Fibre (fiber) diffraction and Macro-molecular fiber crystallography

Joseph Orgel[1,2] and Tom Irving[1,2]

[1]Dept. BCPS, Illinois Institute of Technology
[2]BioCAT APS, Argonne National Lab.

Scope of Lecture

- Why do Fiber Diffraction?
- Physical Principles
- Experimental Methods
- Applications
- Advantages of / Need for Third Generation Synchrotrons for Fiber Diffraction

References

- **Basics**:

  C. Cantor and P. Schimmel "Biophysical Chemistry part II: Techniques for the study of Biological Structure and Function" Chapter 14. Freeman, 1980
- **A terrific introduction**:

  John Squire "The Structural Basis of Muscular Contraction" Plenum, 1981
- **Definitive Reference**:

  B.K. Vainshtein "Diffraction of X-rays by Chain Molecules" Elsevier, 1966.

More references:

Good introduction to Fiber crystallography :

Chandrasekaran, R. and Stubbs, G. (2001). Fiber diffraction. in *International Tables for Crystallography, Vol. F: Crystallography of Biological Macromolecules* (Rossman, M.G. and Arnold, E., eds.), Kluwer Academic Publishers, The Netherlands, 444-450.

Twenty Years Hard Labor as a Fiber Diffractionist

STRUTHER ARNOTT

Department of Biological Sciences, Purdue University, West Lafayette, IN 47907

X-ray diffraction can be used to help determine the molecular geometry of polymers that prefer to be long helices rather than more complexly folded structures. It is usually possible to prepare specimens in which such helical molecules are aligned with their long axes parallel. Often further lateral organization occurs, but rarely to the degree of a three-dimensionally ordered single crystal. *Potentially* this is an advantage, since there is more information (about the Fourier transform (1, 2, 3) of a molecular structure) in the continuous intensity distribution in the diffraction pattern from a less well-ordered system than there is the "sampled" distribution characteristic of a single crystal. But, since "sampling" also implies local amplification of the molecular transform (at reciprocal lattice points), its absence results in much weaker diffraction signals and the theoretical advantage of knowing the continuous intensity variation is offset by the experimental difficulty of recording it accurately. A further complication is that there are a great many *kinds* of partially-ordered systems of helical molecules, each giving rise to different types of diffraction pattern in which both continuous intensity and Bragg maxima occur. If we wish quantitatively to analyze a diffraction pattern we, of course, have to succeed in modelling not only the molecular structure but also the molecular packing. This is true

0-8412-0589-2/80/47-141-001$07.50/0

Why Fiber Diffraction ?

- Atomic level structures from crystallography or NMR = "gold standard" for structural inferences

- But there is a large class of "fibrous proteins"

  e.g: Actin, myosin, intermediate filaments, microtubules, bacterial flagella, filamentous viruses, amyloid, collagenous connective tissue

- Will not crystallize but can be induced to form oriented assemblies

- Some systems *naturally* form ordered systems

Dimensional hierarchy of Biophysical (X-ray) techniques

1D — Low angle solution or powder diffraction

large macromolecular assembles

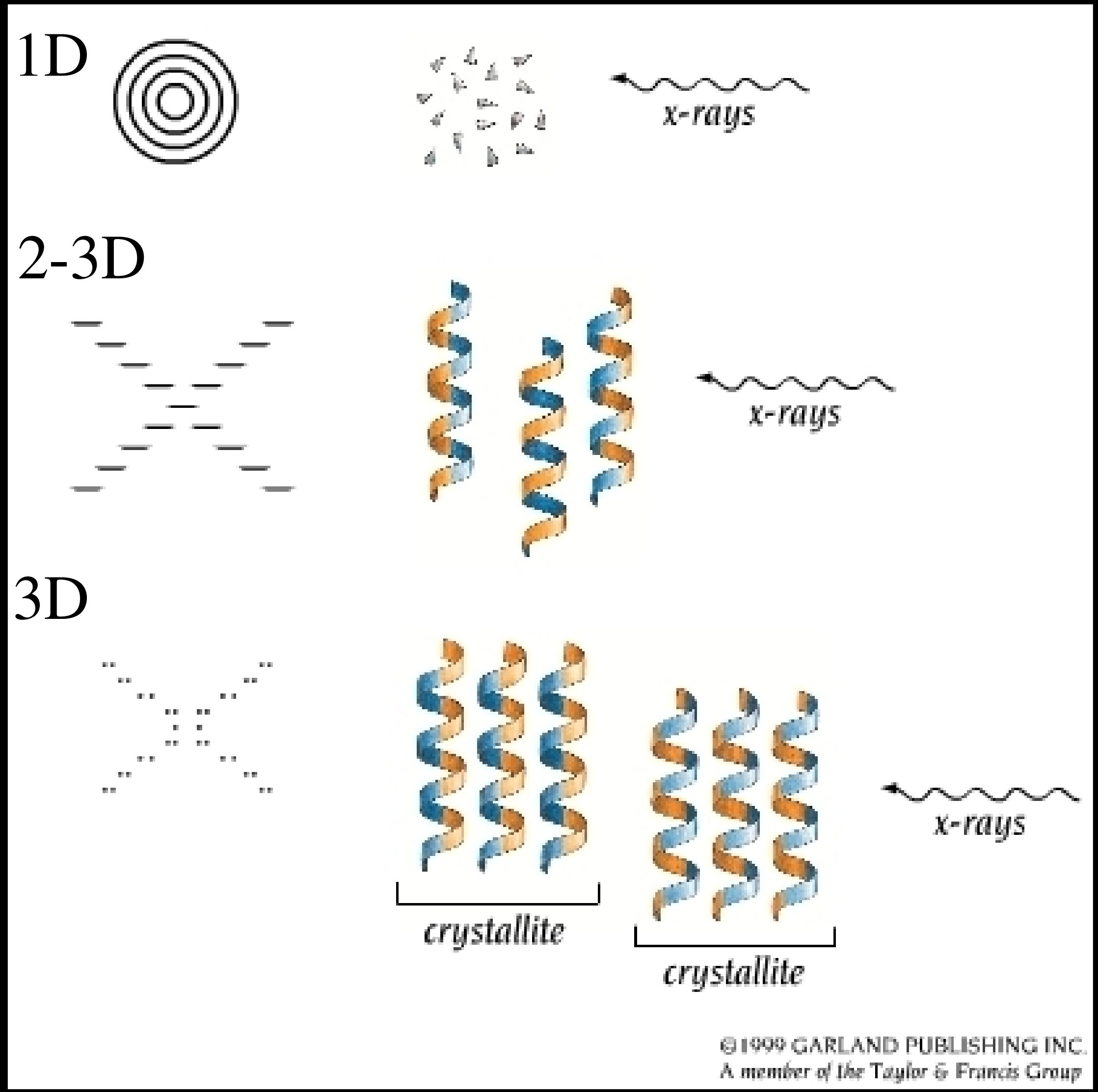
1-2-3D — Fiber diffraction

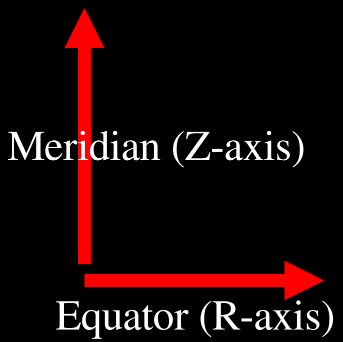
fiber forming arrays – muscle, collagen, DNA, amyloids, various carbohydrates, often *super-macromolecular scale*

3D — Single crystal X-ray diffraction

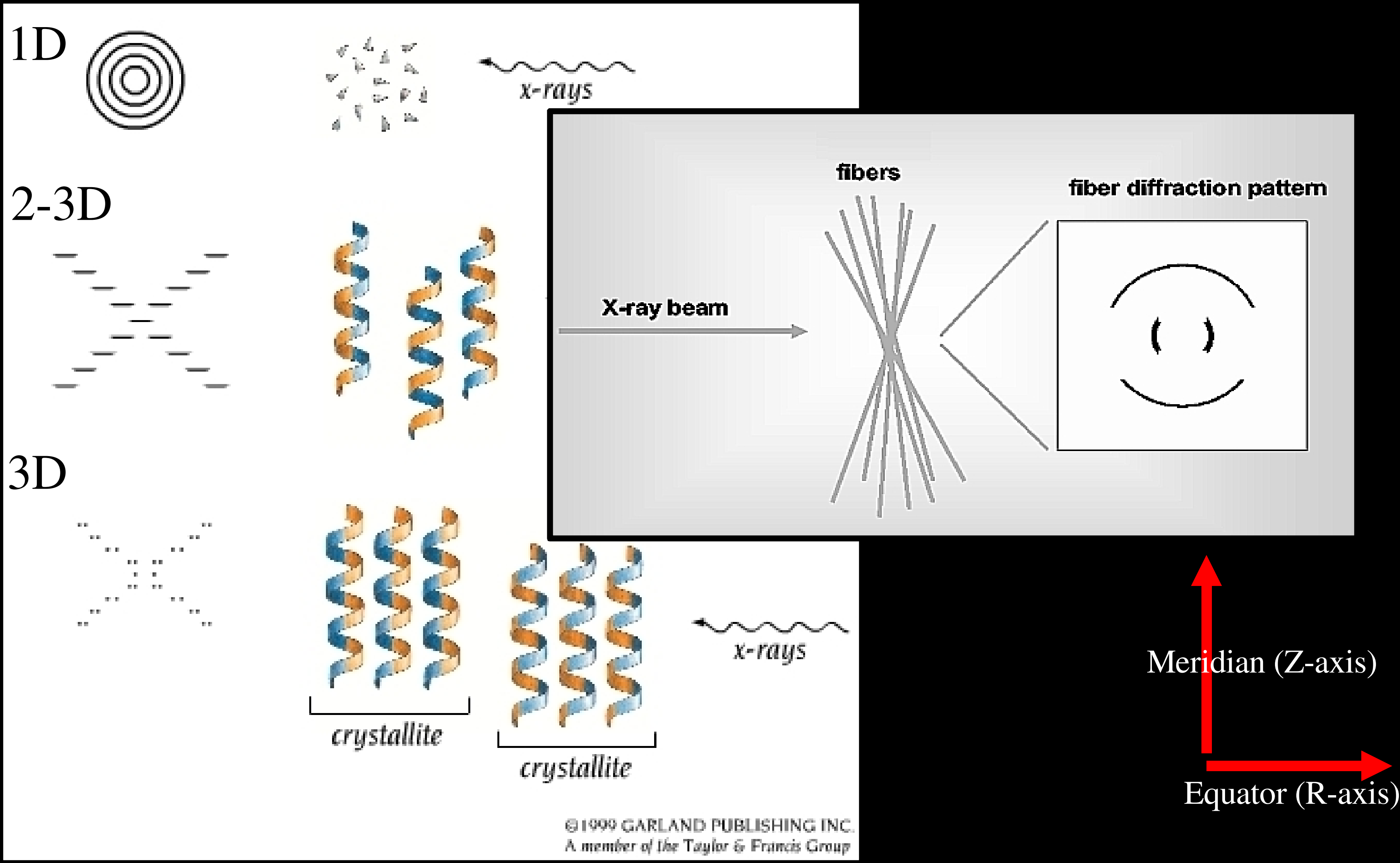
anything that can crystallize, must be (initially) soluble, usually relatively small in comparison to the above, *molecular to macromolecular scale*

Dimensional hierarchy of fiber diffraction patterns



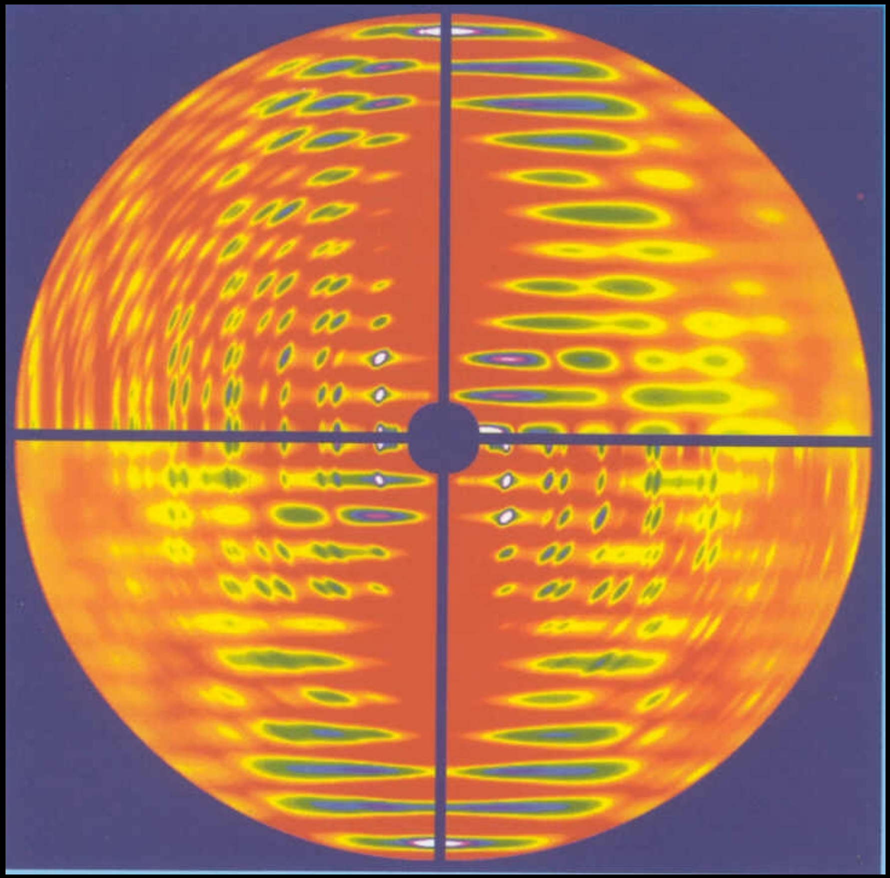


Dimensional hierarchy of fiber diffraction patterns



polycrystalline

noncrystalline



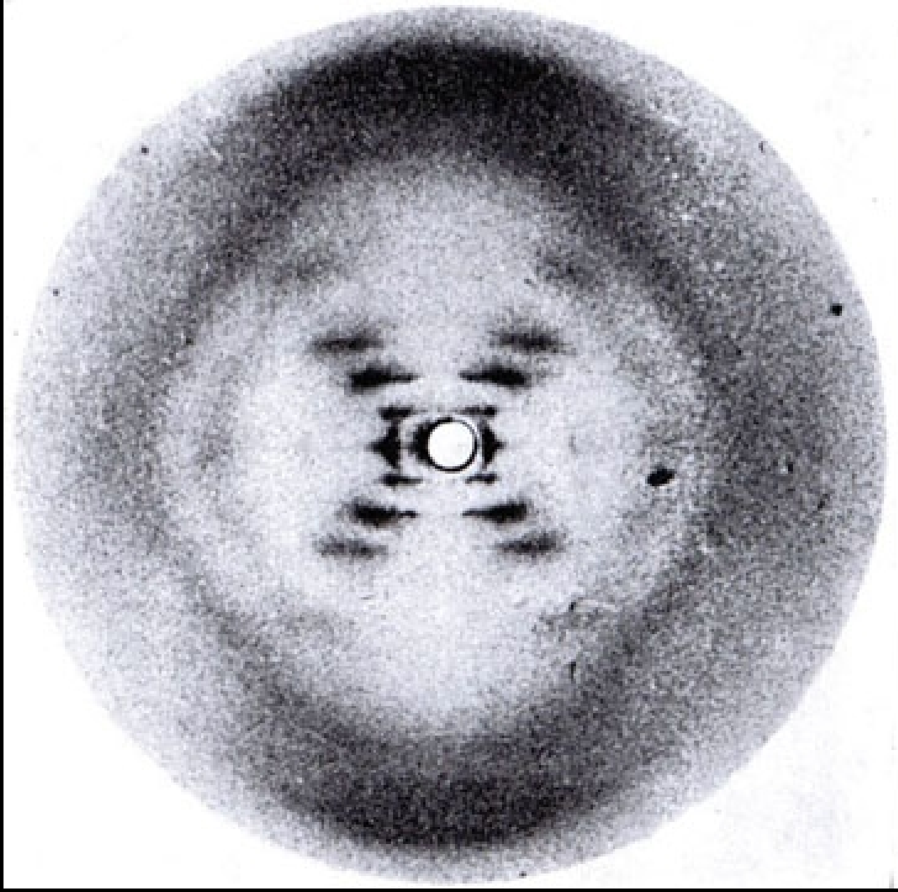
Discrete rotations +
correlated lattice disorder

