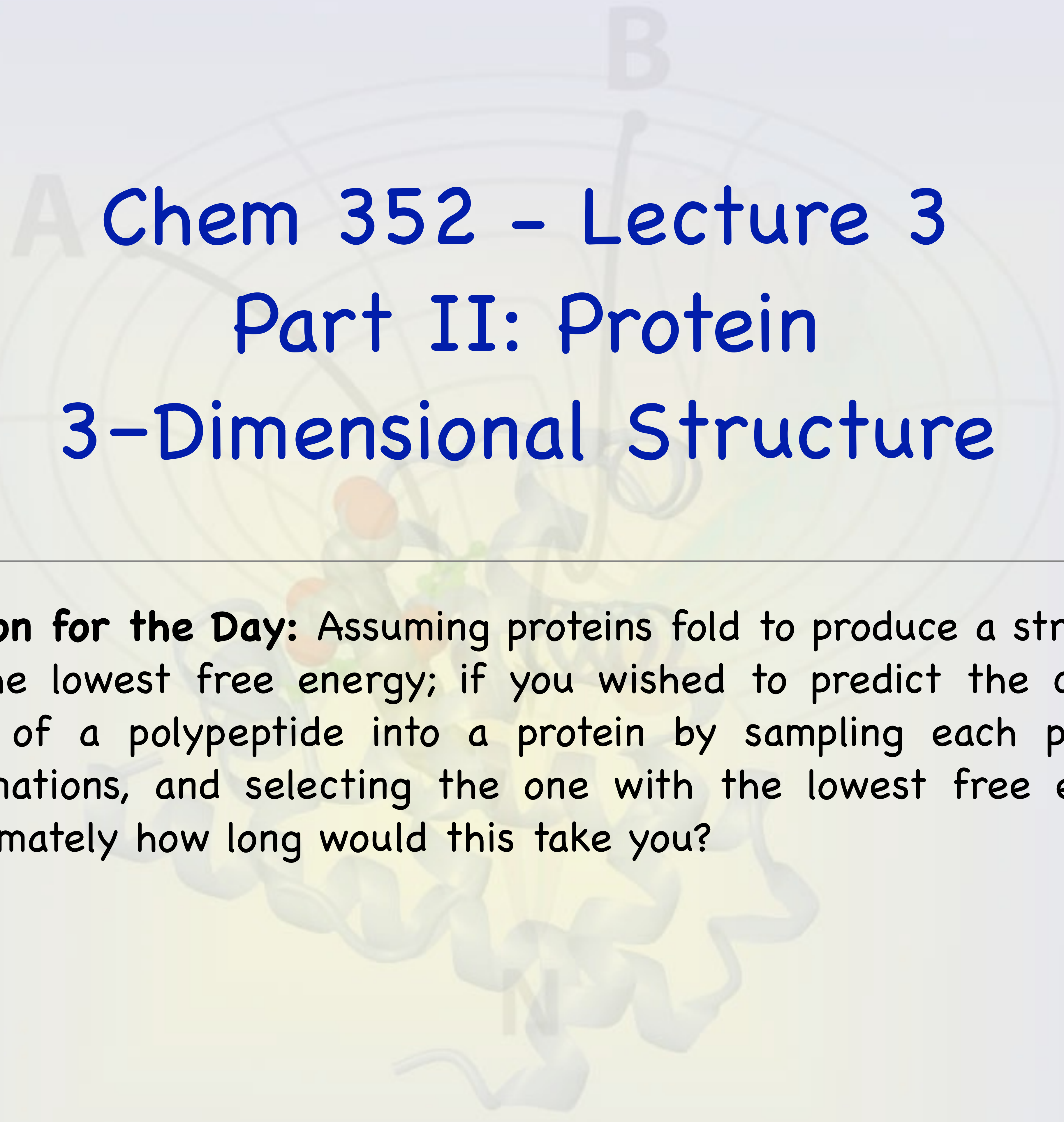
Chem 352 - Lecture 3
Part II: Protein
3-Dimensional Structure



Question for the Day: Assuming proteins fold to produce a structure with the lowest free energy; if you wished to predict the correct folding of a polypeptide into a protein by sampling each possible conformations, and selecting the one with the lowest free energy, approximately how long would this take you?

Introduction to Protein Structure

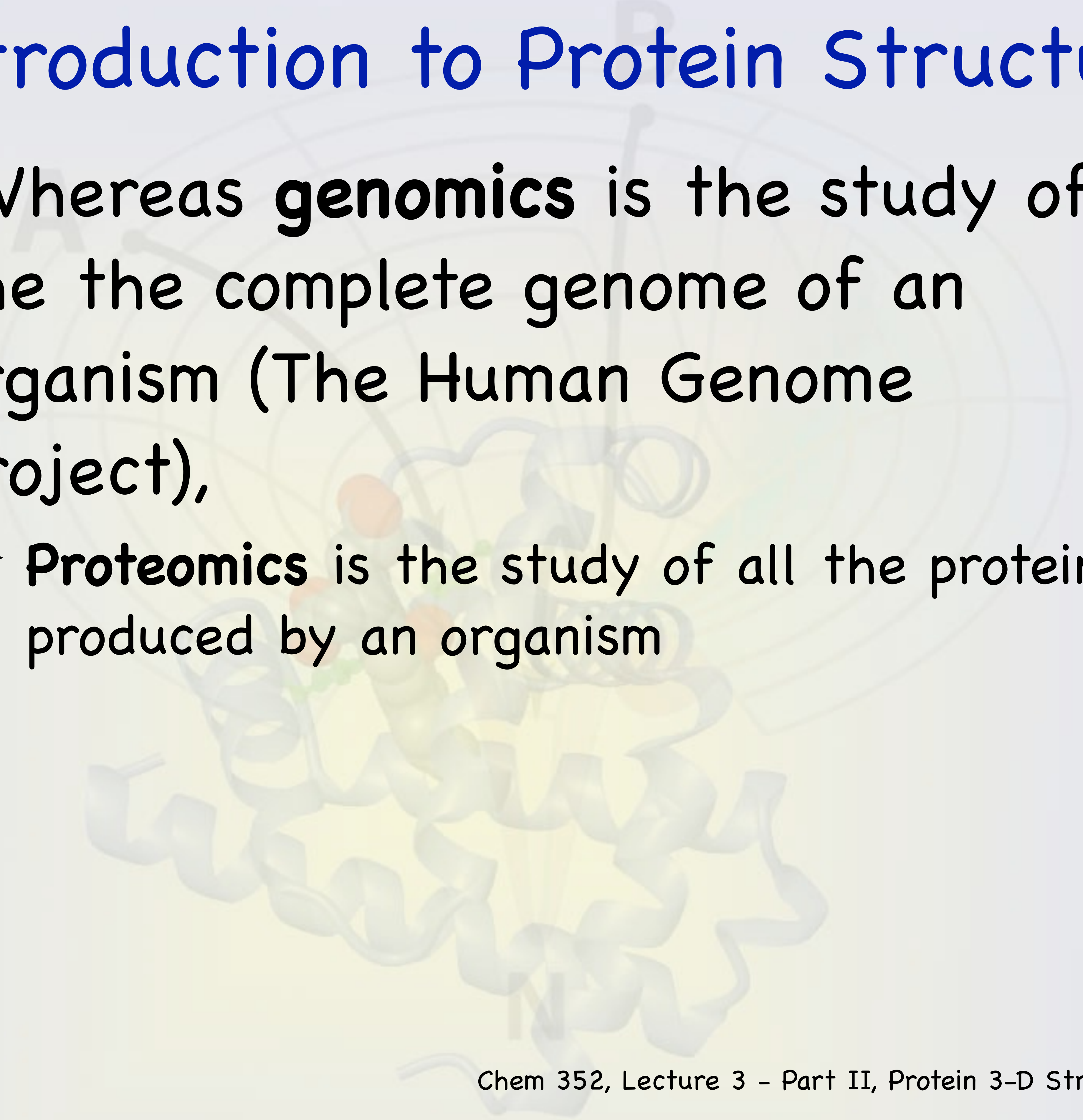
A polypeptide has a lot of flexibility.

- This allows them to adopt numerous shapes or conformations.
- To be functional, proteins usually need to adopt a particular conformation, referred to as the native conformation.

Polypeptide can have wide range of lengths

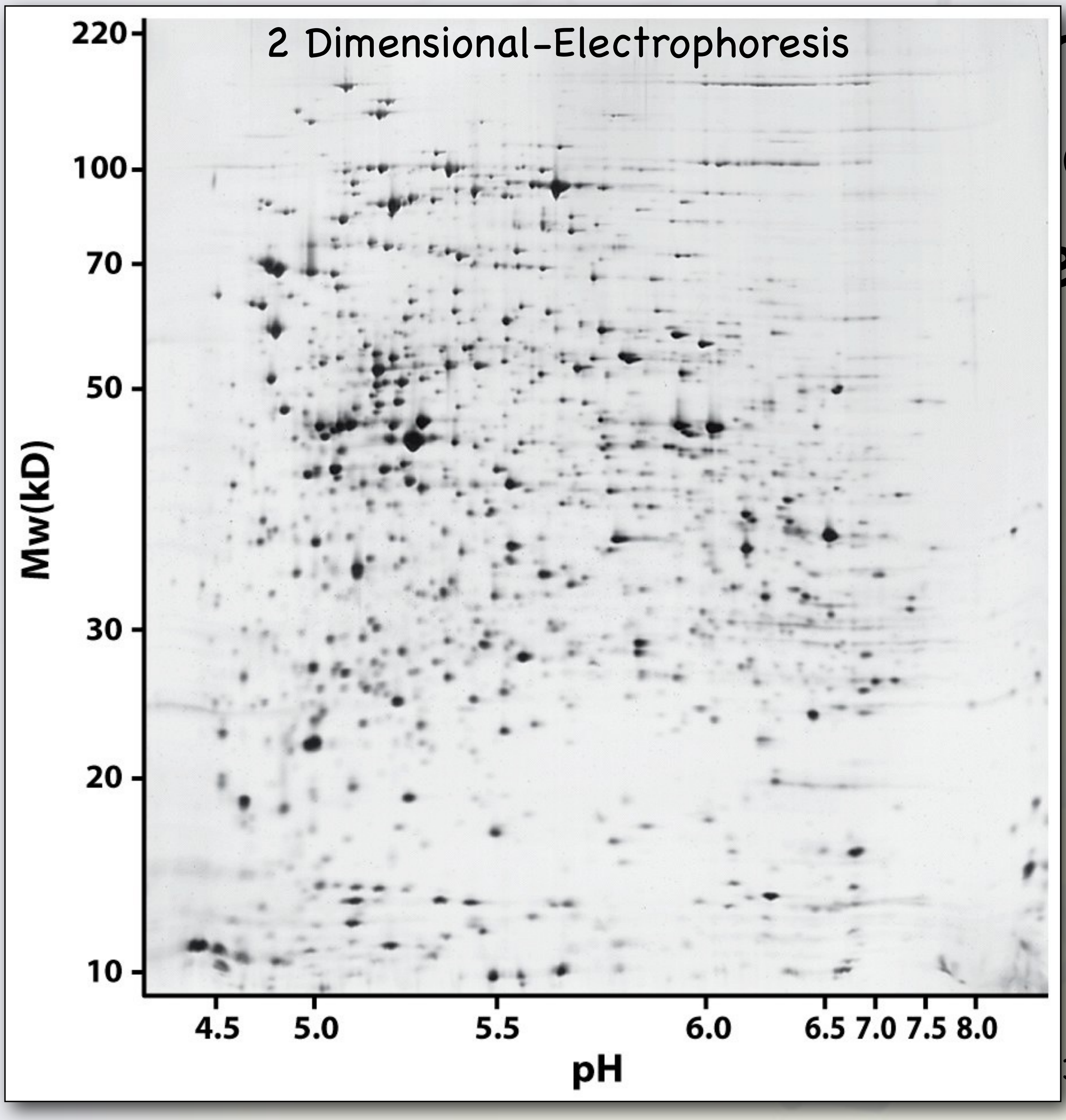
- <100 amino acids to >2000 amino acids

Introduction to Protein Structure



- Whereas **genomics** is the study of the the complete genome of an organism (The Human Genome Project),
  - **Proteomics** is the study of all the proteins produced by an organism

Introduction to Protein Structure



2 Dimensional-Electrophoresis

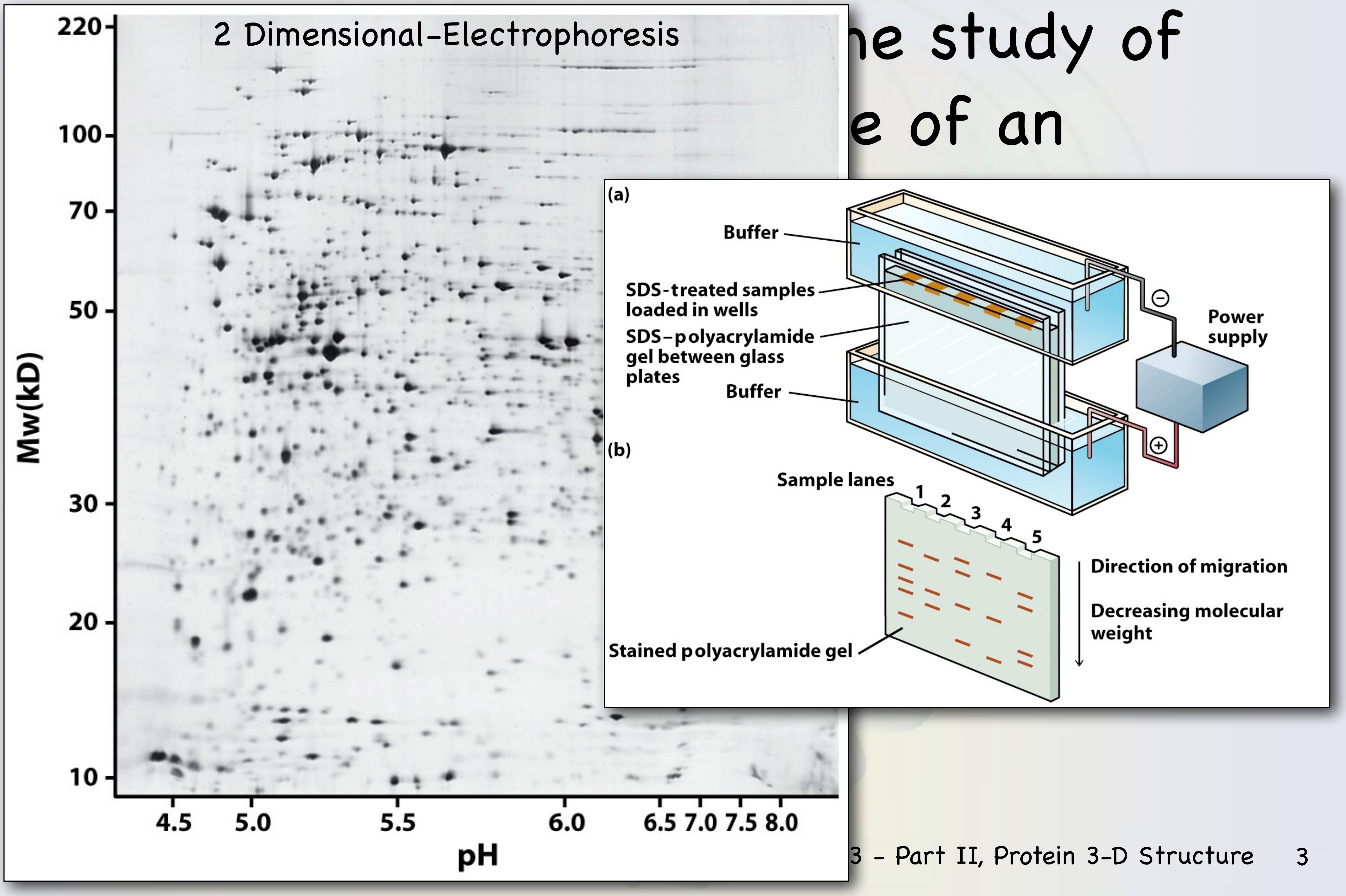
ne study of
e of an
enome

all the proteins

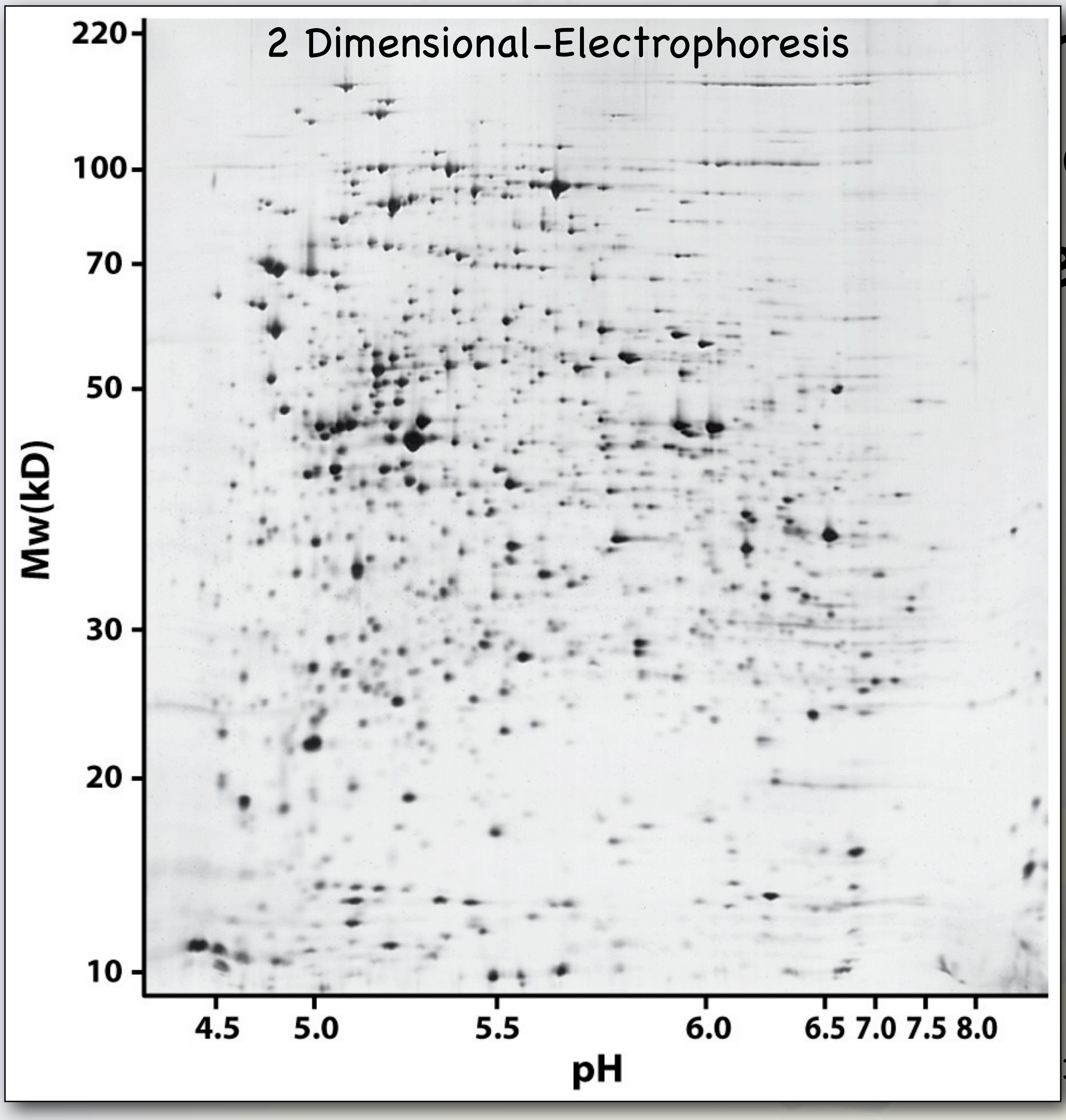
Introduction to Protein Structure

he study of

e of an



Introduction to Protein Structure



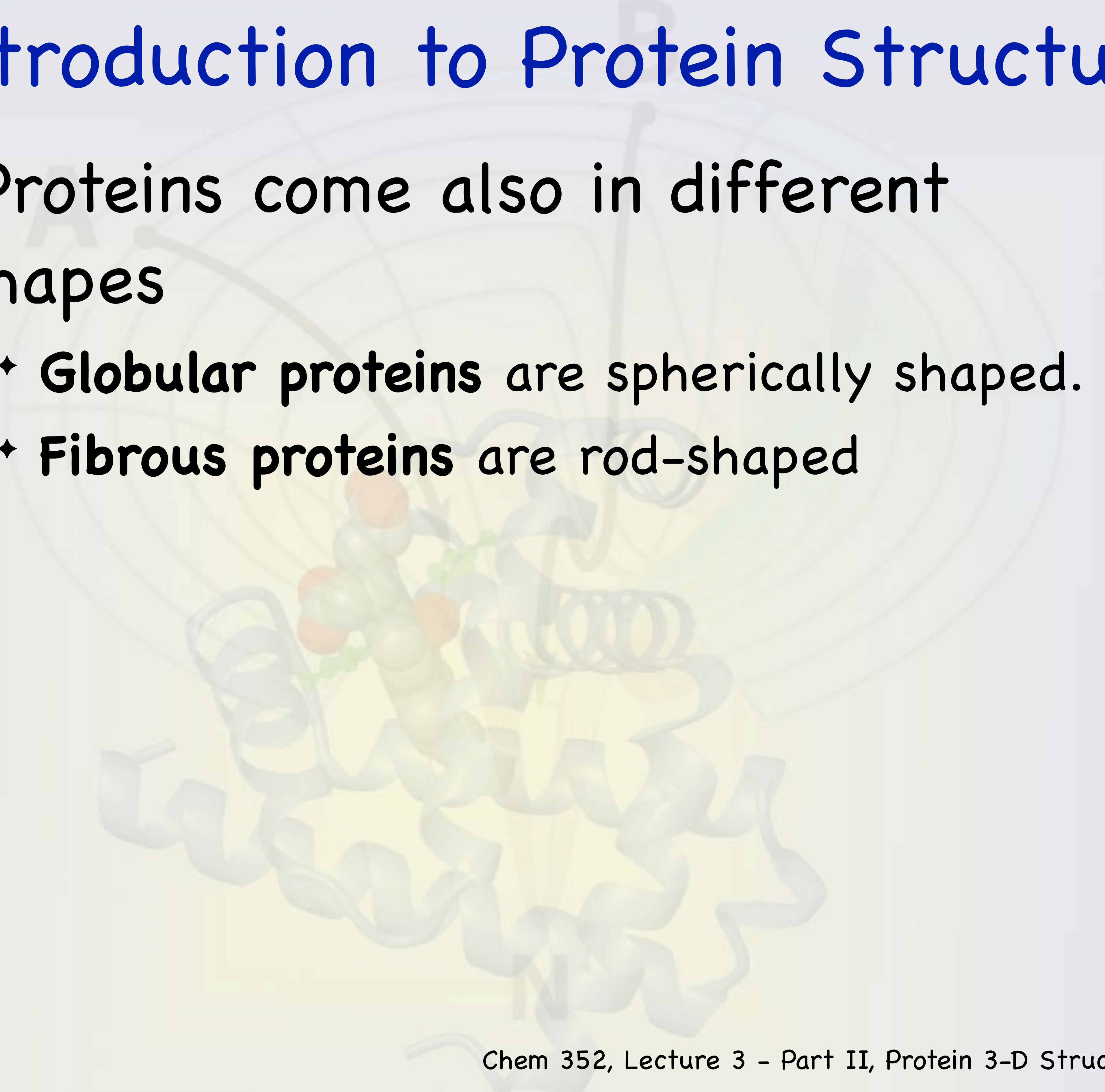
ne study of
e of an
enome

all the proteins

Introduction to Protein Structure

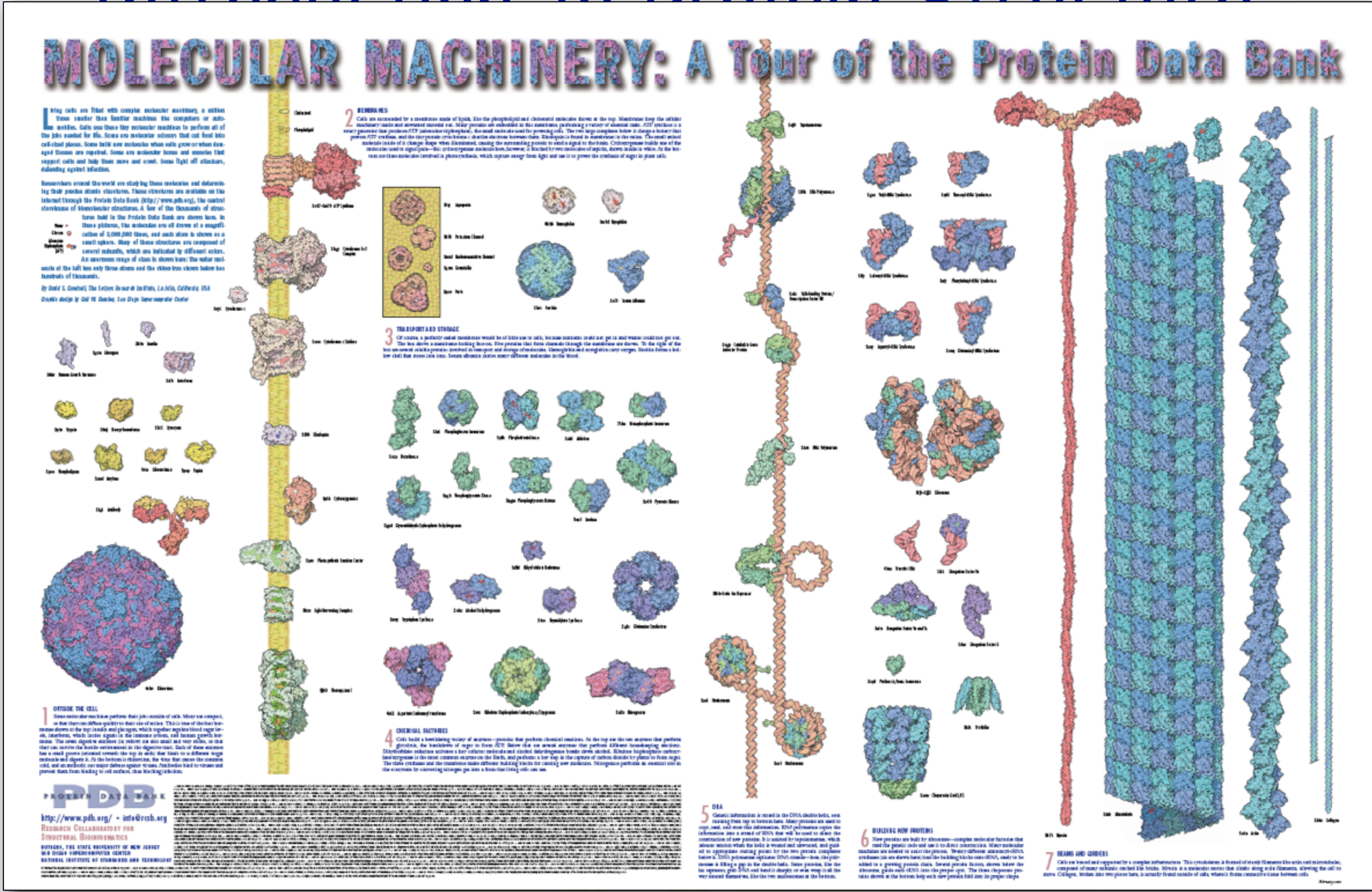
- Whereas **genomics** is the study of the the complete genome of an organism (The Human Genome Project),
  - ✦ **Proteomics** is the study of all the proteins produced by an organism

Introduction to Protein Structure

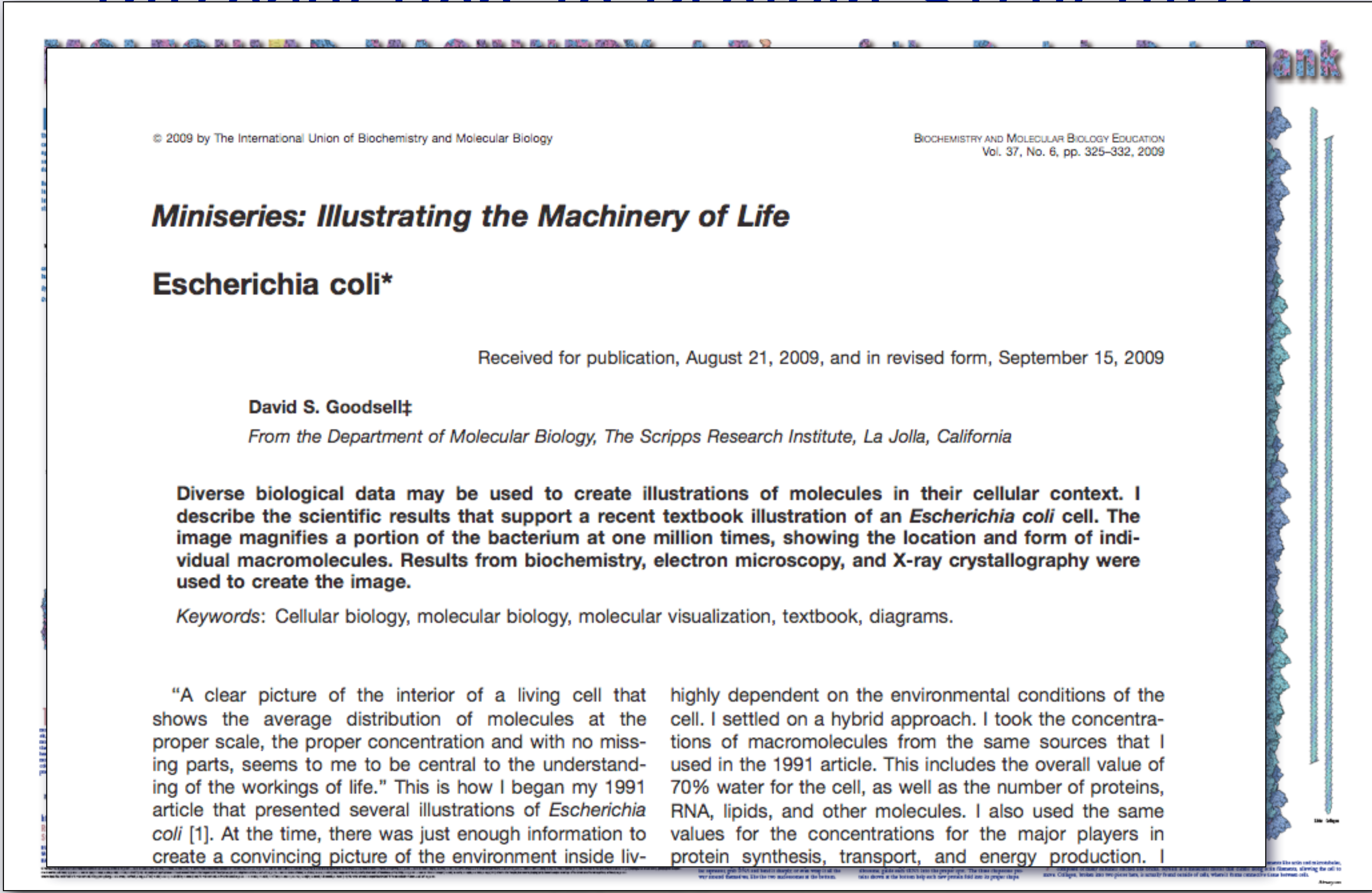


- Proteins come also in different shapes
  - **Globular proteins** are spherically shaped.
  - **Fibrous proteins** are rod-shaped

Introduction to Protein Structure



Introduction to Protein Structure



© 2009 by The International Union of Biochemistry and Molecular Biology

BIOCHEMISTRY AND MOLECULAR BIOLOGY EDUCATION
Vol. 37, No. 6, pp. 325–332, 2009

Miniseries: Illustrating the Machinery of Life

Escherichia coli*

Received for publication, August 21, 2009, and in revised form, September 15, 2009

David S. Goodsell‡

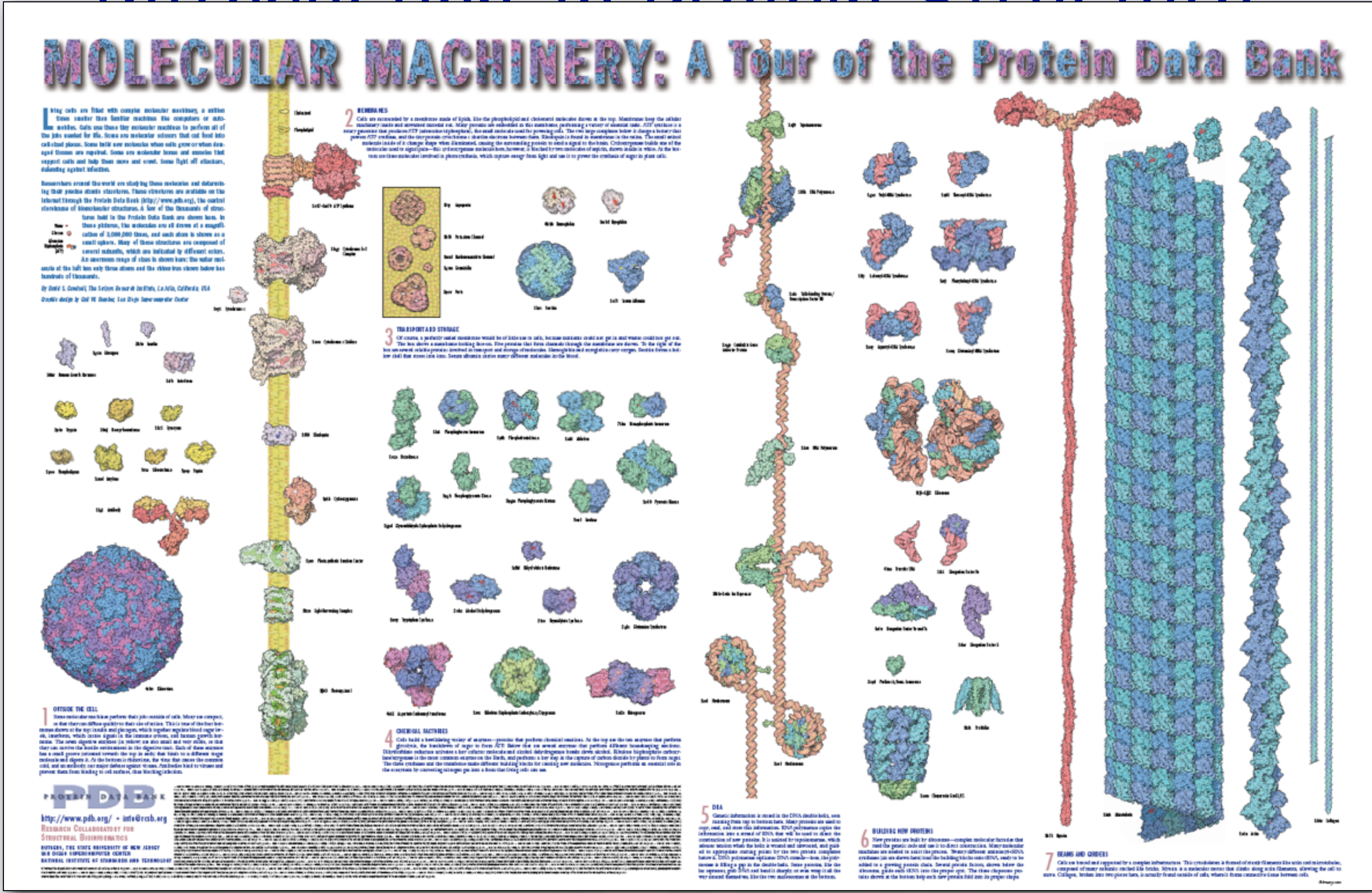
From the Department of Molecular Biology, The Scripps Research Institute, La Jolla, California

Diverse biological data may be used to create illustrations of molecules in their cellular context. I describe the scientific results that support a recent textbook illustration of an *Escherichia coli* cell. The image magnifies a portion of the bacterium at one million times, showing the location and form of individual macromolecules. Results from biochemistry, electron microscopy, and X-ray crystallography were used to create the image.

Keywords: Cellular biology, molecular biology, molecular visualization, textbook, diagrams.

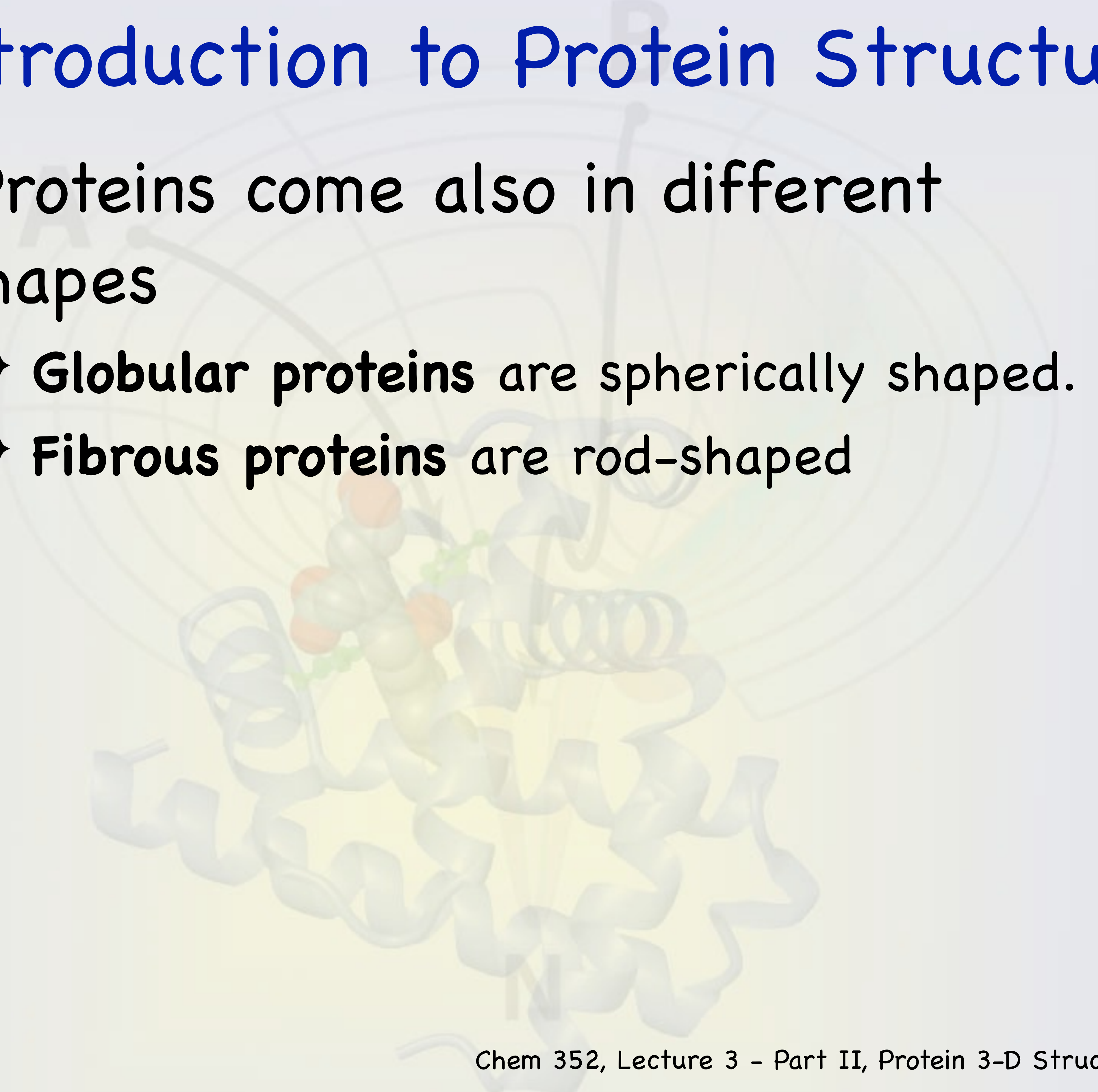
"A clear picture of the interior of a living cell that shows the average distribution of molecules at the proper scale, the proper concentration and with no missing parts, seems to me to be central to the understanding of the workings of life." This is how I began my 1991 article that presented several illustrations of *Escherichia coli* [1]. At the time, there was just enough information to create a convincing picture of the environment inside liv- highly dependent on the environmental conditions of the cell. I settled on a hybrid approach. I took the concentrations of macromolecules from the same sources that I used in the 1991 article. This includes the overall value of 70% water for the cell, as well as the number of proteins, RNA, lipids, and other molecules. I also used the same values for the concentrations for the major players in protein synthesis, transport, and energy production. I

Introduction to Protein Structure



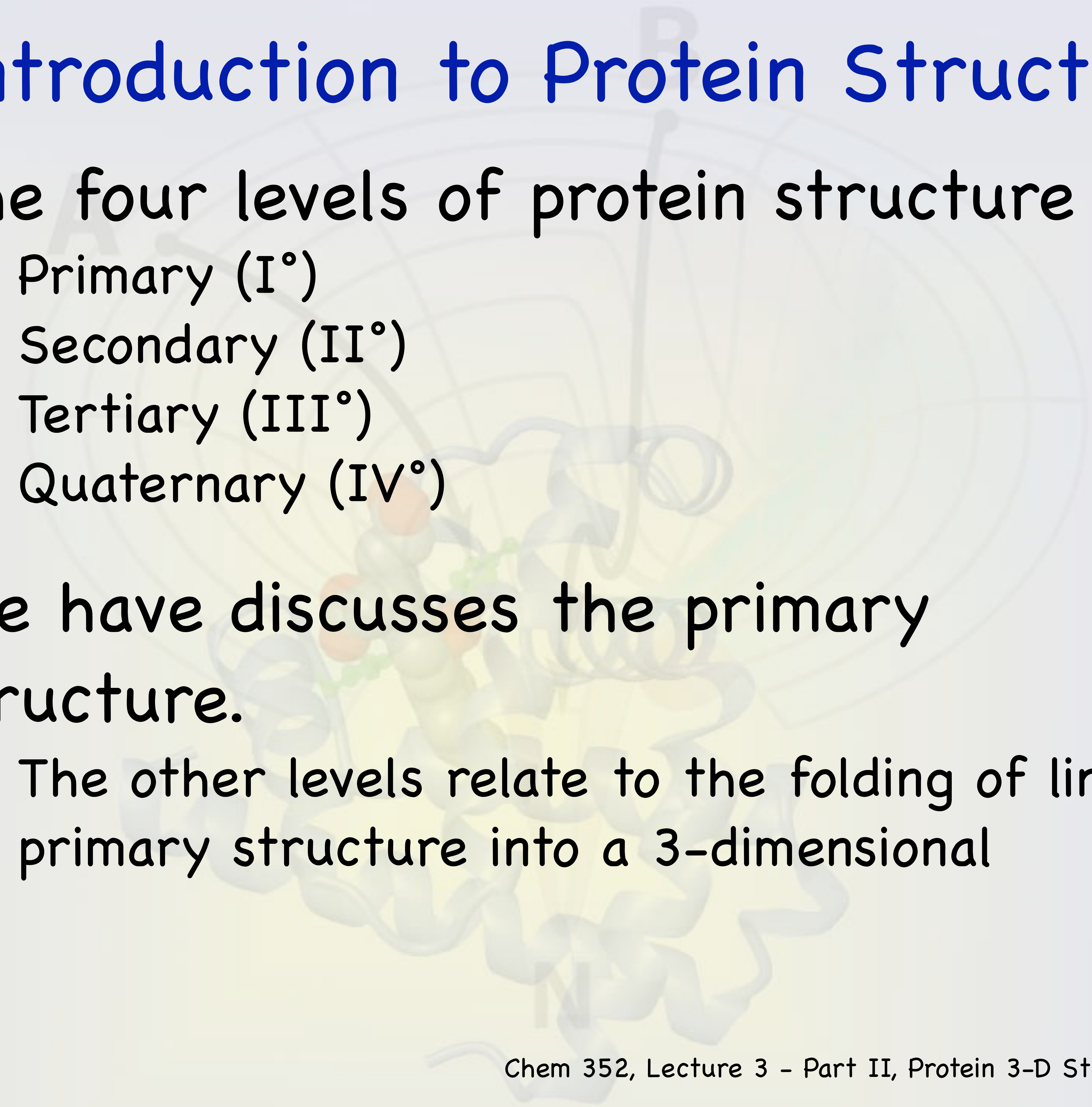
MOLECULAR MACHINERY: A Tour of the Protein Data Bank

Introduction to Protein Structure



- Proteins come also in different shapes
  - **Globular proteins** are spherically shaped.
  - **Fibrous proteins** are rod-shaped

Introduction to Protein Structure



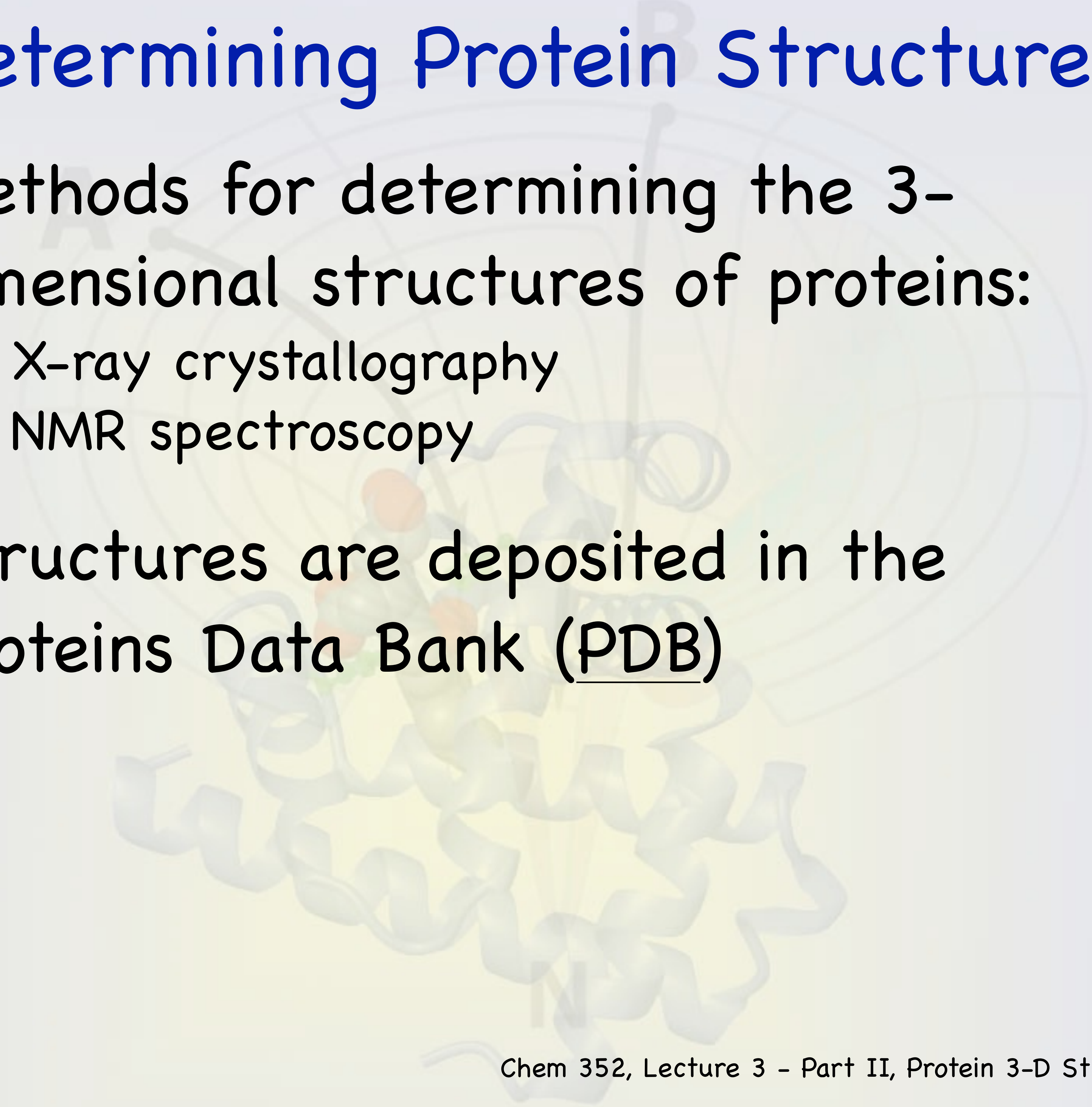
The four levels of protein structure

- Primary (I°)
- Secondary (II°)
- Tertiary (III°)
- Quaternary (IV°)

We have discusses the primary structure.

