Biological molecules: Order, Disorder, Chaos

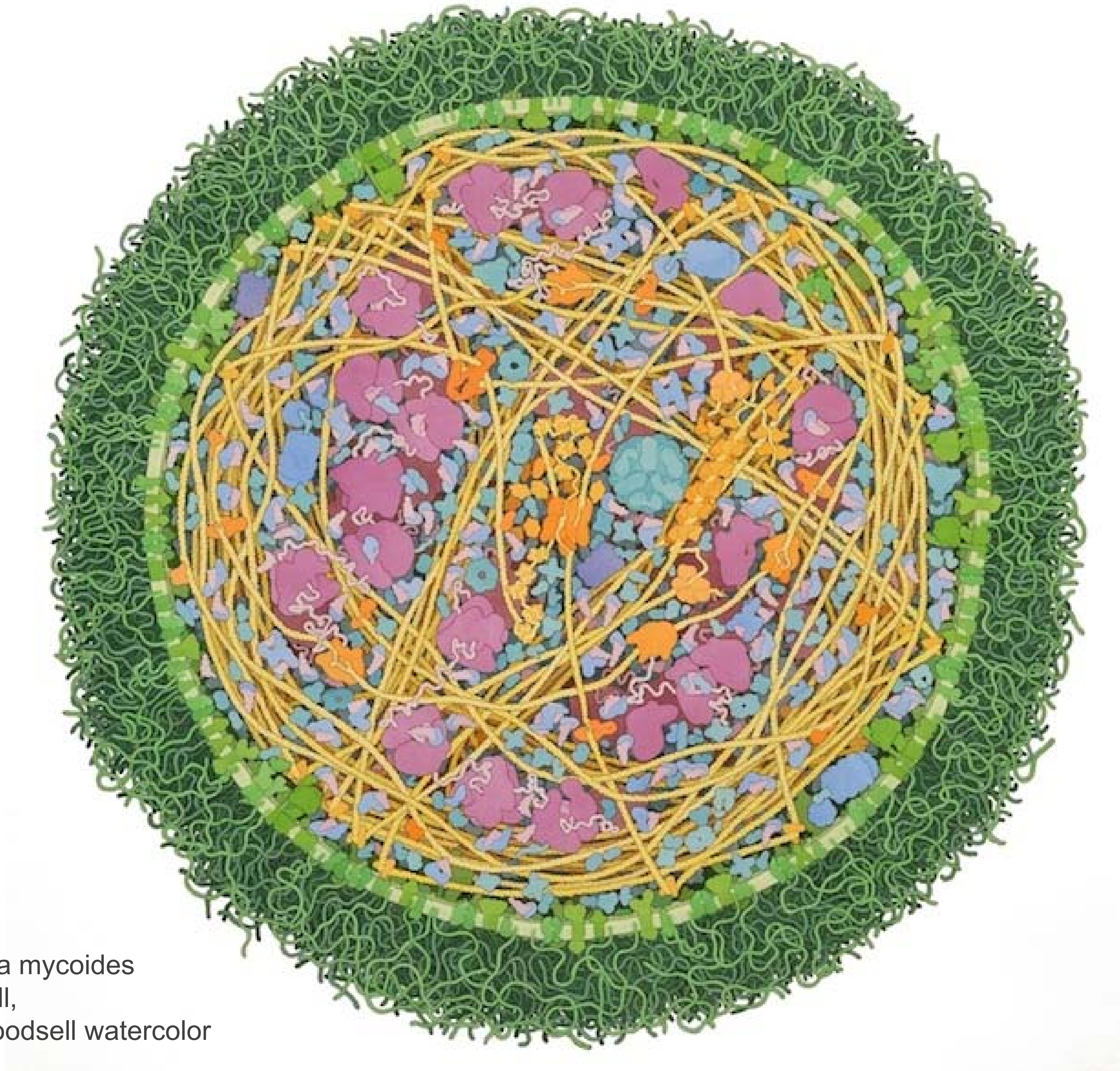
Jim Clarage
Department of Chemistry and Physics
University of St. Thomas

Biological molecules: Order, Disorder, Chaos

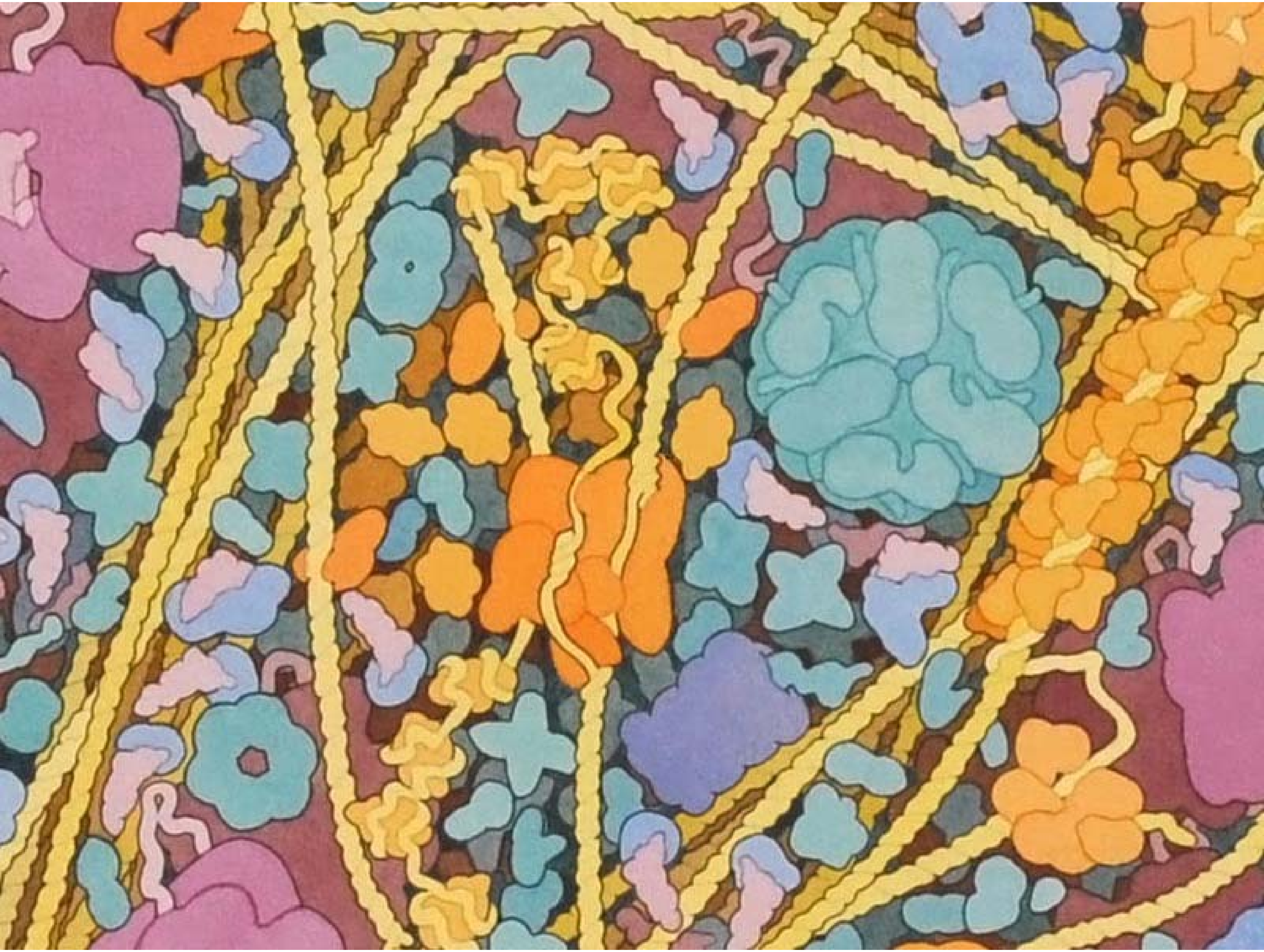
... A physicist's paean to proteins

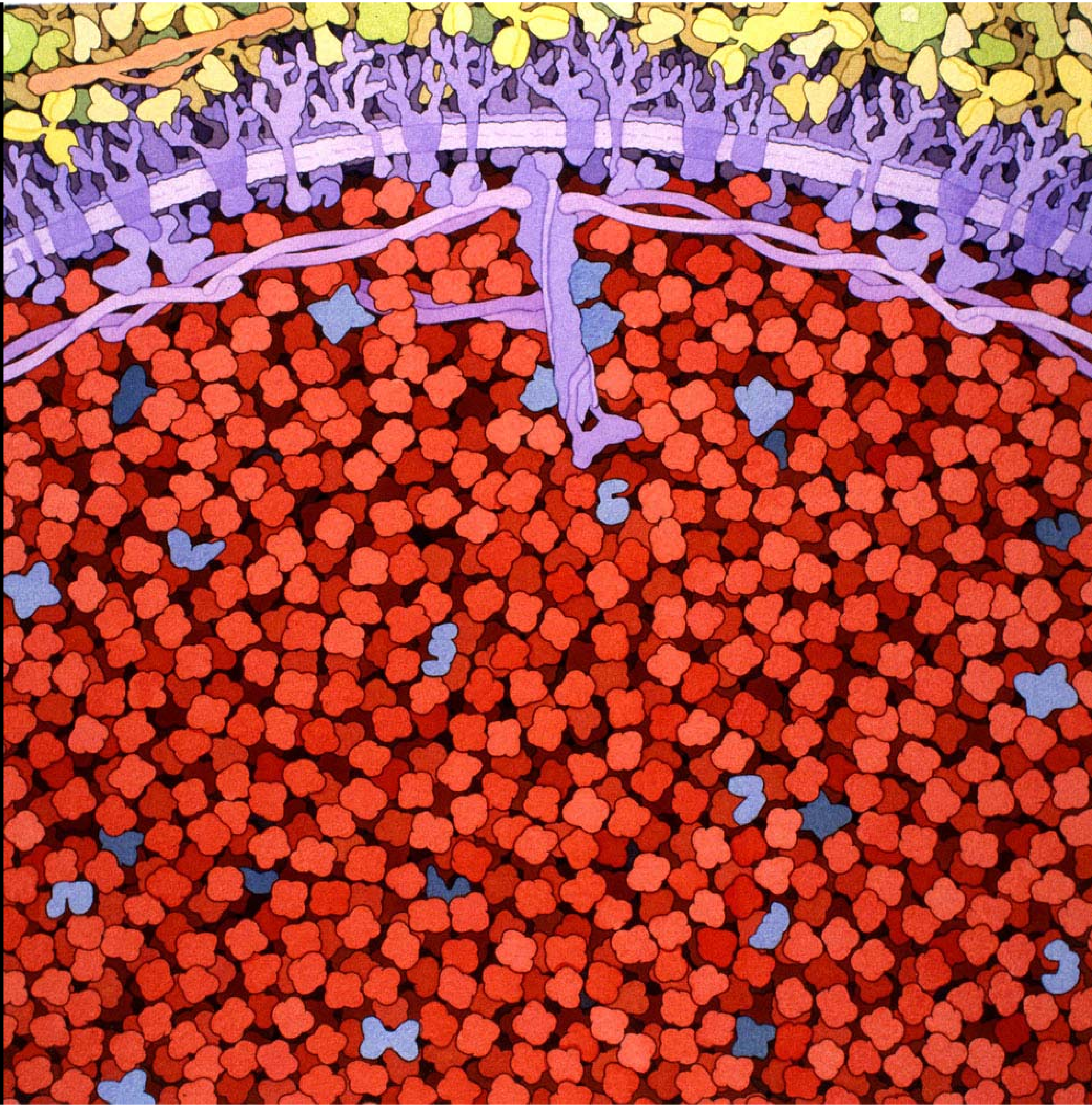
paean |ˈpēən|, Greek. A song of praise

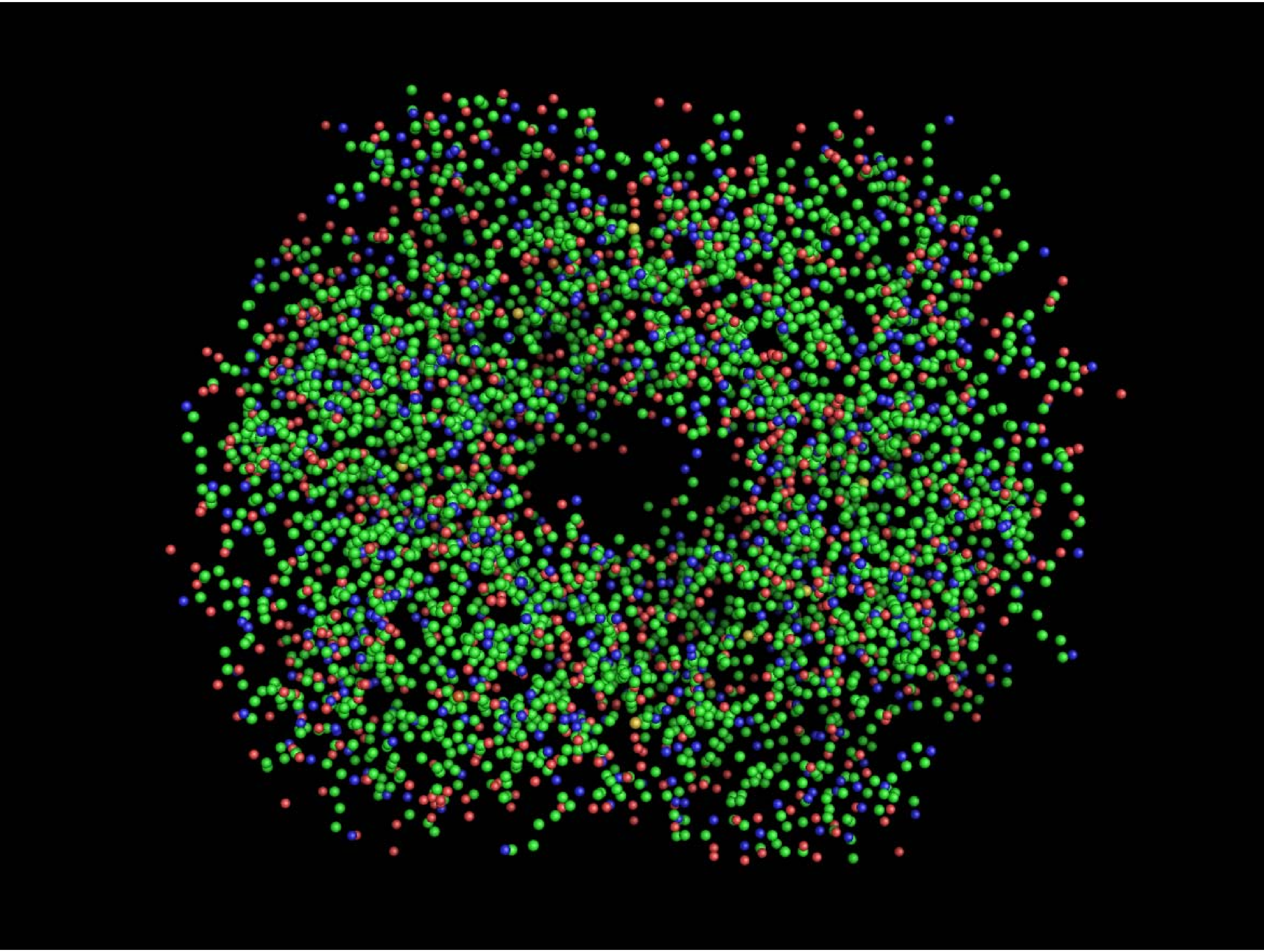
Introduction

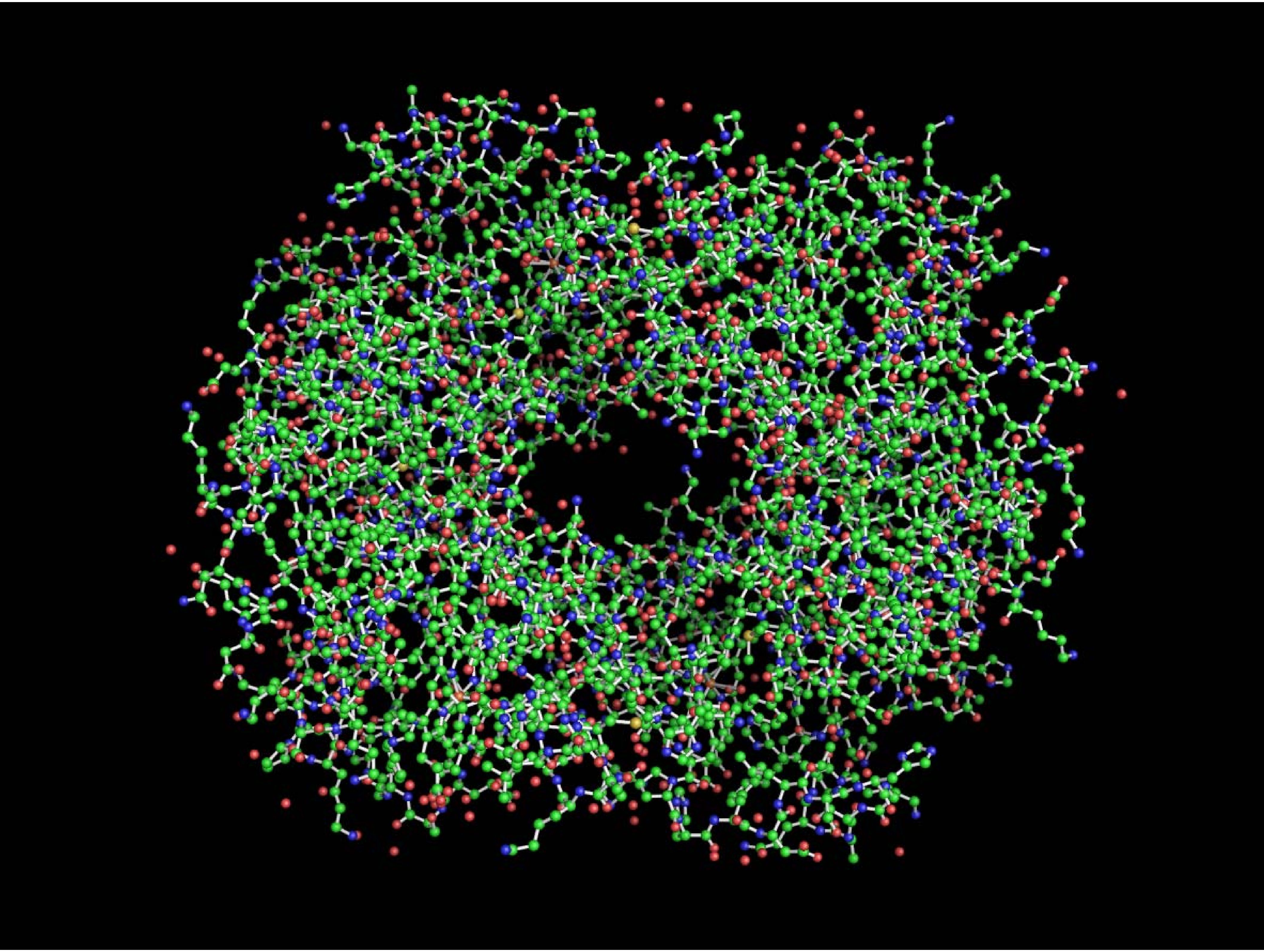


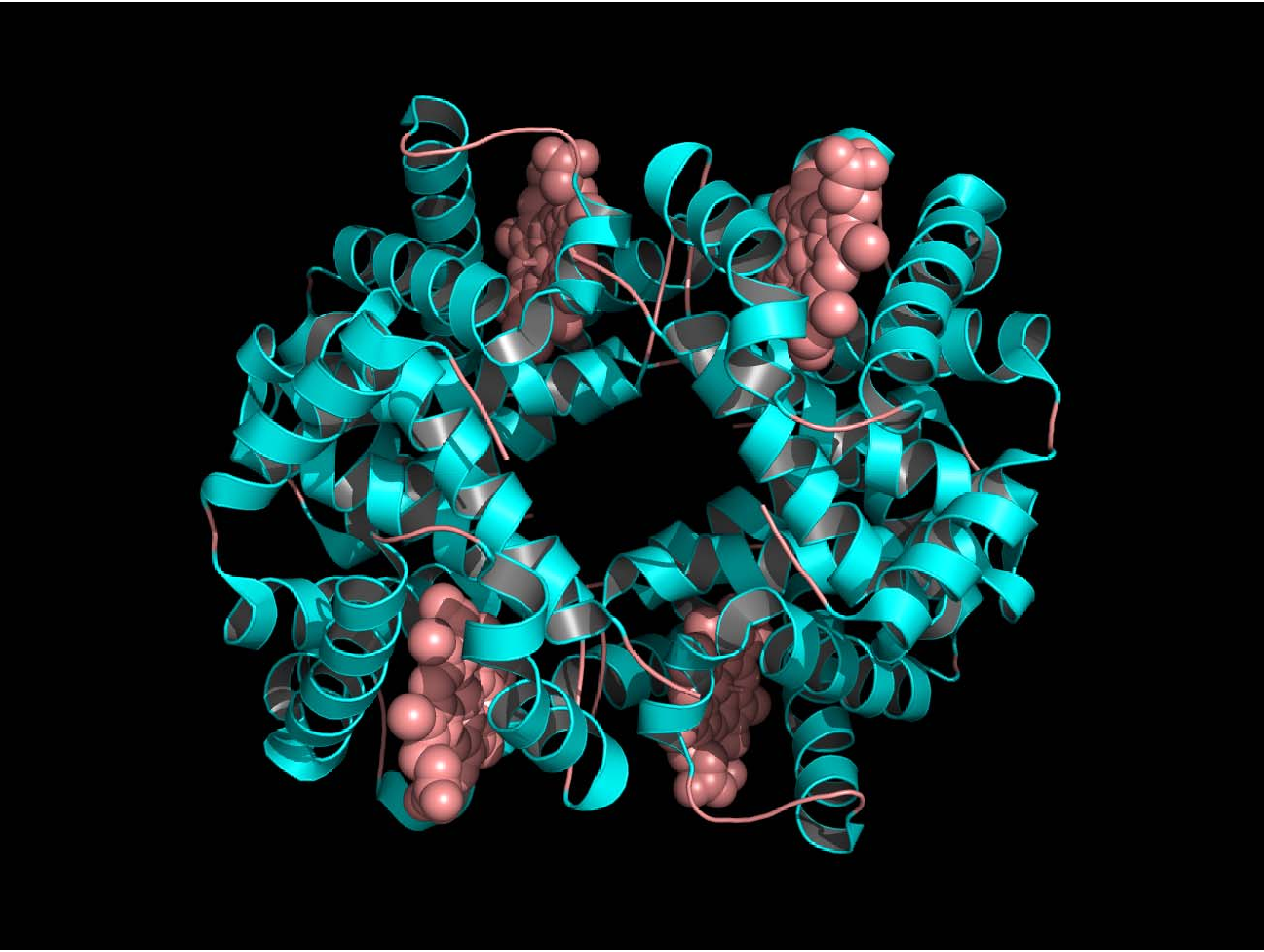
Mycoplasma mycoides
bacterial cell,
David S. Goodsell watercolor



