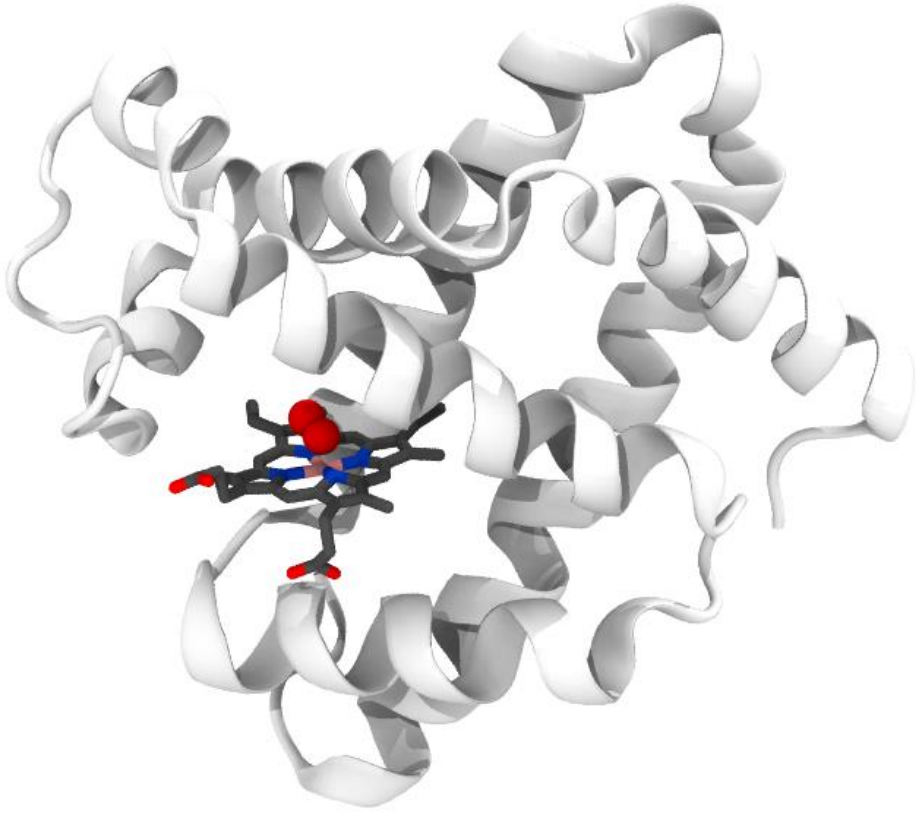
Molecular Biophysics

Protein assembly, function and malfunction

Summary of last lessons

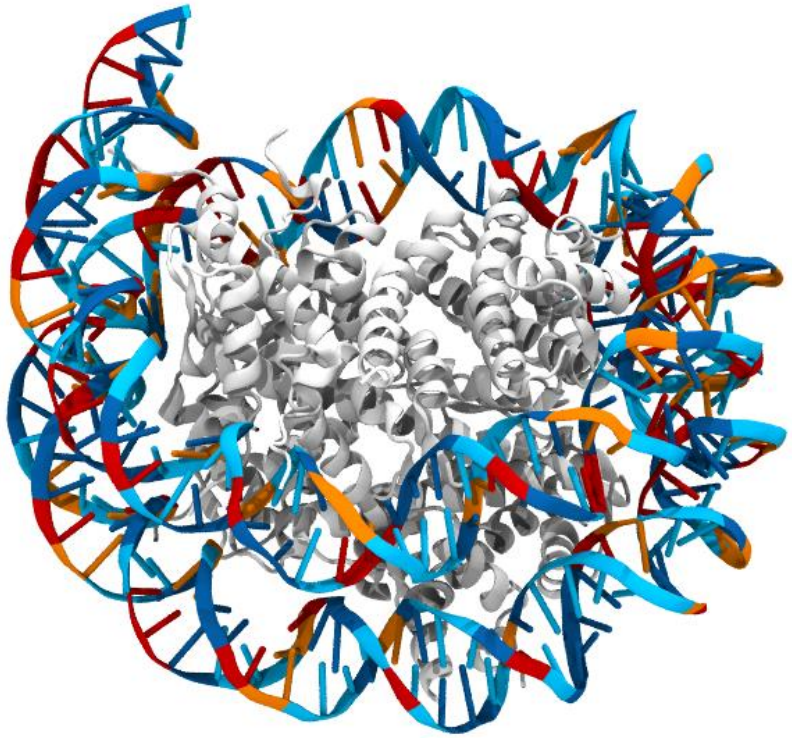
- Proteins are amino acid polymers that fold in a 3D structure
- amino acid sequence + environment = conformational space
- *The structure determines the function*: proteins interact with specific binding partners
- Means to study protein structure: X-ray crystallography, Nuclear Magnetic Resonance (NMR), Electron Microscopy, Modelling
- *Our view of the proteome is biased by the techniques we use to observe it*

Proteins interact with specific binding partners



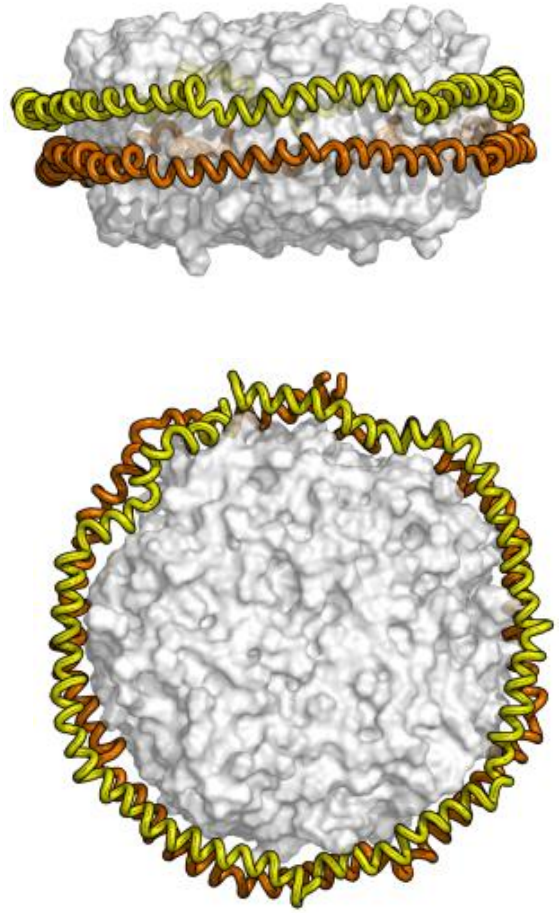
Myoglobin (PDB: 1MBO)

J.C. Kendrew et al., *A three-dimensional Model of the Myoglobin Molecule obtained by X-Ray Analysis,* Nature, 1958



Nucleosome (PDB: 5CPI)

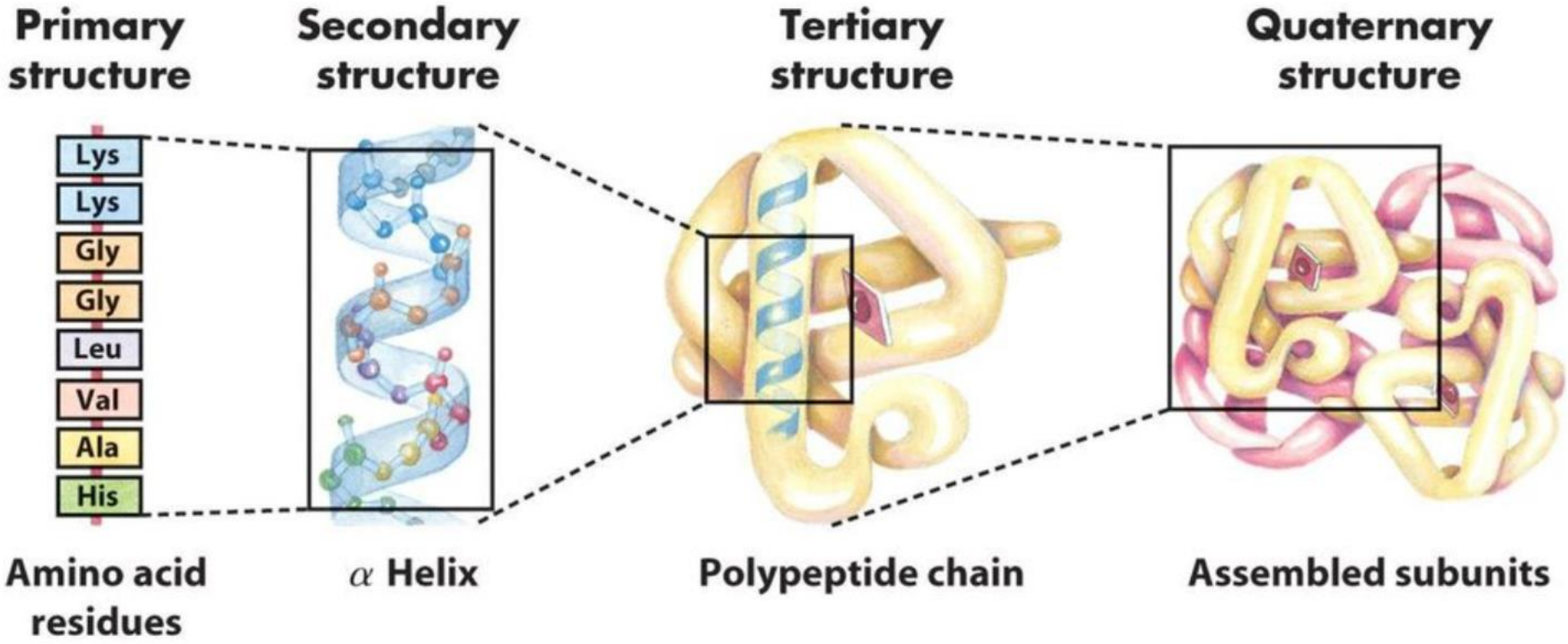
K. Luger et al., *Crystal Structure of the nucleosome core particle at 2.8 A resolution,* Nature, 1997



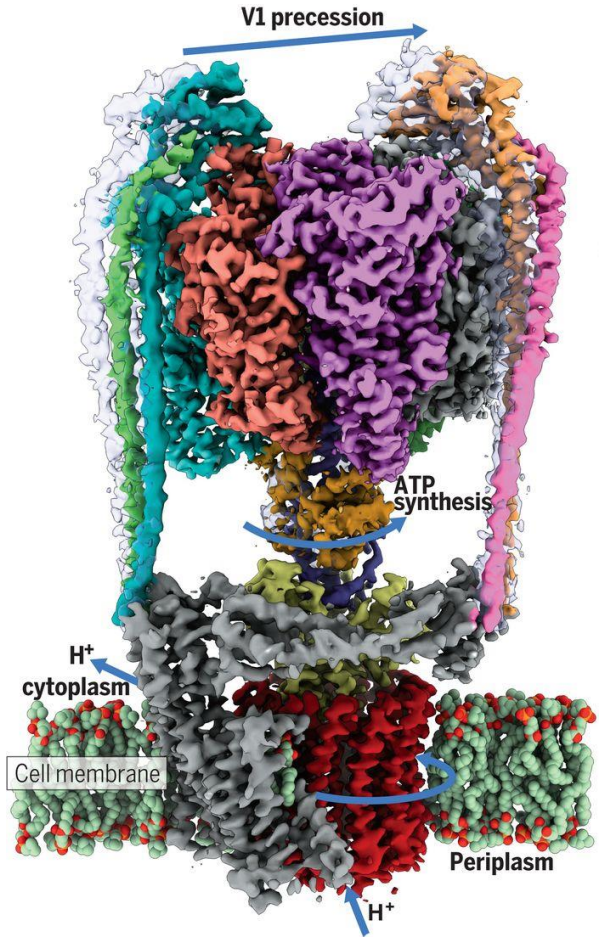
Lipoprotein (PDB: 1AV1)

D.W. Bohrani et al., *Crystal structure of truncated human apolipoprotein A-I suggests a lipid-bound conformation*, PNAS, 1997