- The other levels relate to the folding of linear primary structure into a 3-dimensional

Determining Protein Structures



Methods for determining the 3-dimensional structures of proteins:

- X-ray crystallography
- NMR spectroscopy

Structures are deposited in the Proteins Data Bank (PDB)

Determining Protein Structures

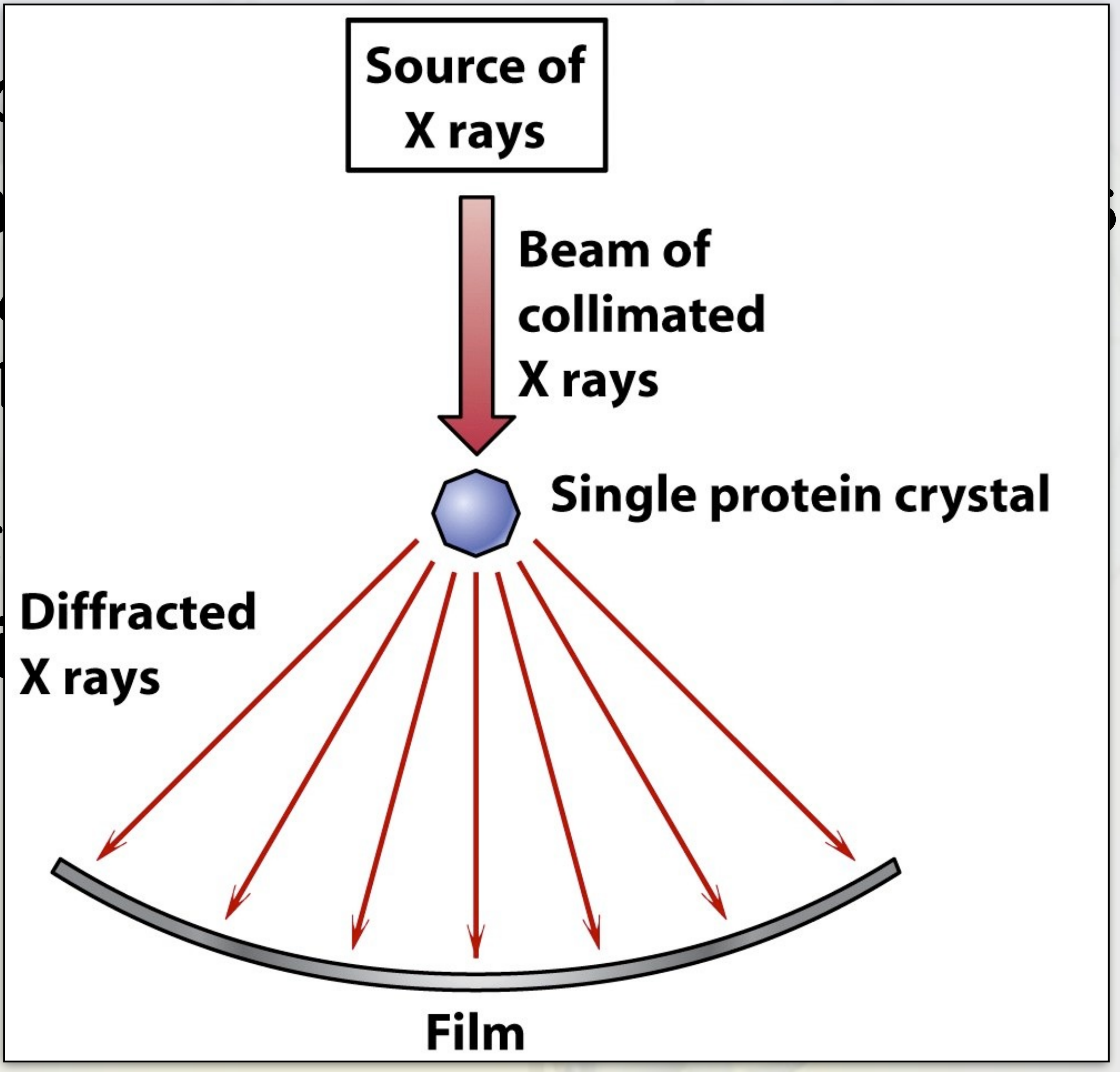
Meth

dimen :

- X-r
- NM

Struc

Protei



Determining Protein Structures

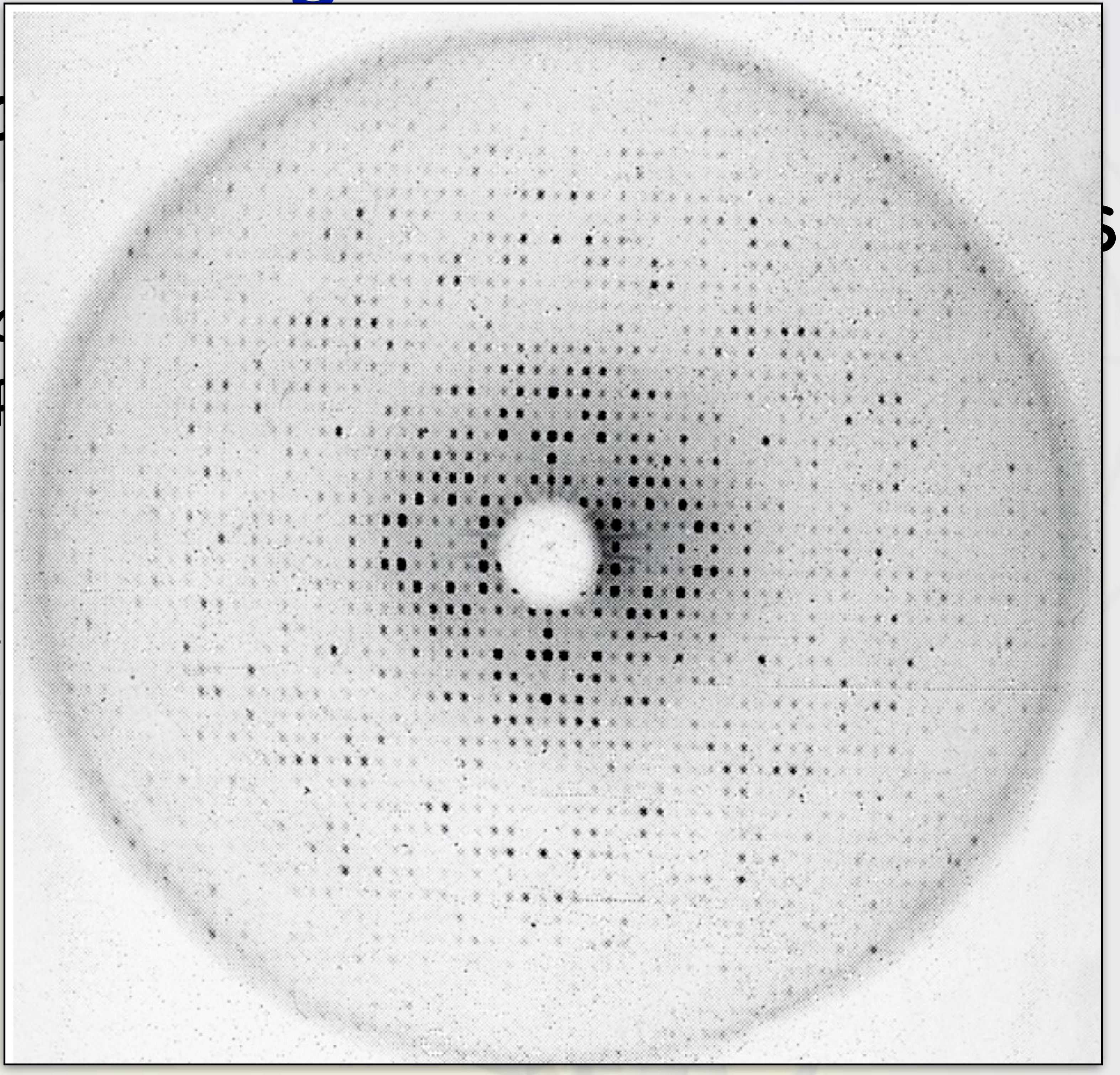
Meth

dimen s:

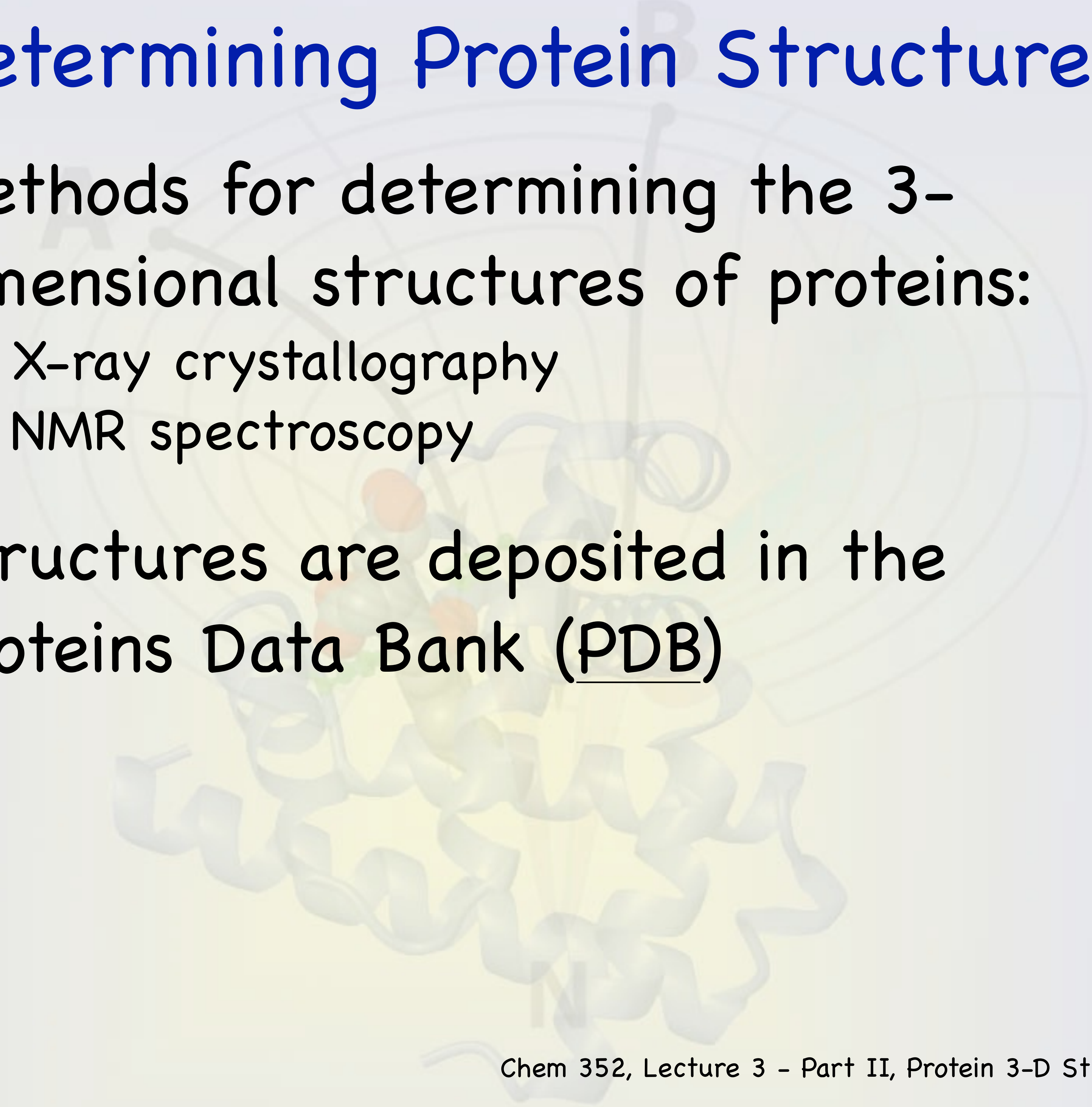
- X-r
- NM

Struc

Protei



Determining Protein Structures



Methods for determining the 3-dimensional structures of proteins:

- X-ray crystallography
- NMR spectroscopy

Structures are deposited in the Proteins Data Bank (PDB)

Determining Protein Structures

Met

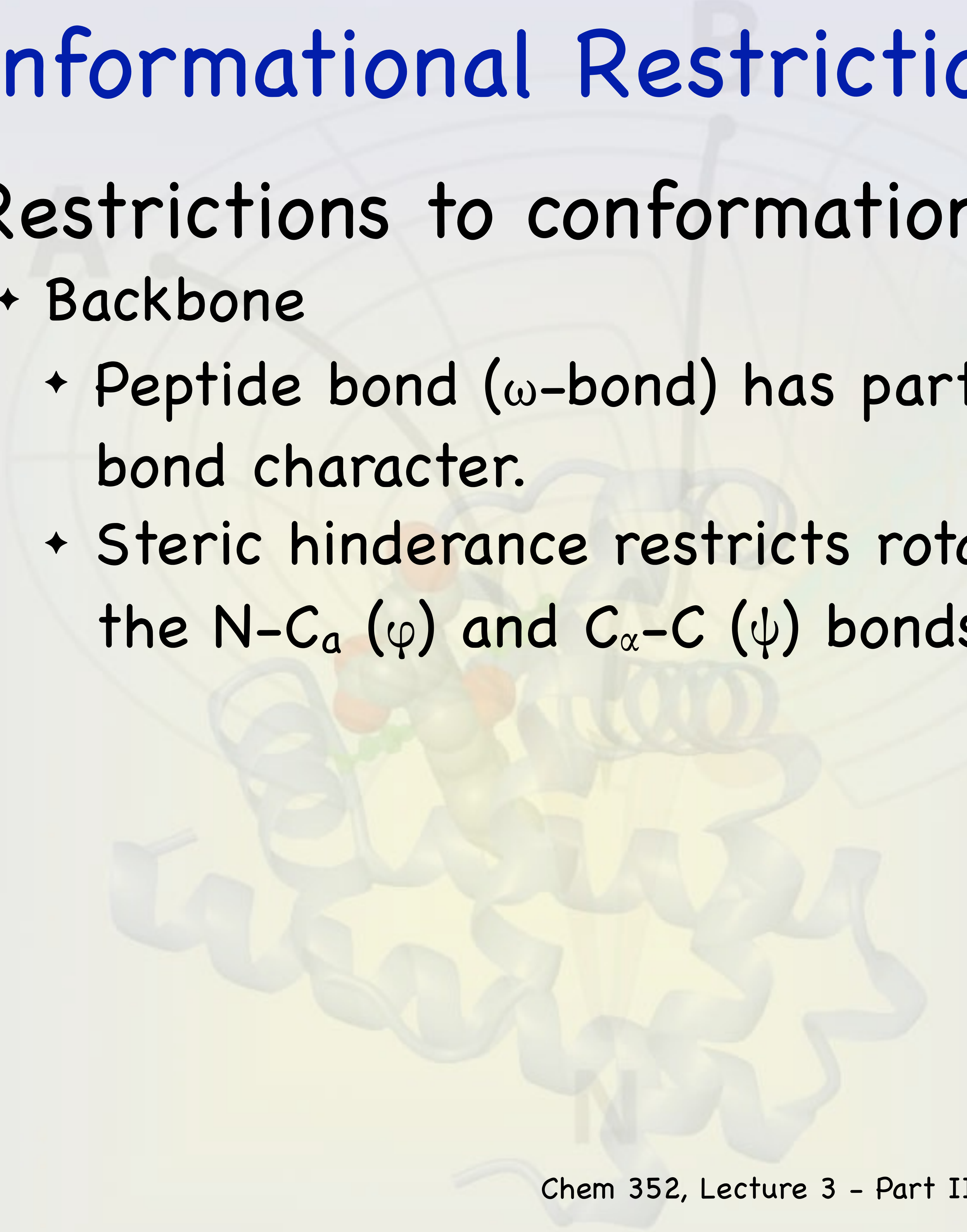
dim

- X
- N

Str

Pro

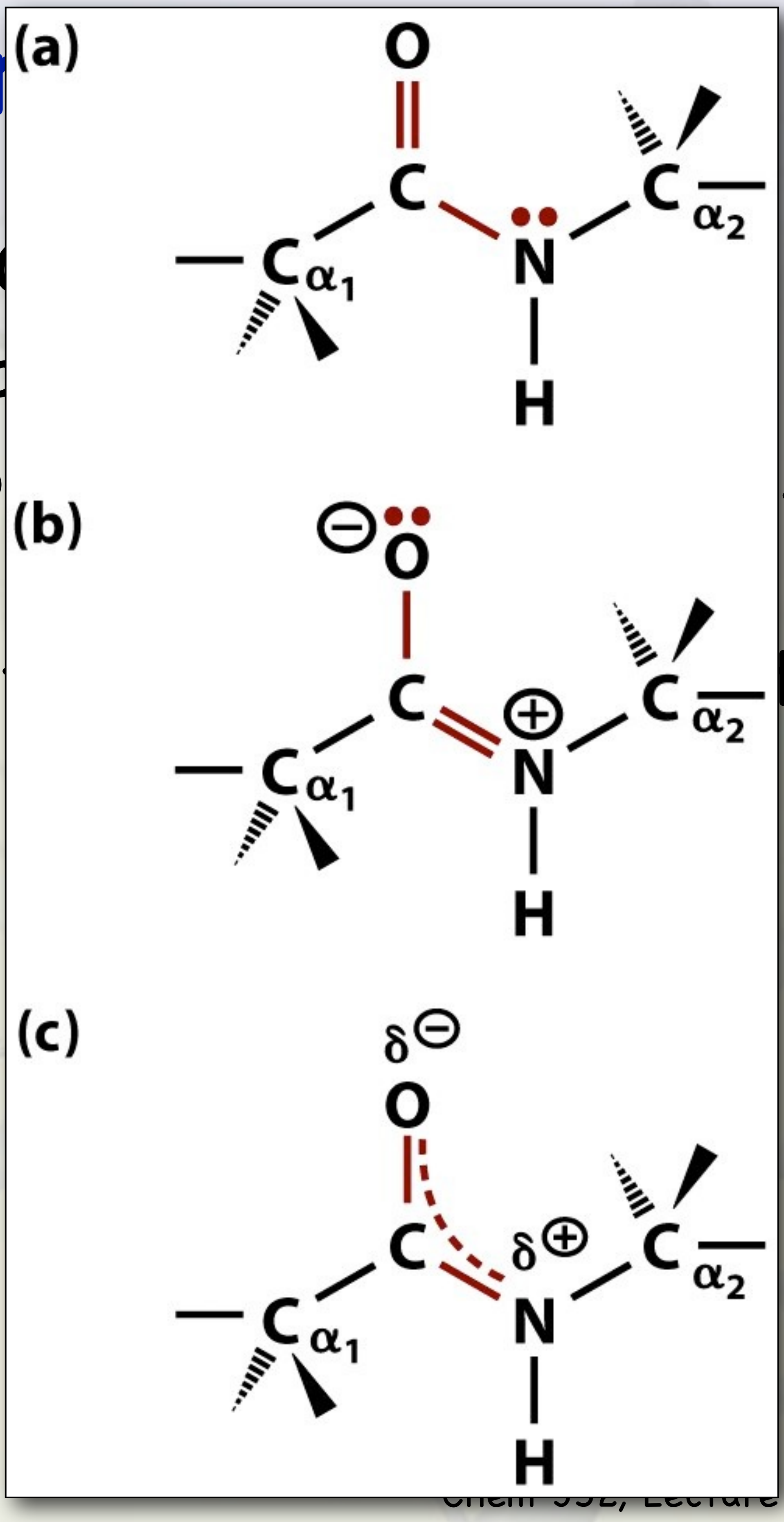
Conformational Restrictions



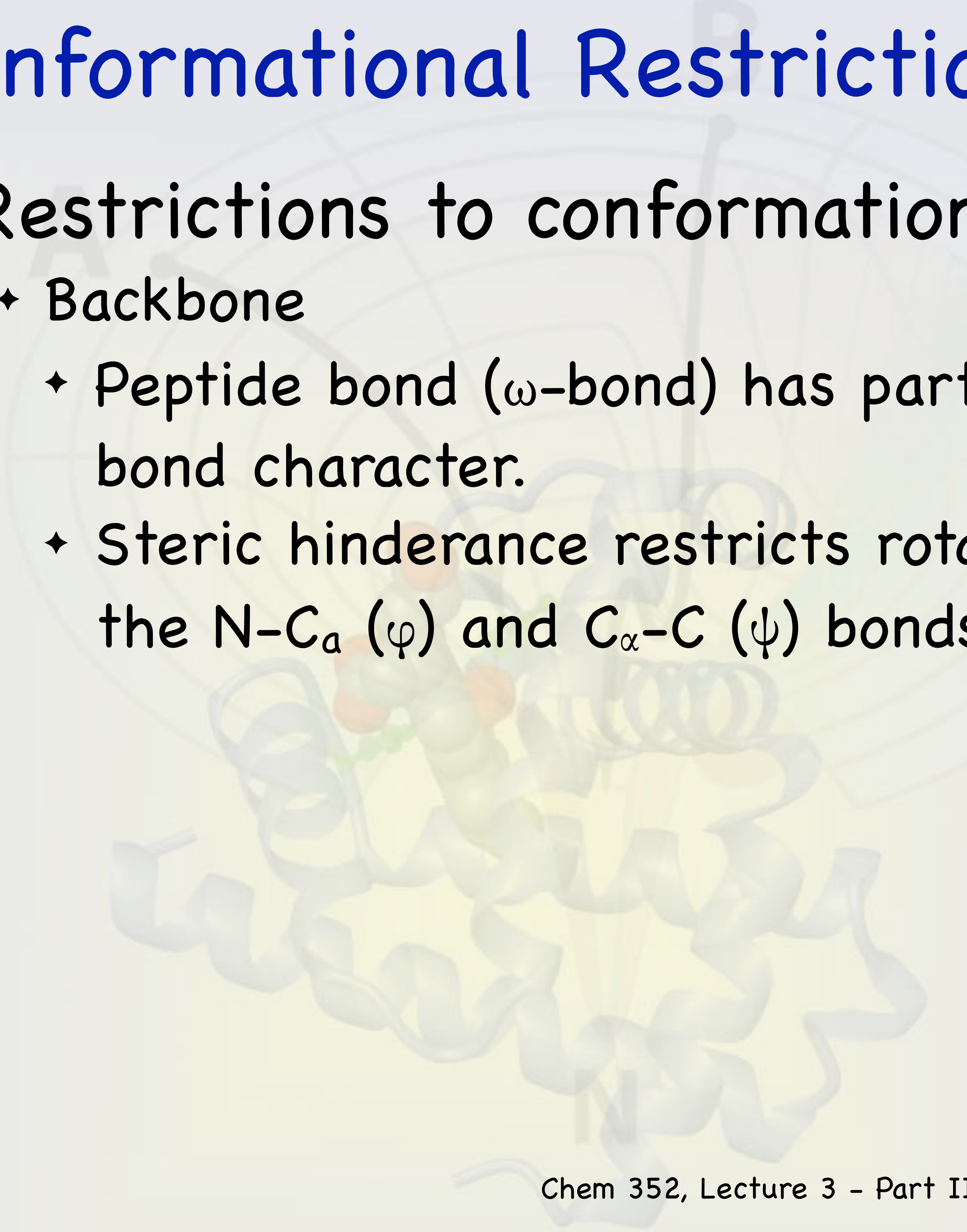
- Restrictions to conformations
  - Backbone
    - Peptide bond (ω-bond) has partial double-bond character.
    - Steric hinderance restricts rotation about the N–C_a (φ) and C_α–C (ψ) bonds

Confor ictions

- •Restri ations
  - ✦ Backb
    - ✦ Pep s partial double-bon
    - ✦ Ste ts rotation about the bonds



Conformational Restrictions

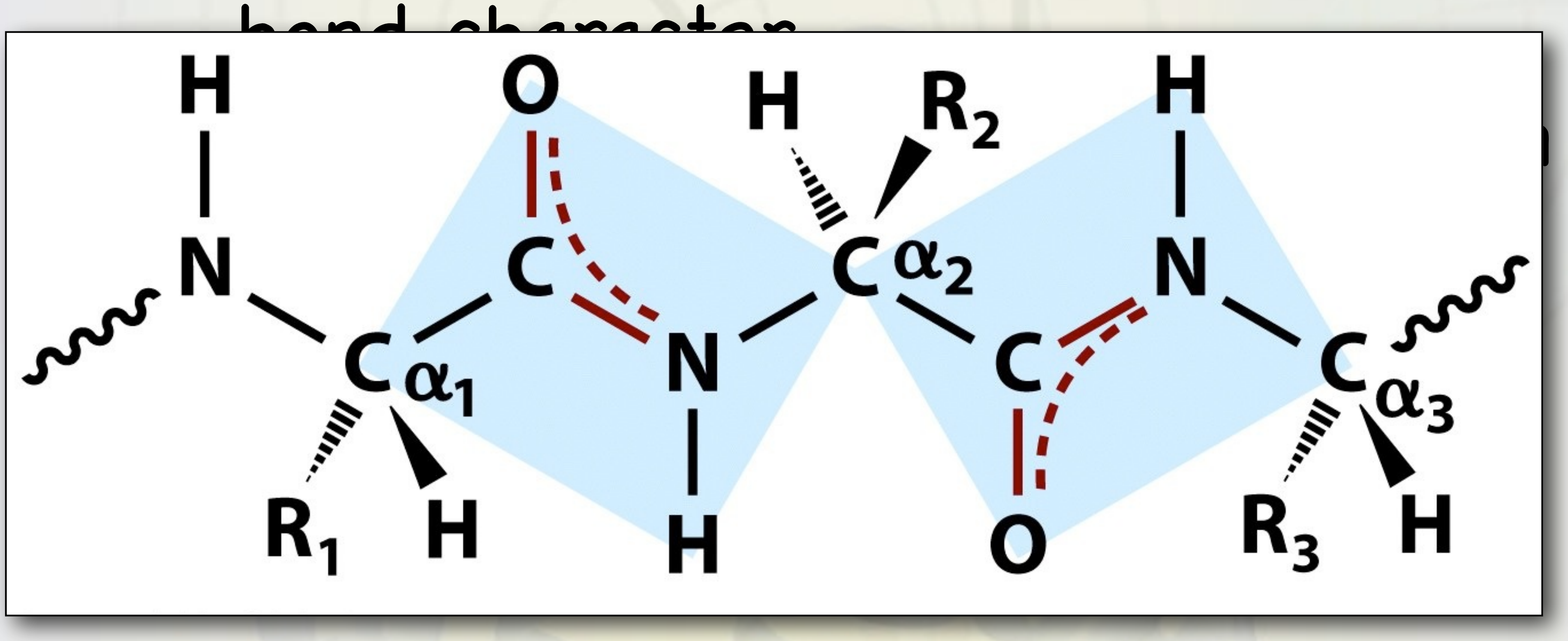


- Restrictions to conformations
  - Backbone
    - Peptide bond (ω-bond) has partial double-bond character.
    - Steric hinderance restricts rotation about the $N-C_a$ (φ) and $C_\alpha-C$ (ψ) bonds

Conformational Restrictions

- Restrictions to conformations
  - Backbone
    - Peptide bond (ω-bond) has partial double-bond character

about

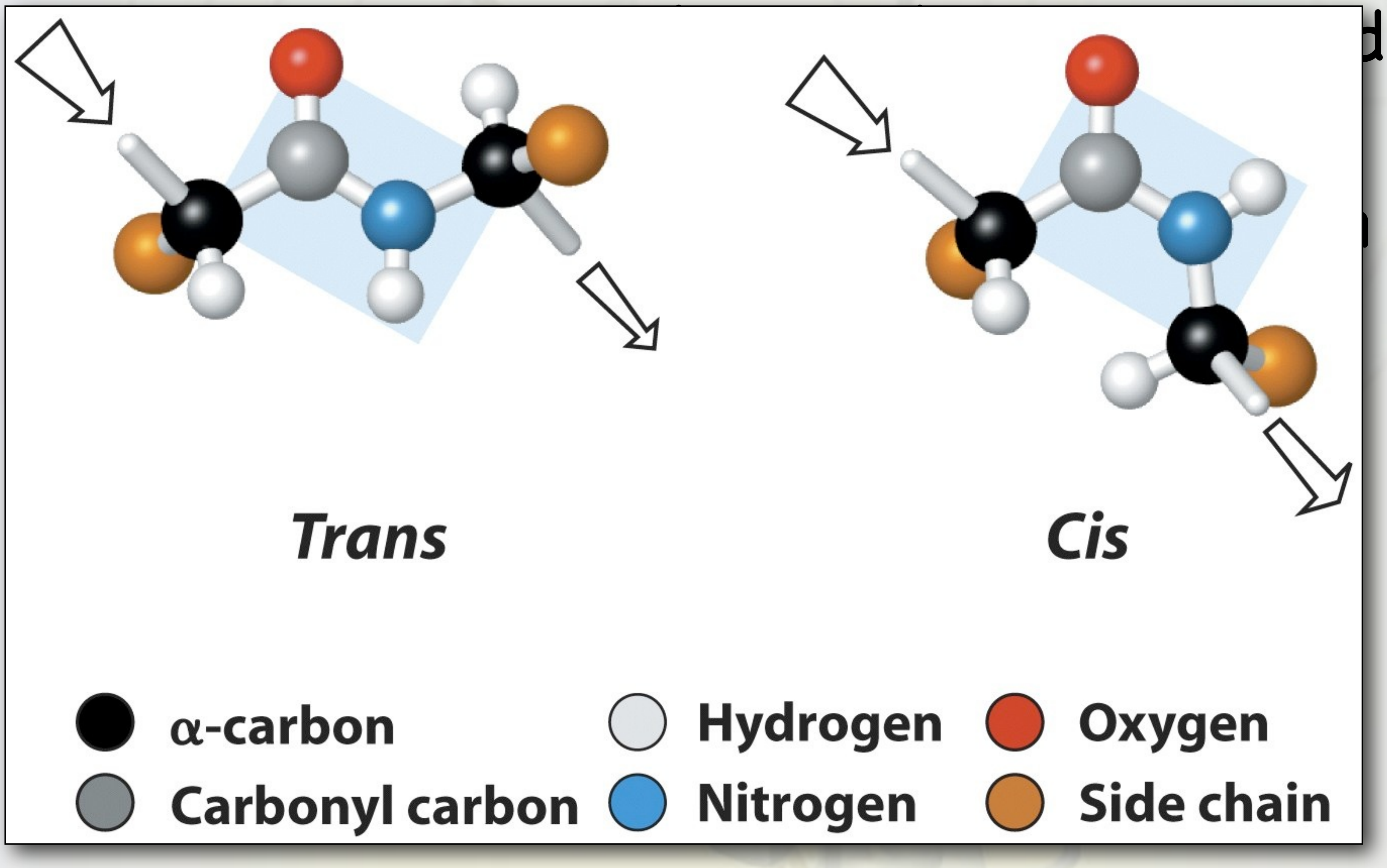


Conformational Restrictions

- Restrictions to conformations
  - Backbone
    - Peptide bond (ω-bond) has partial double-bond character.
    - Steric hinderance restricts rotation about the N-C_a (φ) and C_α-C (ψ) bonds

Conformational Restrictions

- Restrictions to conformations
  - Backbone



double-

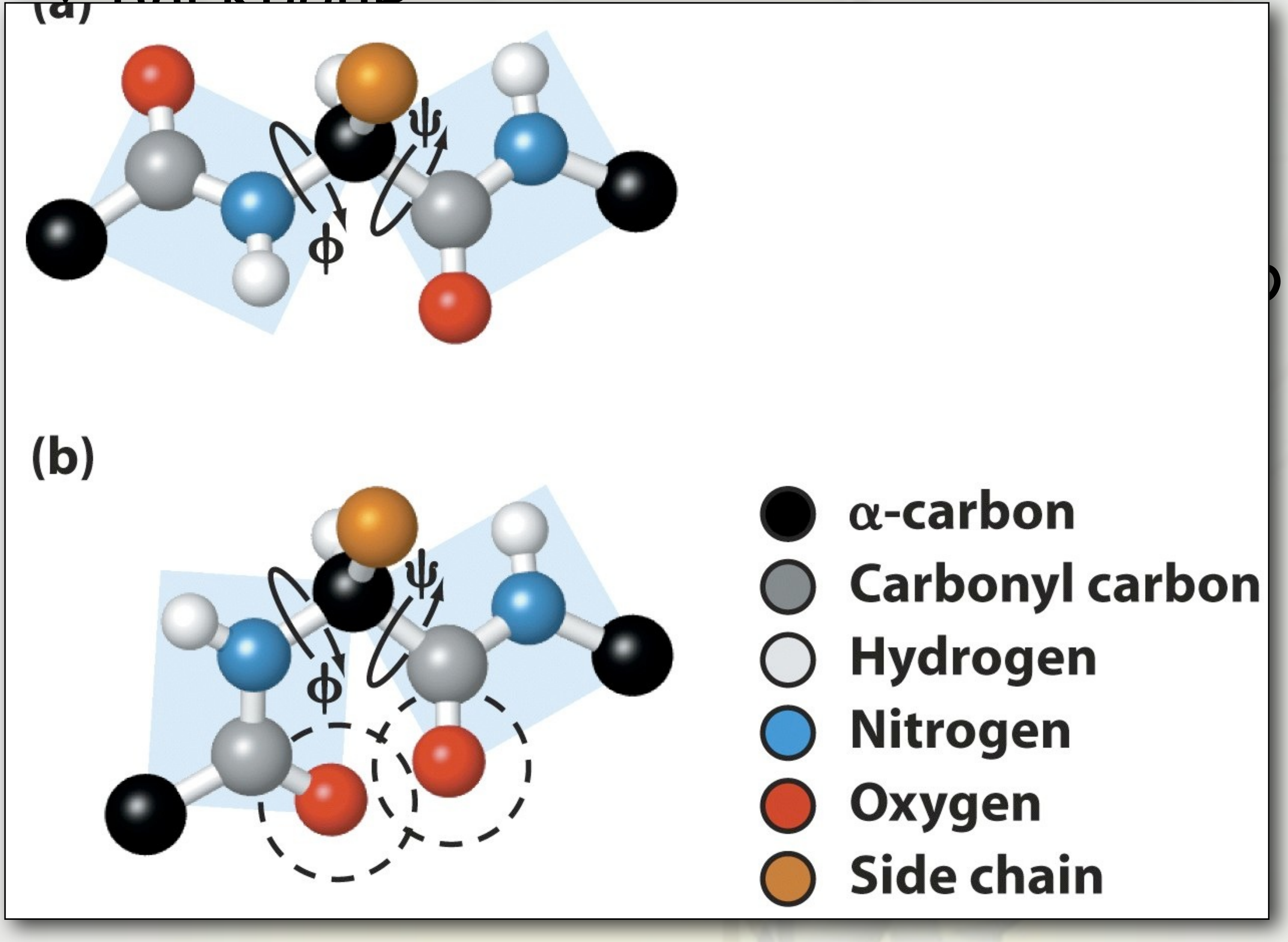
about

Conformational Restrictions

- Restrictions to conformations
  - Backbone
    - Peptide bond (ω-bond) has partial double-bond character.
    - Steric hinderance restricts rotation about the N-C_a (φ) and C_α-C (ψ) bonds

Conformational Restrictions

- Restrictions to conformations
  - ✦ Backbone



double-

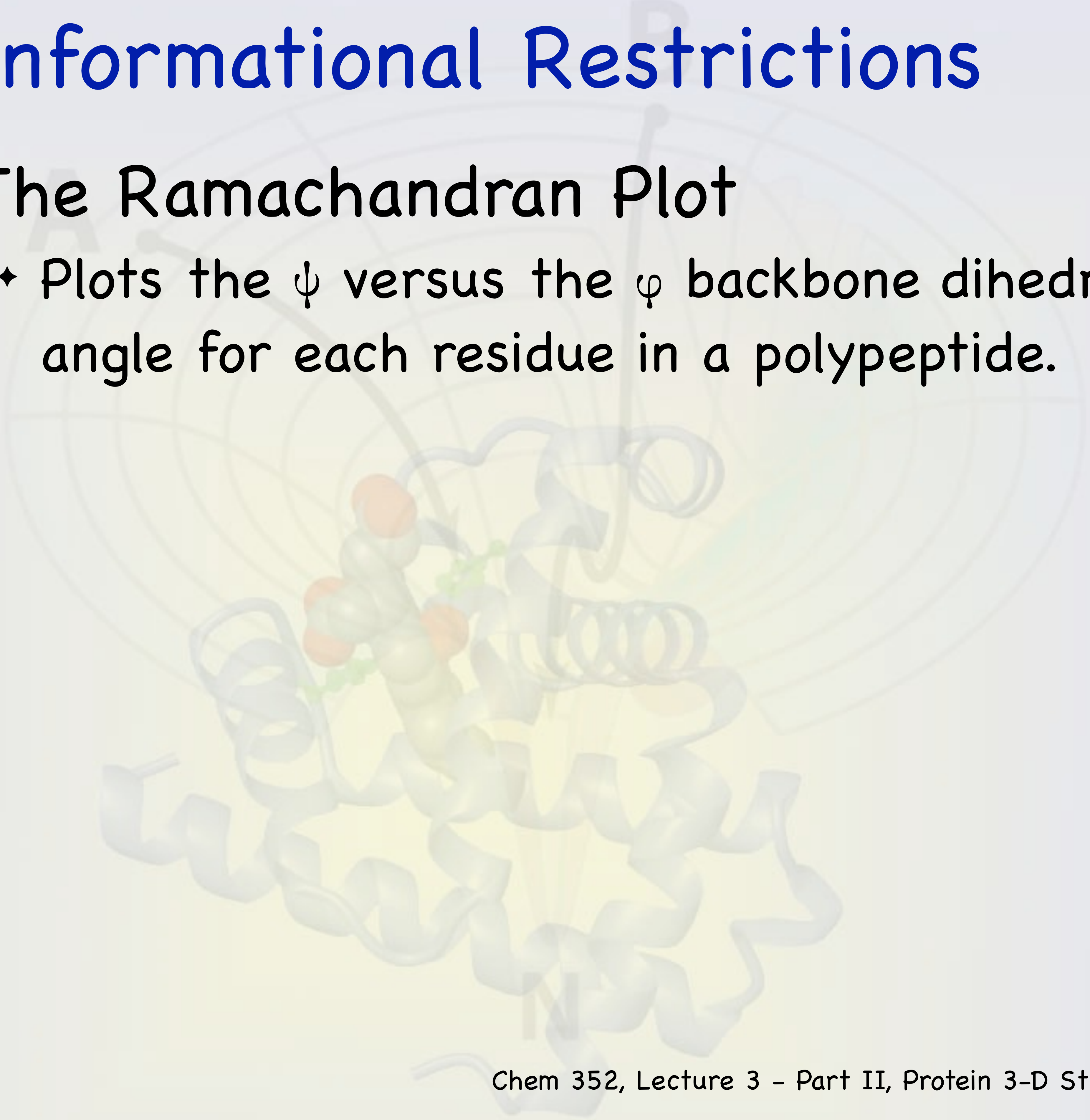
on about

Conformational Restrictions

- Restrictions to conformations
  - Backbone
    - Peptide bond (ω-bond) has partial double-bond character.
    - Steric hinderance restricts rotation about the N-C_a (φ) and C_α-C (ψ) bonds

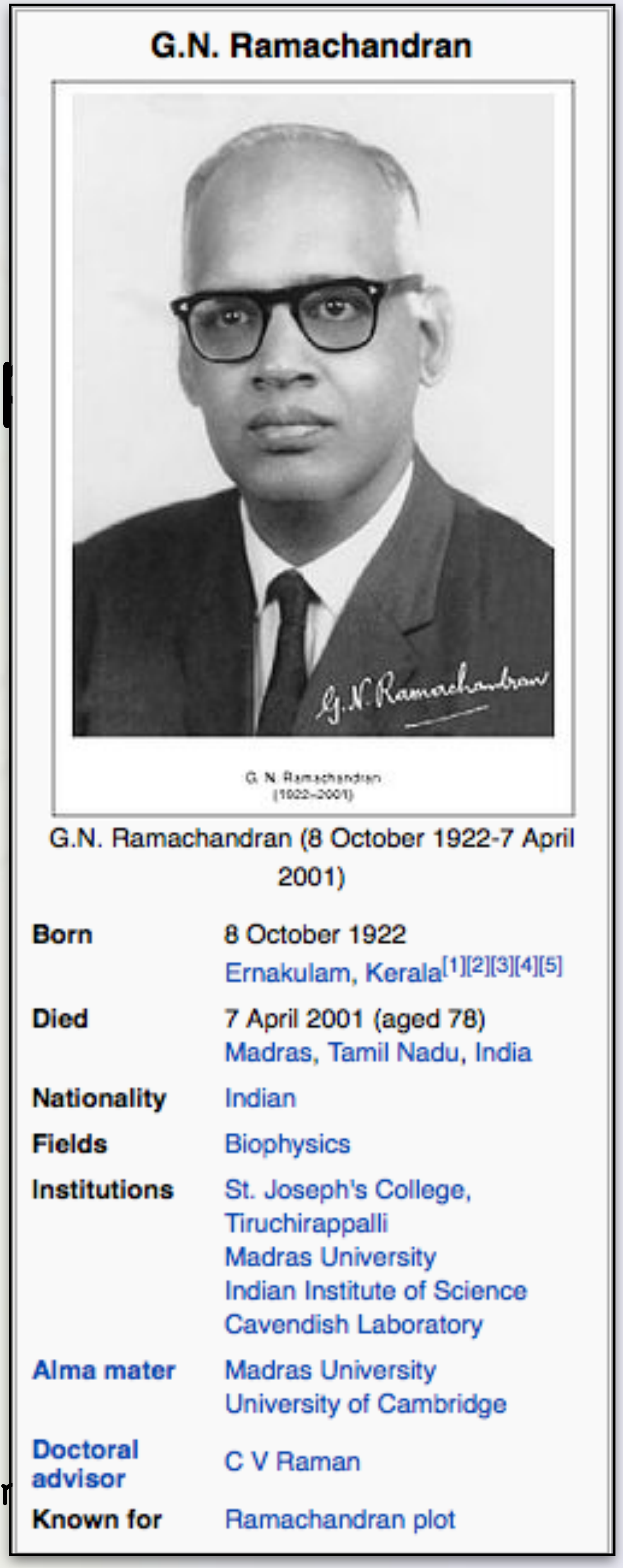
Conformational Restrictions

- The Ramachandran Plot
  - Plots the ψ versus the φ backbone dihedral angle for each residue in a polypeptide.



Conformational Restrictions

- The Ramachandran Plot
  - Plots the ψ versus the φ backbone angle for each residue in a polype



G.N. Ramachandran

G.N. Ramachandran

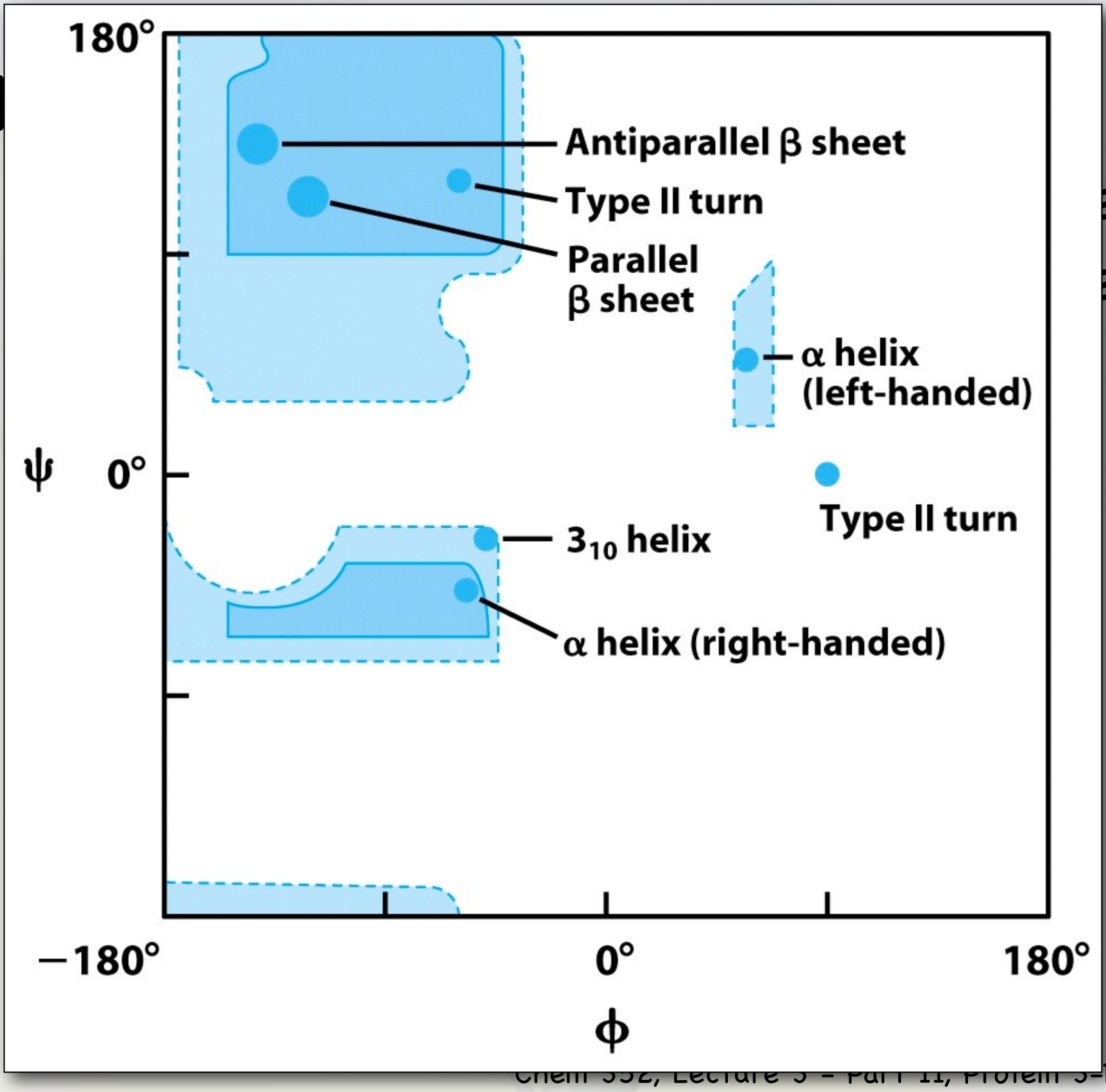
G. N. Ramachandran (1922–2001)

G.N. Ramachandran (8 October 1922-7 April 2001)

Born	8 October 1922 Ernakulam, Kerala[1][2][3][4][5]
Died	7 April 2001 (aged 78) Madras, Tamil Nadu, India
Nationality	Indian
Fields	Biophysics
Institutions	St. Joseph's College, Tiruchirappalli Madras University Indian Institute of Science Cavendish Laboratory
Alma mater	Madras University University of Cambridge
Doctoral advisor	C V Raman
Known for	Ramachandran plot

Conformational Restrictions

- Th
  - edral
    e.

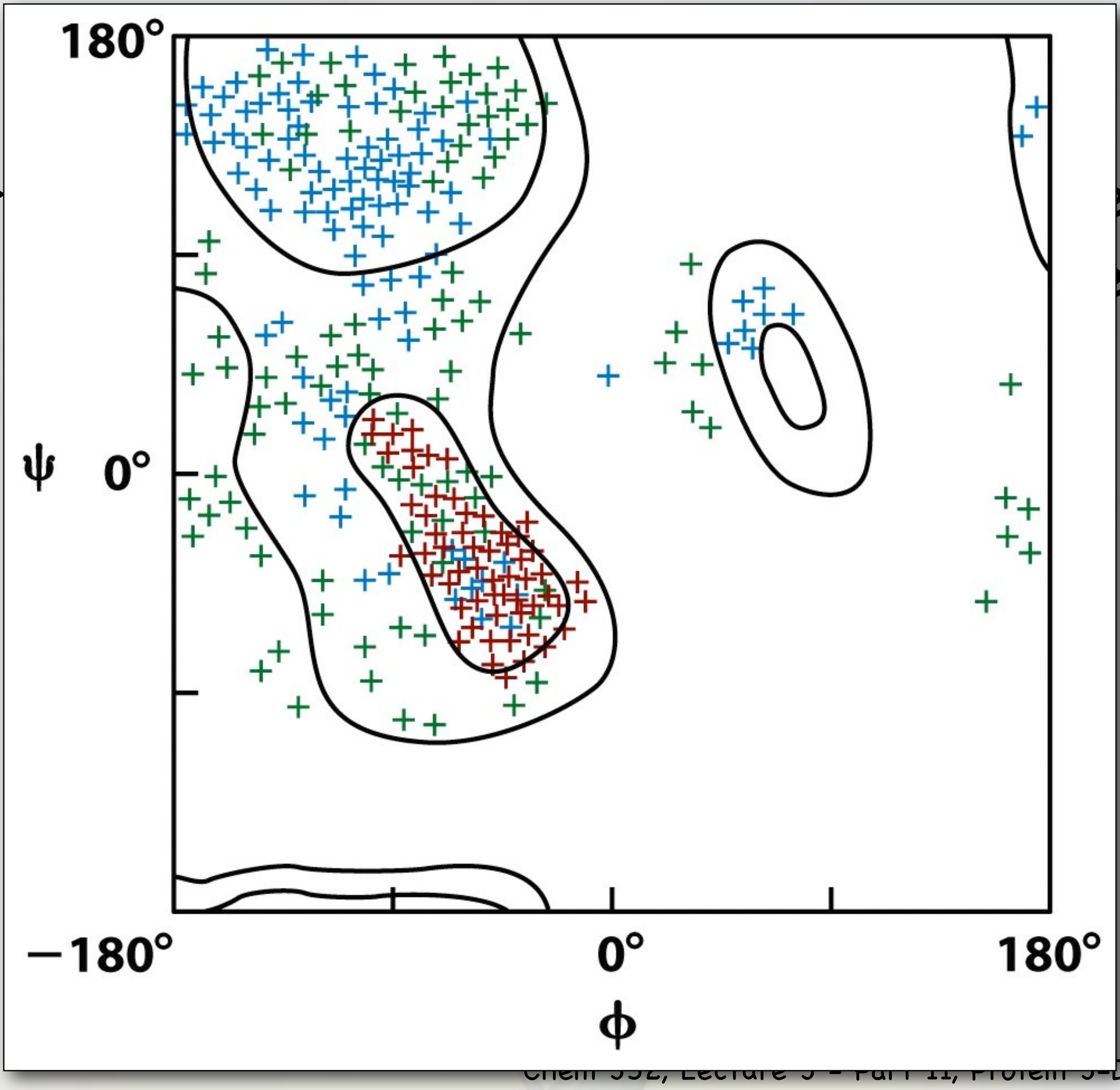


Conformational Restrictions

- T

  ✦ edral

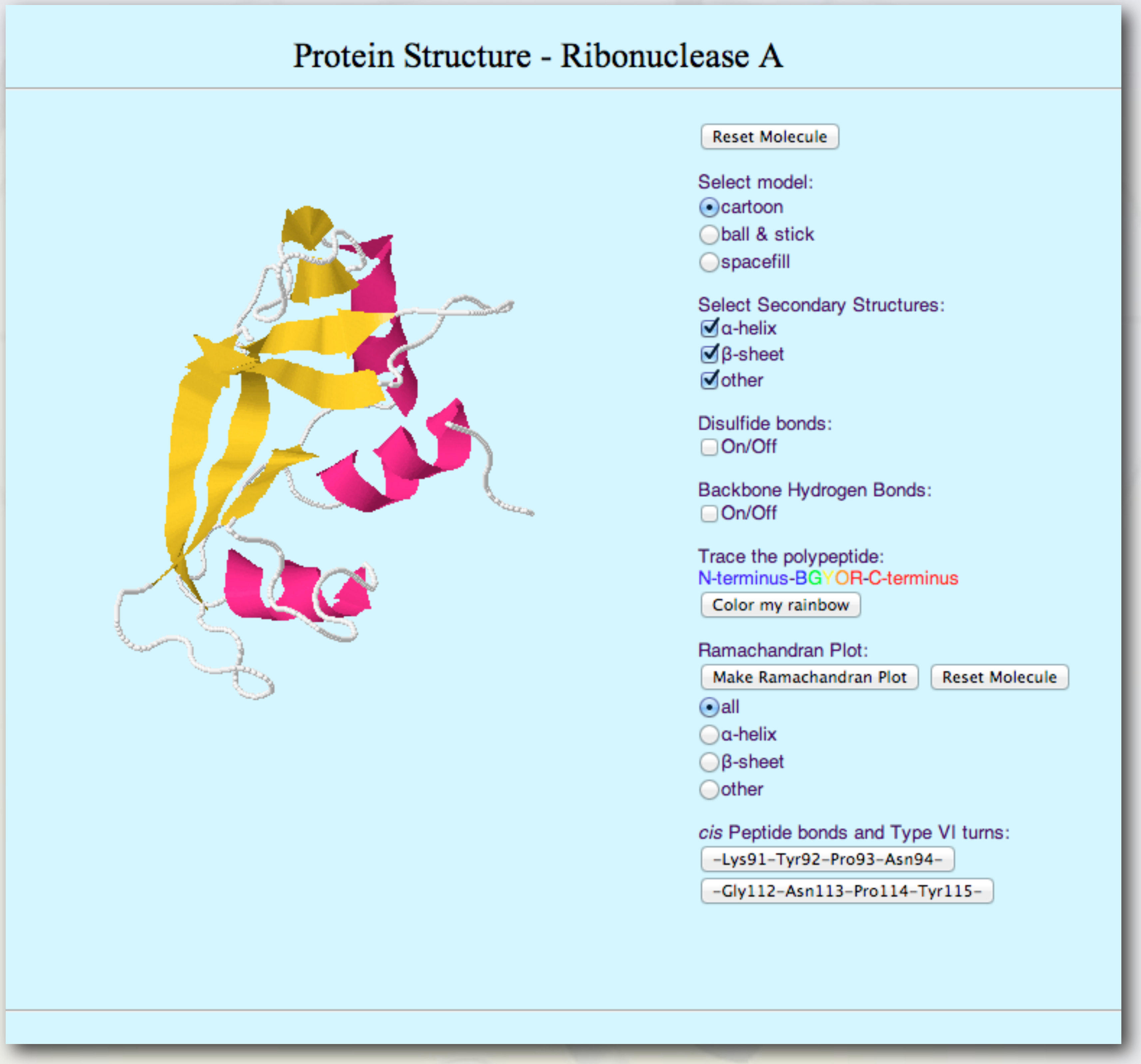
  e.



Conformational Restrictions

- The Ramachandran Plot
  - Plots the ψ versus the φ backbone dihedral angle for each residue in a polypeptide.