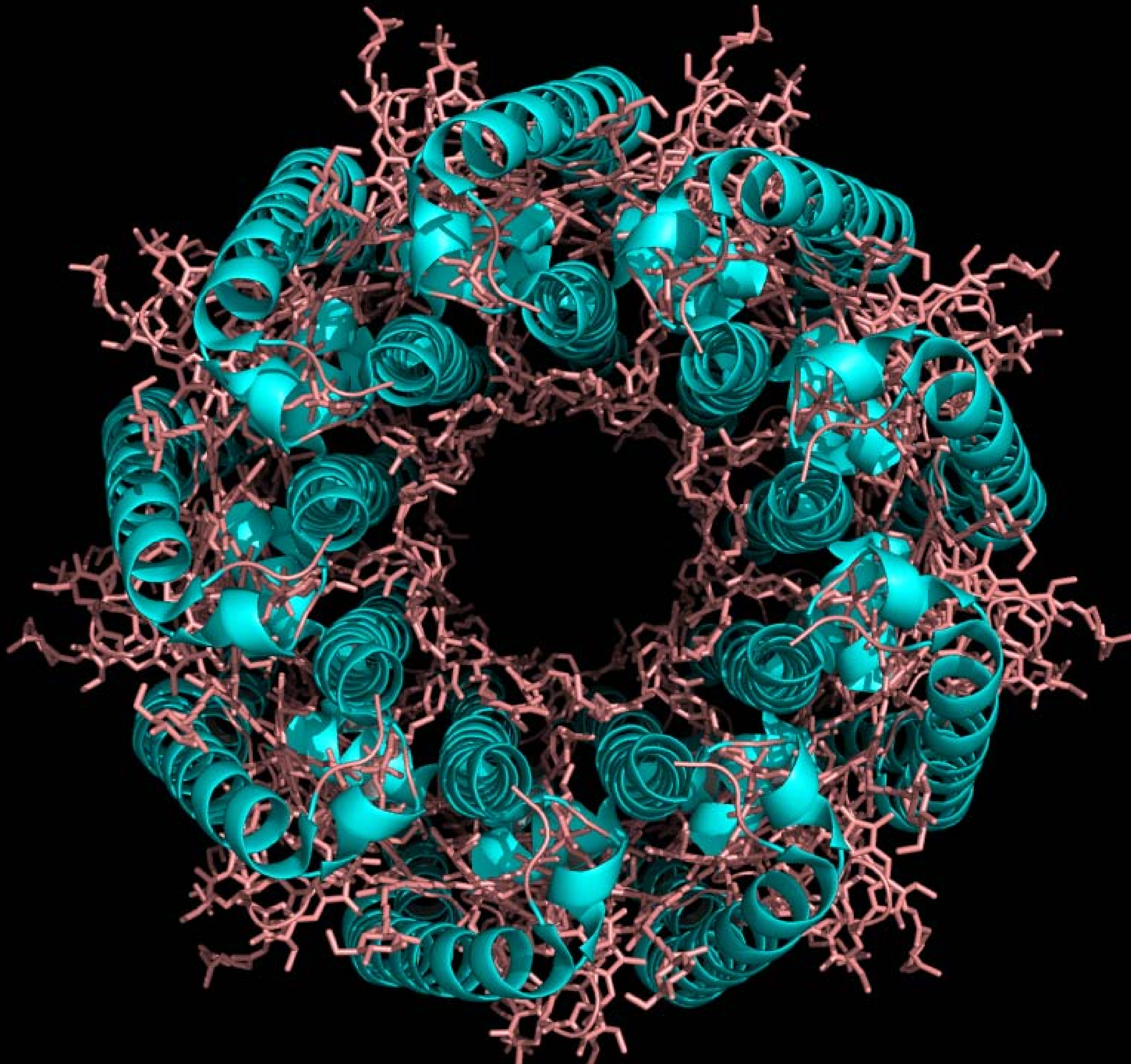


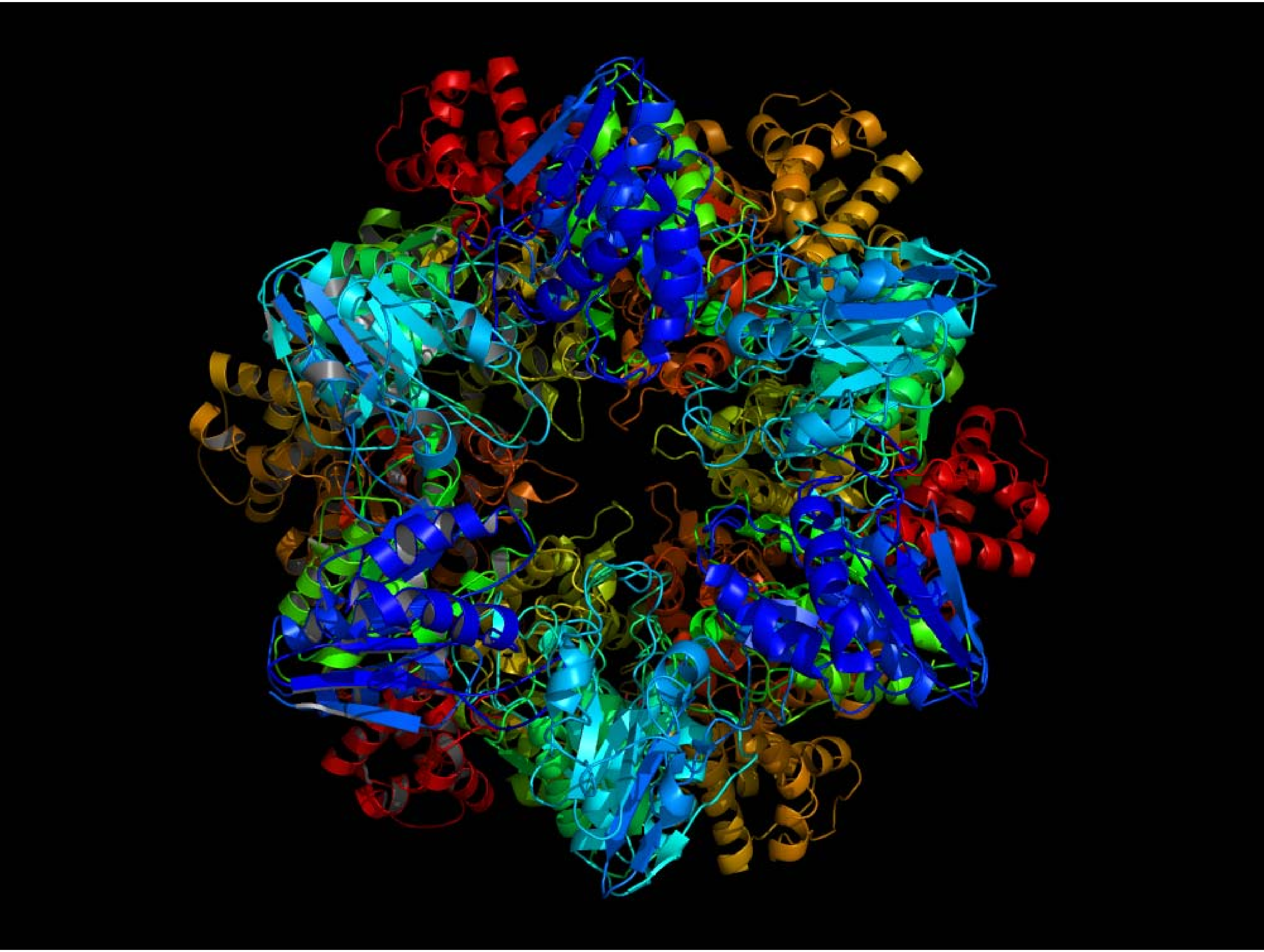


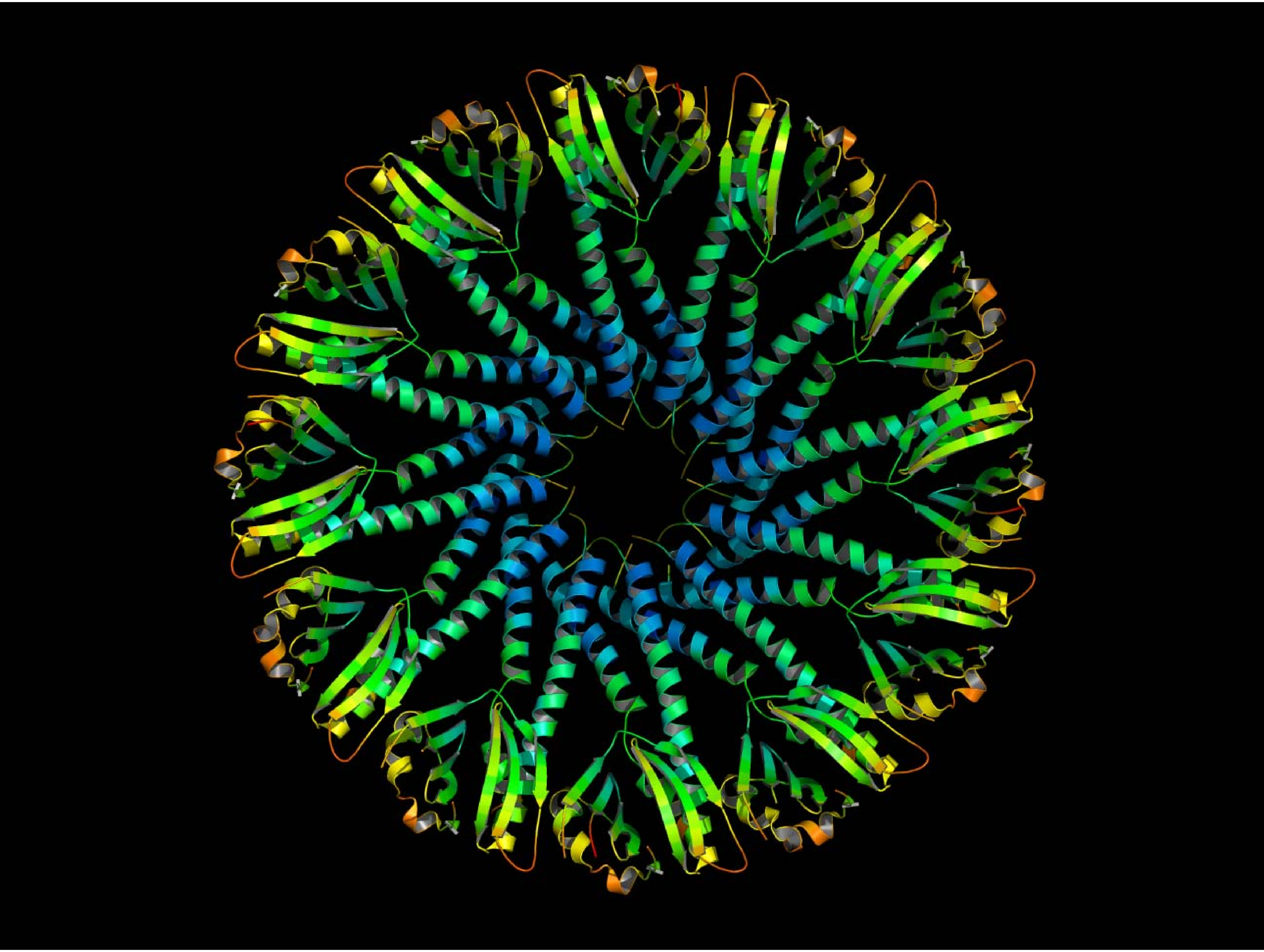




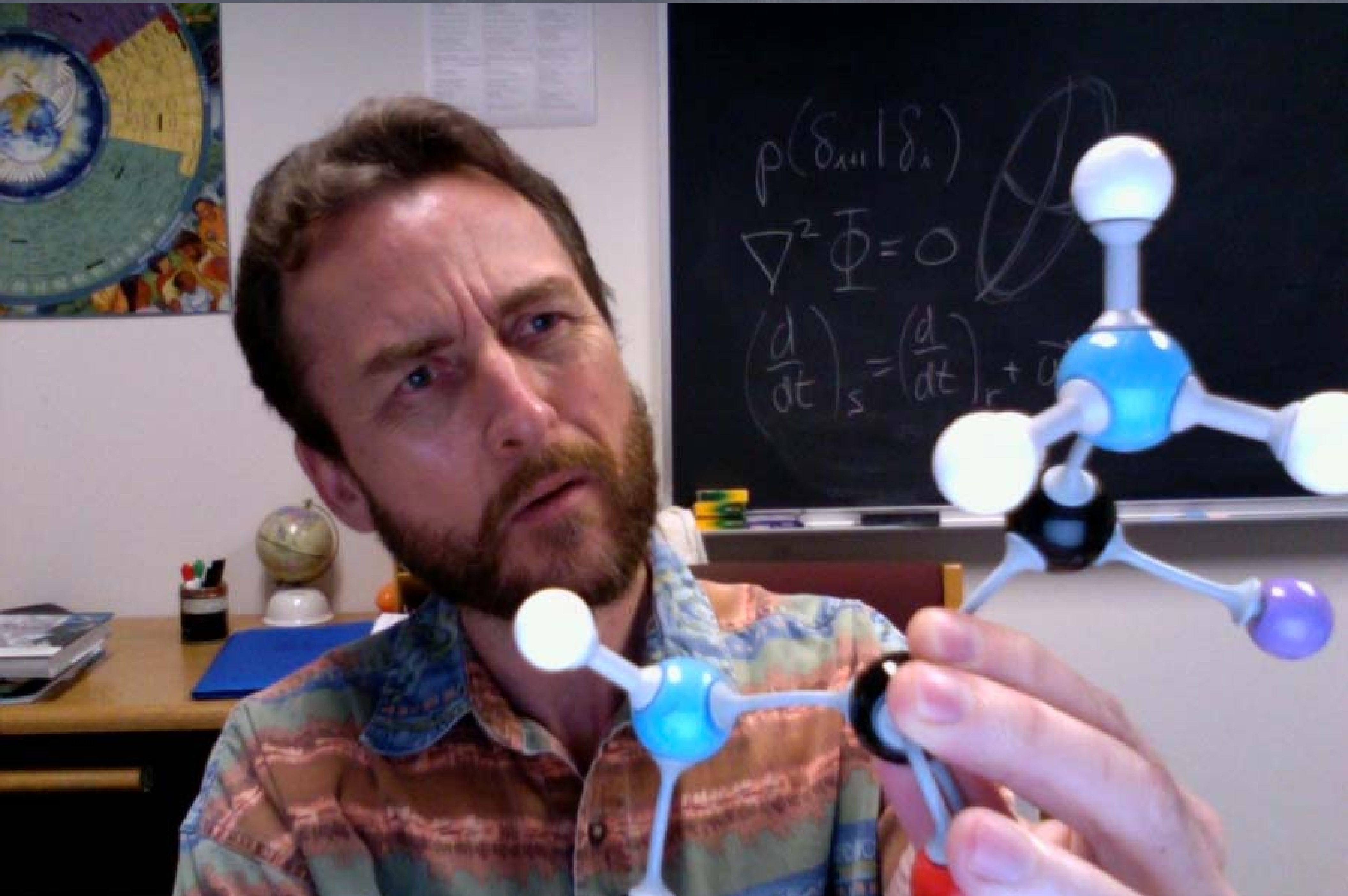
Form ever follows function - Louis Sullivan (1896)