Protein quaternary structure: assembly

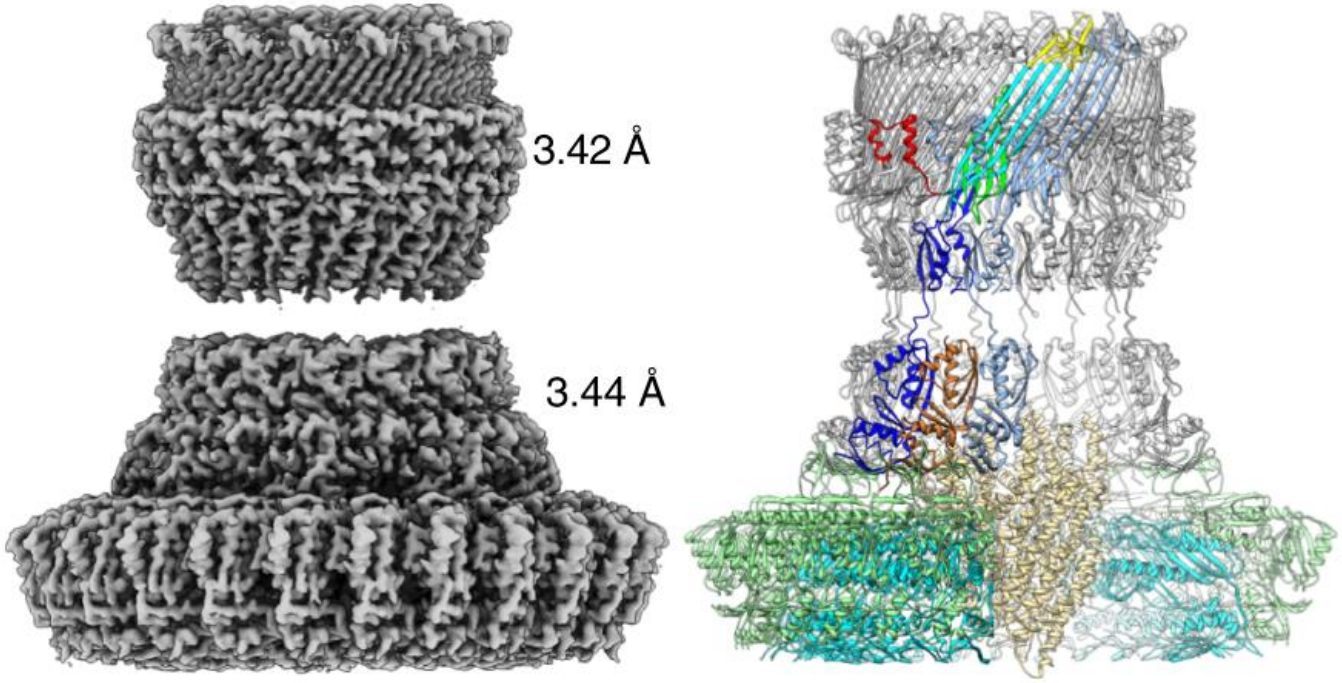


Protein quaternary structure: assembly



ATPAse

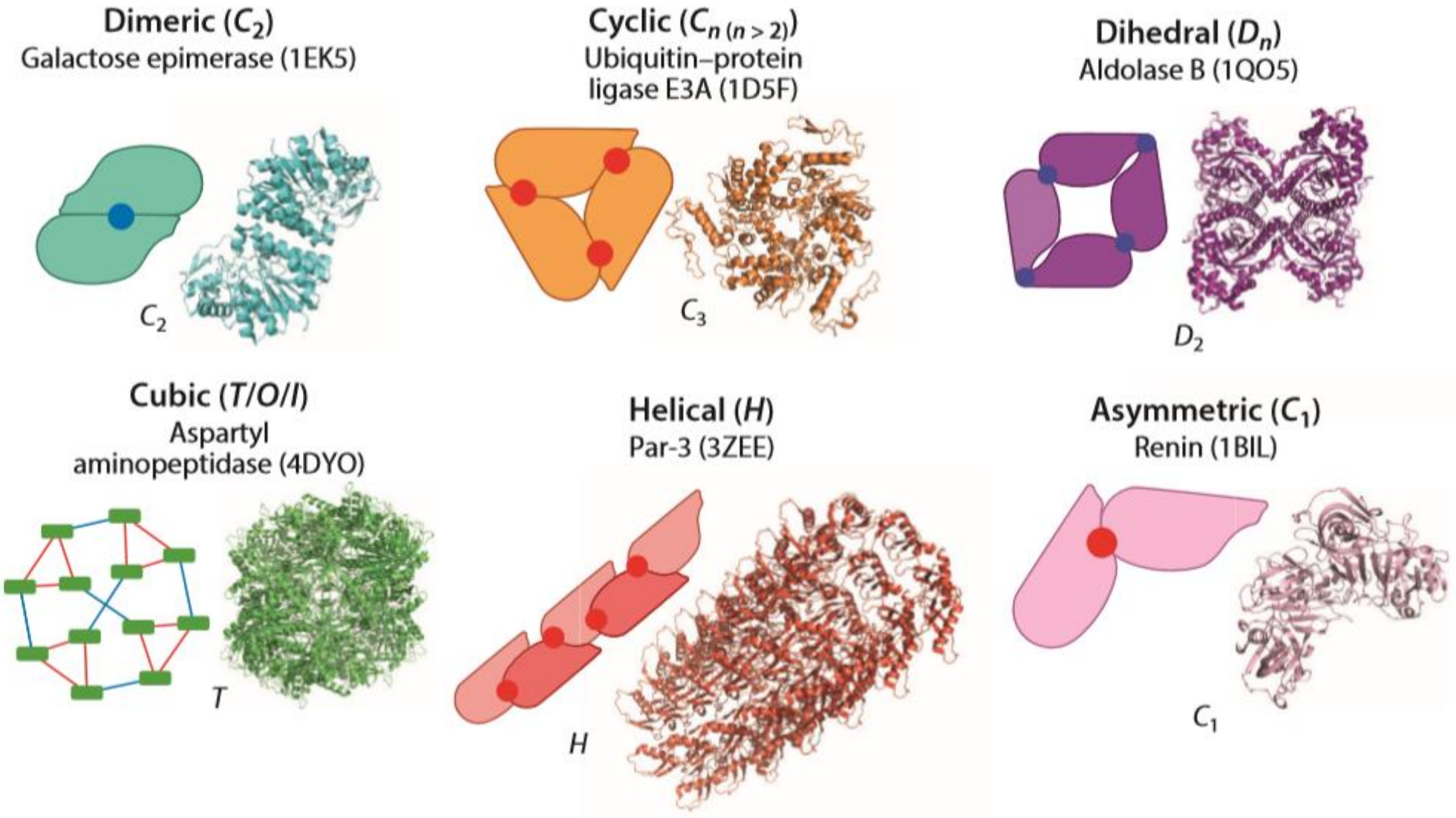
L. Zhou and L. A. Sazanov, *Structure and conformational plasticity of the intact Thermus thermophilus V/A-type ATPase, Science, 2019*



Injectisome
(evolutionarily related to the flagellum)

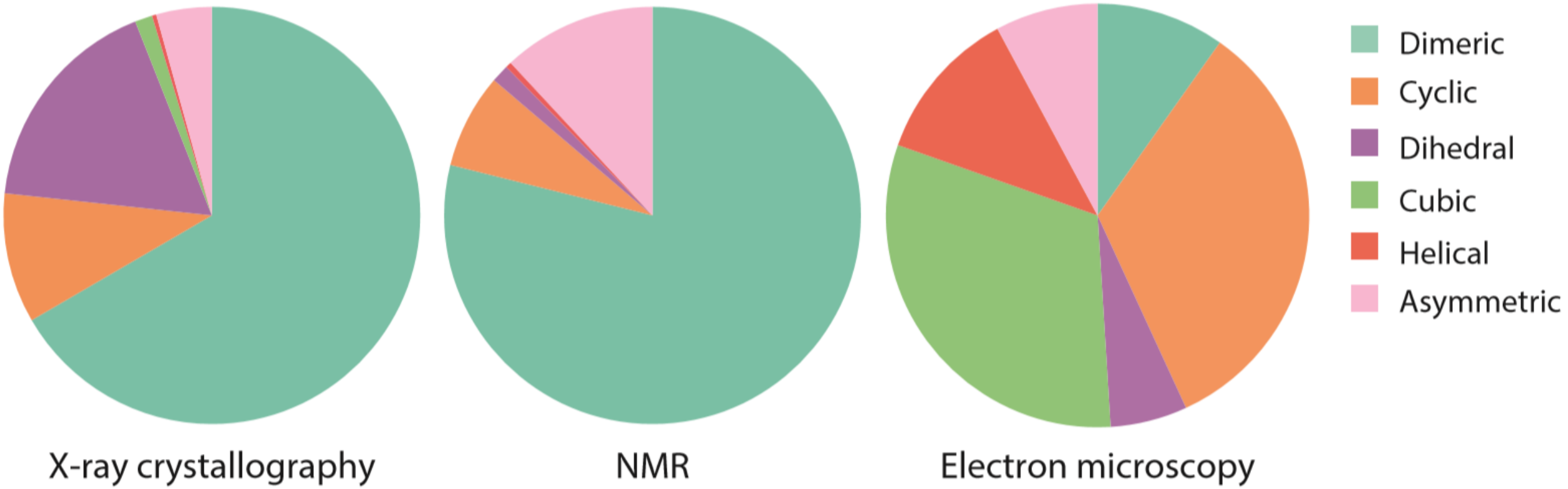
J.Hu et al., *Cryo-EM analysis of the T3S injectisome reveals the structure of the needle and open secretin,* Nature comms, 2018

Assembly stoichiometry and symmetry



J.A. Marsh and S.A. Teichmann, *Structure, Dynamics, Assembly and Evolution of Protein Complexes,* Annu. Rev. Biochem., 2015

Assembly stoichiometry and symmetry

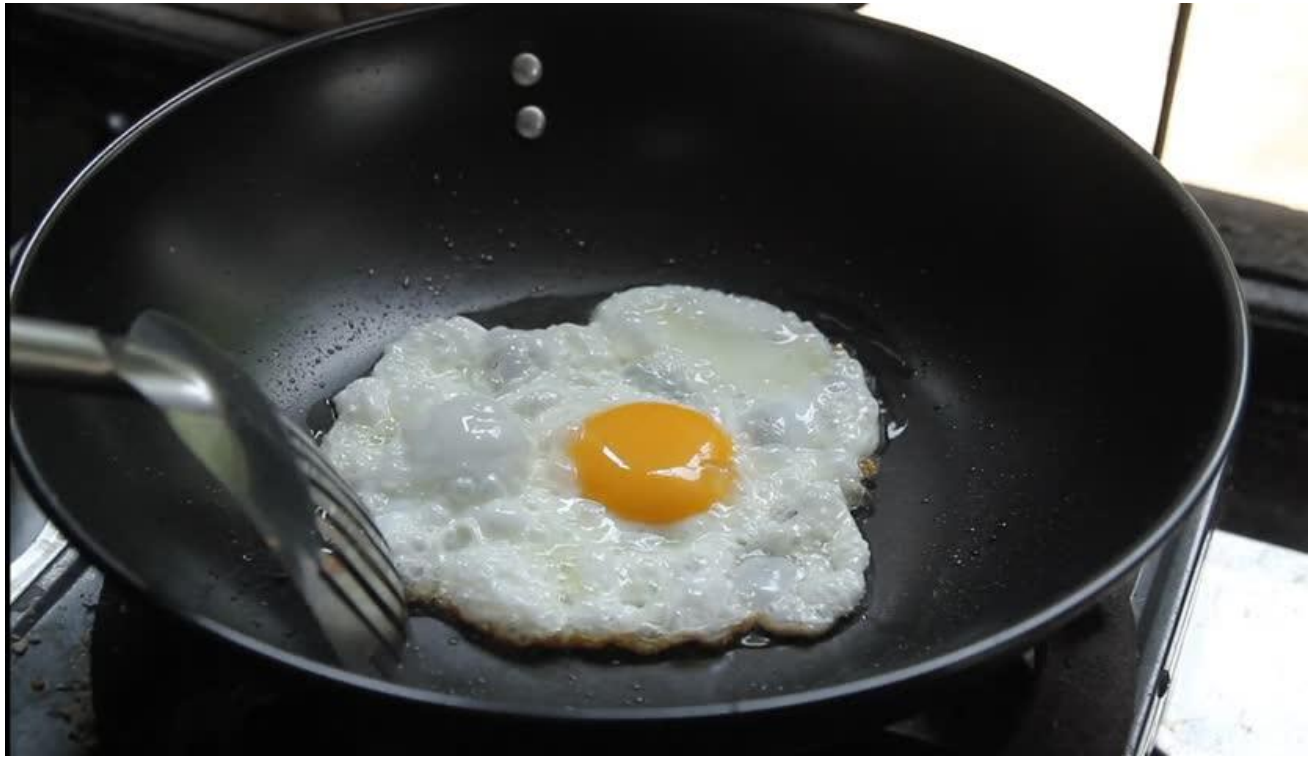


J.A. Marsh and S.A. Teichmann, *Structure, Dynamics, Assembly and Evolution of Protein Complexes,* Annu. Rev. Biochem., 2015