Conformational Restrictions



Protein Structure - Ribonuclease A

Select model:
● cartoon
○ ball & stick
○ spacefill

Select Secondary Structures:
☑ α-helix
☑ β-sheet
☑ other

Disulfide bonds:
☐ On/Off

Backbone Hydrogen Bonds:
☐ On/Off

Trace the polypeptide:
N-terminus-BGYOR-C-terminus

Ramachandran Plot:
● all
○ α-helix
○ β-sheet
○ other

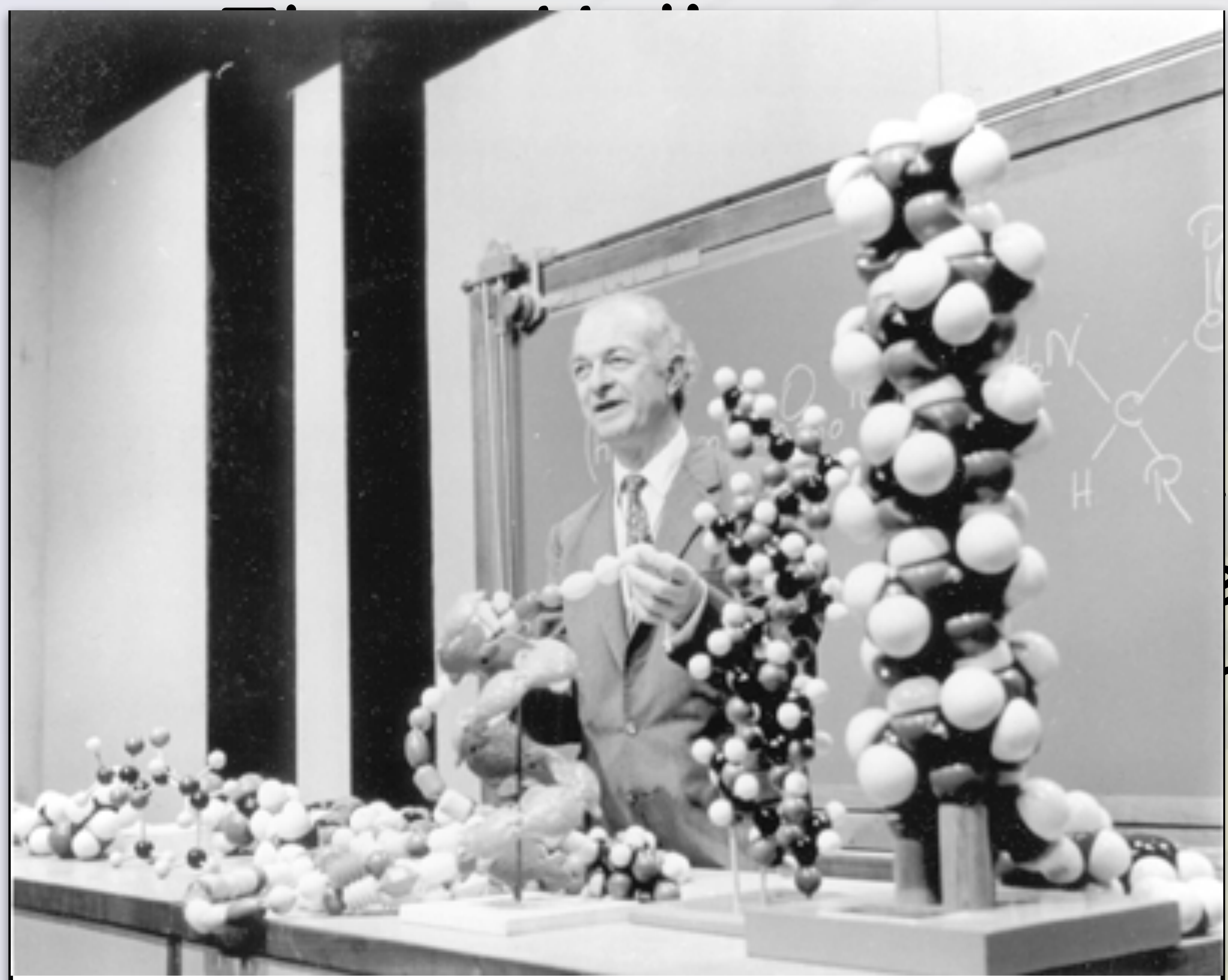
cis Peptide bonds and Type VI turns:

Protein Secondary Structures

The α-Helix

- The backbone is wound into a right-handed cork screw
  - 5.4 Angstroms/turn (the pitch)
  - 1.5 Angstroms/aa (the rise)
- Peptide amide from one turn, hydrogen bonds to the peptide amide from the next turn
- φ and ψ angles are in a favorable region of the Ramachandran plot
- Side chains extend out from the helix axis

Protein Secondary Structures



Linus Pauling (1901-1994)

into a right-handed

he pitch)

rise)

turn, hydrogen bonds

om the next turn

favorable region of

from the helix axis

Protein Secondary Structures

Linus Pauling (1901–1994)

Nobel Prize in Chemistry, 1954
Nobel Prize in Peace, 1962

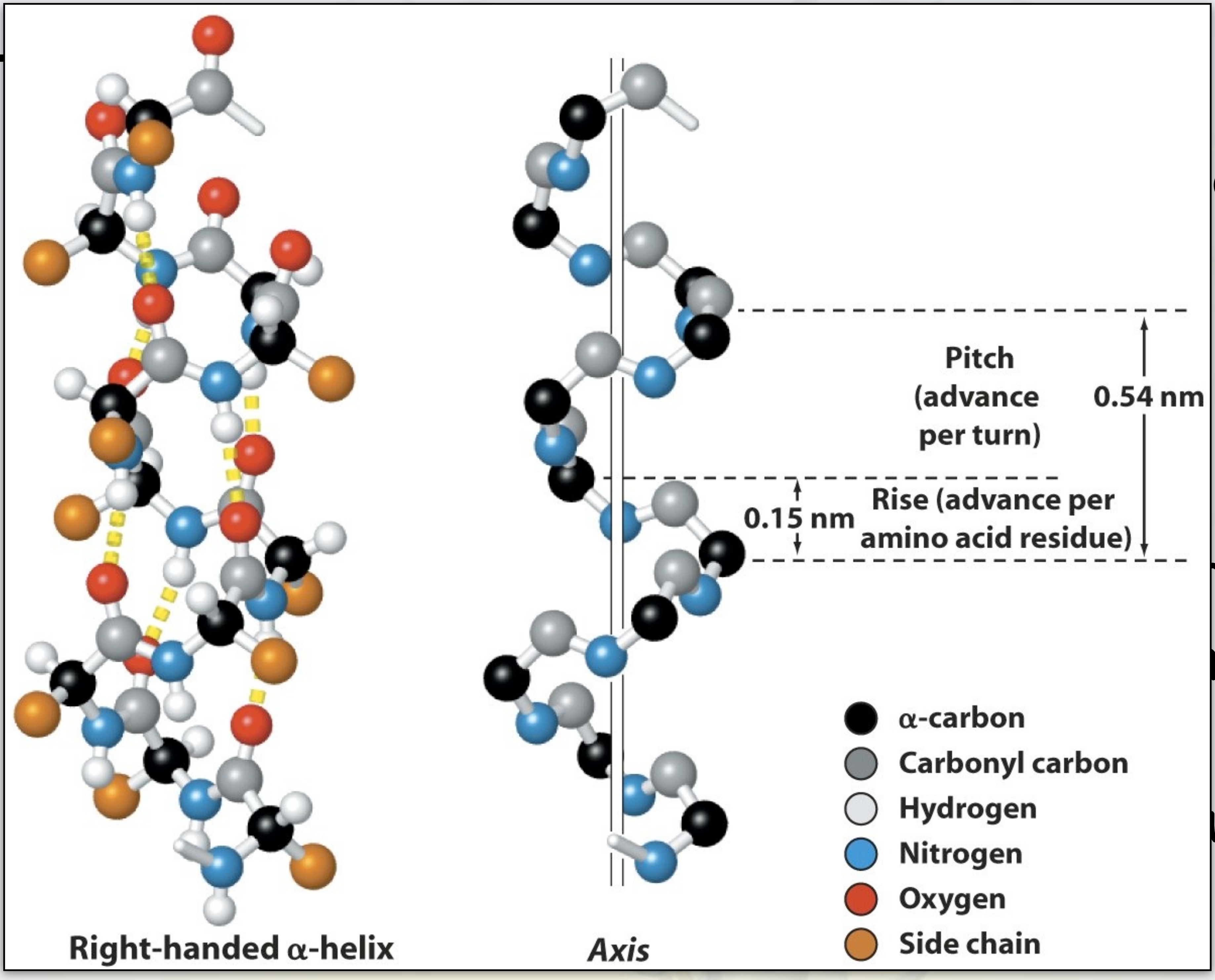
Protein Secondary Structures

The α-Helix

- The backbone is wound into a right-handed cork screw
  - 5.4 Angstroms/turn (the pitch)
  - 1.5 Angstroms/aa (the rise)
- Peptide amide from one turn, hydrogen bonds to the peptide amide from the next turn
- φ and ψ angles are in a favorable region of the Ramachandran plot
- Side chains extend out from the helix axis

Protein Secondary Structures

T



ded

bonds

rn

of

xis

Protein Secondary Structures

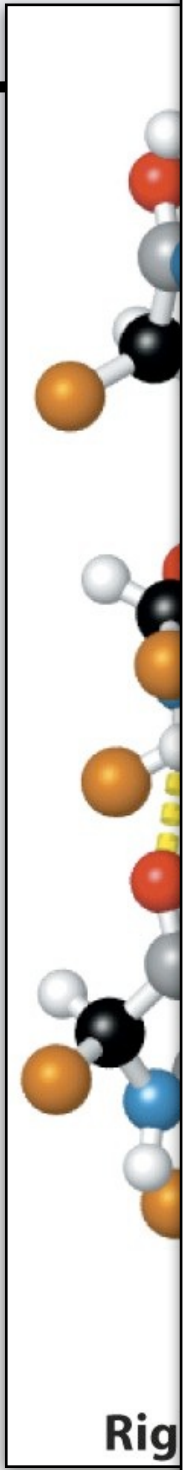
T

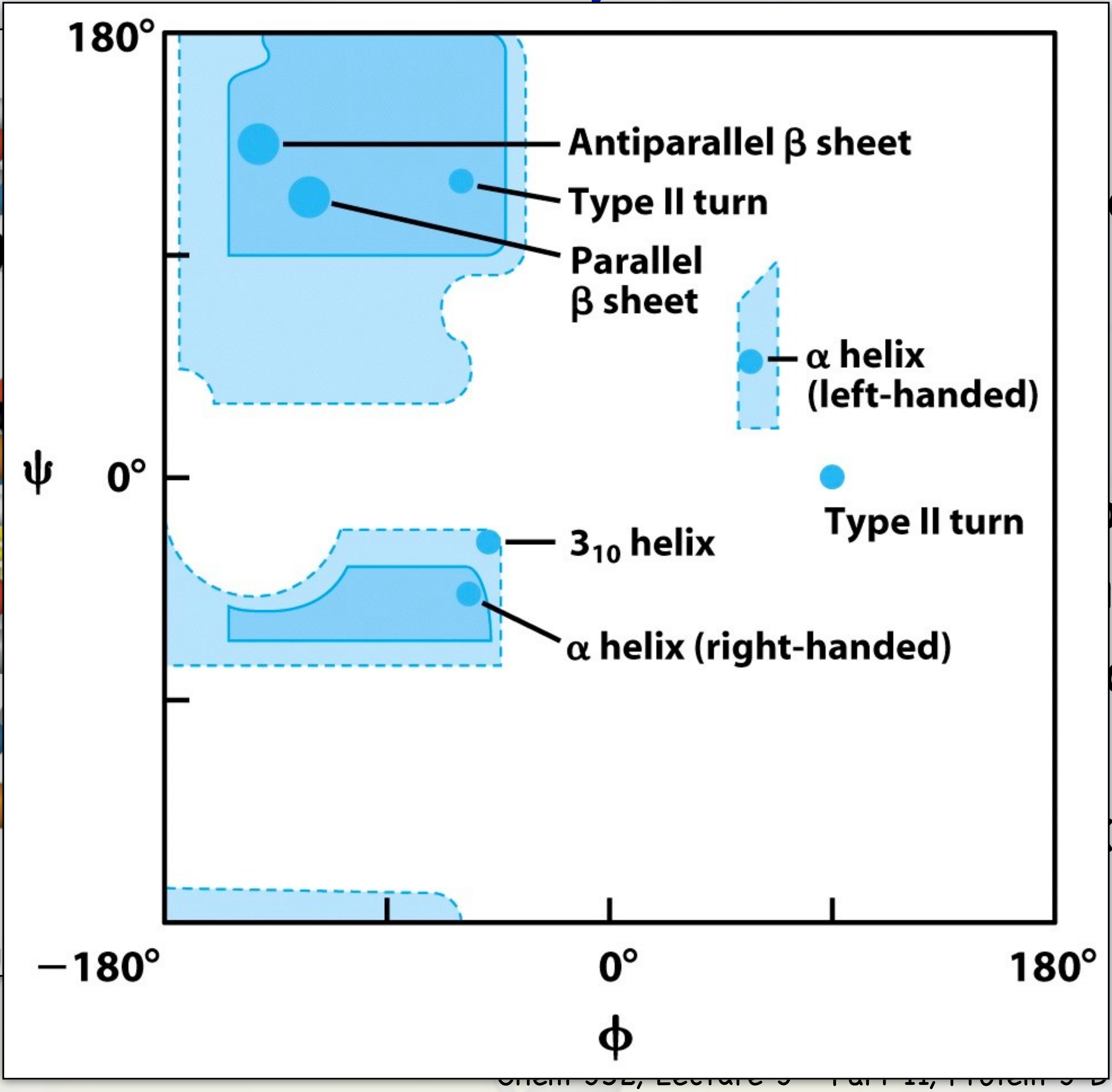
ed

onds

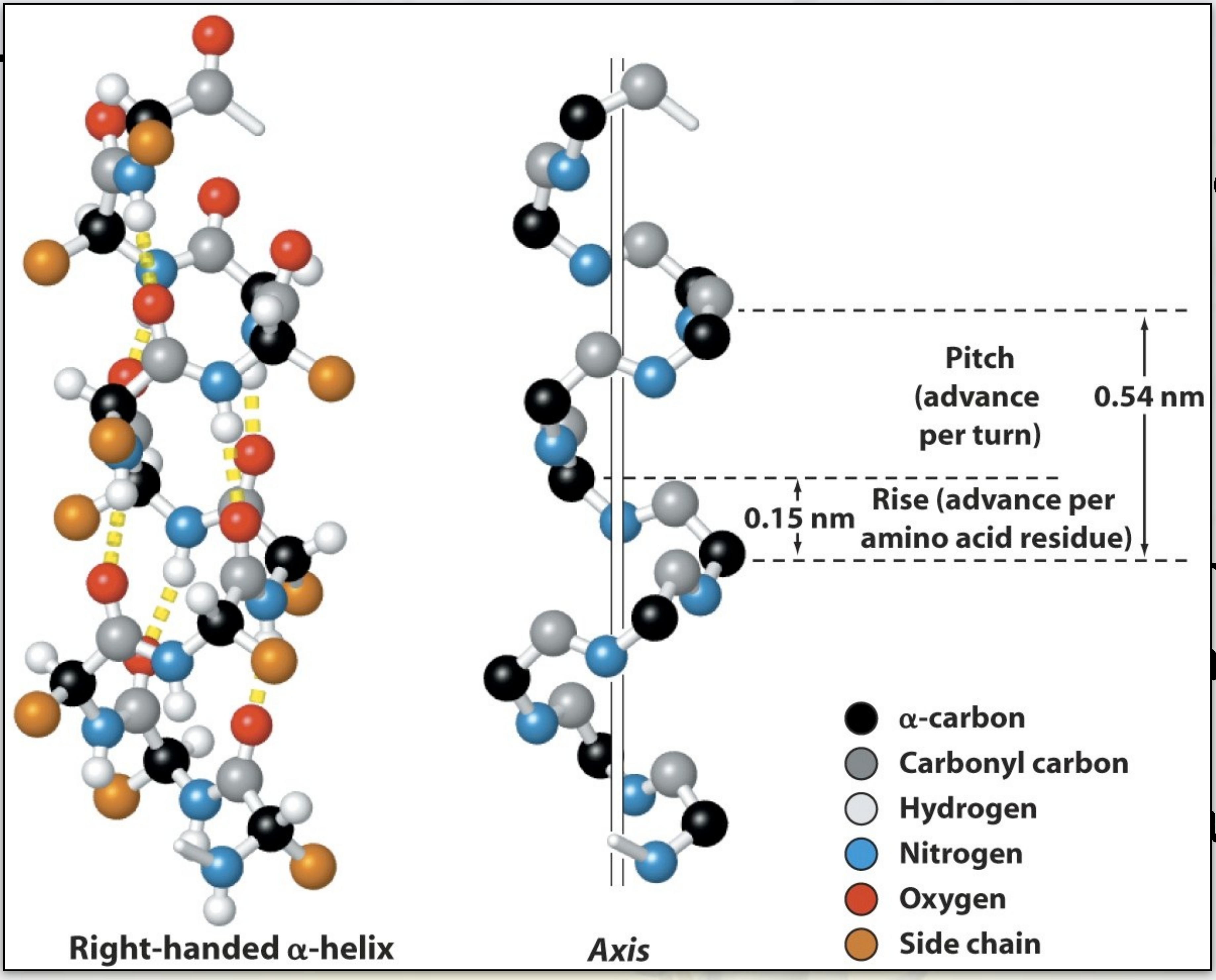
of

is


Rig



Protein Secondary Structures



T

ded

bonds

rn

of

xis

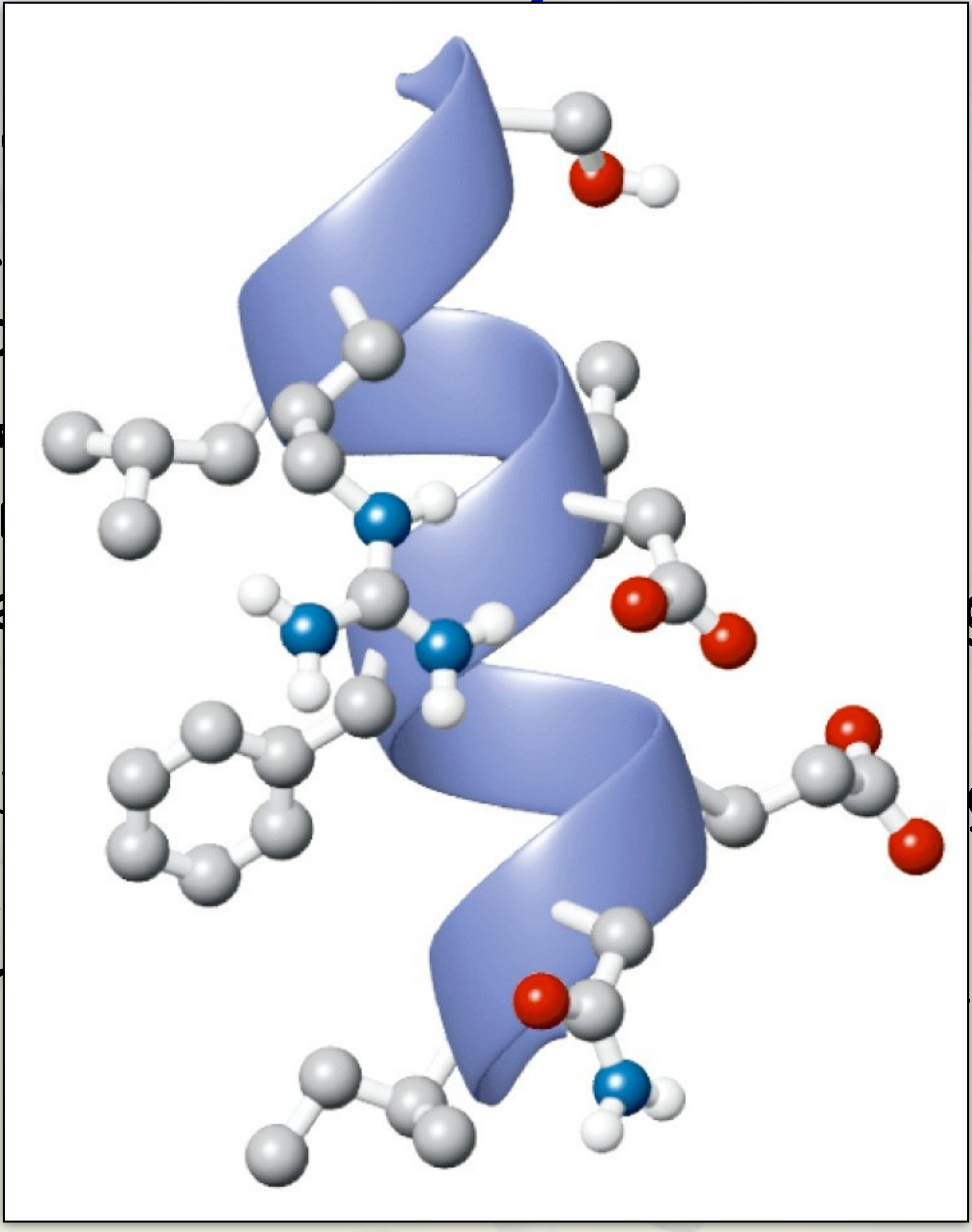
Protein Secondary Structures

The α-Helix

- The backbone is wound into a right-handed cork screw
  - 5.4 Angstroms/turn (the pitch)
  - 1.5 Angstroms/aa (the rise)
- Peptide amide from one turn, hydrogen bonds to the peptide amide from the next turn
- φ and ψ angles are in a favorable region of the Ramachandran plot
- Side chains extend out from the helix axis

Protein Secondary Structures

The α-H



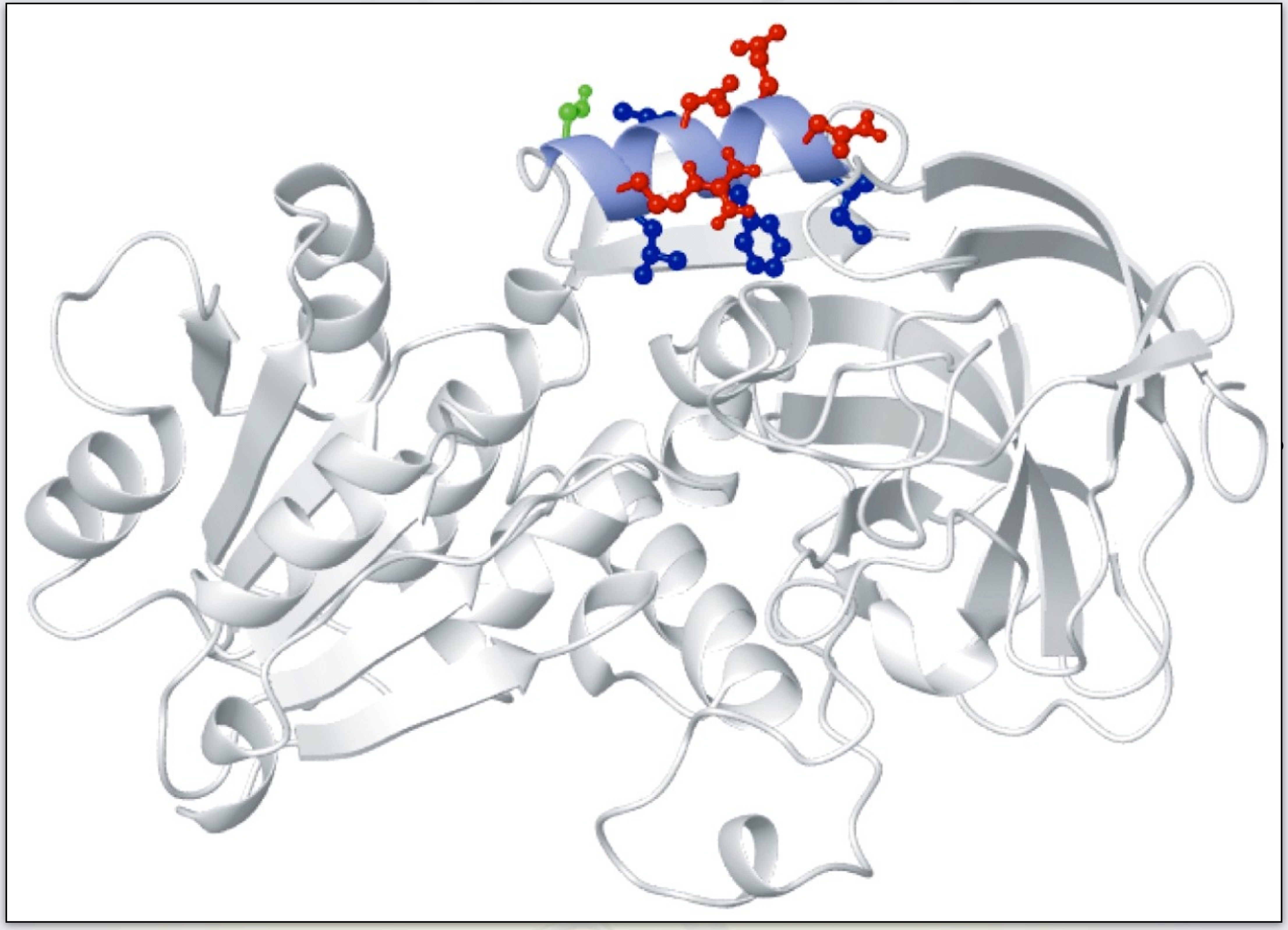
- The ba handed cork sc
  - 5.4 A
  - 1.5 A
- Peptide gen bonds to the turn
- φ and gion of the Ra
- Side ch ix axis

Protein Secondary Structures

The α-Helix

- The backbone is wound into a right-handed cork screw
  - 5.4 Angstroms/turn (the pitch)
  - 1.5 Angstroms/aa (the rise)
- Peptide amide from one turn, hydrogen bonds to the peptide amide from the next turn
- φ and ψ angles are in a favorable region of the Ramachandran plot
- Side chains extend out from the helix axis

Protein Secondary Structures



Protein Secondary Structures

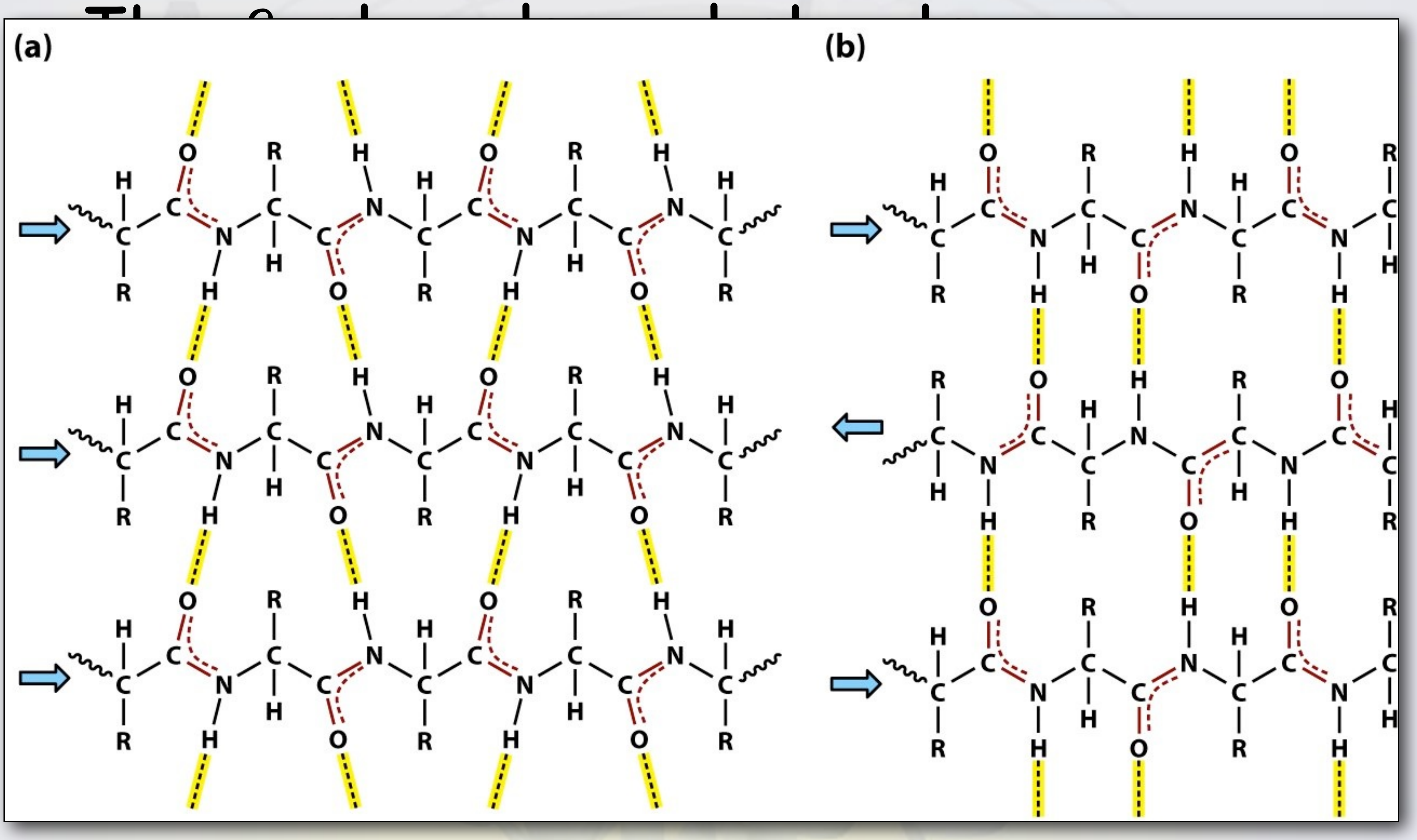
The α-Helix

- The backbone is wound into a right-handed cork screw
  - 5.4 Angstroms/turn (the pitch)
  - 1.5 Angstroms/aa (the rise)
- Peptide amide from one turn, hydrogen bonds to the peptide amide from the next turn
- φ and ψ angles are in a favorable region of the Ramachandran plot
- Side chains extend out from the helix axis

Protein Secondary Structures

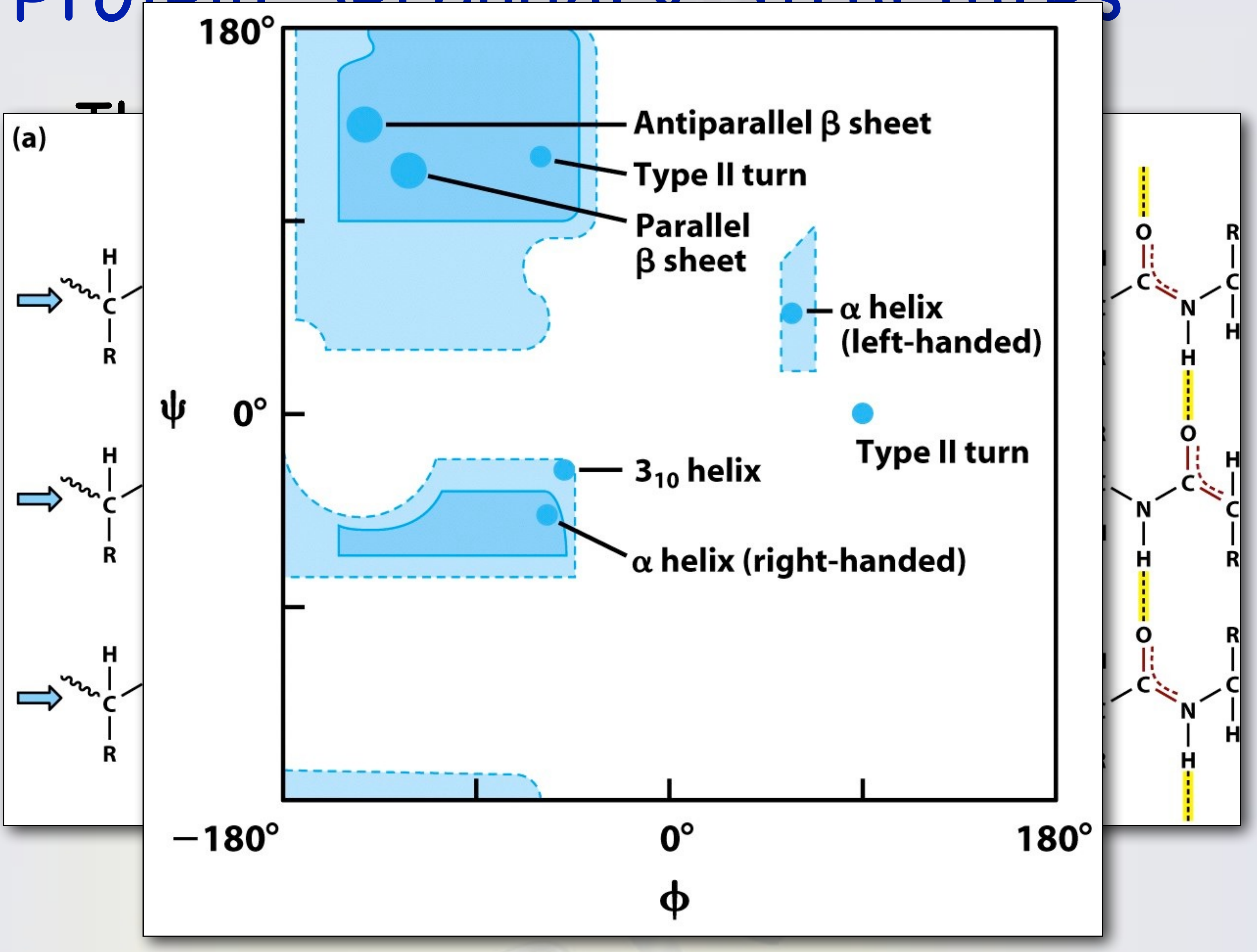
- The β-strands and sheets
  - The polypeptide is nearly fully extended
  - Peptide bond amides form hydrogen bonds with neighboring strands
  - The φ and ψ angles are in a favorable region of the the Ramachandran plot
  - Side chains extend from both sides of the sheet

Protein Secondary Structures

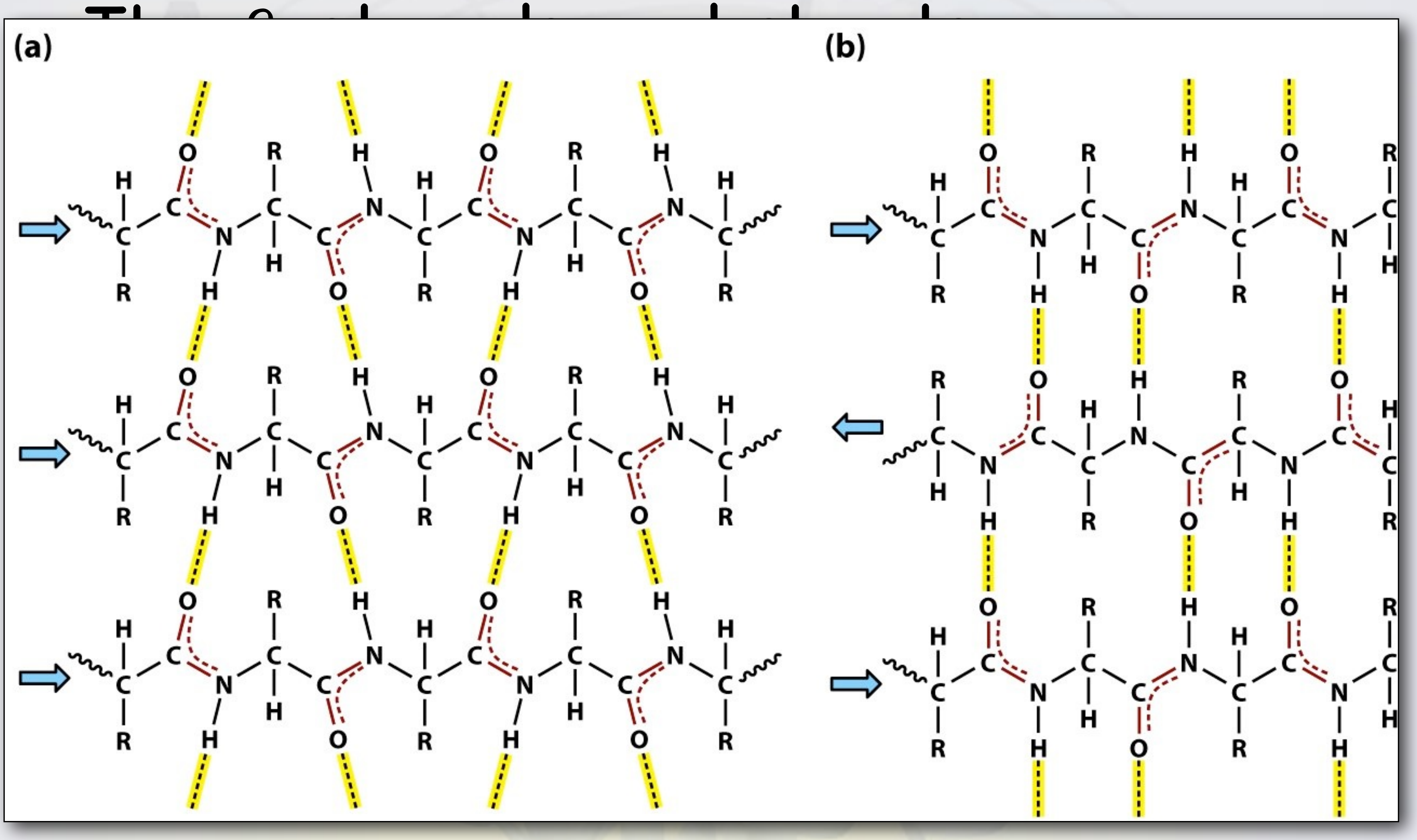


Protein Secondary Structures

Th



Protein Secondary Structures

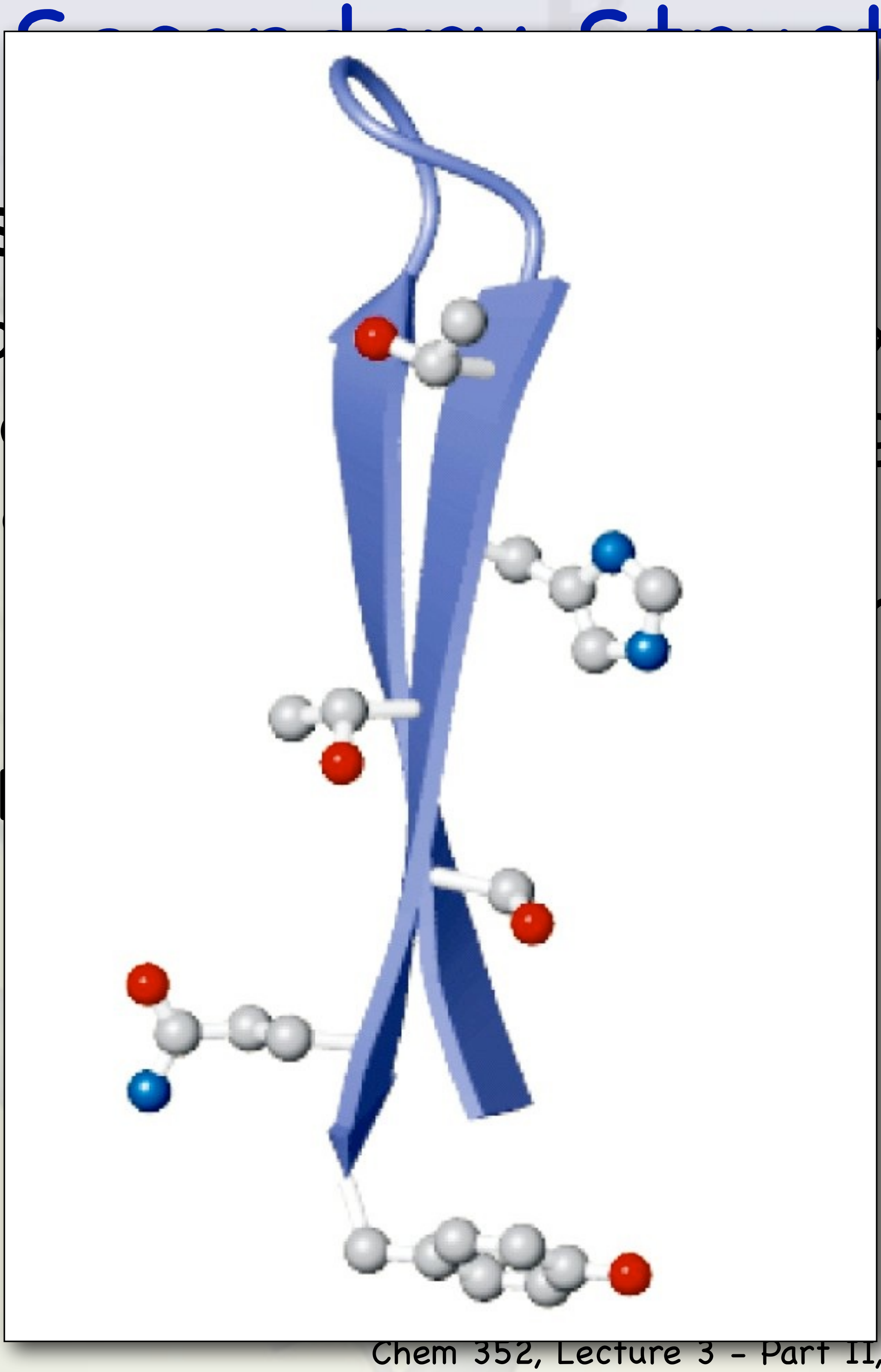


Protein Secondary Structures

- The β-strands and sheets
  - The polypeptide is nearly fully extended
  - Peptide bond amides form hydrogen bonds with neighboring strands
  - The φ and ψ angles are in a favorable region of the the Ramachandran plot
  - Side chains extend from both sides of the sheet

Protein tures

- The β-s
  - The po xtended
  - Peptid gen bonds with n
  - The φ rable region of the
  - Side cl es of the sheet

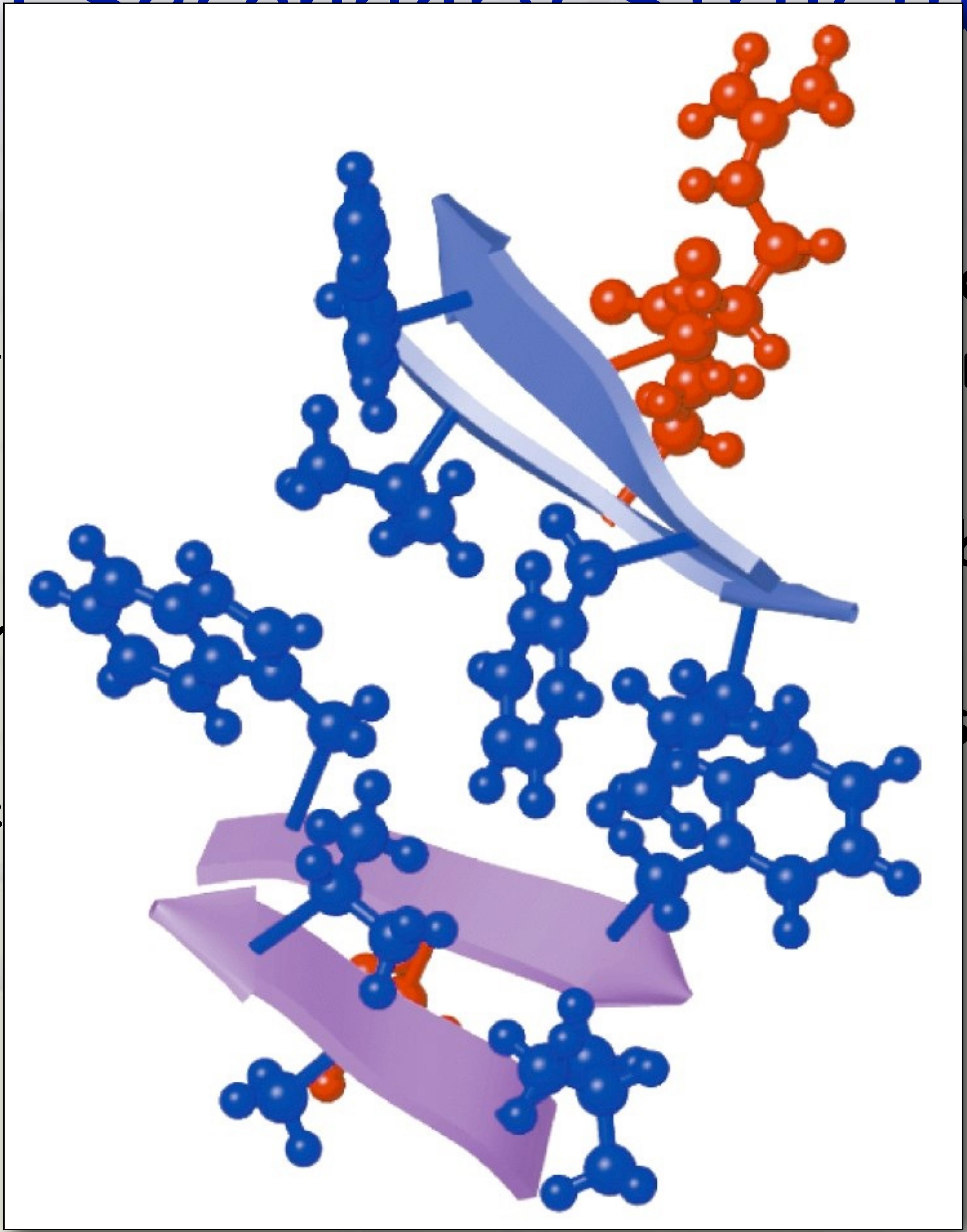


Protein Secondary Structures

- The β-strands and sheets
  - The polypeptide is nearly fully extended
  - Peptide bond amides form hydrogen bonds with neighboring strands
  - The φ and ψ angles are in a favorable region of the the Ramachandran plot
  - Side chains extend from both sides of the sheet

Protein Secondary Structures

- The β
  - The … ended
  - Pept … n bonds with
  - The … ble region of th
  - Side … of the shee

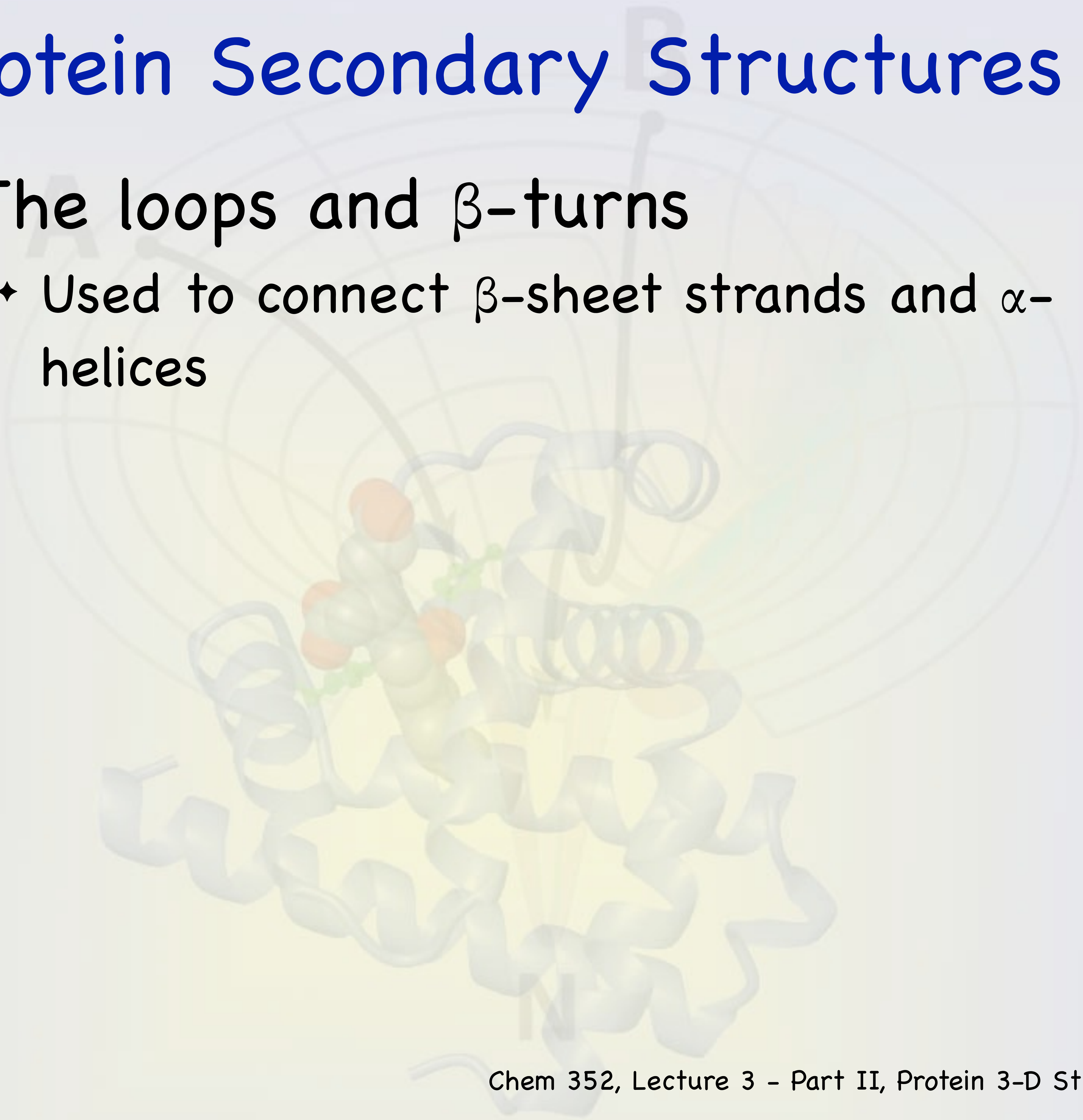


Protein Secondary Structures

- The β-strands and sheets
  - The polypeptide is nearly fully extended
  - Peptide bond amides form hydrogen bonds with neighboring strands
  - The φ and ψ angles are in a favorable region of the the Ramachandran plot
  - Side chains extend from both sides of the sheet

