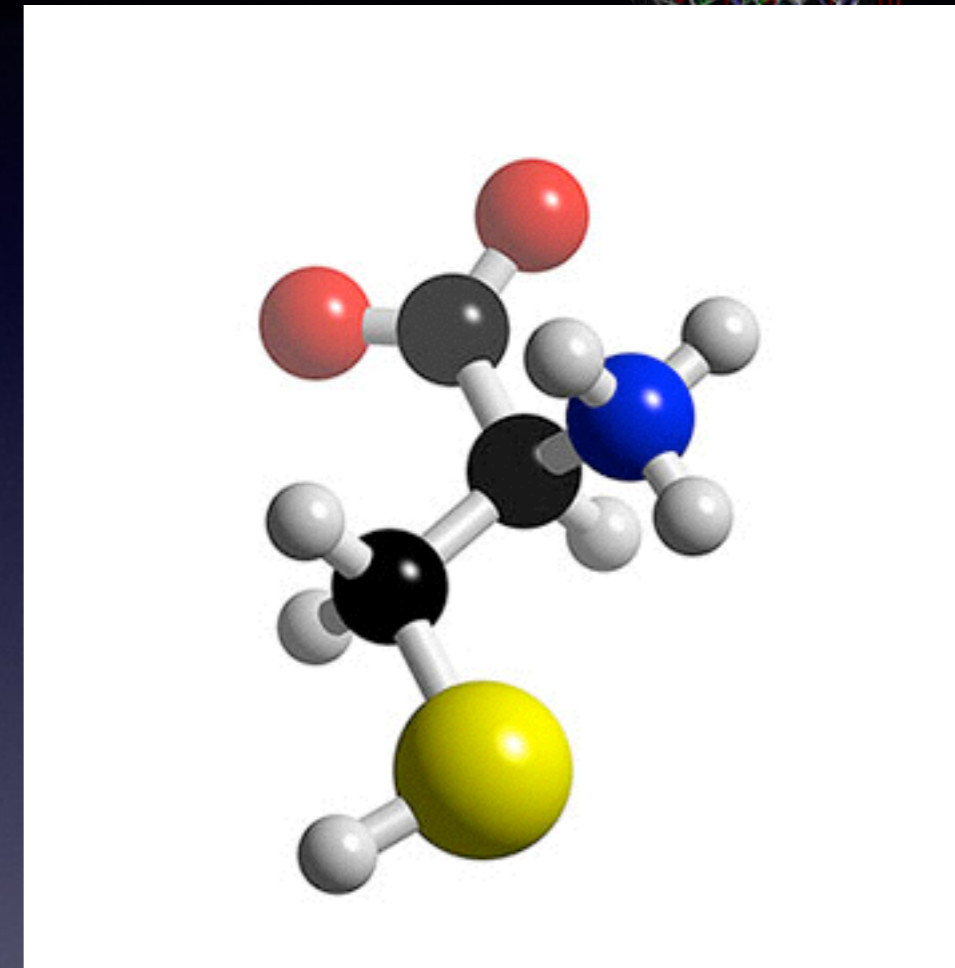
Labeling starts from backbone:
 $\alpha, \beta, \gamma, \delta, \epsilon, \zeta$



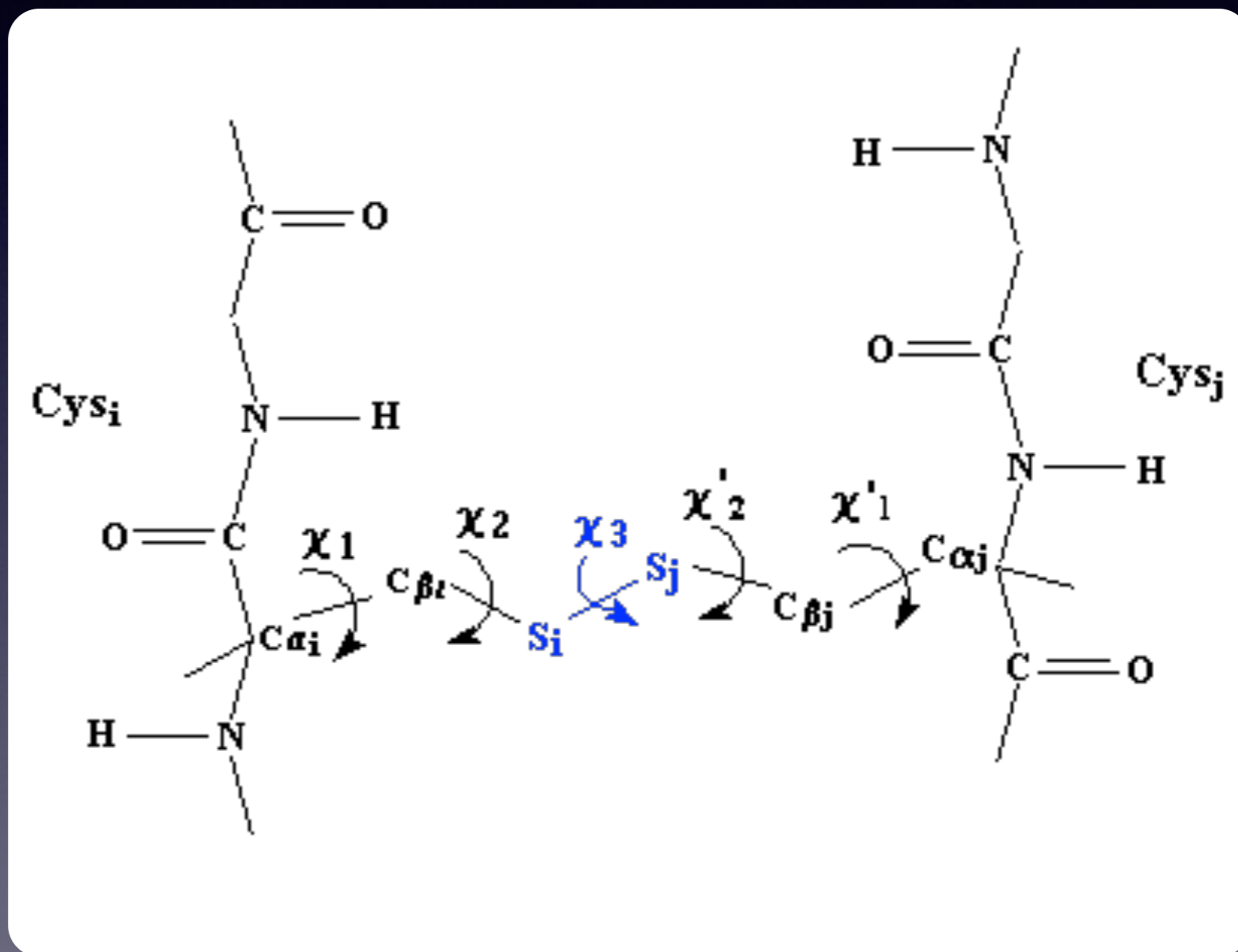
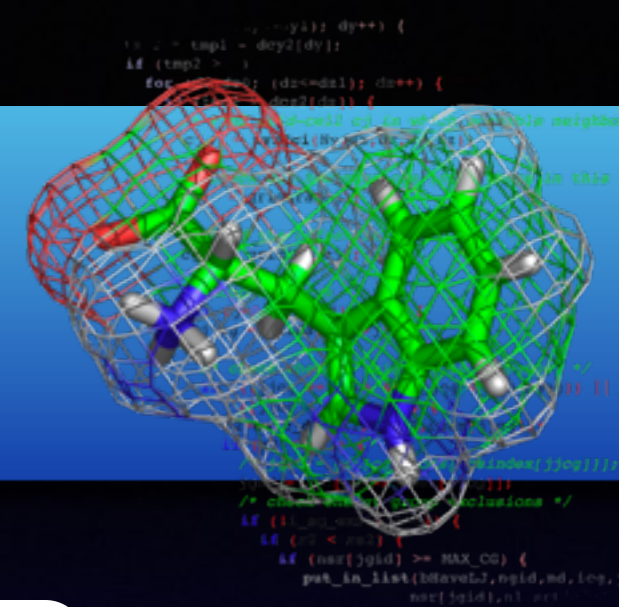
Cysteine