Small heat-shock protein (HSP): countering protein unfolding

Denaturation: protein unfolding caused by its exposition to non-native conditions, e.g.:

- heat
- pH
- reducing/oxydising agents

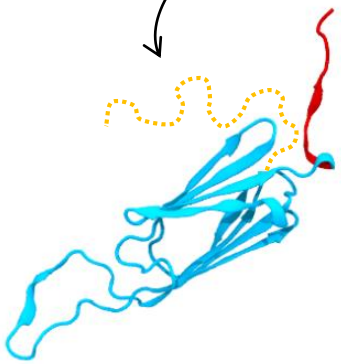


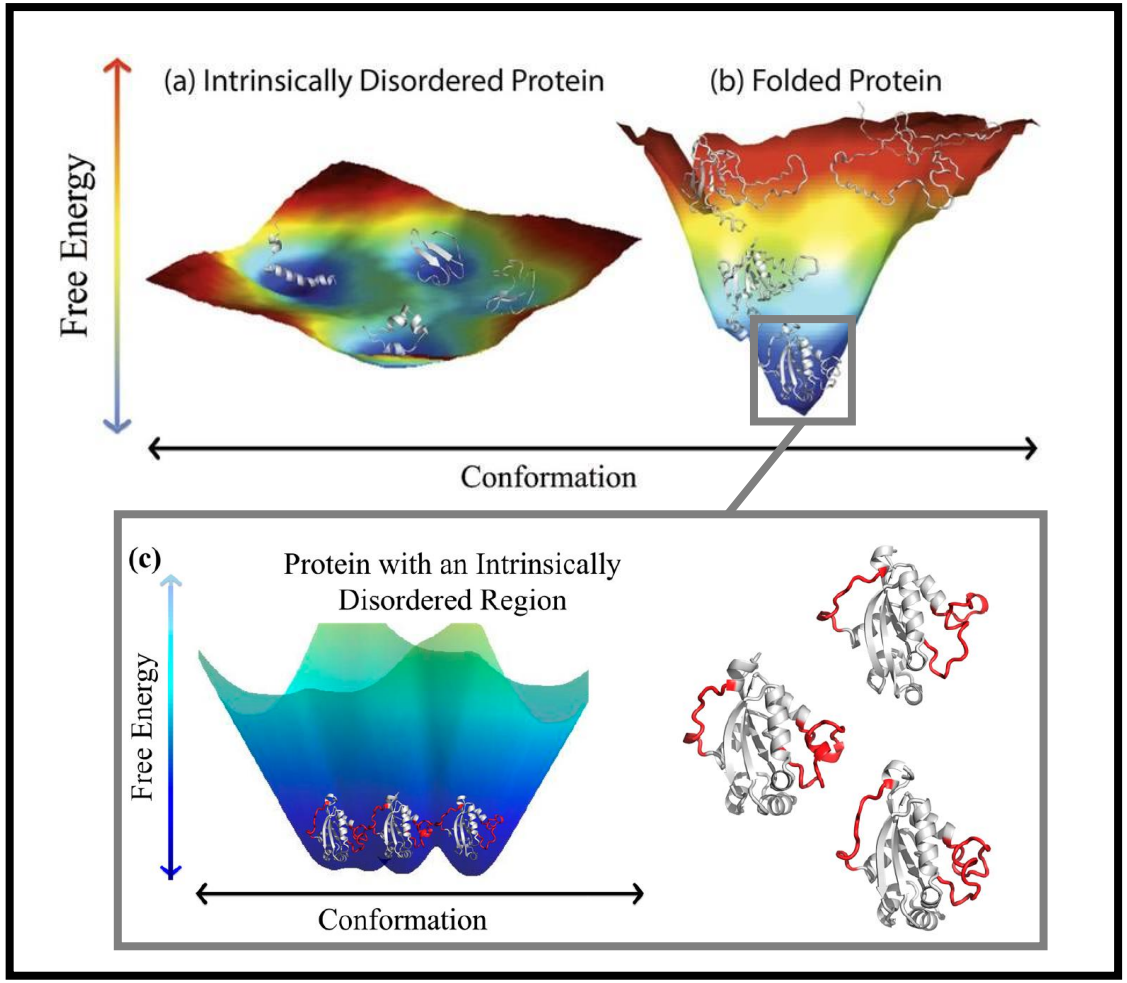
Protein denaturation is usually harmful for an organism

small heat-shock proteins (HSP) – self assembly

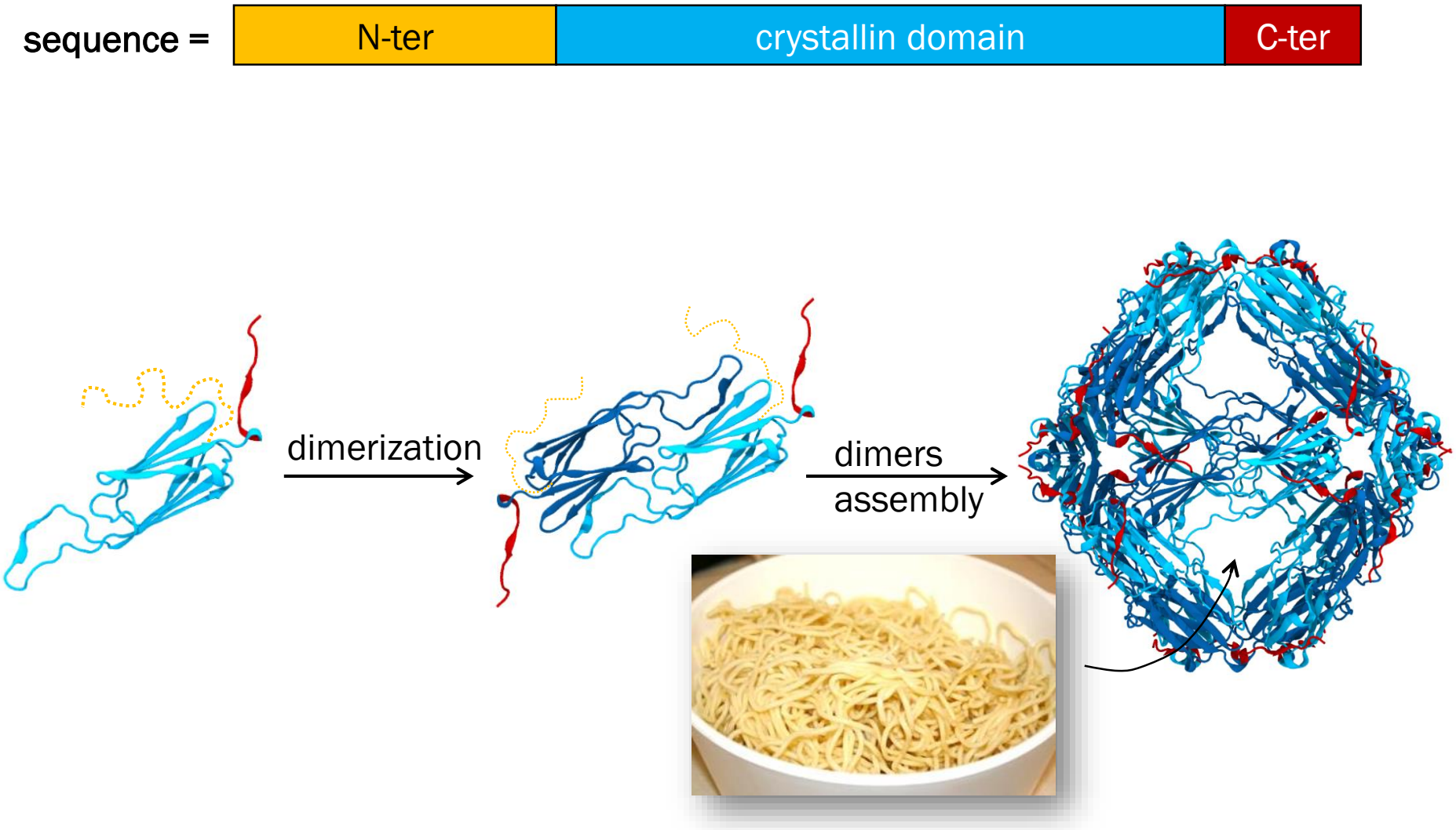
sequence = N-ter | crystallin domain | C-ter