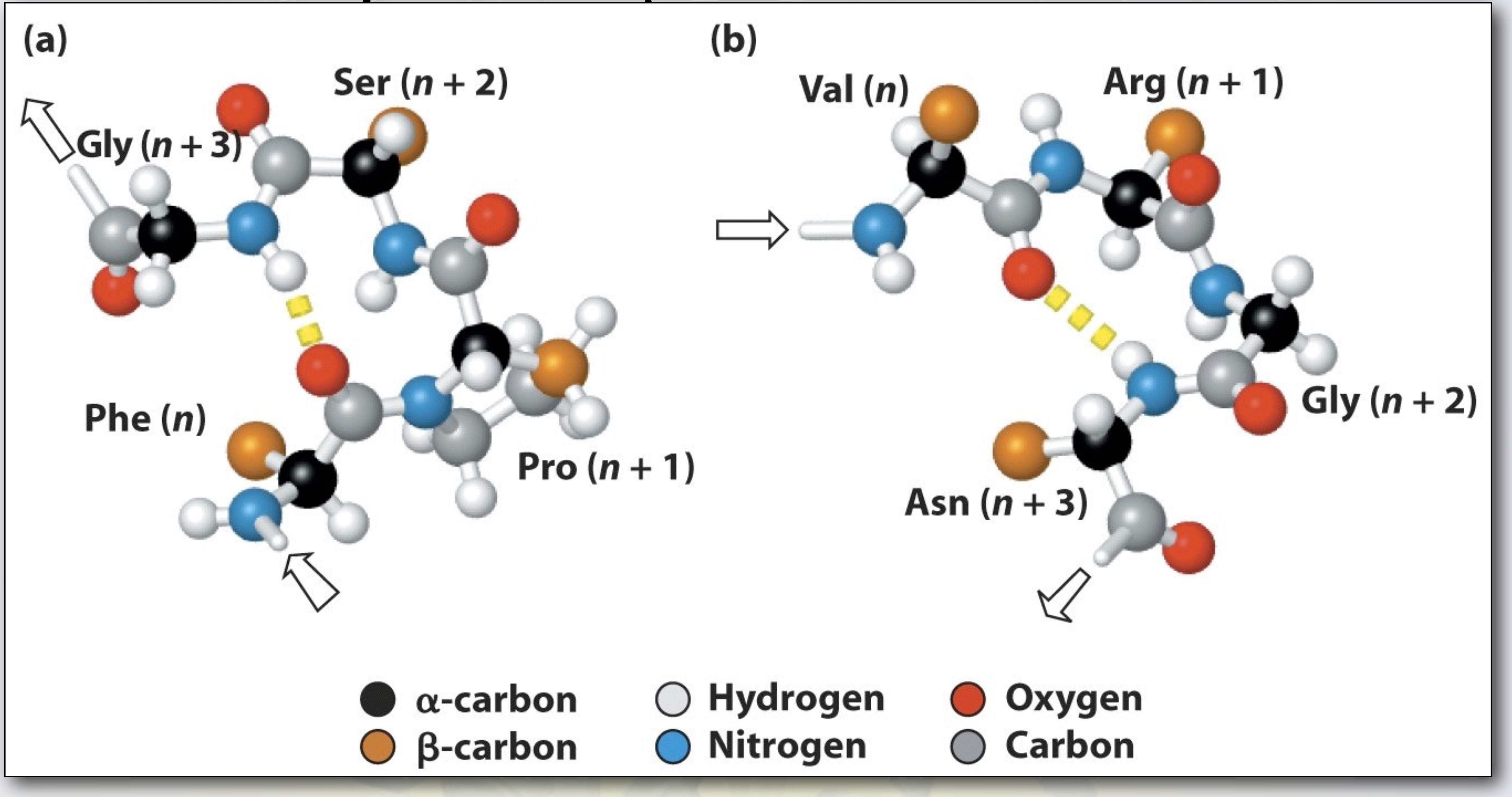
Protein Secondary Structures

- The loops and β-turns
  - Used to connect β-sheet strands and α-helices



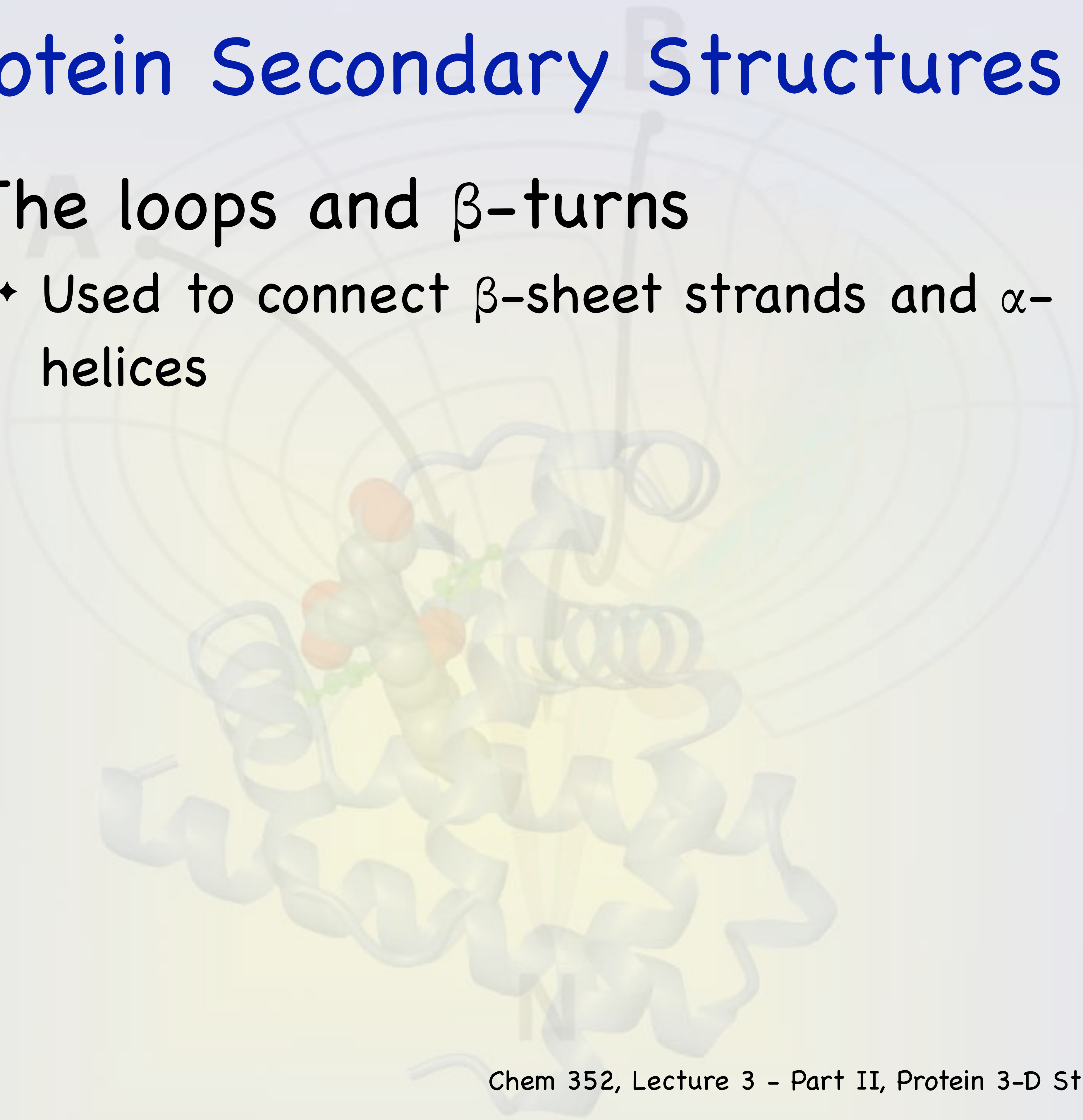
Protein Secondary Structures

- The loops and β-turns



Protein Secondary Structures

- The loops and β-turns
  - Used to connect β-sheet strands and α-helices

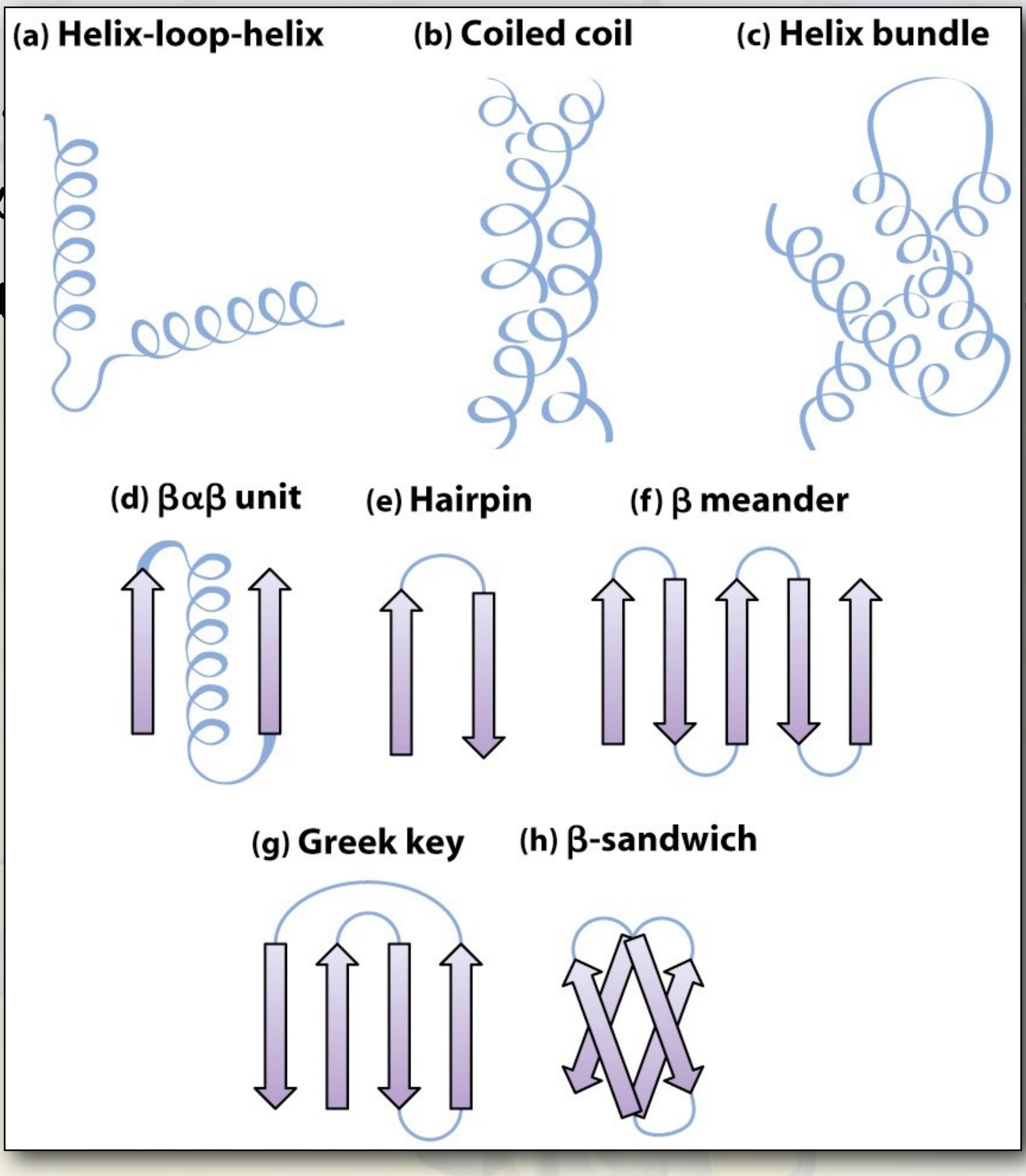


Protein Secondary Structure

- Motifs,
  - The α-helices and β-sheet strands can produce recognizable supersecondary patterns called **motifs**.

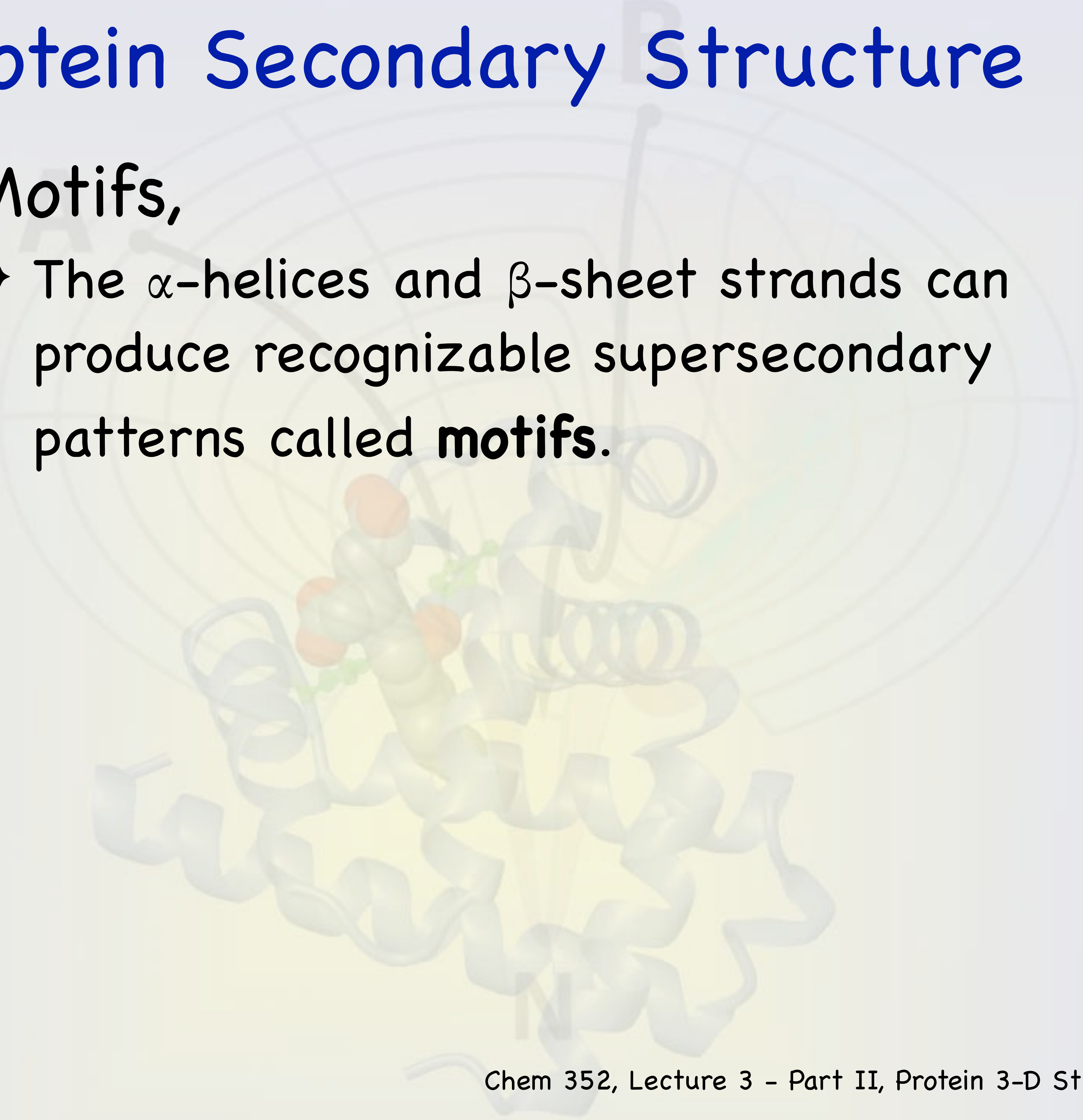
Protein Secondary Structure

- Motifs
  - The α … can
    produ … ary
    patte

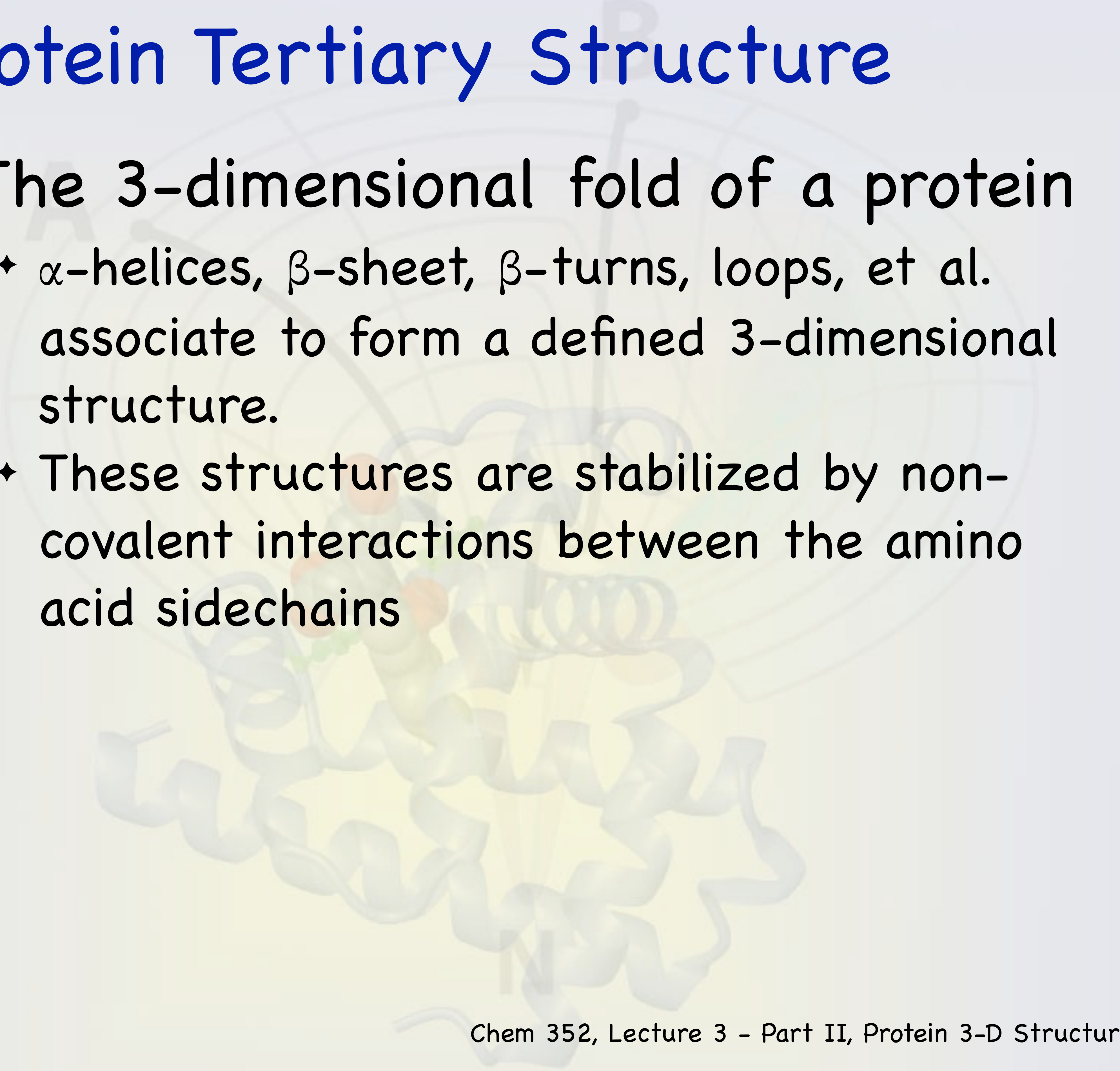


Protein Secondary Structure

- Motifs,
  - The α-helices and β-sheet strands can produce recognizable supersecondary patterns called **motifs**.



Protein Tertiary Structure



- The 3-dimensional fold of a protein
  - α-helices, β-sheet, β-turns, loops, et al. associate to form a defined 3-dimensional structure.
  - These structures are stabilized by non-covalent interactions between the amino acid sidechains

Protein Tertiary Structure

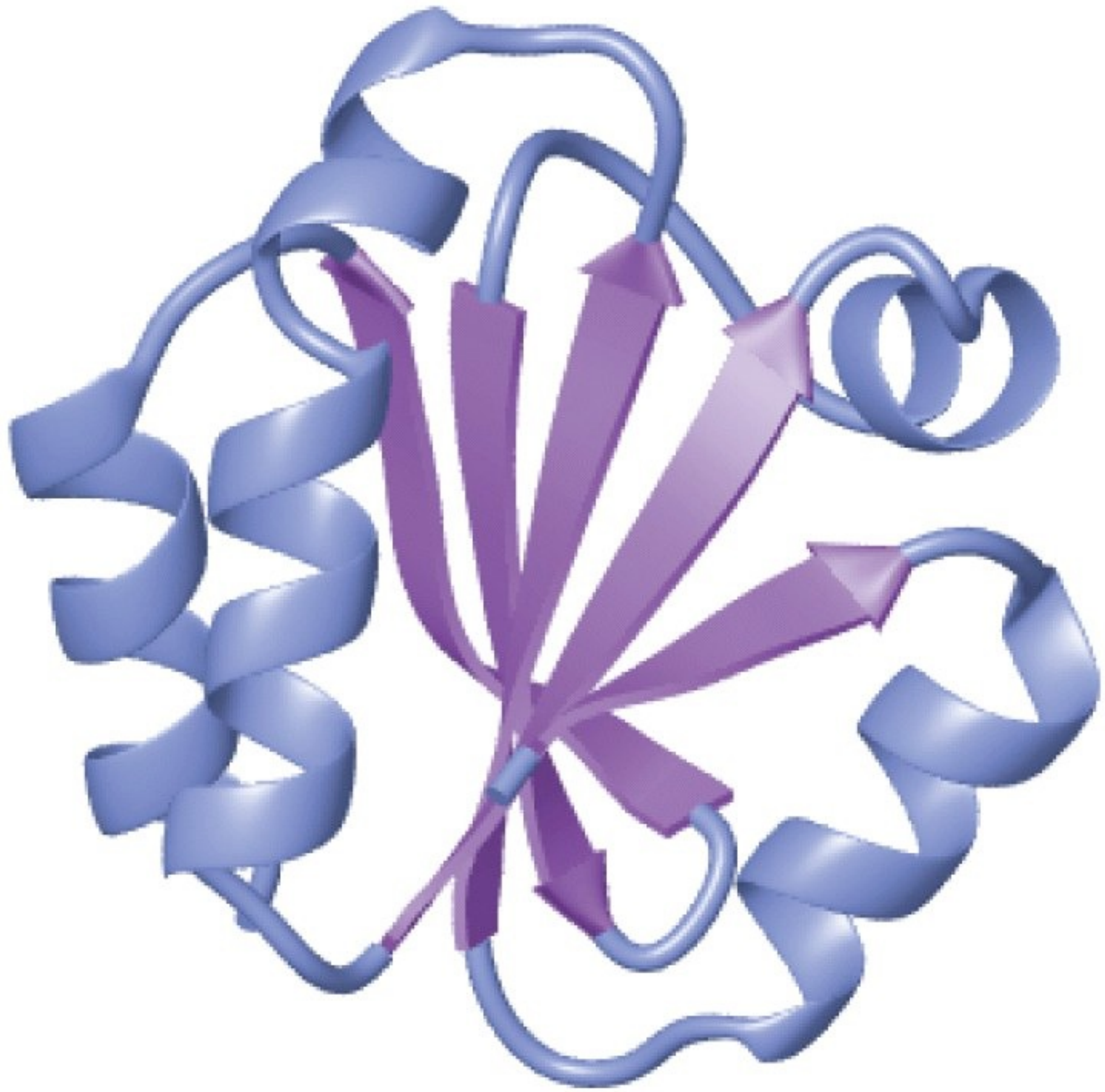
- The otein
  - α-h al.

    ass sional

    str
  - The on-

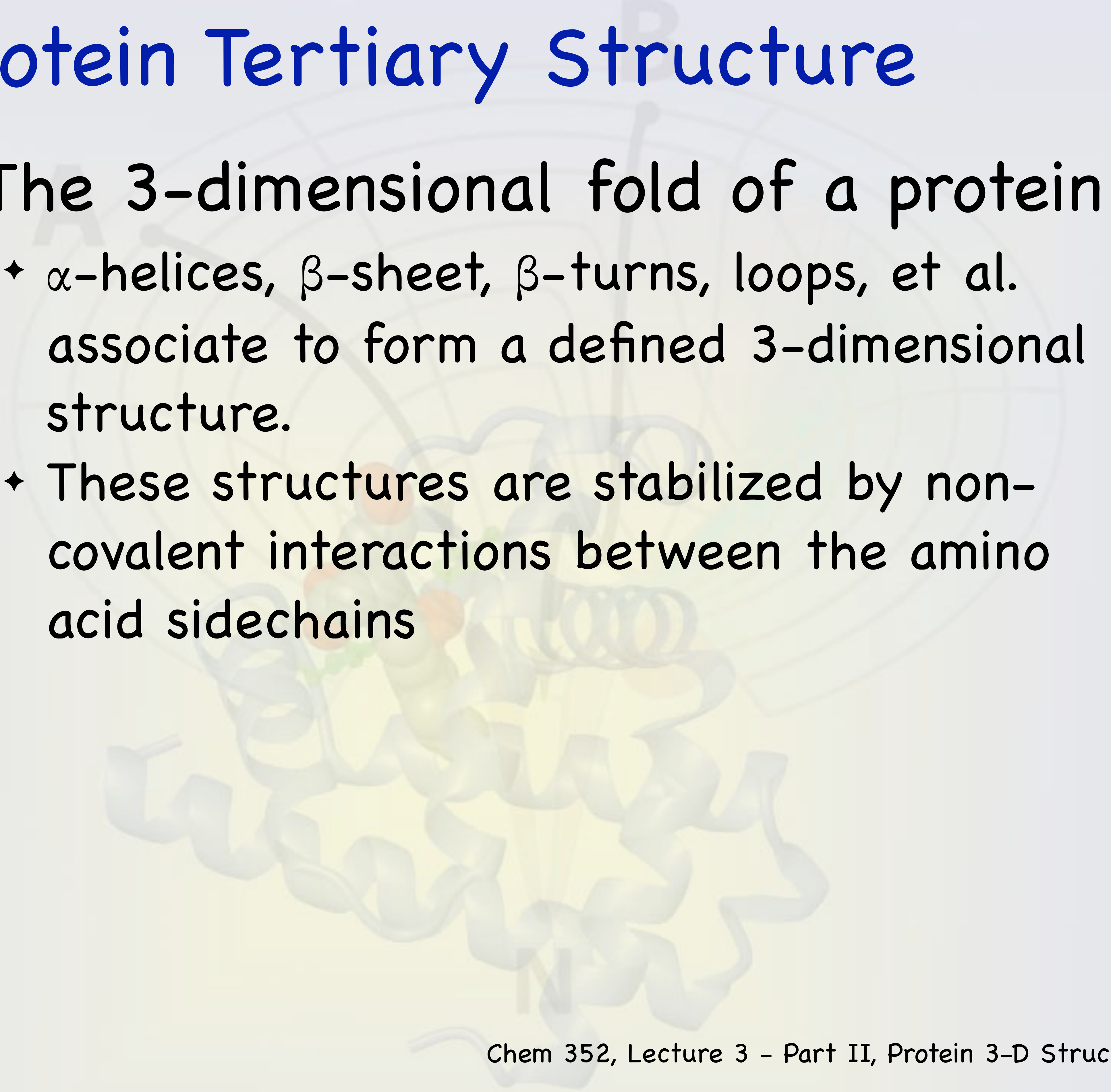
    cov mino

    aci



Human thioredoxin (1ERU)

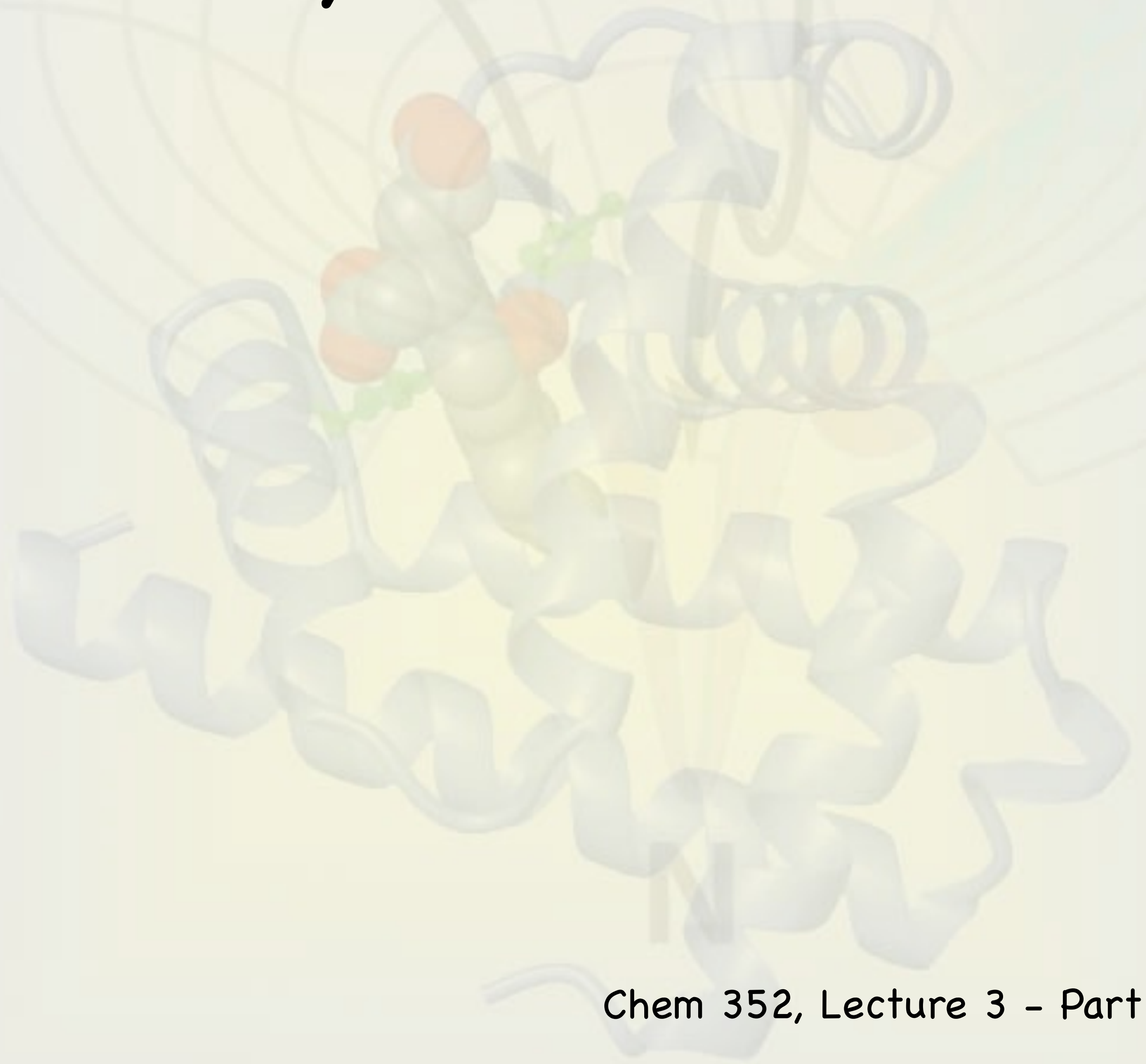
Protein Tertiary Structure

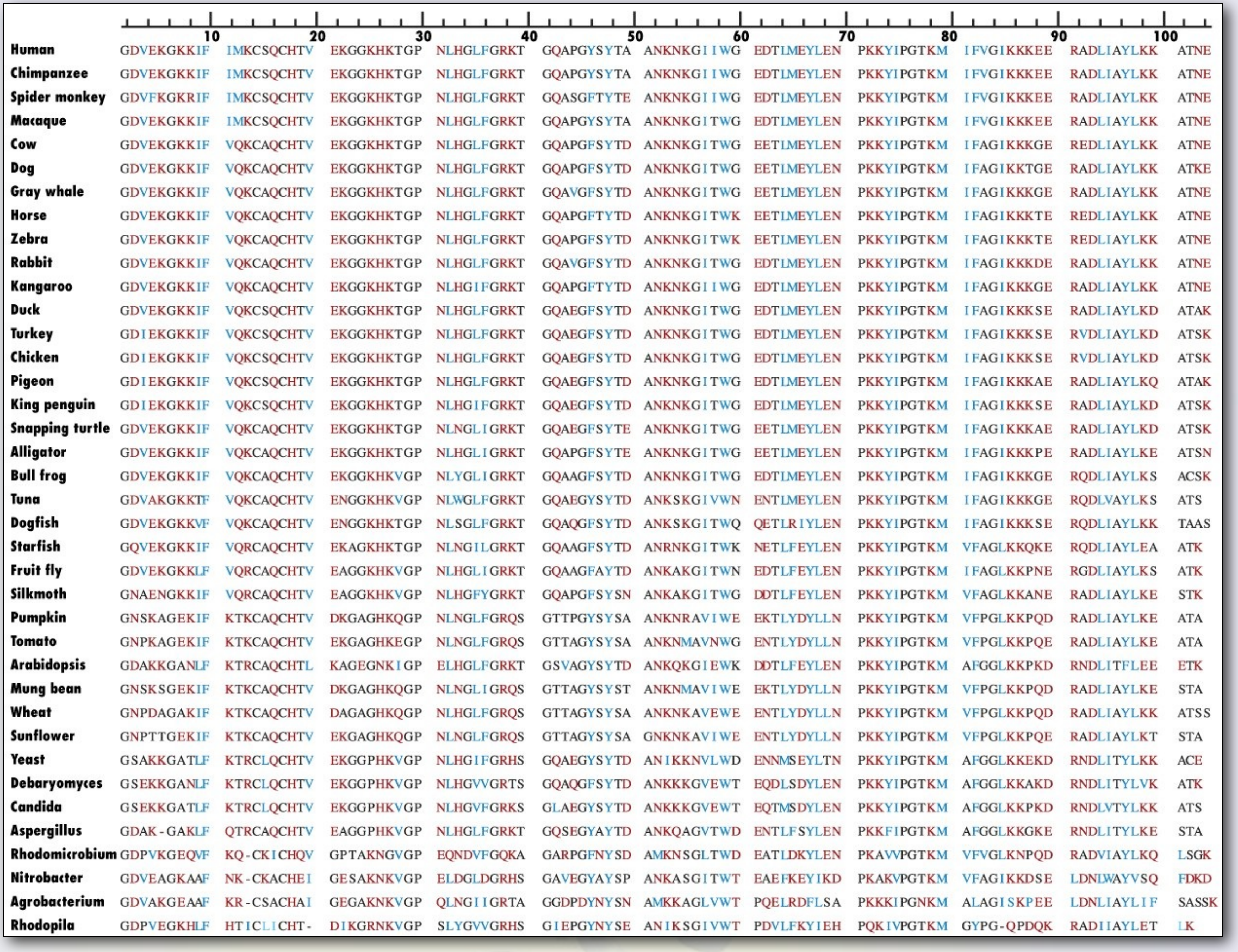


- The 3-dimensional fold of a protein
  - α-helices, β-sheet, β-turns, loops, et al. associate to form a defined 3-dimensional structure.
  - These structures are stabilized by non-covalent interactions between the amino acid sidechains

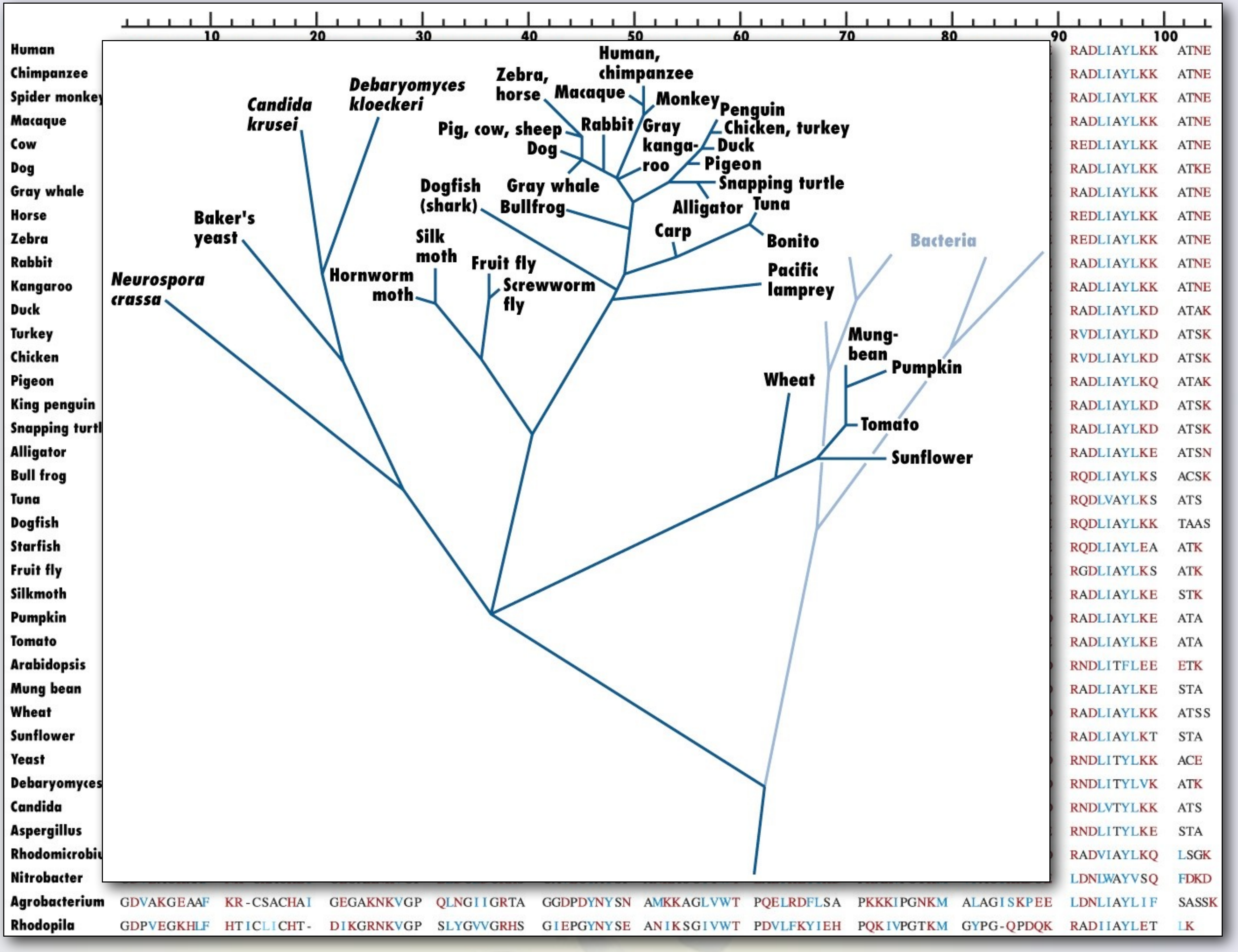
Protein Tertiary Structure

Homologous proteins have similar tertiary structures, which are evolutionarily conserved.



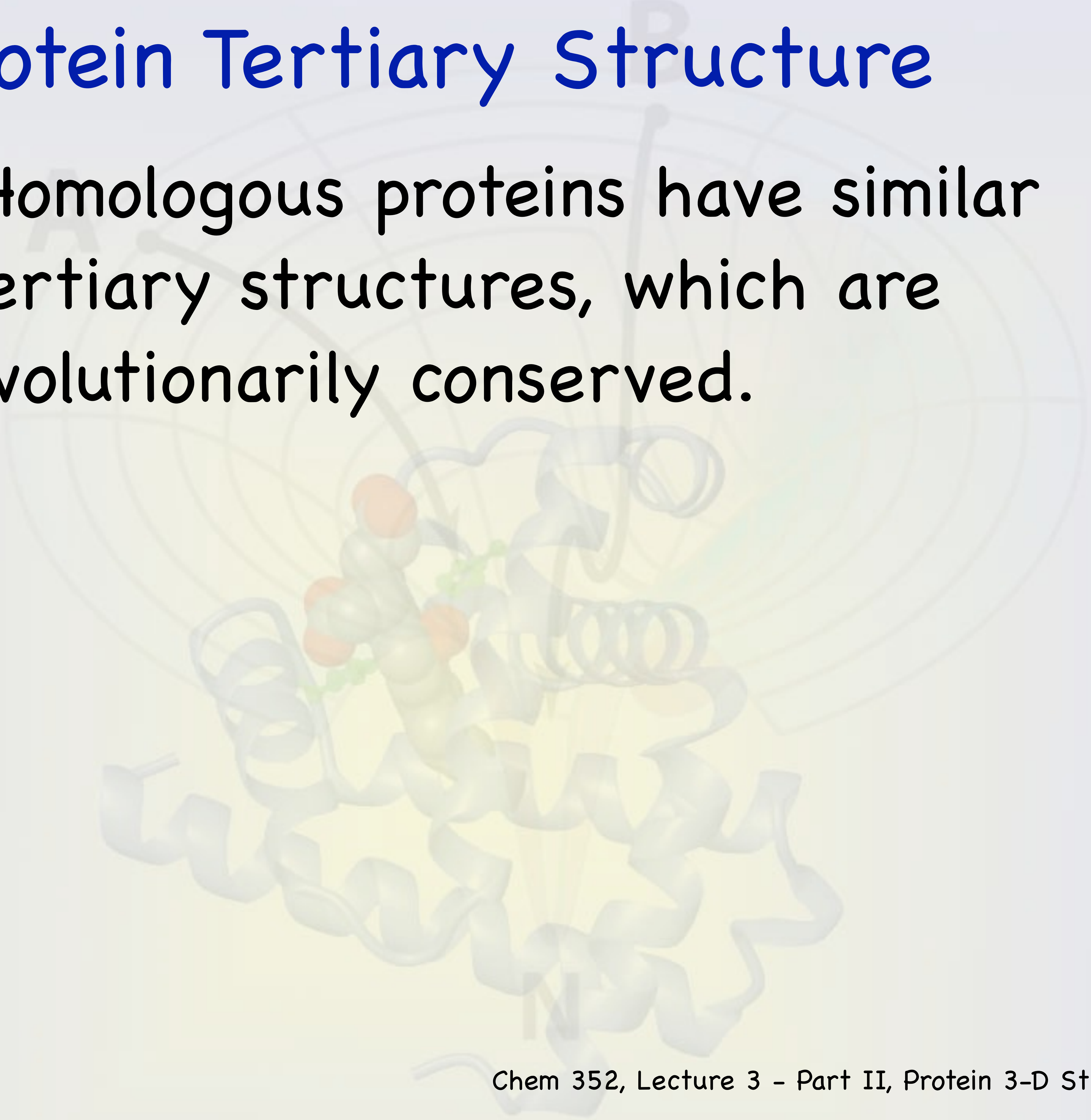


	10	20	30	40	50	60	70	80	90	100	
Human	GDVEKGKKIF	IMKCSQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAPGYSYTA	ANKNKGIIWG	EDTLMEYLEN	PKKYIPGTKM	IFVGIKKKEE	RADLIAYLKK	ATNE
Chimpanzee	GDVEKGKKIF	IMKCSQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAPGYSYTA	ANKNKGIIWG	EDTLMEYLEN	PKKYIPGTKM	IFVGIKKKEE	RADLIAYLKK	ATNE
Spider monkey	GDVFKGKRIF	IMKCSQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQASGFTYTE	ANKNKGIIWG	EDTLMEYLEN	PKKYIPGTKM	IFVGIKKKEE	RADLIAYLKK	ATNE
Macaque	GDVEKGKKIF	IMKCSQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAPGYSYTA	ANKNKGITWG	EDTLMEYLEN	PKKYIPGTKM	IFVGIKKKEE	RADLIAYLKK	ATNE
Cow	GDVEKGKKIF	VQKCAQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAPGFSYTD	ANKNKGITWG	EETLMEYLEN	PKKYIPGTKM	IFAGIKKKGE	REDLIAYLKK	ATNE
Dog	GDVEKGKKIF	VQKCAQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAPGFSYTD	ANKNKGITWG	EETLMEYLEN	PKKYIPGTKM	IFAGIKKTGE	RADLIAYLKK	ATKE
Gray whale	GDVEKGKKIF	VQKCAQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAVGFSYTD	ANKNKGITWG	EETLMEYLEN	PKKYIPGTKM	IFAGIKKKGE	RADLIAYLKK	ATNE
Horse	GDVEKGKKIF	VQKCAQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAPGFTYTD	ANKNKGITWK	EETLMEYLEN	PKKYIPGTKM	IFAGIKKKTE	REDLIAYLKK	ATNE
Zebra	GDVEKGKKIF	VQKCAQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAPGFSYTD	ANKNKGITWK	EETLMEYLEN	PKKYIPGTKM	IFAGIKKKTE	REDLIAYLKK	ATNE
Rabbit	GDVEKGKKIF	VQKCAQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAVGFSYTD	ANKNKGITWG	EDTLMEYLEN	PKKYIPGTKM	IFAGIKKKDE	RADLIAYLKK	ATNE
Kangaroo	GDVEKGKKIF	VQKCAQCHTV	EKGGKHKTGP	NLHGIFGRKT	GQAPGFTYTD	ANKNKGIIWG	EDTLMEYLEN	PKKYIPGTKM	IFAGIKKKGE	RADLIAYLKK	ATNE
Duck	GDVEKGKKIF	VQKCSQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAEGFSYTD	ANKNKGITWG	EDTLMEYLEN	PKKYIPGTKM	IFAGIKKKSE	RADLIAYLKD	ATAK
Turkey	GDIEKGKKIF	VQKCSQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAEGFSYTD	ANKNKGITWG	EDTLMEYLEN	PKKYIPGTKM	IFAGIKKKSE	RVDLIAYLKD	ATSK
Chicken	GDIEKGKKIF	VQKCSQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAEGFSYTD	ANKNKGITWG	EDTLMEYLEN	PKKYIPGTKM	IFAGIKKKSE	RVDLIAYLKD	ATSK
Pigeon	GDIEKGKKIF	VQKCSQCHTV	EKGGKHKTGP	NLHGLFGRKT	GQAEGFSYTD	ANKNKGITWG	EDTLMEYLEN	PKKYIPGTKM	IFAGIKKKAE	RADLIAYLKQ	ATAK
King penguin	GDIEKGKKIF	VQKCSQCHTV	EKGGKHKTGP	NLHGIFGRKT	GQAEGFSYTD	ANKNKGITWG	EDTLMEYLEN	PKKYIPGTKM	IFAGIKKKSE	RADLIAYLKD	ATSK
Snapping turtle	GDVEKGKKIF	VQKCAQCHTV	EKGGKHKTGP	NLNGLIGRKT	GQAEGFSYTE	ANKNKGITWG	EETLMEYLEN	PKKYIPGTKM	IFAGIKKKAE	RADLIAYLKD	ATSK
Alligator	GDVEKGKKIF	VQKCAQCHTV	EKGGKHKTGP	NLHGLIGRKT	GQAPGFSYTE	ANKNKGITWG	EETLMEYLEN	PKKYIPGTKM	IFAGIKKKPE	RADLIAYLKE	ATSN
Bull frog	GDVEKGKKIF	VQKCAQCHTV	EKGGKHKVGP	NLYGLIGRKT	GQAAGFSYTD	ANKNKGITWG	EDTLMEYLEN	PKKYIPGTKM	IFAGIKKKGE	RQDLIAYLKS	ACSK
Tuna	GDVAKGKKTF	VQKCAQCHTV	ENGGKHKVGP	NLWGLFGRKT	GQAEGYSYTD	ANKSKGIVWN	ENTLMEYLEN	PKKYIPGTKM	IFAGIKKKGE	RQDLVAYLKS	ATS
Dogfish	GDVEKGKKVF	VQKCAQCHTV	ENGGKHKTGP	NLSGLFGRKT	GQAQGFSYTD	ANKSKGITWQ	QETLRIYLEN	PKKYIPGTKM	IFAGIKKKSE	RQDLIAYLKK	TAAS
Starfish	GQVEKGKKIF	VQRCAQCHTV	EKAGKHKTGP	NLNGILGRKT	GQAAGFSYTD	ANRNKGITWK	NETLFEYLEN	PKKYIPGTKM	VFAGLKKQKE	RQDLIAYLEA	ATK
Fruit fly	GDVEKGKKLF	VQRCAQCHTV	EAGGKHKVGP	NLHGLIGRKT	GQAAGFAYTD	ANKAKGITWN	EDTLFEYLEN	PKKYIPGTKM	IFAGLKKPNE	RGDLIAYLKS	ATK
Silkmoth	GNAENGKKIF	VQRCAQCHTV	EAGGKHKVGP	NLHGFYGRKT	GQAPGFSYSN	ANKAKGITWG	DDTLFEYLEN	PKKYIPGTKM	VFAGLKKANE	RADLIAYLKE	STK
Pumpkin	GNSKAGEKIF	KTKCAQCHTV	DKGAGHKQGP	NLNGLFGRQS	GTTPGYSYSA	ANKNRAVIWE	EKTLYDYLLN	PKKYIPGTKM	VFPGLKKPQD	RADLIAYLKE	ATA
Tomato	GNPKAGEKIF	KTKCAQCHTV	EKGAGHKEGP	NLNGLFGRQS	GTTAGYSYSA	ANKNMAVNWG	ENTLYDYLLN	PKKYIPGTKM	VFPGLKKPQE	RADLIAYLKE	ATA
Arabidopsis	GDAKKGANLF	KTRCAQCHTL	KAGEGNKIGP	ELHGLFGRKT	GSVAGYSYTD	ANKQKGIEWK	DDTLFEYLEN	PKKYIPGTKM	AFGGLKKPKD	RNDLITFLEE	ETK
Mung bean	GNSKSGEKIF	KTKCAQCHTV	DKGAGHKQGP	NLNGLIGRQS	GTTAGYSYST	ANKNMAVIWE	EKTLYDYLLN	PKKYIPGTKM	VFPGLKKPQD	RADLIAYLKE	STA
Wheat	GNPDAGAKIF	KTKCAQCHTV	DAGAGHKQGP	NLHGLFGRQS	GTTAGYSYSA	ANKNKAVEWE	ENTLYDYLLN	PKKYIPGTKM	VFPGLKKPQD	RADLIAYLKK	ATSS
Sunflower	GNPTTGEKIF	KTKCAQCHTV	EKGAGHKQGP	NLNGLFGRQS	GTTAGYSYSA	GNKNKAVIWE	ENTLYDYLLN	PKKYIPGTKM	VFPGLKKPQE	RADLIAYLKT	STA
Yeast	GSAKKGATLF	KTRCLQCHTV	EKGGPHKVGP	NLHGIFGRHS	GQAEGYSYTD	ANIKKNVLWD	ENNMSEYLTN	PKKYIPGTKM	AFGGLKKEKD	RNDLITYLKK	ACE
Debaryomyces	GSEKKGANLF	KTRCLQCHTV	EKGGPHKVGP	NLHGVVGRTS	GQAQGFSYTD	ANKKKGVEWT	EQDLSDYLEN	PKKYIPGTKM	AFGGLKKAKD	RNDLITYLVK	ATK
Candida	GSEKKGATLF	KTRCLQCHTV	EKGGPHKVGP	NLHGVFGRKS	GLAEGYSYTD	ANKKKGVEWT	EQTMSDYLEN	PKKYIPGTKM	AFGGLKKPKD	RNDLVTYLKK	ATS
Aspergillus	GDAK-GAKLF	QTRCAQCHTV	EAGGPHKVGP	NLHGLFGRKT	GQSEGYAYTD	ANKQAGVTWD	ENTLFSYLEN	PKKFIPGTKM	AFGGLKKGKE	RNDLITYLKE	STA
Rhodomicrobium	GDPVKGEQVF	KQ-CKICHQV	GPTAKNGVGP	EQNDVFGQKA	GARPGFNYSD	AMKNSGLTWD	EATLDKYLEN	PKAVVPGTKM	VFVGLKNPQD	RADVIAYLKQ	LSGK
Nitrobacter	GDVEAGKAAF	NK-CKACHEI	GESAKNKVGP	ELDGLDGRHS	GAVEGYAYSP	ANKASGITWT	EAEFKEYIKD	PKAKVPGTKM	VFAGIKKDSE	LDNLWAYVSQ	FDKD
Agrobacterium	GDVAKGEAAF	KR-CSACHAI	GEGAKNKVGP	QLNGIIGRTA	GGDPDYNYSN	AMKKAGLVWT	PQELRDFLSA	PKKKIPGNKM	ALAGISKPEE	LDNLIAYLIF	SASSK
Rhodopila	GDPVEGKHLF	HTICLICHT-	DIKGRNKVGP	SLYGVVGRHS	GIEPGYNYSE	ANIKSGIVWT	PDVLFKYIEH	PQKIVPGTKM	GYPG-QPDQK	RADIIAYLET	LK


10
20
30
40
50
60
70
80
90
100
Human RADLIAYLKK ATNE
Chimpanzee RADLIAYLKK ATNE
Spider monkey RADLIAYLKK ATNE
Macaque RADLIAYLKK ATNE
Cow REDLIAYLKK ATNE
Dog RADLIAYLKK ATKE
Gray whale RADLIAYLKK ATNE
Horse REDLIAYLKK ATNE
Zebra REDLIAYLKK ATNE
Rabbit RADLIAYLKK ATNE
Kangaroo RADLIAYLKK ATNE
Duck RADLIAYLKD ATAK
Turkey RVDLIAYLKD ATSK
Chicken RVDLIAYLKD ATSK
Pigeon RADLIAYLKQ ATAK
King penguin RADLIAYLKD ATSK
Snapping turtl RADLIAYLKD ATSK
Alligator RADLIAYLKE ATSN
Bull frog RQDLIAYLKS ACSK
Tuna RQDLVAYLKS ATS
Dogfish RQDLIAYLKK TAAS
Starfish RQDLIAYLEA ATK
Fruit fly RGDLIAYLKS ATK
Silkmoth RADLIAYLKE STK
Pumpkin RADLIAYLKE ATA
Tomato RADLIAYLKE ATA
Arabidopsis RNDLITFLEE ETK
Mung bean RADLIAYLKE STA
Wheat RADLIAYLKK ATSS
Sunflower RADLIAYLKT STA
Yeast RNDLITYLKK ACE
Debaryomyces RNDLITYLVK ATK
Candida RNDLVTYLKK ATS
Aspergillus RNDLITYLKE STA
Rhodomicrobiu RADVIAYLKQ LSGK
Nitrobacter LDNLWAYVSQ FDKD
Agrobacterium GDVAKGEAAF KR-CSACHAI GEGAKNKVGP QLNGIIGRTA GGDPDYNYSN AMKKAGLVWT PQELRDFLSA PKKKIPGNKM ALAGISKPEE LDNLIAYLIF SASSK
Rhodopila GDPVEGKHLF HTICLICHT- DIKGRNKVGP SLYGVVGRHS GIEPGYNYSE ANIKSGIVWT PDVLFKYIEH PQKIVPGTKM GYPG-QPDQK RADIIAYLET LK
Human, chimpanzee
Zebra, horse
Macaque
Monkey
Debaryomyces kloeckeri
Candida krusei
Pig, cow, sheep
Rabbit
Gray kangaroo
Penguin
Chicken, turkey
Duck
Dog
Pigeon
Snapping turtle
Dogfish (shark)
Gray whale
Bullfrog
Alligator
Tuna
Bonito
Carp
Baker's yeast
Silk moth
Fruit fly
Screwworm fly
Hornworm moth
Pacific lamprey
Neurospora crassa
Bacteria
Mung-bean
Pumpkin
Wheat
Tomato
Sunflower

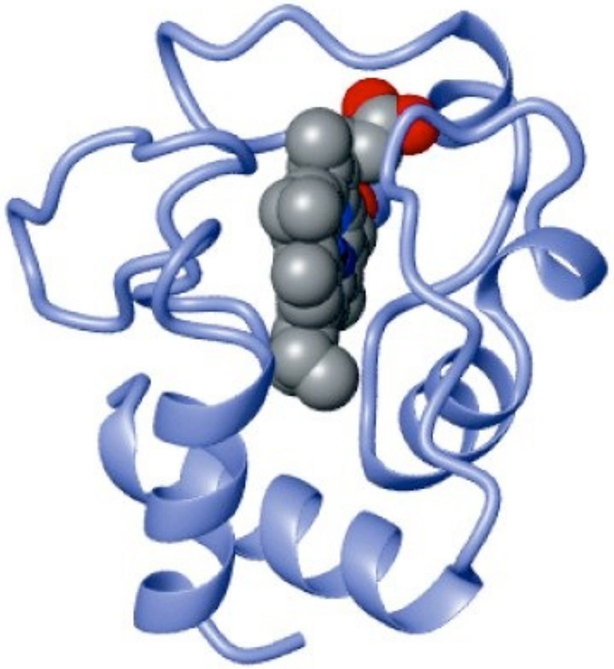
Protein Tertiary Structure

- Homologous proteins have similar tertiary structures, which are evolutionarily conserved.