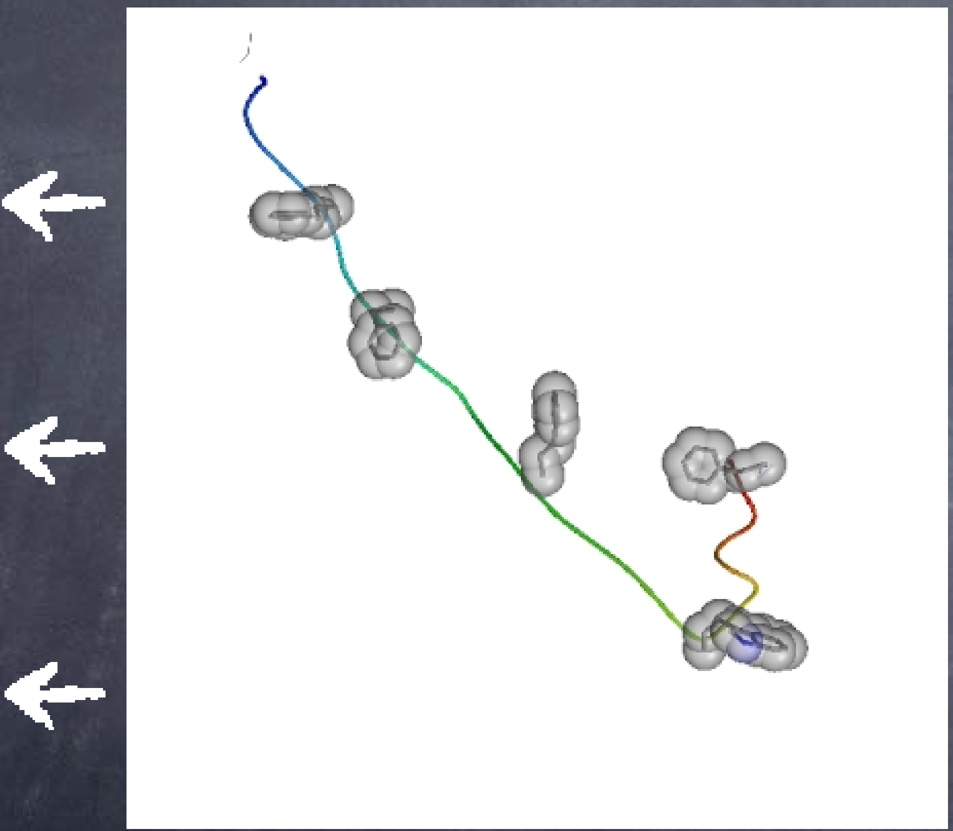


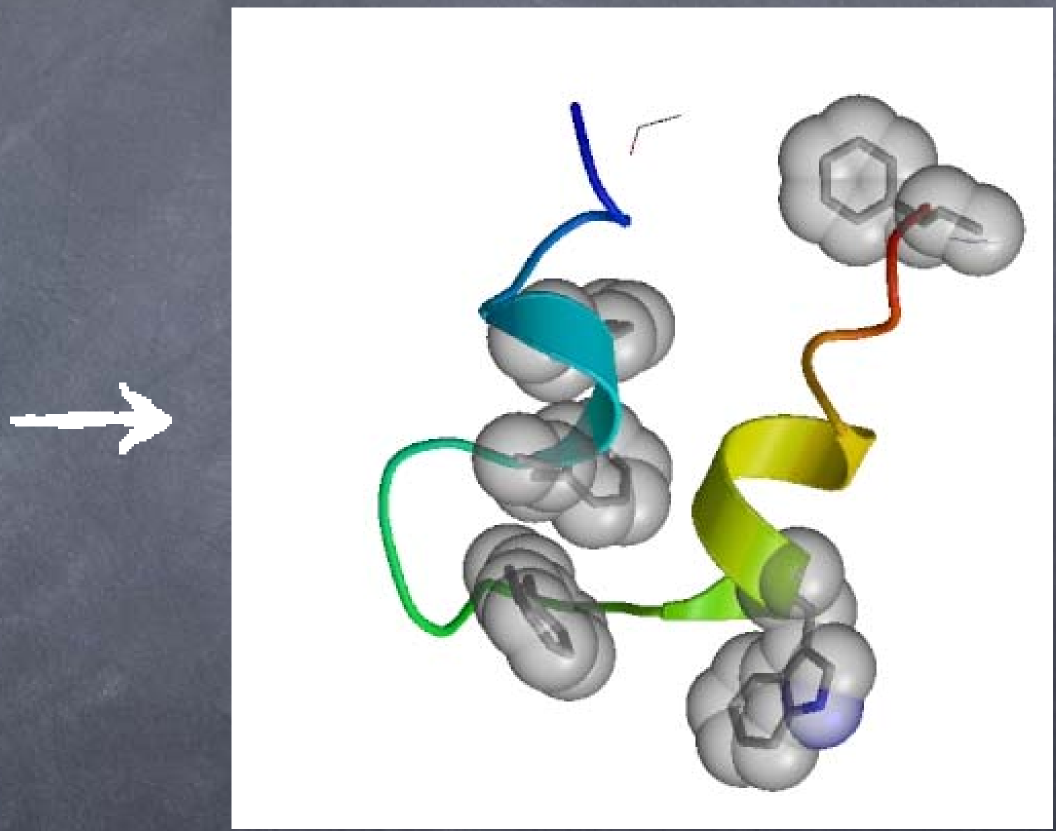
Order



$\rho(\delta_{i+1}|\delta_i)$
$\nabla^2\Phi = 0$

The protein folding problem

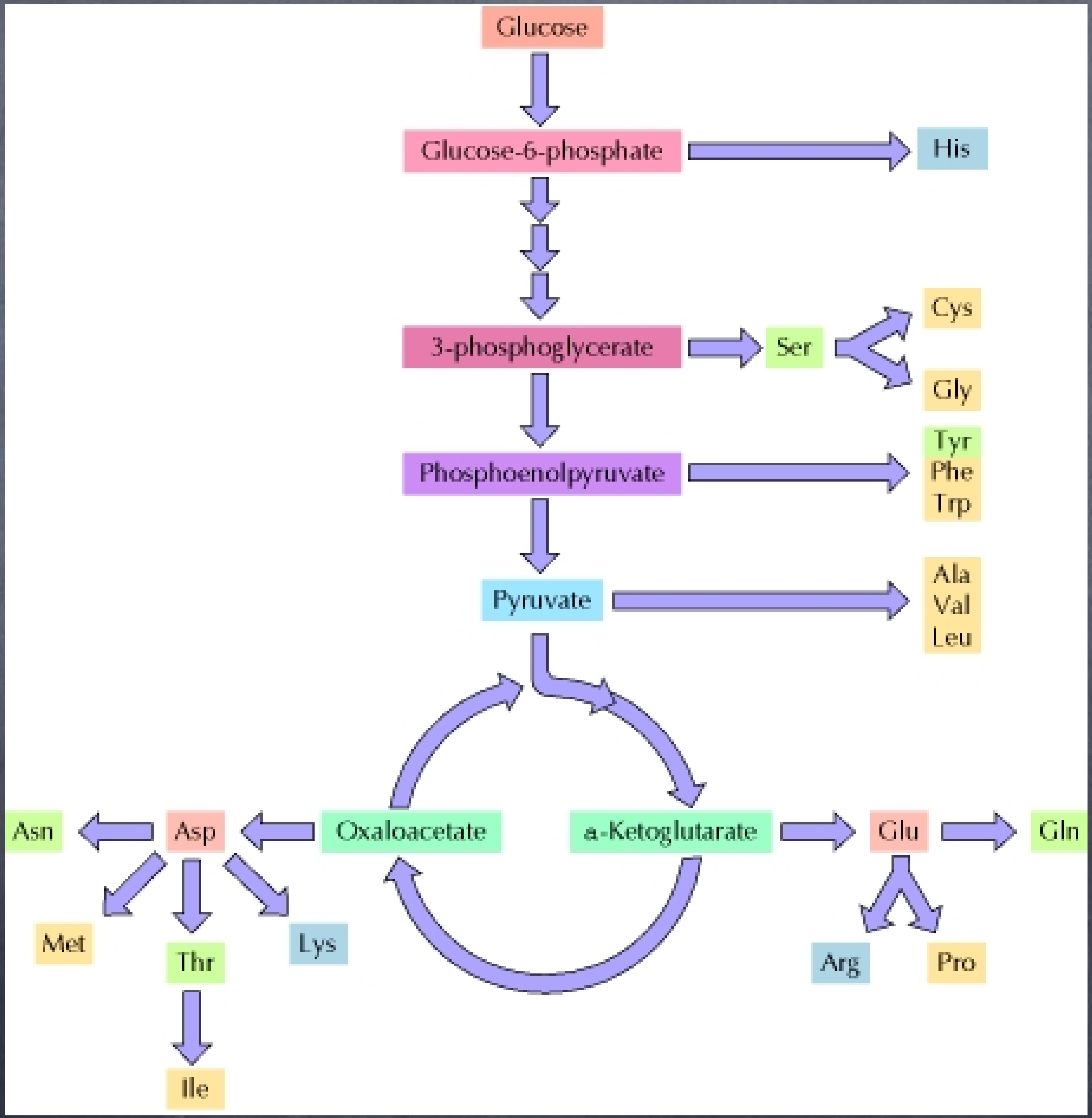




Villin, 36 amino acids, ~10 μs

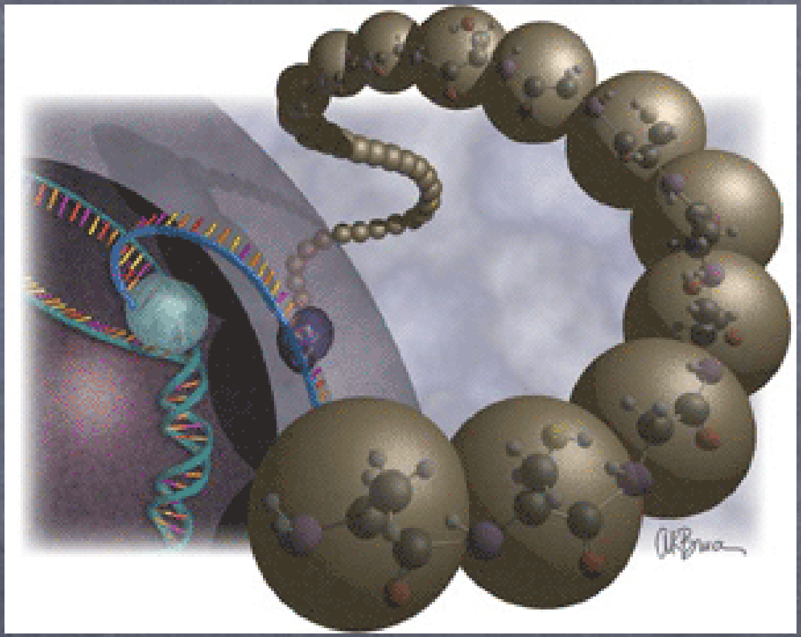
Biosynthesis of amino acids
(glycolysis and in the citric acid cycle (eek!))

C, N, O →

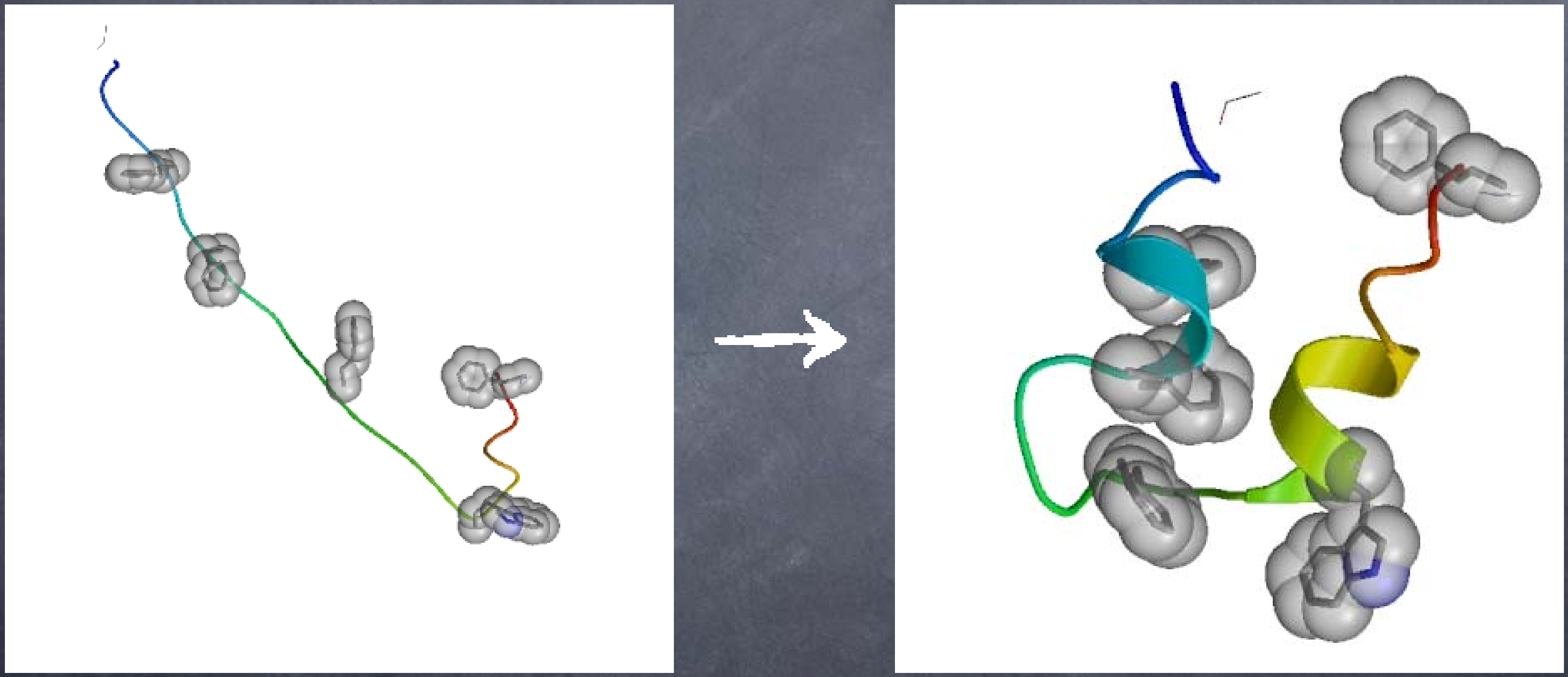


→

Polymerization of amino acids
(protein translation)



The protein folding problem



Villin, 36 amino acids, ~10 μs

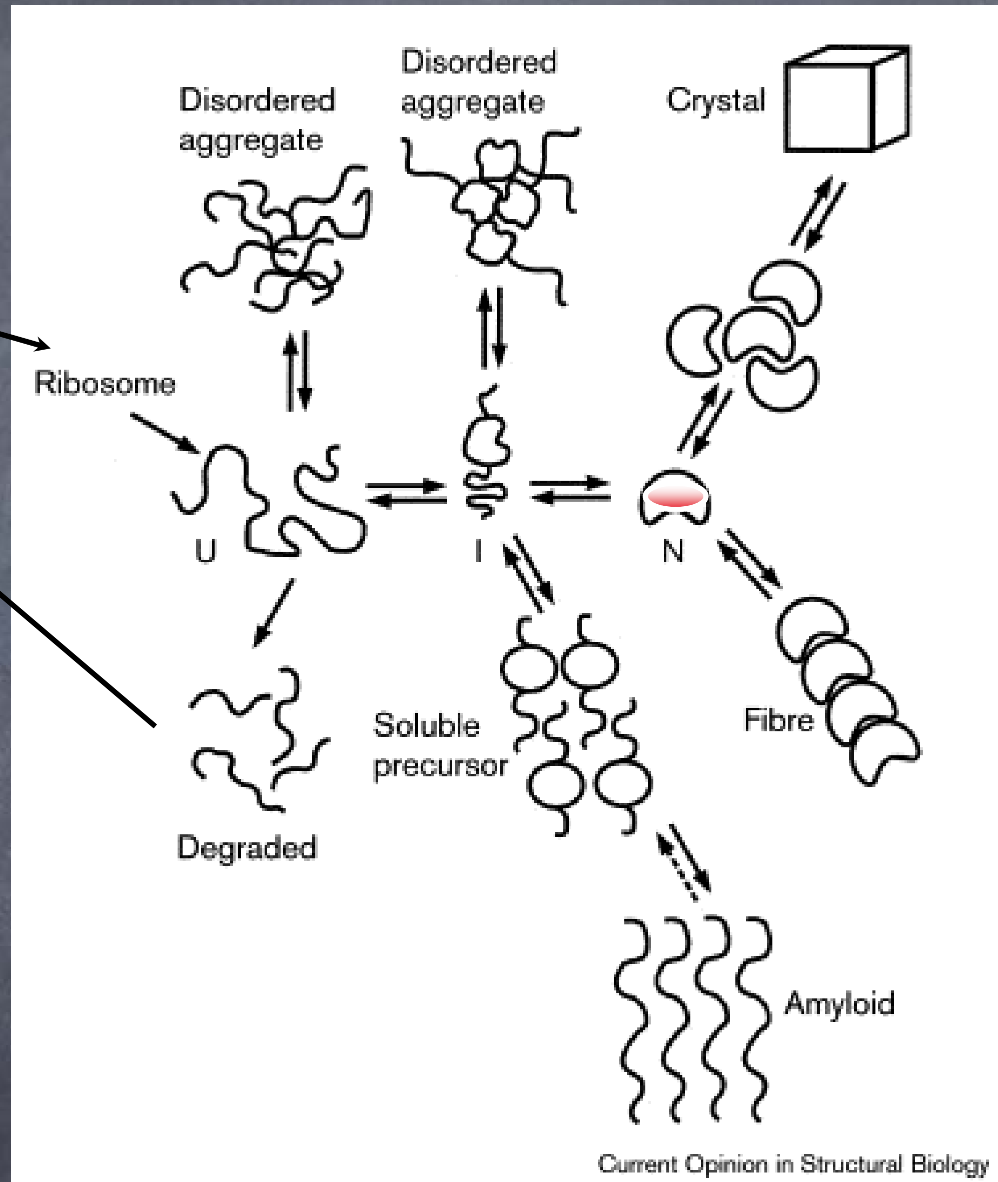
QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

Bowman, G. R., & Pande, V. S. (2010). Protein folded states are kinetic hubs. *Proceedings of the National Academy of Sciences*, *107*(24), 10890. National Acad Sciences.