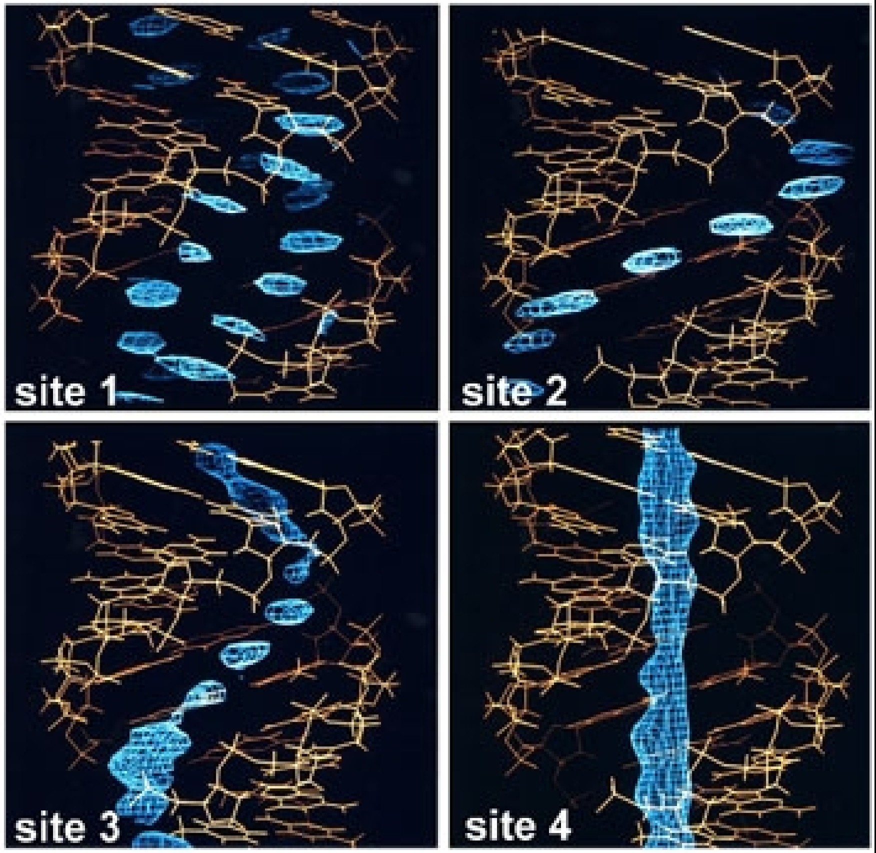
Random screw disorder

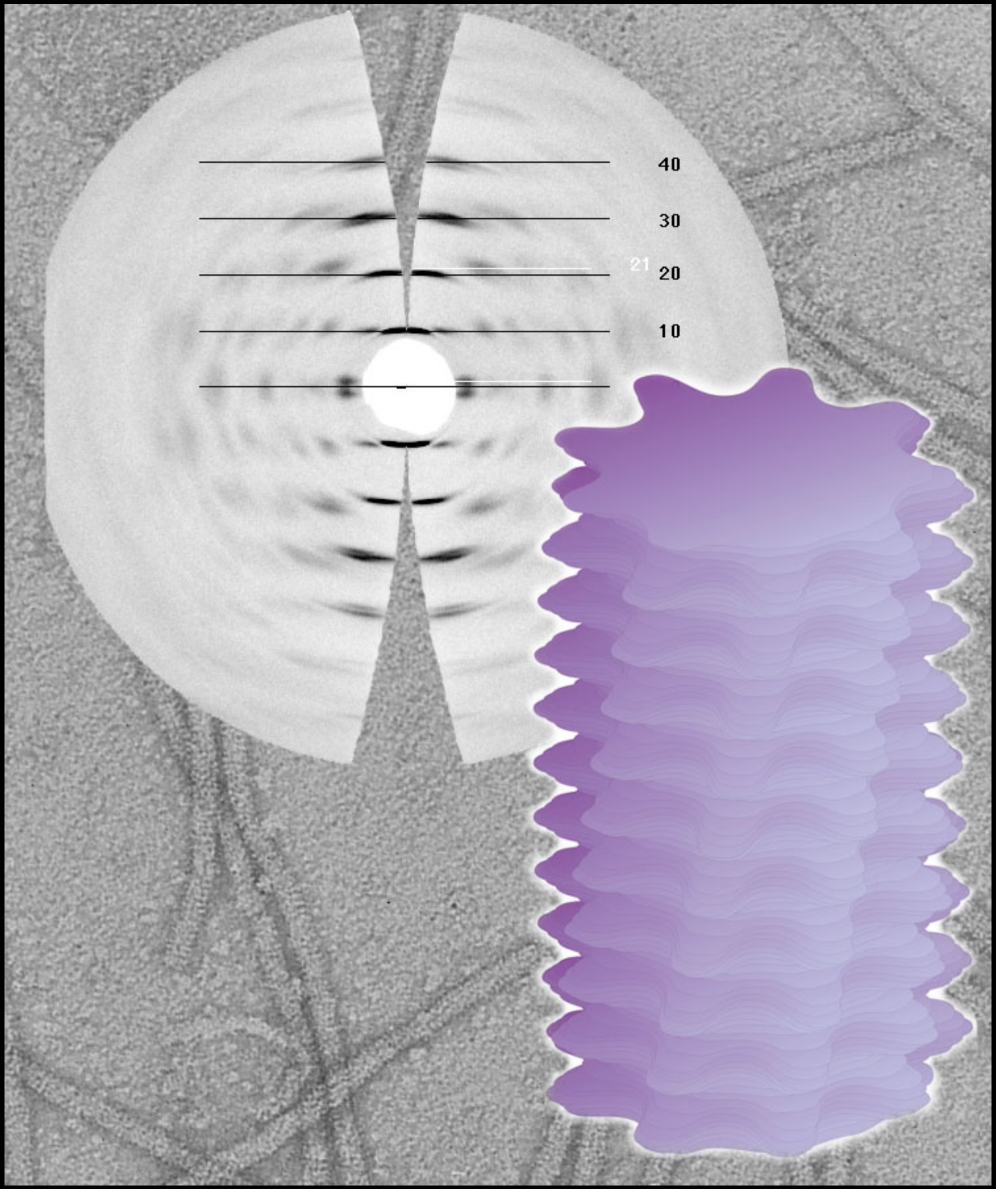
Examples

Some fibers, their diffraction patterns and corresponding molecular structures

B-form DNA

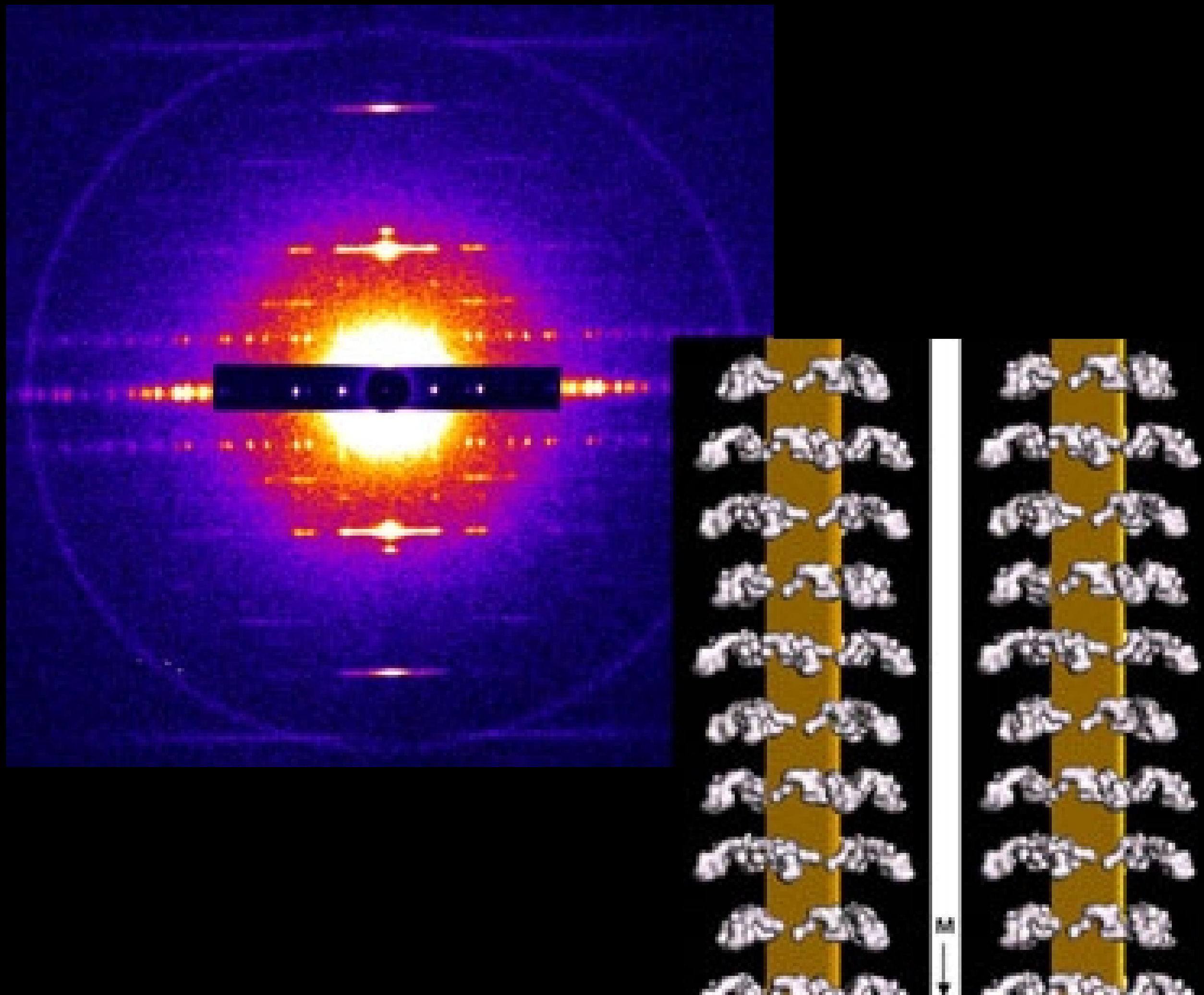


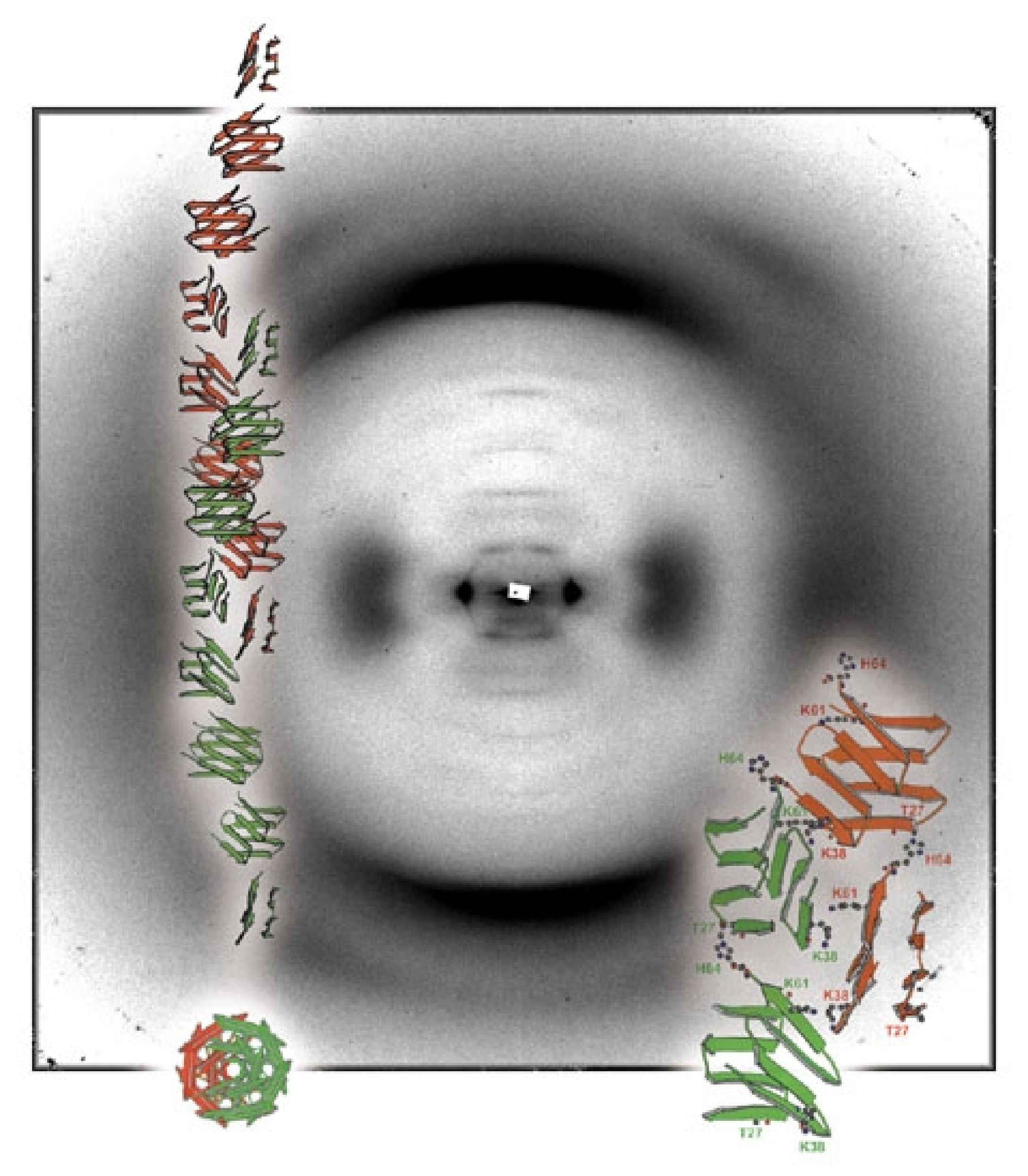




Potato virus X

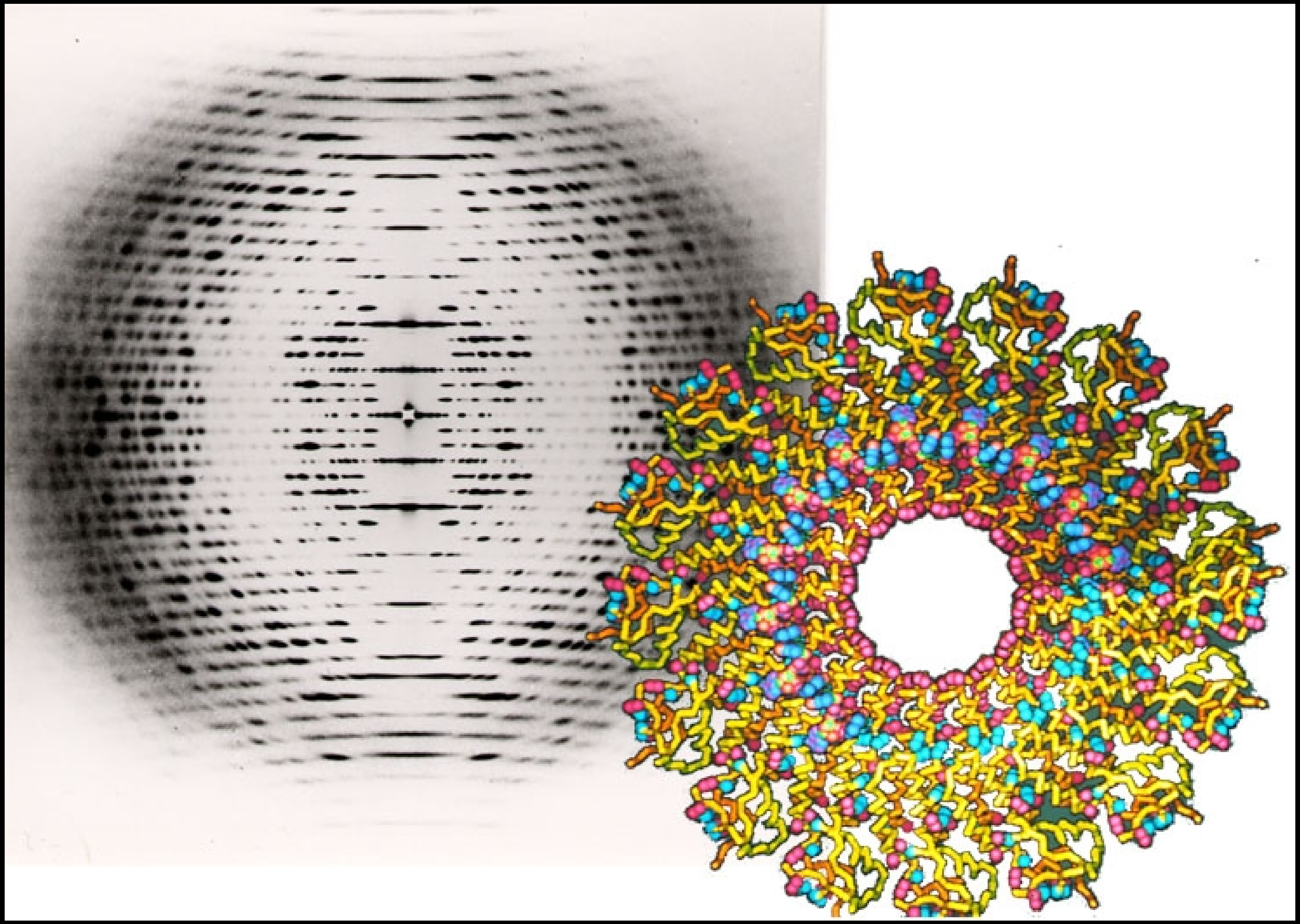
Insect flight-muscle



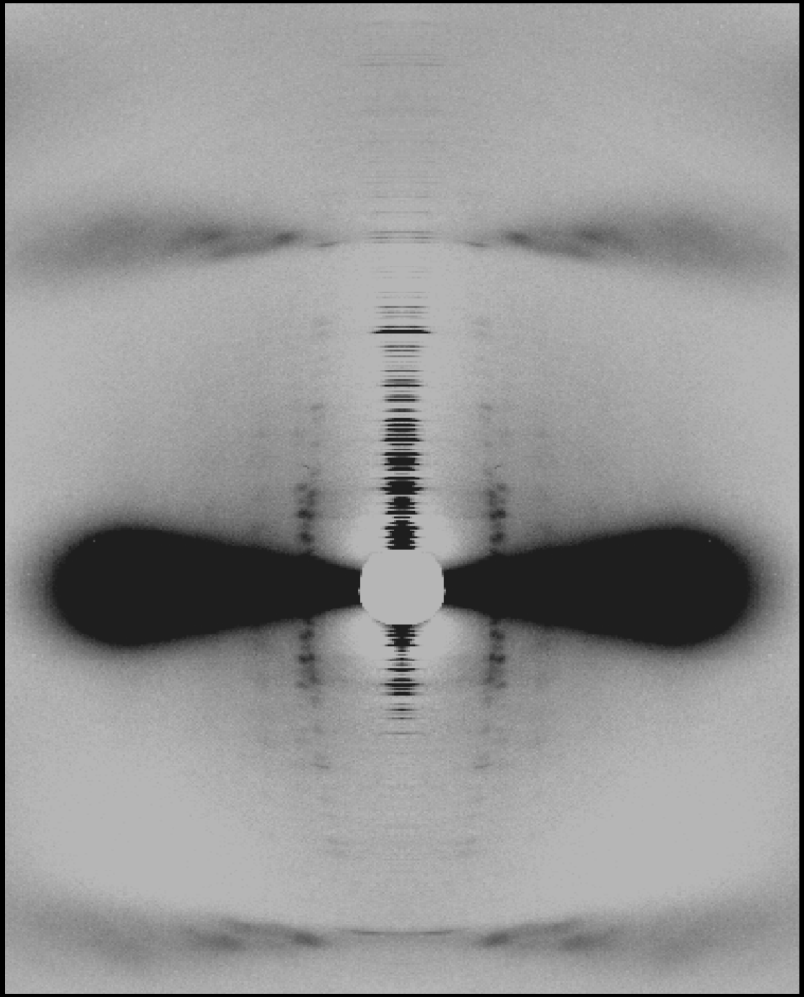


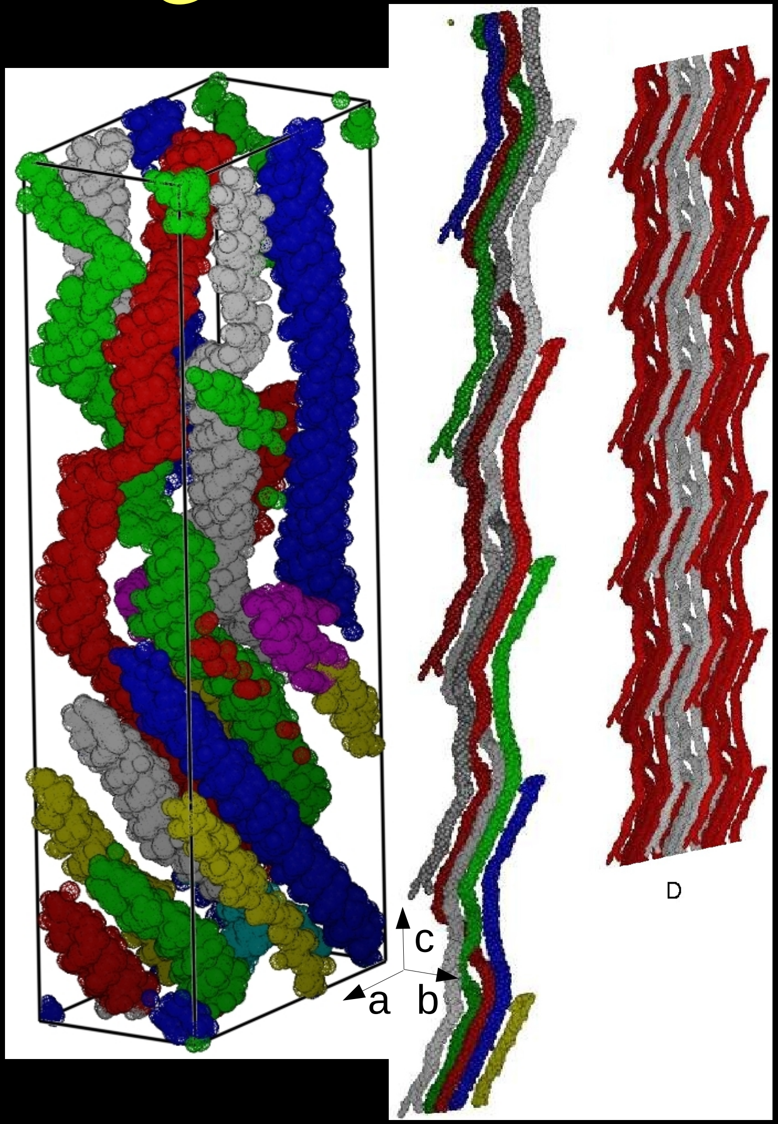
Engineered amyloid fiber