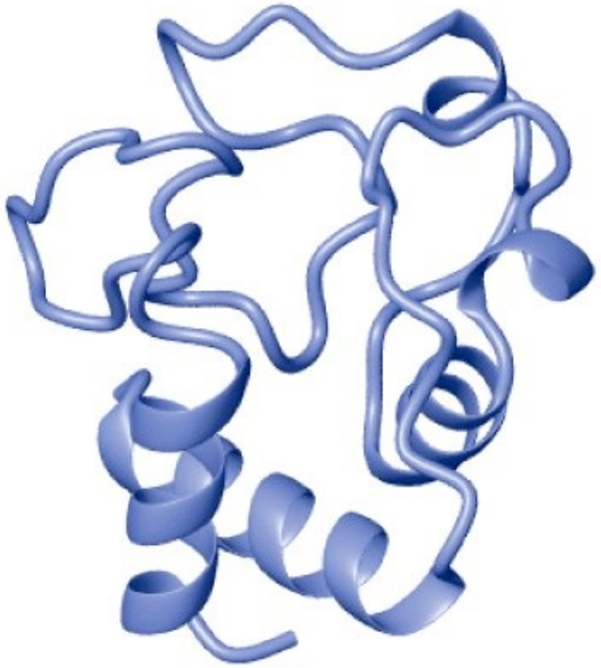
Protein Tertiary Structure

Tuna (/w heme)



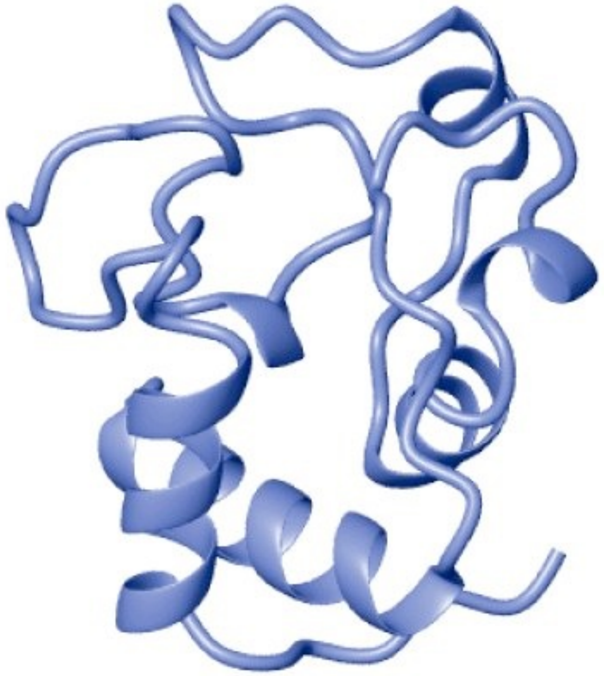
cytochrome c

Tuna (vertebrate)
(5CYT)



Rice (plant)
(1CCR)

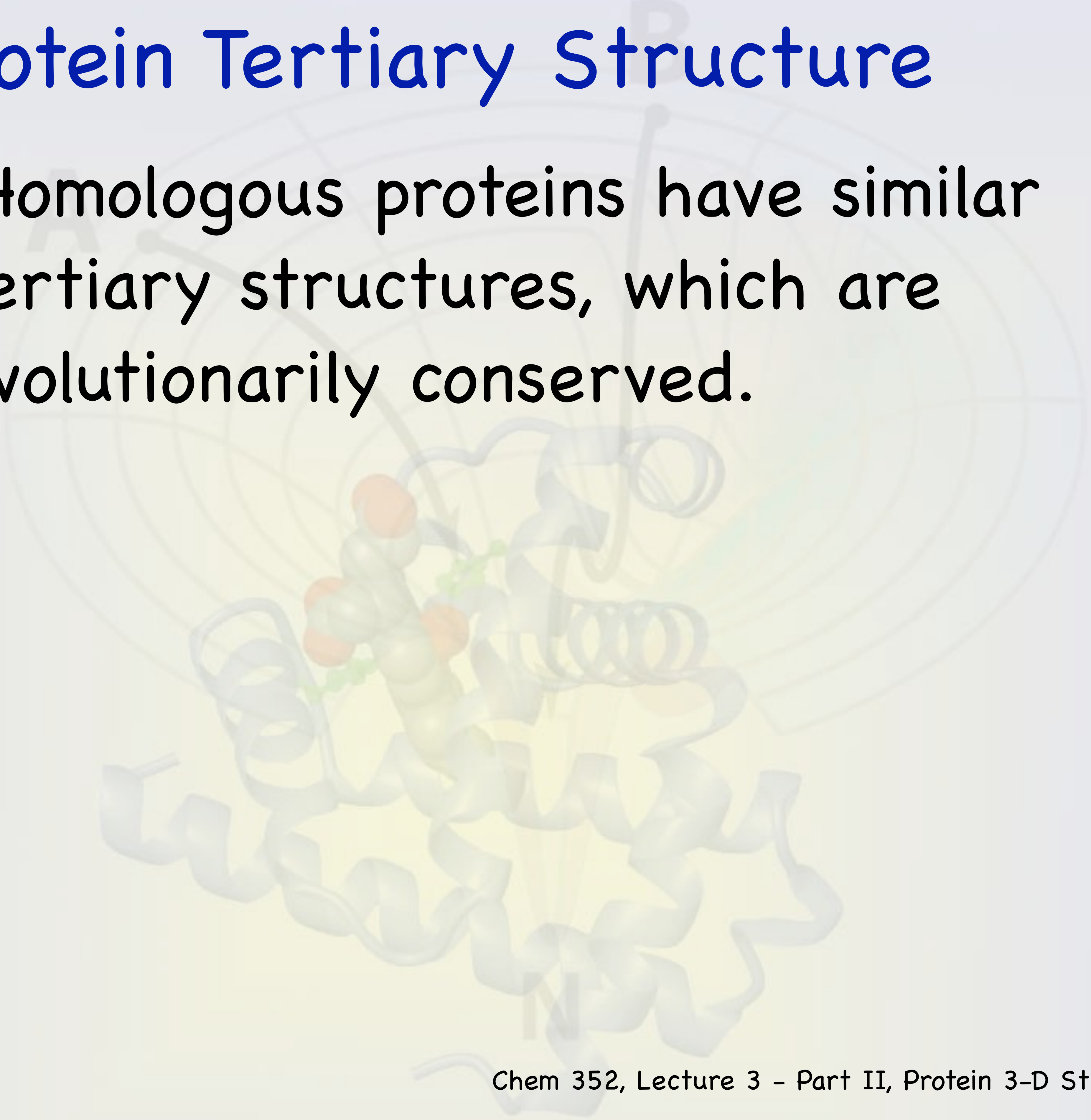
Yeast (fungus)
(1YCC)



Rhodopila (bacteria)
(1HRO)

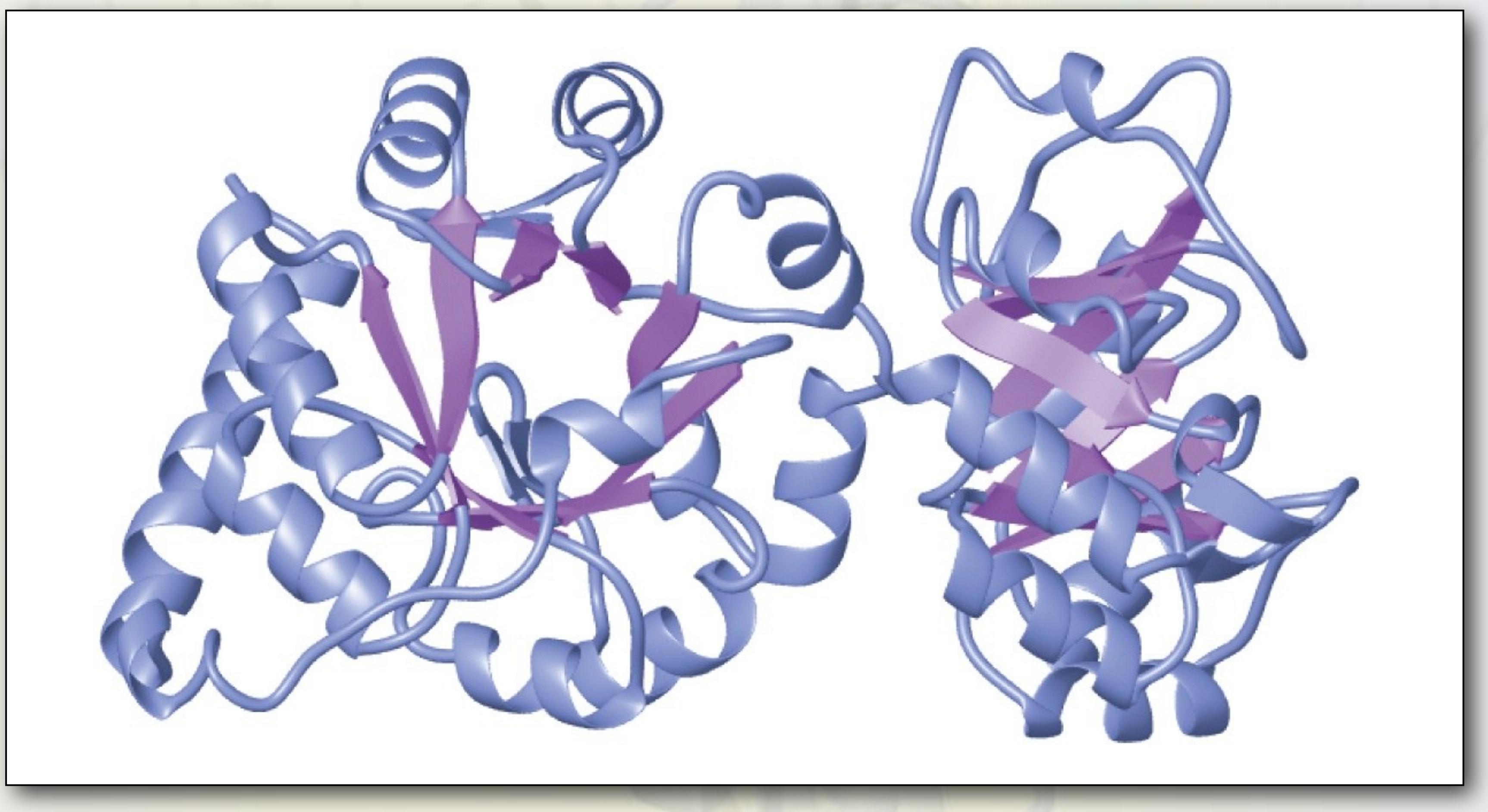
Protein Tertiary Structure

- Homologous proteins have similar tertiary structures, which are evolutionarily conserved.



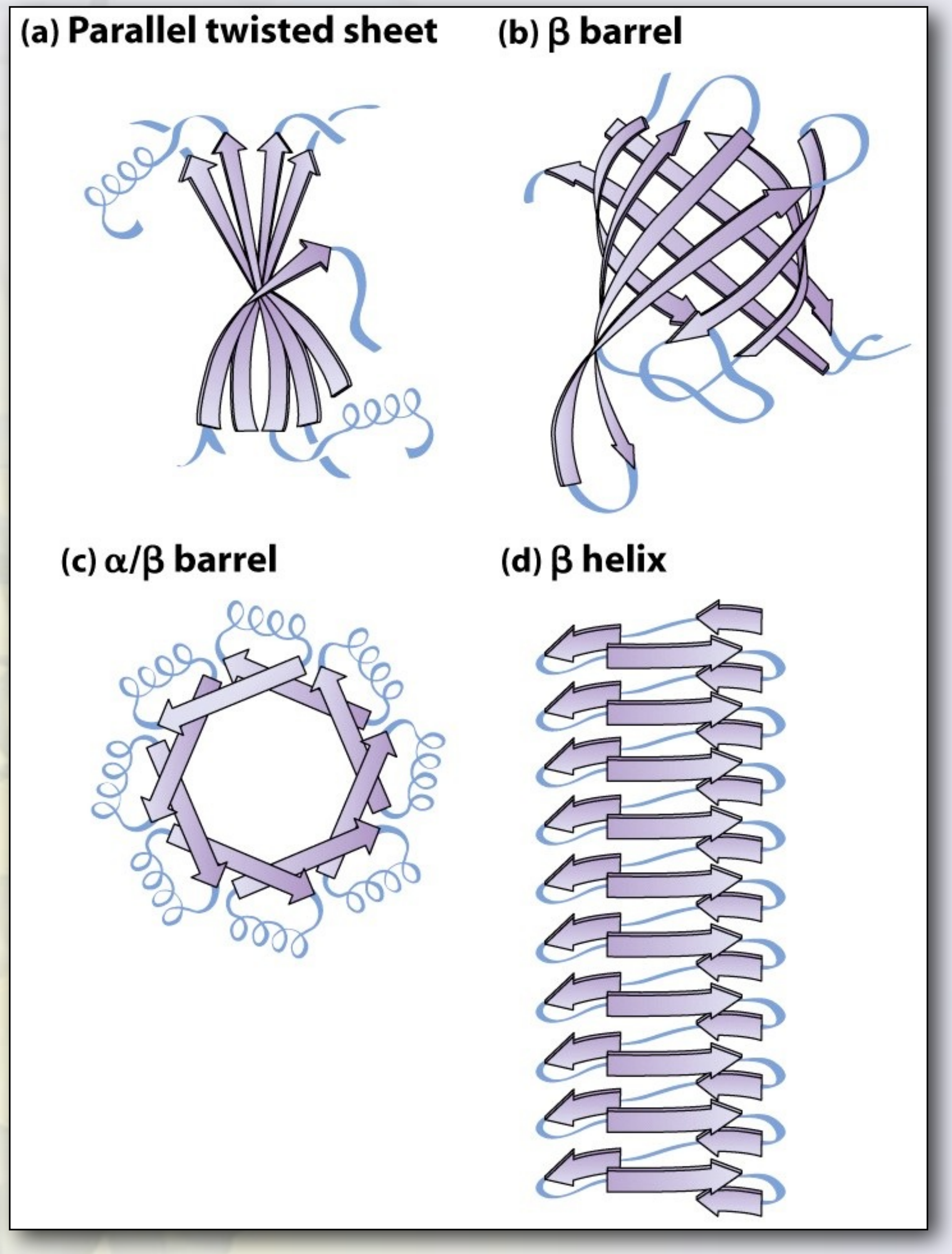
Protein Tertiary Structure

- Domains
  - Domains are independent folding units within a single polypeptide chain.



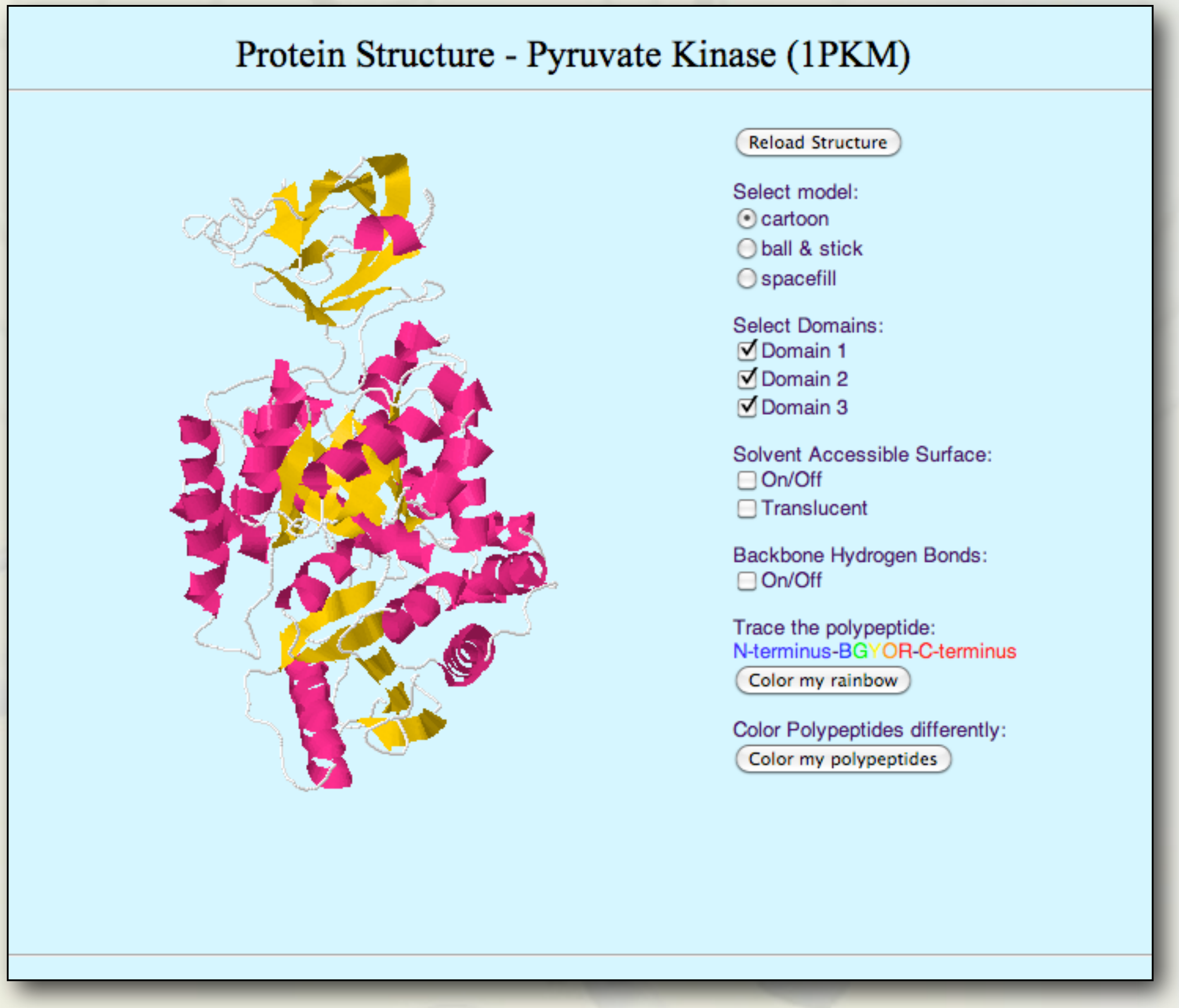
Protein Tertiary Structure

- Domains
  - Some common domain folds



Protein Tertiary Structure

- Domains
  - Domains often have specific functions and provide proteins with a modular design.



Protein Structure - Pyruvate Kinase (1PKM)

Reload Structure

Select model:
● cartoon
○ ball & stick
○ spacefill

Select Domains:
☑ Domain 1
☑ Domain 2
☑ Domain 3

Solvent Accessible Surface:
☐ On/Off
☐ Translucent

Backbone Hydrogen Bonds:
☐ On/Off

Trace the polypeptide:
N-terminus-BGYOR-C-terminus
Color my rainbow

Color Polypeptides differently:
Color my polypeptides

Go to Wikipedia entry on protein domains

Protein Tertiary Structure

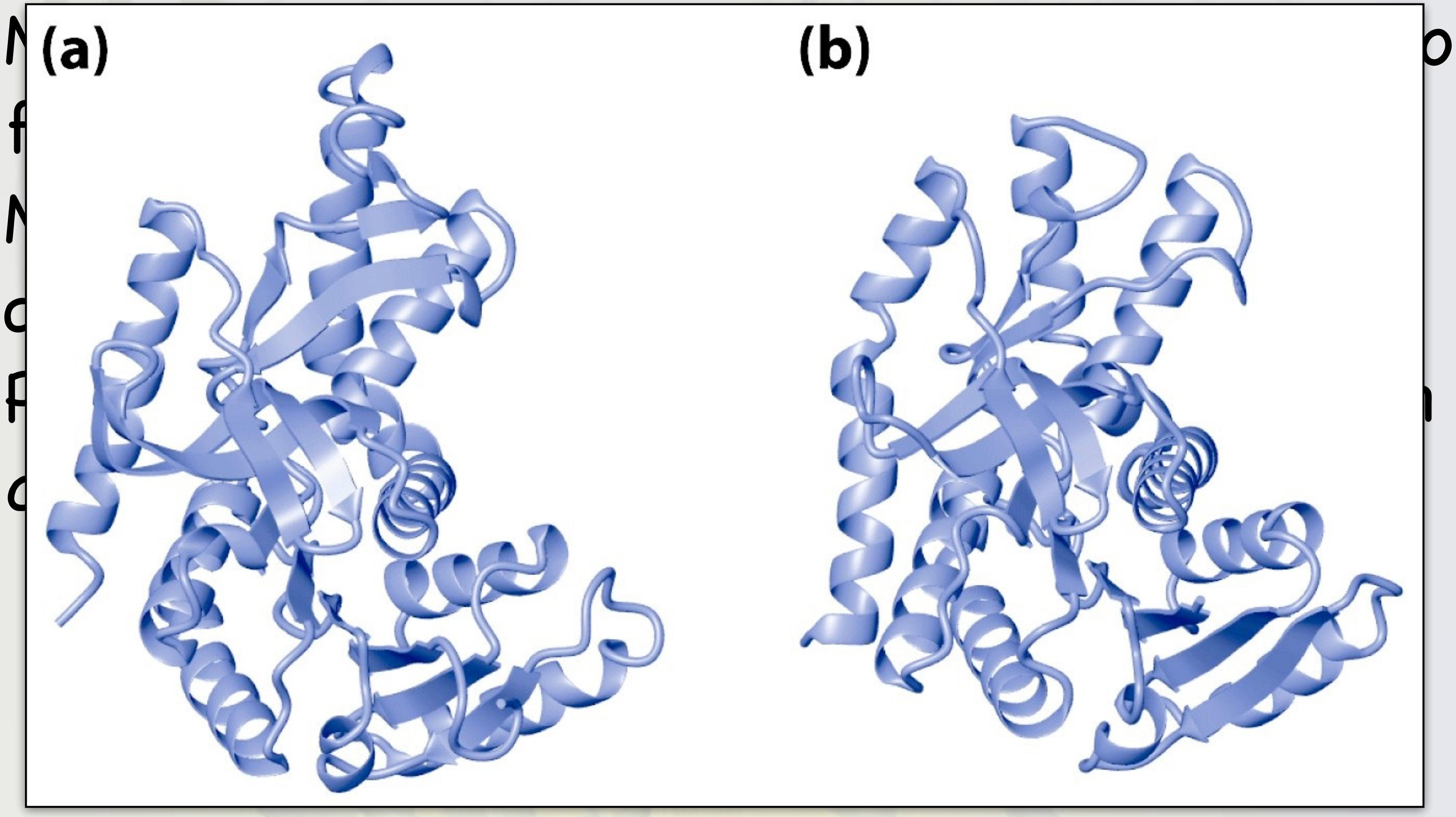
Families

- Many proteins in an organism share a common fold and can be grouped into families.
- Members of a family are believed to have descended from a common ancestor.
- Proteins that are descended from a common ancestor are said to be homologous.

Protein Tertiary Structure

Families

- N on
  f
- N
  c
- F n
  c



Lactate Dehydrogenase

Malate Dehydrogenase

Protein Tertiary Structure

Families

- Many proteins in an organism share a common fold and can be grouped into families.
- Members of a family are believed to have descended from a common ancestor.
- Proteins that are descended from a common ancestor are said to be homologous.

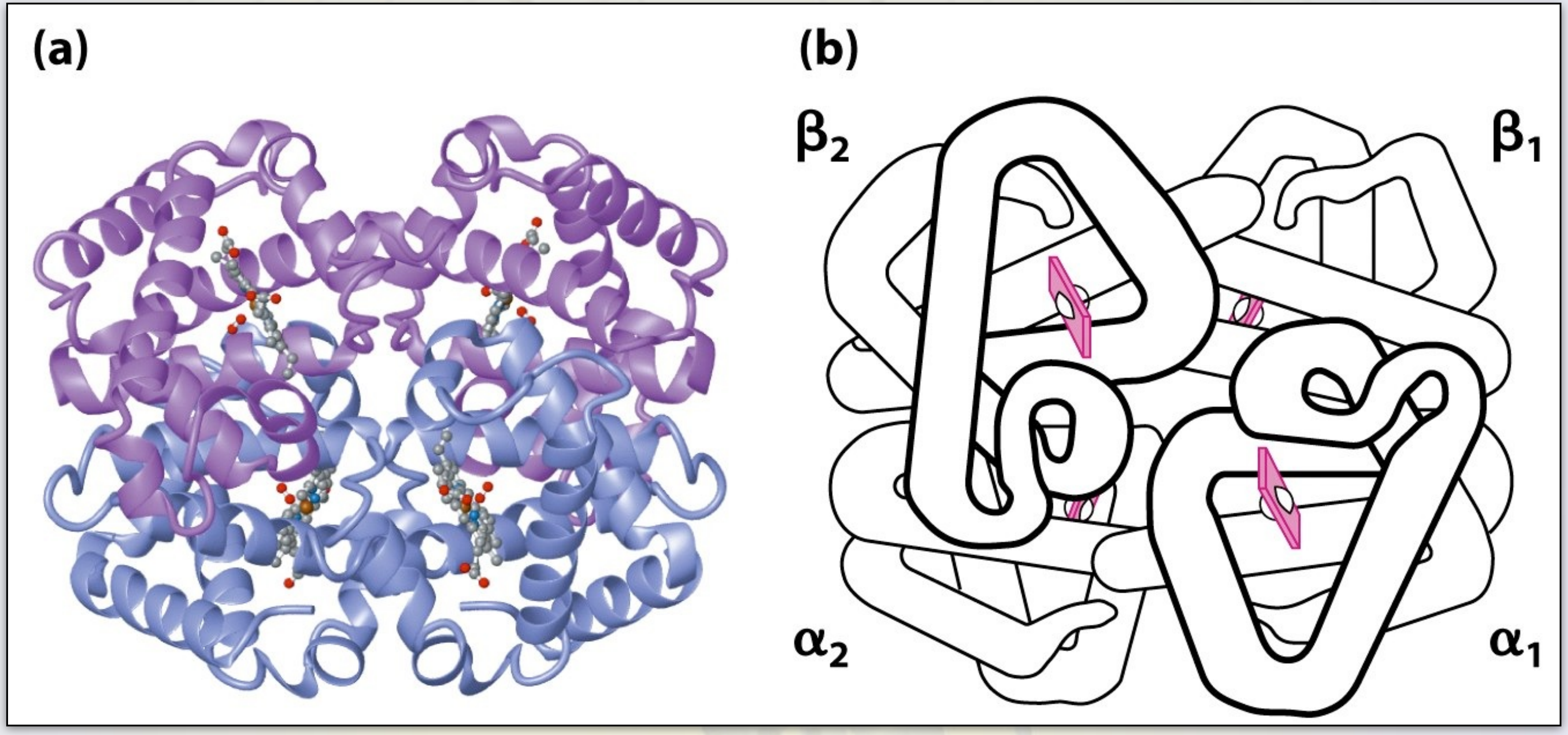
Protein Quaternary Structure

The assembly of multiple polypeptides to form a single protein

- Each polypeptide has its own tertiary structure.
- The individual polypeptides are called subunits.
- The subunits are held together by weak non-covalent interactions, primarily between the amino acid side chains.

Protein Quaternary Structure

The assembly of multiple polypeptides to form a single protein



Protein Quaternary Structure

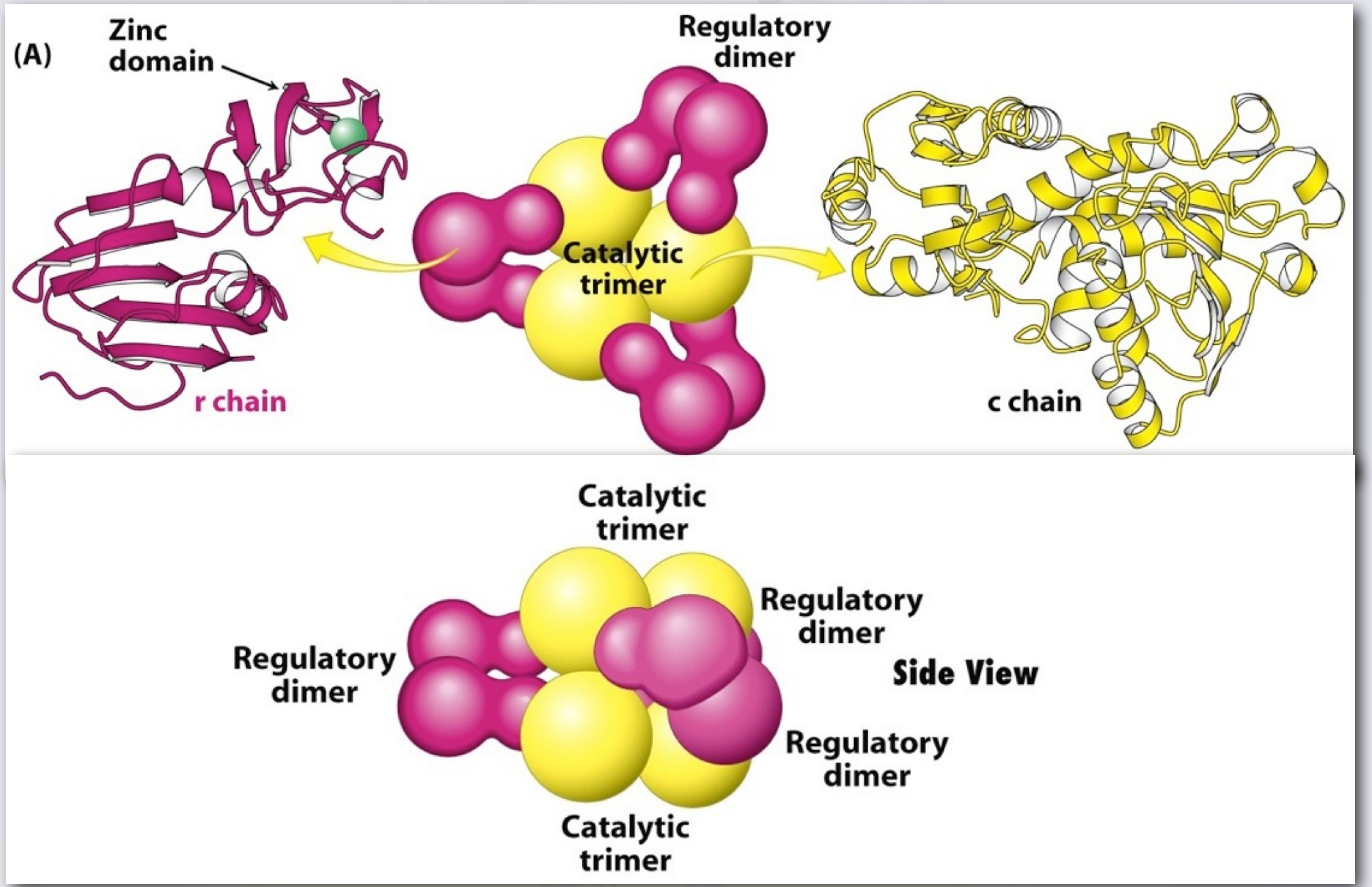
The assembly of multiple polypeptides to form a single protein

- Each polypeptide has its own tertiary structure.
- The individual polypeptides are called subunits.
- The subunits are held together by weak non-covalent interactions, primarily between the amino acid side chains.

Protein Quaternary Structure

- Reasons for forming quaternary structures.
  - Increased stability
  - Creation of active sites at subunit interfaces
  - Regulation of activity by way of subunit interactions
  - Different proteins can share the same subunits

Protein Quaternary Structure



The enzyme aspartate transcarbamoylase

Protein Quaternary Structure

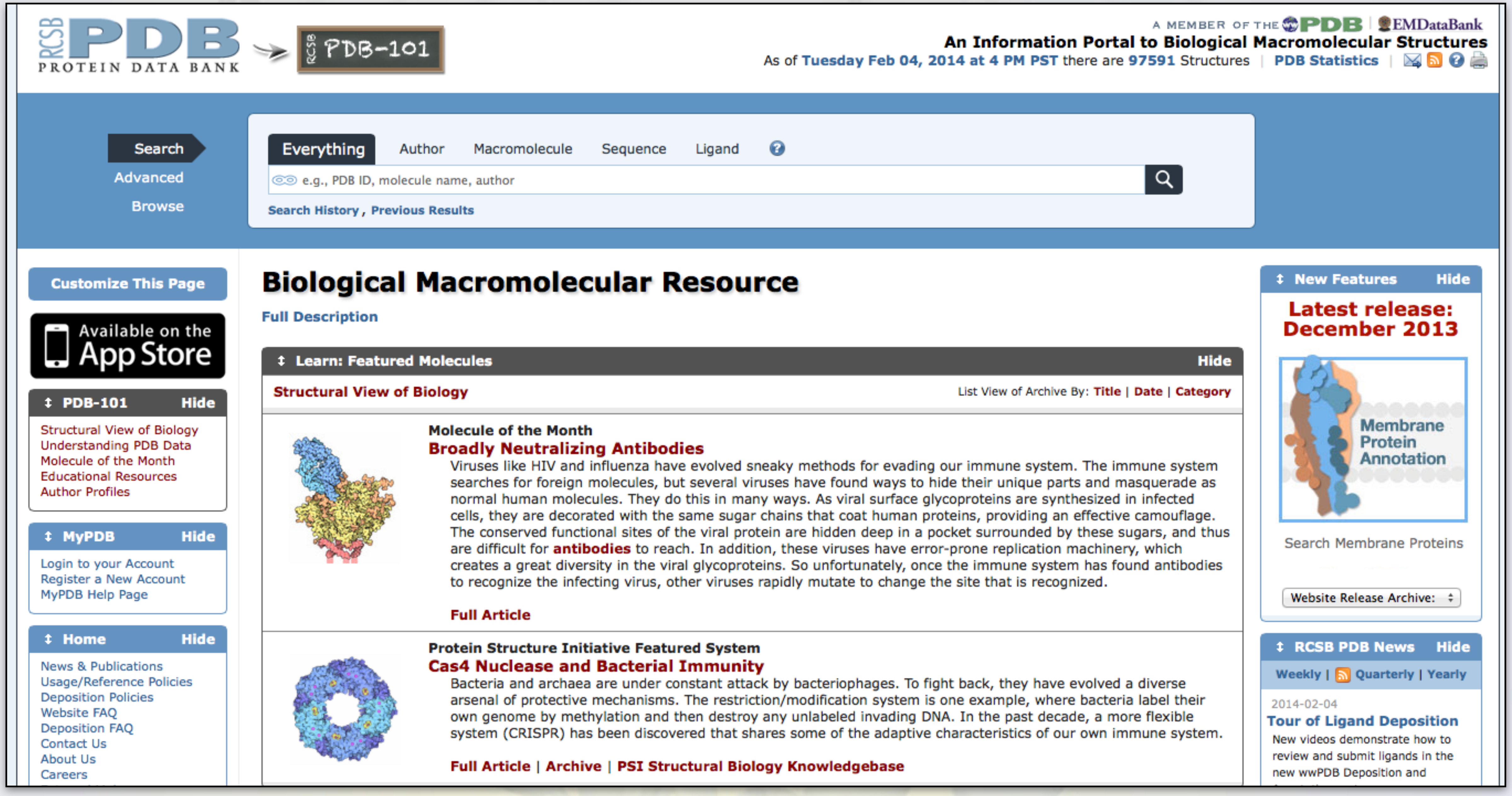
- Reasons for forming quaternary structures.
  - Increased stability
  - Creation of active sites at subunit interfaces
  - Regulation of activity by way of subunit interactions
  - Different proteins can share the same subunits

The Protein Data Bank

- The Protein Data Bank is a repository for the 3-dimensional structures of proteins.
  - In addition to proteins, it also contains structures for nucleic acids and large assemblies, such as viruses and ribosomes.
  - Each entry has a four character code
    - Moran et al. gives you the codes in the figure legends when is shows structures.

The Protein Data Bank

www.pdb.org



An Information Portal to Biological Macromolecular Structures

As of **Tuesday Feb 04, 2014 at 4 PM PST** there are **97591** Structures | **PDB Statistics** |

Structural View of Biology
Understanding PDB Data
Molecule of the Month
Educational Resources
Author Profiles

Login to your Account
Register a New Account
MyPDB Help Page

News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us
About Us
Careers

Biological Macromolecular Resource

Full Description

Structural View of Biology

List View of Archive By: **Title** | **Date** | **Category**

Molecule of the Month

Broadly Neutralizing Antibodies

Viruses like HIV and influenza have evolved sneaky methods for evading our immune system. The immune system searches for foreign molecules, but several viruses have found ways to hide their unique parts and masquerade as normal human molecules. They do this in many ways. As viral surface glycoproteins are synthesized in infected cells, they are decorated with the same sugar chains that coat human proteins, providing an effective camouflage. The conserved functional sites of the viral protein are hidden deep in a pocket surrounded by these sugars, and thus are difficult for **antibodies** to reach. In addition, these viruses have error-prone replication machinery, which creates a great diversity in the viral glycoproteins. So unfortunately, once the immune system has found antibodies to recognize the infecting virus, other viruses rapidly mutate to change the site that is recognized.

Full Article

Protein Structure Initiative Featured System

Cas4 Nuclease and Bacterial Immunity

Bacteria and archaea are under constant attack by bacteriophages. To fight back, they have evolved a diverse arsenal of protective mechanisms. The restriction/modification system is one example, where bacteria label their own genome by methylation and then destroy any unlabeled invading DNA. In the past decade, a more flexible system (CRISPR) has been discovered that shares some of the adaptive characteristics of our own immune system.

Full Article | **Archive** | **PSI Structural Biology Knowledgebase**

Latest release: December 2013

Search Membrane Proteins

2014-02-04

Tour of Ligand Deposition

New videos demonstrate how to review and submit ligands in the new wwPDB Deposition and

The Protein Data Bank

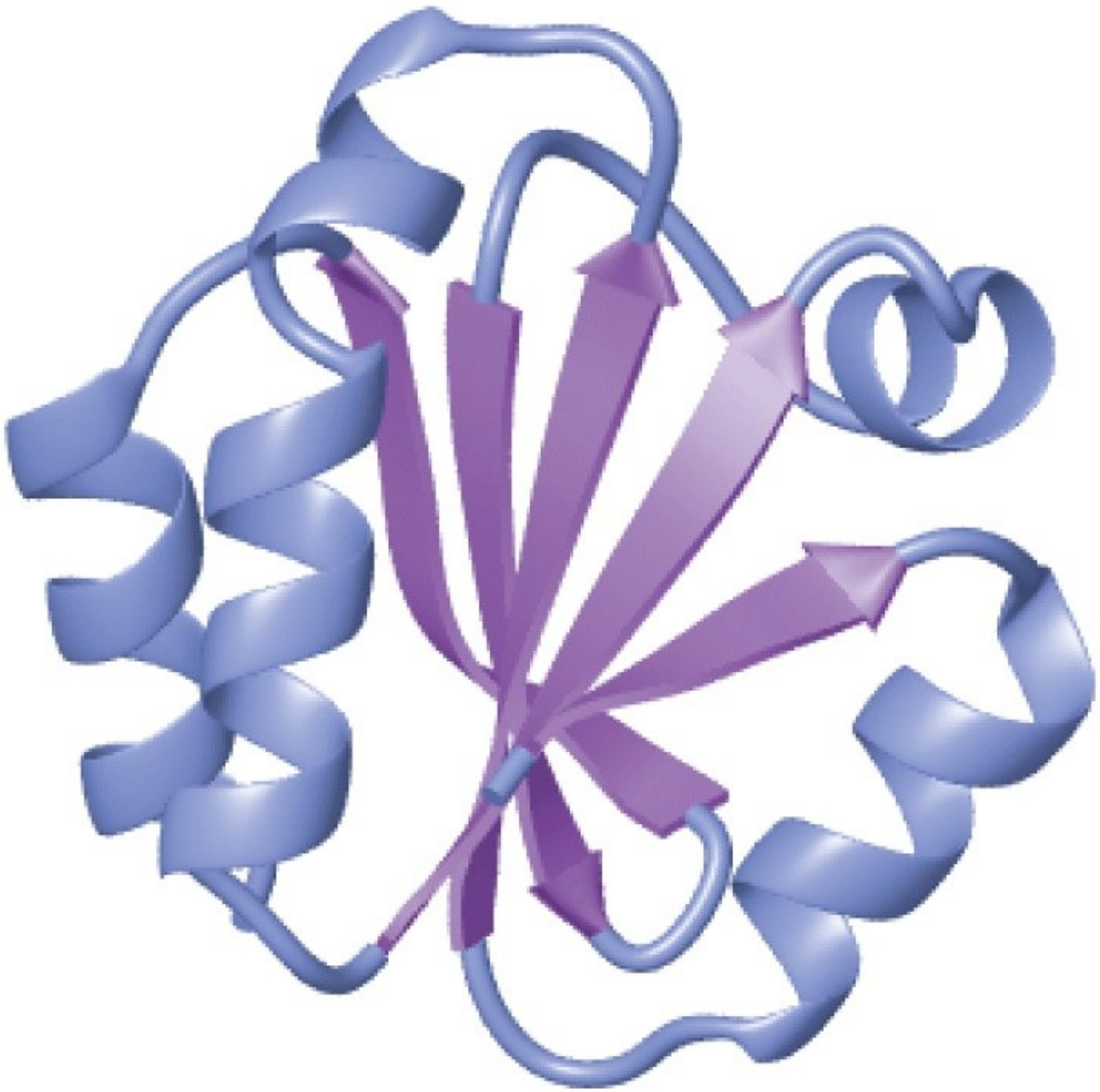
The Protein Data Bank is a repository for the 3-dimensional structures of proteins.

- In addition to proteins, it also contains structures for nucleic acids and large assemblies, such as viruses and ribosomes.
- Each entry has a four character code
  - Moran et al. gives you the codes in the figure legends when is shows structures.

The Protein Data Bank

The Pr … sitory
for the … s of
protein

- In ad … s
  struc
  asse … omes.
- Each … .
  - Mo … the
    figu … ures.

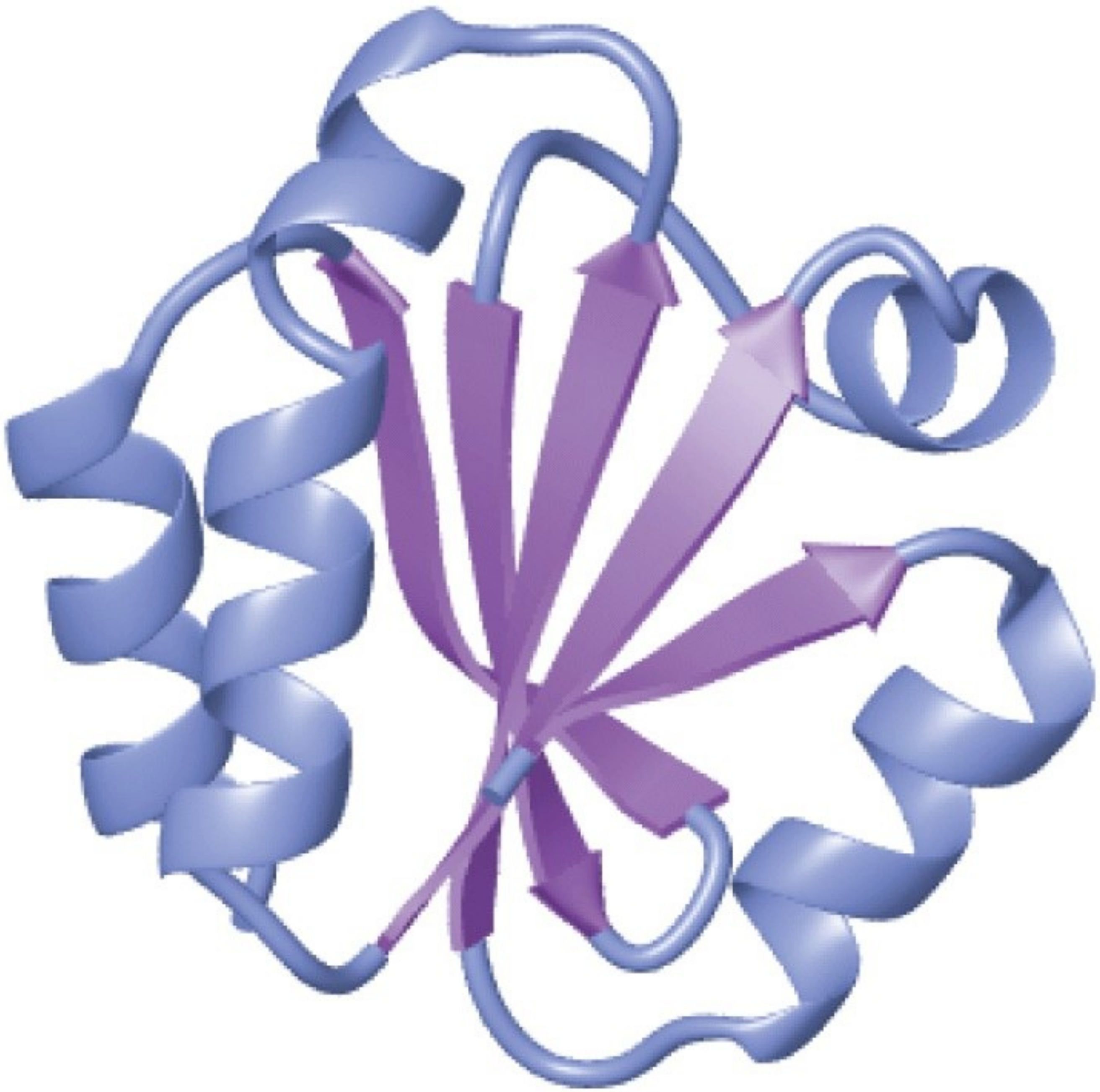


Human thioredoxin (1ERU)

The Protein Data Bank

The Pr sitory
for the s of
protein

- In ad s
  struc
  assem omes.
- Each .
  - Mor the
    figu ures.



Human thioredoxin (1ERU)

http://www.chem.uwec.edu/marvin

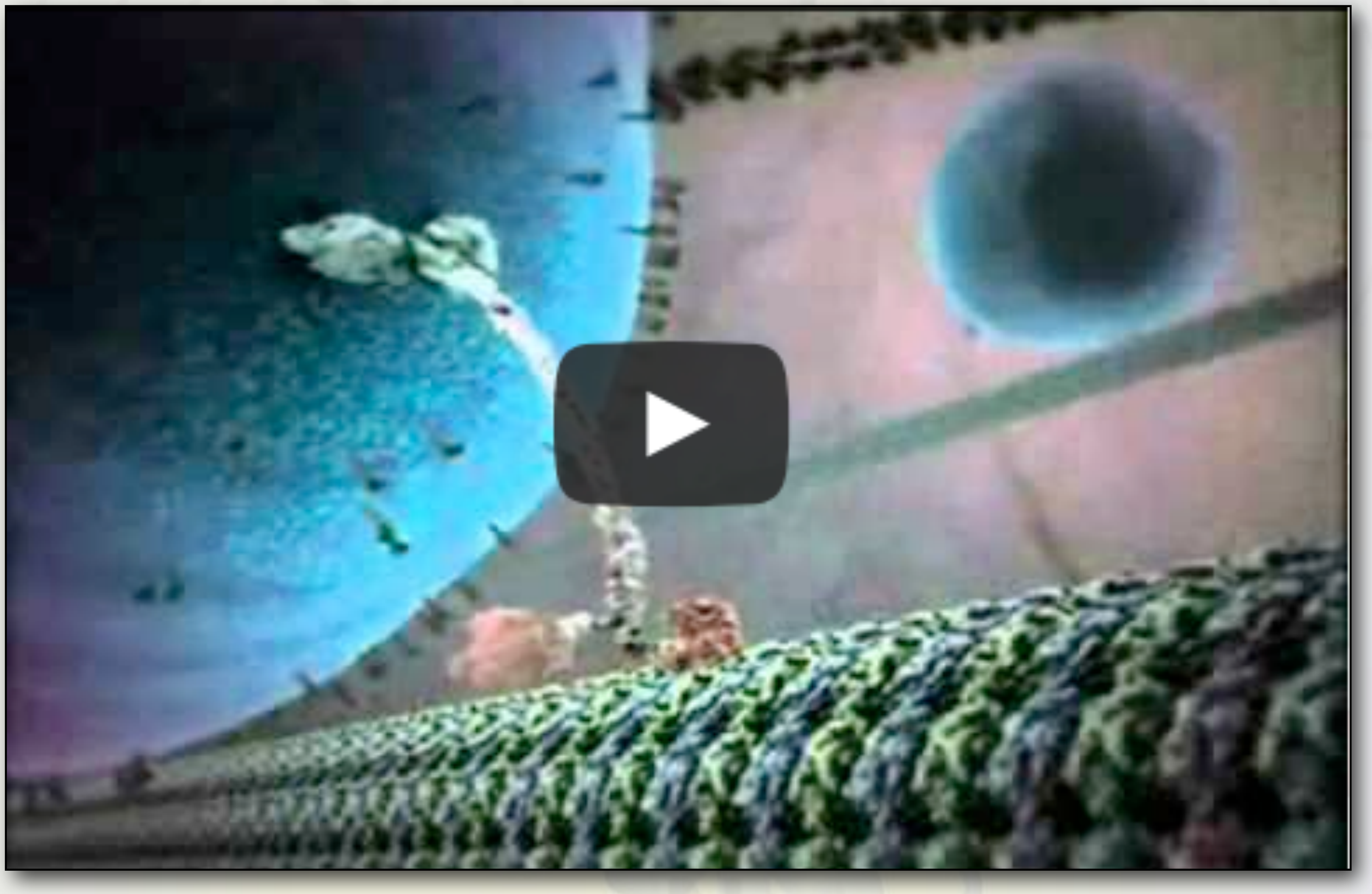
The Protein Data Bank

The Protein Data Bank is a repository for the 3-dimensional structures of proteins.

- In addition to proteins, it also contains structures for nucleic acids and large assemblies, such as viruses and ribosomes.
- Each entry has a four character code
    - Moran et al. gives you the codes in the figure legends when is shows structures.