N-ter region is disordered



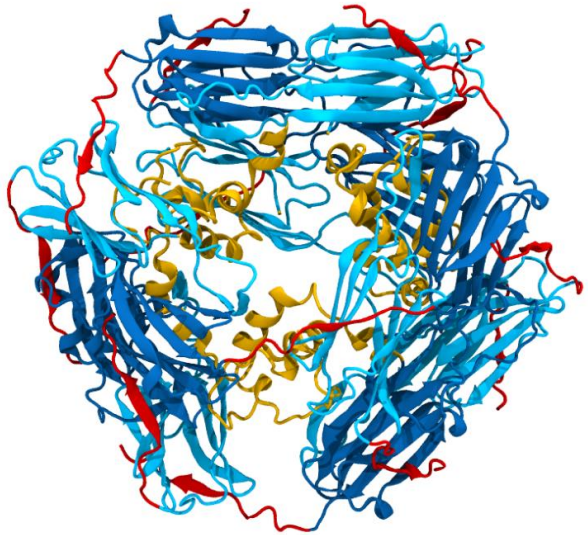


small heat-shock proteins (HSP) – self assembly

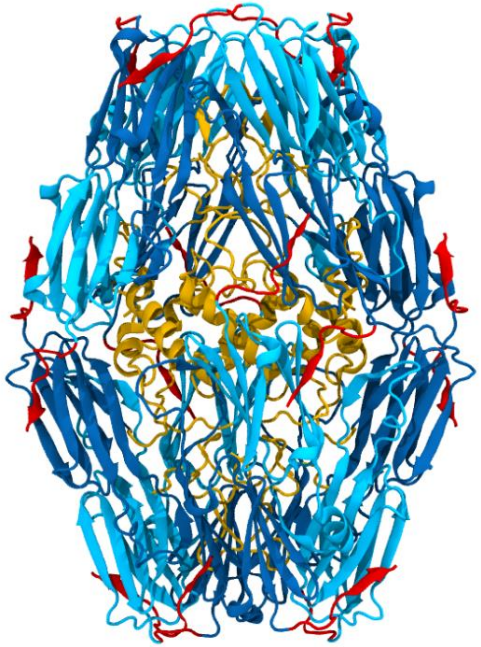


small heat-shock proteins (HSP) – self assembly

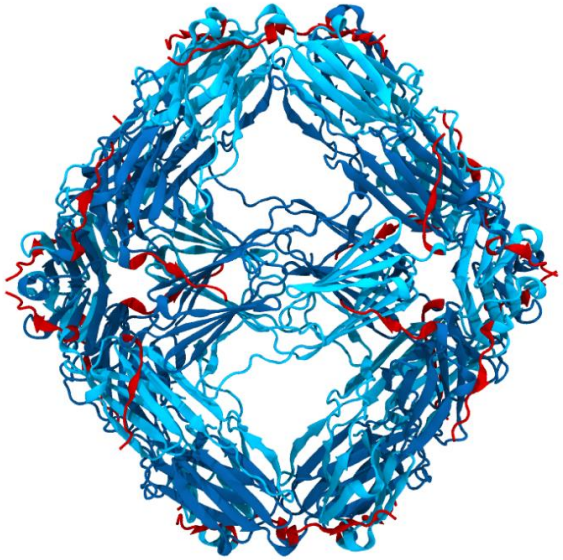
sHSP 16.9



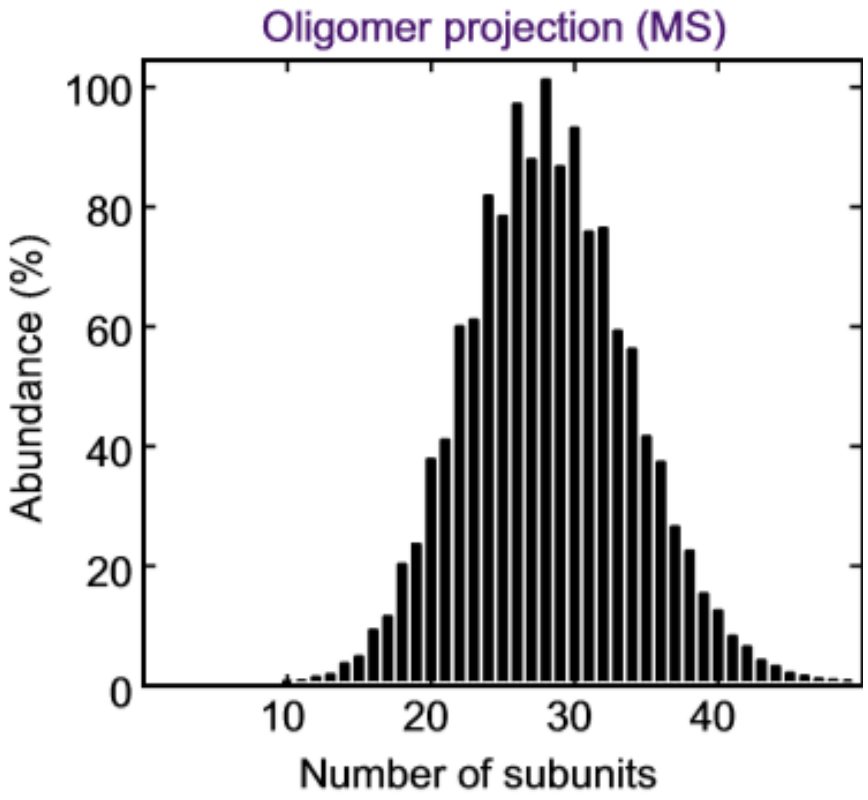
sHSP 16.0

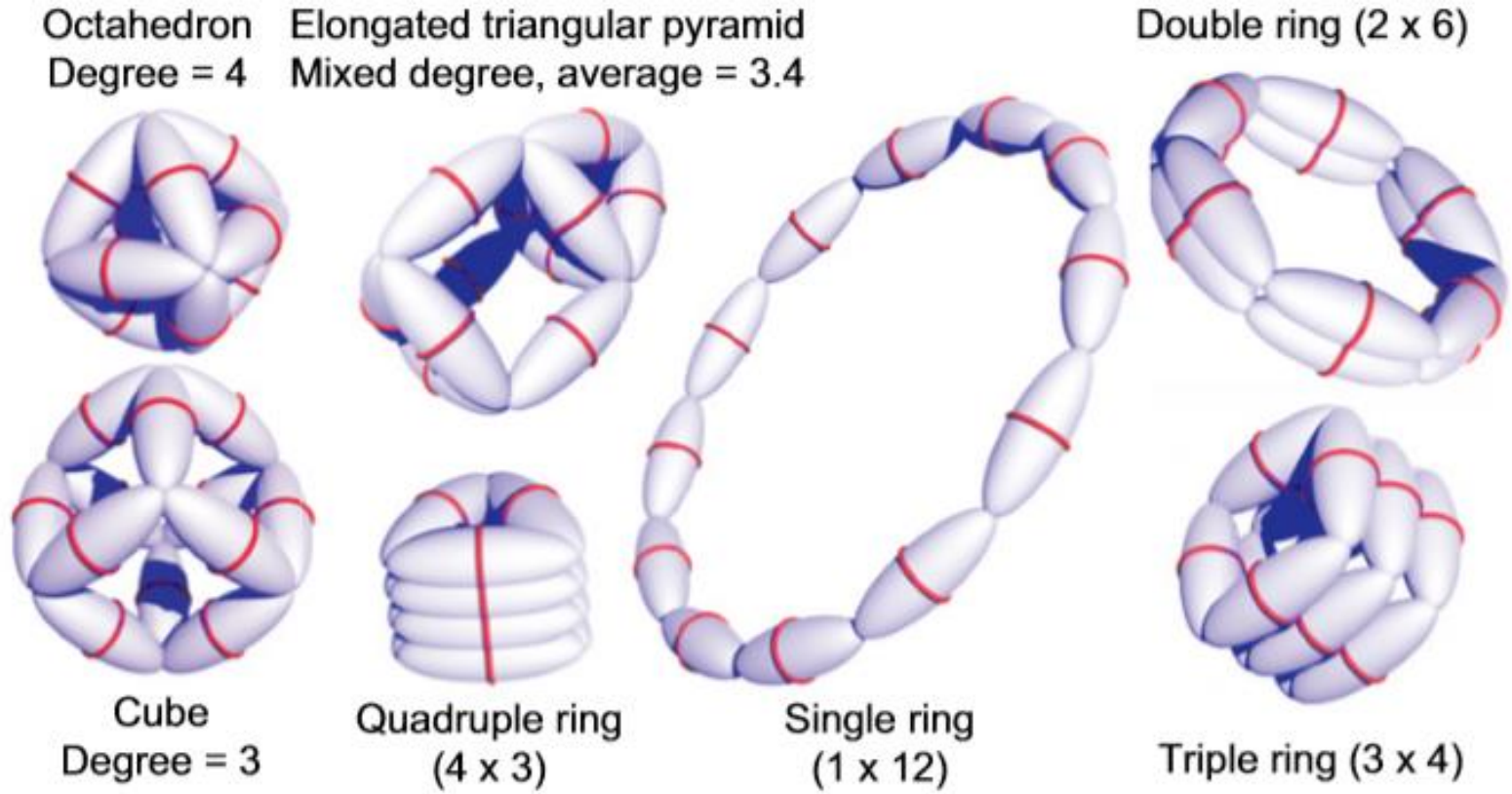


sHSP 16.5



small heat-shock proteins (HSP) – self assembly



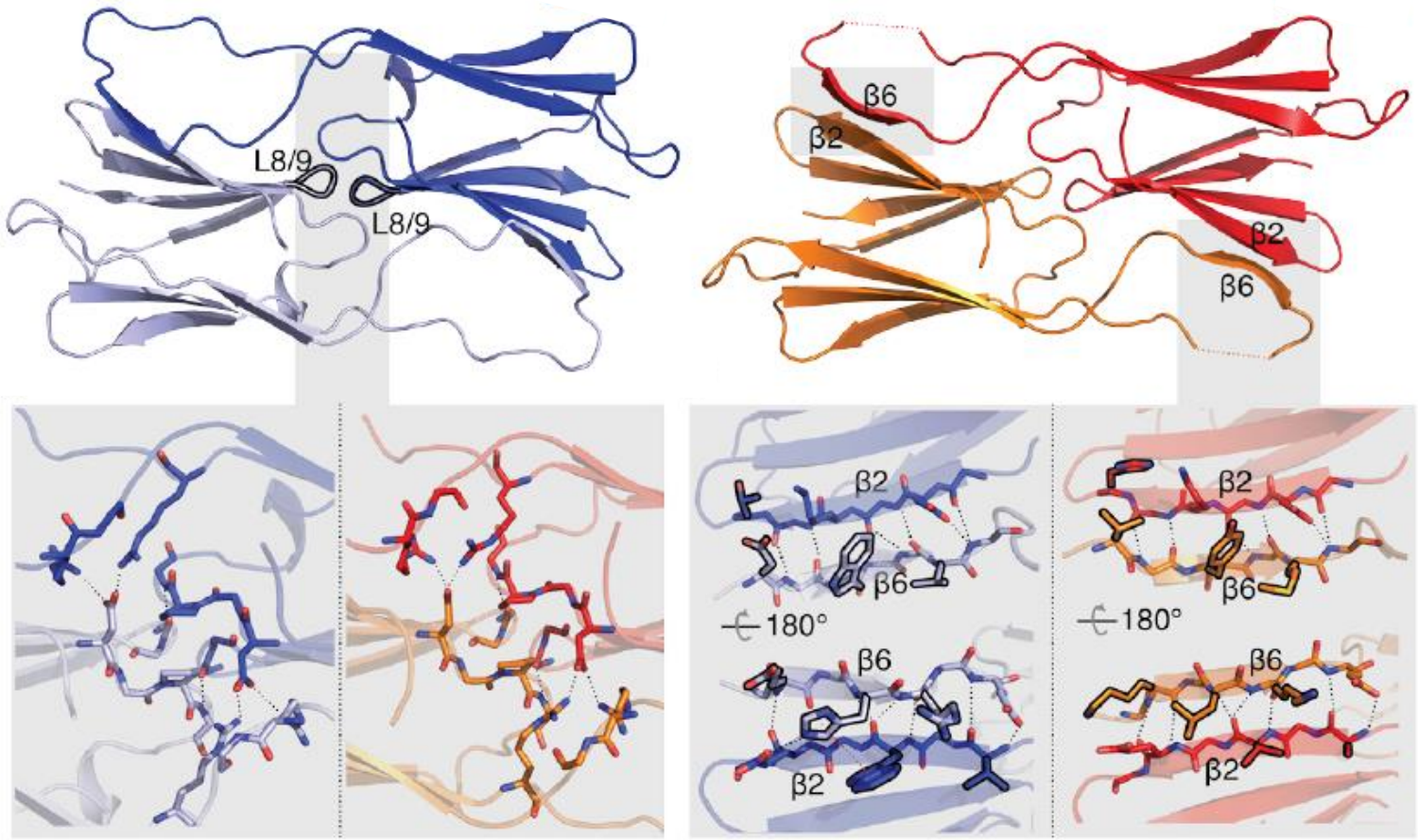


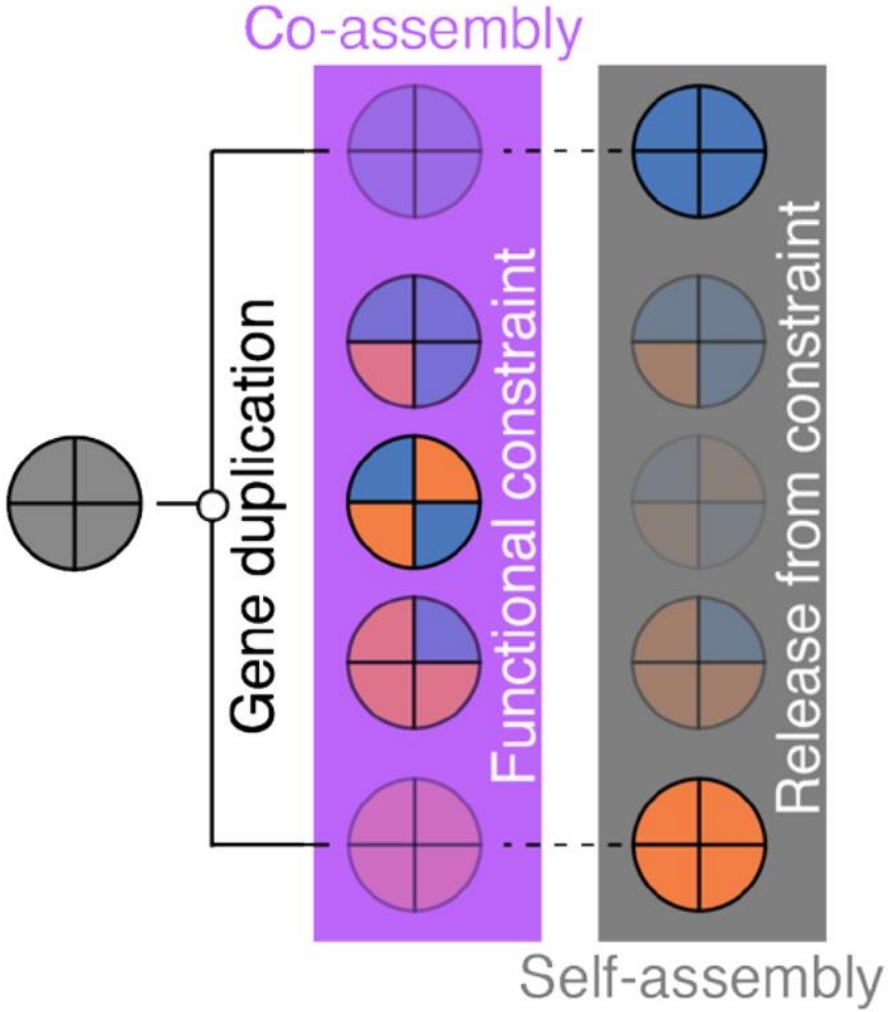
A.J. Baldwin at al., *The Polydispersity of αB-Crystallin Is Rationalized by an Interconverting Polyhedral Architecture*, Structure, 2011