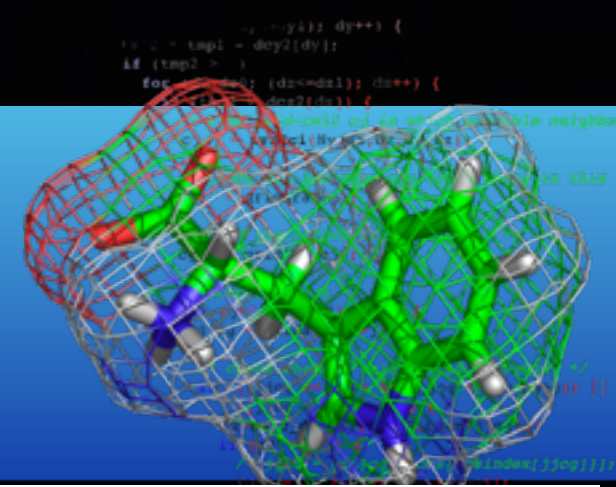
- Relatively small sidechain
 - Contains an -S-H group
 - Polar
- Two Cysteines can form a disulphide bond: -S-S-
- Very tight (covalent) bonds, harder than hydrogen bond
- Fixes structure in space



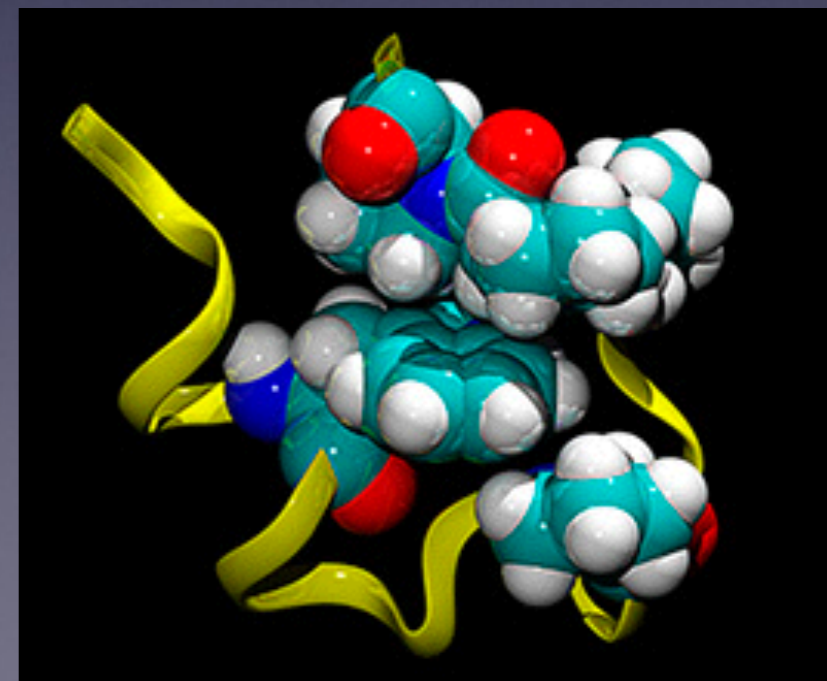
Disulphides

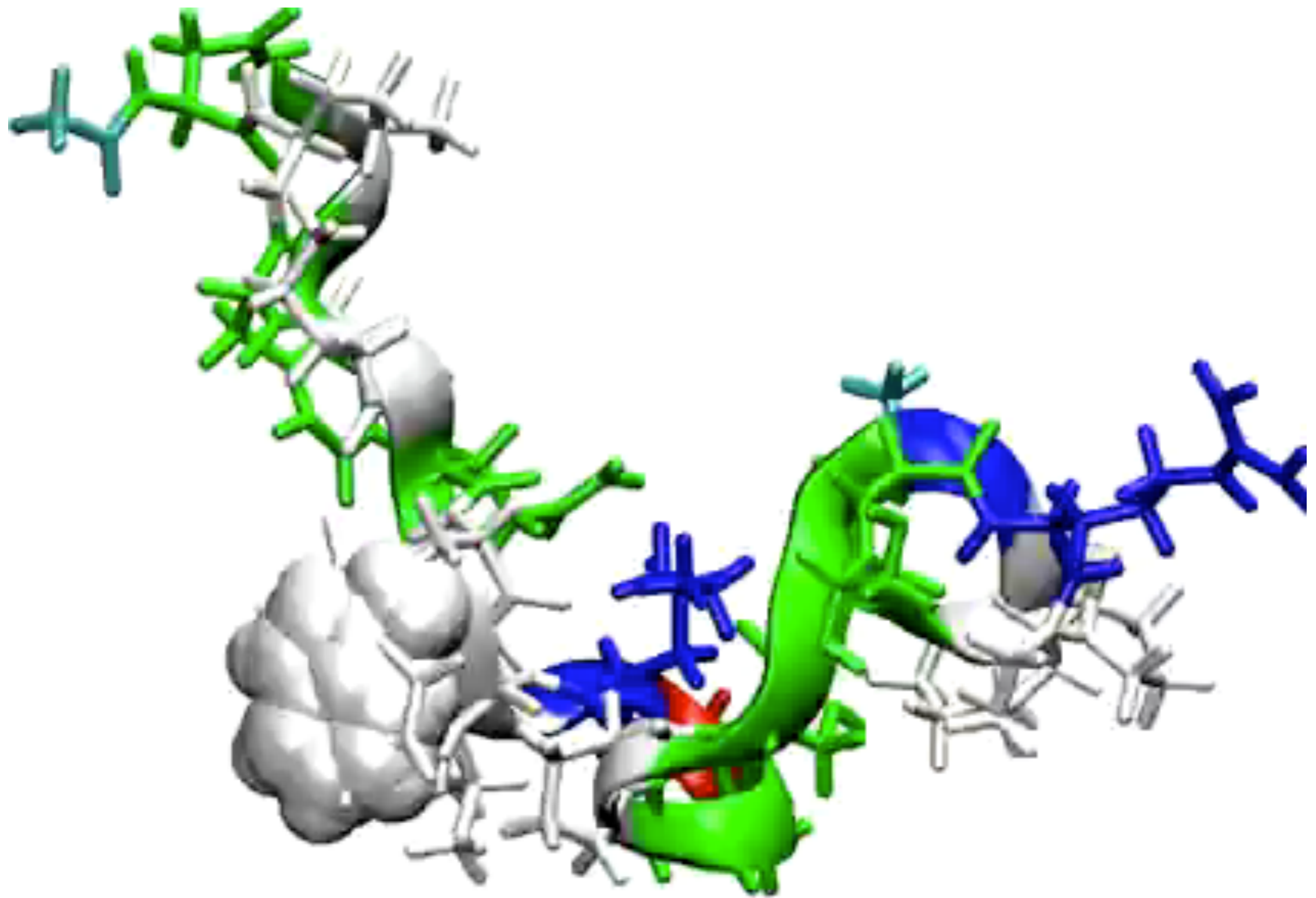


Trp: Big & bulky



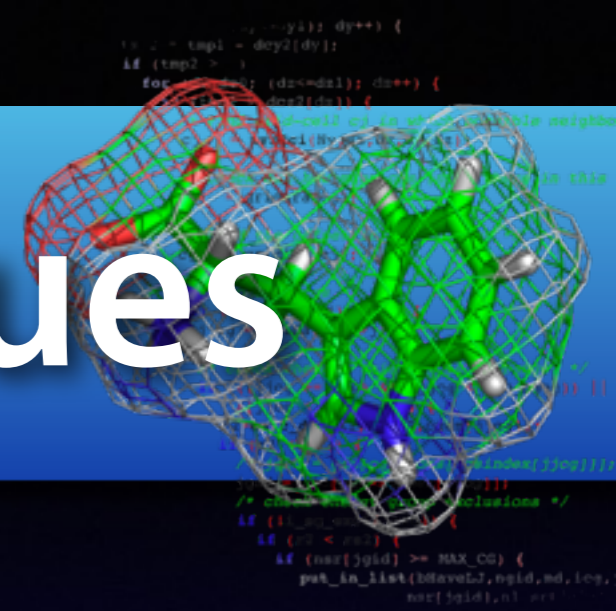
- Tryptophan
- Two rings
 - 5-member ring with indole group
 - Aromatic ring
- Large and stiff side chain
 - Difficult to pack
- World's smallest protein:
 - Trp Cage (Andersen 2002)