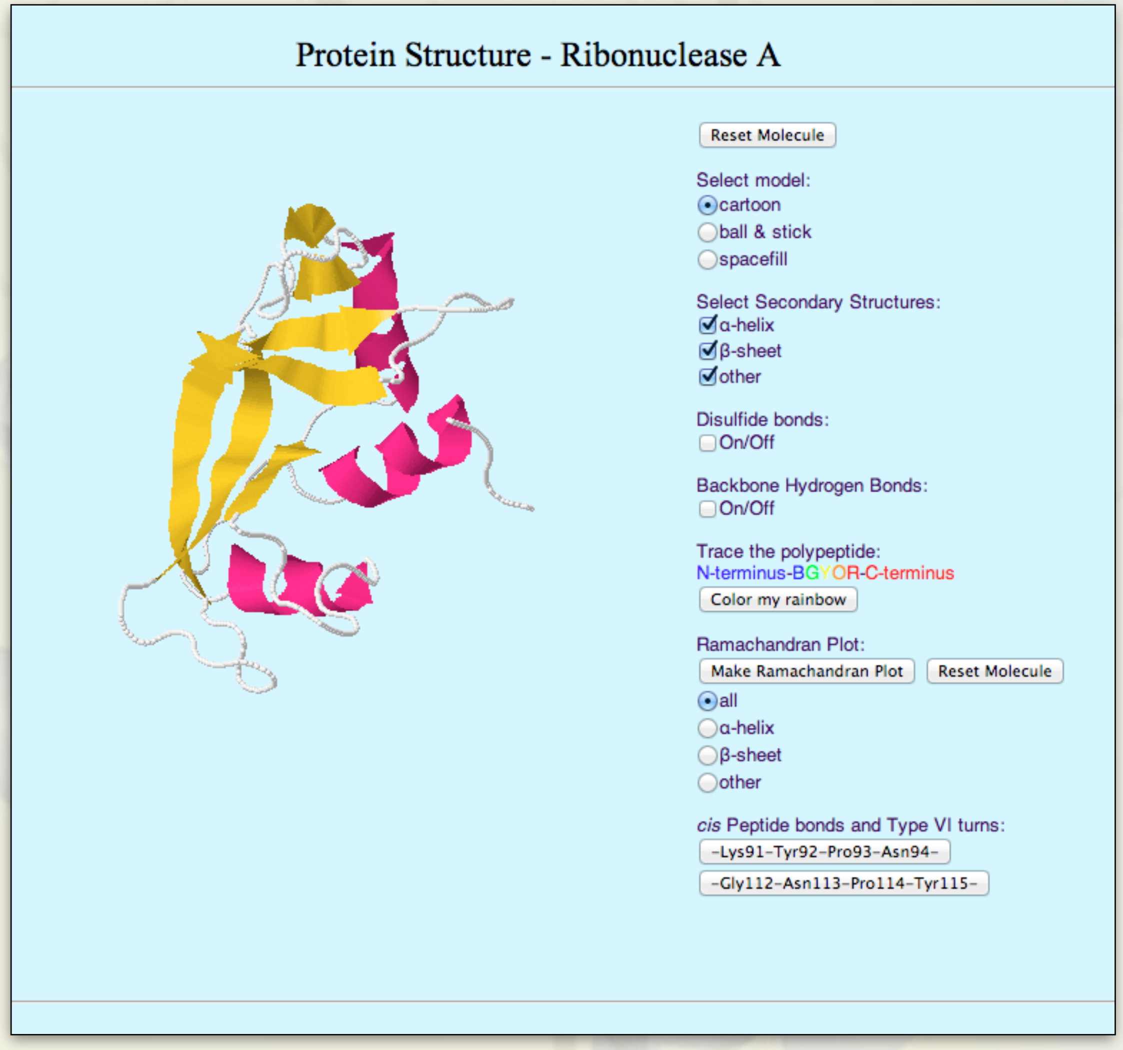
Large Protein Assemblies

- The Molecular Machinery of Life



Recap: Conformational Restrictions

Steric hinderances restrict the φ and ψ to a limited number of combinations.



Protein Structure - Ribonuclease A

Reset Molecule

Select model:
● cartoon
○ ball & stick
○ spacefill

Select Secondary Structures:
☑ α-helix
☑ β-sheet
☑ other

Disulfide bonds:
☐ On/Off

Backbone Hydrogen Bonds:
☐ On/Off

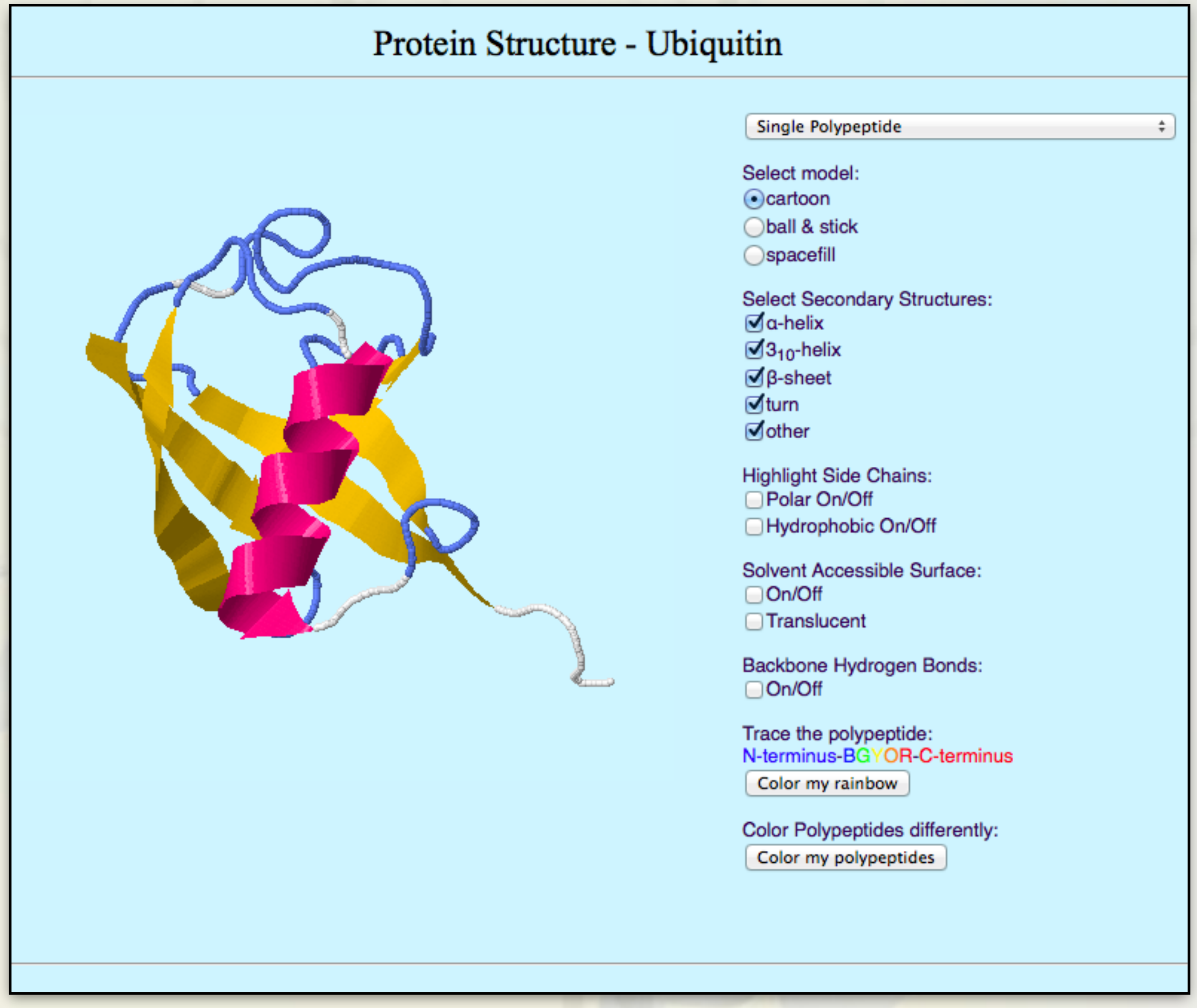
Trace the polypeptide:
N-terminus-BGYOR-C-terminus
Color my rainbow

Ramachandran Plot:
Make Ramachandran Plot
Reset Molecule
● all
○ α-helix
○ β-sheet
○ other

cis Peptide bonds and Type VI turns:
-Lys91-Tyr92-Pro93-Asn94-
-Gly112-Asn113-Pro114-Tyr115-

Recap: Interactions with water

Globular proteins fold to remove as many non-polar side chains from water as possible.



Protein Structure - Ubiquitin

Single Polypeptide

Select model:
● cartoon
○ ball & stick
○ spacefill

Select Secondary Structures:
☑ α-helix
☑ 3_{10}-helix
☑ β-sheet
☑ turn
☑ other

Highlight Side Chains:
☐ Polar On/Off
☐ Hydrophobic On/Off

Solvent Accessible Surface:
☐ On/Off
☐ Translucent

Backbone Hydrogen Bonds:
☐ On/Off

Trace the polypeptide:
N-terminus-BGYOR-C-terminus

Color Polypeptides differently:

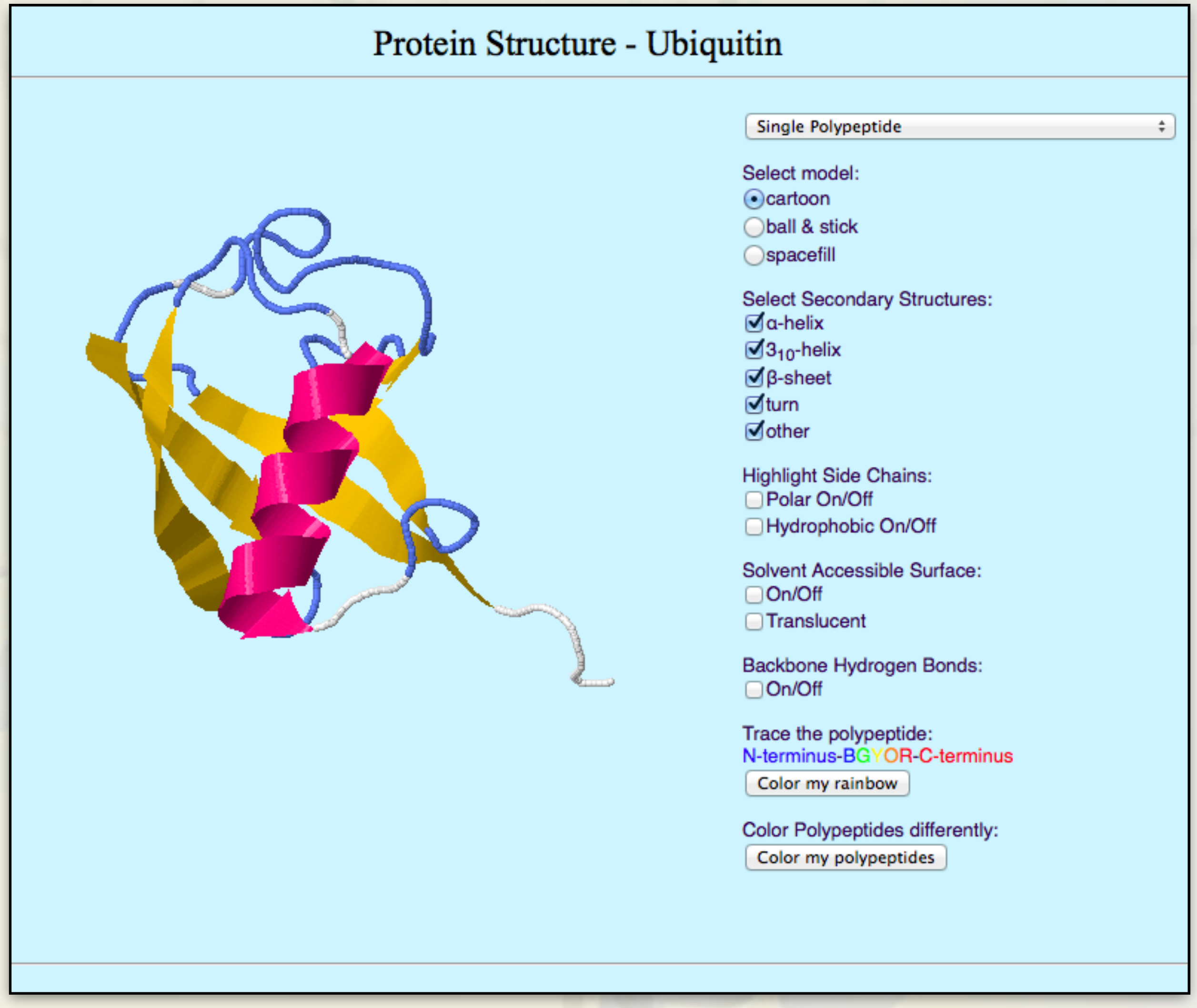
Recap: Interactions with water

Question:

What is the driving force behind this process?

Recap: Interactions with water

Globular proteins fold to remove as many non-polar side chains from water as possible.



Protein Structure - Ubiquitin

Select model:
● cartoon
○ ball & stick
○ spacefill

Select Secondary Structures:
☑ α-helix
☑ 3_{10}-helix
☑ β-sheet
☑ turn
☑ other

Highlight Side Chains:
☐ Polar On/Off
☐ Hydrophobic On/Off

Solvent Accessible Surface:
☐ On/Off
☐ Translucent

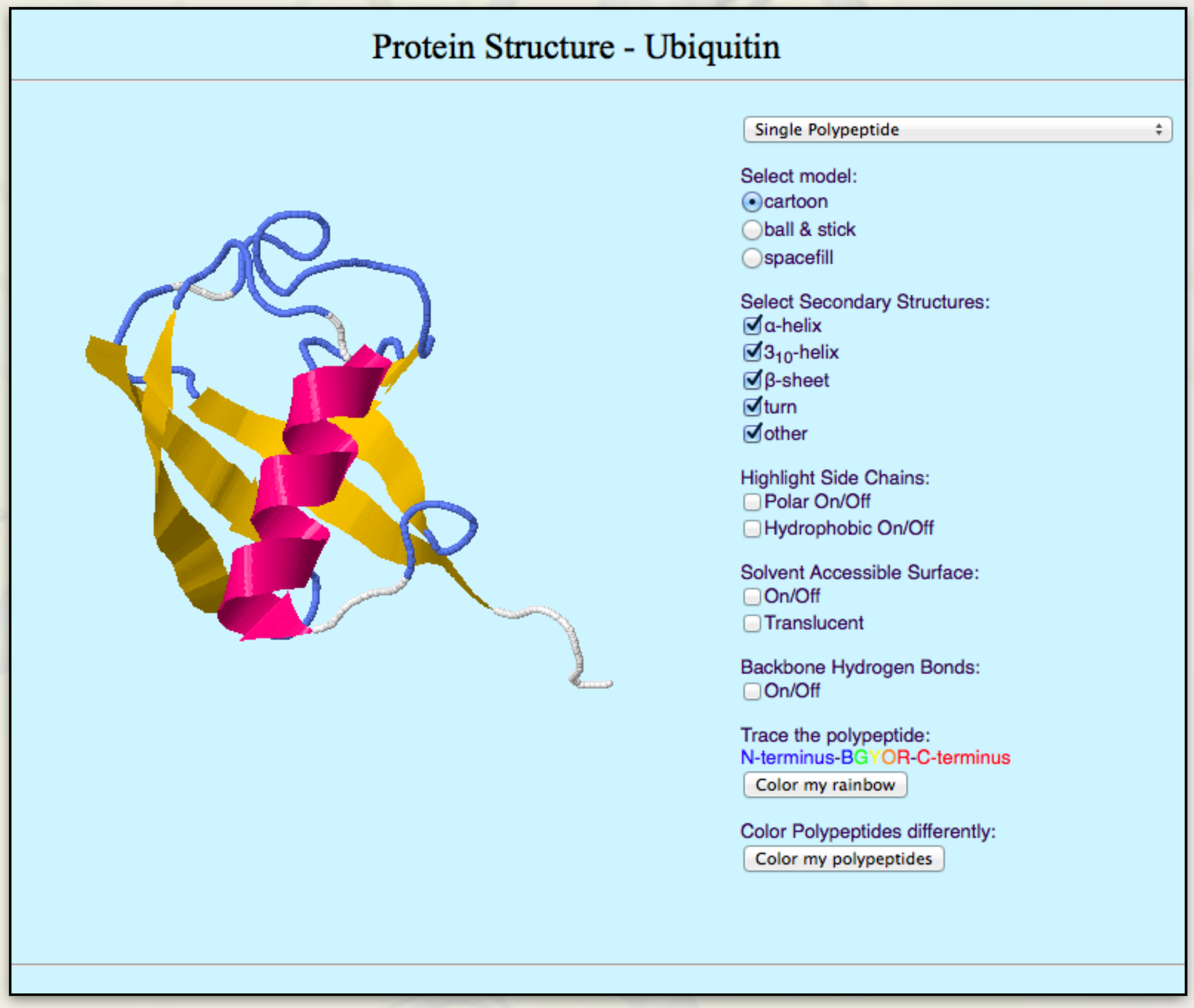
Backbone Hydrogen Bonds:
☐ On/Off

Trace the polypeptide:
N-terminus-BGYOR-C-terminus

Color Polypeptides differently:

Recap: Role of secondary structures

α-helices and β-sheets provide hydrogen bond opportunities for backbone amide groups in the absence of water



Protein Structure - Ubiquitin

Single Polypeptide

Select model:
● cartoon
○ ball & stick
○ spacefill

Select Secondary Structures:
☑ α-helix
☑ 3_{10}-helix
☑ β-sheet
☑ turn
☑ other

Highlight Side Chains:
☐ Polar On/Off
☐ Hydrophobic On/Off

Solvent Accessible Surface:
☐ On/Off
☐ Translucent

Backbone Hydrogen Bonds:
☐ On/Off

Trace the polypeptide:
N-terminus-BGYOR-C-terminus

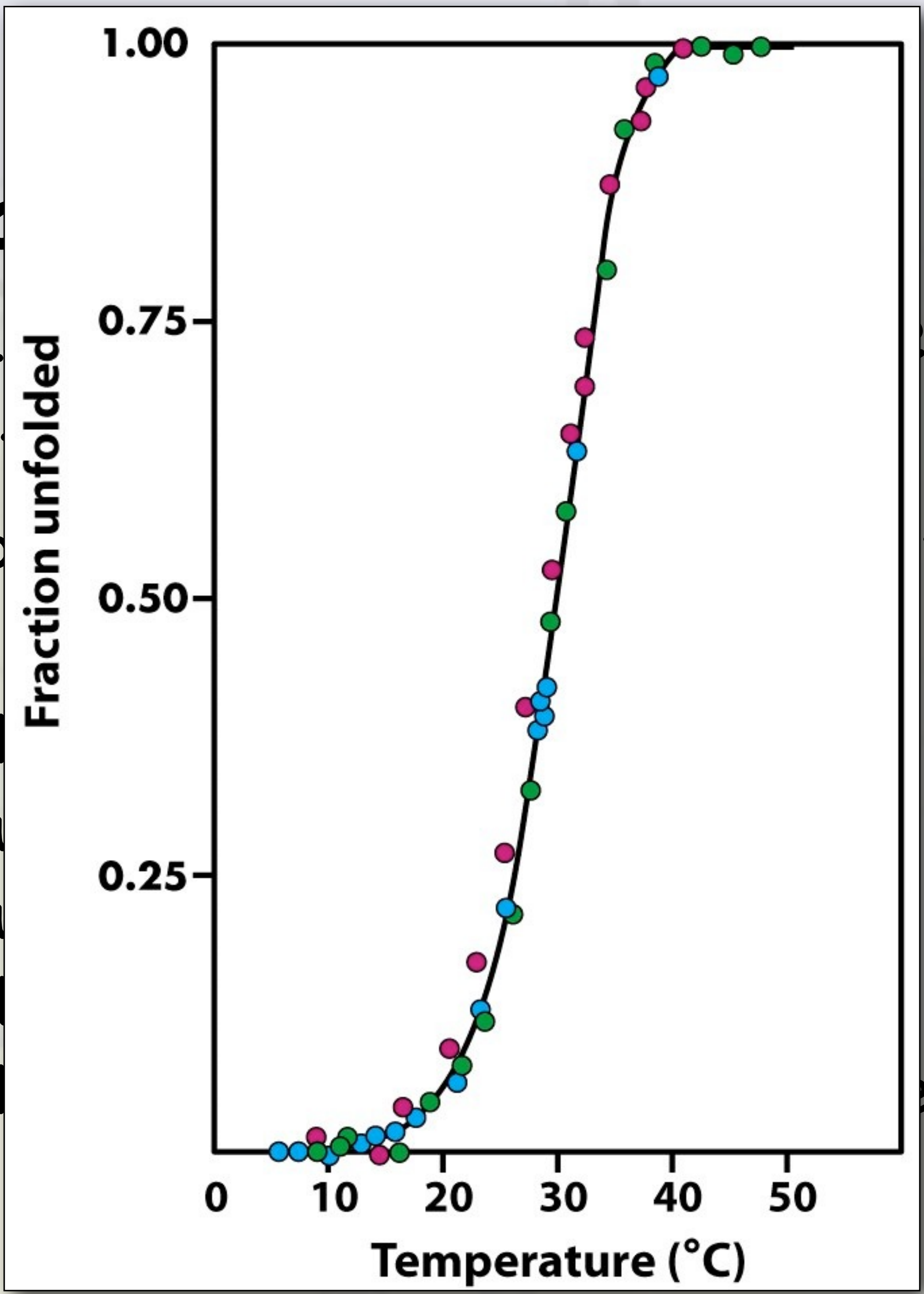
Color Polypeptides differently:

Protein Denaturation

- The **native** state of a protein is stabilized by weak non-covalent interactions.
  - The process of disrupting these interactions is called **denaturation.**
  - Denaturation can be accomplished with
    - Changes in temperature.
    - Changes in pH.
    - Added chemical agents.
  - Denaturation is a cooperative, often two-state, event.

Protein

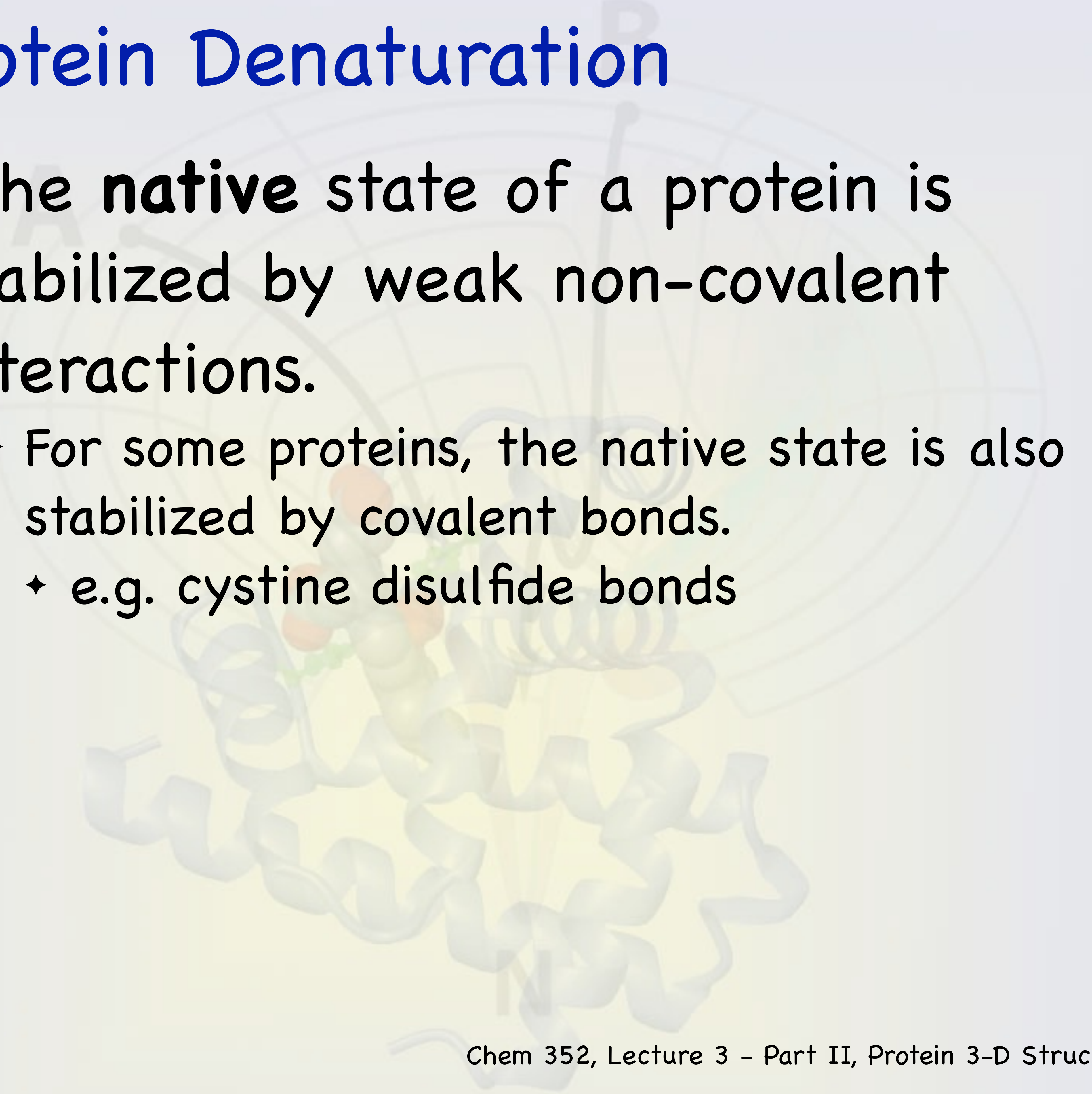
- The **n** is
  stabiliz nt
  interac
  - The p teractions
    is cal
  - Denat with
    - Cha
    - Cha
    - Add
  - Denat n two-
    state,



Protein Denaturation

- The **native** state of a protein is stabilized by weak non-covalent interactions.
  - The process of disrupting these interactions is called **denaturation.**
  - Denaturation can be accomplished with
    - Changes in temperature.
    - Changes in pH.
    - Added chemical agents.
  - Denaturation is a cooperative, often two-state, event.

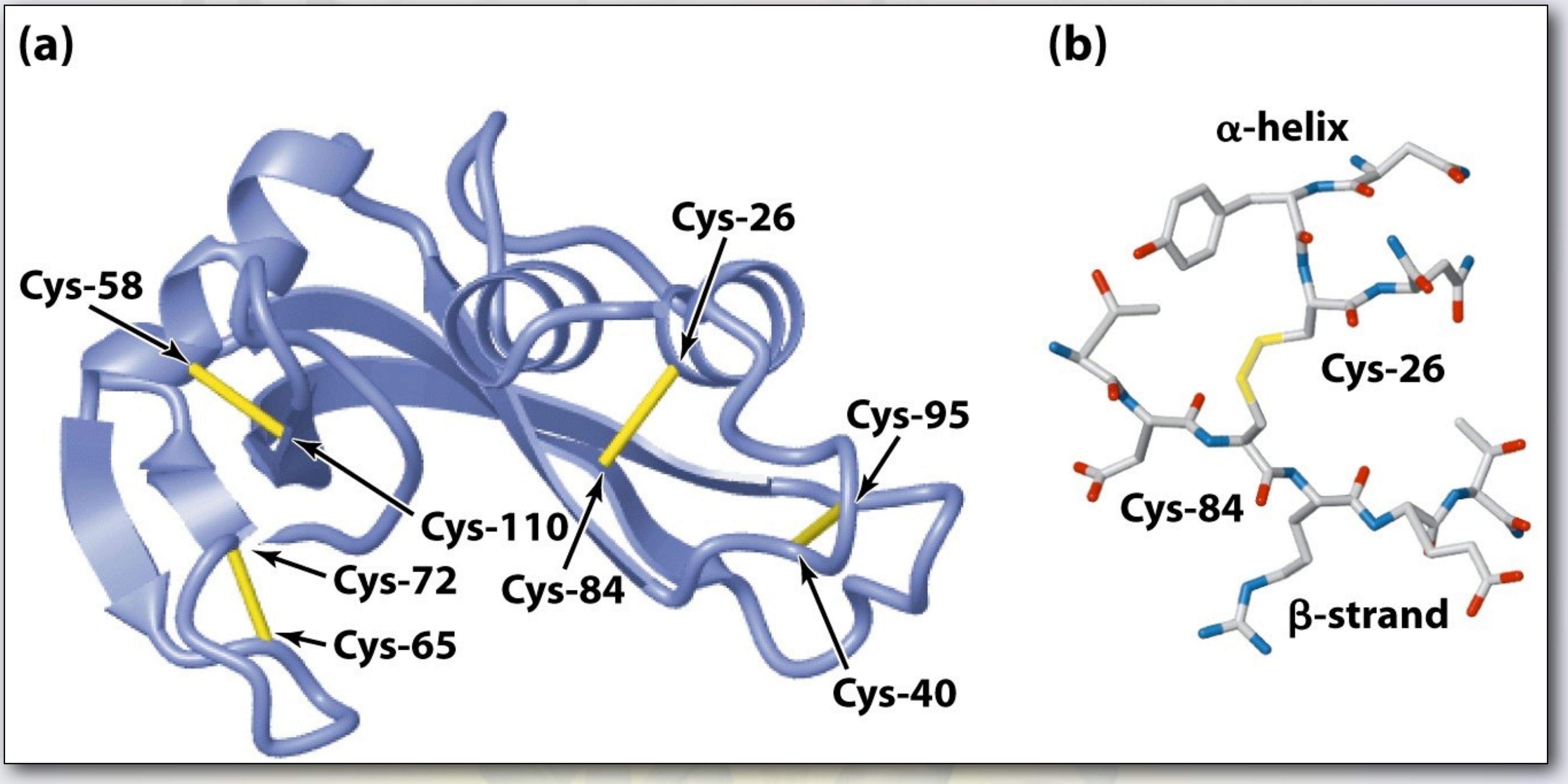
Protein Denaturation



- The **native** state of a protein is stabilized by weak non-covalent interactions.
  - ✦ For some proteins, the native state is also stabilized by covalent bonds.
    - ✦ e.g. cystine disulfide bonds

Protein Denaturation

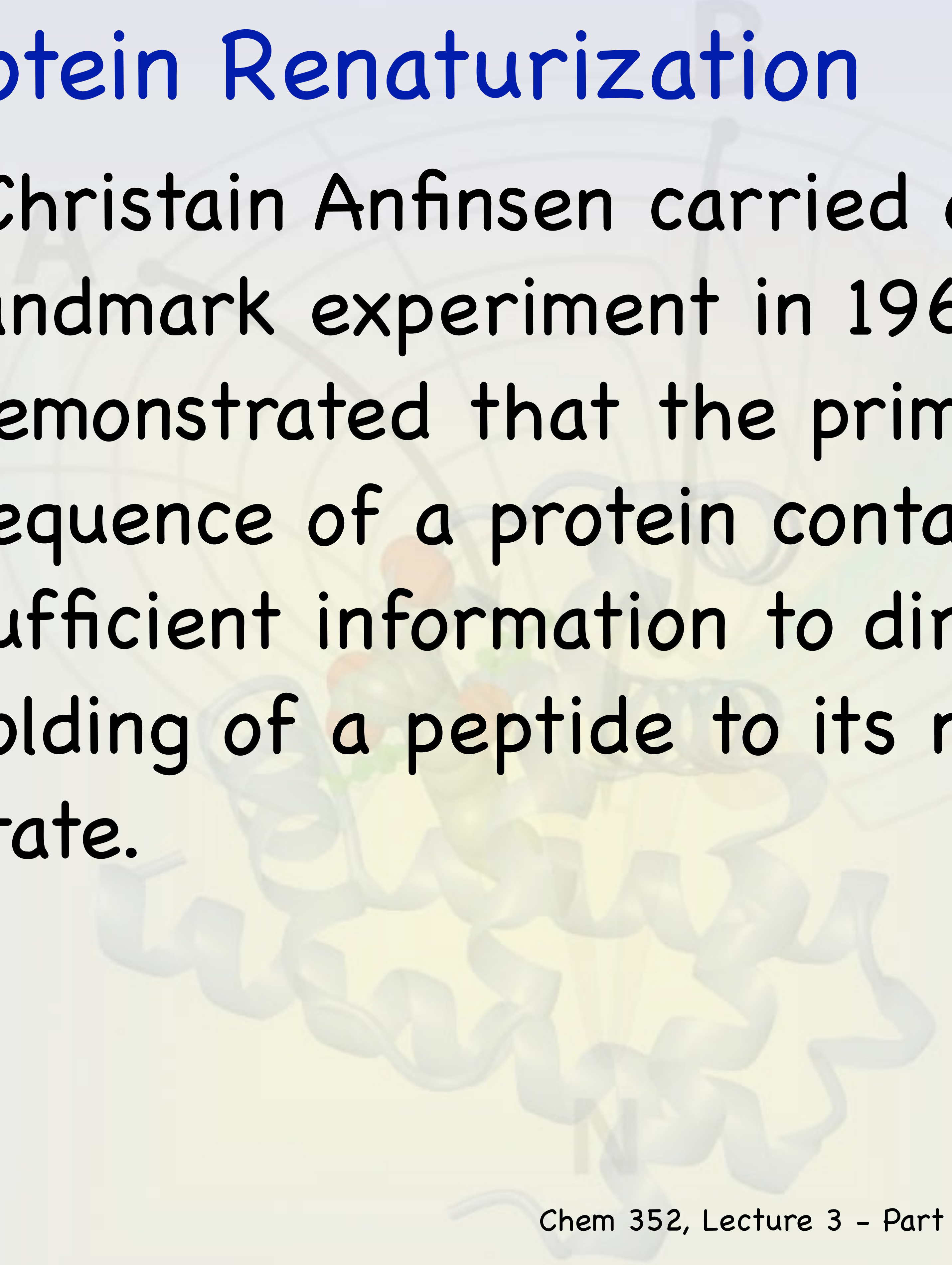
- The **native** state of a protein is



Protein Denaturation

- The **native** state of a protein is stabilized by weak non-covalent interactions.
  - For some proteins, the native state is also stabilized by covalent bonds.
    - e.g. cystine disulfide bonds

Protein Renaturization



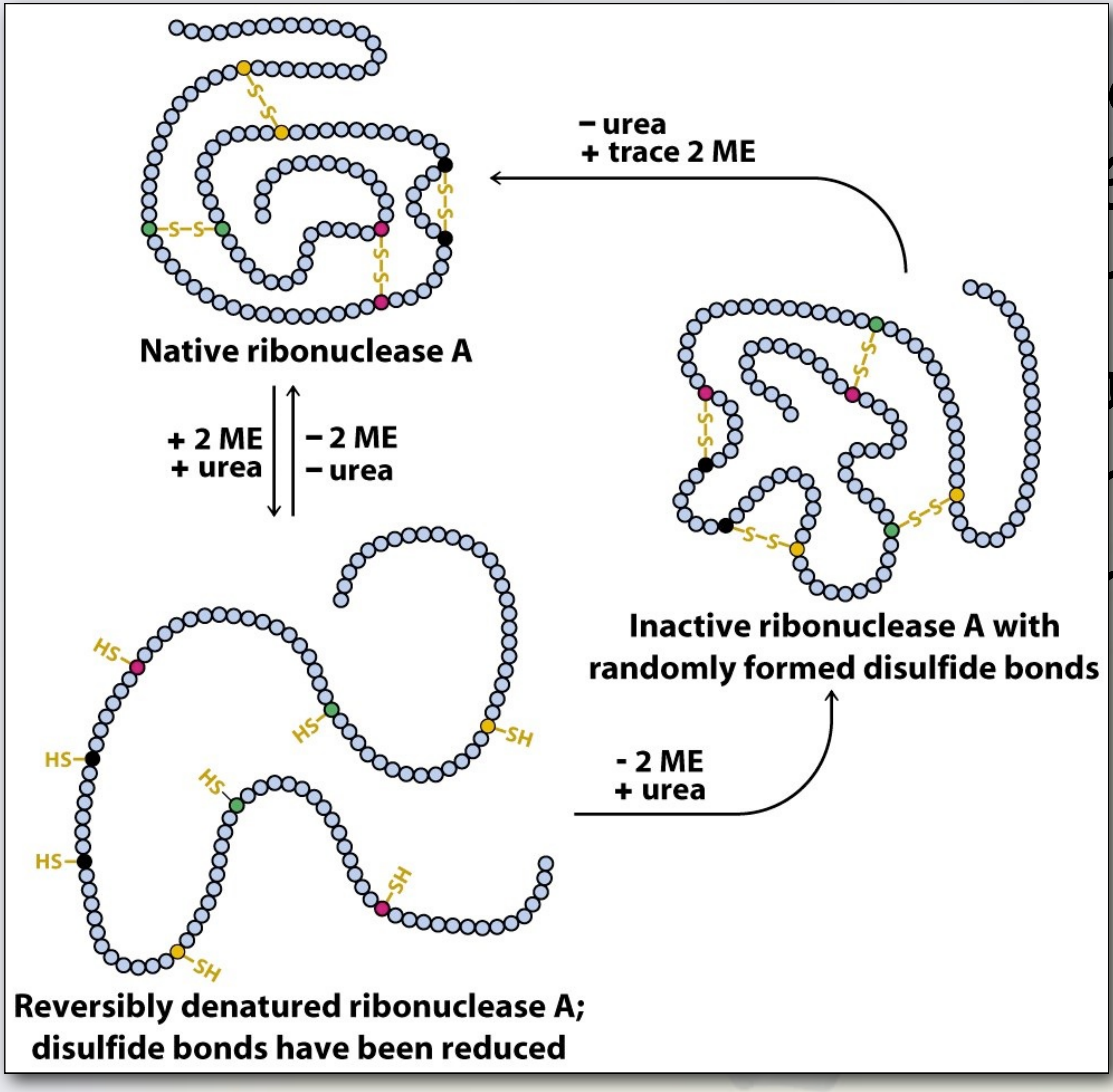
- Christain Anfinsen carried out a landmark experiment in 1961, which demonstrated that the primary sequence of a protein contained sufficient information to direct the folding of a peptide to its native state.

Protein Renaturization

- Christain Anfinsen carried out a landmark experiment in 1961, which demonstrated that the prim sequence of a protein contai sufficient information to dir folding of a peptide to its n state.

Christian Anfinsin
Nobel Prize in Chemistry
1972

Protein Renaturization

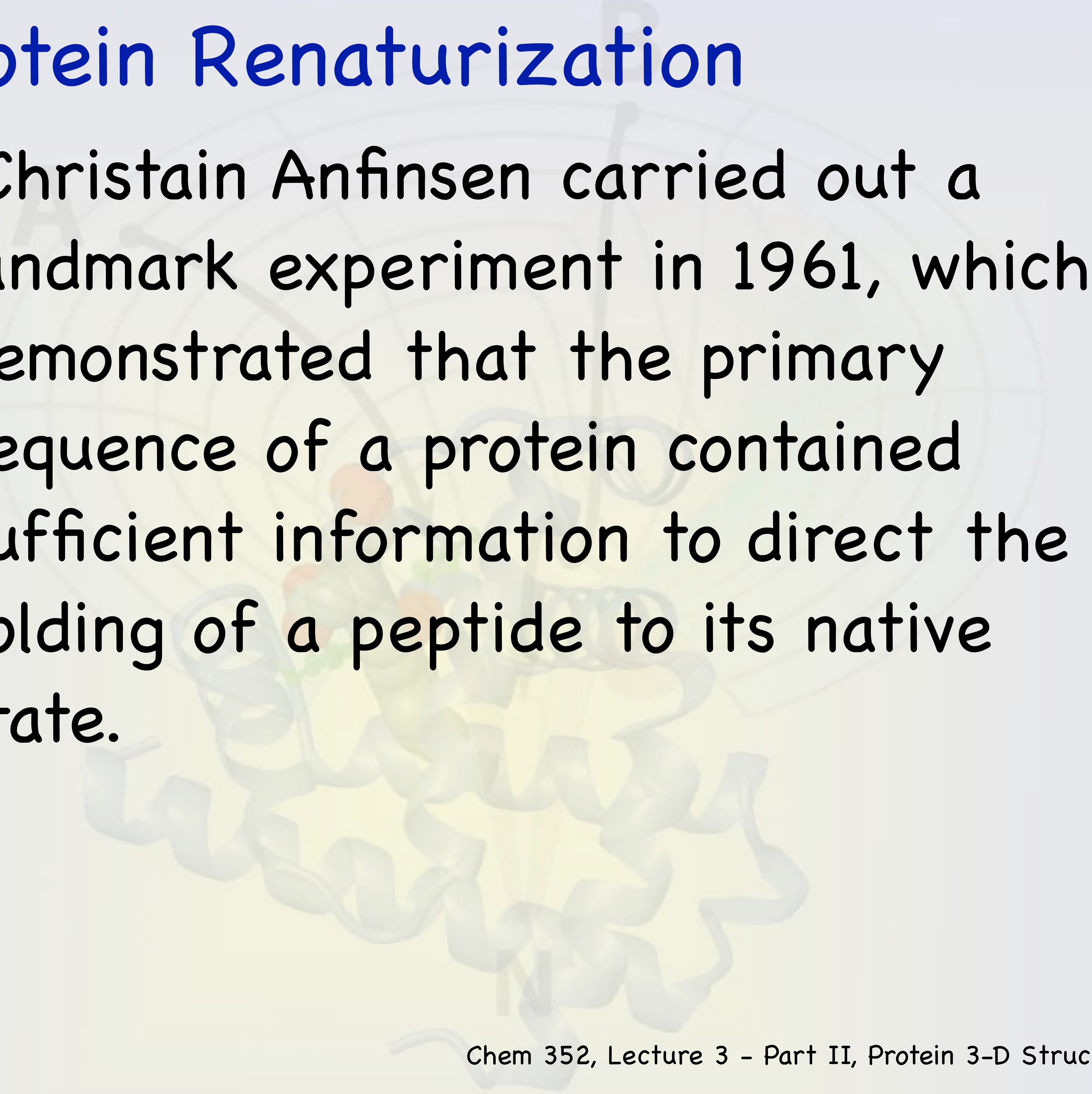


out a

1, which

Christian Anfinsin
Nobel Prize in Chemistry
1972

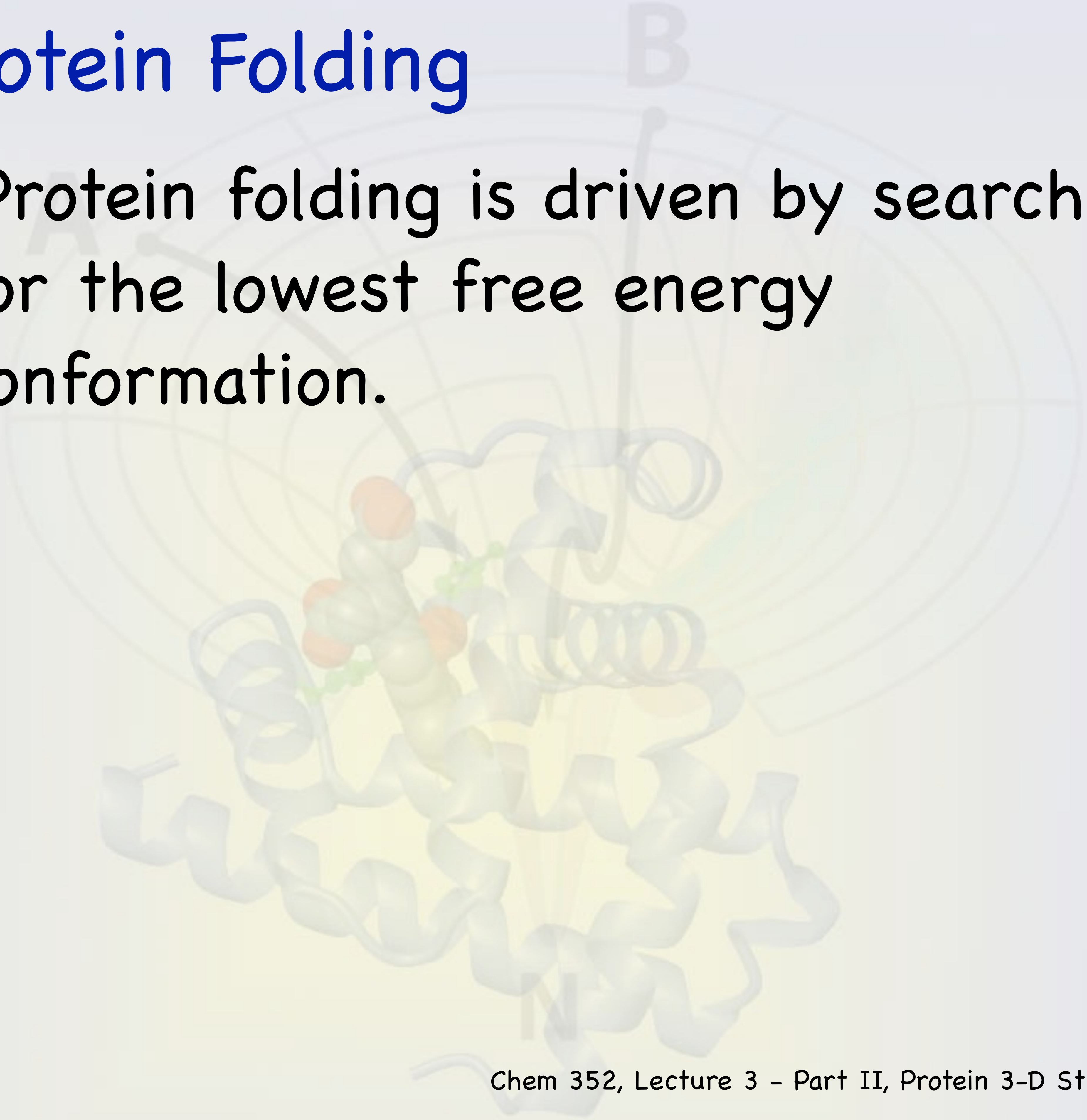
Protein Renaturization



- Christain Anfinsen carried out a landmark experiment in 1961, which demonstrated that the primary sequence of a protein contained sufficient information to direct the folding of a peptide to its native state.

Protein Folding

- Protein folding is driven by search for the lowest free energy conformation.

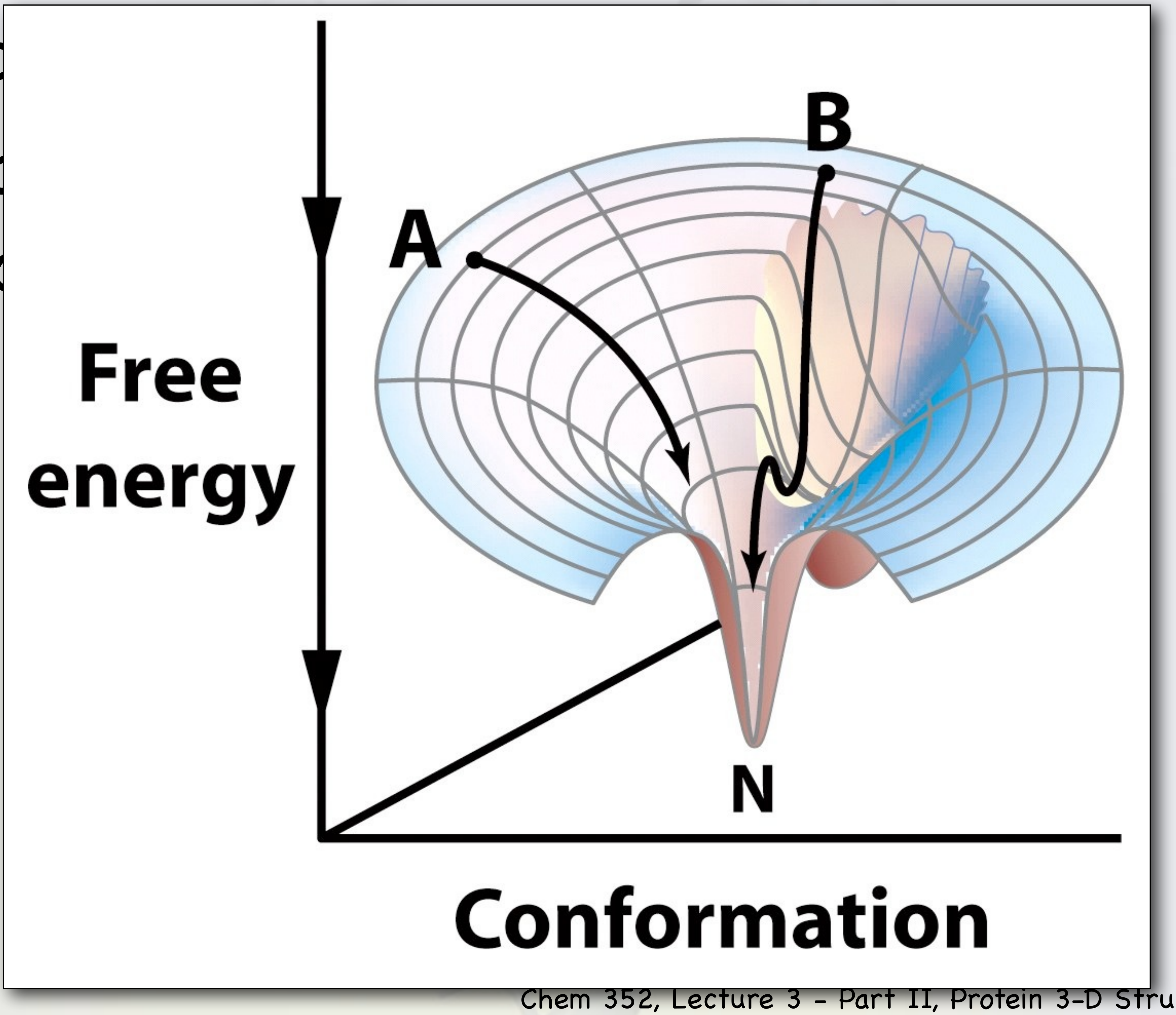


Protein Folding

- P

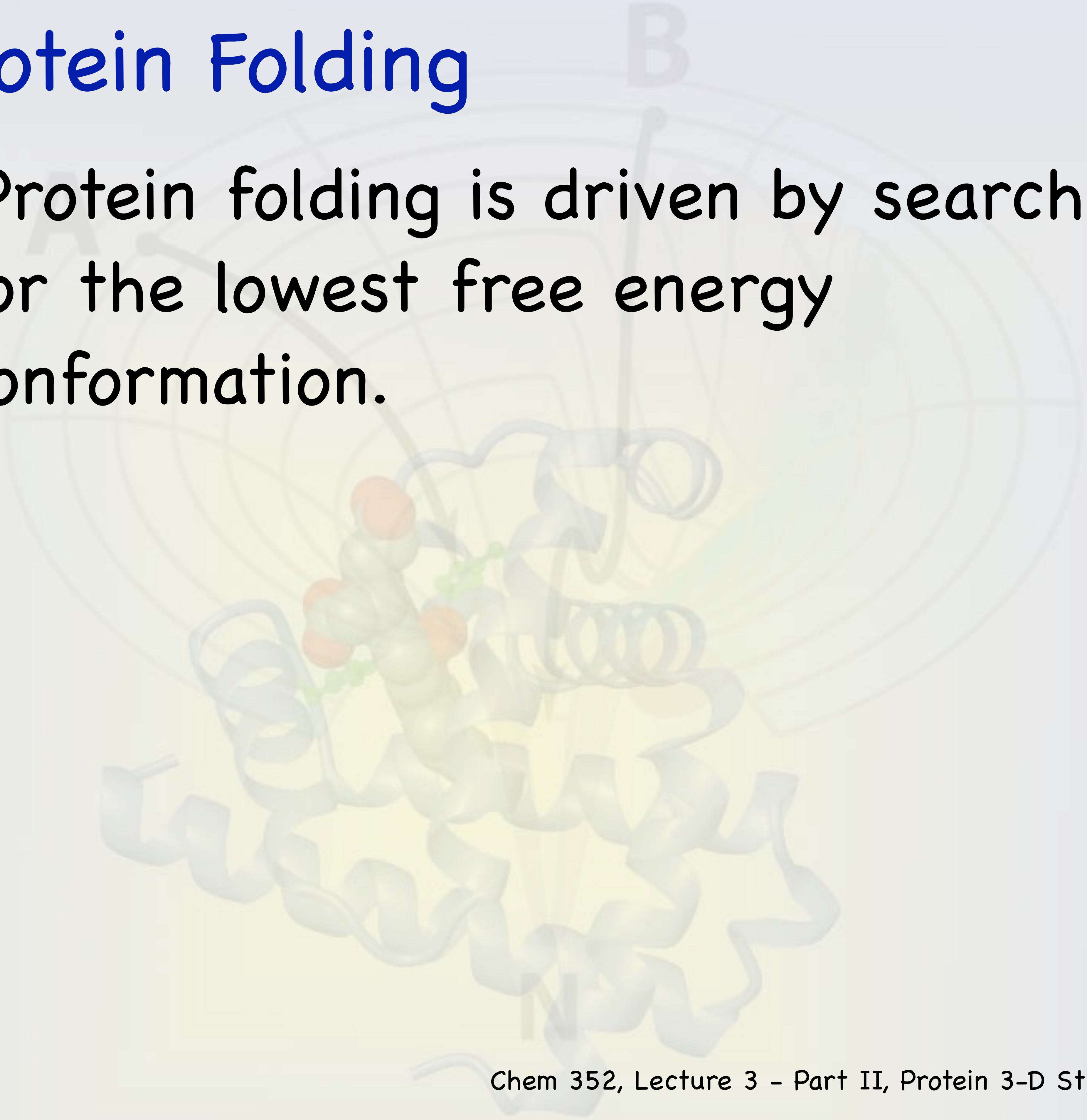
fo

co

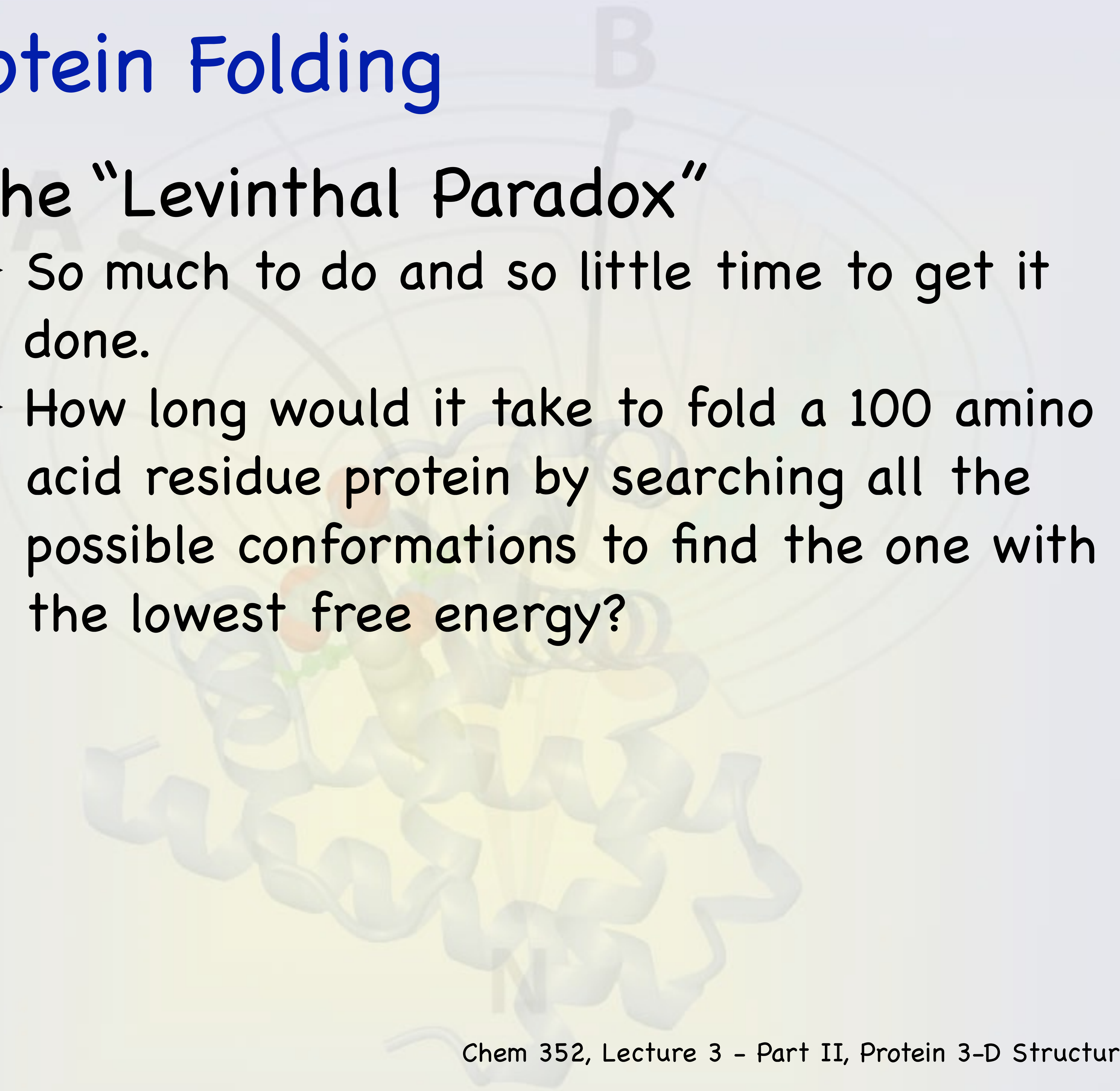


Protein Folding

- Protein folding is driven by search for the lowest free energy conformation.