elements

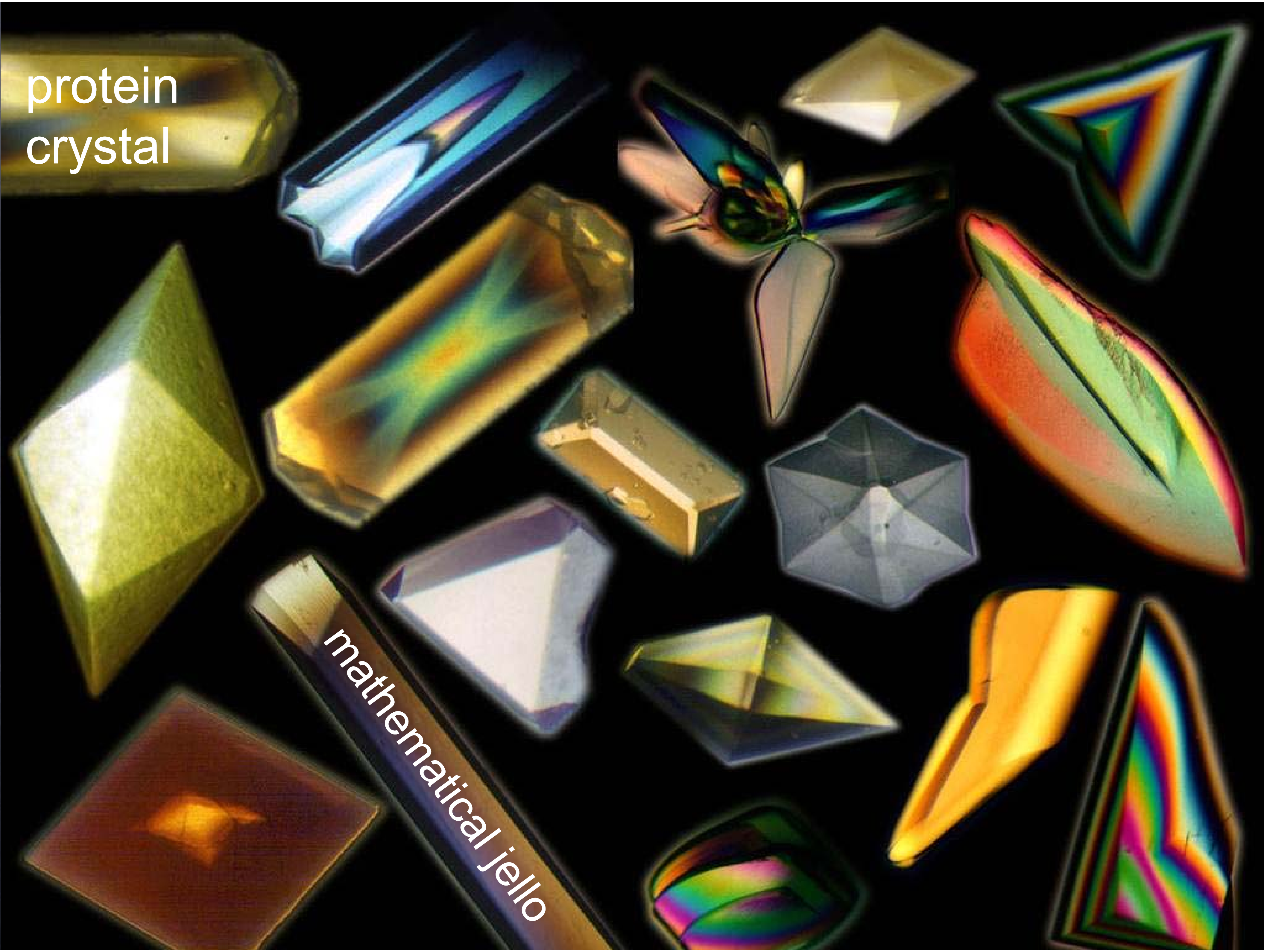
C, N, O

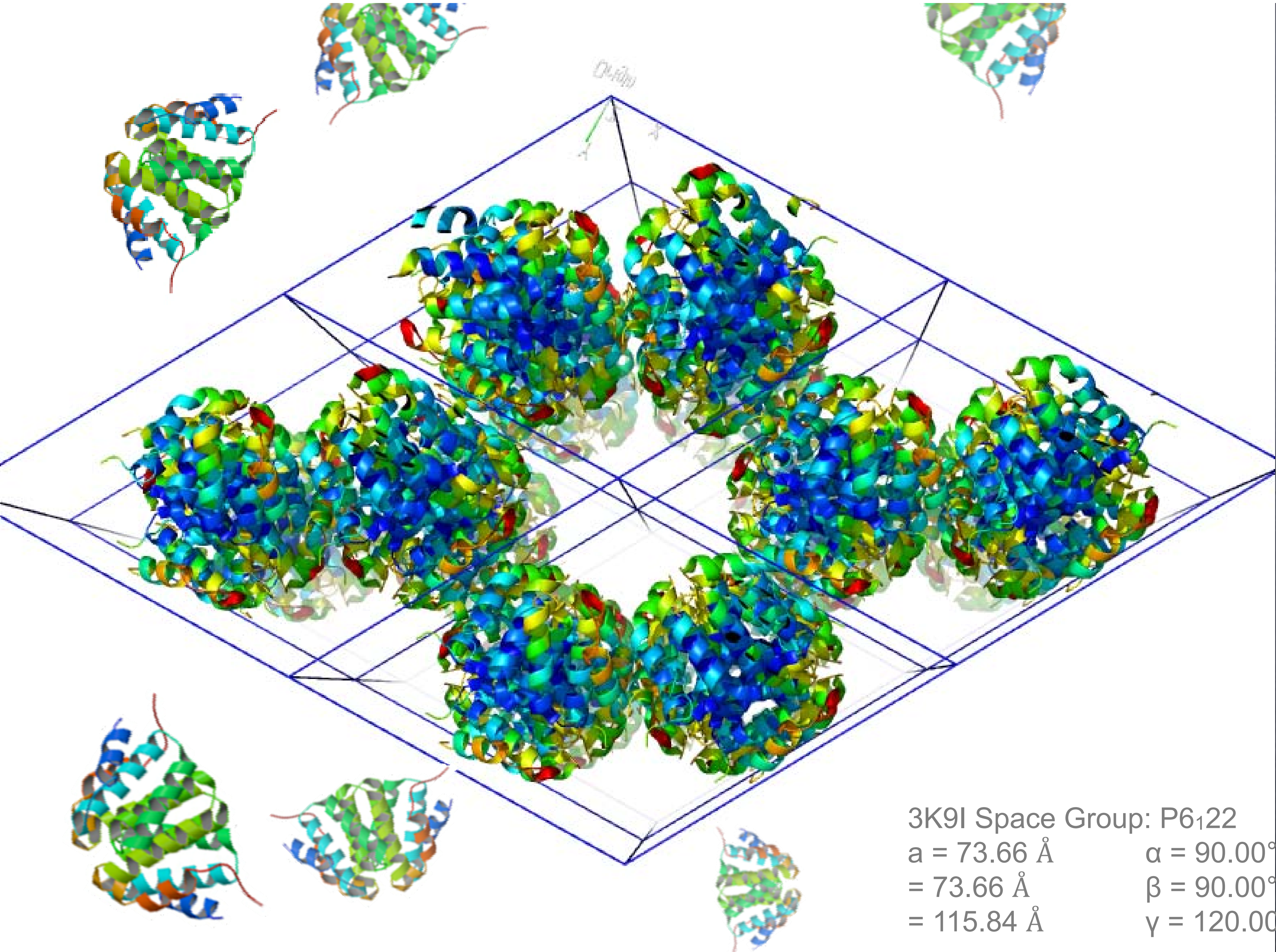
free amino acid monomers

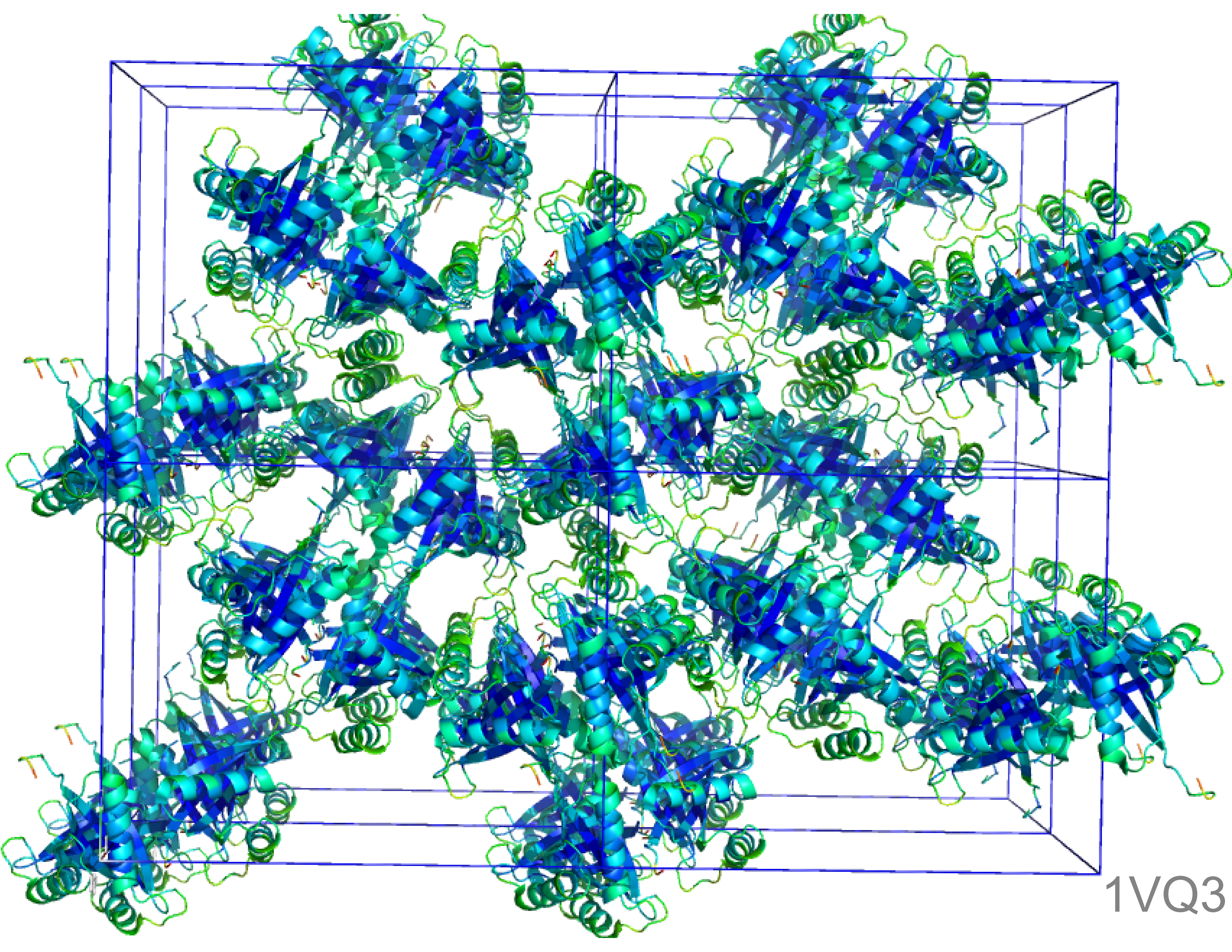


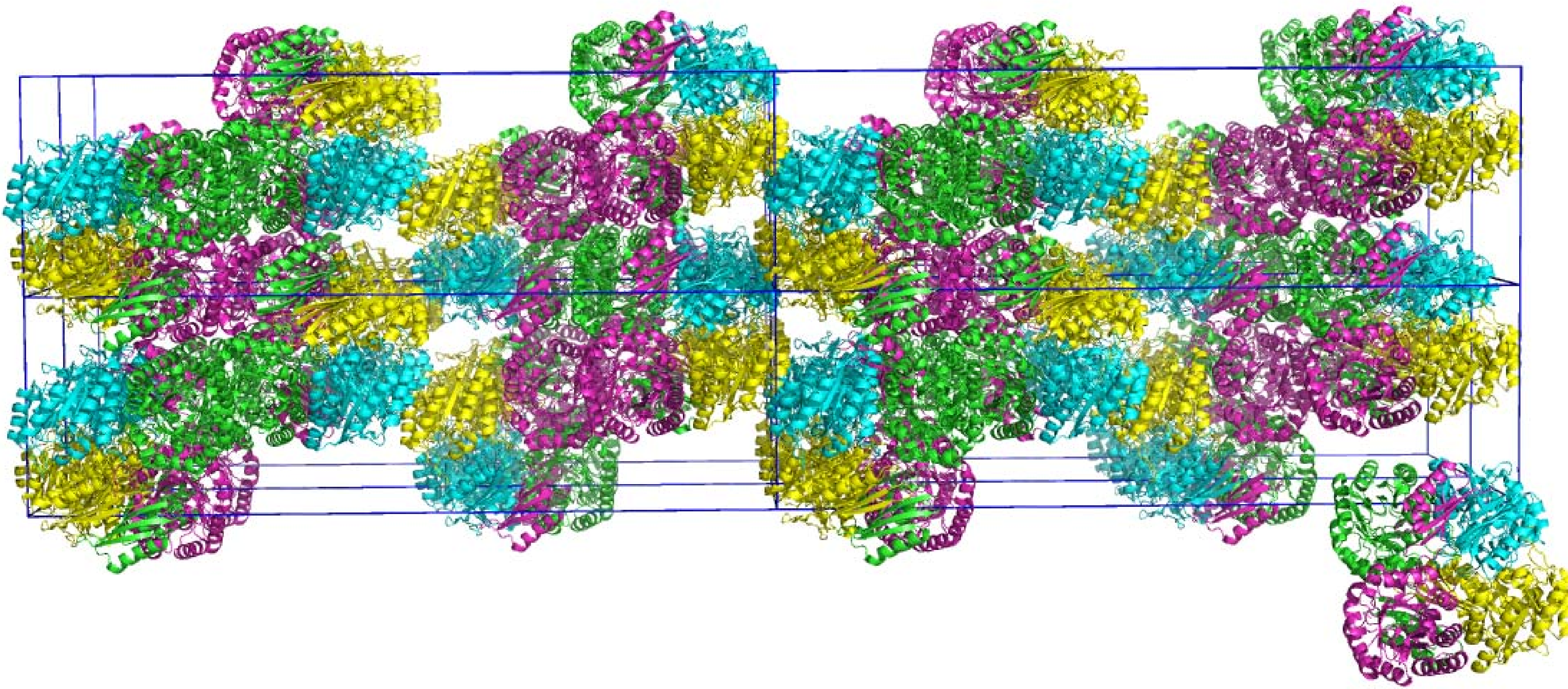
chaperones

Adapted from: Dobson and Karplus, Current Opinion in Structural Biology Volume 9, Issue 1, 1 February 1999, Pages 92–101

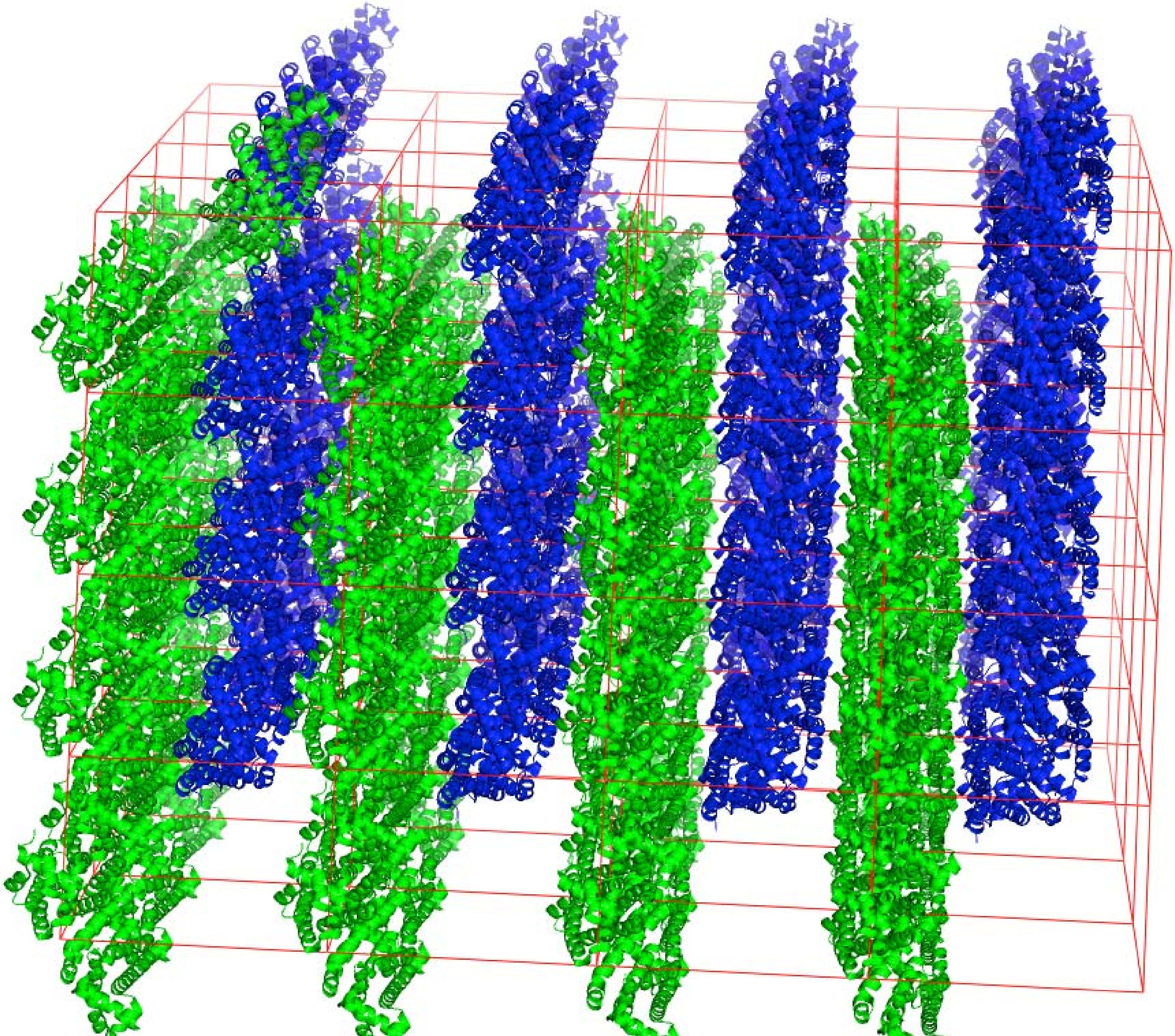

protein
crystal
mathematical jello


3K9I Space Group: $P6_122$
a = 73.66 Å
α = 90.00
= 73.66 Å
β = 90.00
= 115.84 Å
γ = 120.0

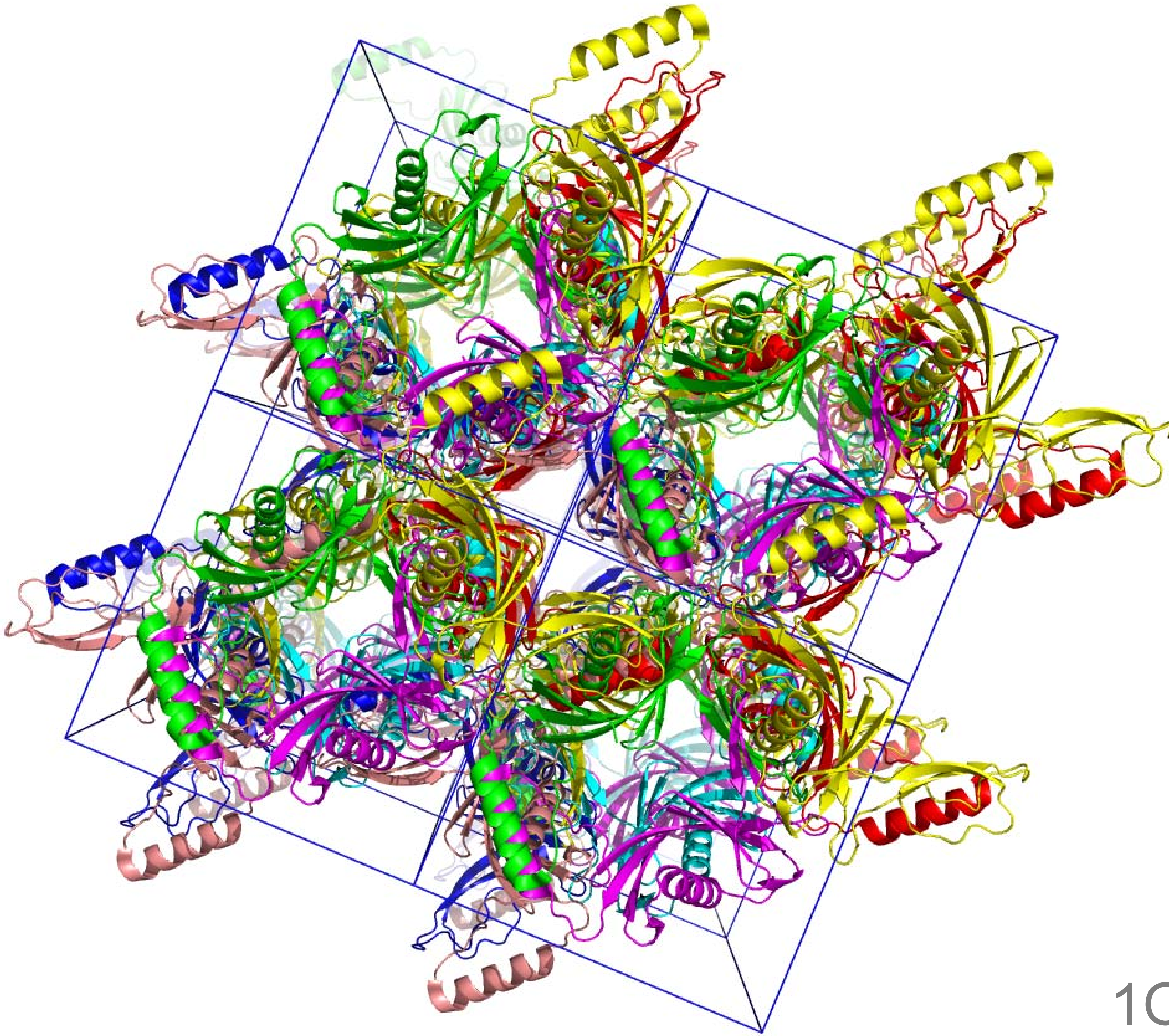

1VQ3



1VR6



1ZKG


1O22

Crystalline order

$$\Delta G_{xtal} - \Delta H_{xtal} - T\Delta S_{xtal} < 0$$

Crystalline order

$$\Delta G_{xtal} = \Delta H_{xtal} - T\Delta S_{xtal} \overset{?}{<} 0$$

$\Delta H_{xtal} \geq 0$ (e.g., +155 kJ/mol for Hemoglobin C)

$\Delta S_{xtal} < 0$ (e.g., -100 to -300 J/mol-K)

Derewenda, Z. S., & Vekilov, P. G. (2005). Entropy and surface engineering in protein crystallization. *Acta Cryst (2006). D62, 116-124 [doi:10.1107/S0907444905035237]*, 1–9. International Union of Crystallography. doi:10.1107/S0907444905035237

Crystalline order

$$\Delta G_{xtal} = \Delta H_{xtal} - T(\Delta S_{protein} + \Delta S_{solvent})_{xtal}$$

$$\Delta S_{solvent} >> 0$$

Derewenda, Z. S., & Vekilov, P. G. (2005). Entropy and surface engineering in protein crystallization. *Acta Cryst (2006). D62, 116-124 [doi:10.1107/S0907444905035237]*, 1–9. International Union of Crystallography. doi:10.1107/S0907444905035237

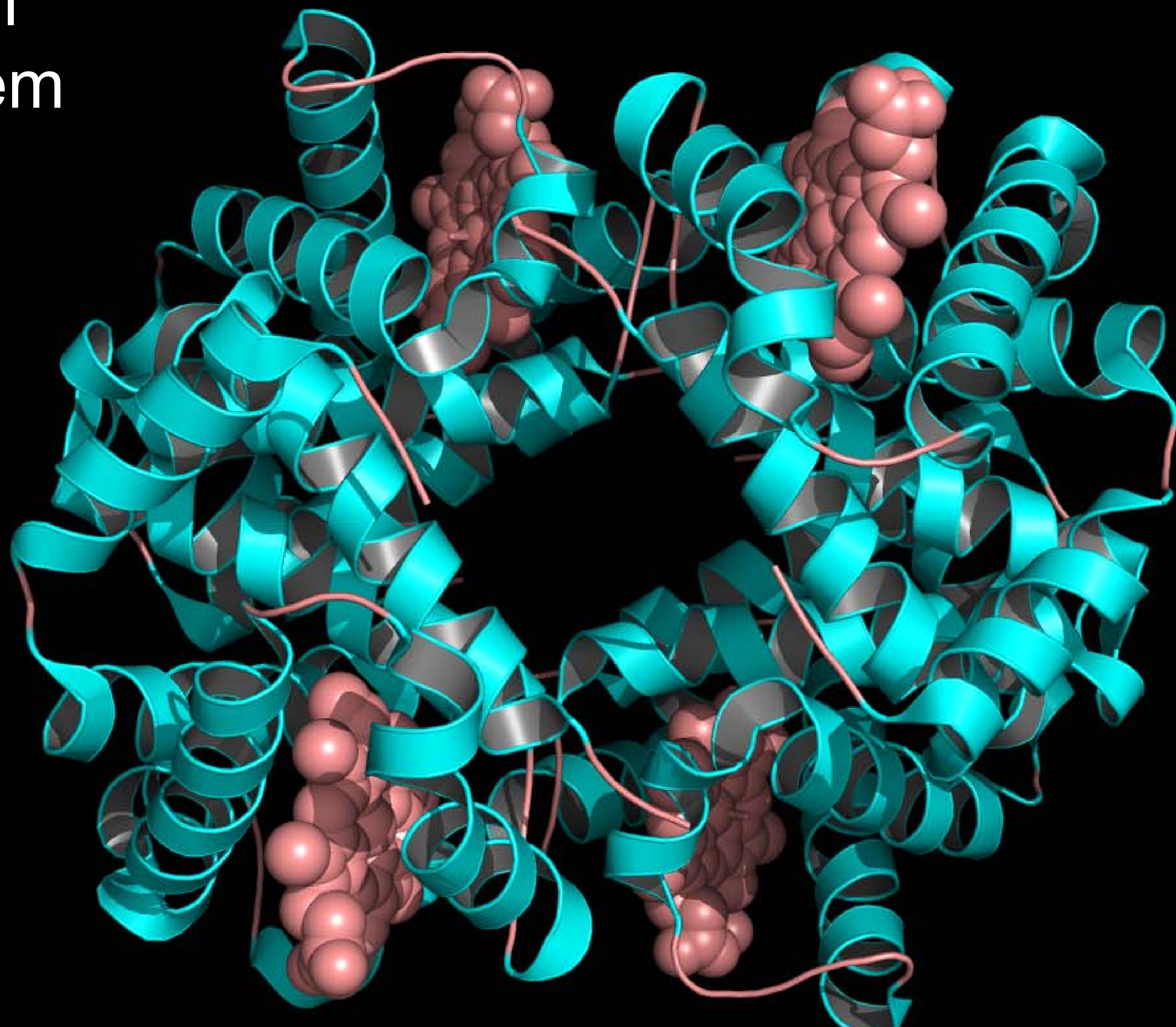
Disorder

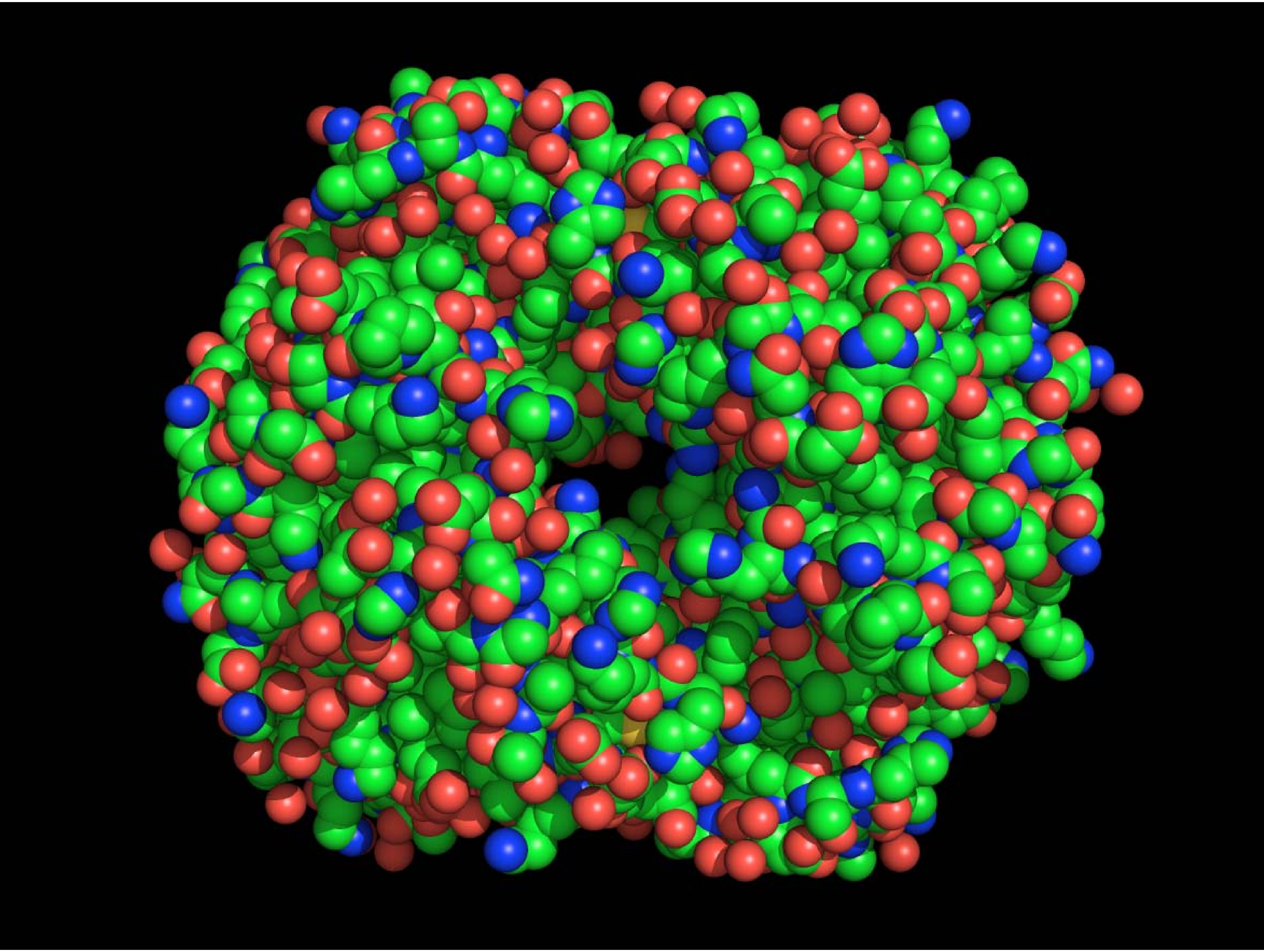
Disorder

protein dynamics

Two proofs...