Tobacco mosaic virus



Type I collagen

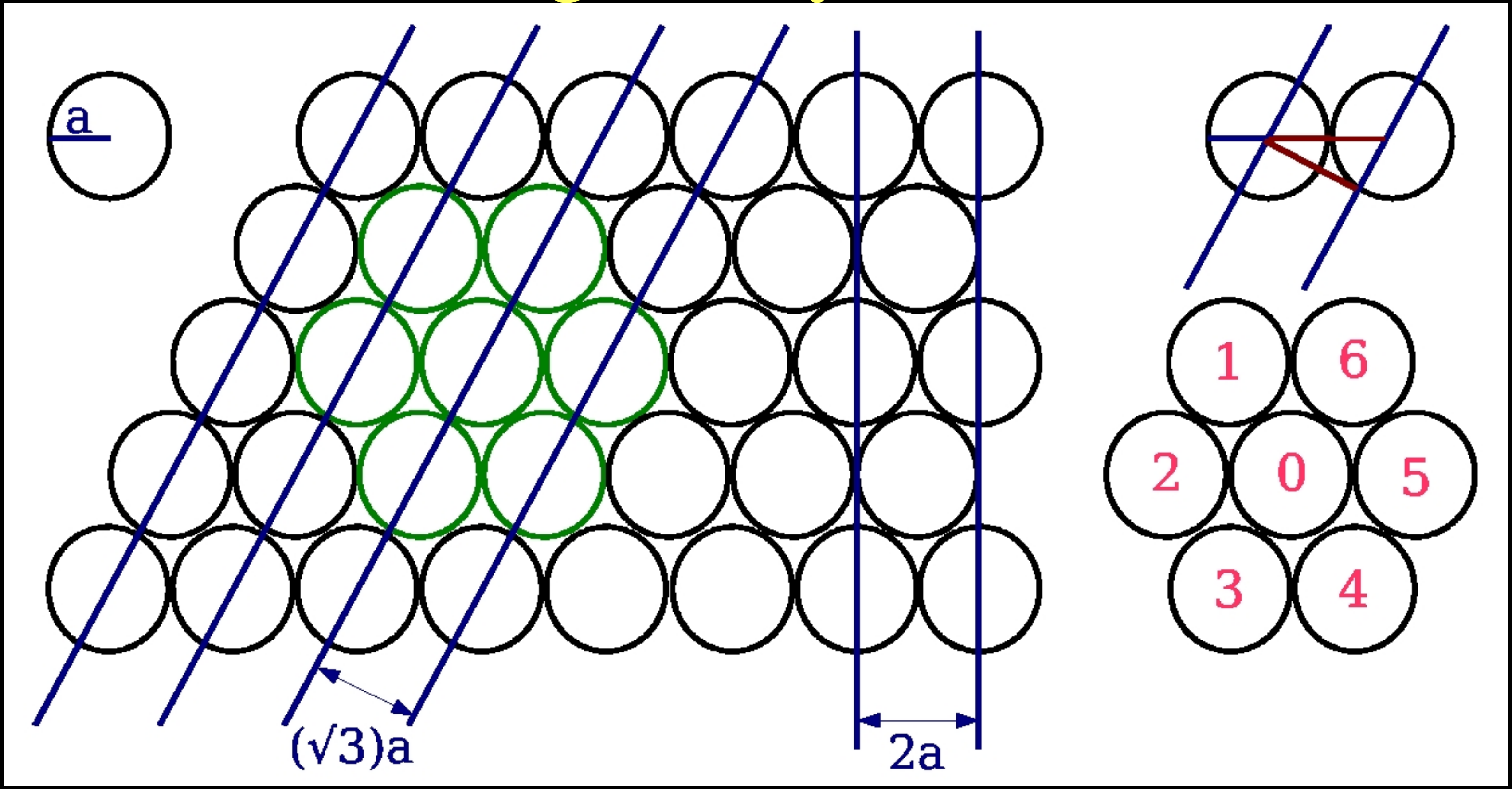




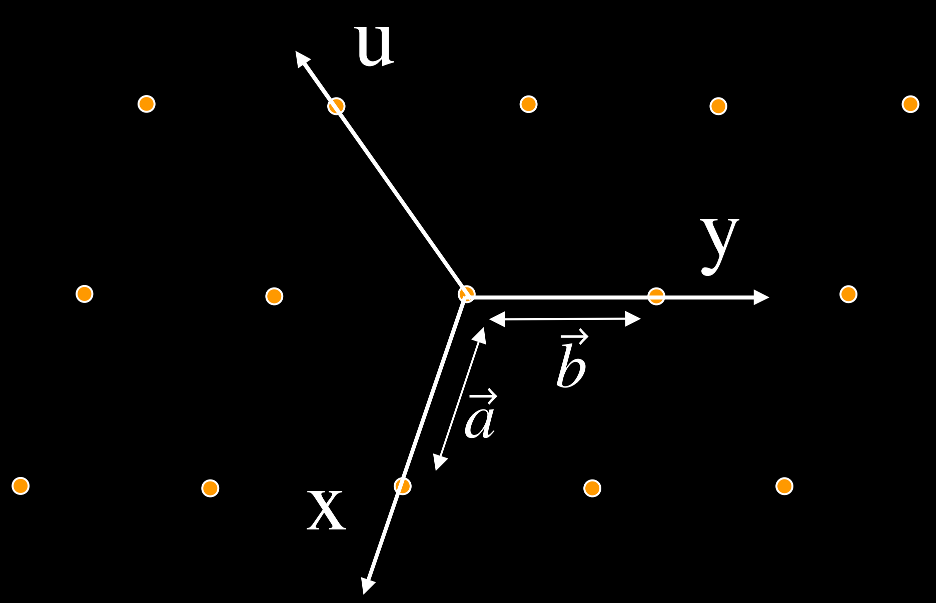
Principles I

packing of fibers and diffraction from the 3D lattice

Fibers (essentially rods/cylinders) Usually Hexagonally Packed

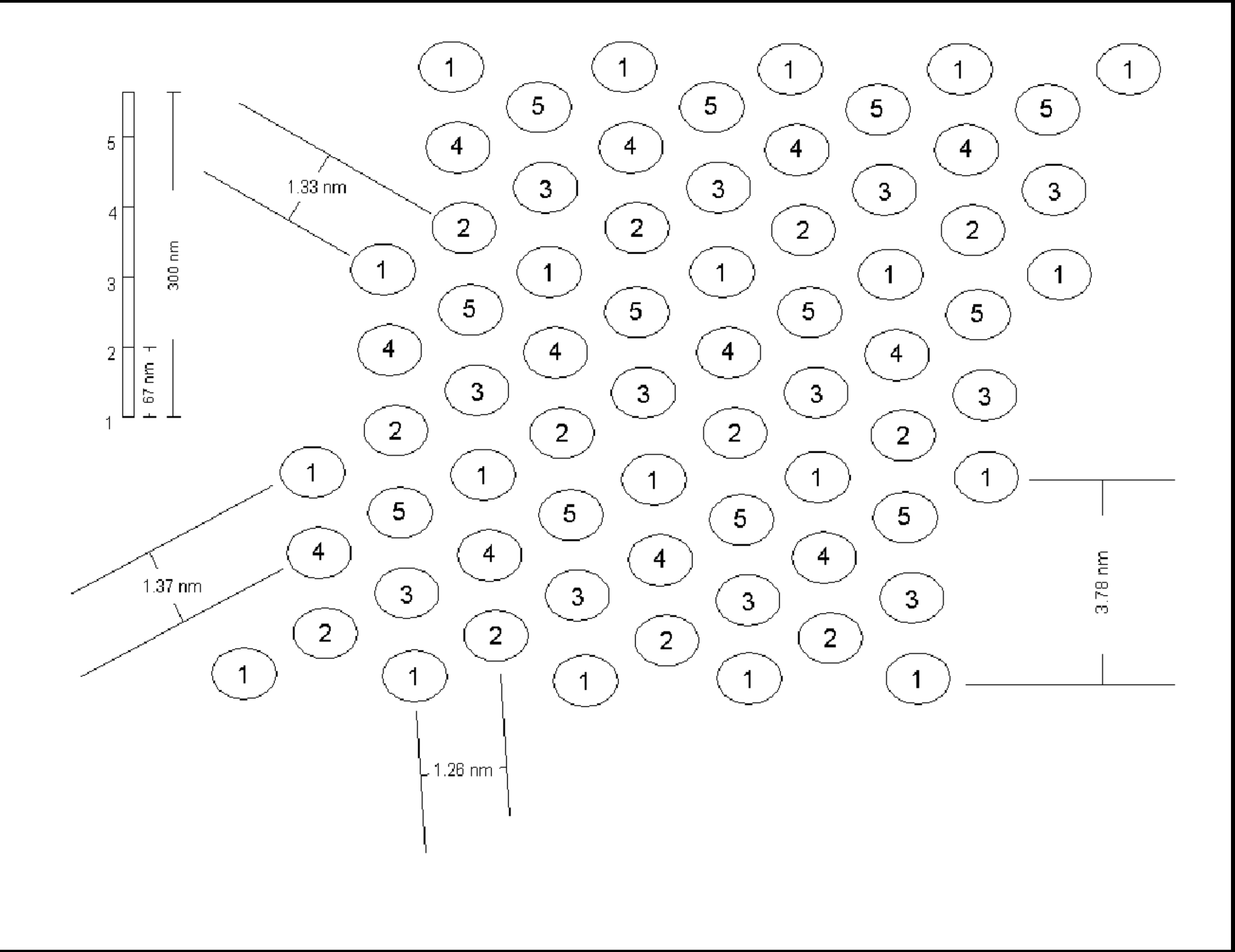


Hexagonal Lattice variables

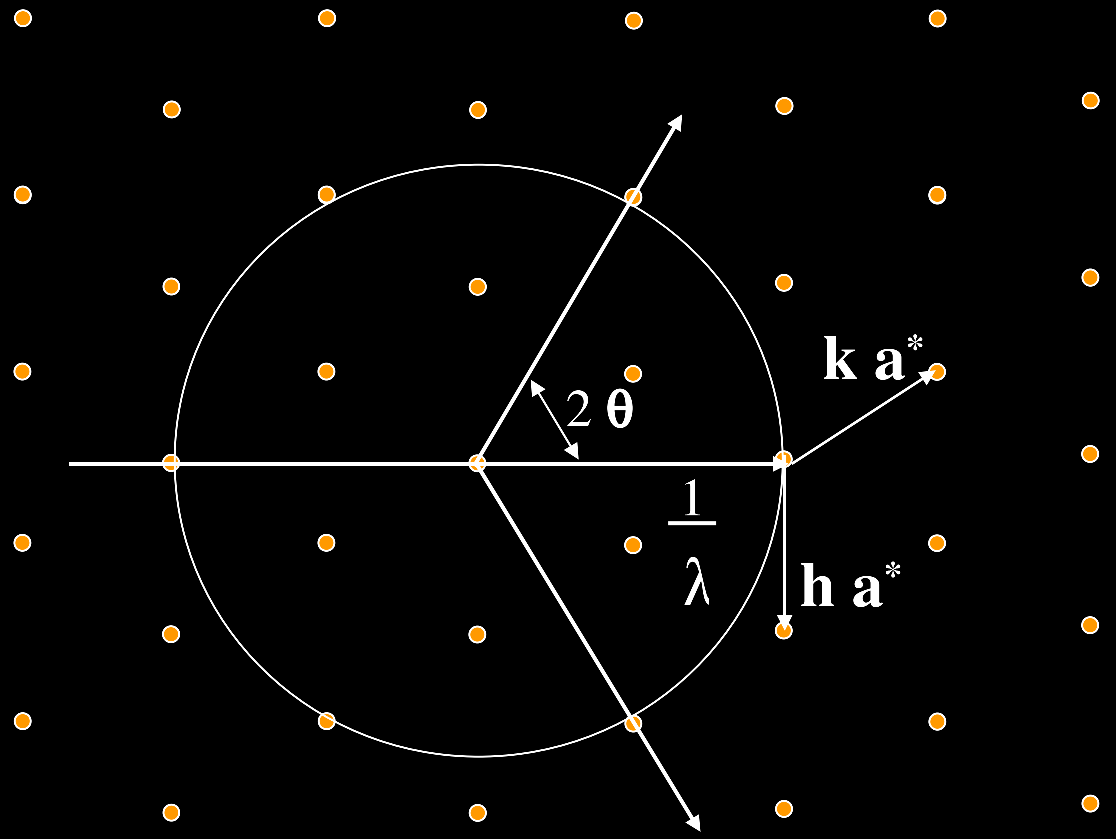


$$\vec{a} = \vec{b}$$

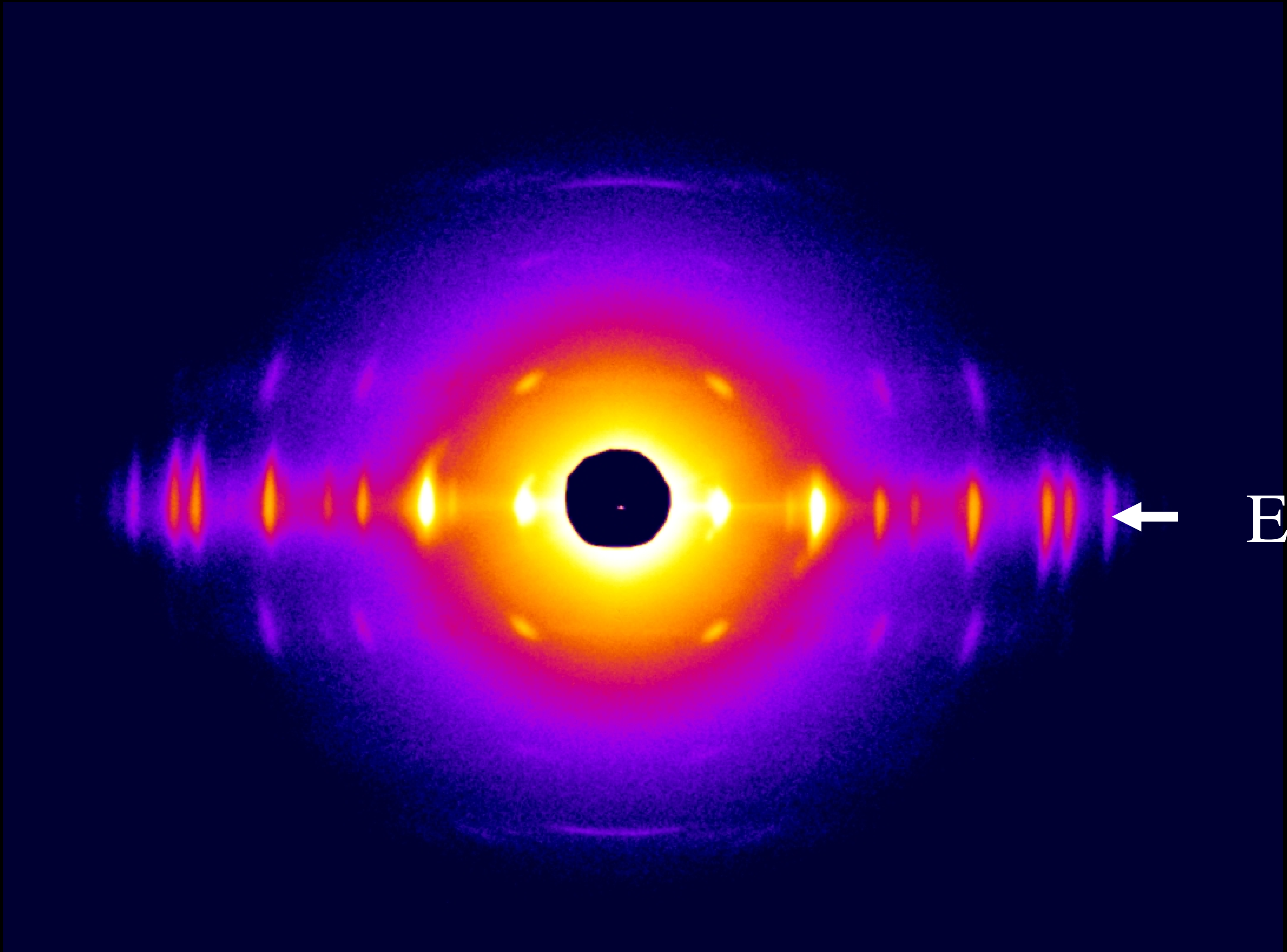
Some fibers not just straight rods: A Quasi-Hexagonal Lattice