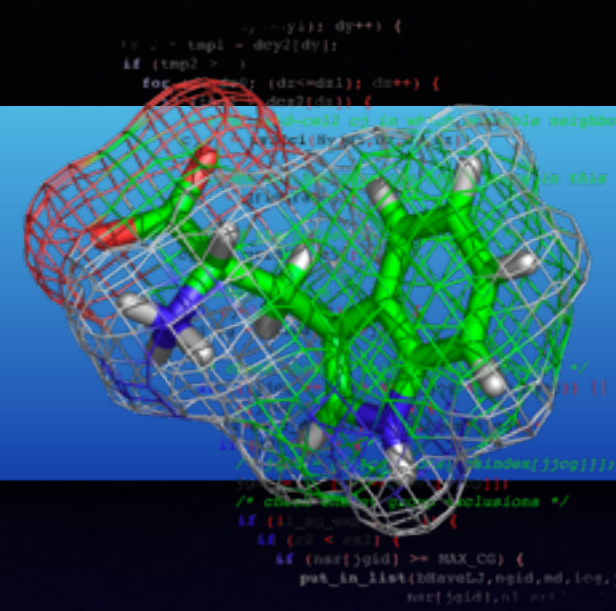
Paschek *et al.*

Polar/charged residues

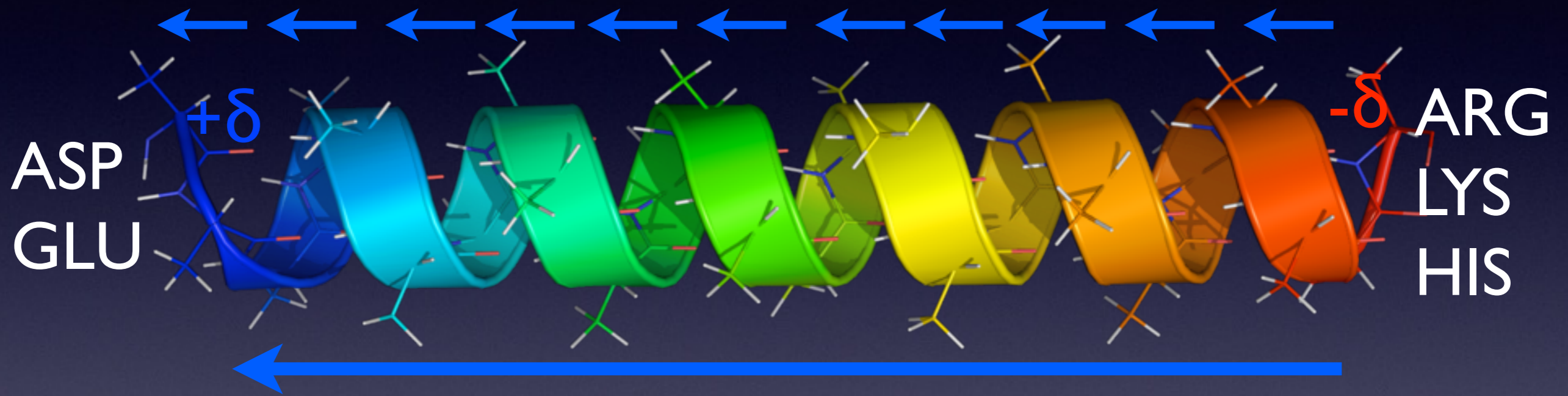


- **Polar:**
 - Prefers turn/loop regions
 - H-bonds to both water and the polypeptide chain
- **Charged:**
 - Occurs on surface, in active sites
 - Negative charges stabilize helix N-terminus
 - Positive charges stabilize helix C-terminus

Helix capping



Remember the helix dipole?