Proof 1
biochem





Proof 2
physics

Protein dynamics

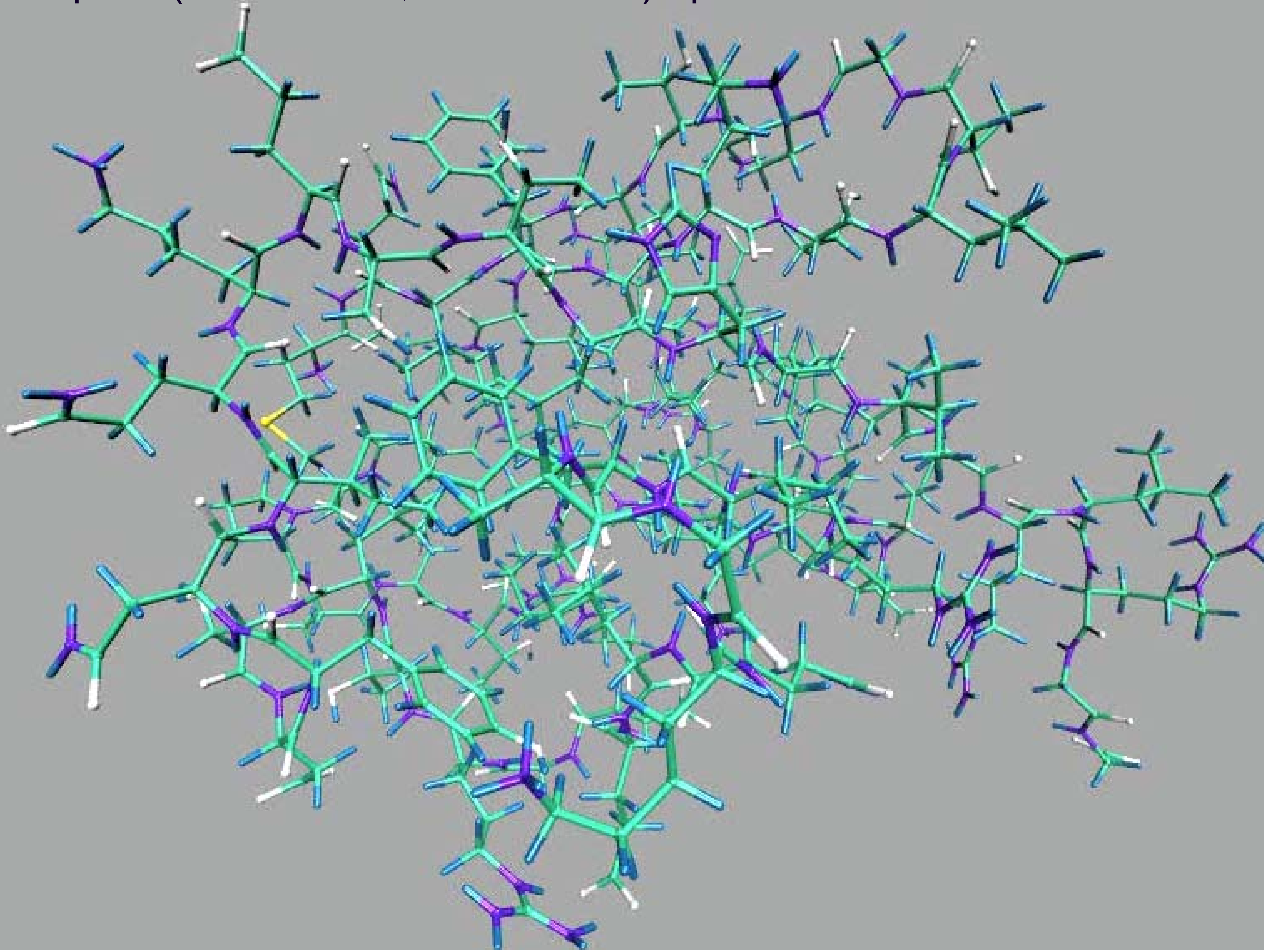
in silico

$$\vec{F}(\mathbf{r}) = -\vec{\nabla}U(\mathbf{r})$$

$$\begin{aligned} U = & \sum_{bonds\ i} k_i^{bond}(r_i - r_{0i})^2 \\ & + \sum_{angles\ i} k_i^{angle}(\theta_i - \theta_{0i})^2 \\ & + \sum_{dihedrals\ i} k_i^{dihedral}(1 + \cos(n_i\phi_i - \gamma_i)) \\ & + \sum_i \sum_{j>i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \\ & + \sum_i \sum_{j>i} \frac{q_i q_j}{4\pi\epsilon r_{ij}} \end{aligned}$$

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., et al. (2005). Scalable molecular dynamics with NAMD. *Journal of computational chemistry*, *26*(16), 1781–1802. doi:10.1002/jcc.20289

Ubiquitin (76 residues, 1231 atoms) 5ps



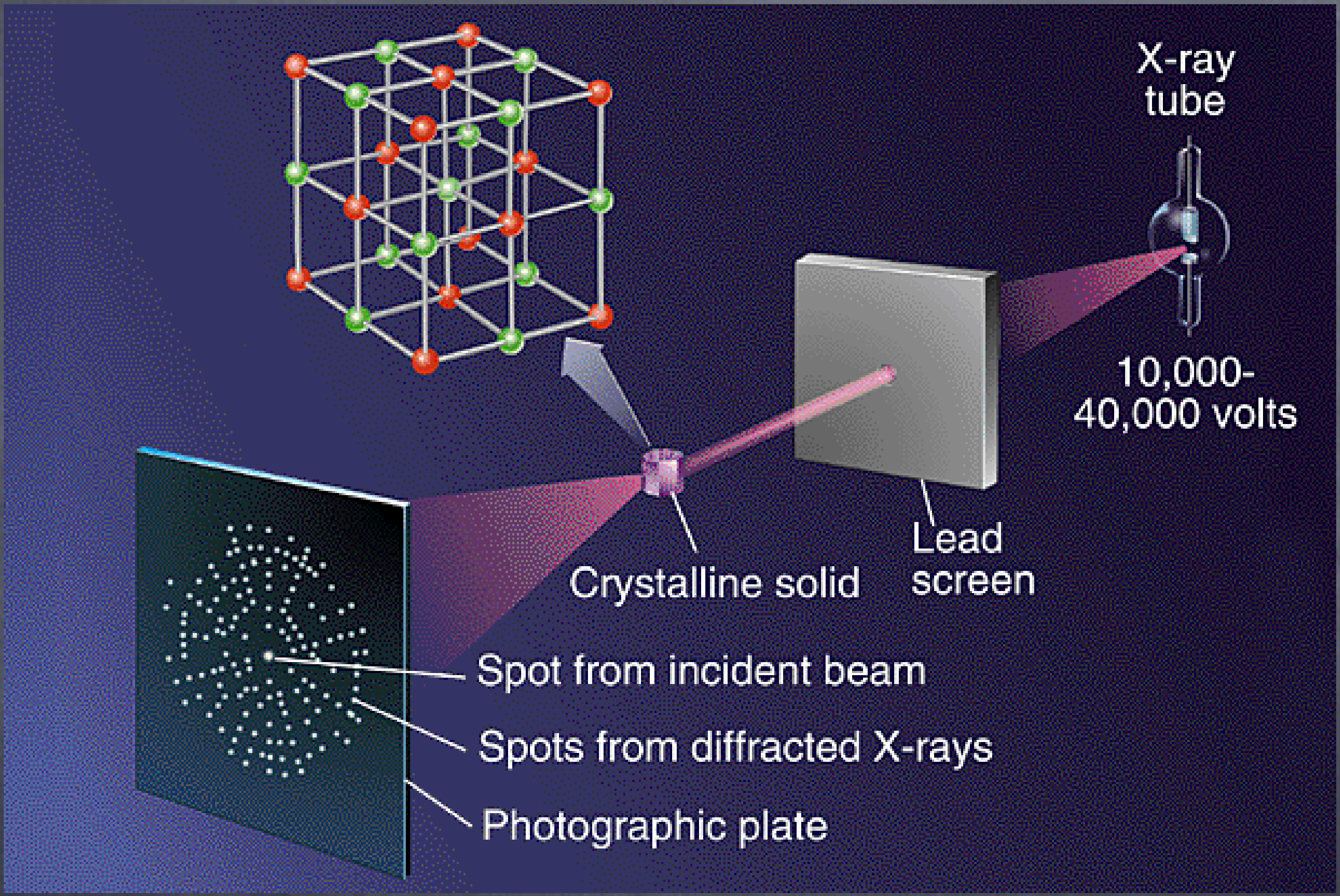
QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

Protein dynamics

in crystallum

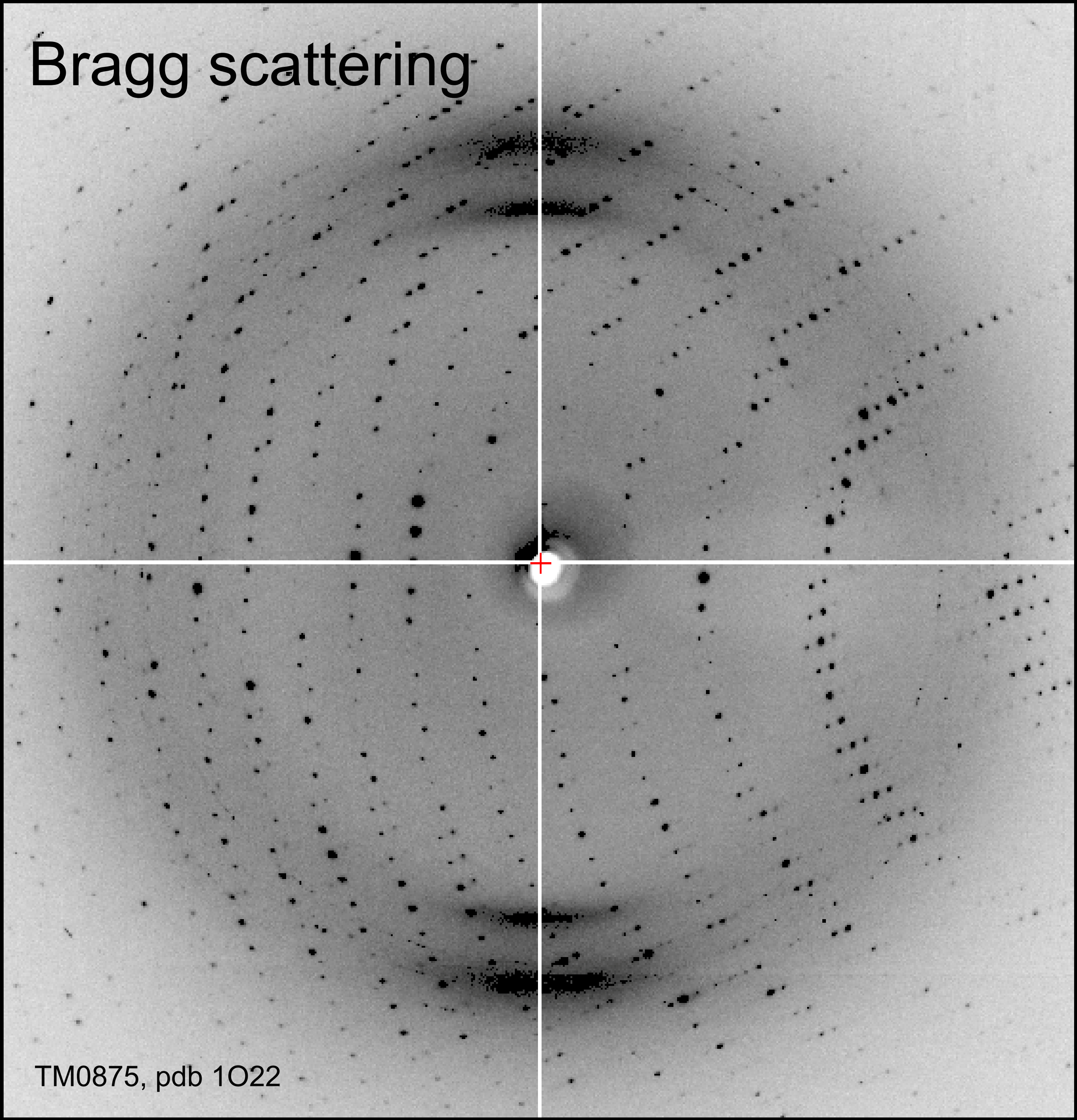
Protein disorder via diffuse X-ray scattering

(irony... study disorder by imposing crystalline order)



PJ Brucat, University of Florida

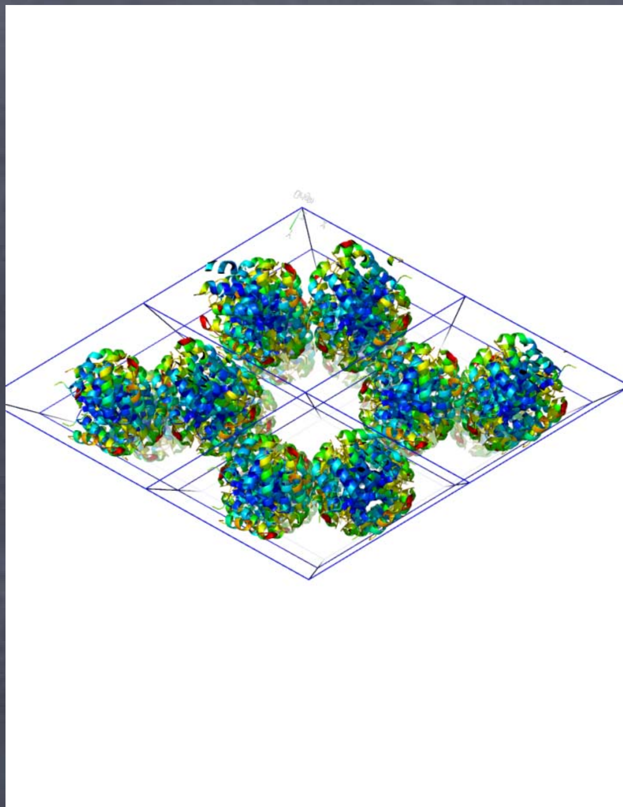
Bragg scattering

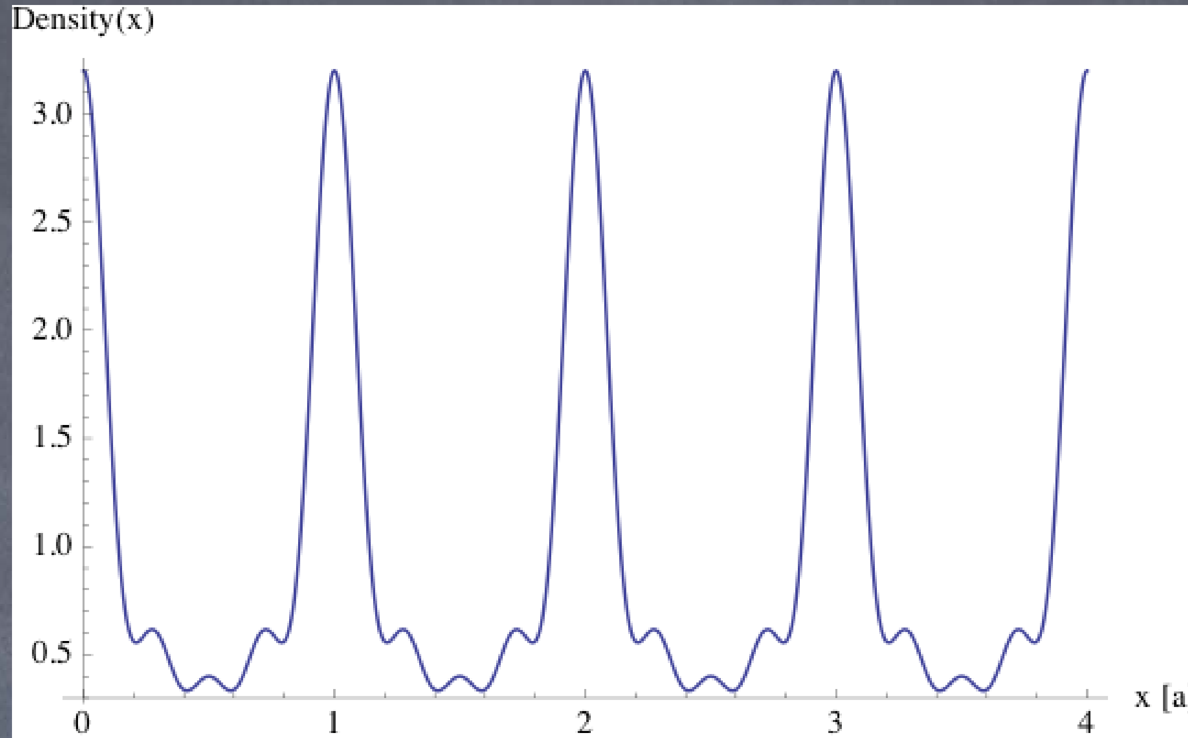


Bragg Scattering

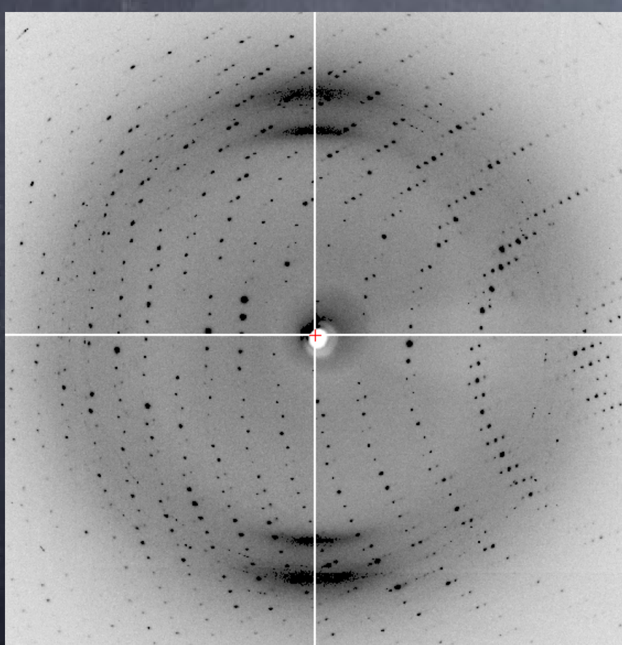
$$F(\mathbf{S}) = \int_{\infty}^{\infty} \rho(\mathbf{r}) e^{2\pi i \mathbf{S}\cdot\mathbf{r}} d\mathbf{r} = \mathcal{F}[\rho](\mathbf{S})$$