Ewald Sphere

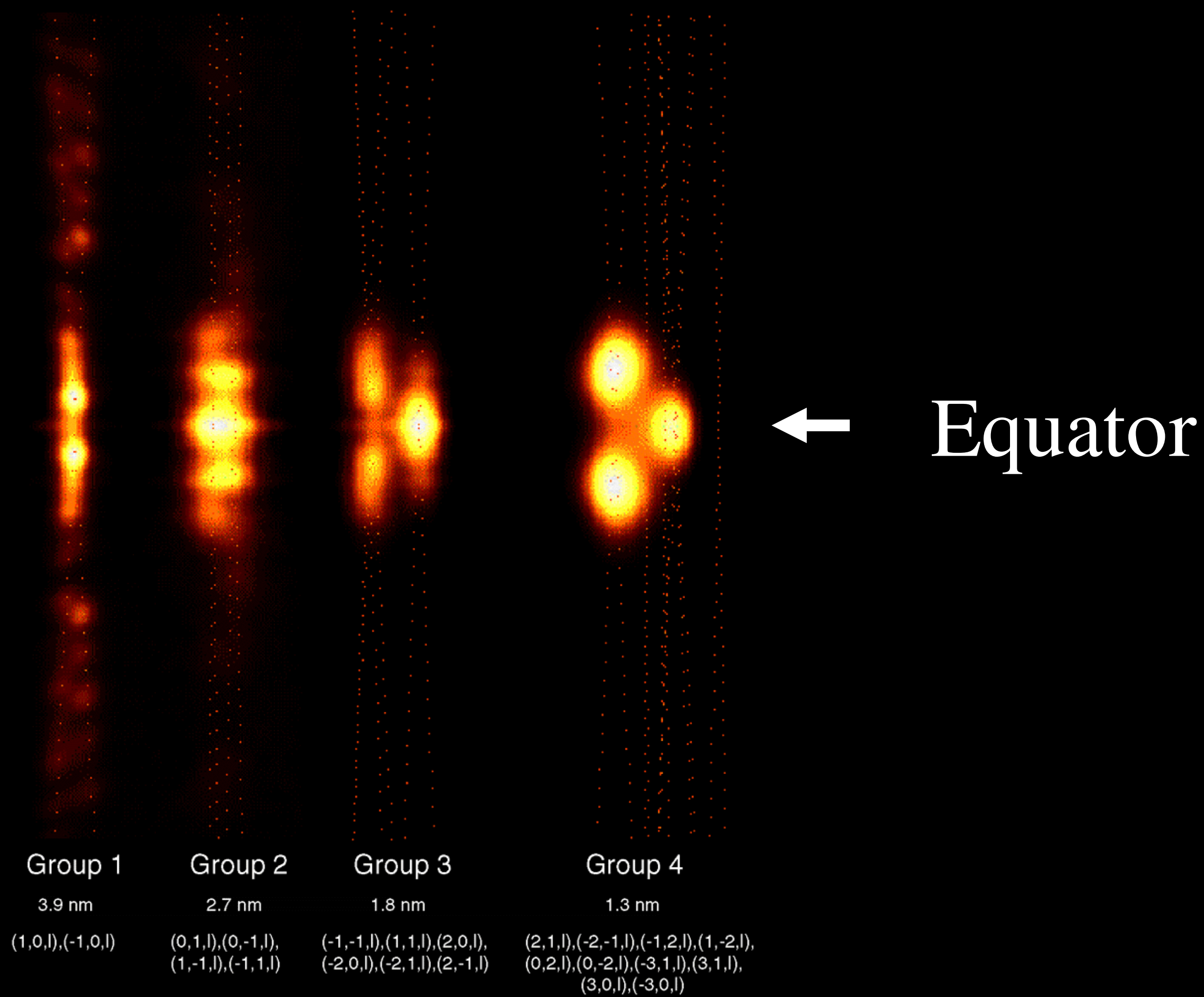


Fiber diagram - Insect Muscle (hexagonal lattice)



Equator

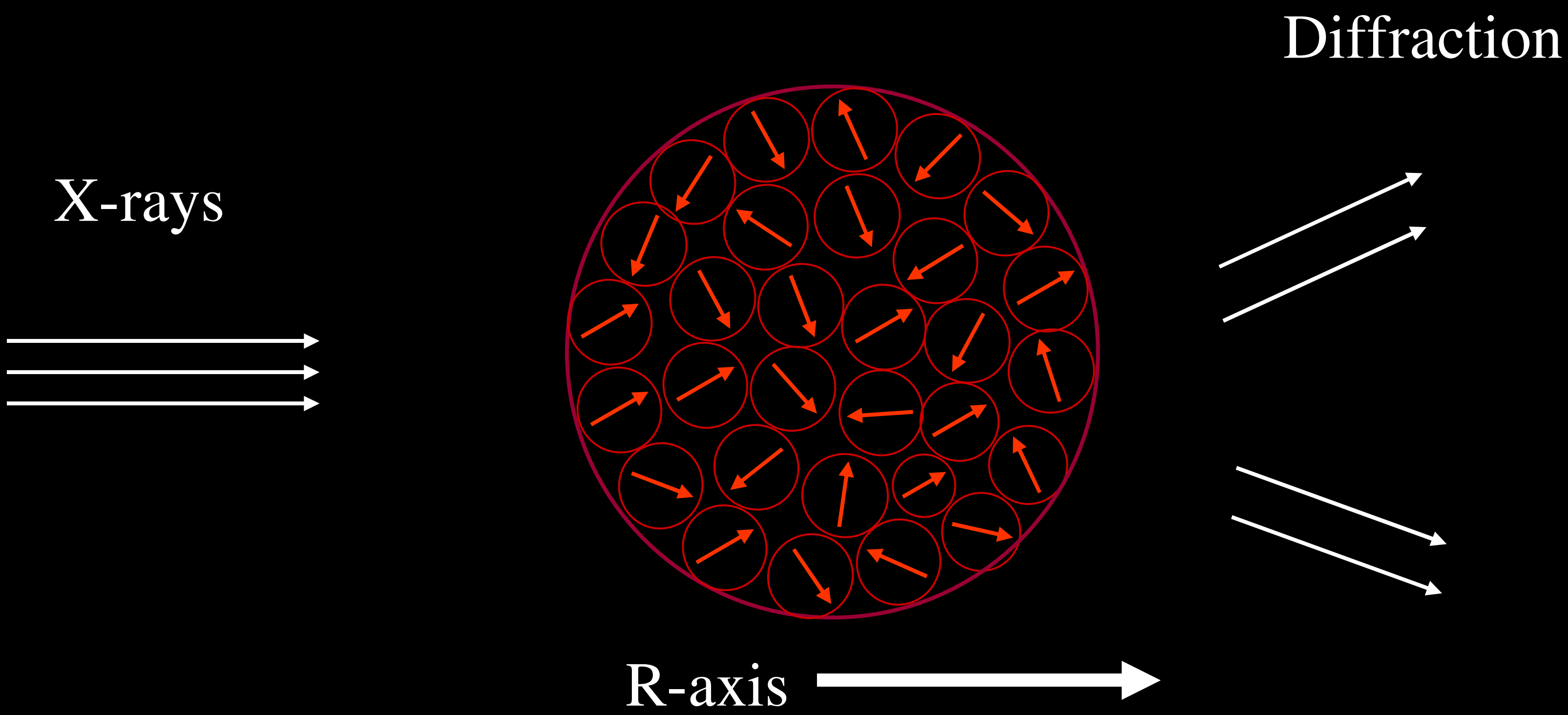
Fiber diagram – Rat tail tendon (quasi-hexagonal lattice)



Principles II

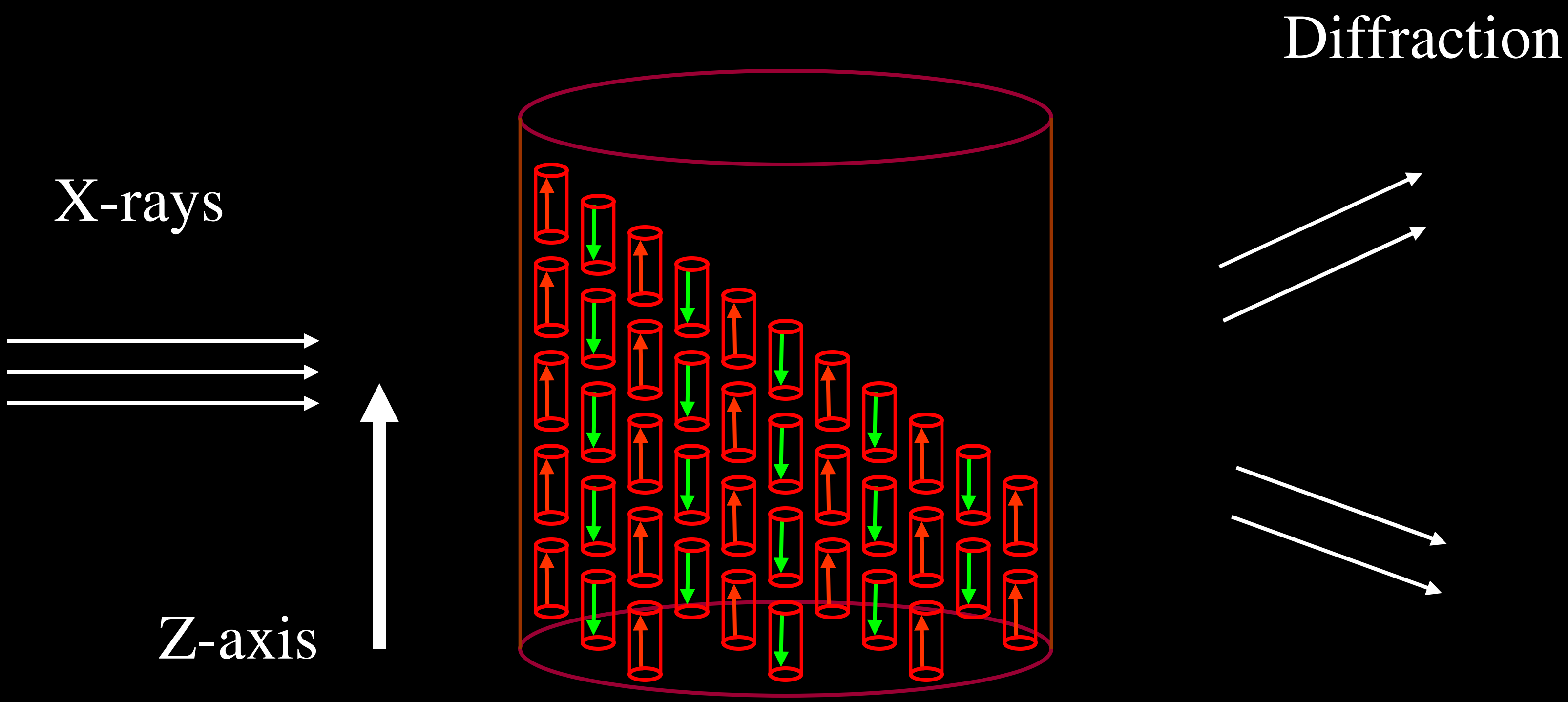
Cylindrical convolution and Orientation effects

Fiber Cross-section



Crystallites randomly orientated around the axis perpendicular to the fiber axis (the '**R**' – axis: Sum of Rotation of crystallites assumed = 360 degrees)

Fiber Cross-section



Crystallites randomly orientated around the fiber axis

the **'Z'** – axis: crystallites assumed to be rotated 0 OR 180 degrees

Therefore:

For every reciprocal-lattice point (h,k,l) there is another $(\bar{h},\bar{k},\bar{l})$ for which:

$$R(h,k,l) = R(\bar{h},\bar{k},\bar{l})$$

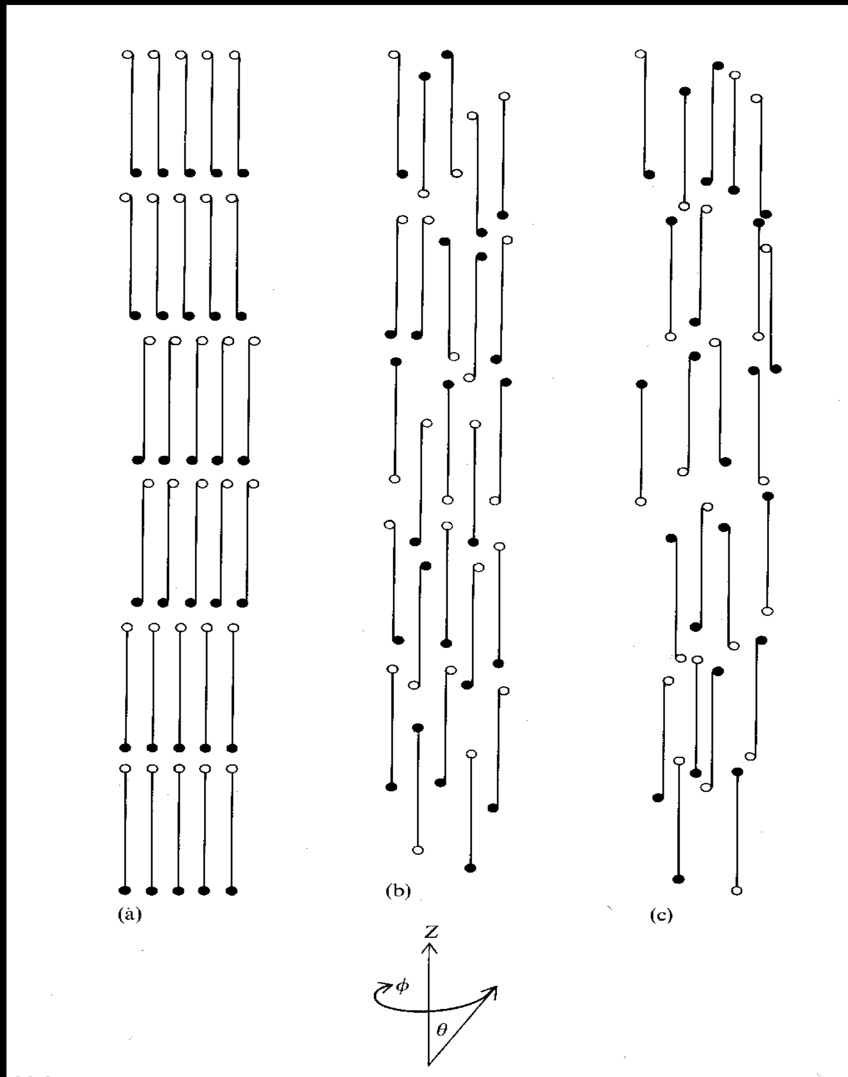
and

$$-Z(h,k,l) = Z(\bar{h},\bar{k},\bar{l})$$

Principles III

Order and disorder in fibrous specimens

A B C



Ordering in Fibers:

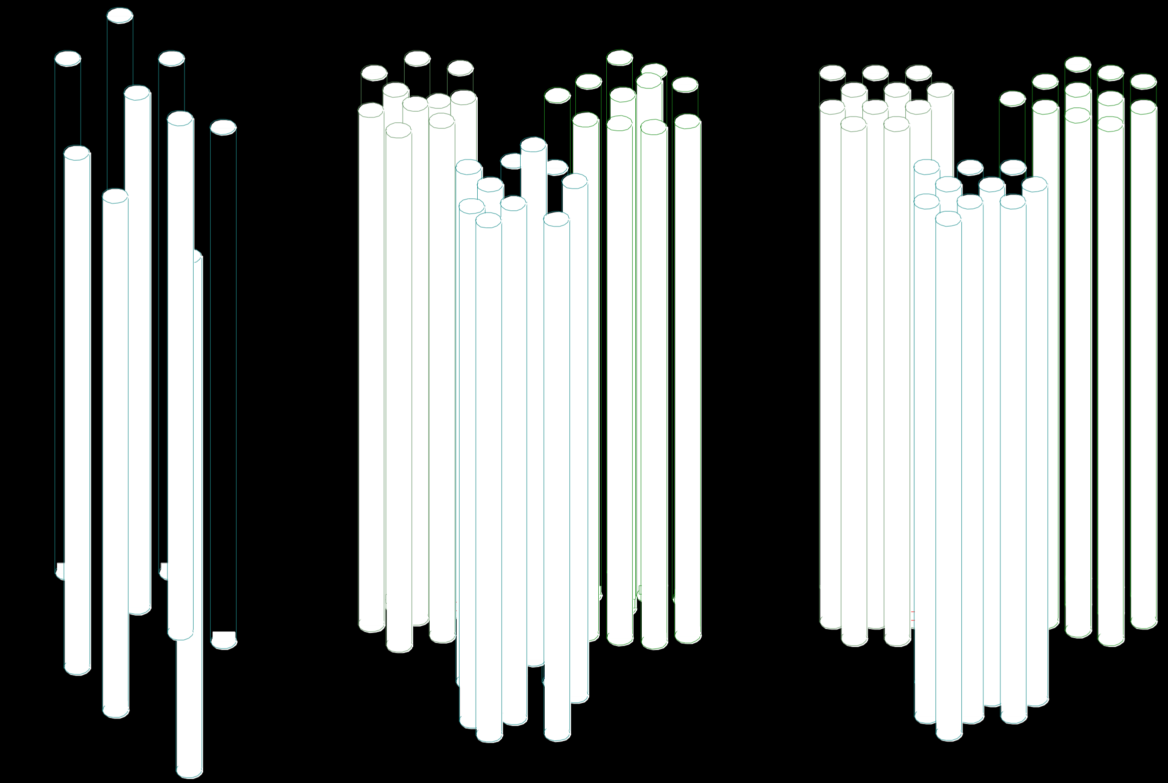
A - Crystalline fiber

B Semicrystalline Fiber

C Non-crystalline fiber

$$\langle I(s)\rangle = \langle |F_m(S)F_L(S)|^2\rangle$$

Average over all molecular and lattice orientations

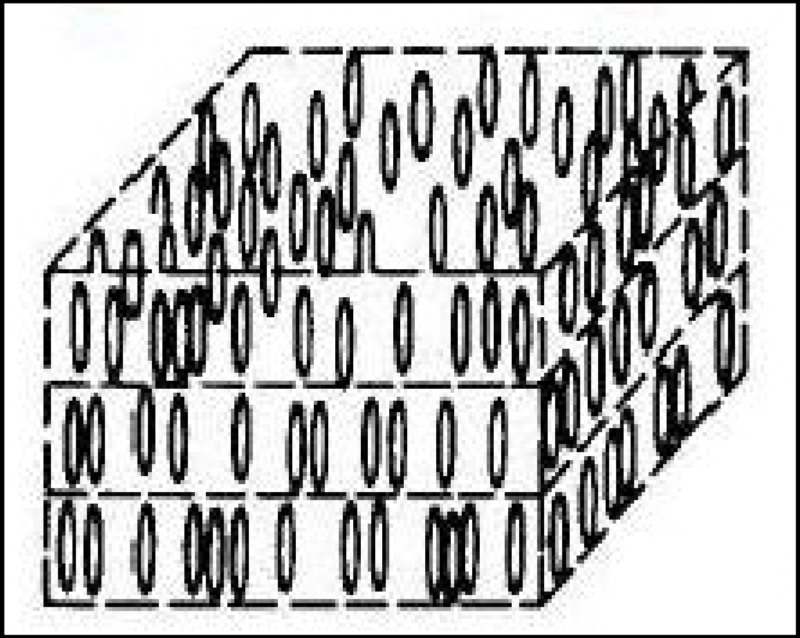


a noncrystalline | b disordered | c polycrystalline

We implicitly incorporate disorientation & crystallite size

One caveat, packing organization (lateral structure)

is often poorer then axial organization

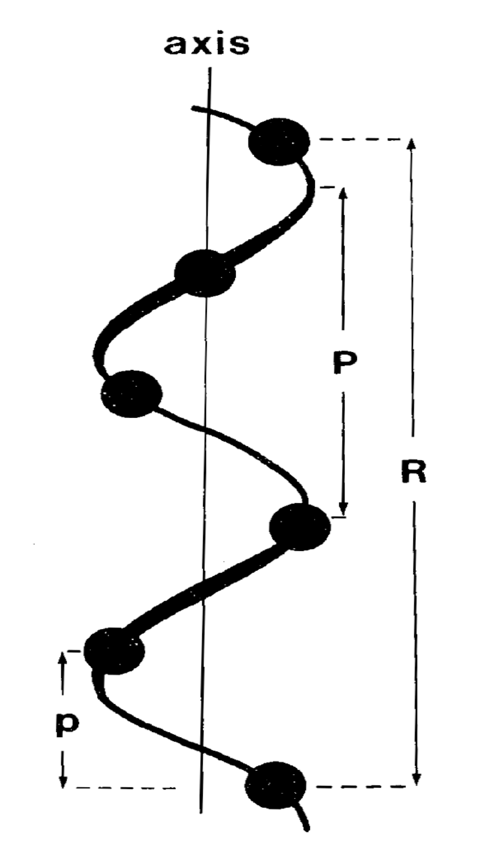


Smetic liquid crystal organization, i.e. The order in one or more dimensions varies from good to poor to random