N-terminus

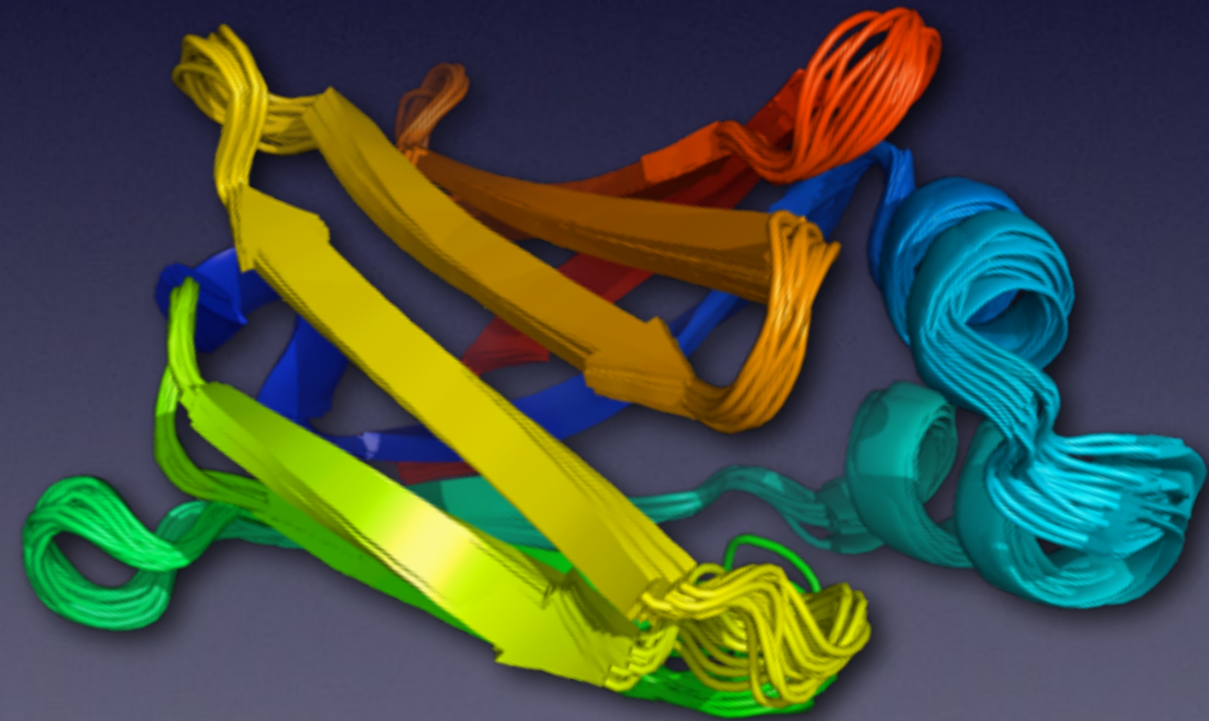
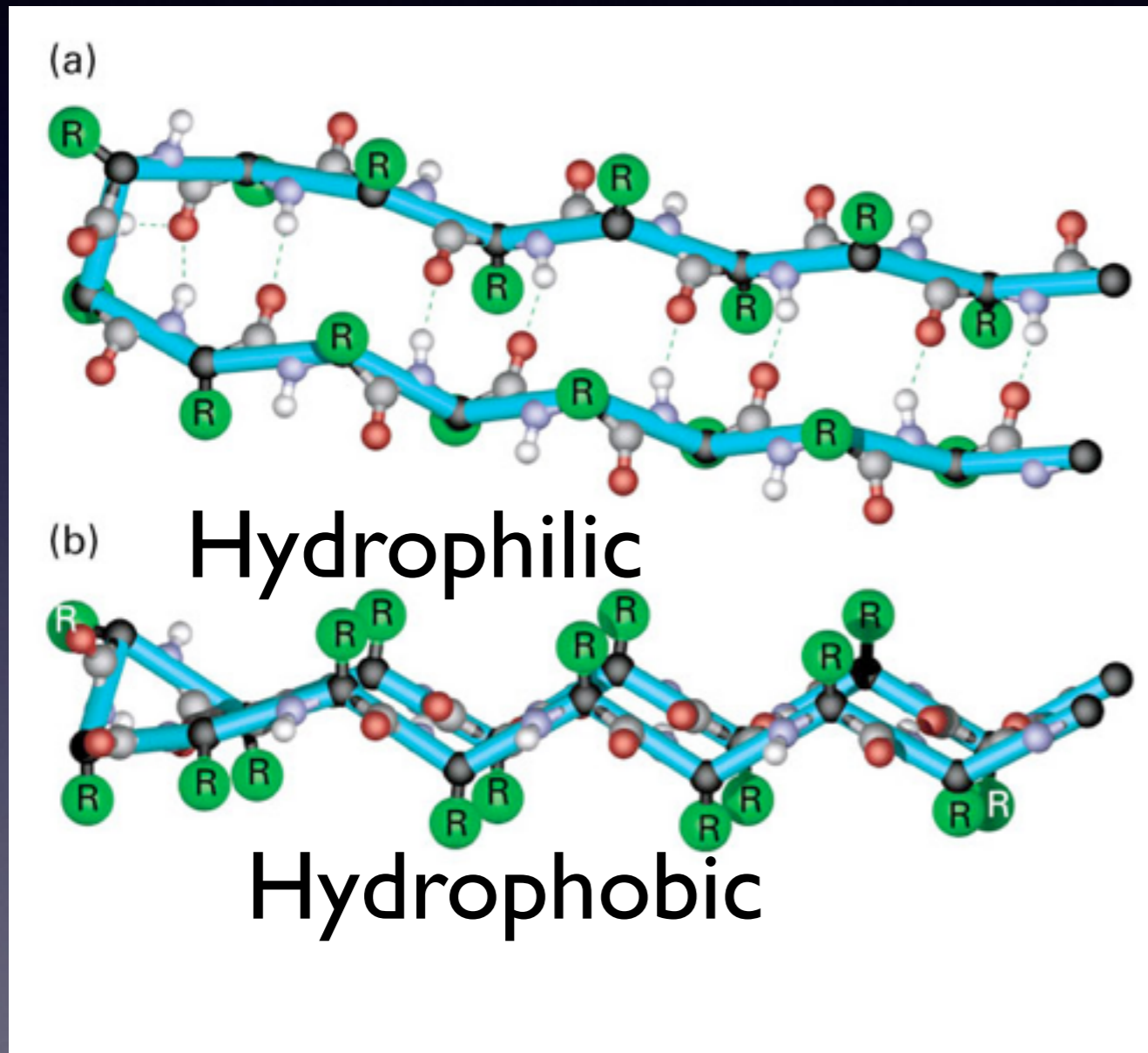
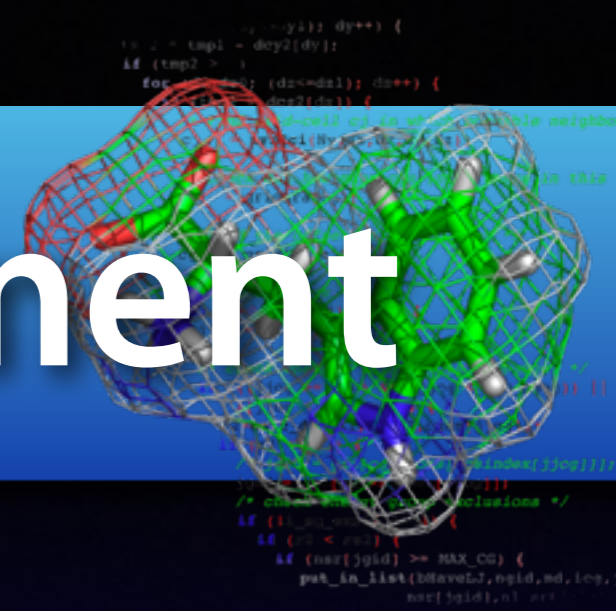
C-terminus

Charged residues act as 'caps' for the helix dipole, which stabilizes both the helix and the charged residue in that position

Residue	Main chain ^a NH	Side chain ^a		Dipole/charge ^b	pK _a ^b	Structural occurrence tendency ^c							
		C ^β	γ Number			before α _N	in helix			after α _C	in		
							α _N	α	α _C		β	loops	core
Gly	+	-					-				-	+	
Ala	+	+					+					-	
Pro	-	+	1				+	-	-	-	-	+	
Glu	+	+	1	COOH → CO ₂ ⁻	4.3	+	+		-	-	-		-
Asp	+	+	1	COOH → CO ₂ ⁻	3.9	+	+		-	-	-	-	+
Gln	+	+	1	OCNH ₂									-
Asn	+	+	1	OCNH ₂		+			-		+	-	+
Ser	+	+	1	OH		+						+	-
His	+	+	1	NH; and N ⇒ NH ⁺	6.5		-			+	+		
Lys	+	+	1	NH ₂ ⇒ NH ₃ ⁺	10.5	-	-			+	+	-	-
Arg	+	+	1	HNC(NH ₂) ₂ ⁺	12.5	-	-			+	+	-	+
Thr	+	+	2	OH		+						+	
Ile	+	+	2									+	-
Val	+	+	2									+	-
Leu	+	+	1						+			+	-
Met	+	+	1						+			+	-
Phe	+	+	1									+	-
Tyr	+	+	1	OH ⇒ O ⁻	10.1				-			+	
Cys	+	+	1	SH ⇒ S ⁻	9.2				-			+	
Trp	+	+	1	NH								+	