I. Santhanagopalan et al, *It takes a dimer to tango: Oligomeric small heat shock proteins dissociate to capture substrate*, JBC, 2018

HSP: similars don’t co-assemble

- High sequence homology, near-identical dimers interfaces (1.2 Å backbone RMSD)

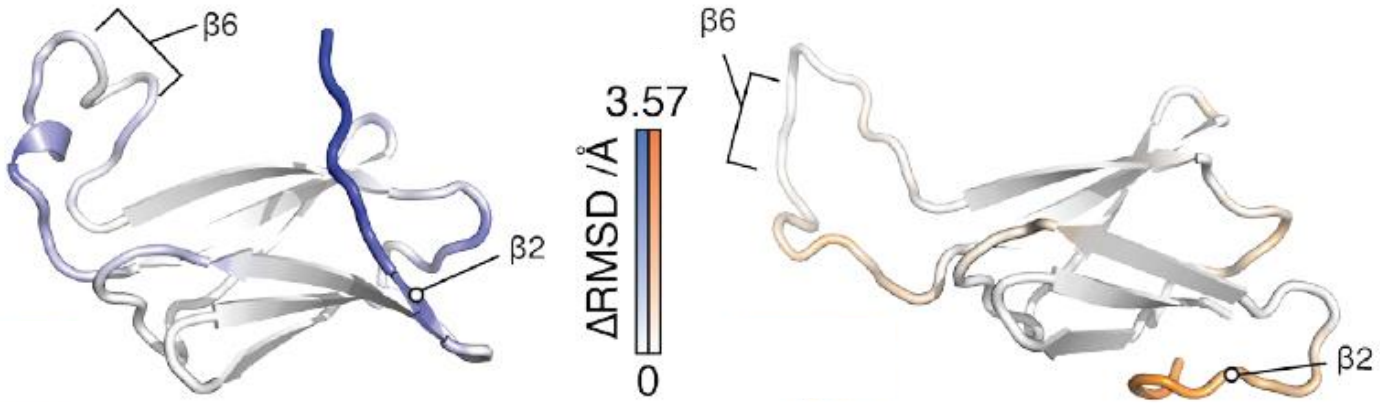


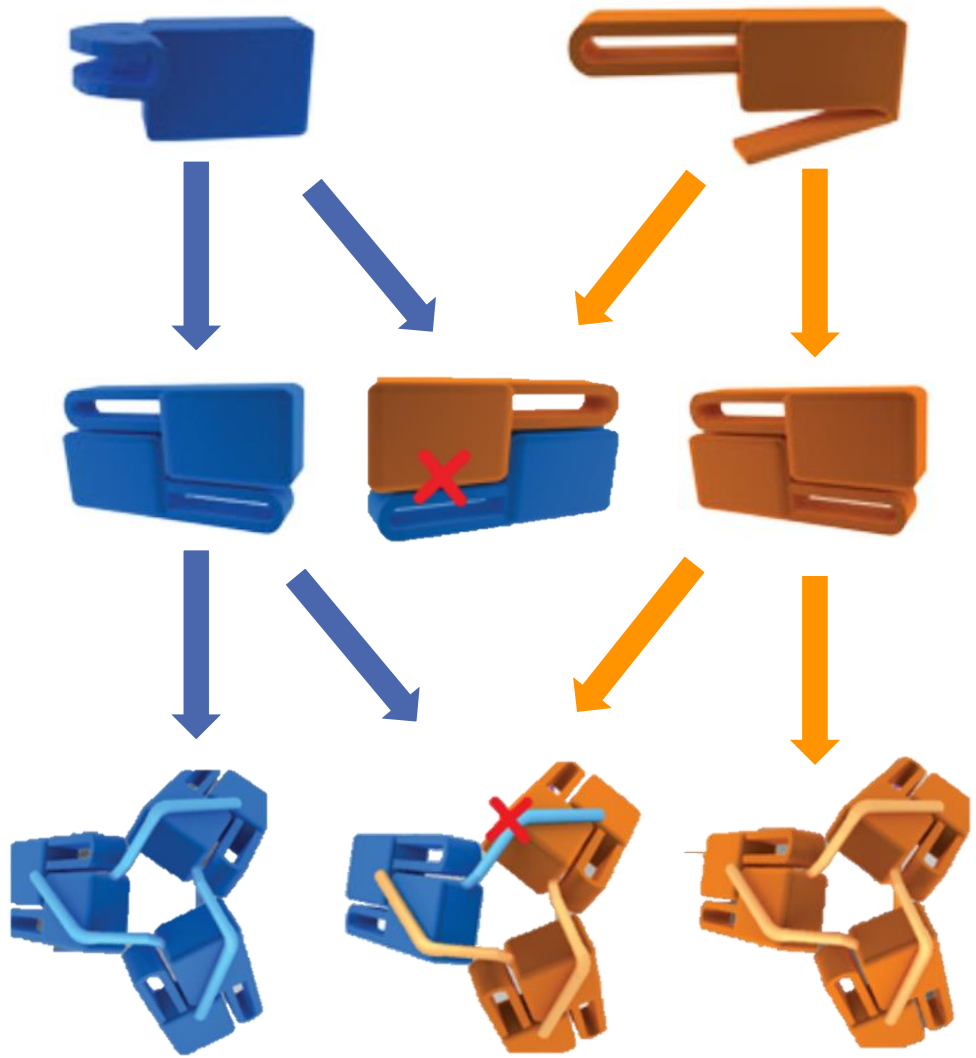


G.K. Hochberg et al. *Structural principles that enable oligomeric small heat-shock protein paralogs to evolve distinct functions.* Science, 2018

HSP: selectivity is mediated by dynamics

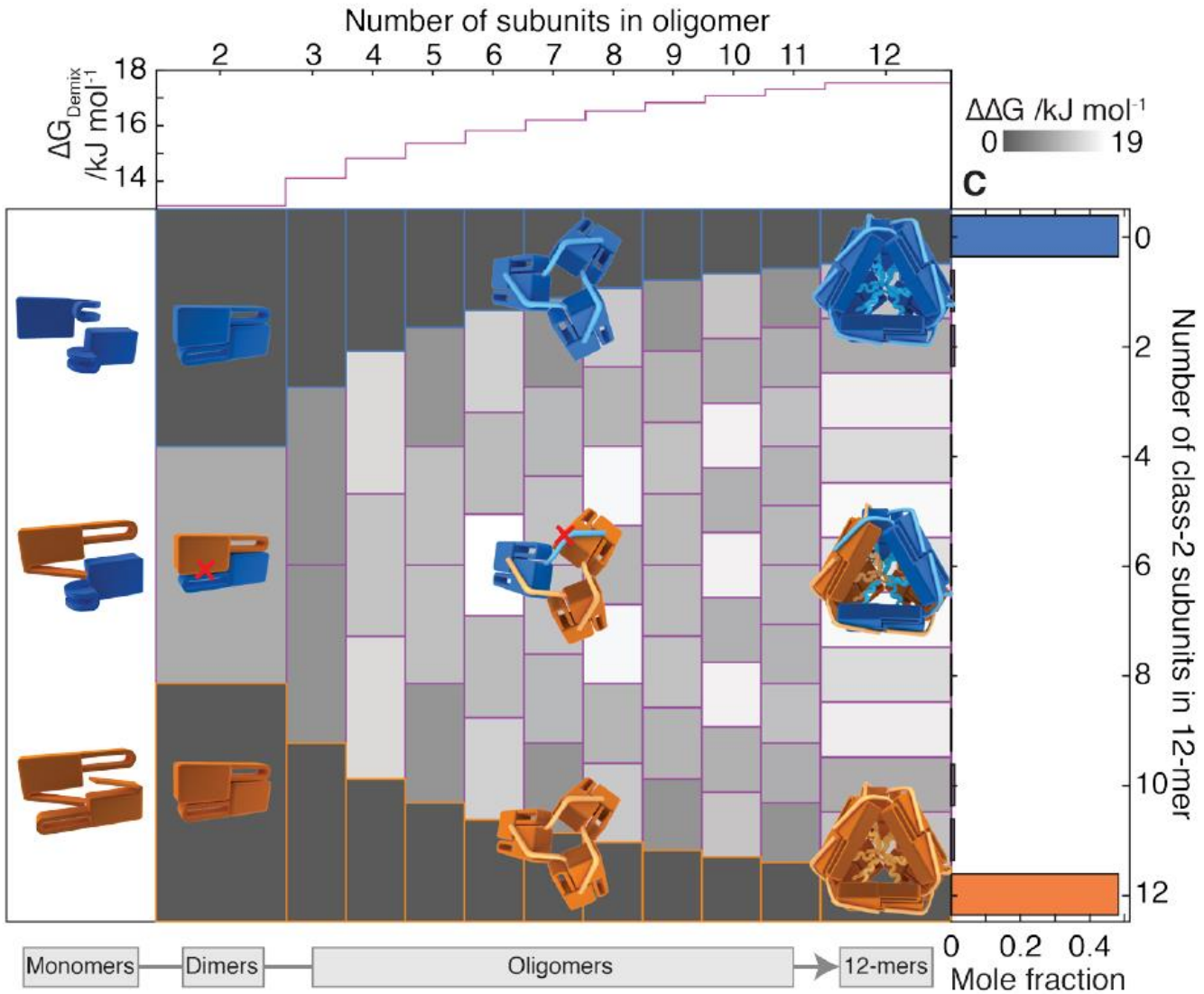
- High sequence homology, near-identical dimers interfaces (1.2 Å backbone RMSD)
- Monomeric subunits feature different disordered regions