Principles IV

Helical diffraction theory

Fibrous Proteins Usually Show Helical Symmetry



P = pitch

p = subunit axial translation distance

R = true repeat distance

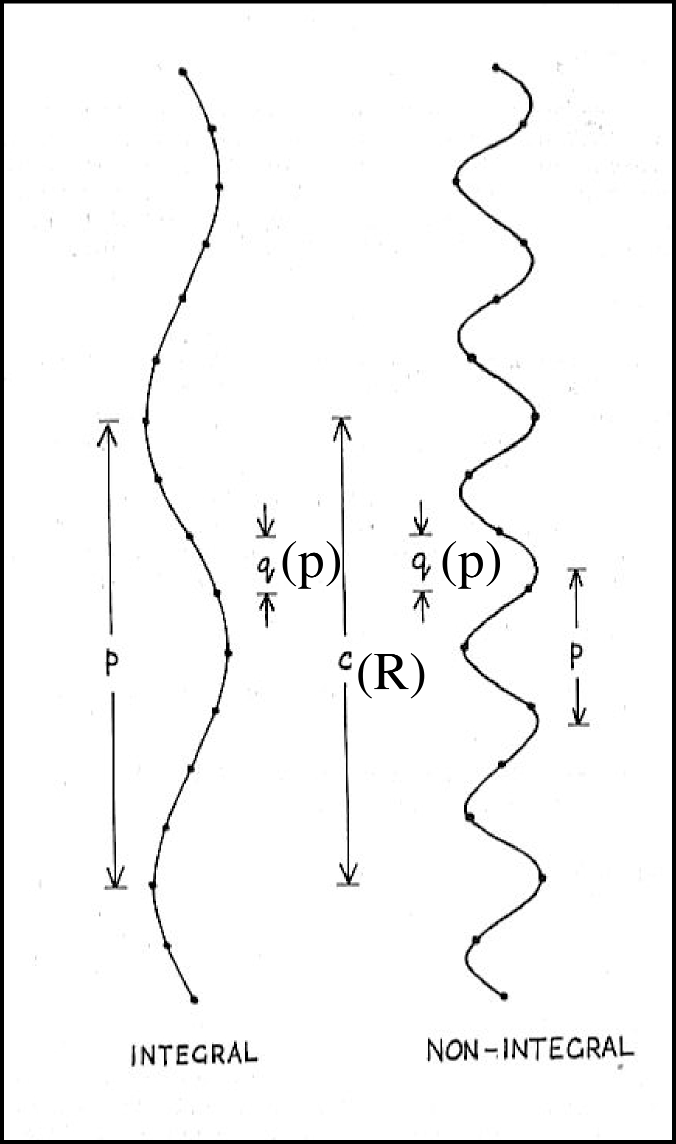
Integral / Non-integral helices

Integral helix:

R = P (true repeat = pitch)

Non-integral helix:

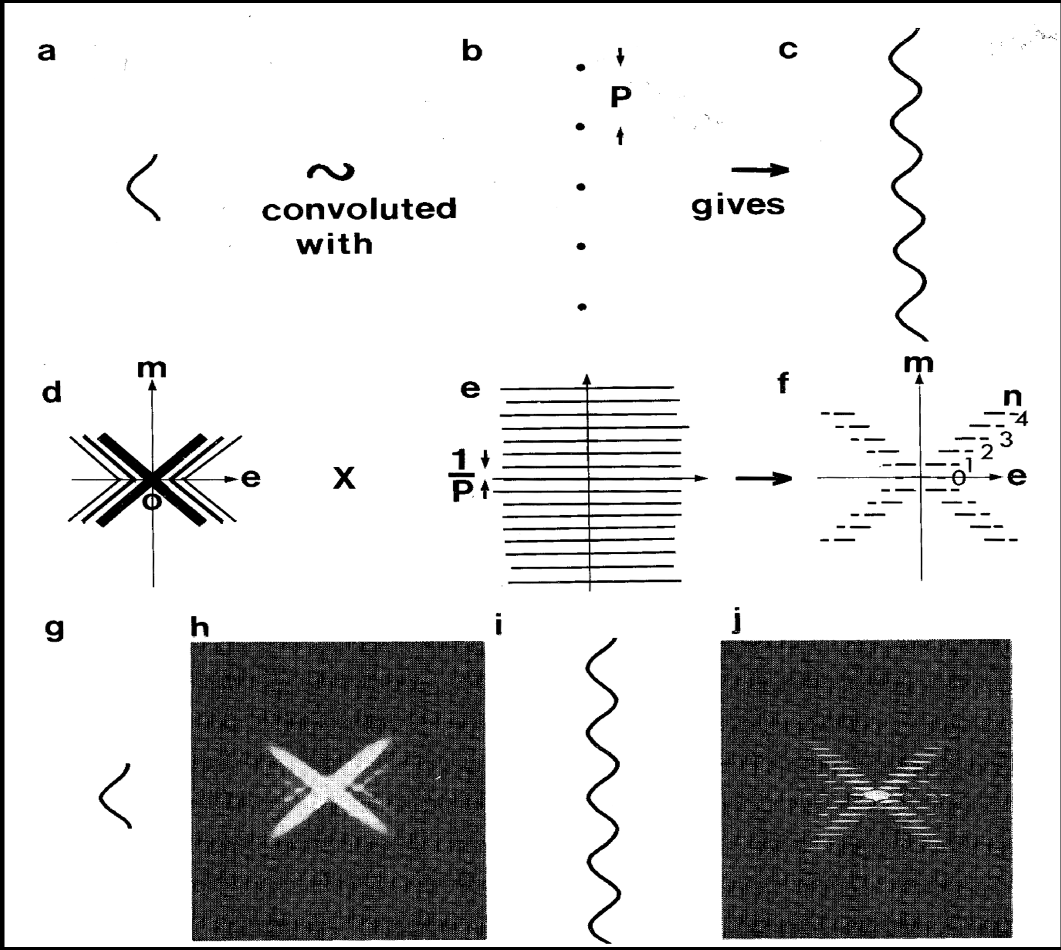
R > P, but must be integral multiple of p or else can not be repeating structure! (R=n.p)



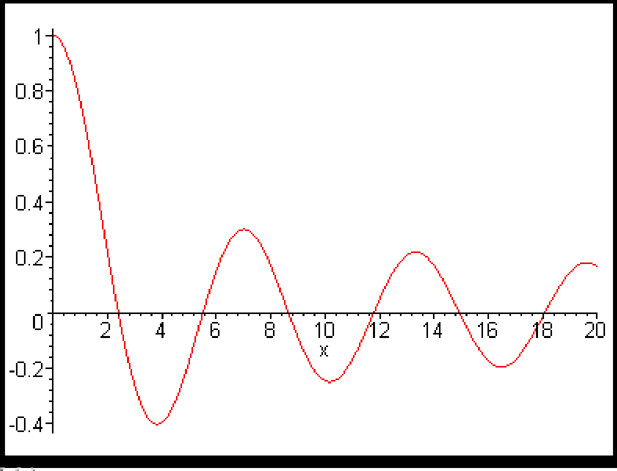
Continuous Helix

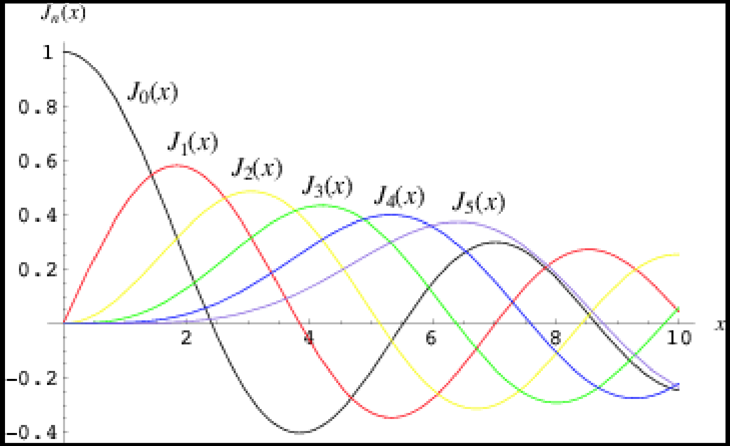
A uniformly smooth object that follows a helical path

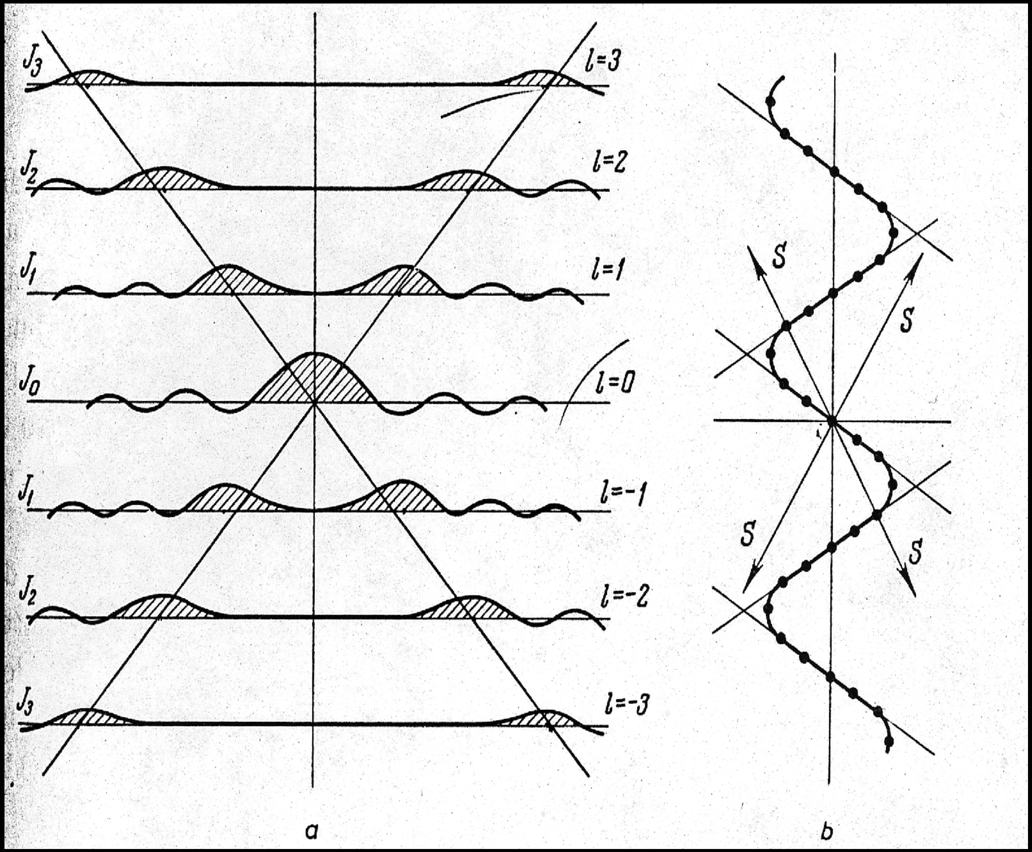
Diffraction from a continuous Helix



Bessel Functions and Layer lines

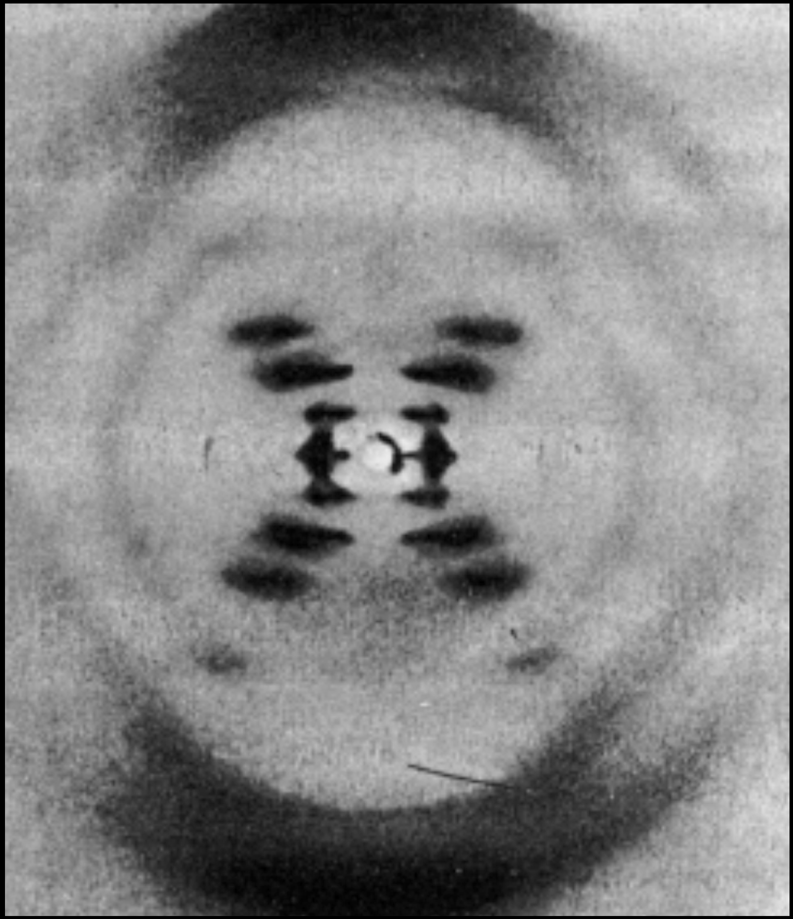






Transform of a cylinder

Rosalind Franklin's Pattern from B-DNA (*not really* a continuous helix)

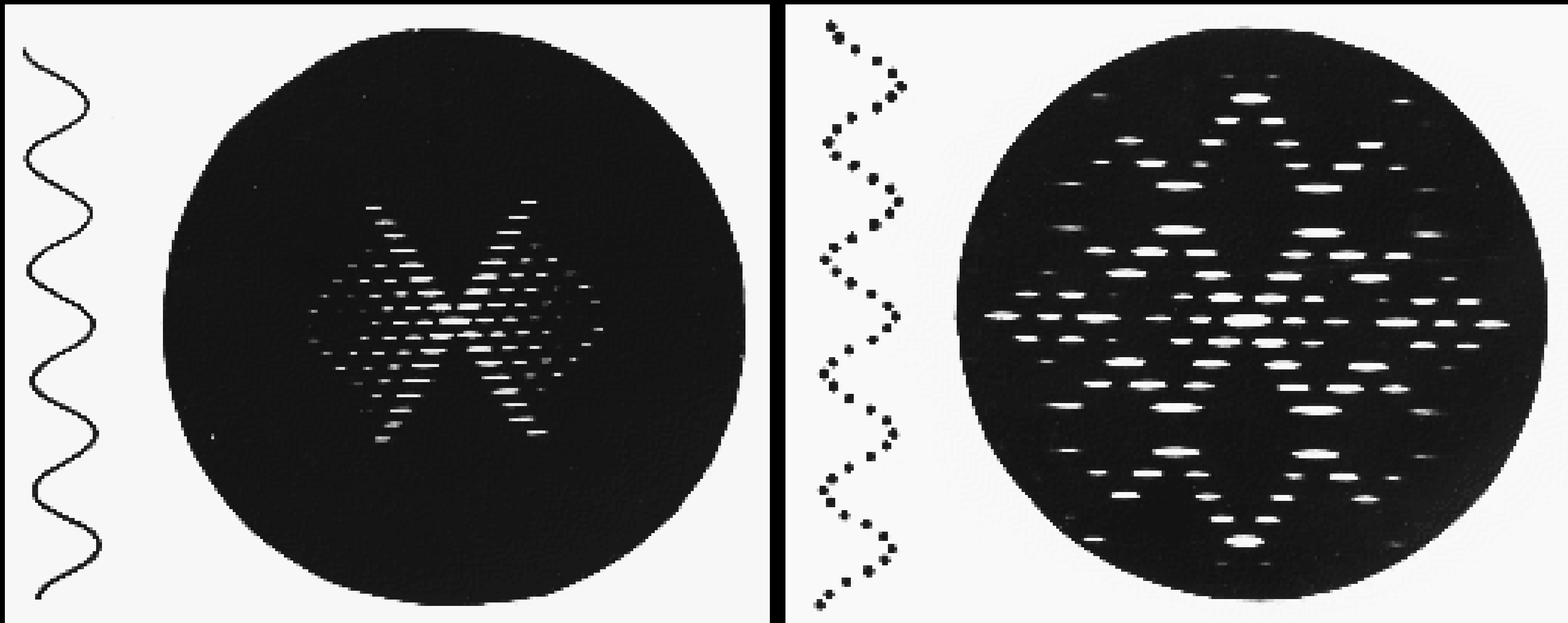


Franklin & Gosling, 1953 Nature 171:740

Discontinuous Helix

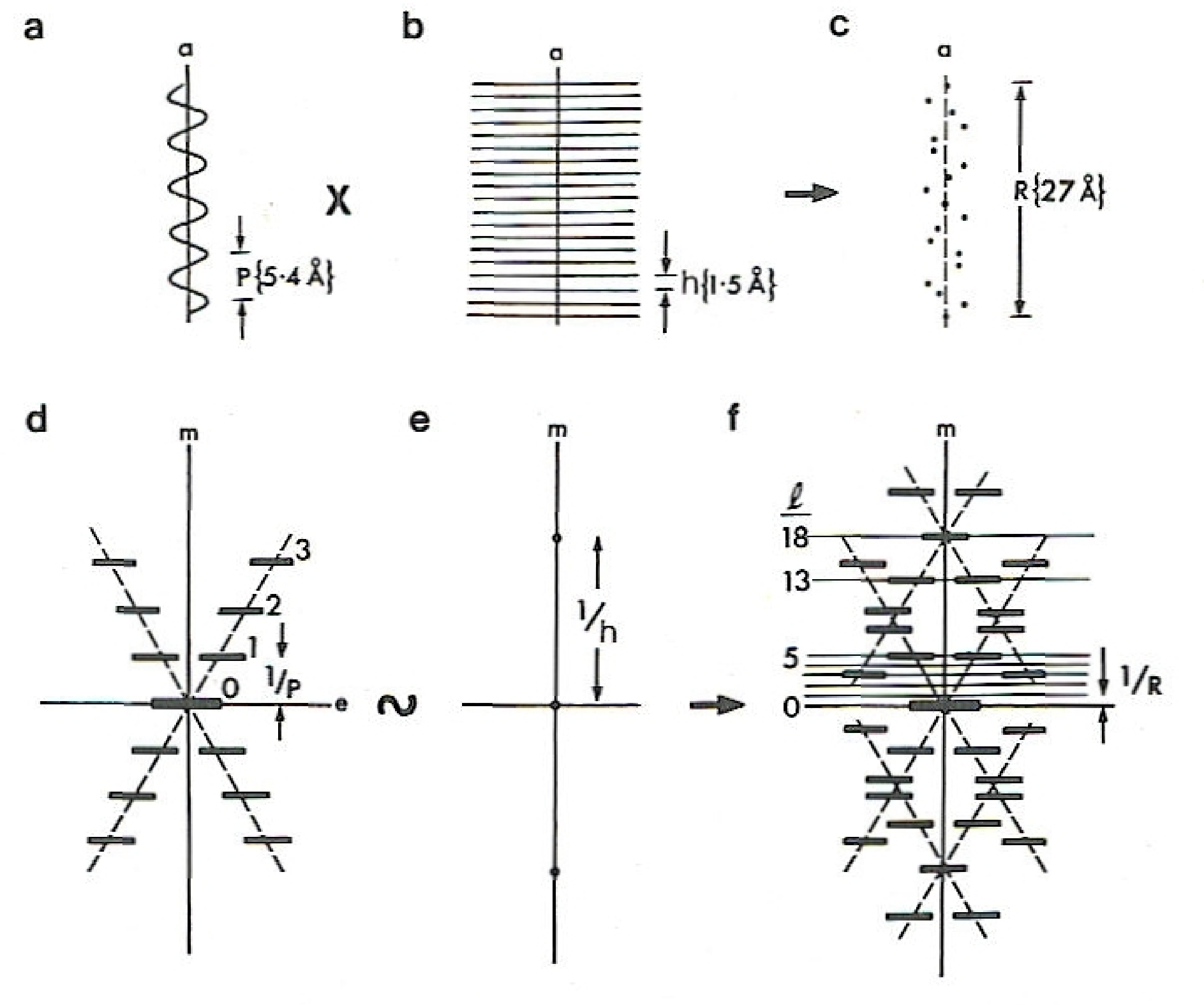
A set of points that are regularly spaced along a helical path