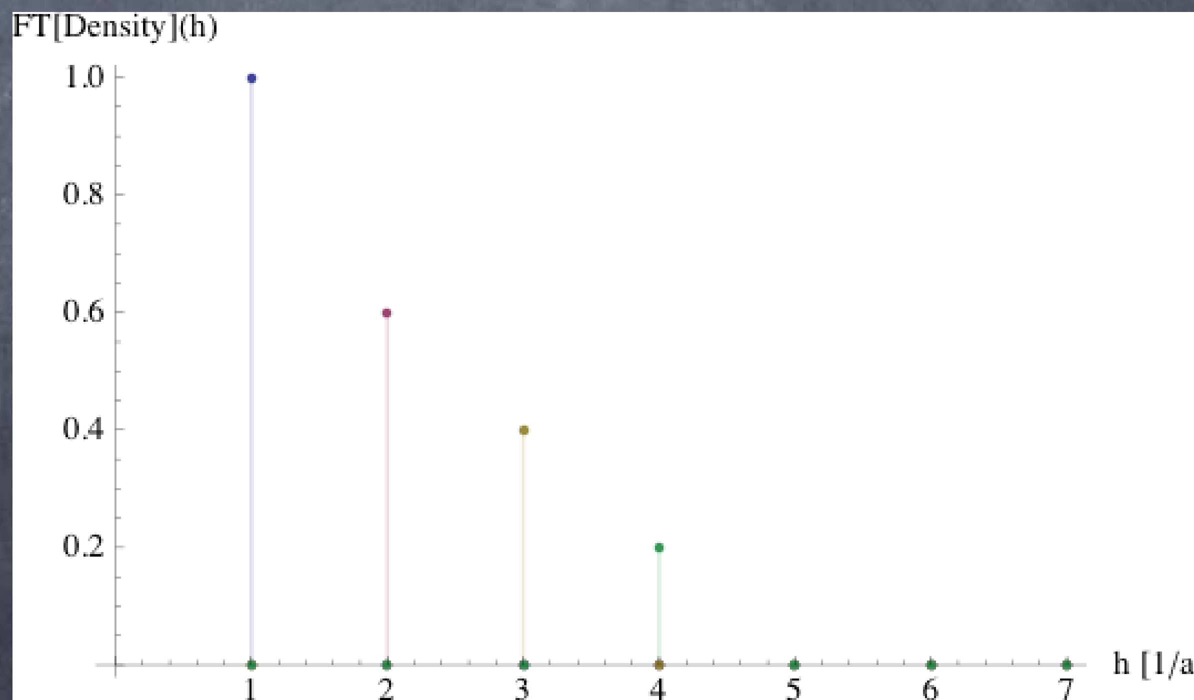
$$I(\mathbf{S}) = |F(\mathbf{S})|^2$$



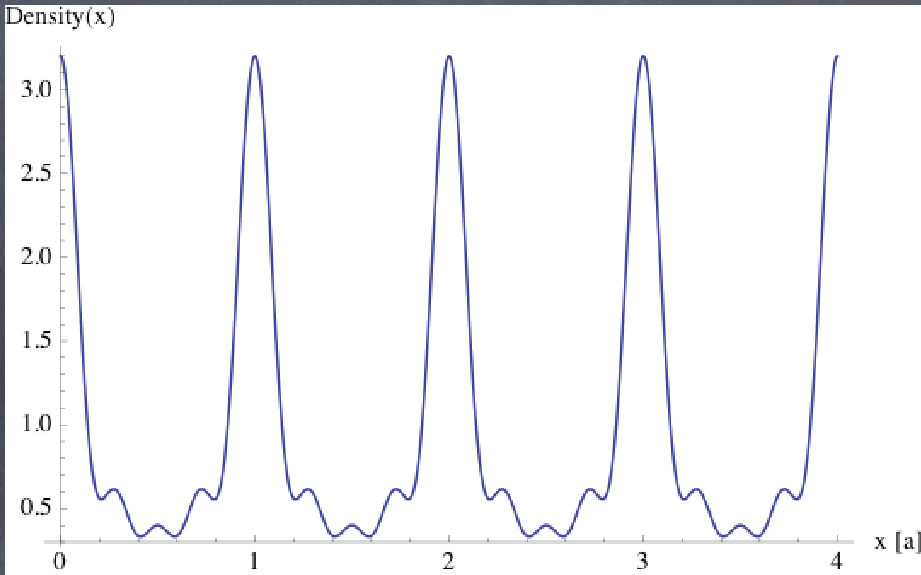


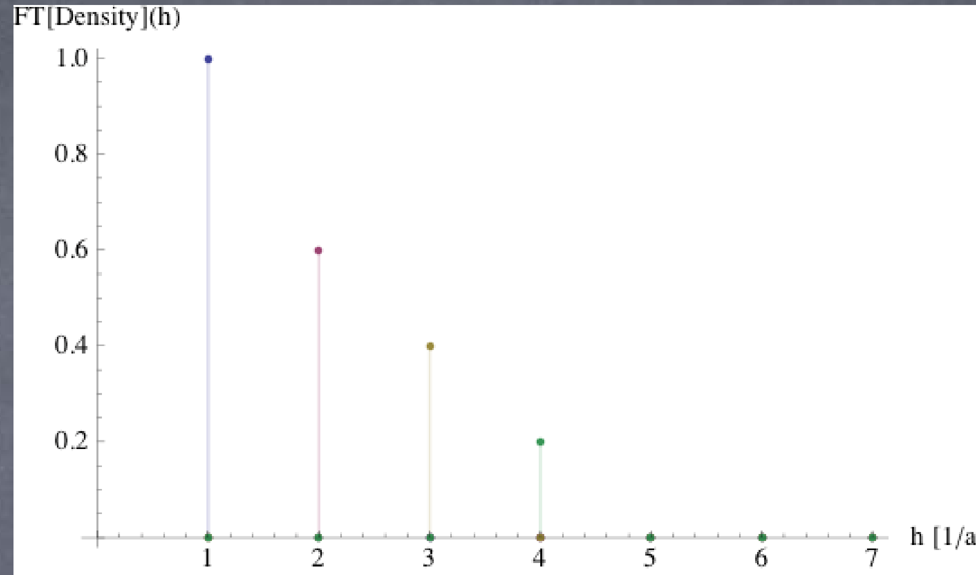
Density(x)
3.0
2.5
2.0
1.5
1.0
0.5
0
1
2
3
4
x [a]

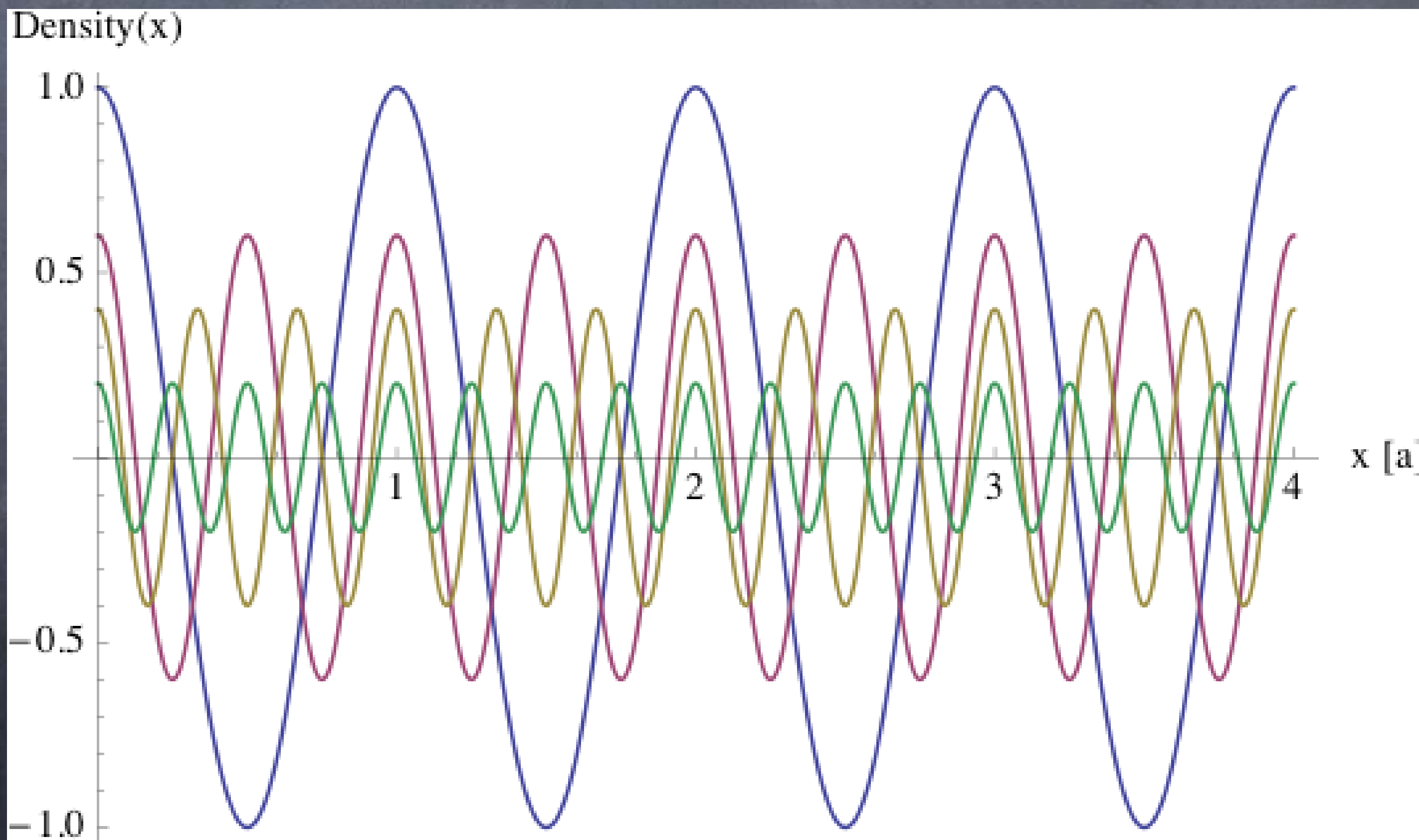


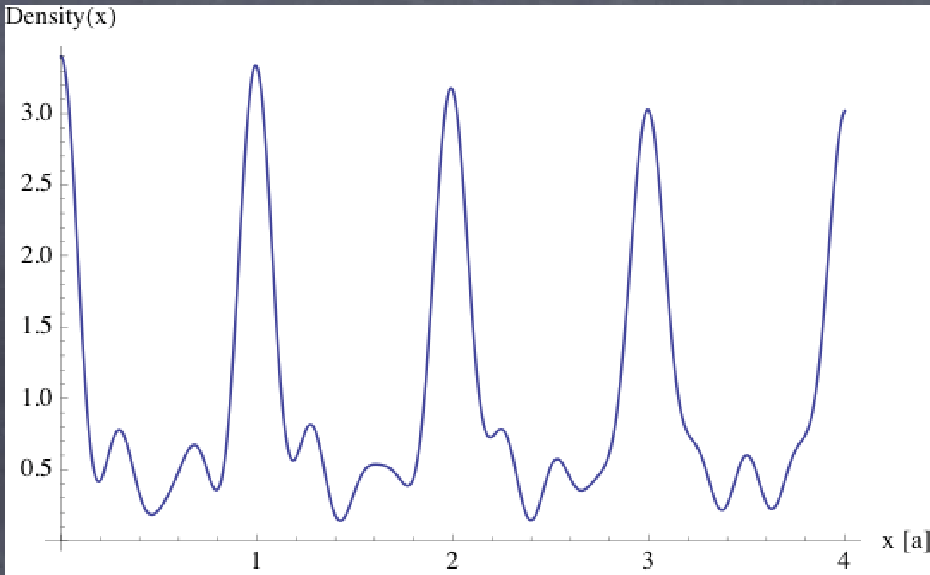


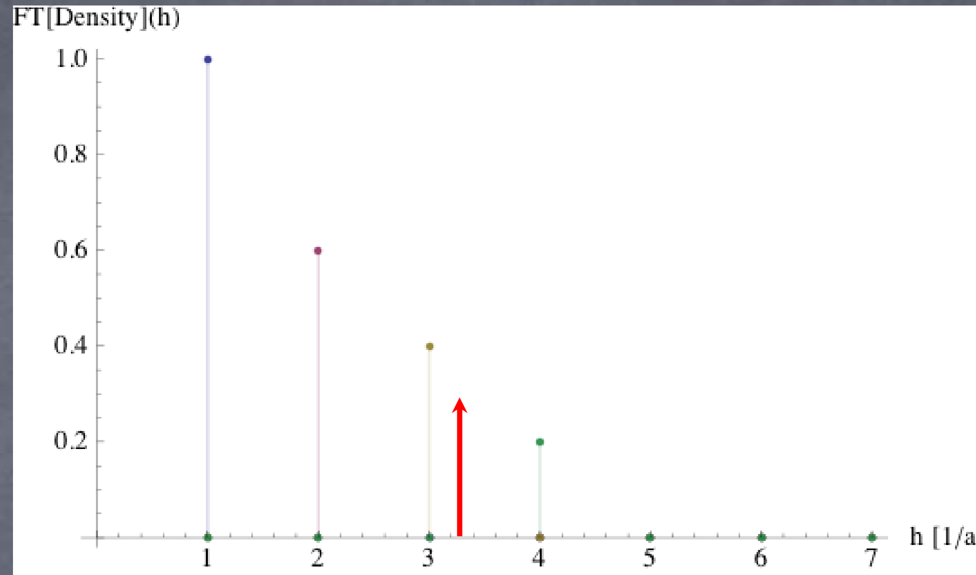
FT[Density](h)
1.0
0.8
0.6
0.4
0.2
1
2
3
4
5
6
7
h [1/a]

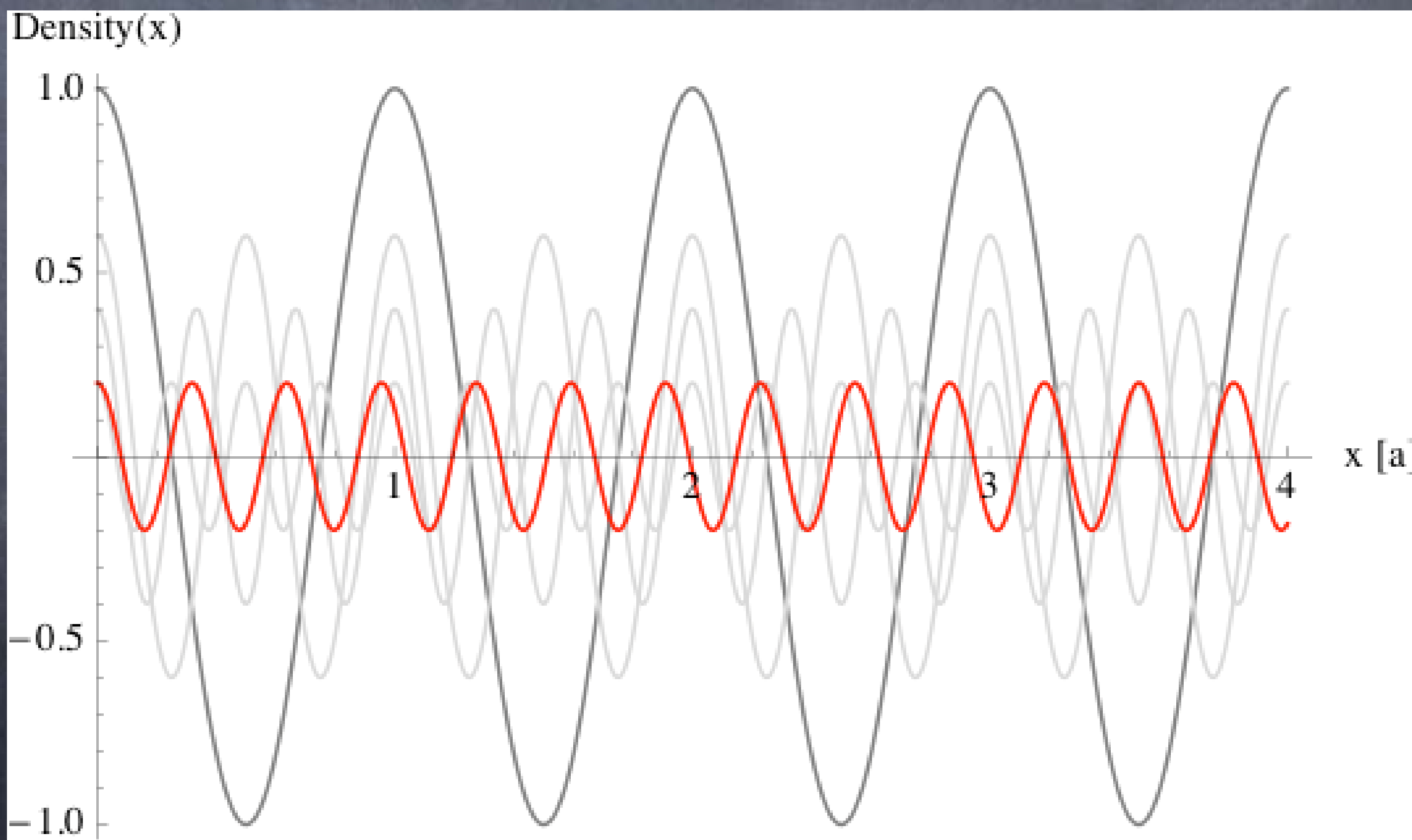

Density(x)
3.0
2.5
2.0
1.5
1.0
0.5
0
1
2
3
4
x [a]


FT[Density](h)
1.0
0.8
0.6
0.4
0.2
1
2
3
4
5
6
7
h [1/a]

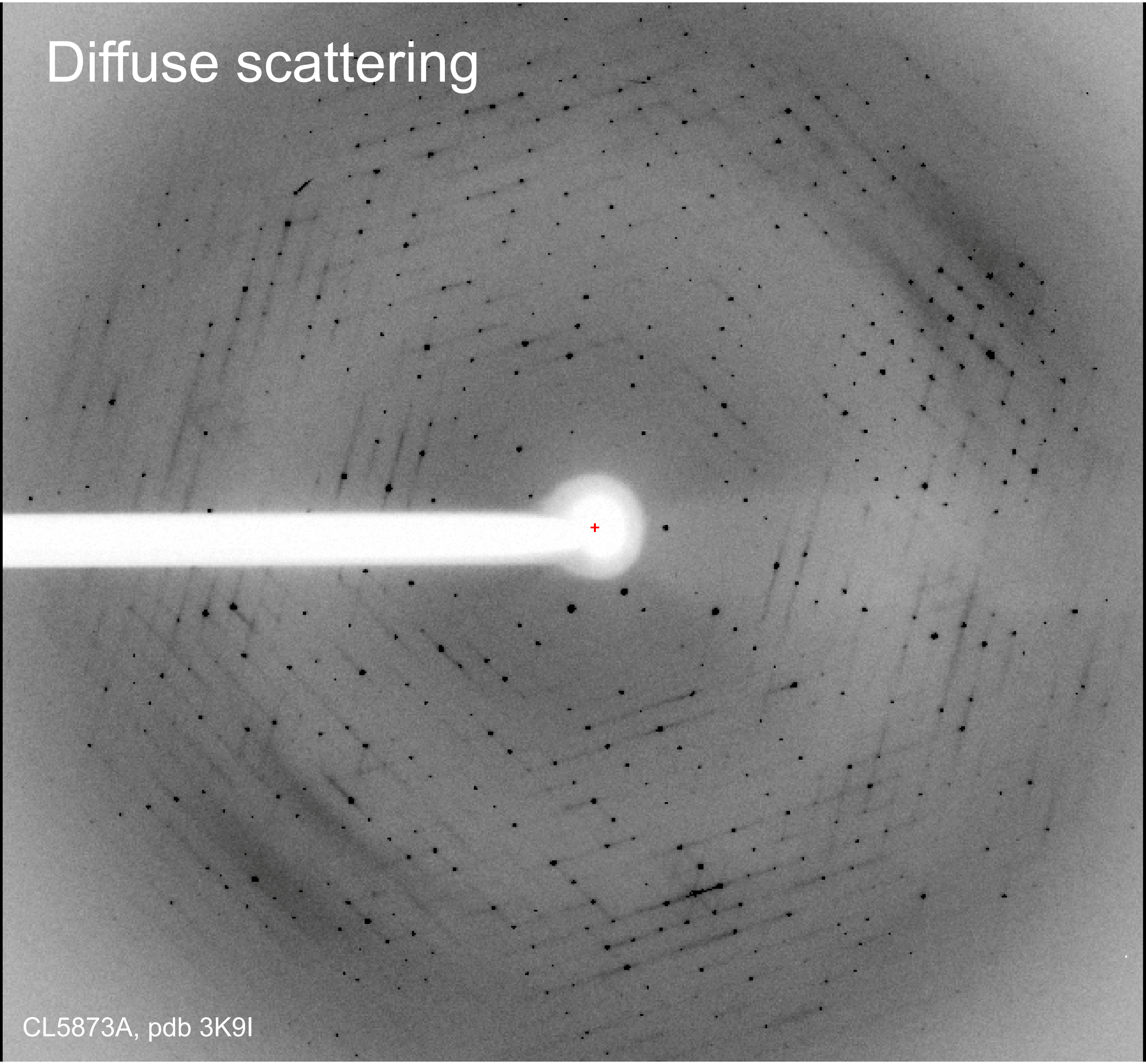

Density(x)
1.0
0.5
−0.5
−1.0
1
2
3
4
x [a]


Density(x)
3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
x [a]


FT[Density](h)
1.0
0.8
0.6
0.4
0.2
1
2
3
4
5
6
7
h [1/a]


Density(x)
1.0
0.5
−0.5
−1.0
1
2
3
4
x [a]

Diffuse scattering



Isn’t the fuzzy photo
often exactly what we need?