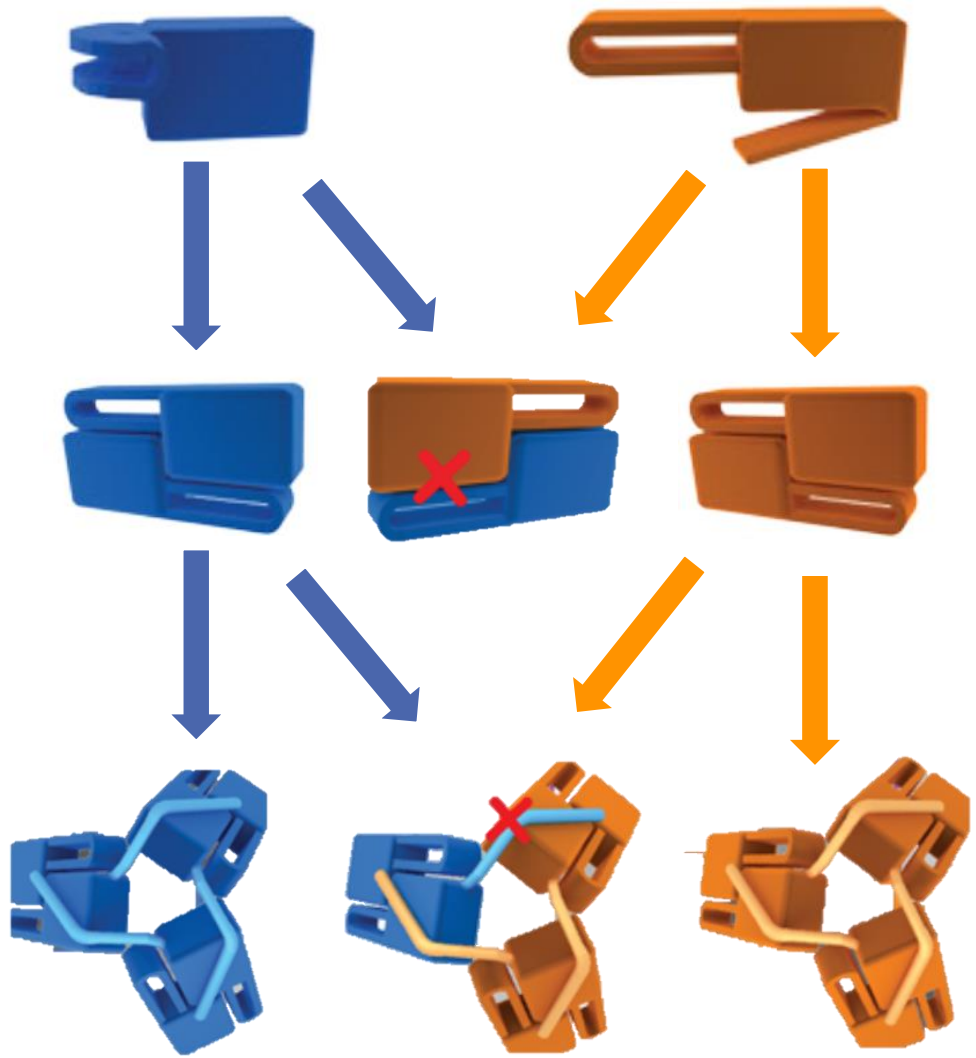




G.K. Hochberg et al. *Structural principles that enable oligomeric small heat-shock protein paralogs to evolve distinct functions.* Science, 2018

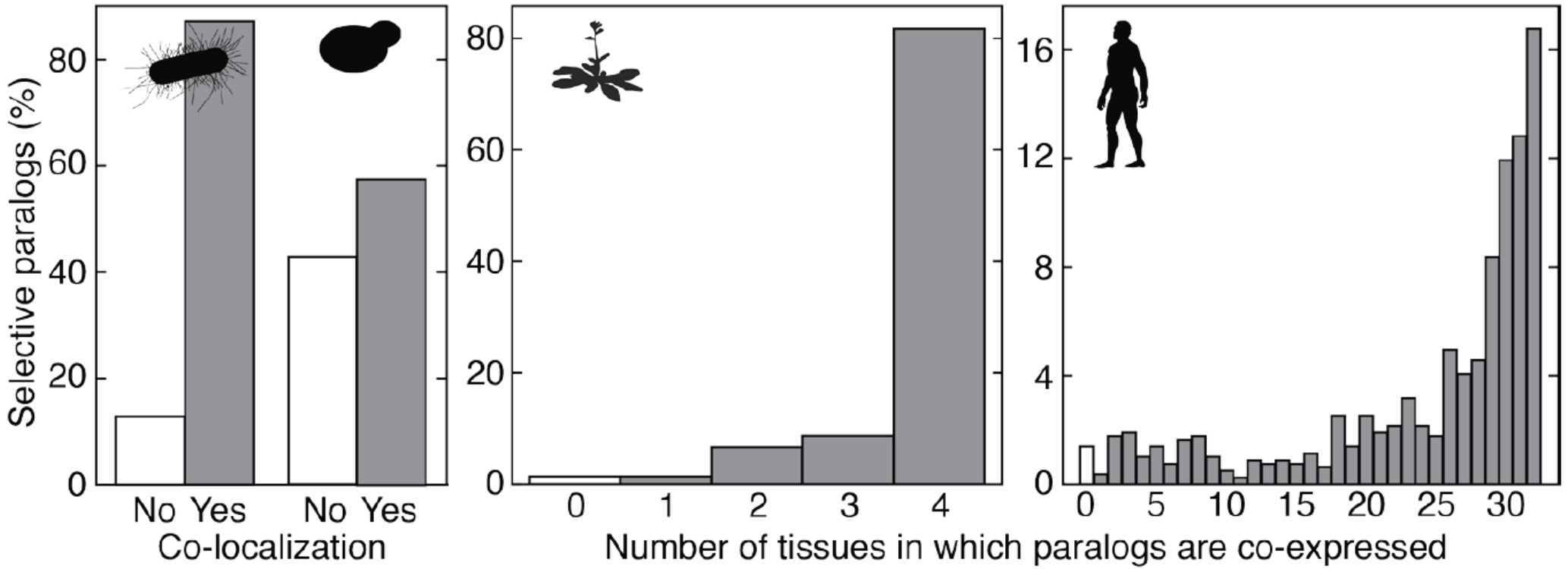
HSP: selectivity overcomes entropy





G.K. Hochberg et al. *Structural principles that enable oligomeric small heat-shock protein paralogs to evolve distinct functions.* Science, 2018

[Extra] Most paralogs co-localise but don’t co-assemble

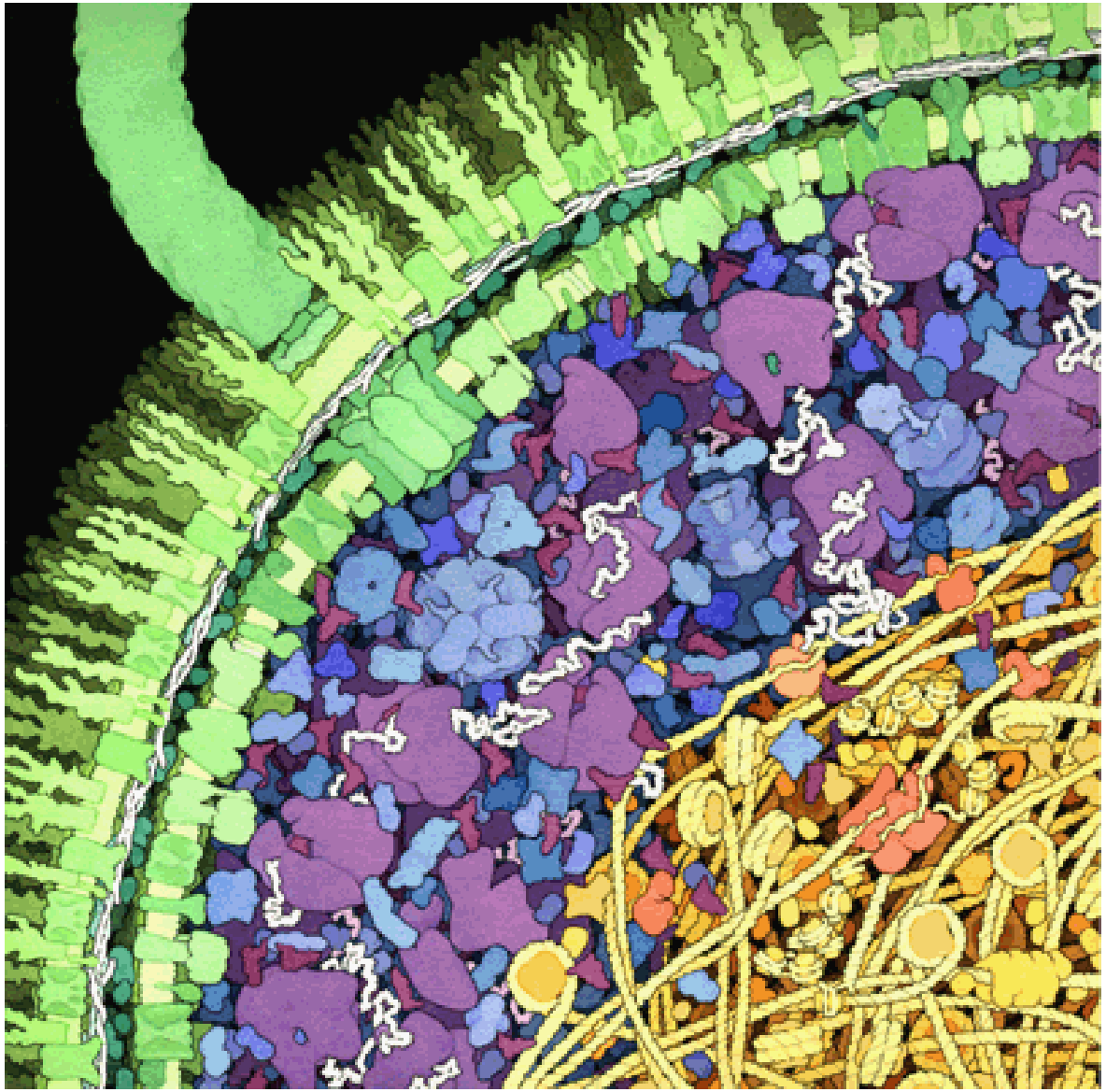


G.K. Hochberg et al. *Structural principles that enable oligomeric small heat-shock protein paralogs to evolve distinct functions.* Science, 2018

The intracellular space

A crowded environment, Brownian motion: proteins bump into each other all the time!

Most contacts are non-specific and short-lived



David G. Goodsell, *The machinery of Life*, Springer

reported proteins per cell	cell volume (μm^3)	proteins per volume ($10^6/\mu m^3$
M. pneumonia		
0.05×10^6	0.015	3
L. interrogans		
$1.0\text{-}1.2\times10^6$ *	0.22	5
E. coli		
2.36×10^6	0.86	2.7
B. subtilis		
2.3×10^6 *	1.13	2.0
1.3×10^6 *	0.62	2.1
1.8×10^6 *	0.85	2.1
S. aureus		
0.35×10^6 *	0.33	1.1
0.27×10^6 *	0.23	1.2
0.26×10^6 *	0.23	1.1
budding yeast (haploid)		
50×10^6	≈30-40	1-2
47×10^6 *	≈30-40	1-2
53×10^6	≈30-40	1-2
fission yeast		
60.3×10^6	≈100	0.6
M. musculus **(NIH3T3 cells)**		
3×10^9 *	≈2000	1.5
H. Sapiens **(U2OS)**		
$0.95\text{-}1.7\times10^9$ *	≈4000	0.2-0.4
H. sapiens **(HeLa)**		
2.0×10^9 *	≈2000	1

book.bionumbers.org/how-many-proteins-are-in-a-cell

Protein density

- Estimate of proteins per unit of volume N_V

$$N_V = \frac{N}{V} = \frac{m_{all}}{m_{prot}} N_A \frac{1}{V} = \frac{N_A}{n_{aa} m_{aa}} \frac{m_{all}}{V}$$

m_{all}: total protein mass in cell (g)
m_{prot}: avg. protein molar mass (g/mole)
N_A : Avogadro number
V: cell volume

n_{aa}: average #aminoacids per protein ≈ **300**
m_{aa}: average mass of one aminoacid ≈ **100 Da**

- Typical protein mass inside cells (m_{all}/V): 200-300 g/L. *~20-30% of total volume*!