Diffraction from a helix: comparison



The main effect of shifting from a continuous to a discontinuous helix is to introduce new helix crosses with their origins displaced up and down the meridian by a distance 1/*p* (little p = h = q !! i.e the sub-unit axial translation distance [or height])

Diffraction from a discontinuous (non-integral) Helix



Helical Selection Rule

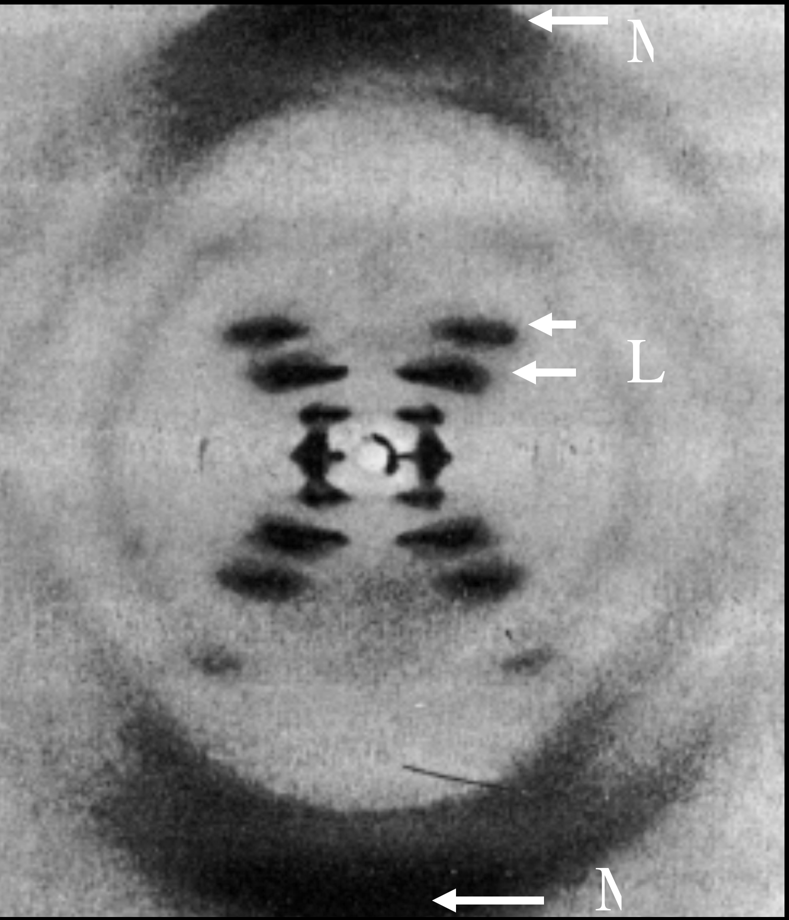
Which Bessel function order will turn up on what layer line for a more complicated helix?

For a non-integral helix (repeats after two or more turns), with **u** subunits in **t** turns, allowed Bessel functions (n) on layer line *l* are:

$$l = \mathbf{m.u} + \mathbf{n.t}$$

m is an integer indicating the translational periodicity index of the helix lattice

Rosalind Franklin's Pattern from B-DNA (integral helix)



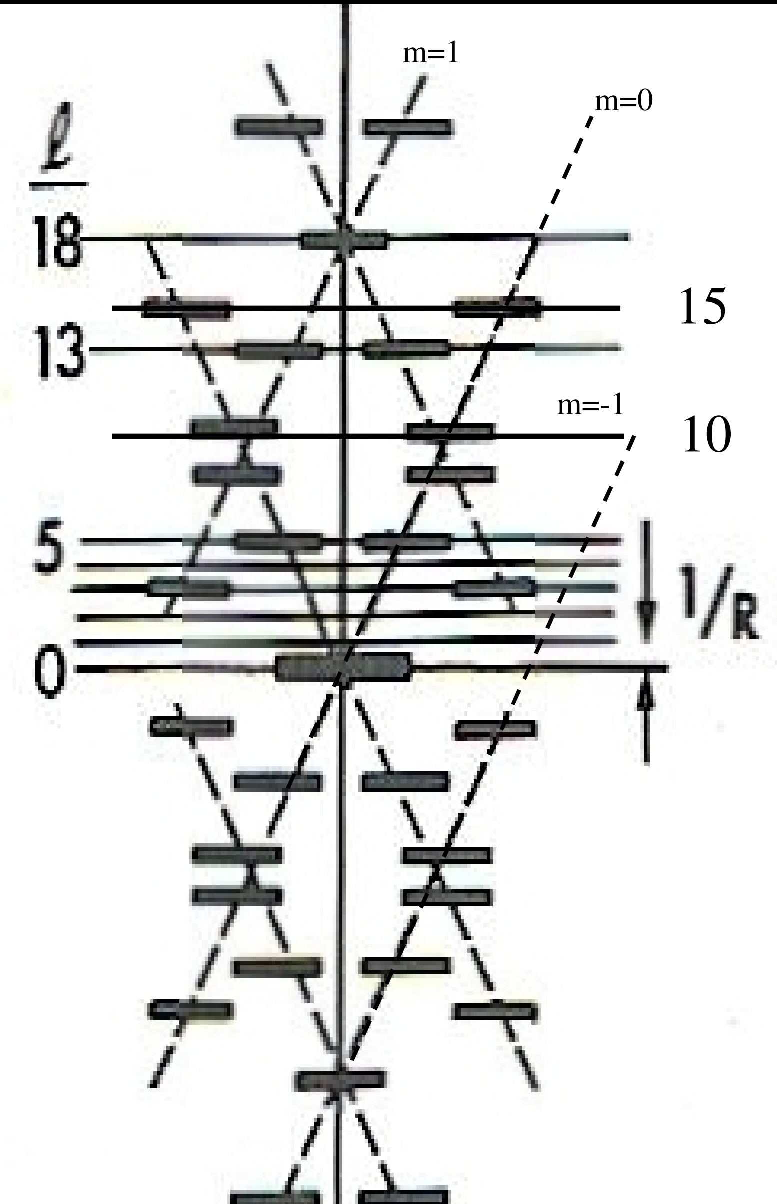
Layer lines (L) separated by 34 Å

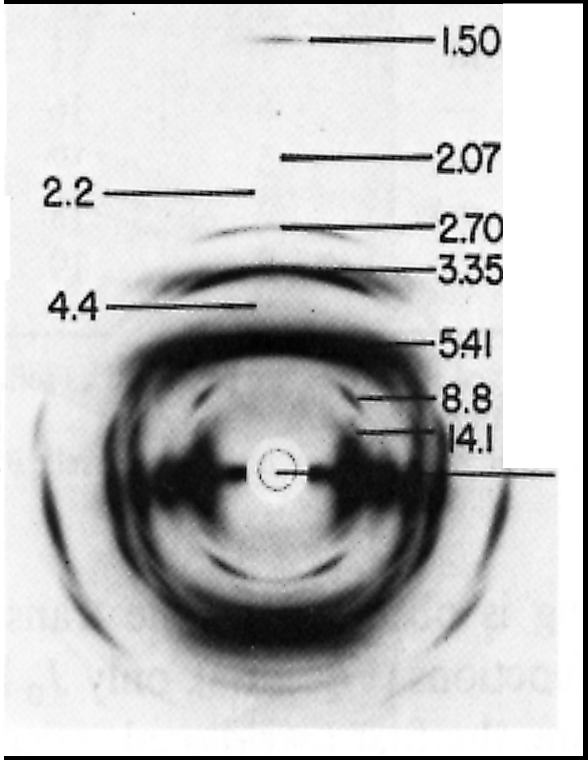
Meridional (M) reflection at 3.4 Å

⇒10 residues/turn

Franklin & Gosling, 1953 Nature 171:740

Diffraction from Poly L-Alanine -α-helix (non-integral helix)





1.5 Å residue trans.

Pitch 5.4Å, R= 27Å

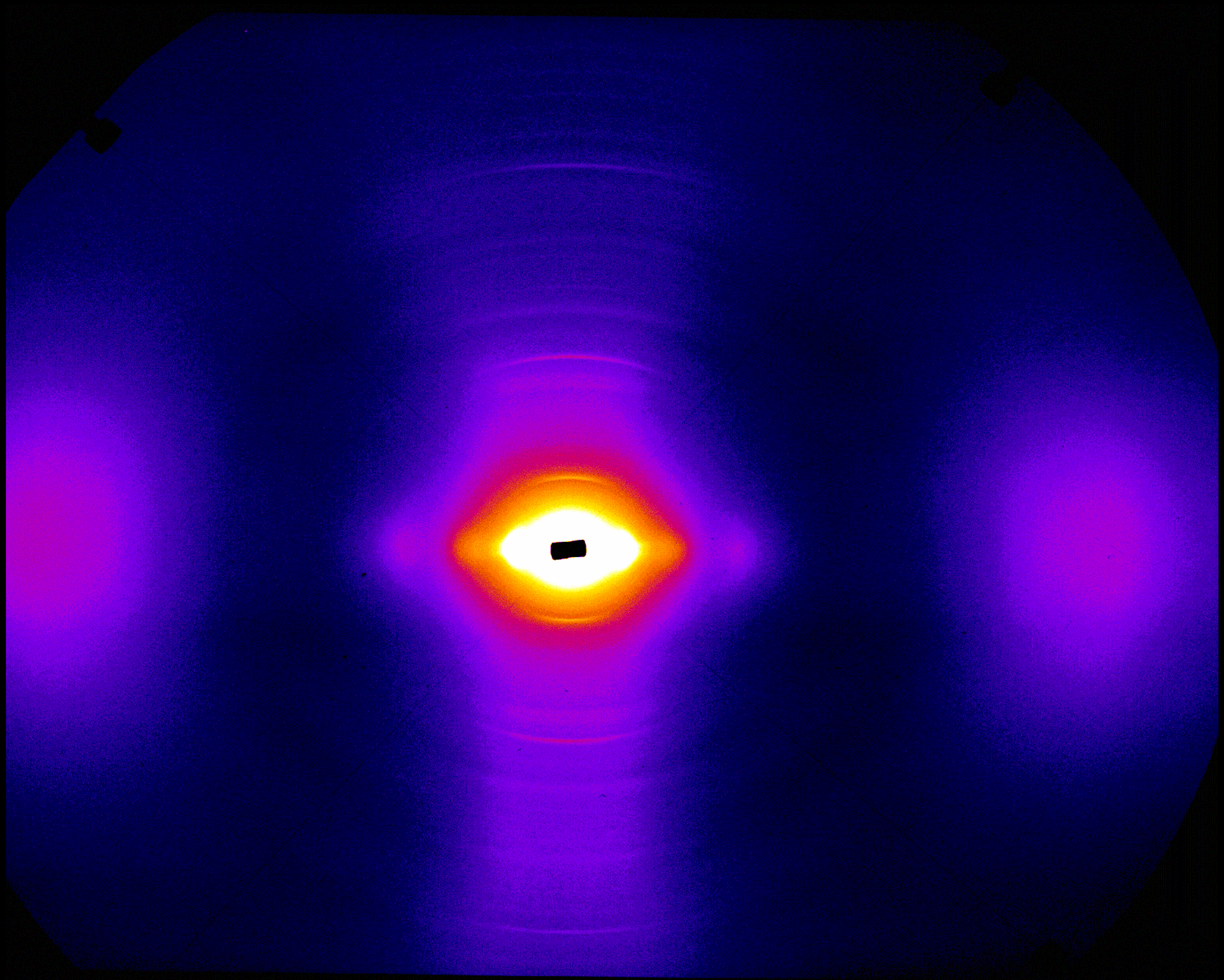
18 residues/5 turns

l=18m+5n

Visible Layer lines (m=0)

every 5th $(27Å / l)^{-1}$

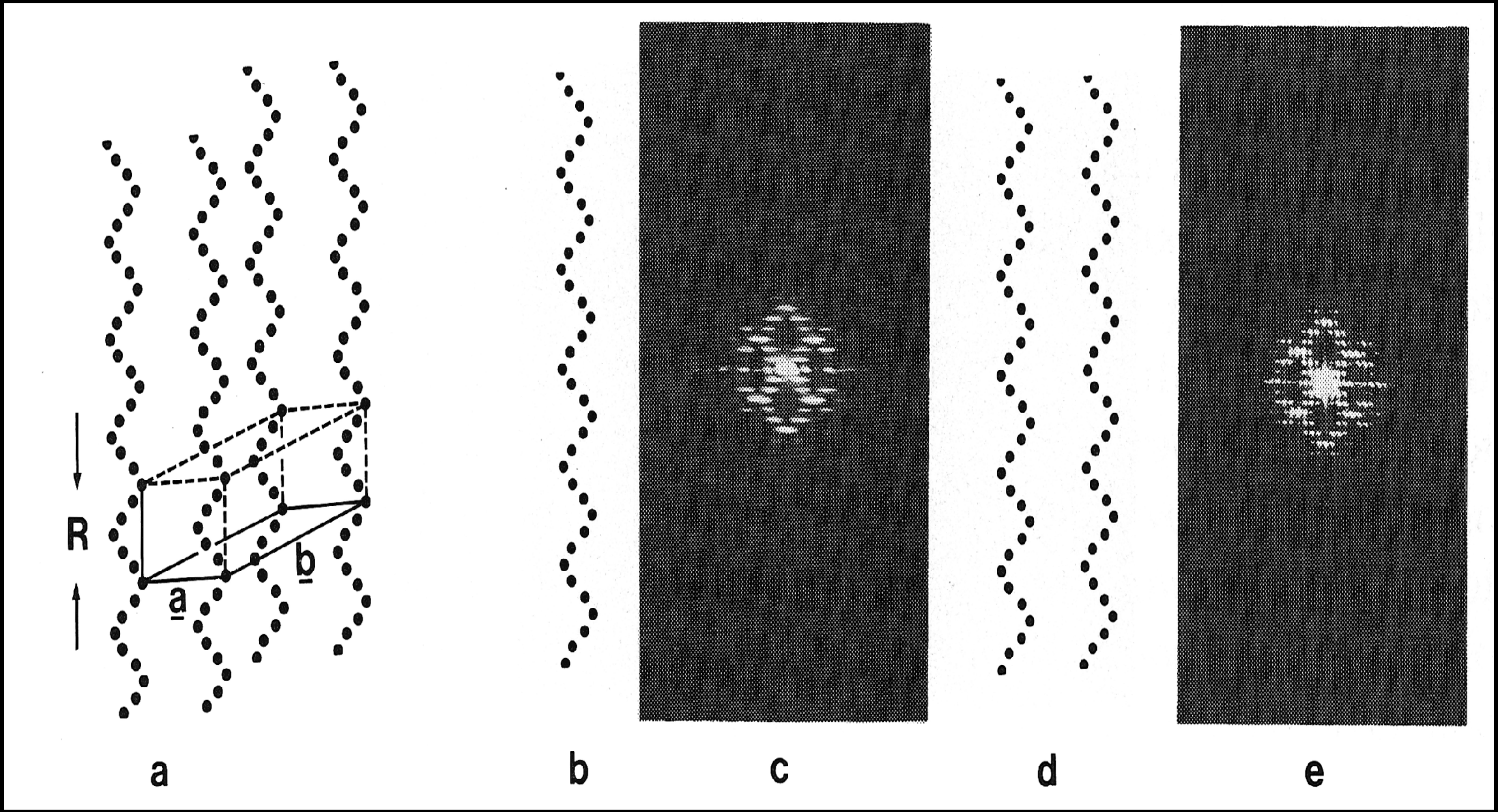
Keratin Pattern from Hair



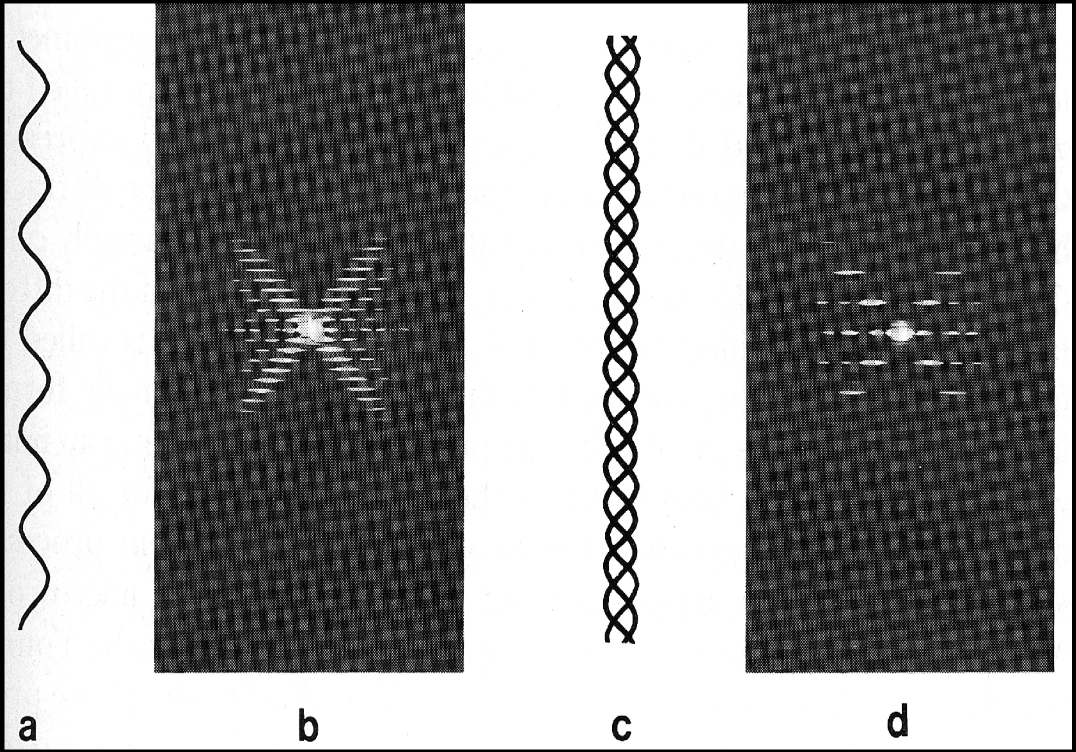
Numerous families of meridionals from different repeating structures