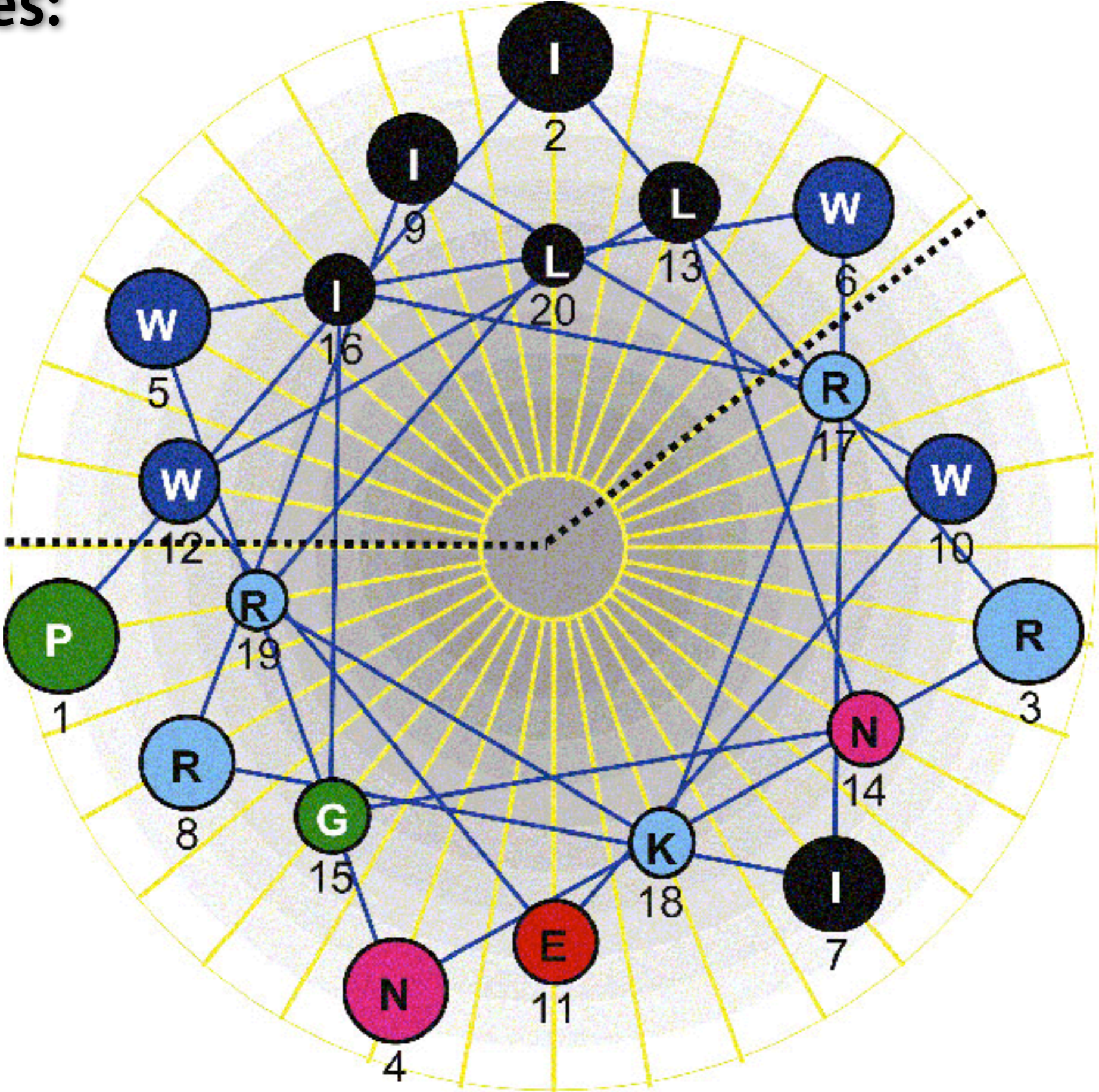
Amino acids tend to occur in places where they stabilize the structure!

Hydrophobicity moment



FABP: Water-soluble surface
Hydrophobic inside cavity

For helices:



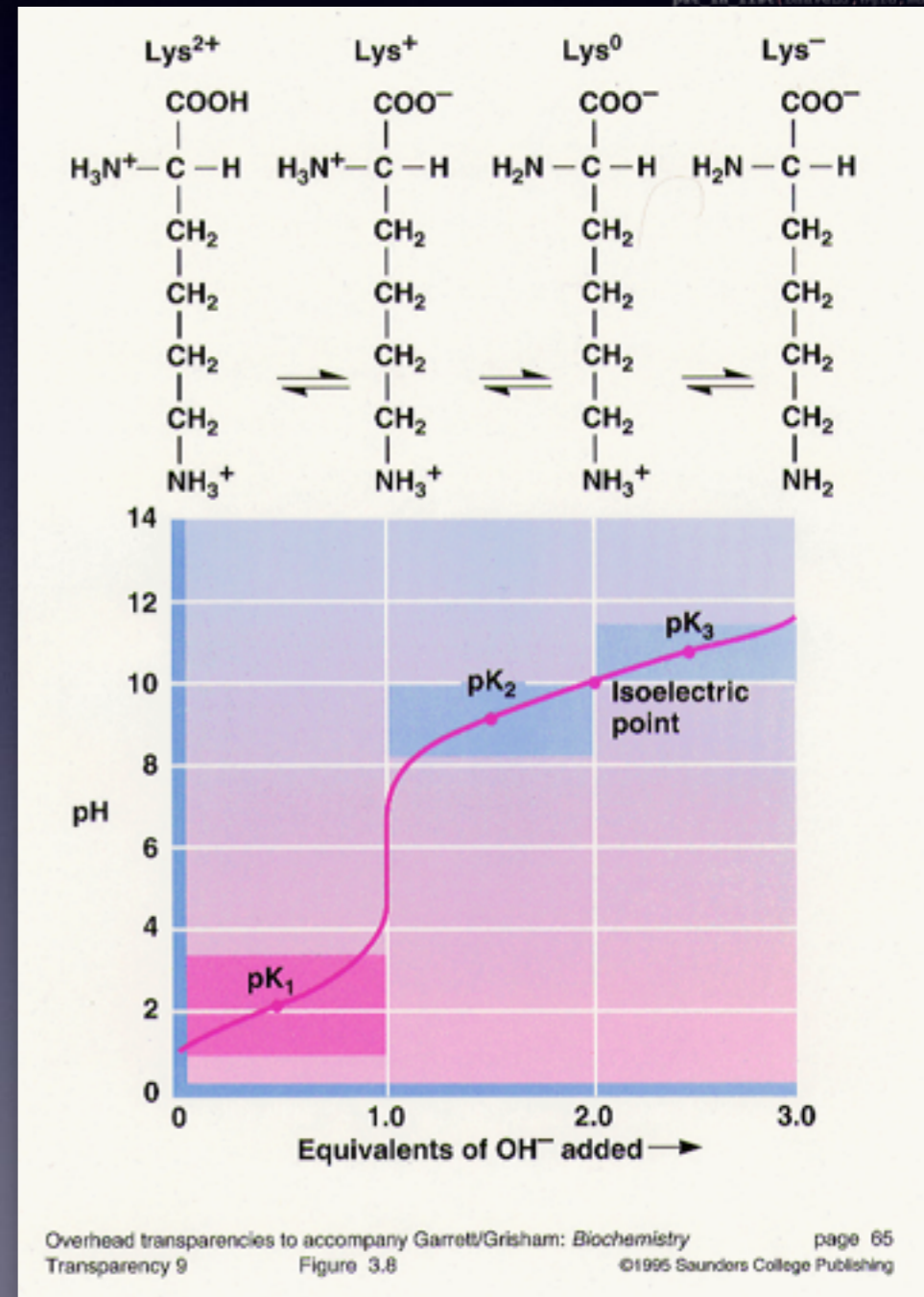
Titratable residues / pKa



- The protonation state of charged/polar amino acids depends on the current pH

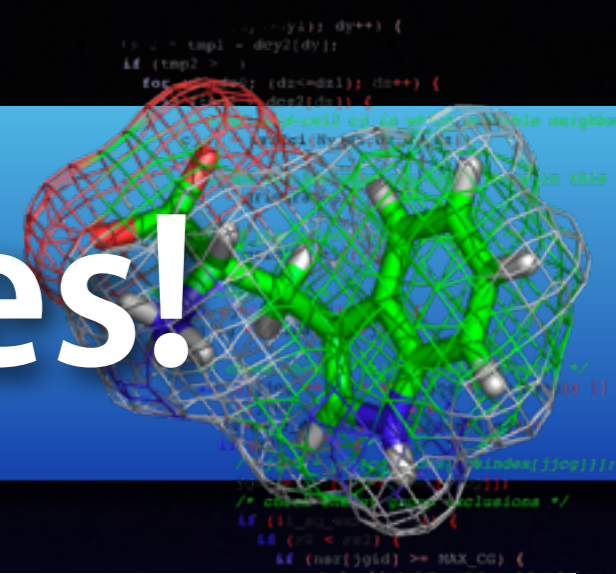
AA	pH 7 charge	pKa
GLU	-1	4,3
ASP	-1	3,9
HIS	0 or +1	6,5
LYS	+1	10,5
ARG	+1	12,5
TYR	0	10,1
CYS	0	9,2