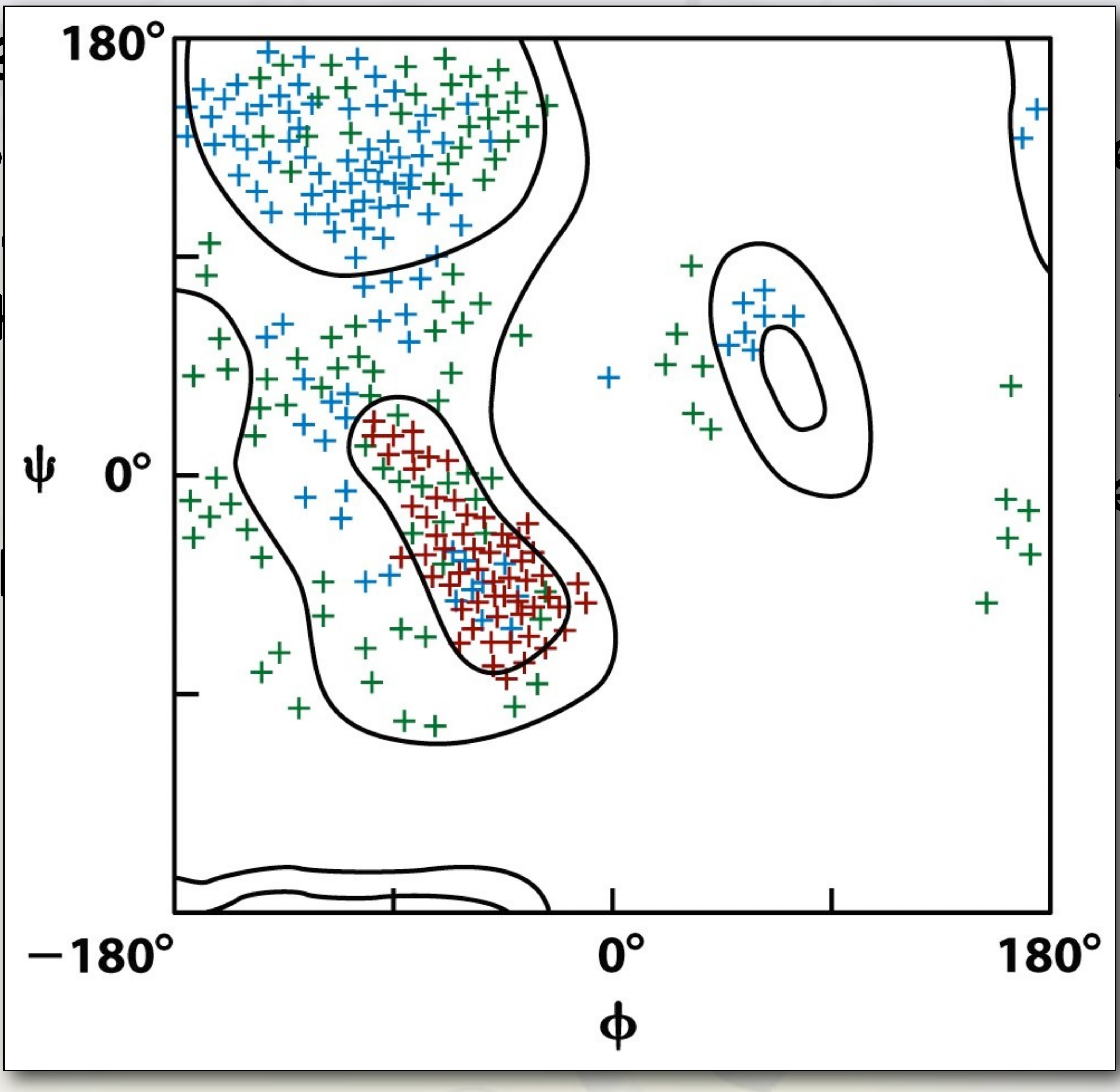
Protein Folding



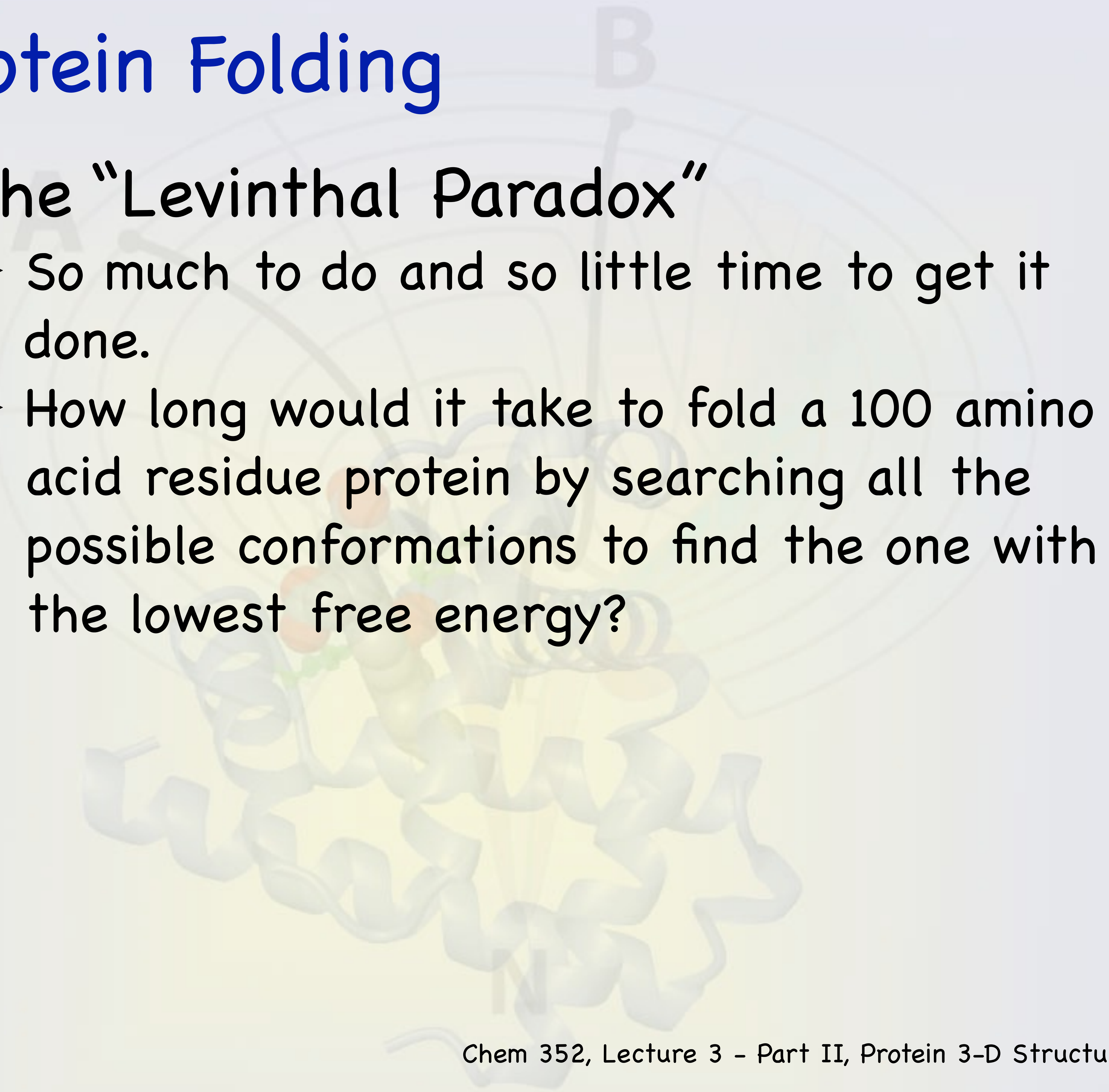
- The "Levinthal Paradox"
  - So much to do and so little time to get it done.
  - How long would it take to fold a 100 amino acid residue protein by searching all the possible conformations to find the one with the lowest free energy?

Protein Folding

- The
  - S
    et it
    d
  - H
    amino
    a
    the
    p
    e with
    t



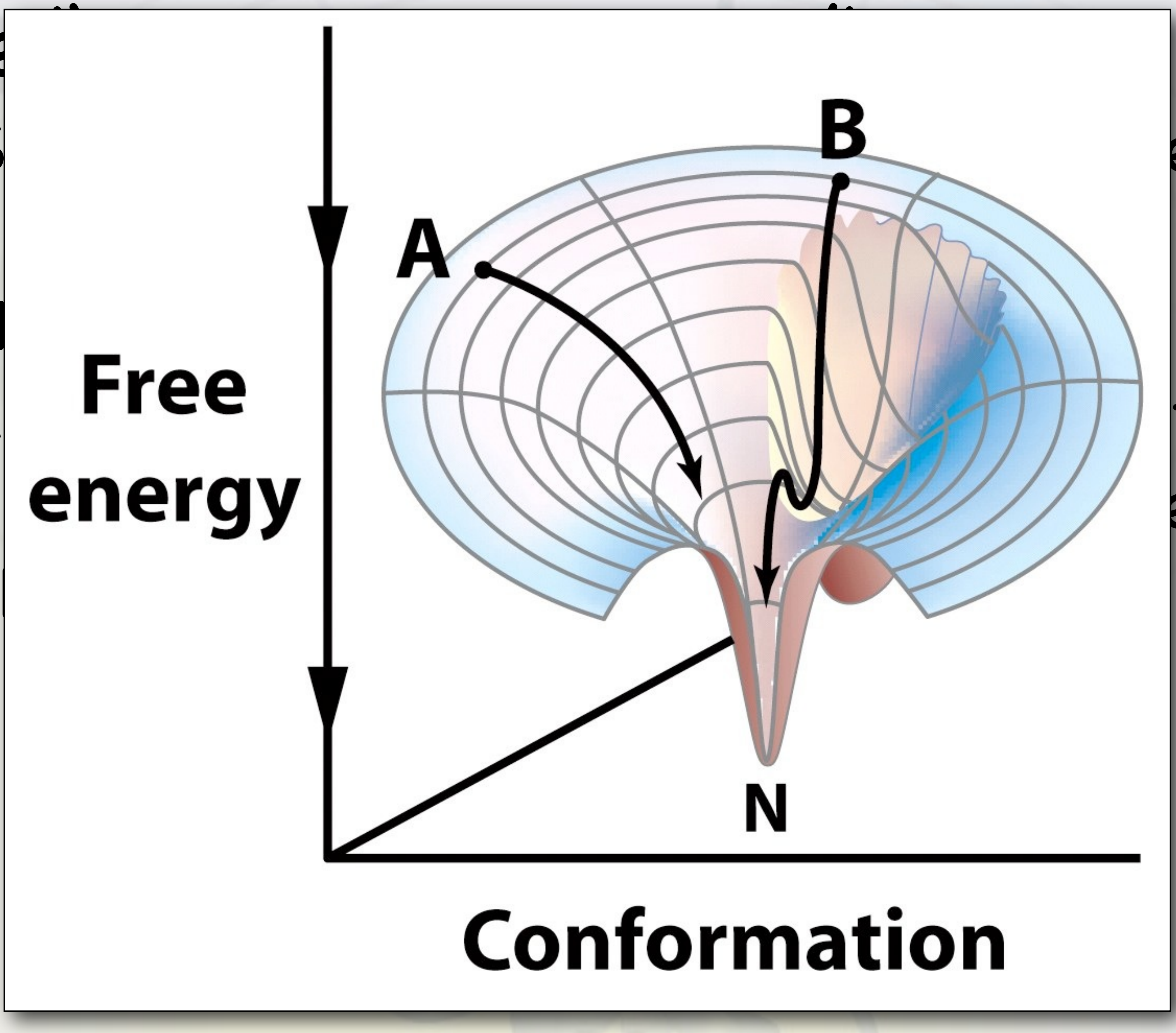
Protein Folding



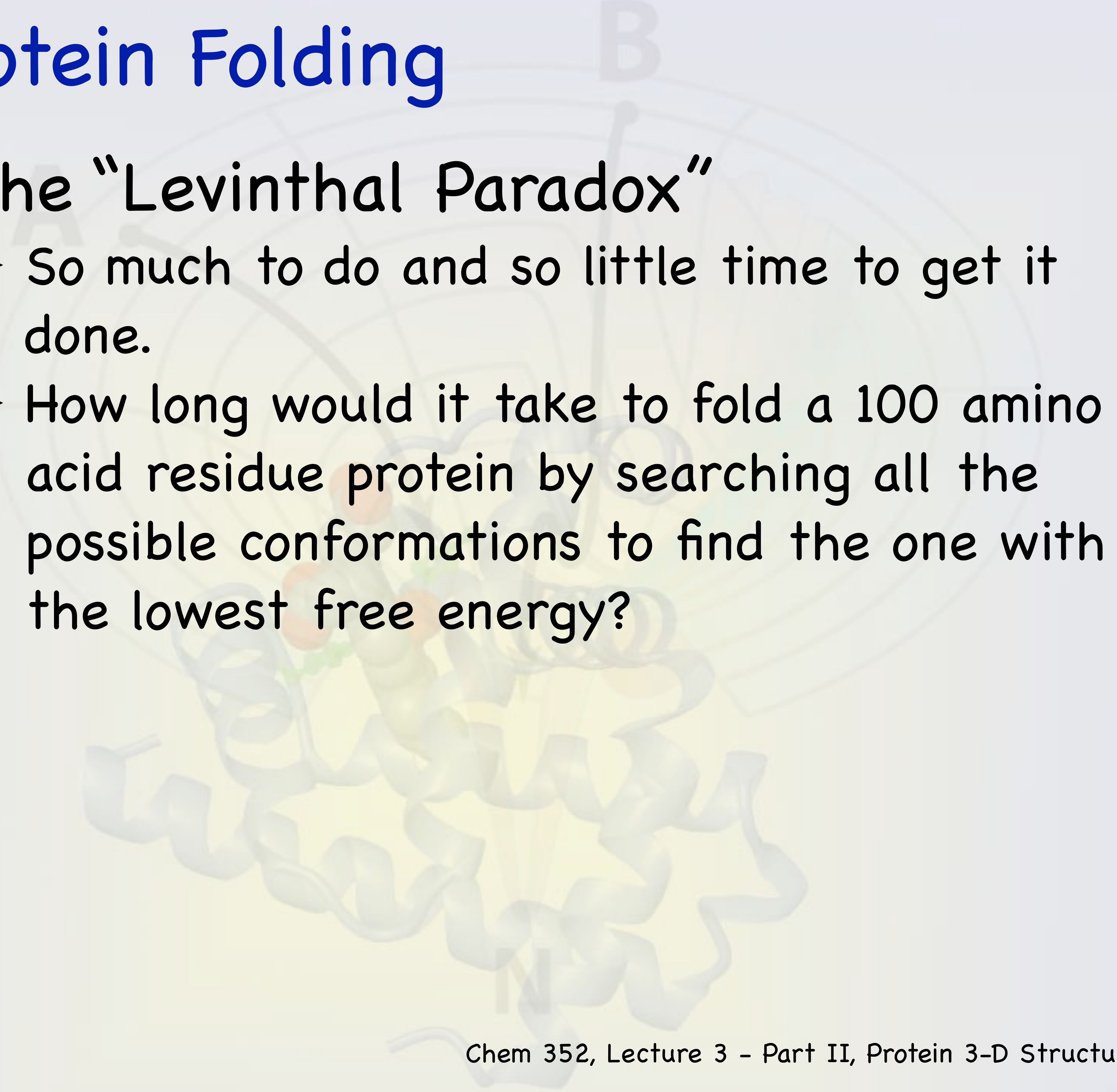
- The "Levinthal Paradox"
  - So much to do and so little time to get it done.
  - How long would it take to fold a 100 amino acid residue protein by searching all the possible conformations to find the one with the lowest free energy?

Protein Folding

- The
  - S
    et it
    d
  - H
    amino
    a
    the
    p
    e with
    t

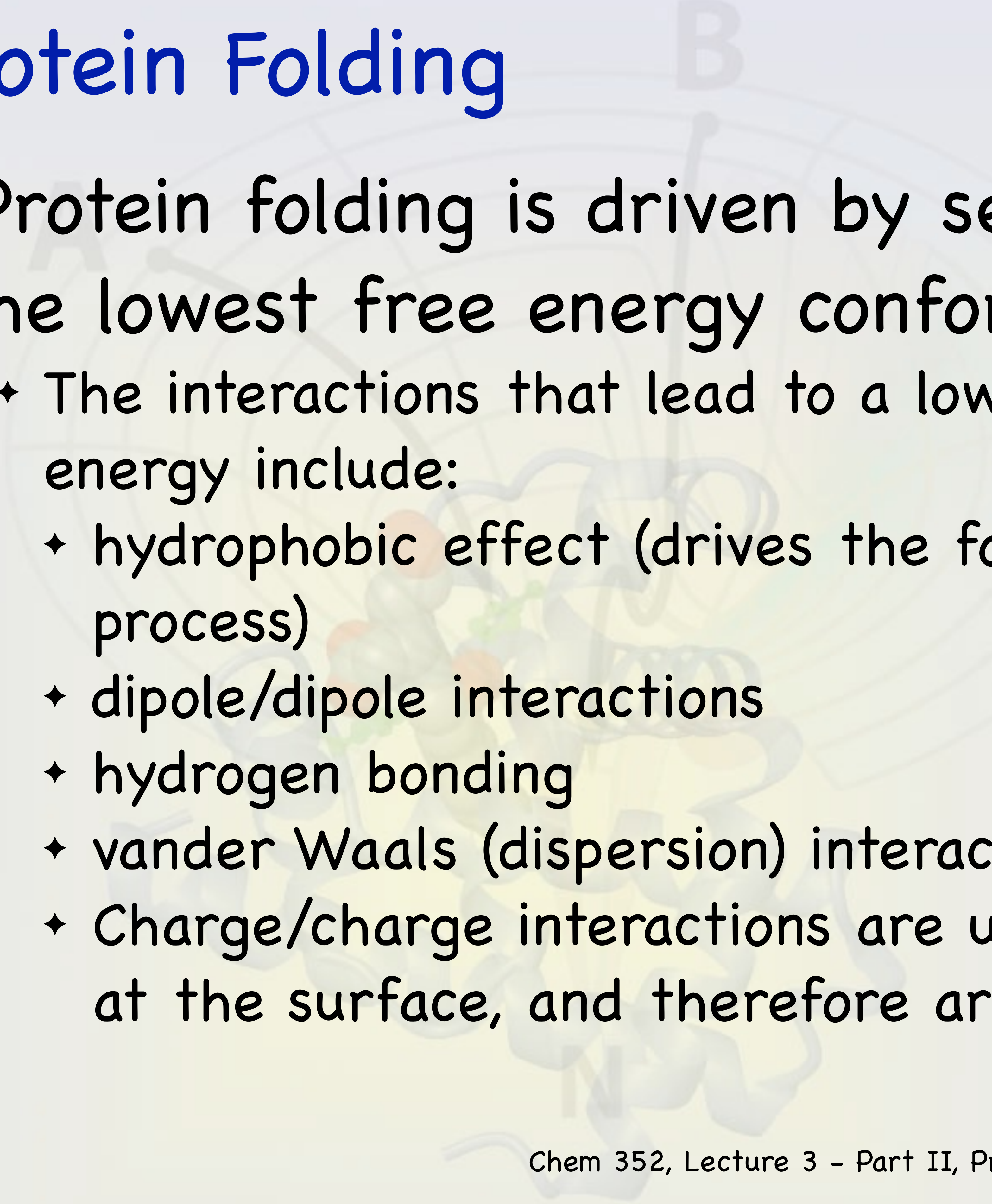


Protein Folding



- The "Levinthal Paradox"
  - So much to do and so little time to get it done.
  - How long would it take to fold a 100 amino acid residue protein by searching all the possible conformations to find the one with the lowest free energy?

Protein Folding



- Protein folding is driven by search for the lowest free energy conformation.
  - The interactions that lead to a lower free energy include:
    - hydrophobic effect (drives the folding process)
    - dipole/dipole interactions
    - hydrogen bonding
    - vander Waals (dispersion) interactions
    - Charge/charge interactions are usually found at the surface, and therefore are strongly

Protein Folding

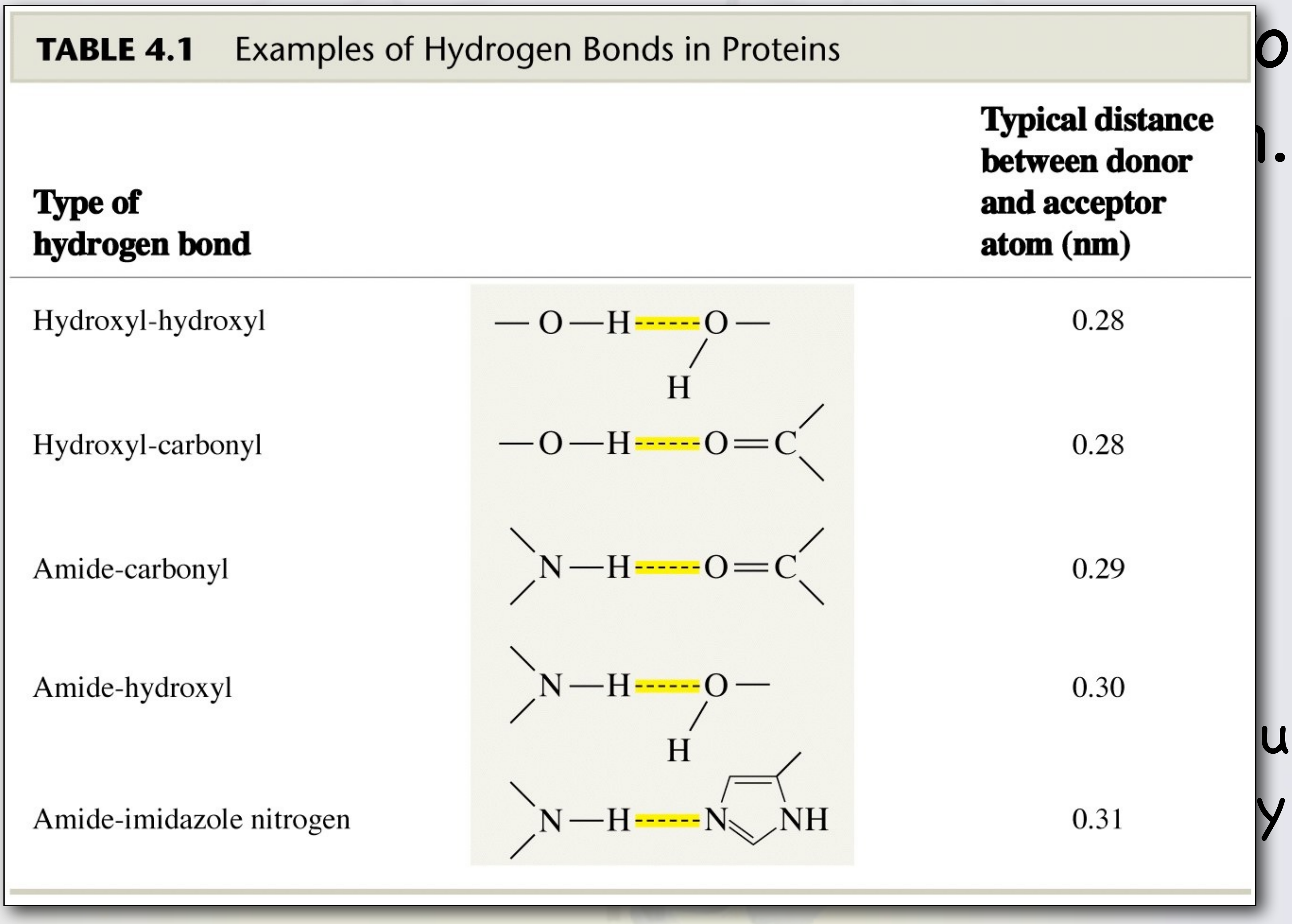


TABLE 4.1 Examples of Hydrogen Bonds in Proteins

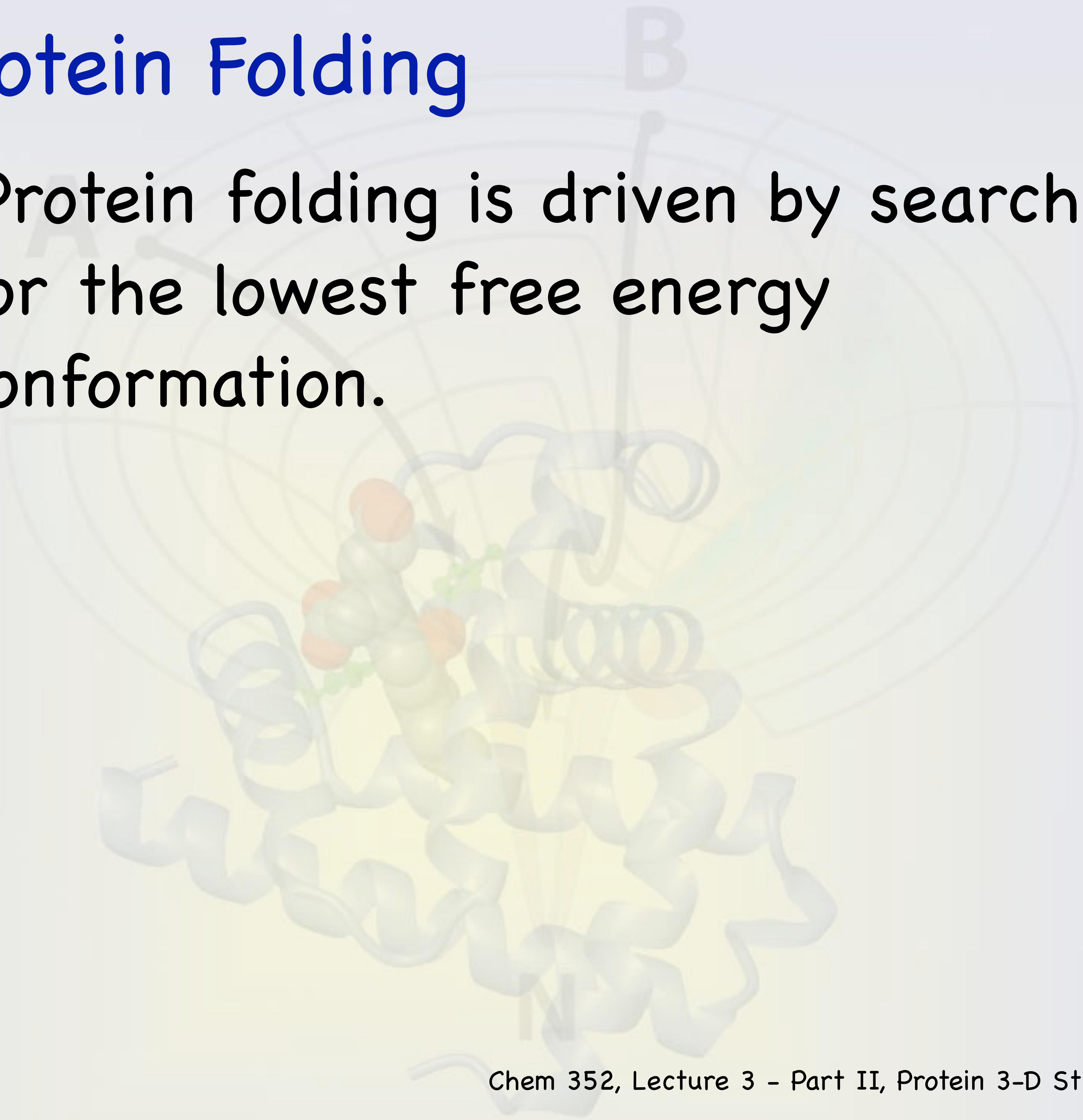
Type of hydrogen bond		**Typical distance between donor and acceptor atom (nm)**
Hydroxyl-hydroxyl		0.28
Hydroxyl-carbonyl		0.28
Amide-carbonyl		0.29
Amide-hydroxyl		0.30
Amide-imidazole nitrogen		0.31

Protein Folding

- Protein folding is driven by search for the lowest free energy conformation.
  - The interactions that lead to a lower free energy include:
    - hydrophobic effect (drives the folding process)
    - dipole/dipole interactions
    - hydrogen bonding
    - vander Waals (dispersion) interactions
    - Charge/charge interactions are usually found at the surface, and therefore are strongly

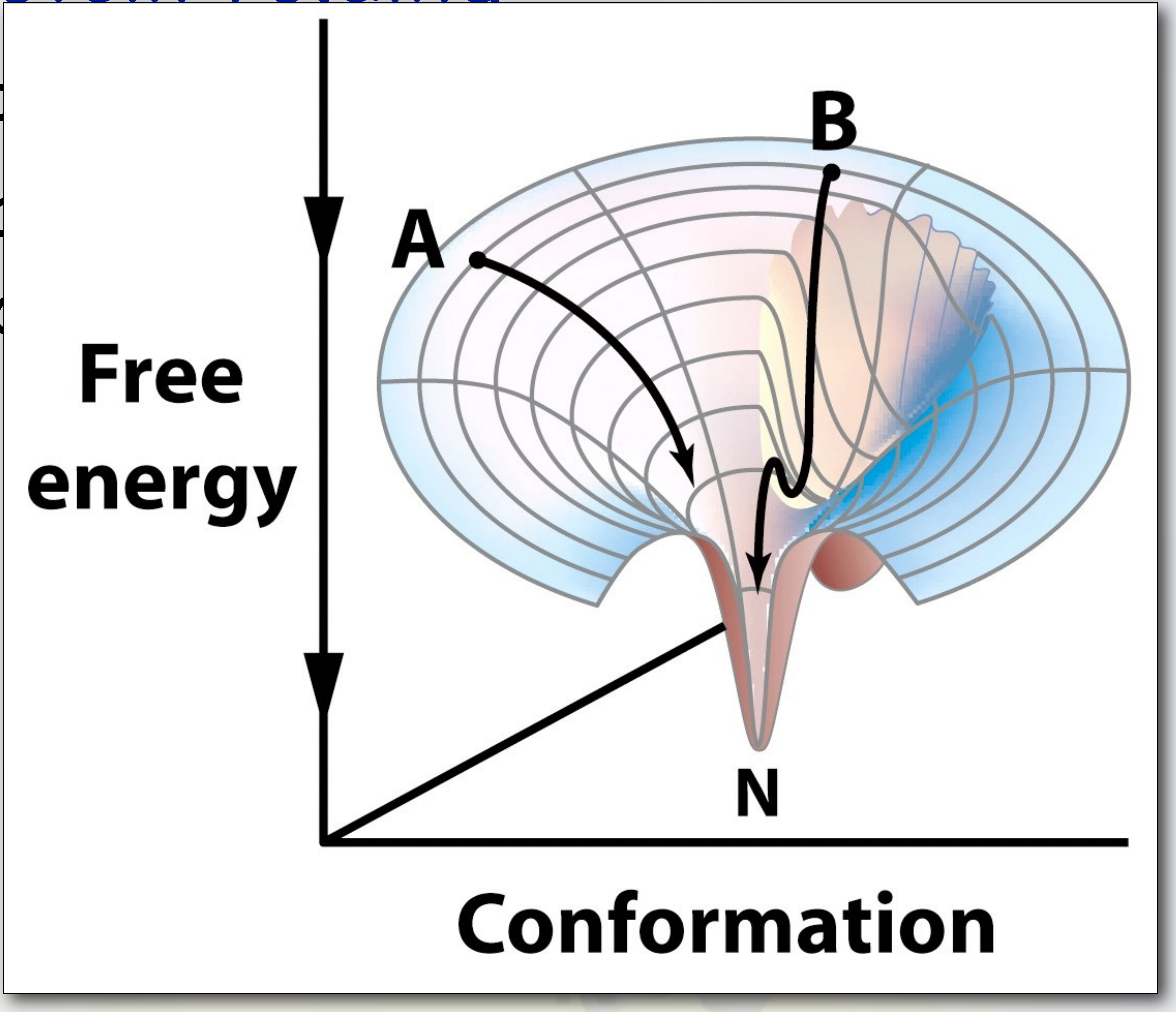
Protein Folding

- Protein folding is driven by search for the lowest free energy conformation.



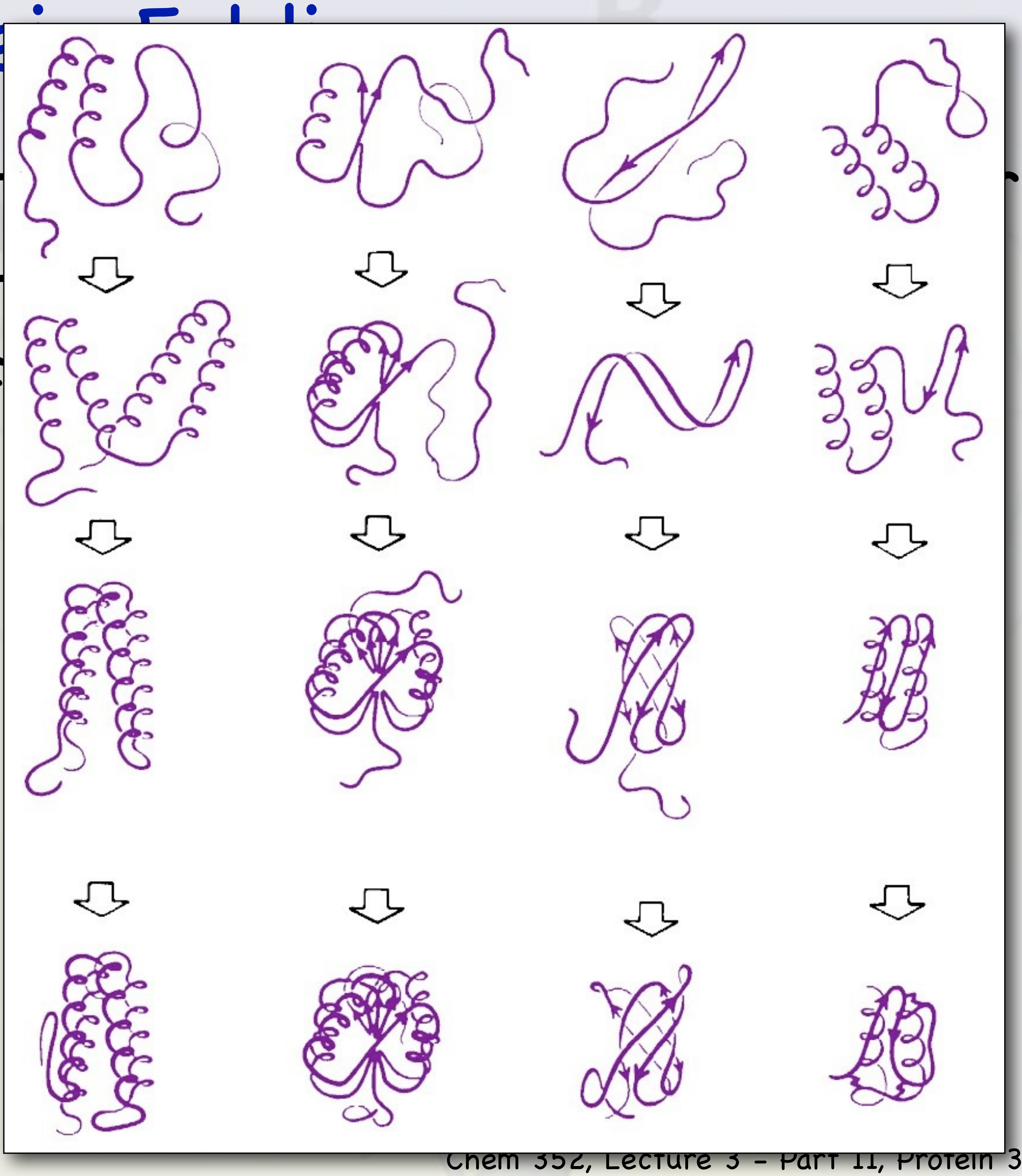
Protein Folding

•P
fo
co



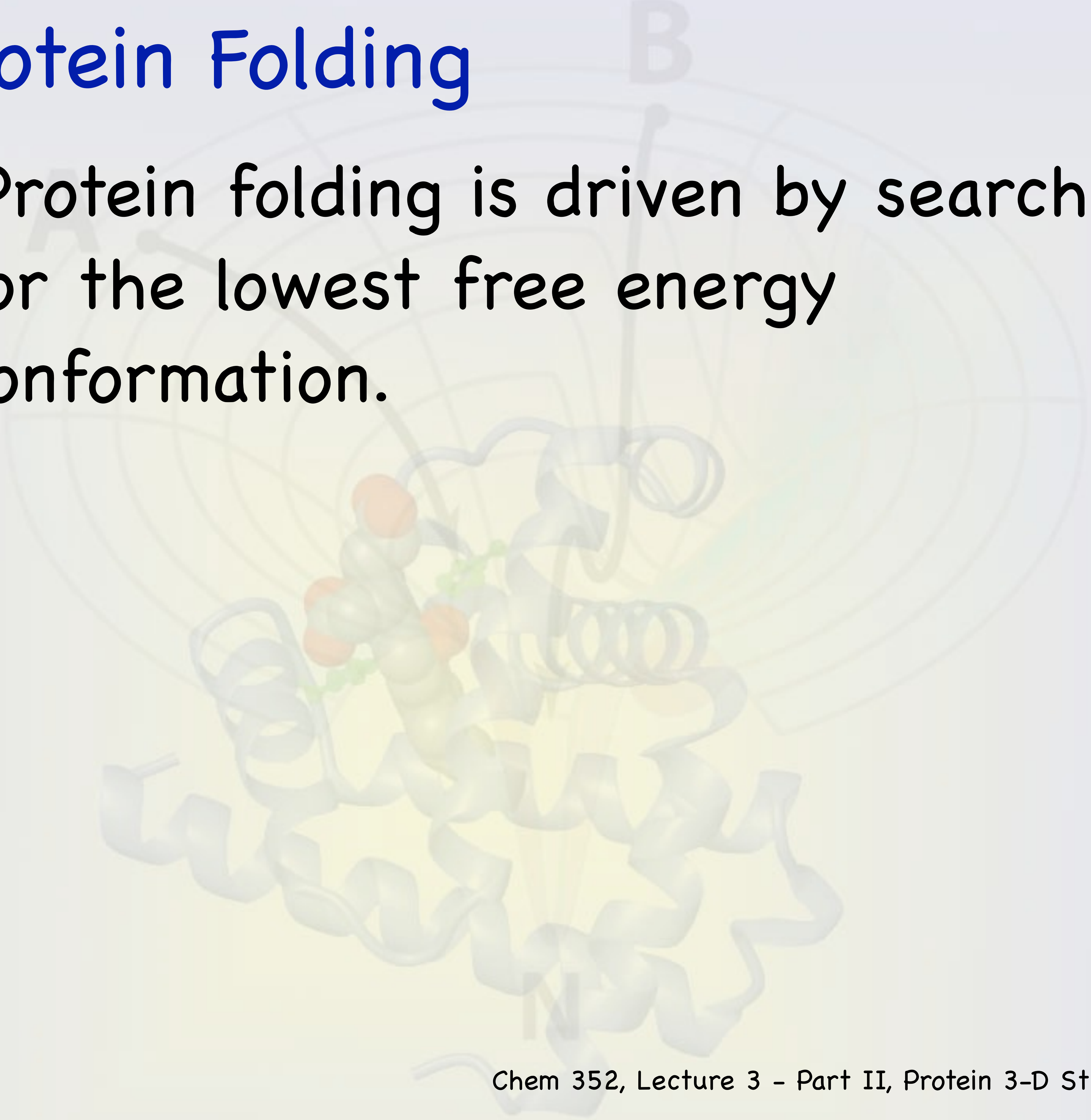
Prote

•Pro rch
for
conf



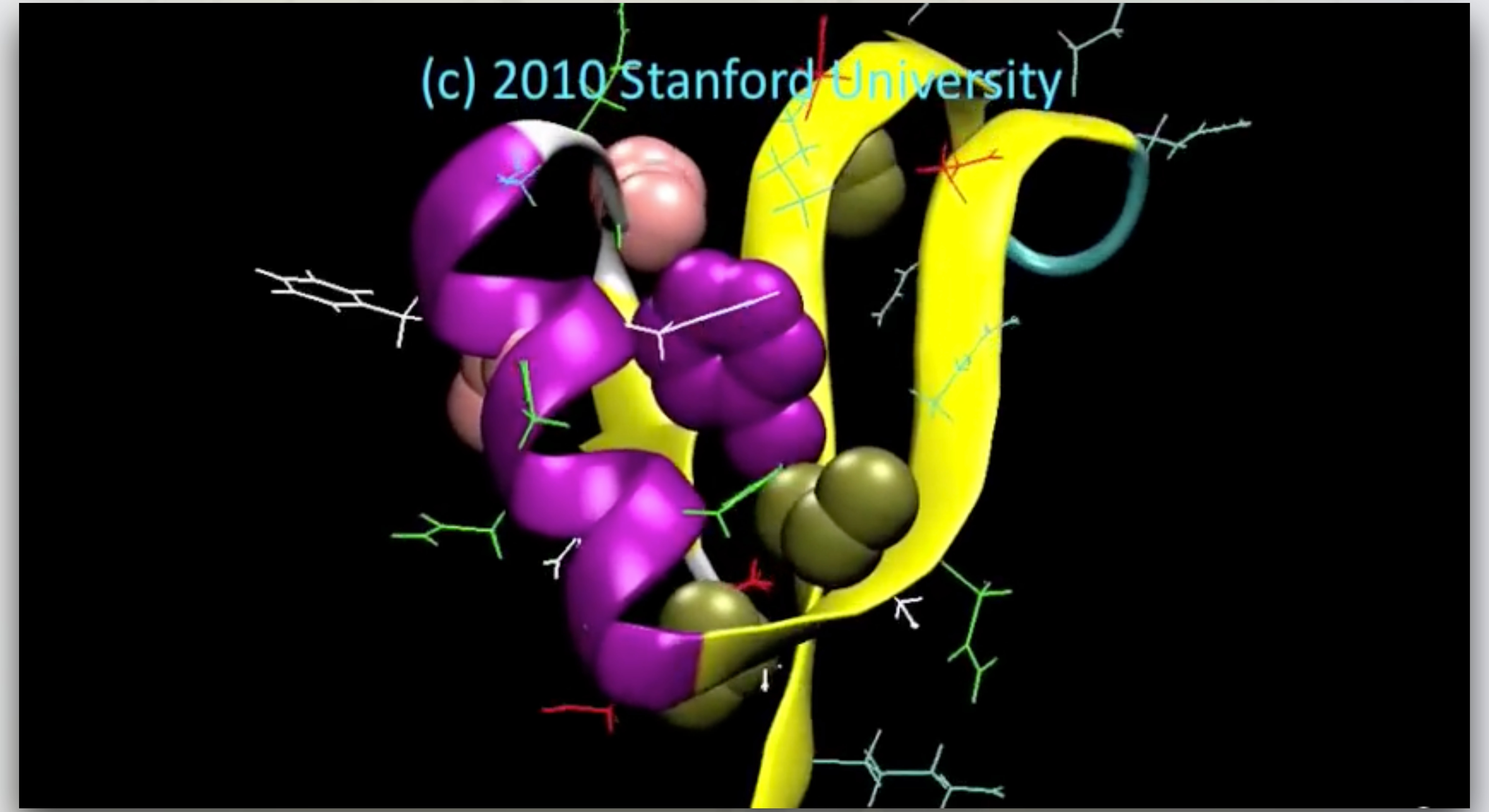
Protein Folding

- Protein folding is driven by search for the lowest free energy conformation.



Protein Folding

- Folding@Home (http://folding.stanford.edu)



Protein Folding

Click to learn how you contribute to science by playing Foldit.

GET STARTED: DOWNLOAD

Windows (XP/Vista/7)

OSX (Intel 10.5 or later)

Linux (64-bit)

Are you new to Foldit? Click here.

Are you an educator? Click here.

SEARCH

☑ Only search fold.it

RECOMMEND FOLDIT

USER LOGIN

Username: *

Password: *

- Create new account
- Request new password

Try our new scientific discovery game!

Be creative and build extraordinary tiny machines!

What's New

Community News, Fall Edition

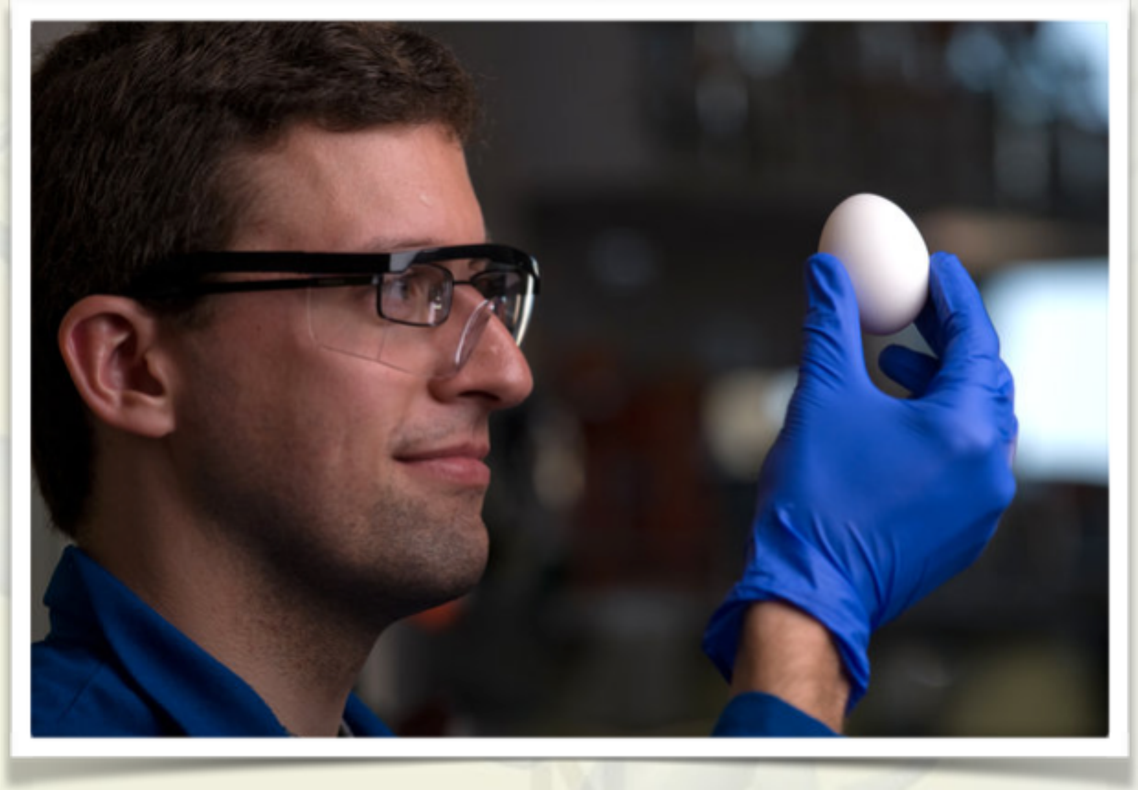
The last fourth of the year has arrived, and with it, there's some community news to share.

First, in October, we're having a Science Chat! This is your opportunity to chat with Foldit scientists about puzzles and general science. Mark down October 16 - you won't want to miss it. Details coming later this week!