- Ludwig Wittgenstein (1953)

non-Bragg "diffuse" Scattering

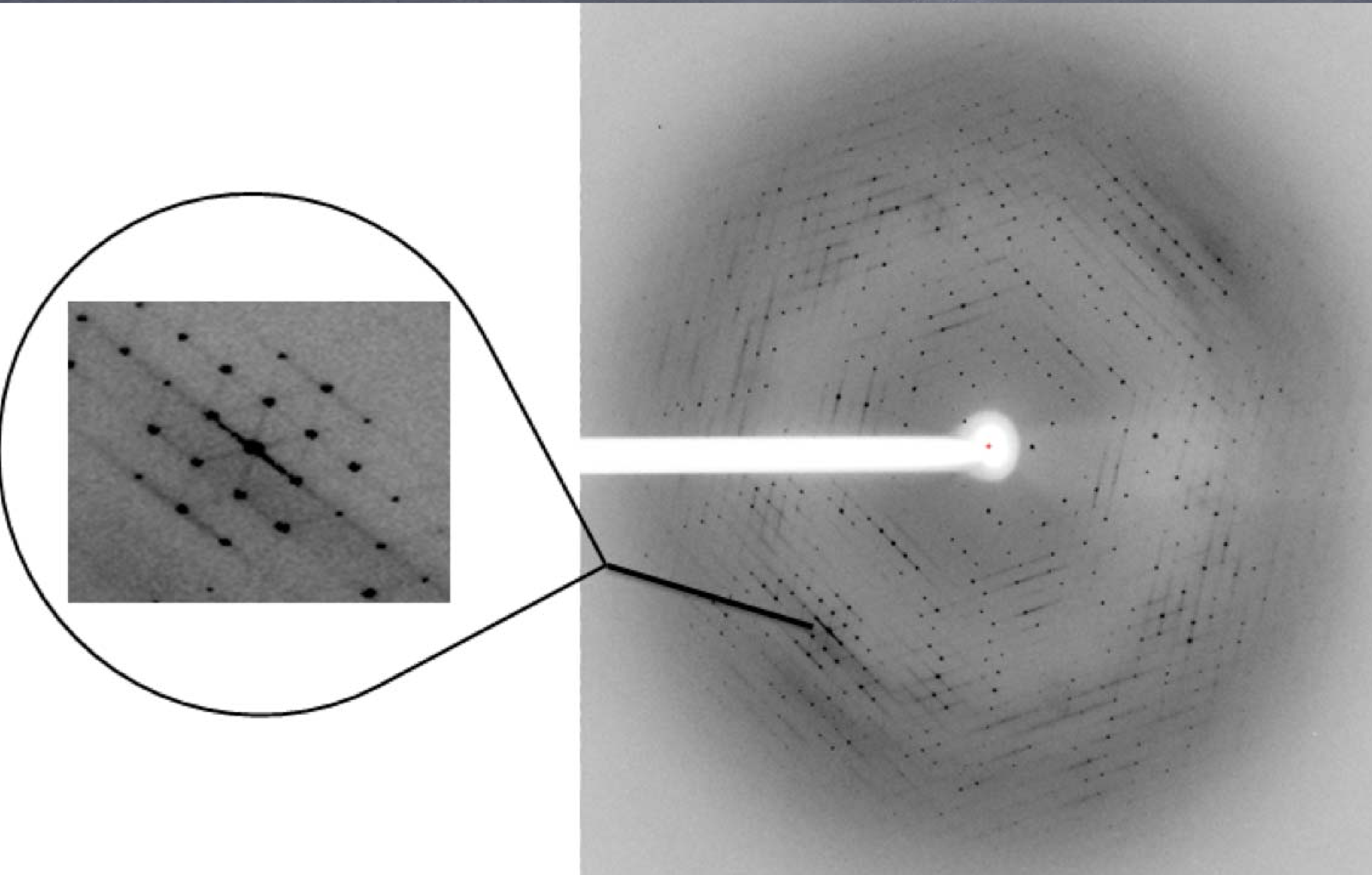
$$F(\mathbf{S}) = \int_{\infty}^{\infty} \rho(\mathbf{r}) e^{2\pi i \mathbf{S}\cdot\mathbf{r}} d\mathbf{r} = \mathcal{F}[\rho](\mathbf{S})$$

$$\longrightarrow \quad \mathbf{r} = \langle \mathbf{r} \rangle + \delta(\mathbf{r})$$

$$I(\mathbf{S}) = |F(\mathbf{S})|^2$$

$$\begin{aligned} I(\mathbf{S}) &\approx e^{-(2\pi \mathbf{S}\cdot\delta)^2} |F_o(\mathbf{S})|^2 \\ &+ e^{-(2\pi \mathbf{S}\cdot\delta)^2} (2\pi \mathbf{S} \cdot \delta)^2 \left\{ |F_o(\mathbf{S})|^2 * \mathcal{F}[\Gamma] \right\} \\ &+ \mathcal{O}(\delta^4) \end{aligned}$$

$$\Gamma(r') = \frac{\langle \delta(r)\delta(r + r') \rangle_r}{\langle \delta^2(r) \rangle_r}$$



CL5873A, pdb 3K9I

Atomic displacements are "liquid-like"
exponentially correlated with correlation length(s) γ

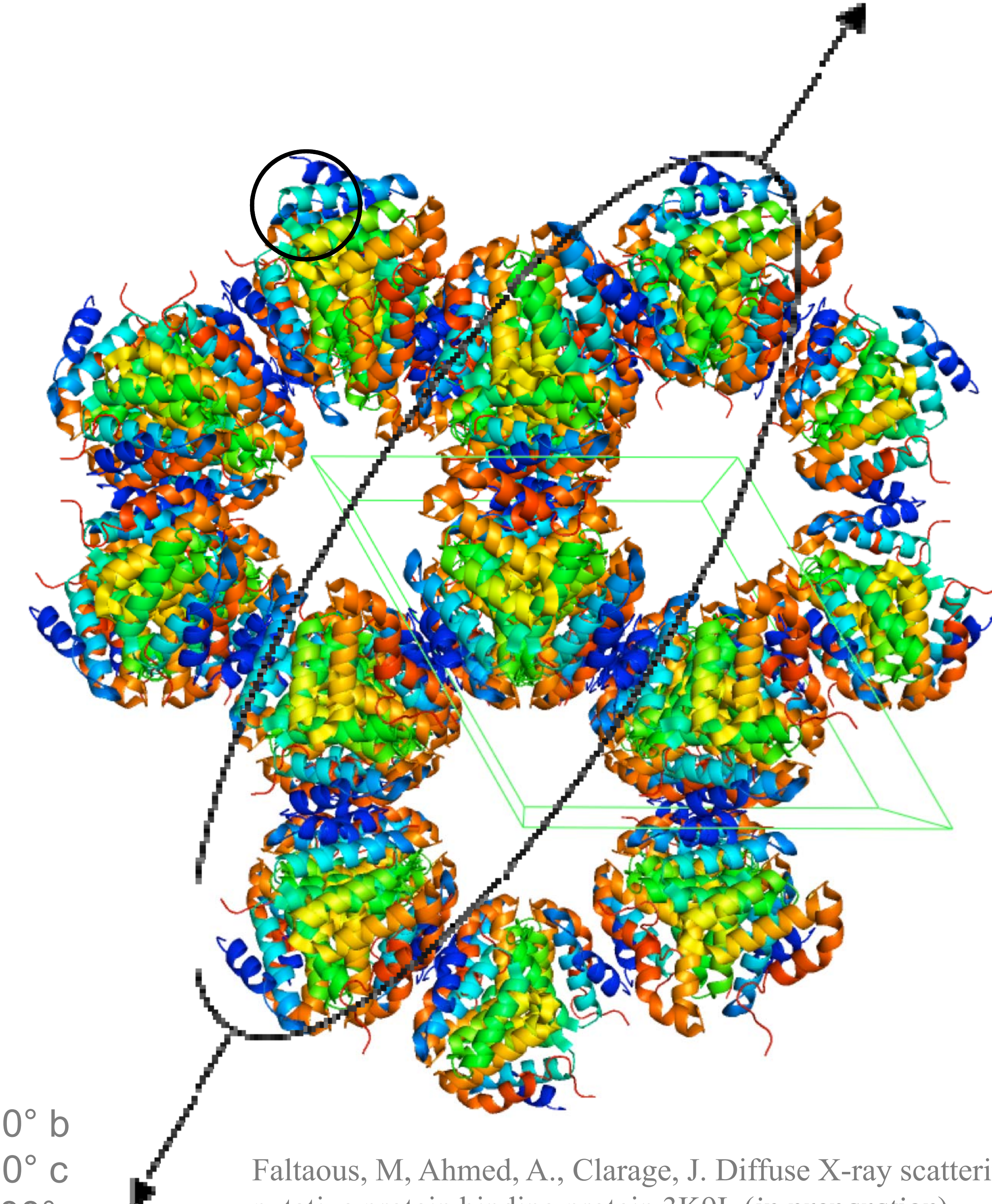
$$\Gamma(r') = \frac{\langle \delta(r)\delta(r + r') \rangle_r}{\langle \delta^2(r) \rangle_r}$$

$$\Gamma(r) = e^{-r/\gamma}$$ (Markov?)

$\gamma_{internal}$ = 6-8 Å

$\gamma_1 = \gamma_2$ = 64 Å

γ_3 = 684 Å



3K9I Space Group: $P6_122$

a = 73.66 Å α = 90.00° b

= 73.66 Å β = 90.00° c

= 115.84 Å γ = 120.00°

Faltaous, M, Ahmed, A., Clarage, J. Diffuse X-ray scattering from a putative protein binding protein 3K9I, (*in preparation*)

Chaos

Laplace (1776)

vs.

Poincare (1903)

Lyapunov exponent λ

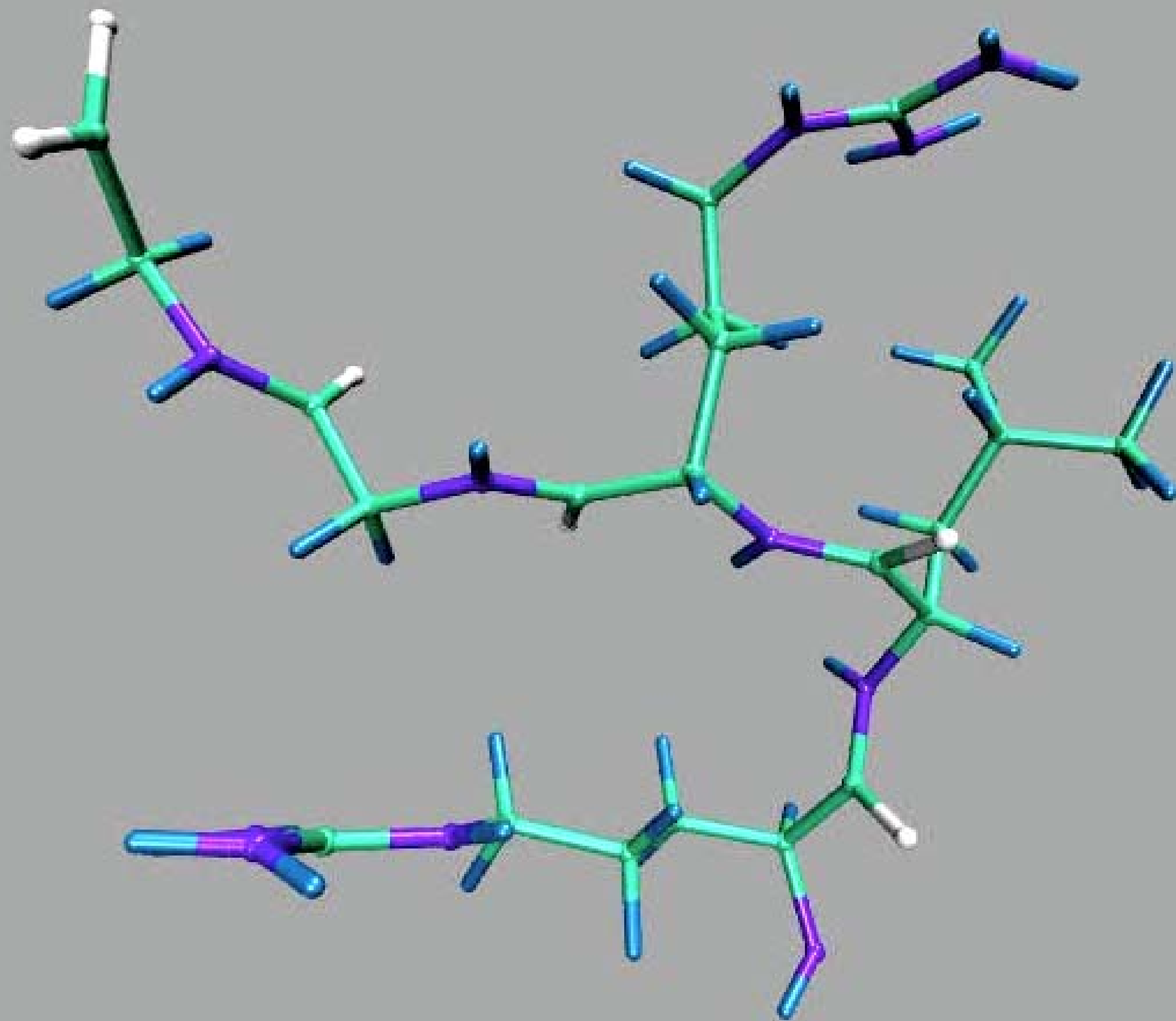
$$\delta(t) = \delta_o e^{\lambda t}$$

$\delta_o = 0.001$ Å

$$
\begin{aligned}
U = & \sum_{bonds\ i} k_i^{bond}(r_i - r_{0i})^2 \\
& + \sum_{angles\ i} k_i^{angle}(\theta_i - \theta_{0i})^2 \\
& + \sum_{dihedrals\ i} k_i^{dihedral}(1 + \cos(n_i\phi_i - \gamma_i)) \\
& + \sum_i \sum_{j>i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \\
& + \sum_i \sum_{j>i} \frac{q_i q_j}{4\pi\epsilon r_{ij}}
\end{aligned}
$$