molecule	measured context	diffusion coefficient ($\mu m^2/s$)
H_2O	water	2000
H_2O	nucleus of chicken erythrocyte	200
H^+ (from H_3O^+ to H_2O)	water	7000
O_2	water	2000
CO_2	water	2000
tRNA (≈20 kDa)	water	100
protein (≈30 kDa GFP)	water	100
protein (≈30 kDa GFP)	eukaryotic cell (CHO) cytoplasm	30
protein (≈30 kDa GFP)	rat liver mitochondria	30
protein (NLS-EGFP)	cytoplasm of *D. melanogaster* embryo	20
protein (≈30 kDa)	*E. coli* cytoplasm	7-8
protein (≈40 kDa)	*E. coli* cytoplasm	2-4
protein (≈70-250 kDa)	*E. coli* cytoplasm	0.4-2
protein (≈140 kDa Tar-YFP)	*E. coli* membrane	0.2
protein (≈70 kDa LacY-YFP)	*E. coli* membrane	0.03
fluorescent dye (carboxy-fluorescein)	*A. thaliana* cell wall	30
fluorescent dye (carboxy-fluorescein)	*A. thaliana* mature root epidermis	3
transcription factor (LacI)	movement along DNA (1D, *in vitro*)	0.04 (4×10^5 bp^2s^{-1})
morphogen (bicoid-GFP)	cytoplasm of *D. melanogaster* embryo	7
morphogen (wingless)	wing imaginal disk of *D. melanogaster*	0.05
mRNA	HeLa nucleus	0.03-0.10
mRNA	various localizations and sizes	0.005-1
ribosome	*E. coli*	0.04

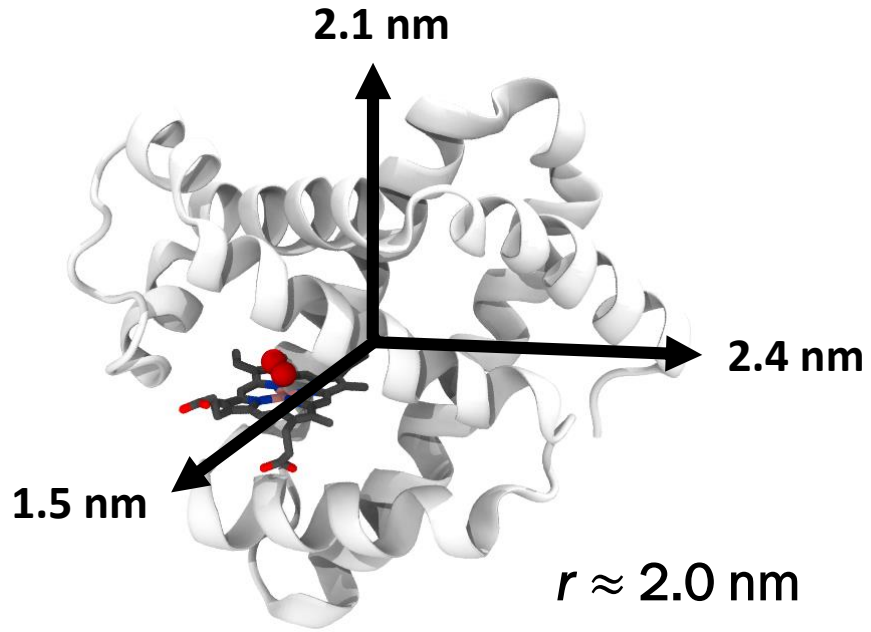
book.bionumbers.org/what-are-the-time-scales-for-diffusion-in-cells

Protein diffusion

Stokes-Einstein relation:

$$D = \frac{k_B T}{6\pi\eta r}$$

r: radius of a protein (nm)
T: temperature (K)
η : viscosity (N s/m²)



<u>In the cytoplasm:</u>

T ≈ 300 K

η ≈ 10^{-2} N s/m²

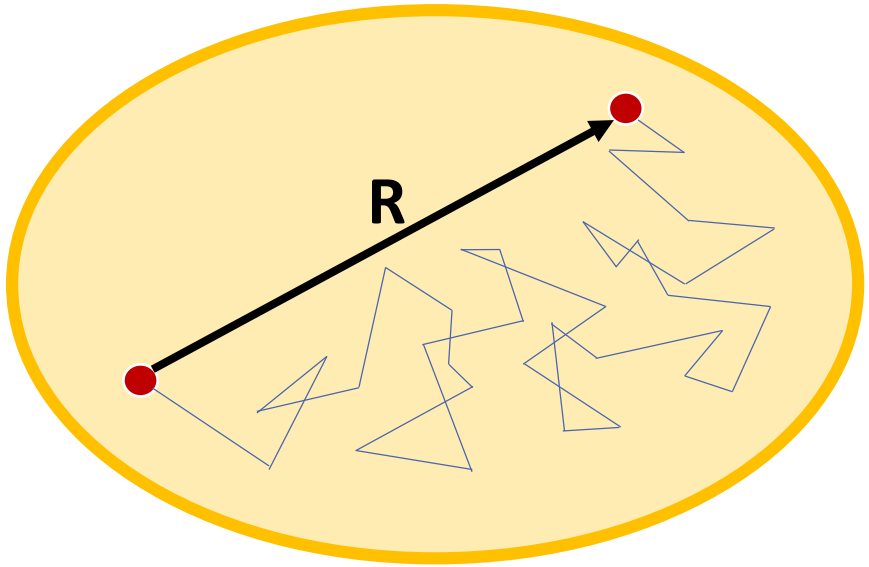
$$D \approx 10\ \mu m^2/s$$

Protein diffusion

molecule	measured context	diffusion coefficient ($\mu m^2/s$)
H_2O	water	2000
H_2O	nucleus of chicken erythrocyte	200
H^+ (from H_3O^+ to H_2O)	water	7000
O_2	water	2000
CO_2	water	2000
tRNA (≈20 kDa)	water	100
protein (≈30 kDa GFP)	water	100
protein (≈30 kDa GFP)	eukaryotic cell (CHO) cytoplasm	30
protein (≈30 kDa GFP)	rat liver mitochondria	30
protein (NLS-EGFP)	cytoplasm of *D. melanogaster* embryo	20
protein (≈30 kDa)	*E. coli* cytoplasm	7-8
protein (≈40 kDa)	*E. coli* cytoplasm	2-4
protein (≈70-250 kDa)	*E. coli* cytoplasm	0.4-2
protein (≈140 kDa Tar-YFP)	*E. coli* membrane	0.2
protein (≈70 kDa LacY-YFP)	*E. coli* membrane	0.03
fluorescent dye (carboxy-fluorescein)	*A. thaliana* cell wall	30
fluorescent dye (carboxy-fluorescein)	*A. thaliana* mature root epidermis	3
transcription factor (LacI)	movement along DNA (1D, *in vitro*)	0.04 (4×10^5 bp^2s^{-1})
morphogen (bicoid-GFP)	cytoplasm of *D. melanogaster* embryo	7
morphogen (wingless)	wing imaginal disk of *D. melanogaster*	0.05
mRNA	HeLa nucleus	0.03-0.10
mRNA	various localizations and sizes	0.005-1
ribosome	*E. coli*	0.04