Protein Folding

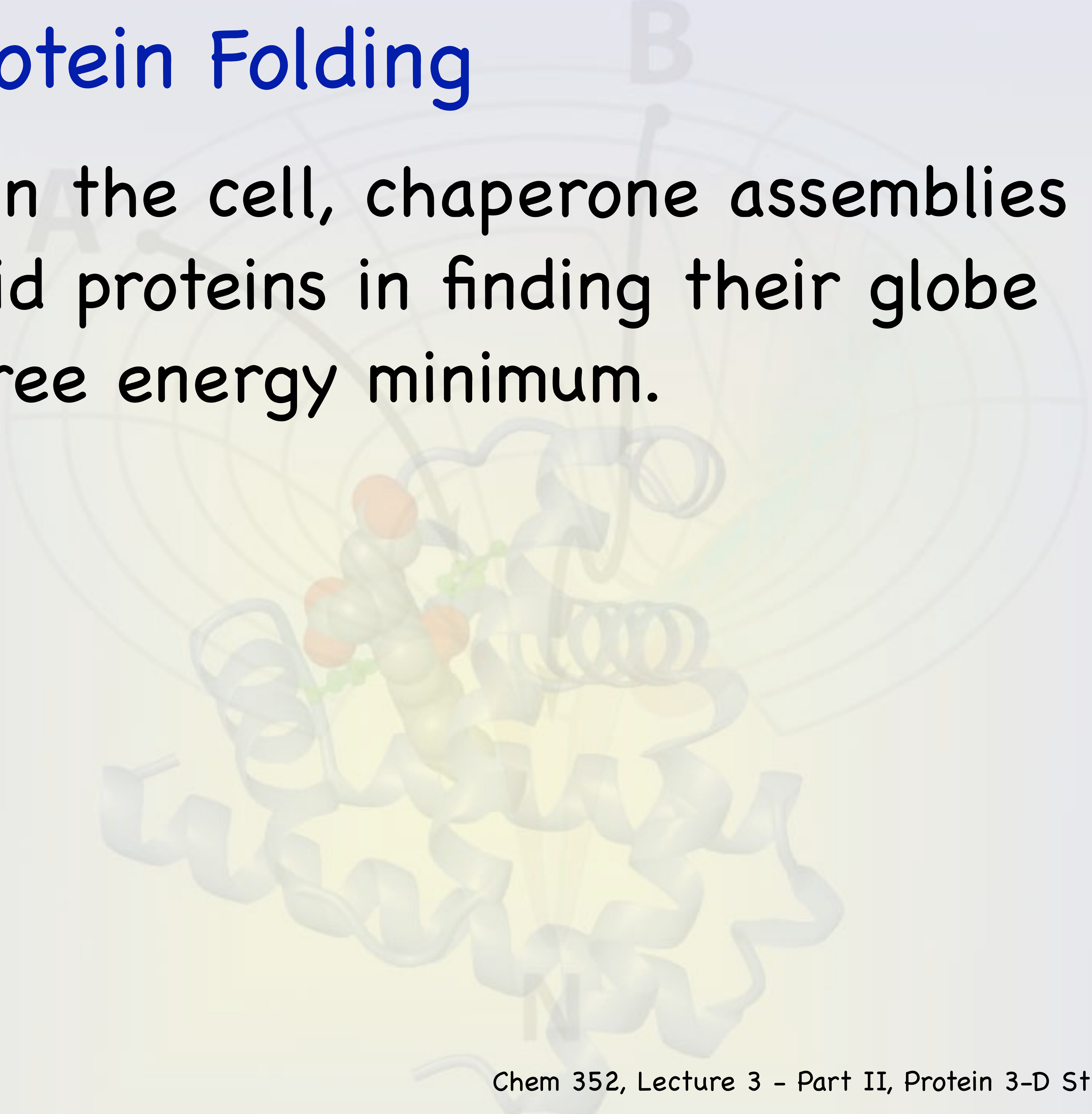
Unboiling an egg

- NPR report
- Yuan, T. Z. et al. (2015) Shear-Stress-Mediated Refolding of Proteins from Aggregates and Inclusion Bodies. (2015) Shear-Stress-Mediated Refolding of Proteins from Aggregates and Inclusion Bodies. ChemBioChem 16, 393-396.



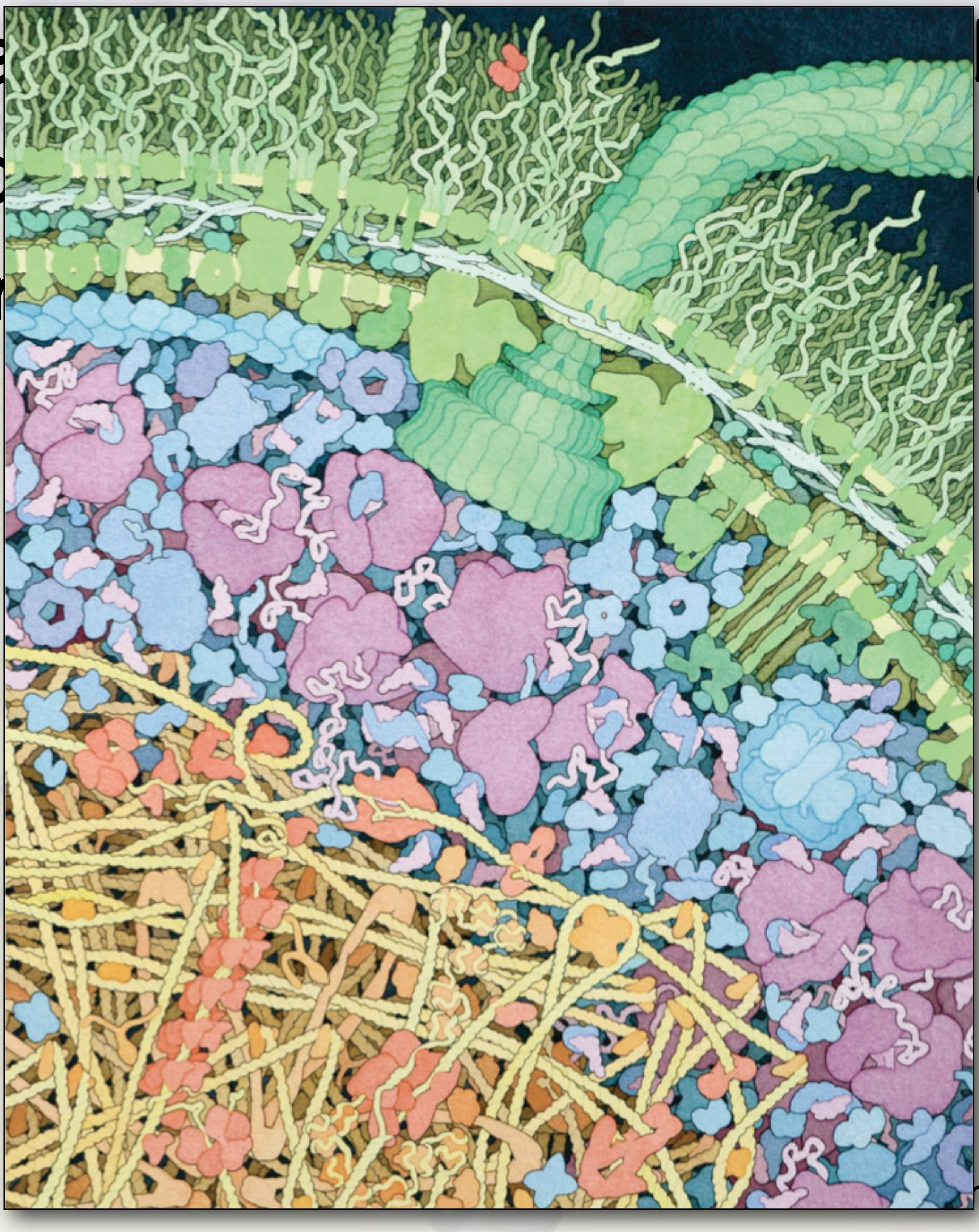
Protein Folding

- In the cell, chaperone assemblies aid proteins in finding their globe free energy minimum.



Protein Folding

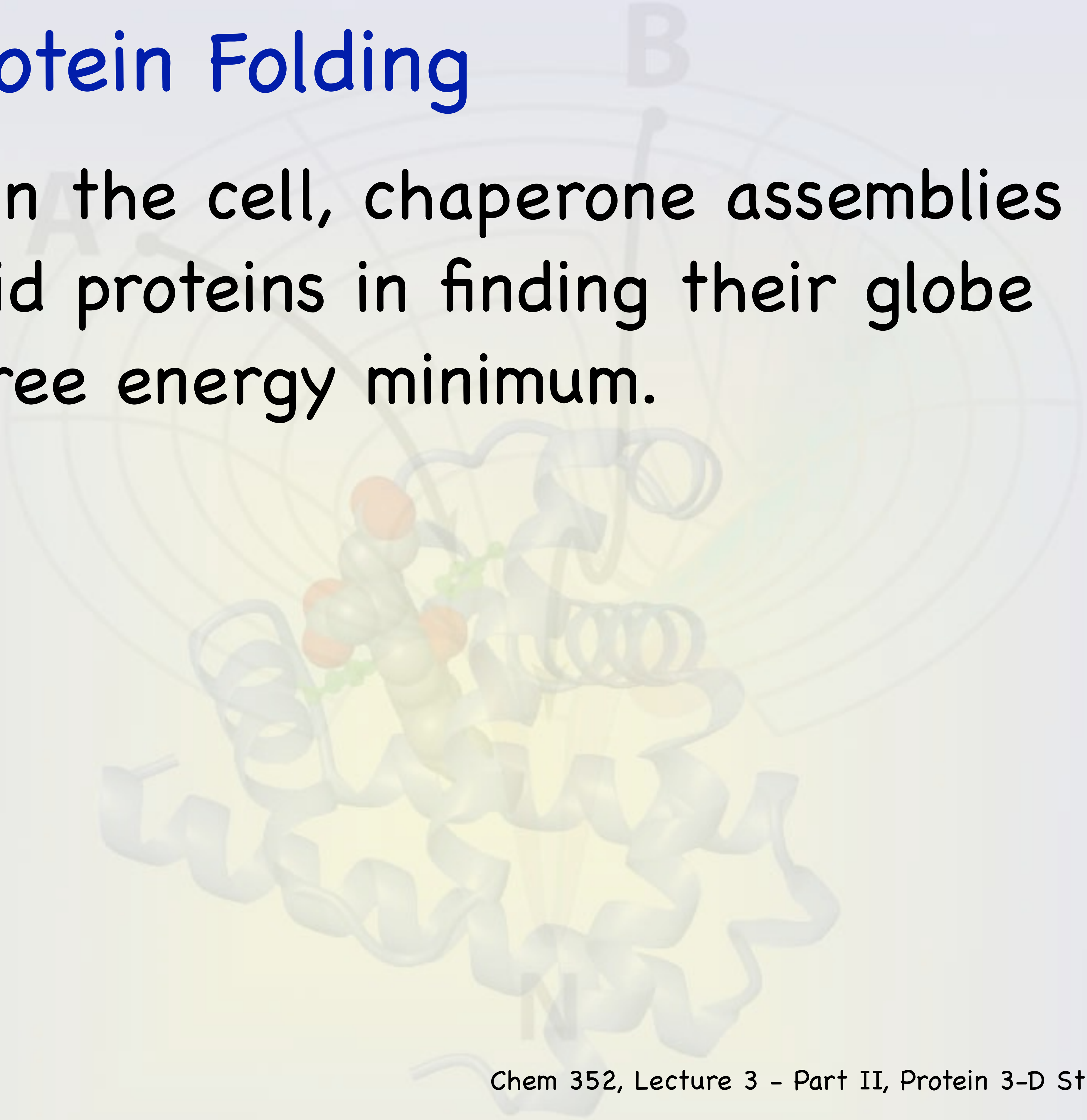
- In the
aid pro
free en



blies
obe

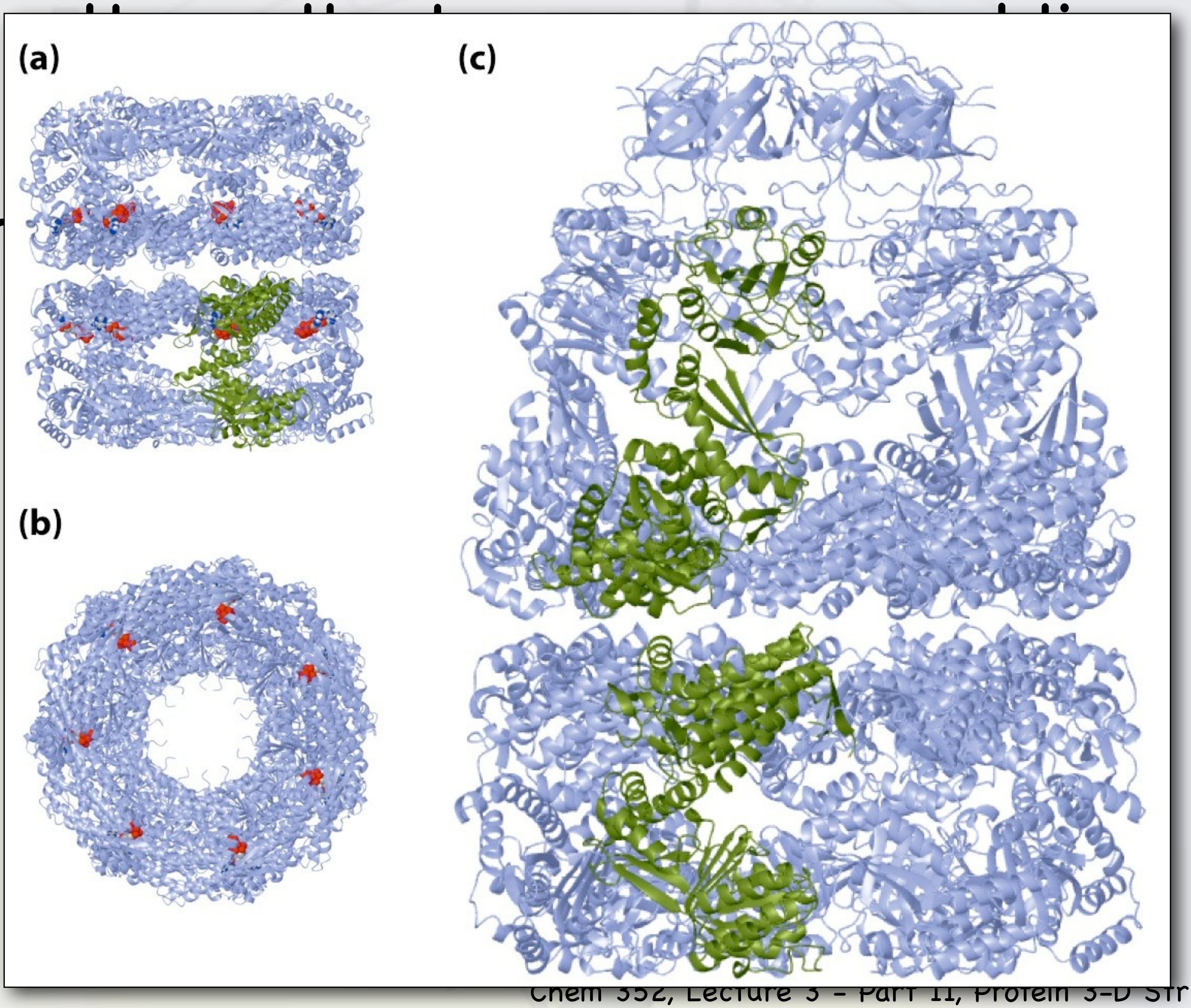
Protein Folding

- In the cell, chaperone assemblies aid proteins in finding their globe free energy minimum.



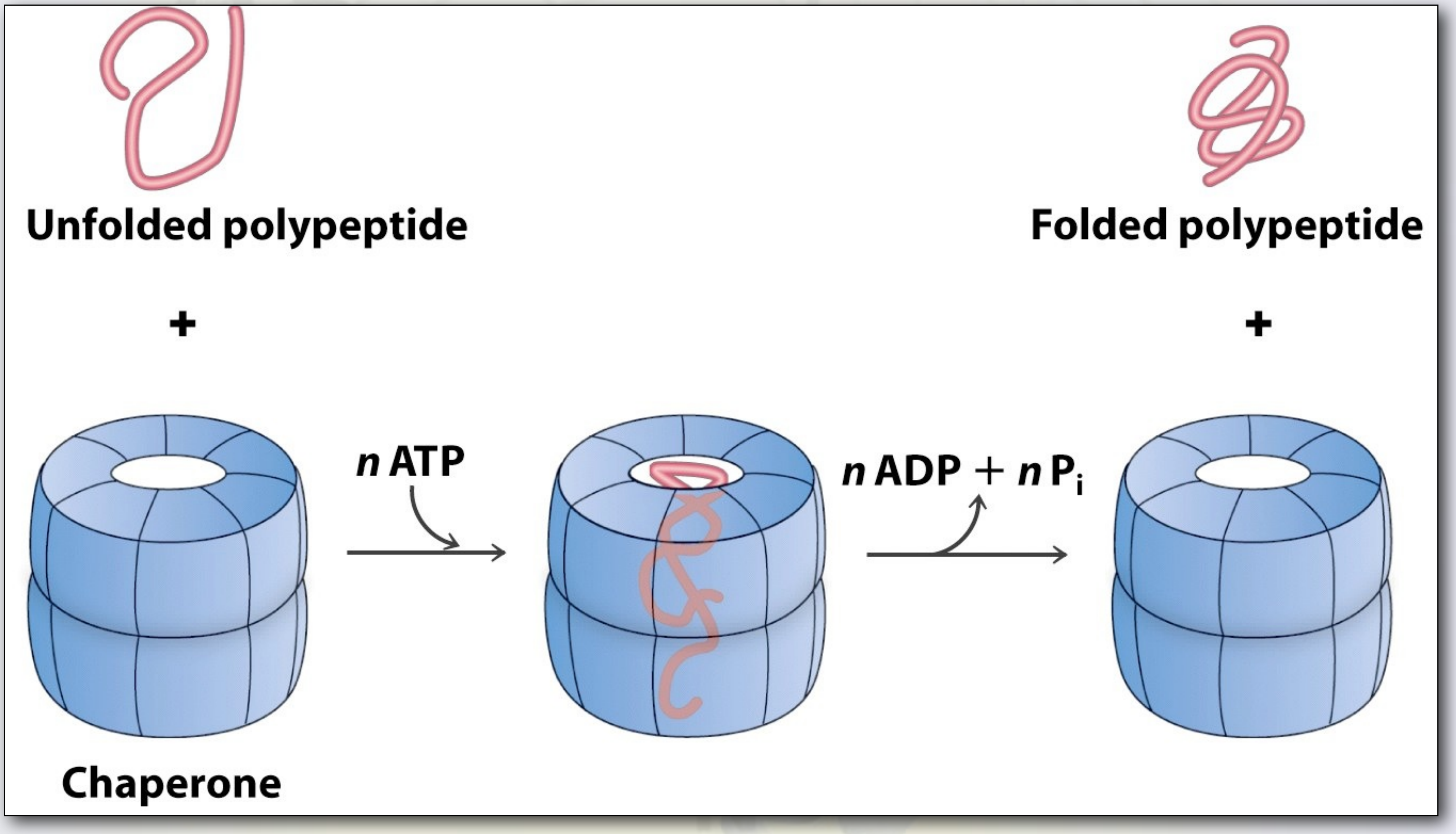
Protein Folding

•I
ai
fr



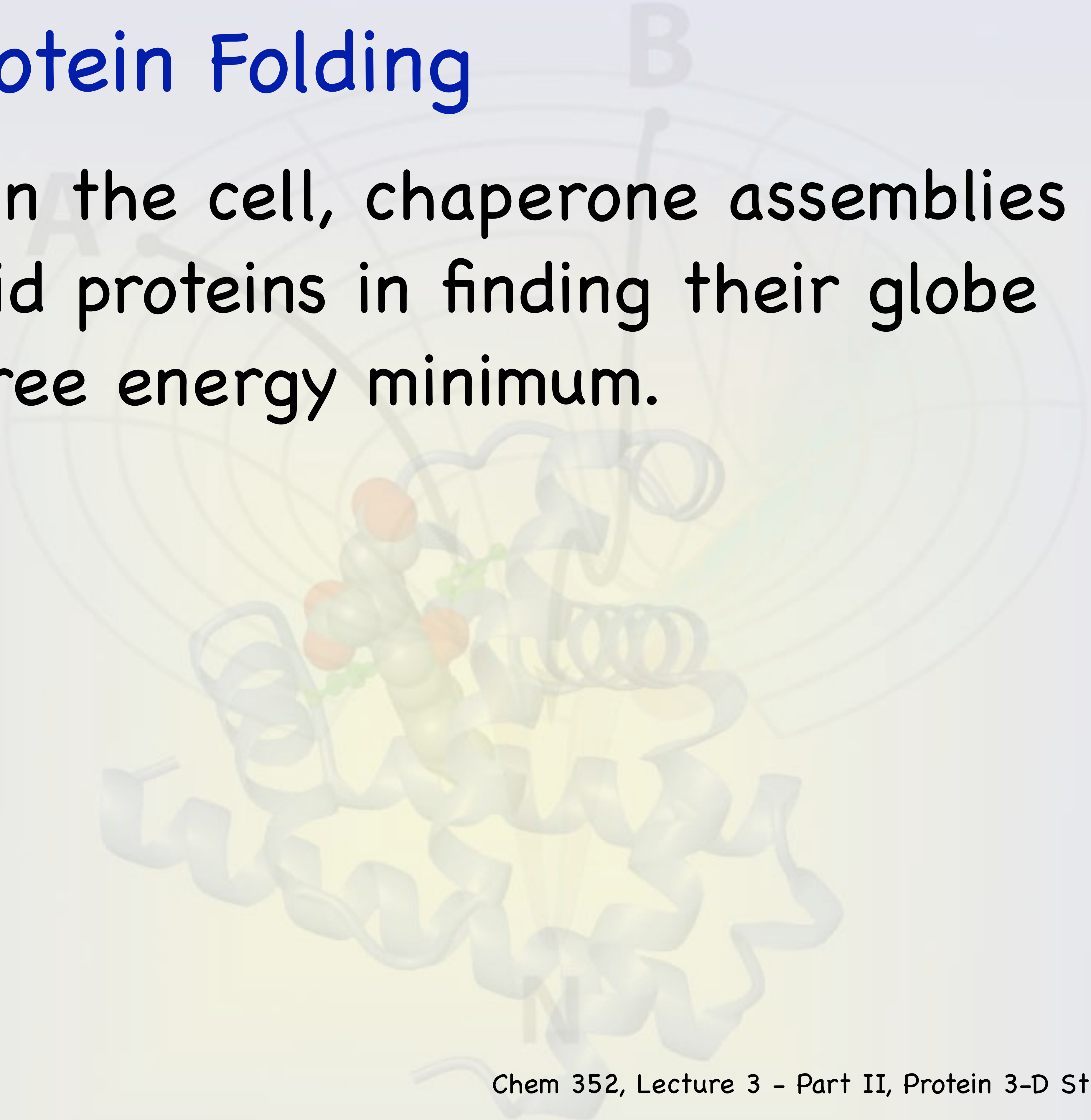
Protein Folding

- In the cell, chaperone assemblies



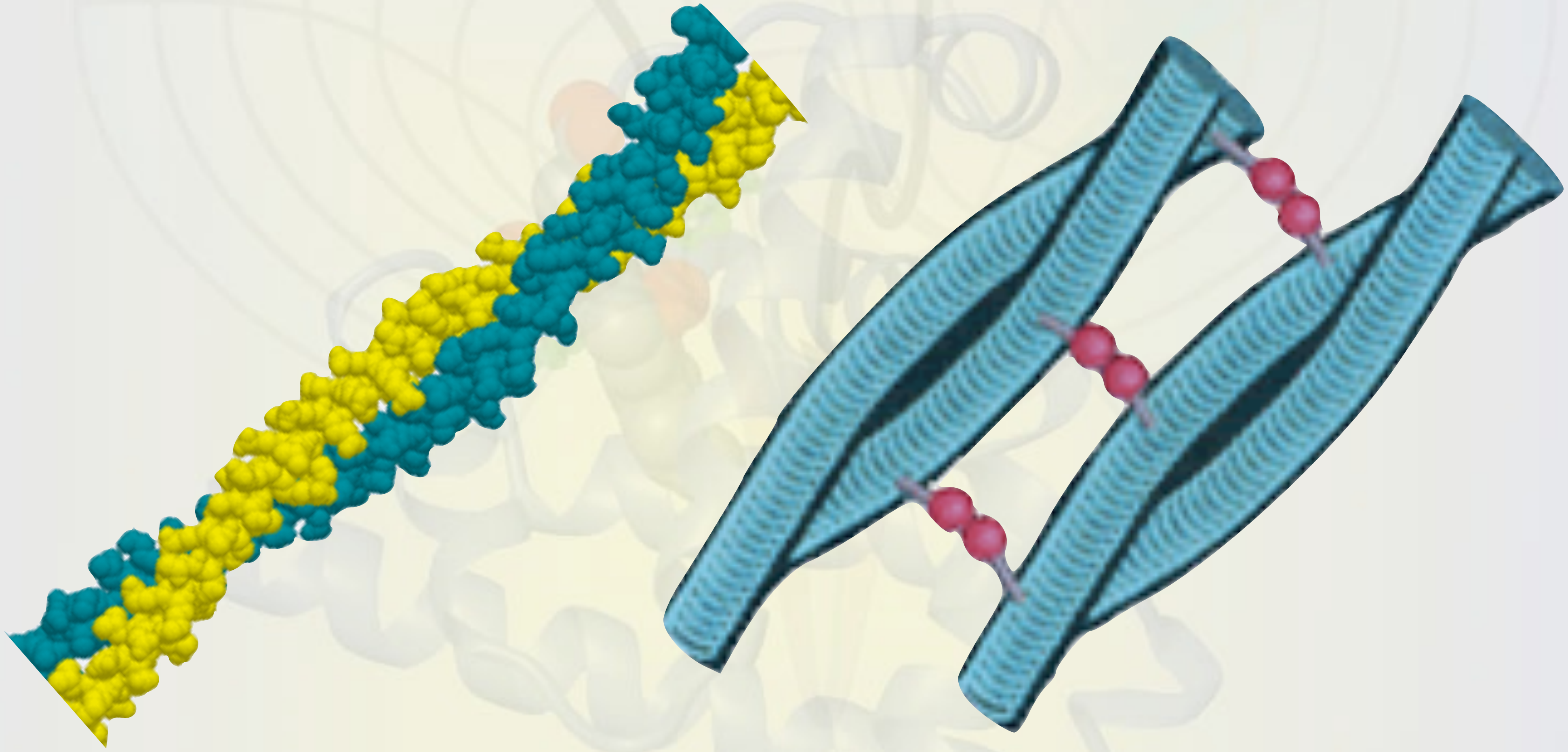
Protein Folding

- In the cell, chaperone assemblies aid proteins in finding their globe free energy minimum.



Fibrous Proteins

- Fibrous proteins have primary, secondary and quaternary structure.
  - The lack tertiary structure



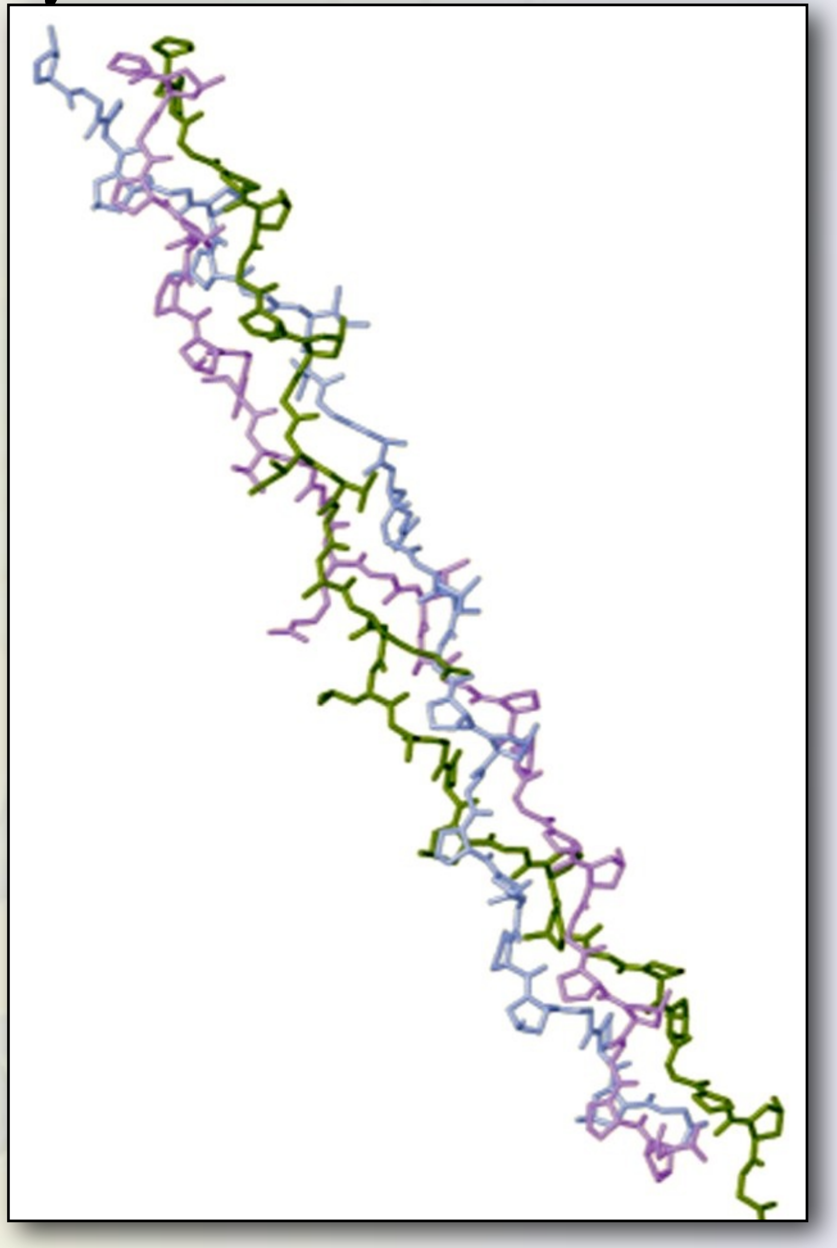
α-keratin is made of “coiled-coiled” α-helices

Fibrous Proteins

- Fibrous proteins have primary, secondary and quaternary structure.
  - The lack tertiary structure

Fibrous Proteins

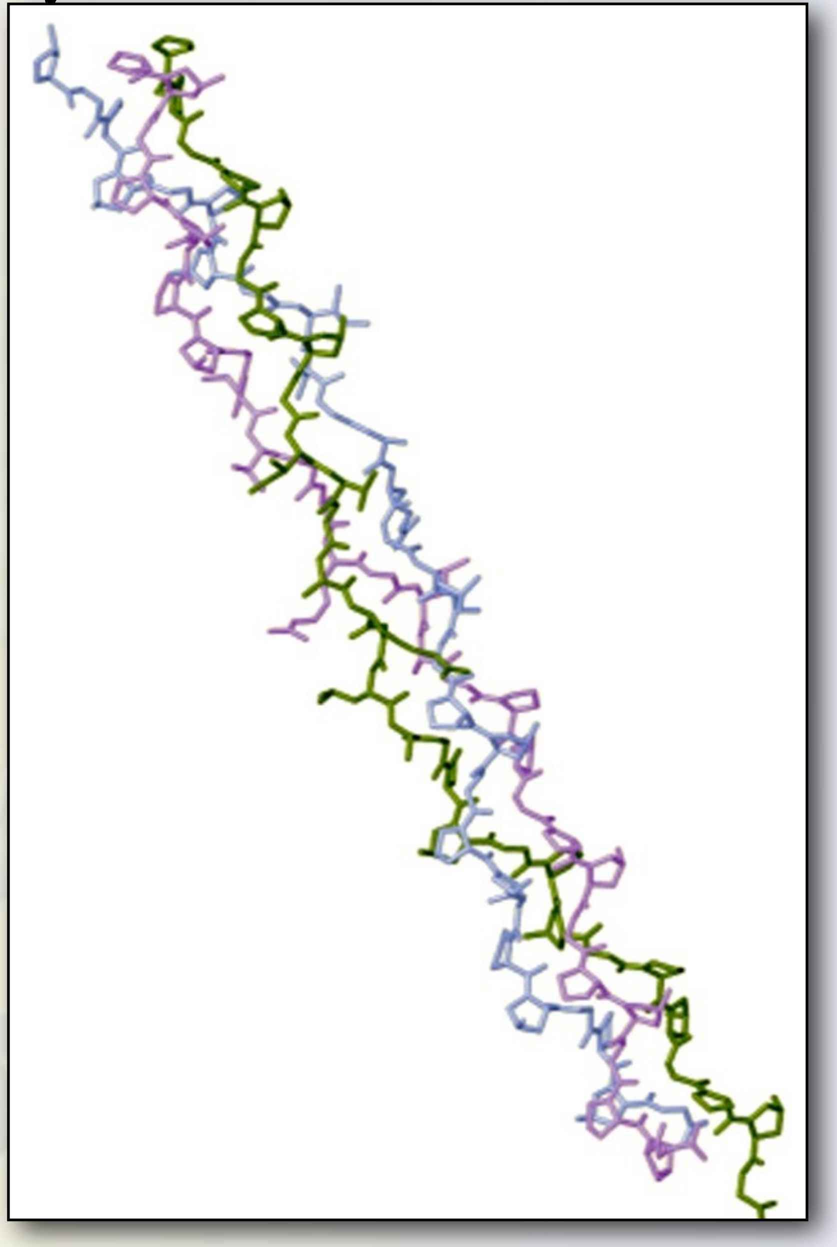
- Fibrous proteins have primary, secondary and quaternary structure.
  - The lack tertiary structure



collagen is made of polyproline-type triple helices

Fibrous Proteins

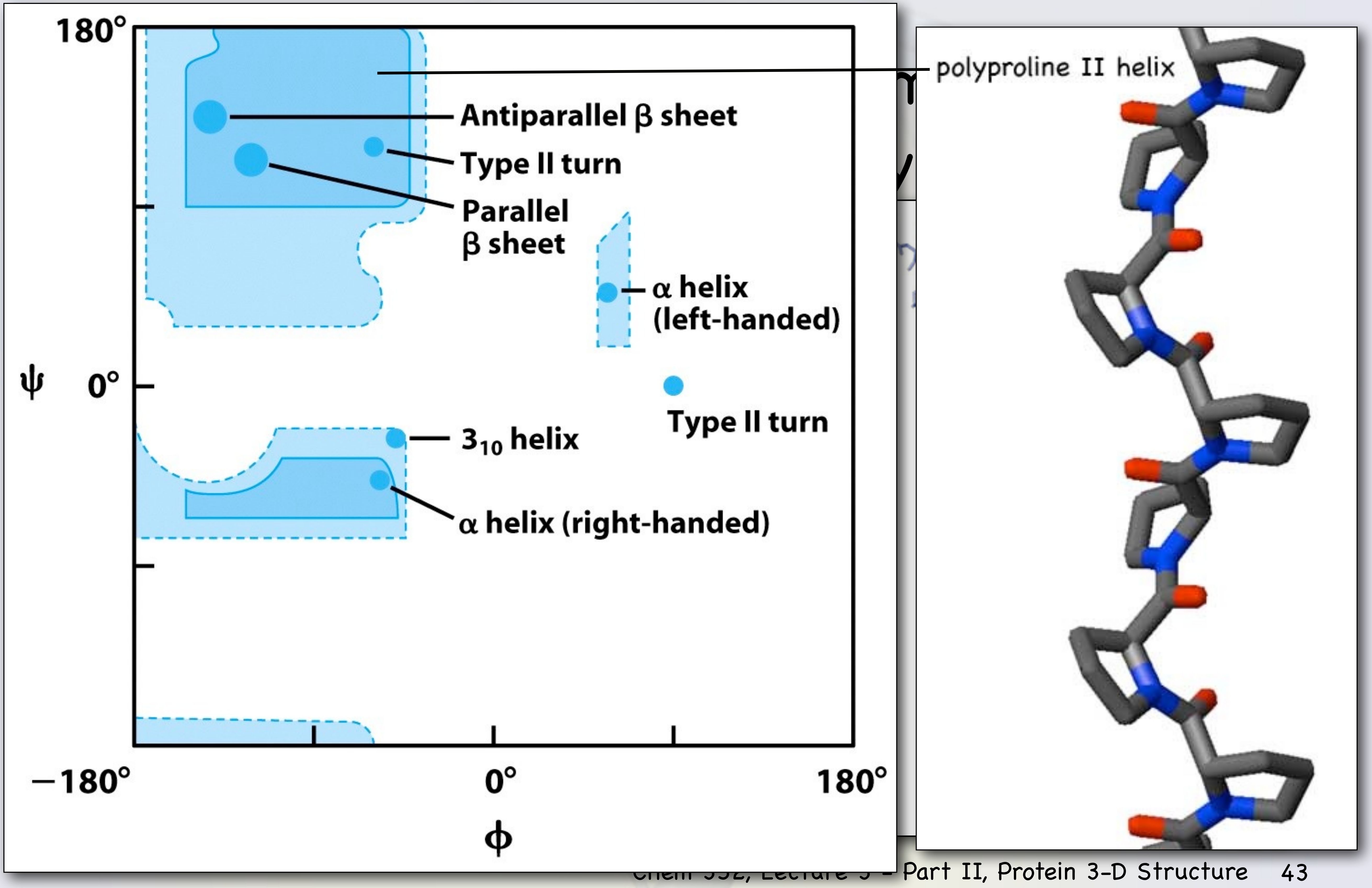
- Fibrous proteins have primary, secondary and quaternary structure.
  - The lack tertiary structure



collagen is made of polyproline-type triple helices

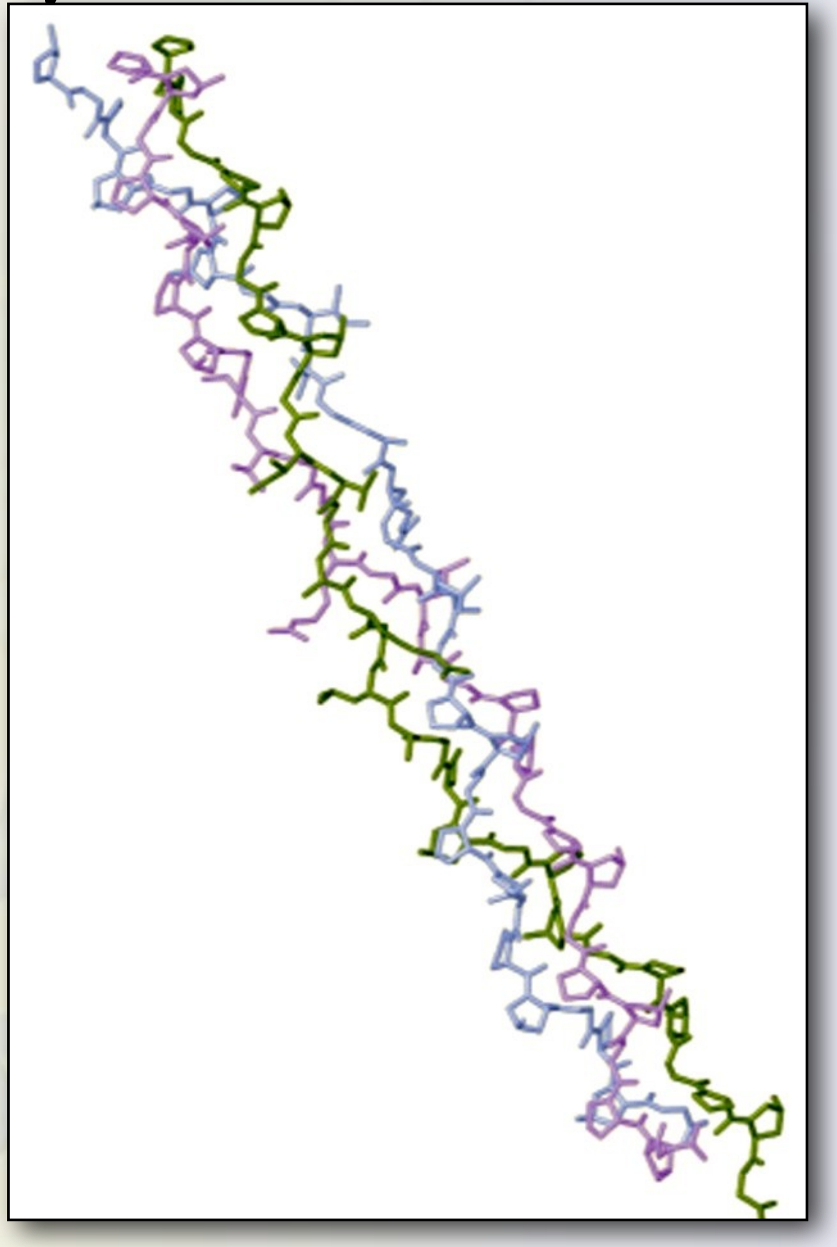
-Gly-Pro-Hydroxyproline-

Fibrous Proteins



Fibrous Proteins

- Fibrous proteins have primary, secondary and quaternary structure.
  - ✦ The lack tertiary structure



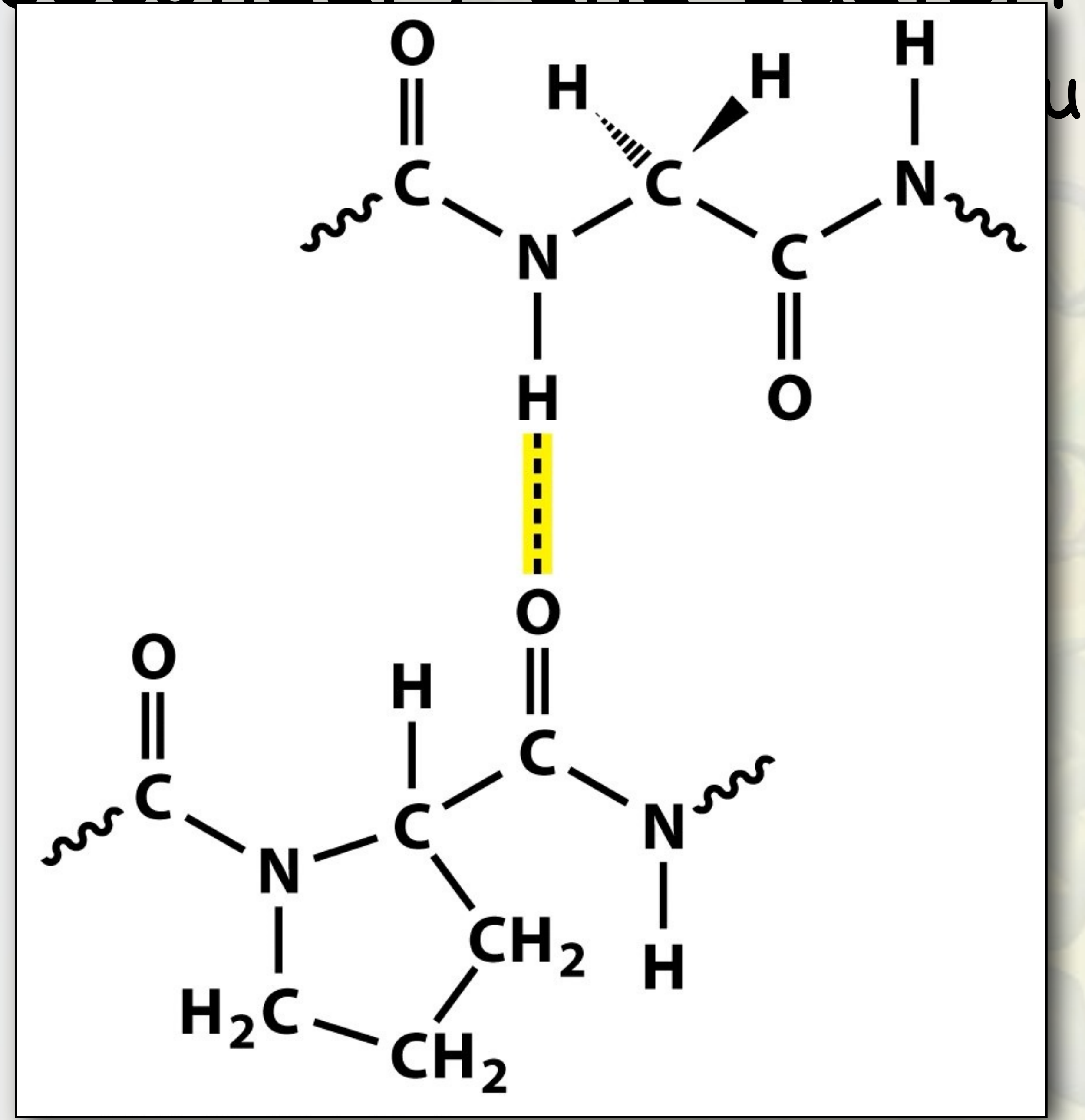
collagen is made of polyproline-type triple helices

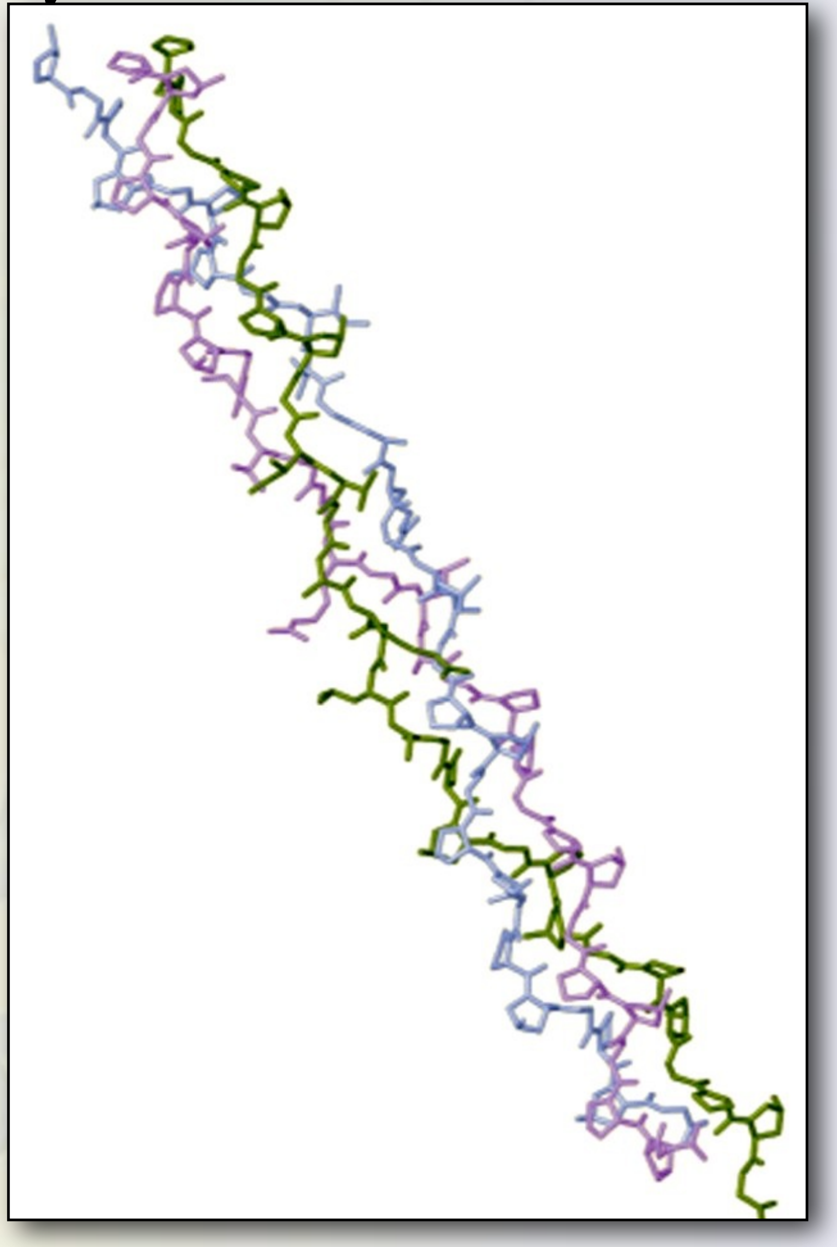
-Gly-Pro-Hydroxyproline-

Fibrous Proteins

- Fibrous proteins have primary, secondary and quaternary structure.

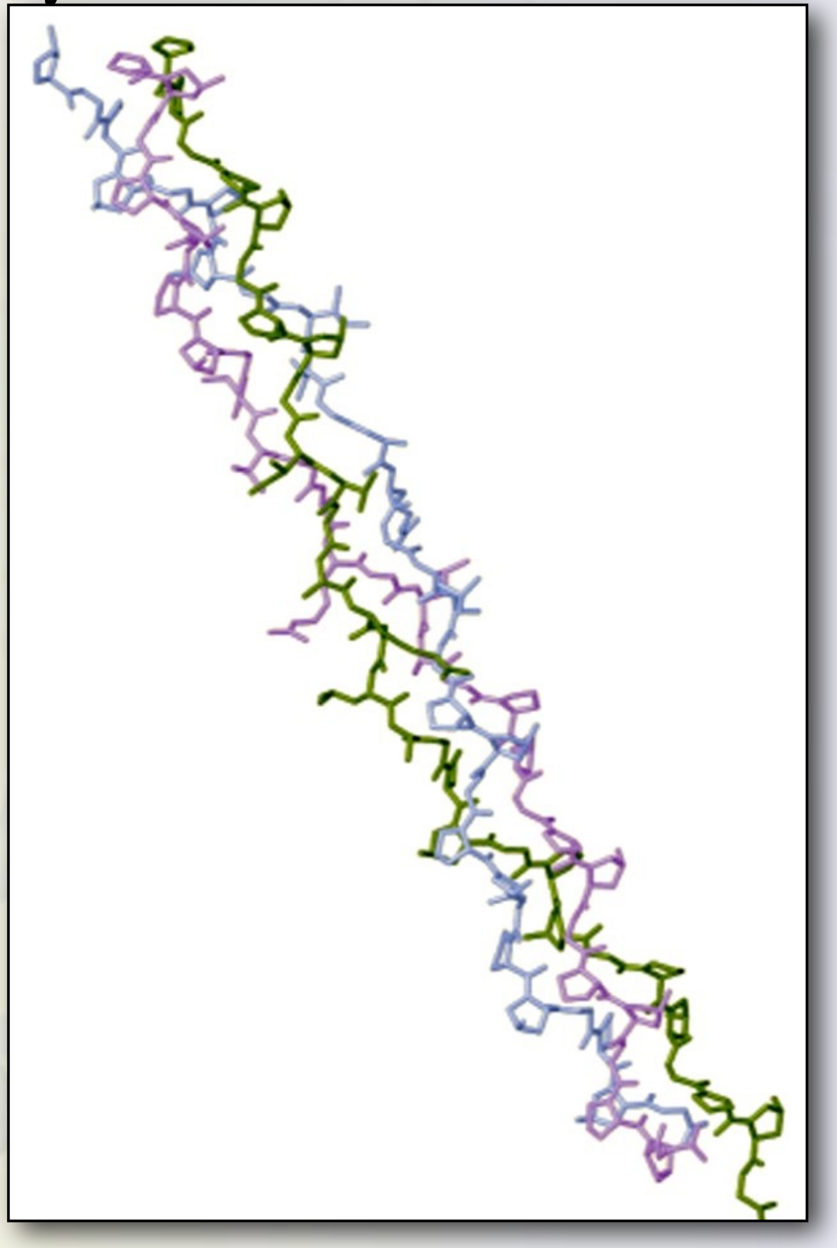
ure





Fibrous Proteins

- Fibrous proteins have primary, secondary and quaternary structure.
  - The lack tertiary structure



collagen is made of
polyproline-type triple helices

-Gly-Pro-Hydroxyproline-

Up next

- Exam I - Lecture 1-3

- Lecture 4, Part I - Enzymes
  - Read Chapter 5 in Moran et al.