Tricky; very close to neutral pH
Depends on environment too

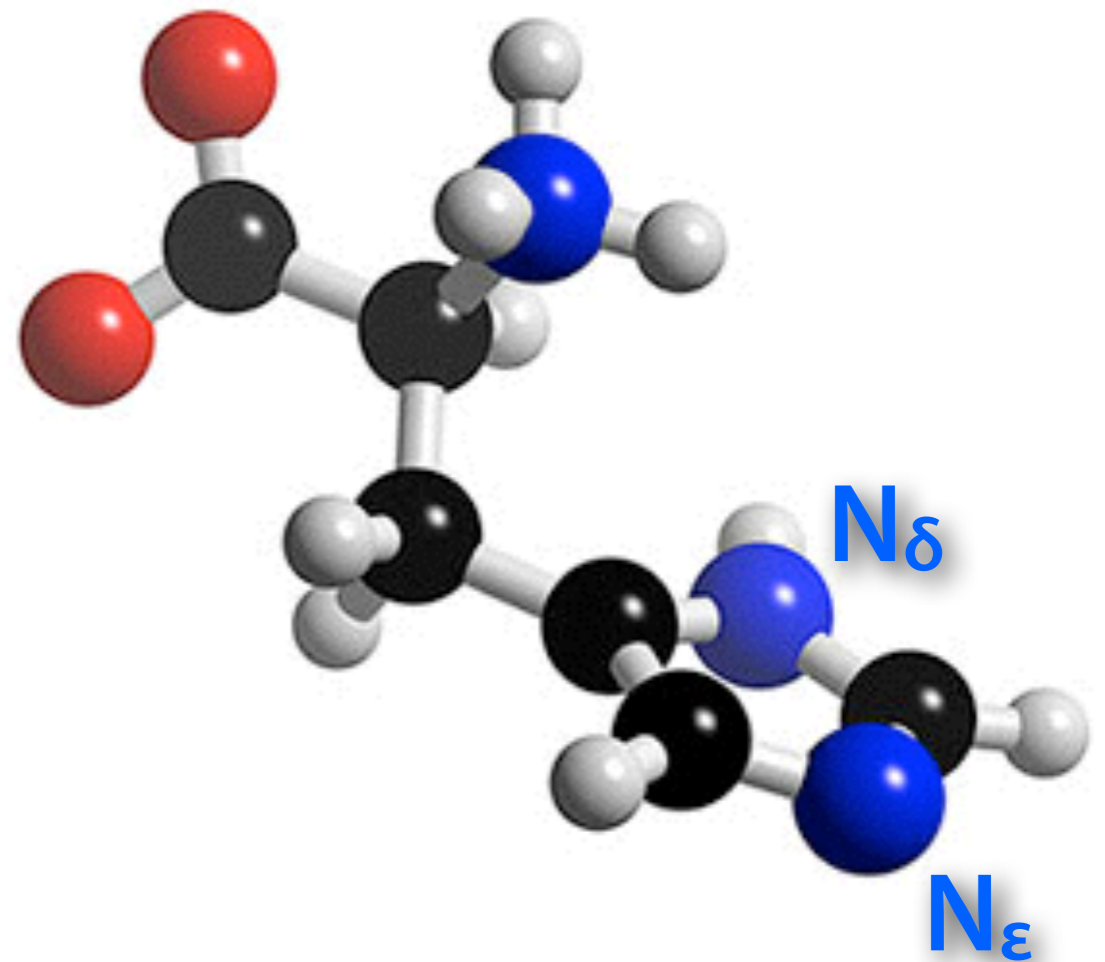


Overhead transparencies to accompany Garrett/Grisham: Biochemistry page 65
Transparency 9 Figure 3.8 ©1995 Saunders College Publishing

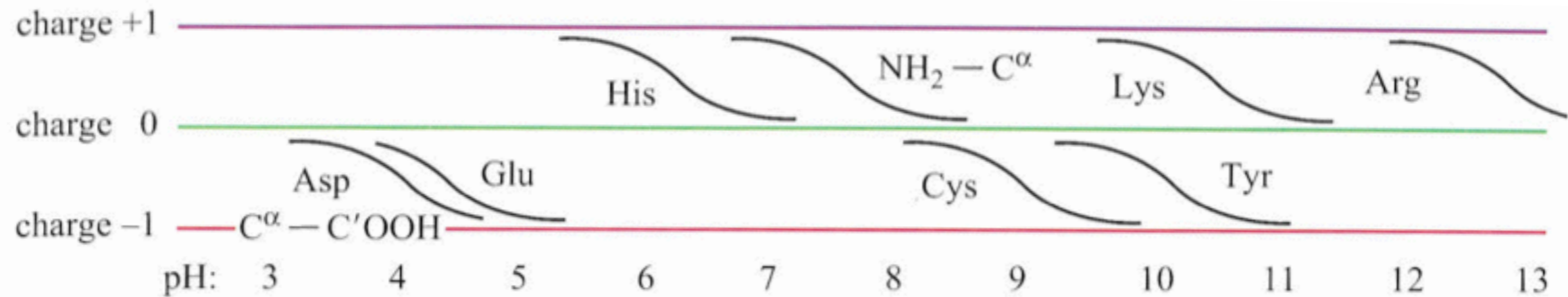
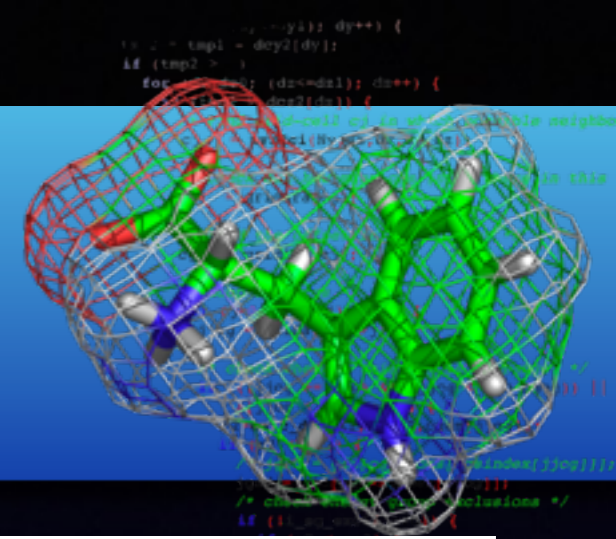
Histidine: Two sites!



- N_{δ} & N_{ϵ}
- Three possibilities
- Neutral:
 - H_{δ}
 - H_{ϵ}
- Charged:
 - H_{δ} & H_{ϵ}



Charge vs. pH



pH-regulated properties

Ion channels: opening, gating

Protein stability,
Salt bridges

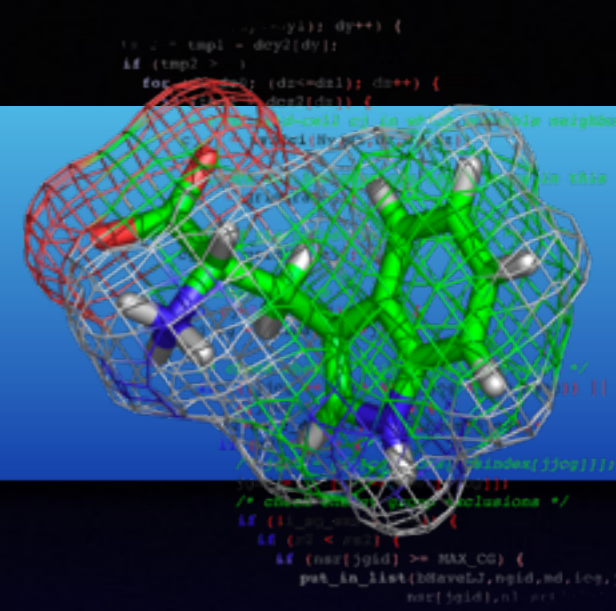
DNA-protein

interaction

Binding of charged
molecules

*Can be difficult
to predict!*

Summary



- Read chapters 7 & 10 of “Protein Physics”
- What are the fundamental differences between helices and sheets in terms of stabilization properties?
- How do you think they might form?