book.bionumbers.org/what-are-the-time-scales-for-diffusion-in-cells

$$D \approx 10\ \mu m^2/s$$



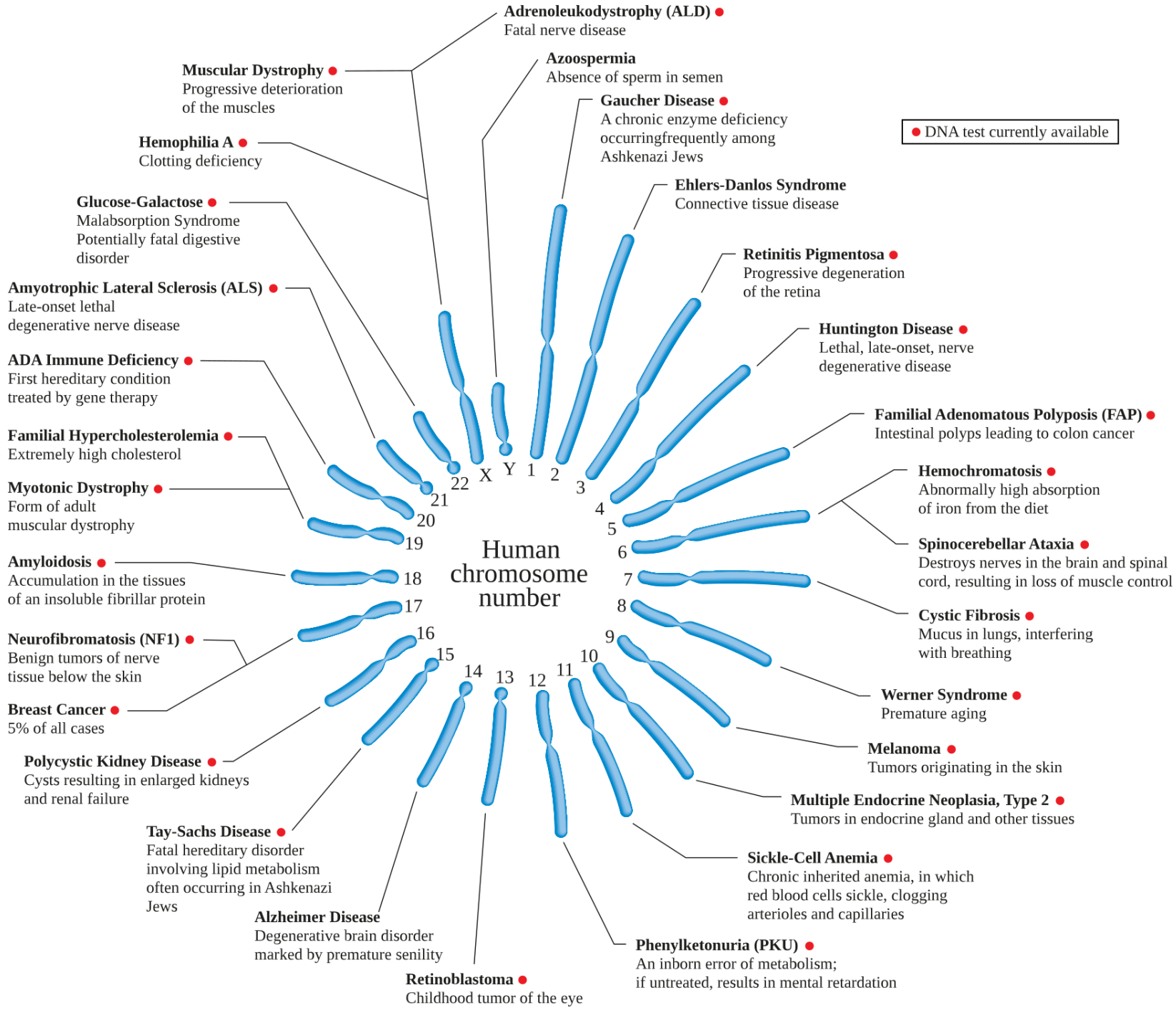
E. coli R ≈ 1 µm

$$t = \frac{R^2}{6D} \approx 17\ ms$$

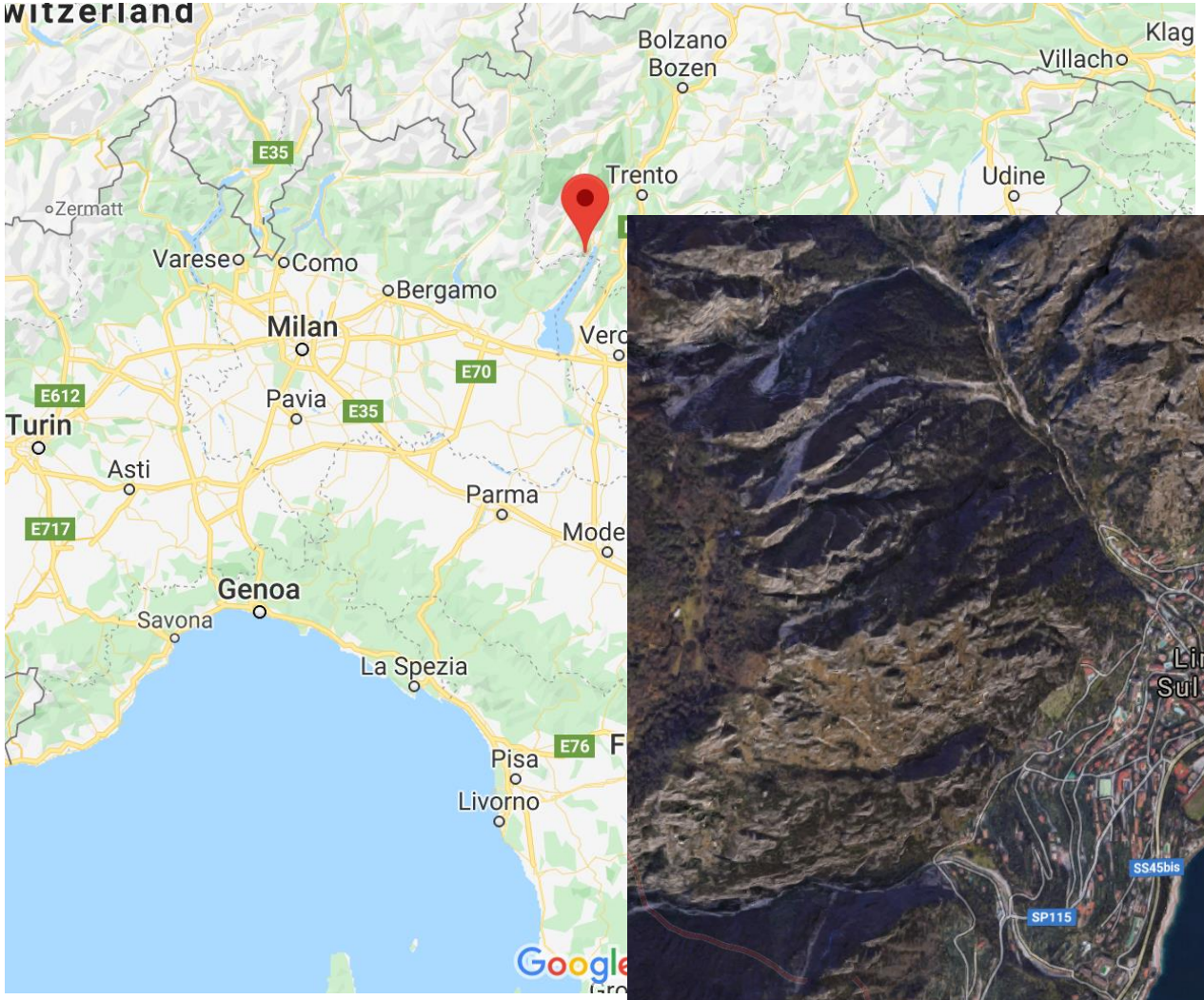
Protein function and malfunction

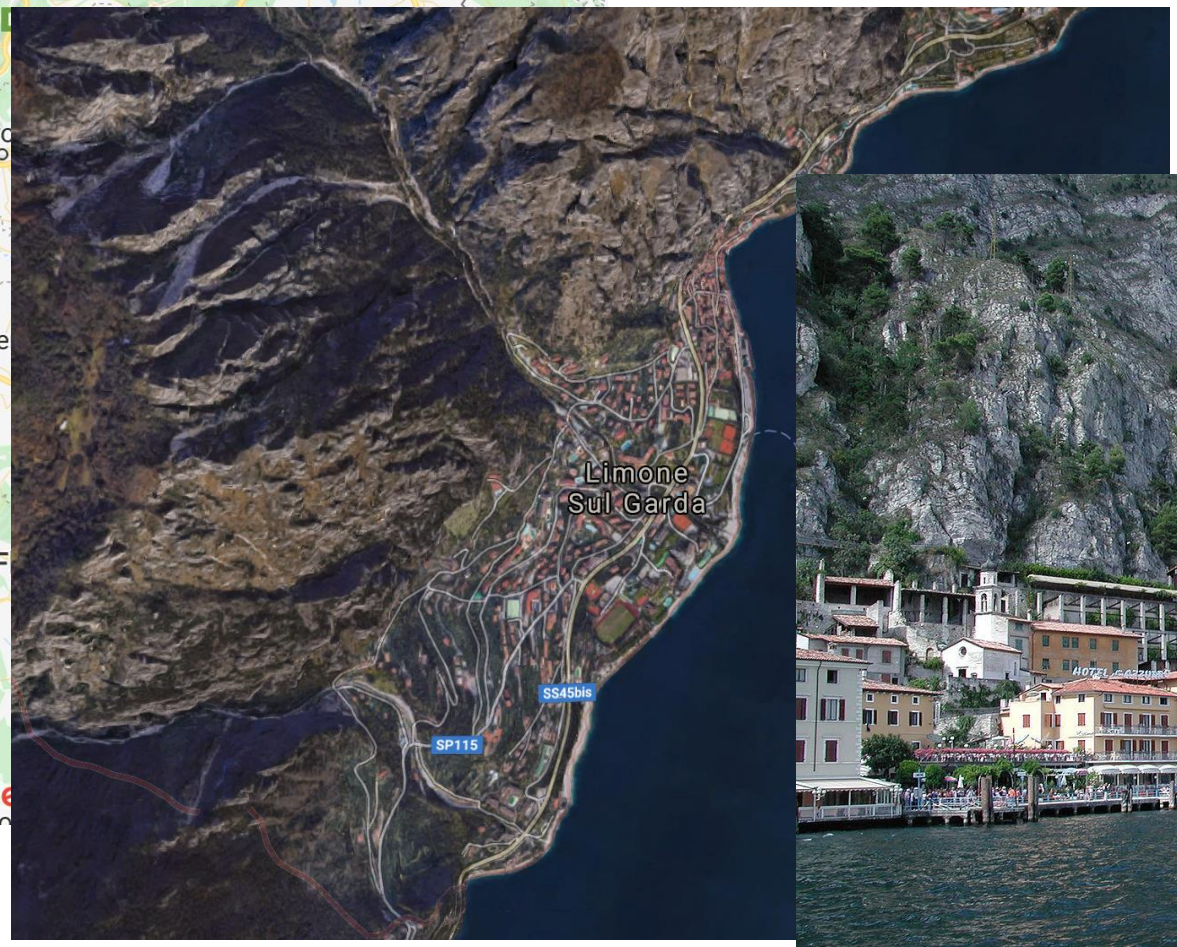
- **Effect of mutations**
  - Protein cannot fold
  - Proteins misfolds
  - Protein folds, but has different dynamics

- **Consequences of mutations**
  - Protein cannot bind to anything
  - Protein interacts with one or more different binding partners


Adrenoleukodystrophy (ALD) ●
Fatal nerve disease
Azoospermia
Absence of sperm in semen
Muscular Dystrophy ●
Progressive deterioration of the muscles
Gaucher Disease ●
A chronic enzyme deficiency occurringfrequently among Ashkenazi Jews
● DNA test currently available
Hemophilia A ●
Clotting deficiency
Ehlers-Danlos Syndrome
Connective tissue disease
Glucose-Galactose ●
Malabsorption Syndrome
Potentially fatal digestive disorder
Retinitis Pigmentosa ●
Progressive degeneration of the retina
Amyotrophic Lateral Sclerosis (ALS) ●
Late-onset lethal degenerative nerve disease
Huntington Disease ●
Lethal, late-onset, nerve degenerative disease
ADA Immune Deficiency ●
First hereditary condition treated by gene therapy
Familial Adenomatous Polyposis (FAP) ●
Intestinal polyps leading to colon cancer
Familial Hypercholesterolemia ●
Extremely high cholesterol
Hemochromatosis ●
Abnormally high absorption of iron from the diet
Myotonic Dystrophy ●
Form of adult muscular dystrophy
Spinocerebellar Ataxia ●
Destroys nerves in the brain and spinal cord, resulting in loss of muscle control
Amyloidosis ●
Accumulation in the tissues of an insoluble fibrillar protein
Cystic Fibrosis ●
Mucus in lungs, interfering with breathing
Human chromosome number
X Y 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Neurofibromatosis (NF1) ●
Benign tumors of nerve tissue below the skin
Werner Syndrome ●
Premature aging
Breast Cancer ●
5% of all cases
Melanoma ●
Tumors originating in the skin
Polycystic Kidney Disease ●
Cysts resulting in enlarged kidneys and renal failure
Multiple Endocrine Neoplasia, Type 2 ●
Tumors in endocrine gland and other tissues
Tay-Sachs Disease ●
Fatal hereditary disorder involving lipid metabolism often occurring in Ashkenazi Jews
Sickle-Cell Anemia ●
Chronic inherited anemia, in which red blood cells sickle, clogging arterioles and capillaries
Alzheimer Disease
Degenerative brain disorder marked by premature senility
Phenylketonuria (PKU) ●
An inborn error of metabolism; if untreated, results in mental retardation
Retinoblastoma ●
Childhood tumor of the eye

Sometimes mutations are beneficial



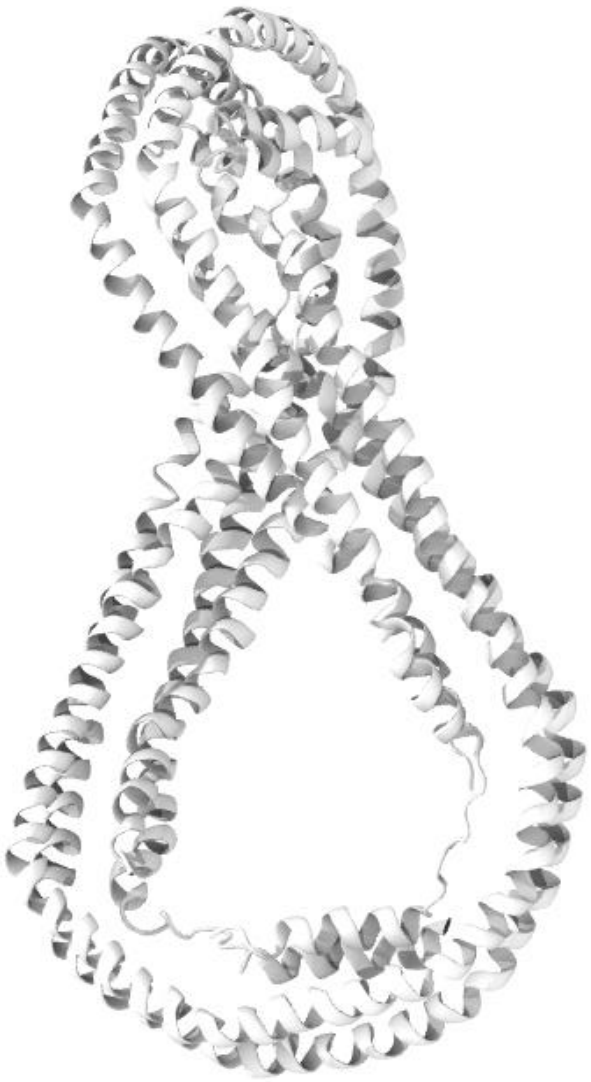


Limone Sul Garda, Italy

Sometimes mutations are beneficial

apolipoprotein A-1 Milano:

- carries cholesterol from tissues to the liver
- Phenotype:
  - reduction in HDL levels and an increase in triglyceride levels
  - positive effects against atherosclerosis
- mutation R173C developed by Giovanni Pomarelli (1780)
- gene therapy works well on animal models
- research as medical treatment stalled (Esperion, Pfizer, SemBioSys, ...)



blogs.sciencemag.org/pipeline/archives/2016/11/16/the-long-saga-of-apo-a1-milano

Sirtori et al., *Cardiovascular status of carriers of the apolipoprotein A-I(Milano) mutant: the Limone sul Garda study*, Circulation, 2001

K.-Y. Chyu and P.K. Shah, *HDL/ApoA-1 infusion and ApoA-1 gene therapy in atherosclerosis*, Frontiers in Pharmacology, 2015

Take home messages

- Proteins often assemble into complexes to carry out their task
- The intracellular space is crowded
- Mutation may alter protein capacity of interacting with its dedicated partner