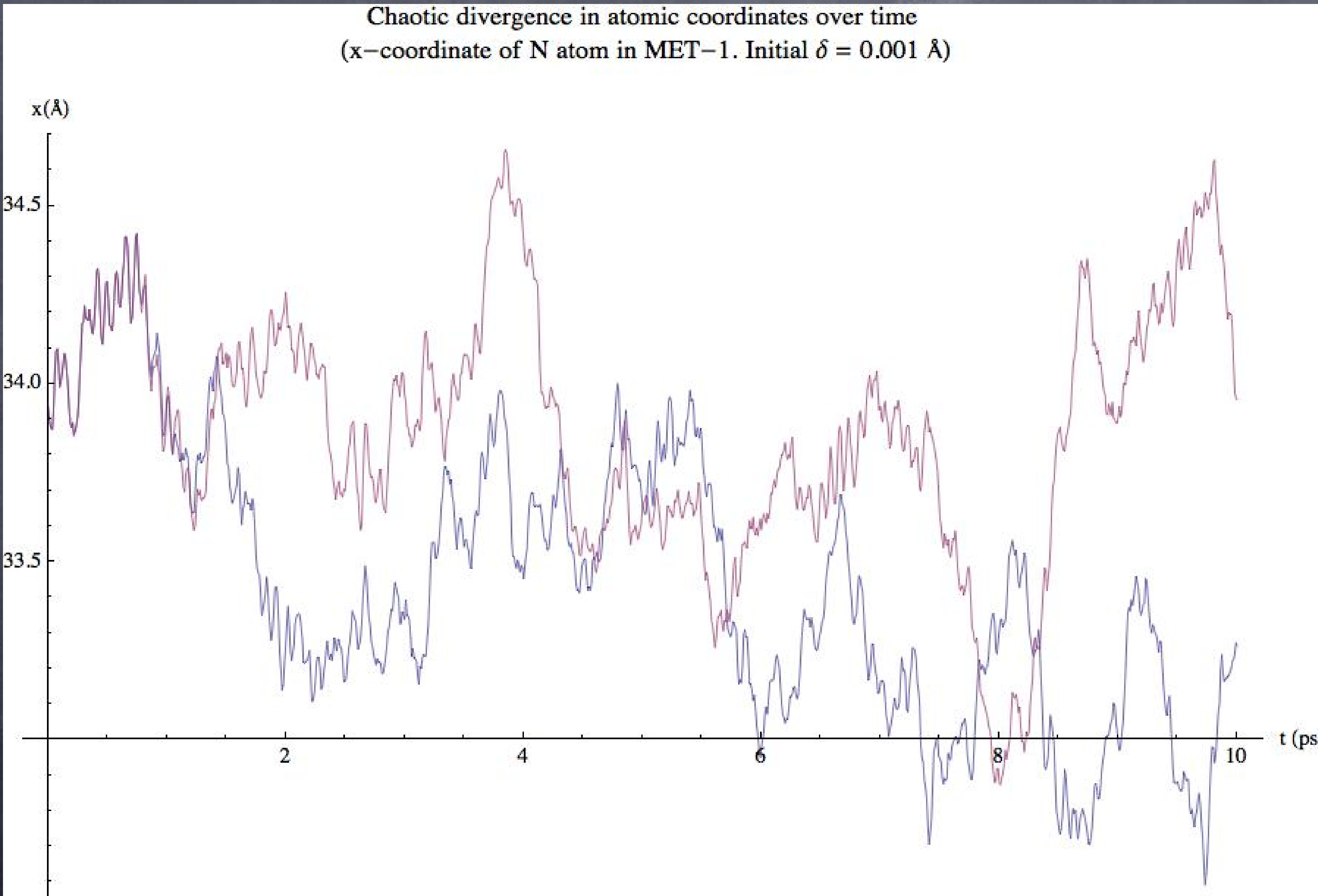
Ubiquitin (76 residues, 1231 atoms) 5ps

QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

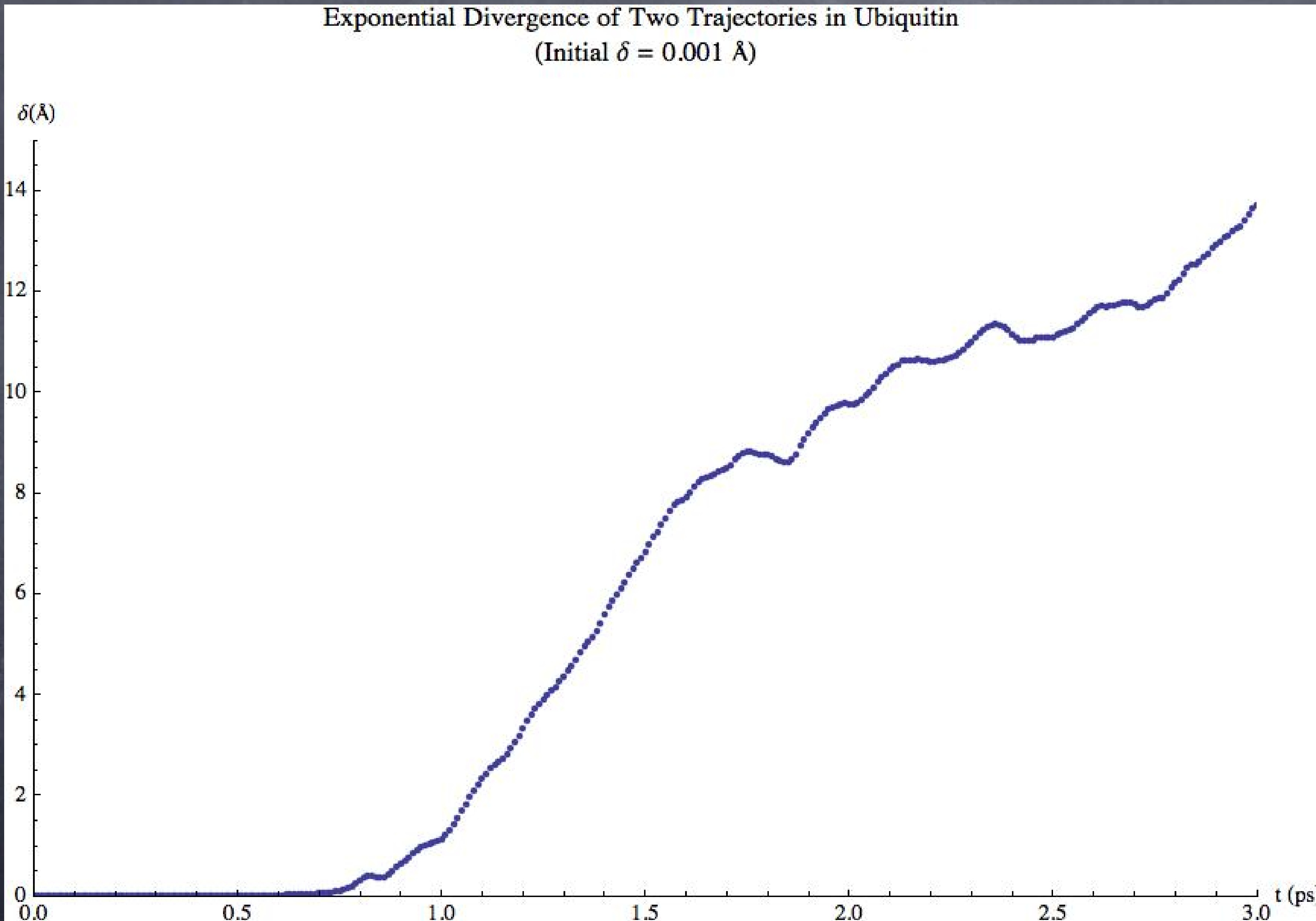


N-terminus 72-76

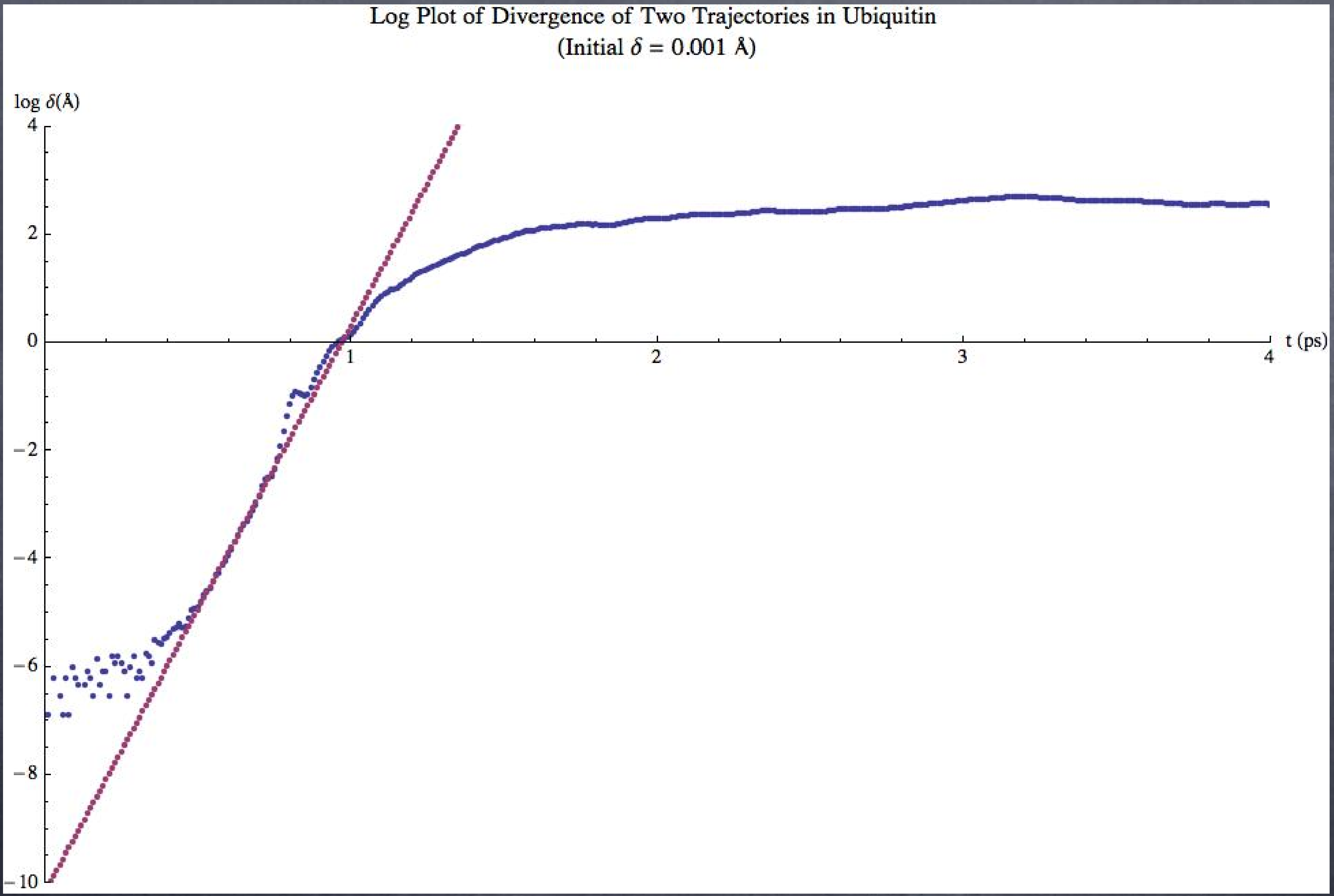
QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.


Chaotic divergence in atomic coordinates over time
(x−coordinate of N atom in MET−1. Initial δ = 0.001 Å)
x(Å)
34.5
34.0
33.5
2
4
6
8
10
t (ps)

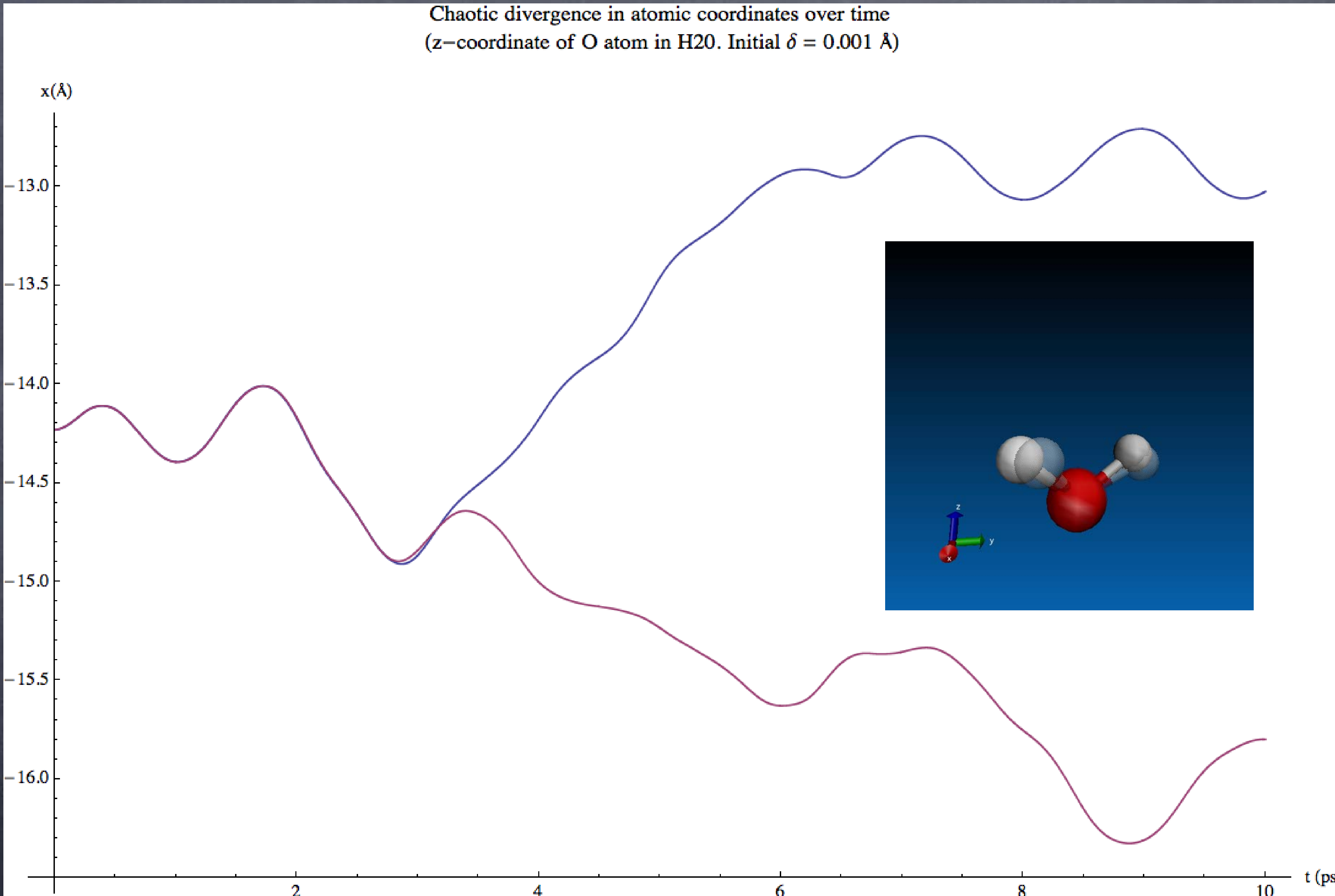
$$\mathbf{r}_{3N}(t) = \begin{pmatrix} x_1(t) \\ y_1(t) \\ z_1(t) \\ x_2(t) \\ y_2(t) \\ z_2(t) \\ \cdot \\ \cdot \\ \cdot \\ x_N(t) \\ y_N(t) \\ z_N(t) \end{pmatrix}$$



Exponential Divergence of Two Trajectories in Ubiquitin
(Initial δ = 0.001 Å)
δ(Å)
14
12
10
8
6
4
2
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
t (ps)



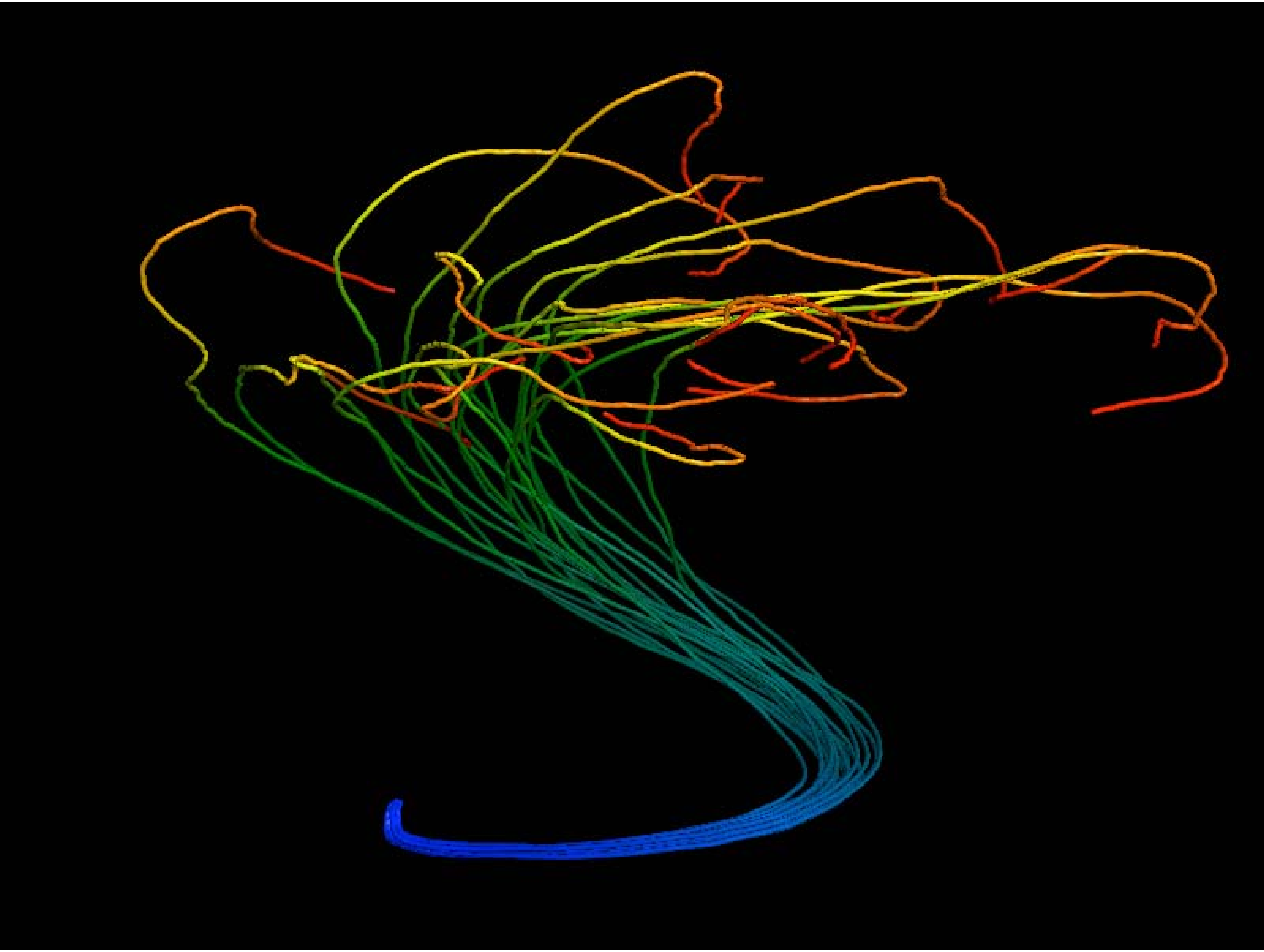
$\lambda = 9.4\ ps^{-1} = 1/(0.11\ ps)$

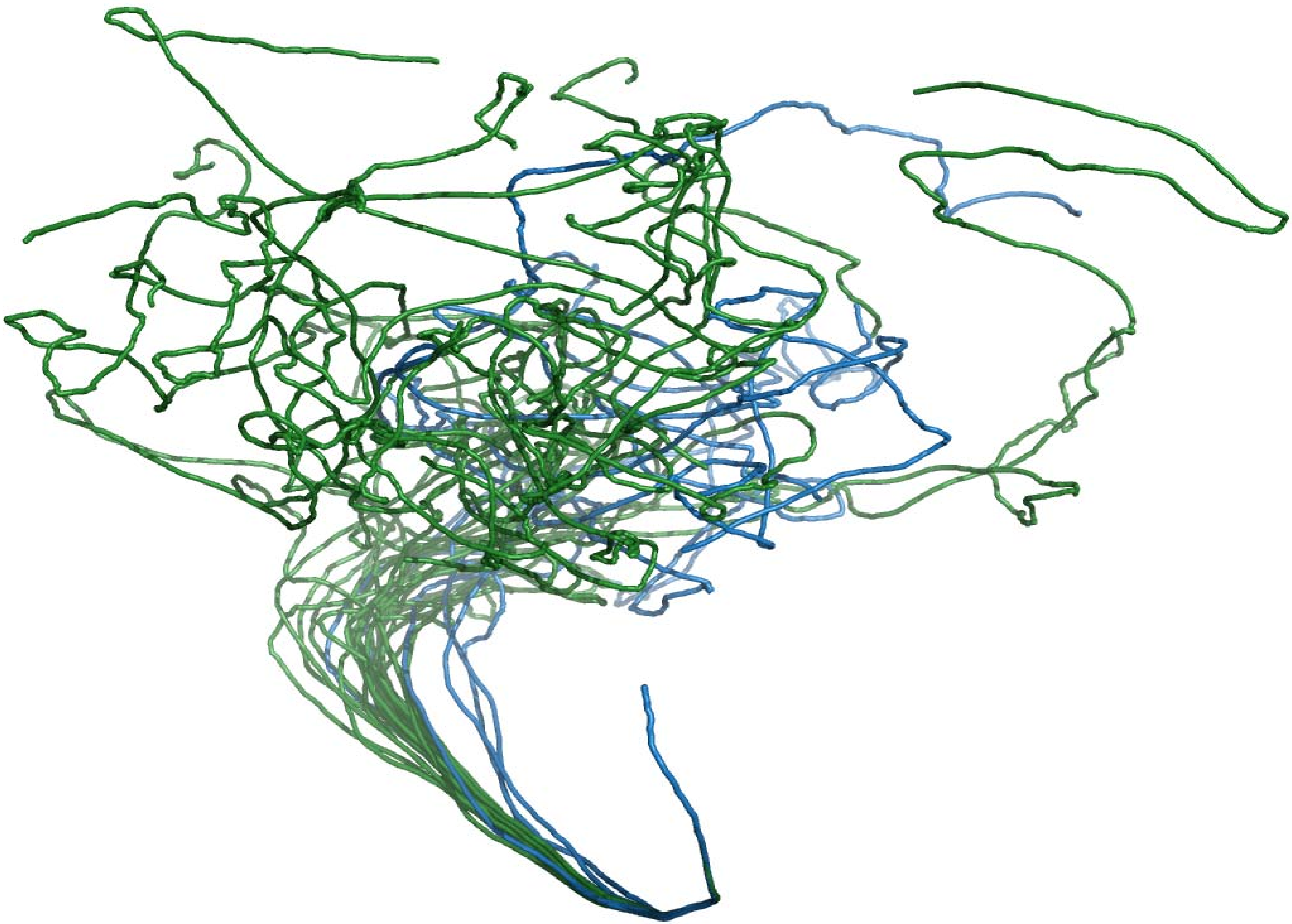


$\lambda = 2.25\ ps^{-1} = 1/(0.44\ ps)$

State space portraits

$$\mathbf{r}_{3N}(t) = \begin{pmatrix} x_1(t) \\ y_1(t) \\ z_1(t) \\ x_2(t) \\ y_2(t) \\ z_2(t) \\ \cdot \\ \cdot \\ \cdot \\ x_N(t) \\ y_N(t) \\ z_N(t) \end{pmatrix}$$





Bui, J., Romo, T., Clarage, J. Visualizing chaos in proteins using state-space portraits. (*in preparation*)

- Mary Faltaous,
  James Bui,
  Anam Ahmed,
  Joseph Bedell,
  Jennifer Doan,
  Brian Kneeland,
  Lukas Caras

- Tod Romo,
  University of Rochester Medical Center

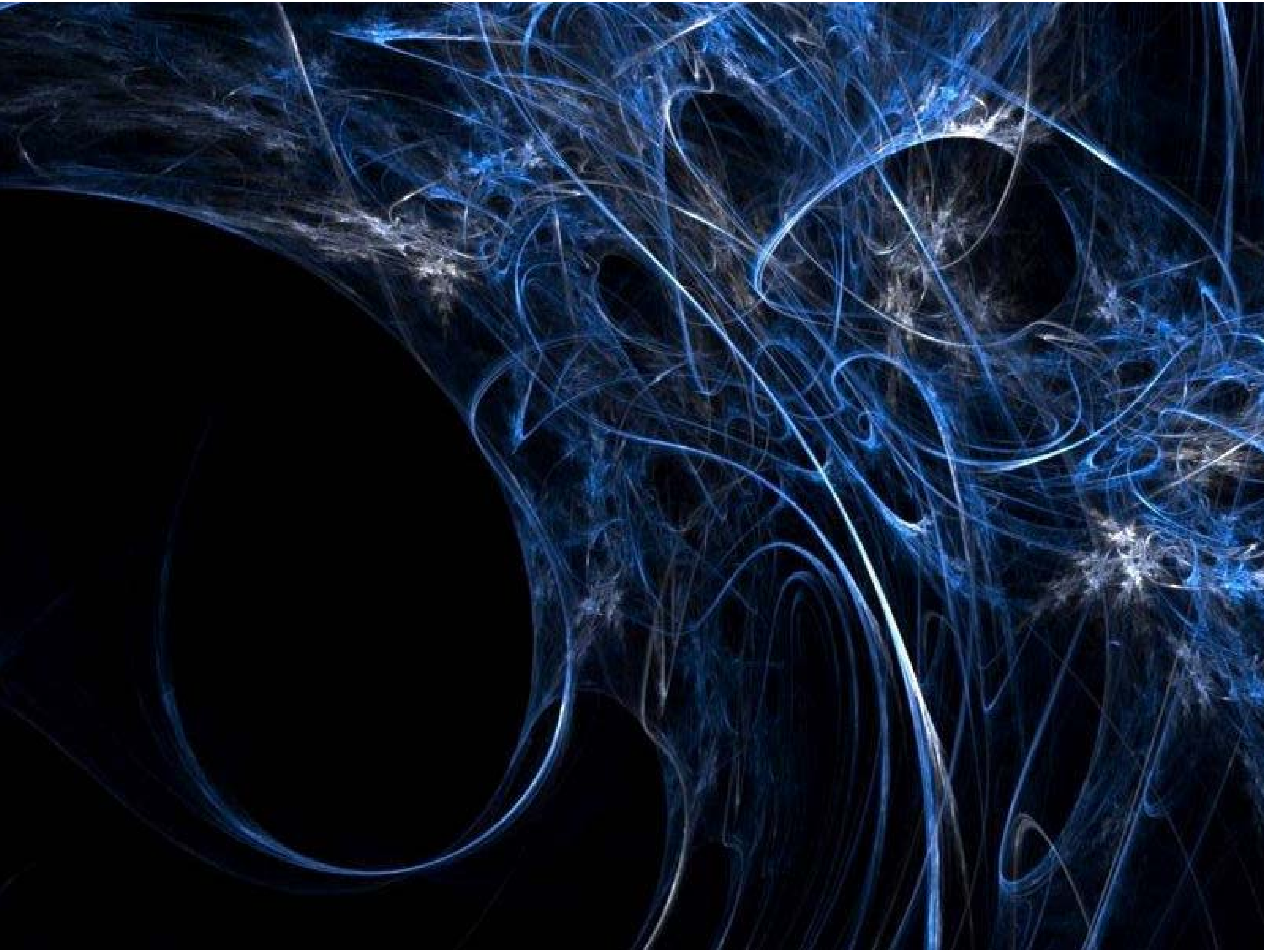
- Adam Godzik,
  Joint Center for Structural Genomics

misc...

non-Ergodicity

Structural: only small fraction of all possible protein sequences sampled in biosphere (Kauffman).

Dynamical: protein lifetimes less than needed to fill out attractor



$$F(\mathbf{S}) = \int_{\infty}^{\infty} \rho(\mathbf{r}) e^{2\pi i \mathbf{S}\cdot\mathbf{r}} d\mathbf{r} = \mathcal{F}[\rho](\mathbf{S})$$

$$I(\mathbf{S}) = |F(\mathbf{S})|^2$$

$$\mathbf{r} = \langle \mathbf{r} \rangle + \delta(\mathbf{r})$$

$$\begin{aligned} I(\mathbf{S}) &\approx e^{-(2\pi\mathbf{S}\cdot\delta)^2} |F_o(\mathbf{S})|^2 \\ &+ e^{-(2\pi\mathbf{S}\cdot\delta)^2} (2\pi\mathbf{S}\cdot\delta)^2 \left\{ |F_o(\mathbf{S})|^2 * \mathcal{F}[\Gamma] \right\} \\ &+ \mathcal{O}(\delta^4) \end{aligned}$$

$$\Gamma(r') = \frac{\langle \delta(r)\delta(r+r') \rangle_r}{\langle \delta^2(r) \rangle_r}$$