Strong equatorial from Bessel functions from 2-3 nm protofilaments

Crystals of Helical Molecules



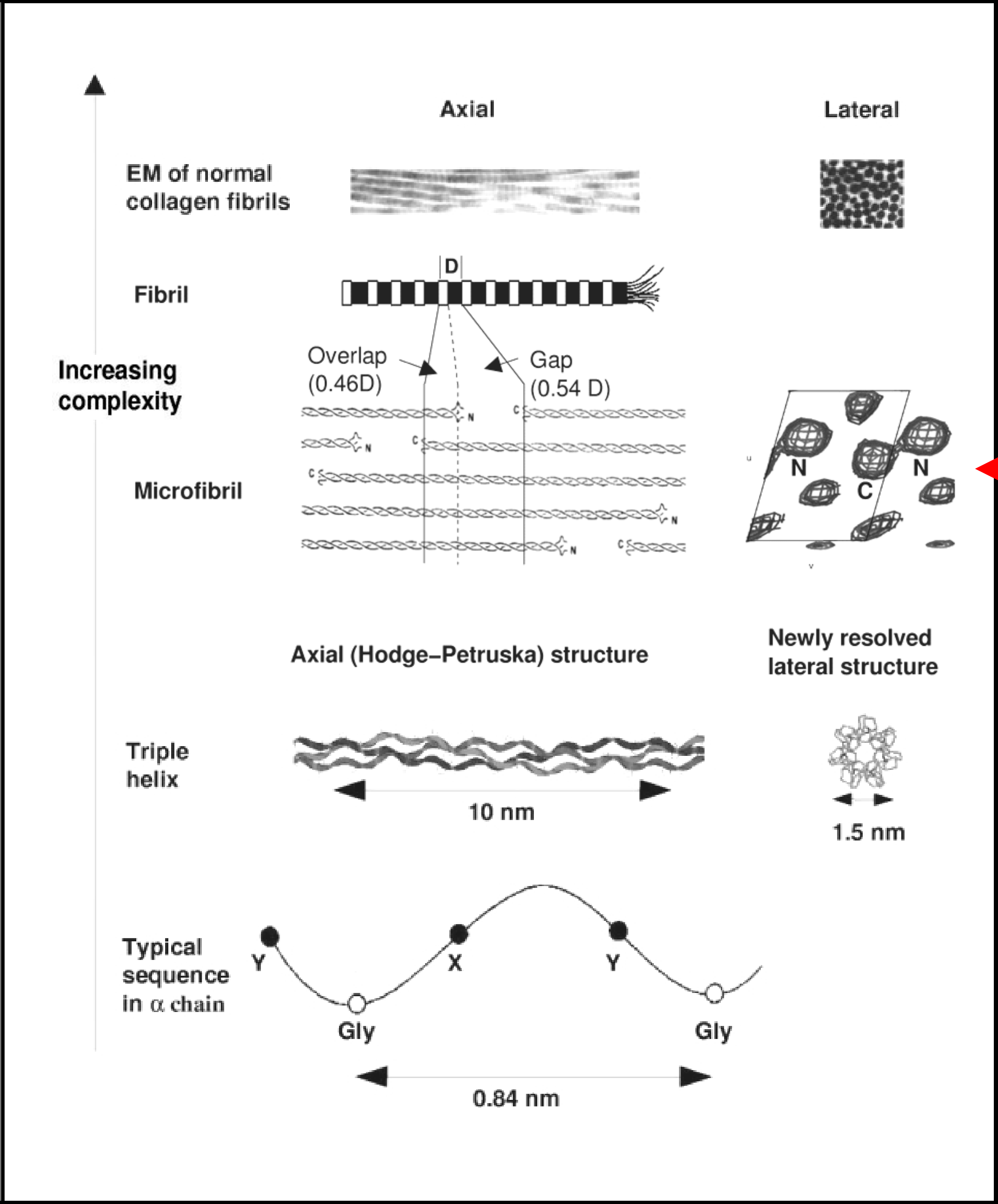
Multi-Stranded (coiled coil) Helices



If N strands
Only every
Nth Layer
-line allowed

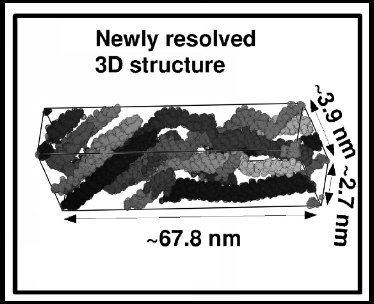
Fiber Crystallography

- Most fiber "structures" result of model building studies
- There have been a small number of Fiber "structure solutions".
- High resolution structure by Keichi Namba on bacterial flagella (Yamashita et al., 1998 Nature SB) aligned by high magnetic fields
- Orgel et al. (2001 Structure) published first MIR structure of a natural fiber - Type I Collagen from rat tail collagen (2006 PNAS gave full structure)



Supermolecular

or



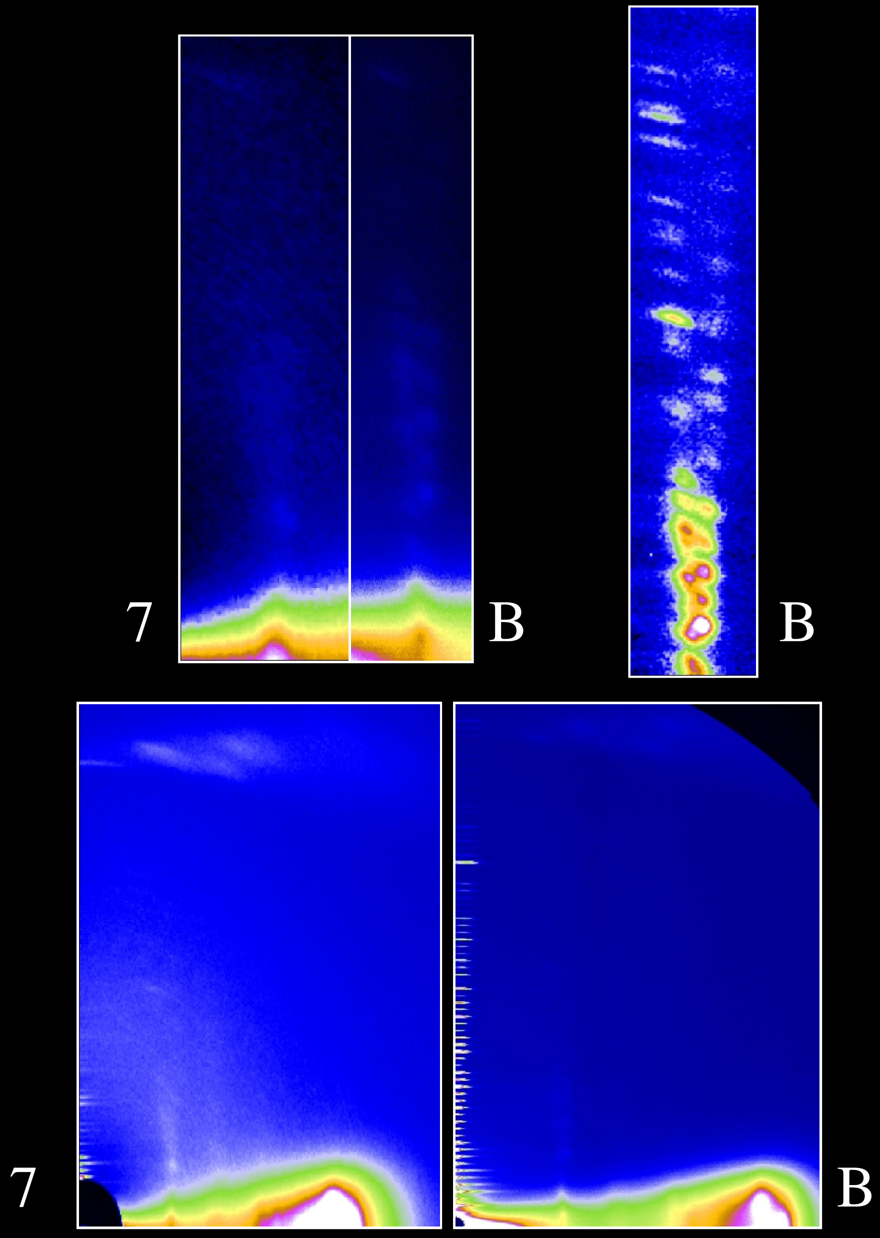
Lateral

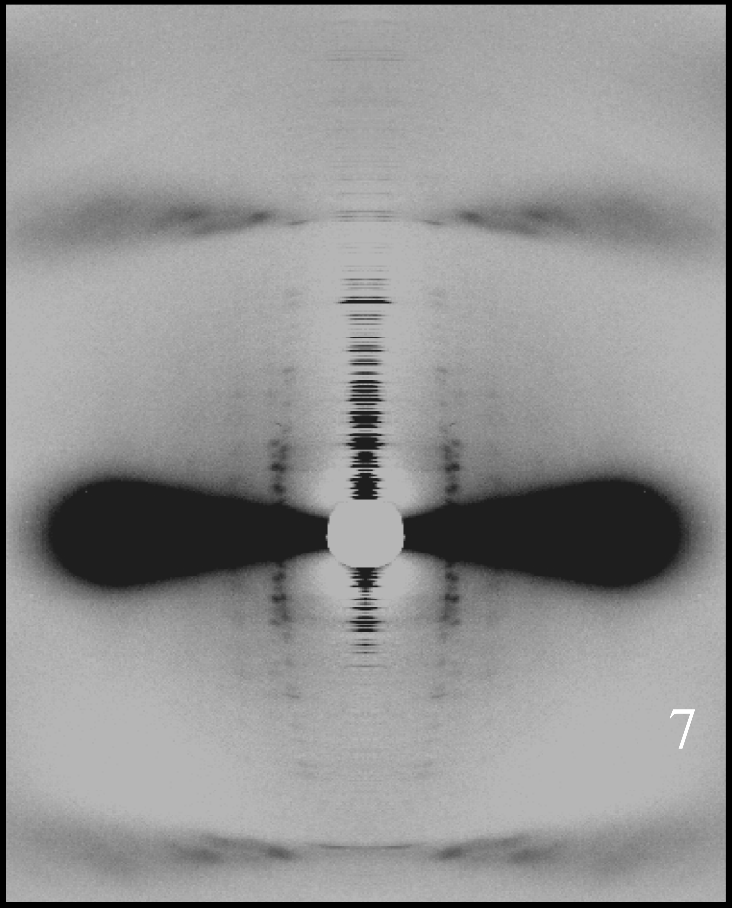
or

Packing structure

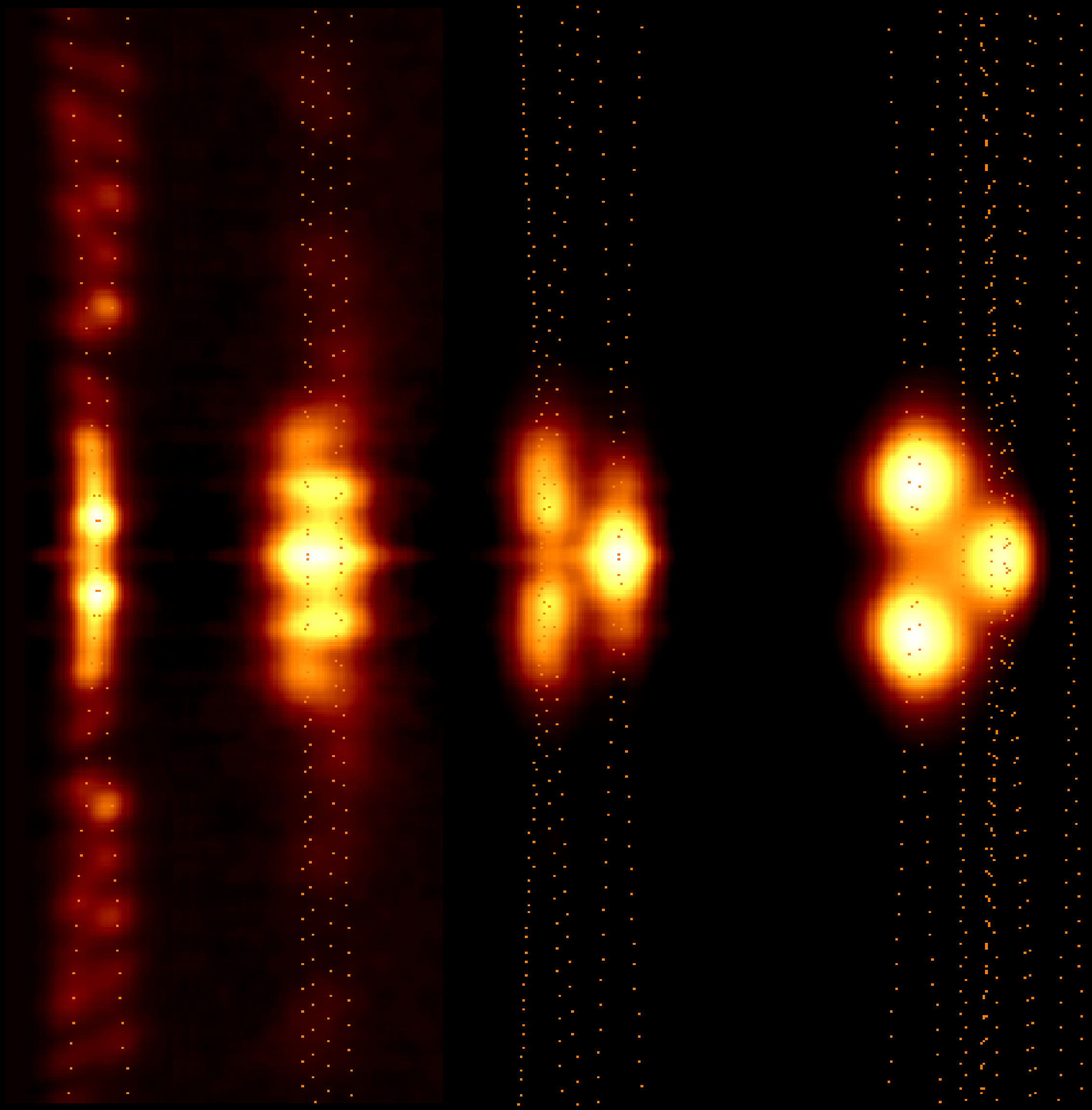
Orgel *et al.*, 2001, 2002

Data collection:

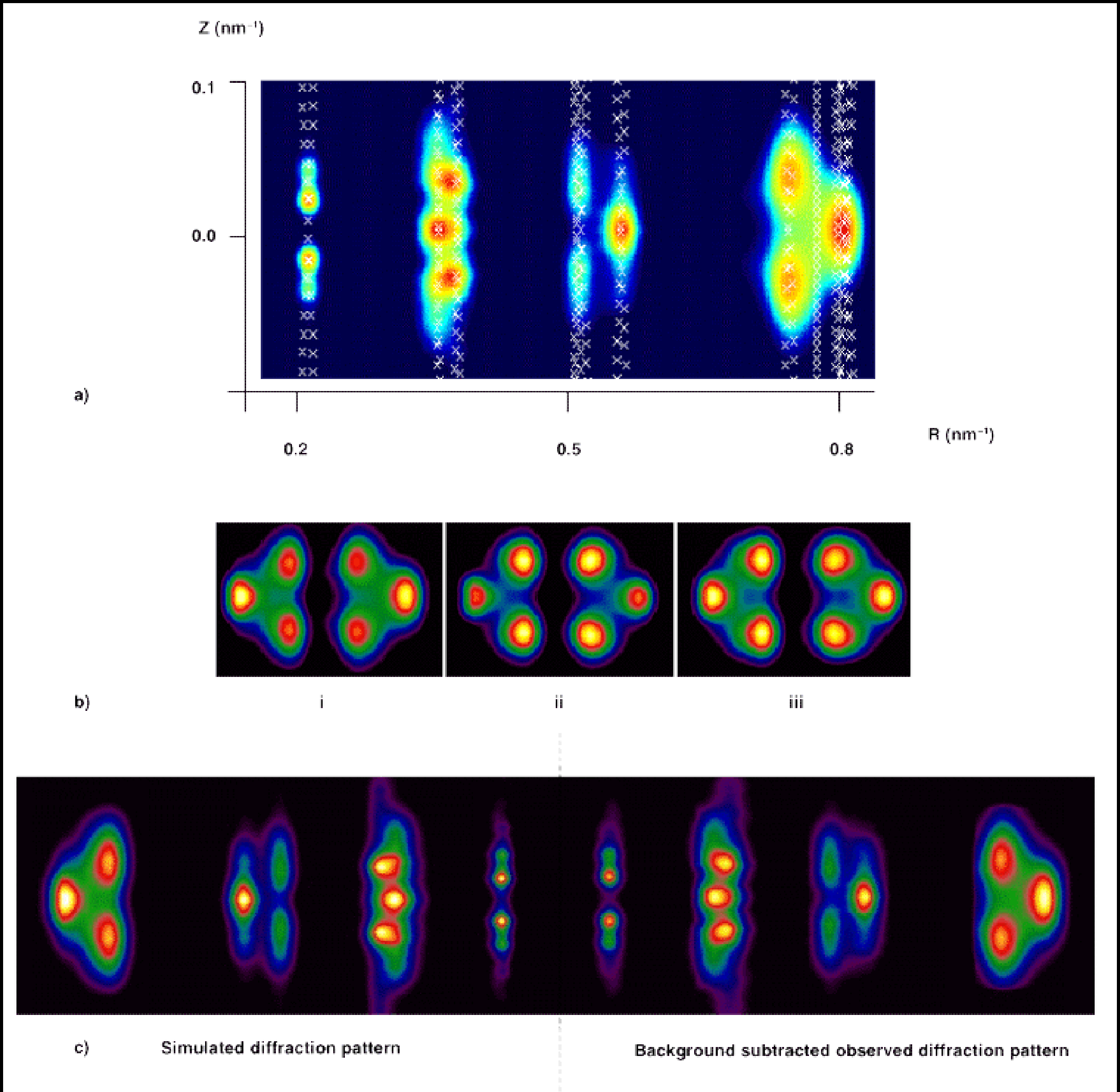




B=Bio-CATid
7=7.2bm

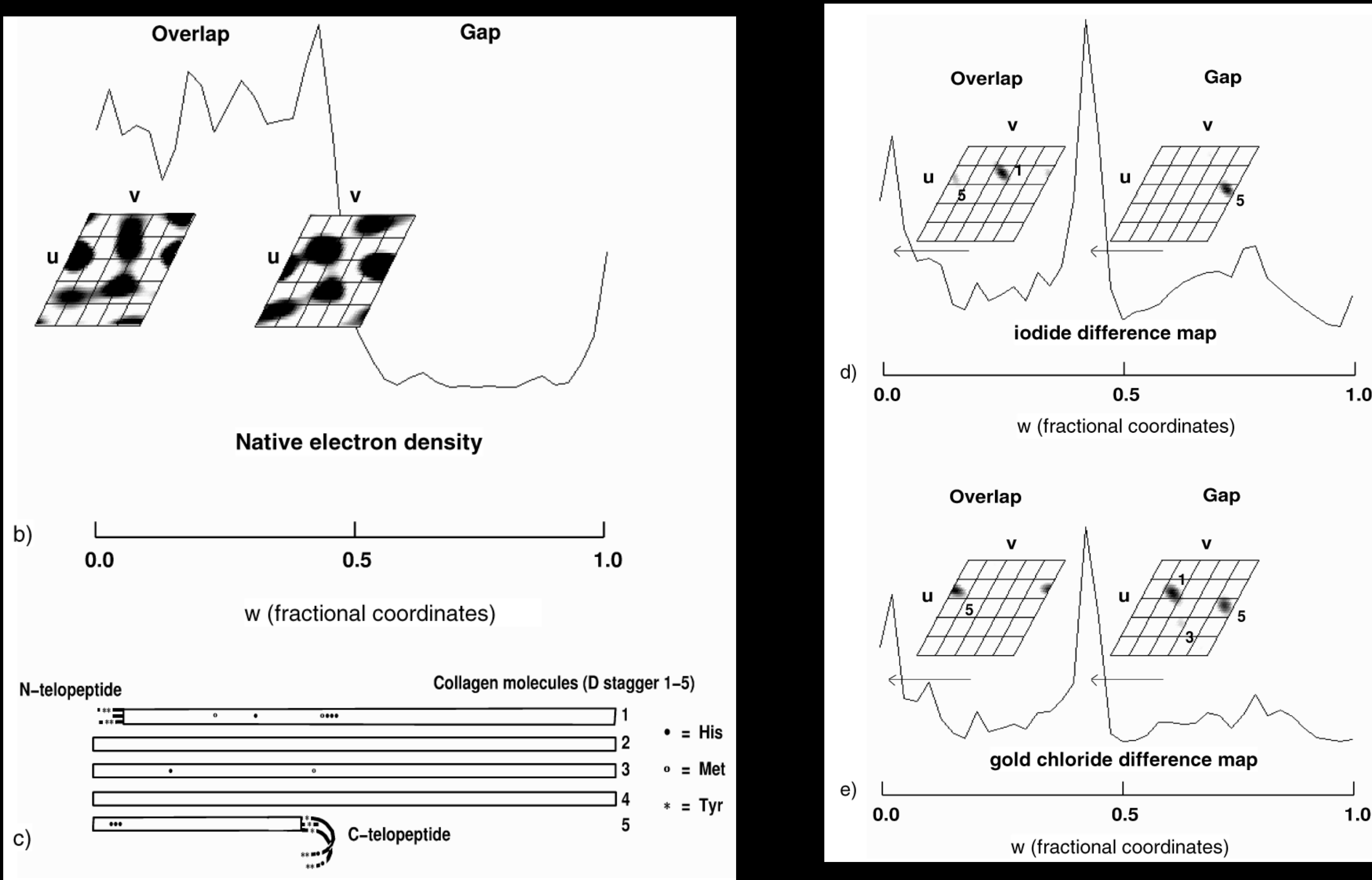


Background subtracted image of row lines showing closely-spaced but resolved diffraction spots

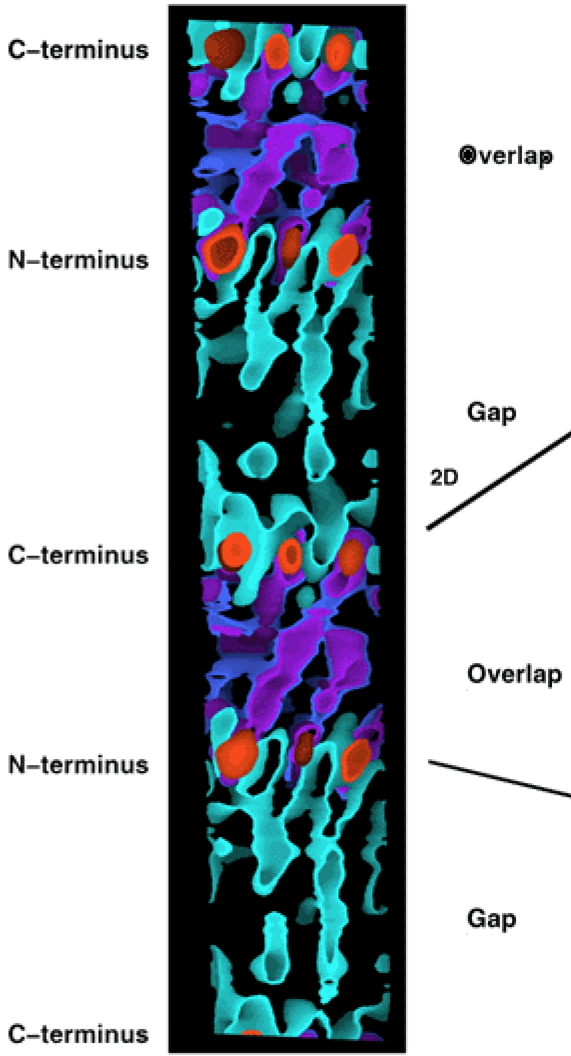


Data Extraction

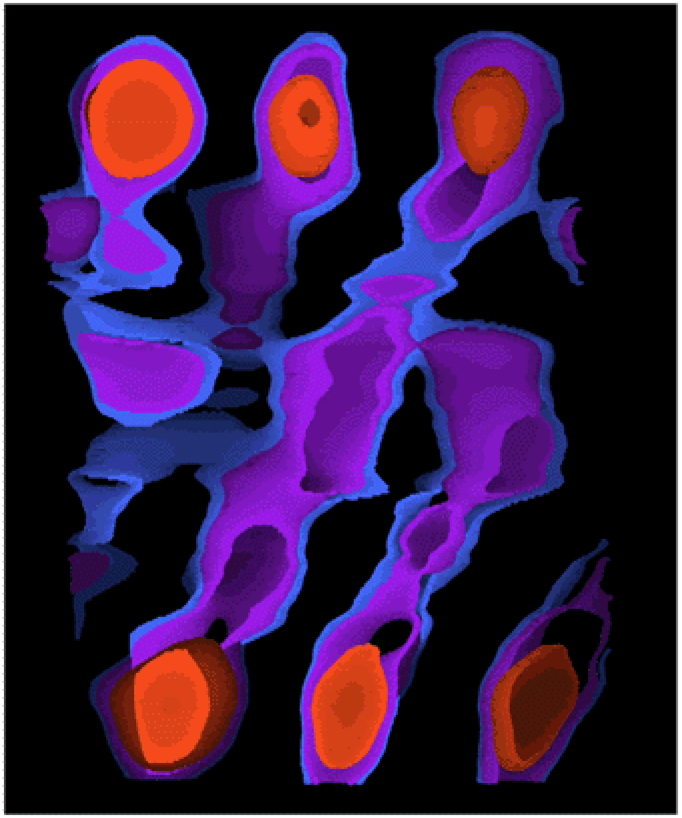
from native and derivative tendon diffraction patterns



Overlap
Gap
v
u
v
u
Native electron density
b)
0.0
0.5
1.0
w (fractional coordinates)
N-telopeptide
Collagen molecules (D stagger 1-5)
1
2
3
4
5
• = His
o = Met
* = Tyr
c)
C-telopeptide
Overlap
Gap
v
u
1
5
v
u
5
iodide difference map
d)
0.0
0.5
1.0
w (fractional coordinates)
Overlap
Gap
v
u
5
v
u
1
5
3
gold chloride difference map
e)
0.0
0.5
1.0
w (fractional coordinates)

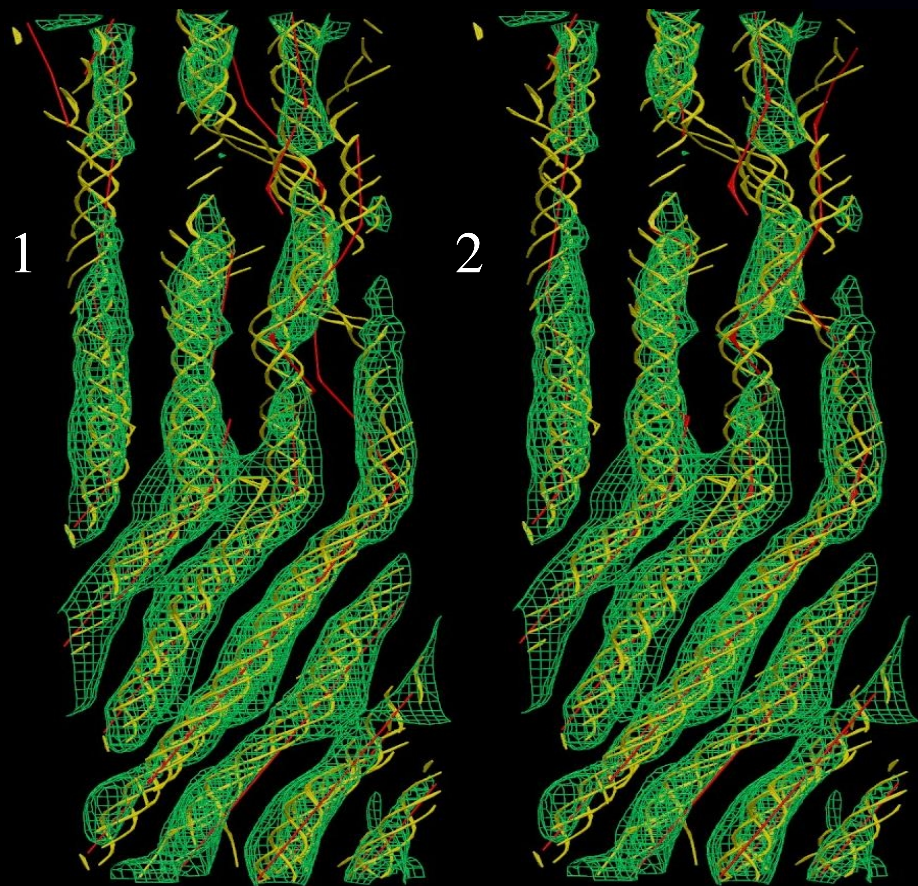

C–terminus
Overlap
N–terminus
Gap
2D
C–terminus
Overlap
N–terminus
Gap
C–terminus

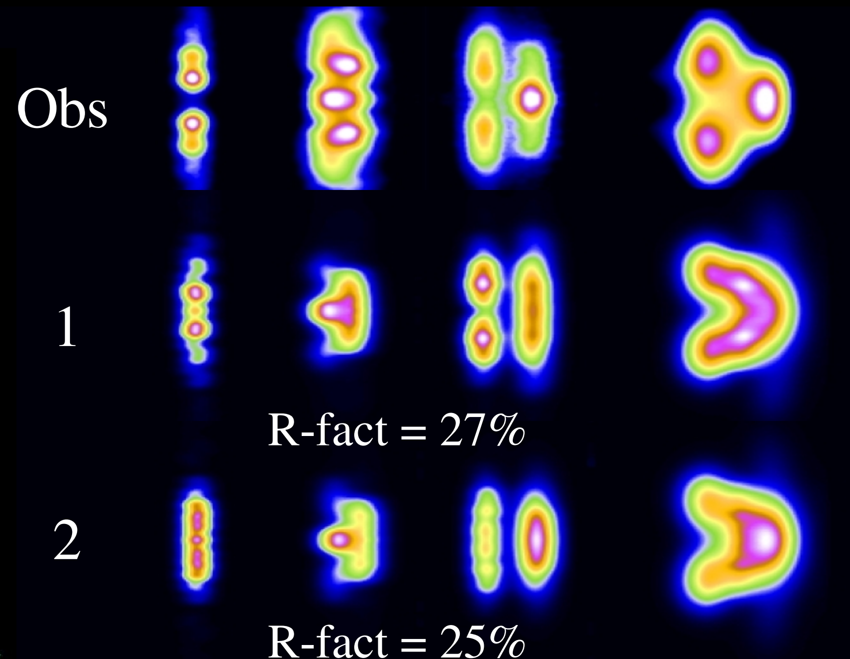
a)



b)

Model to data fitting stage I

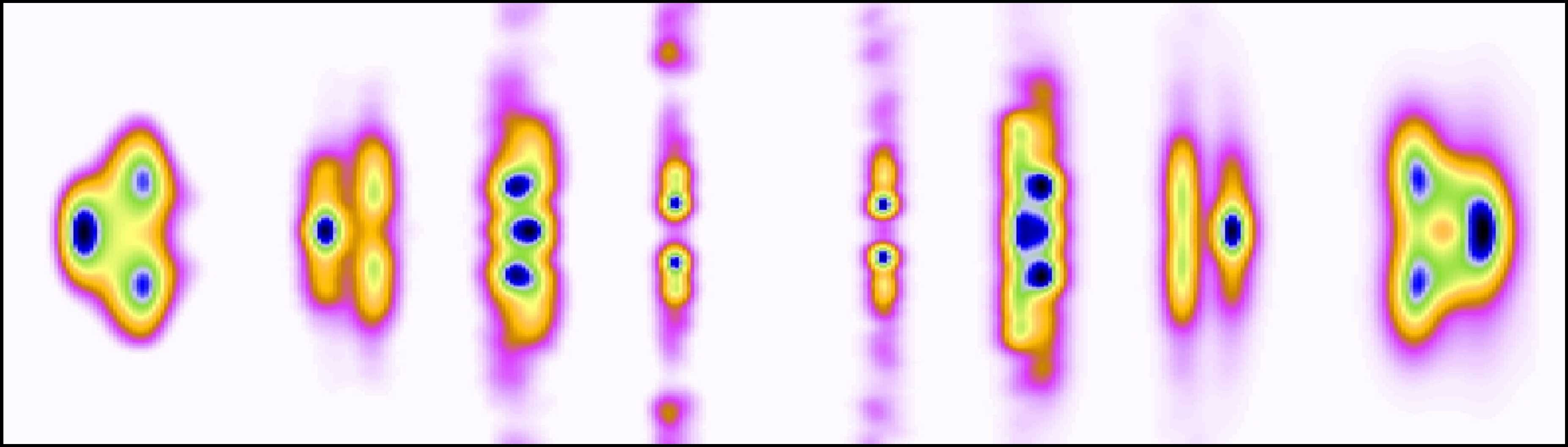




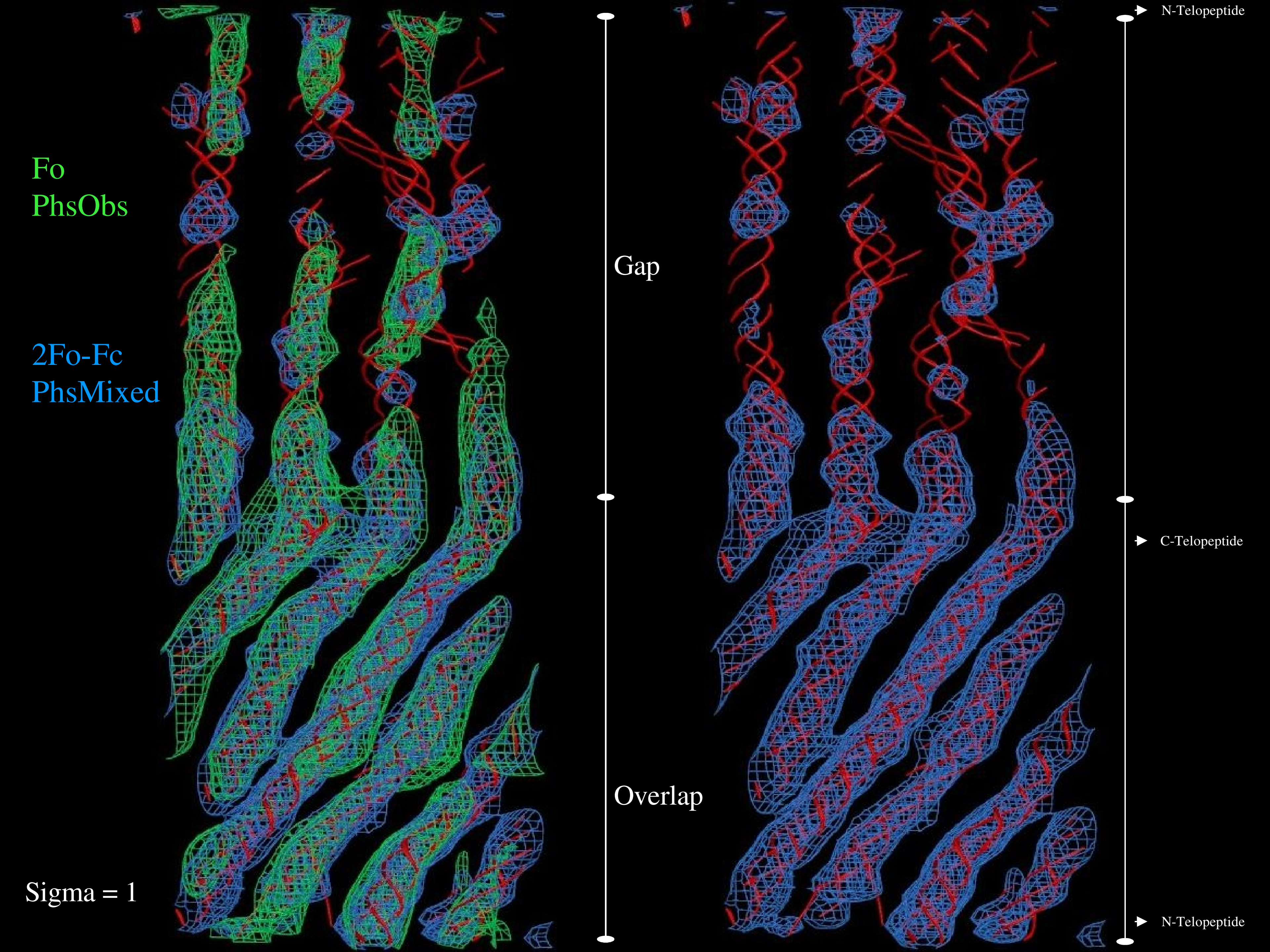
Test model path

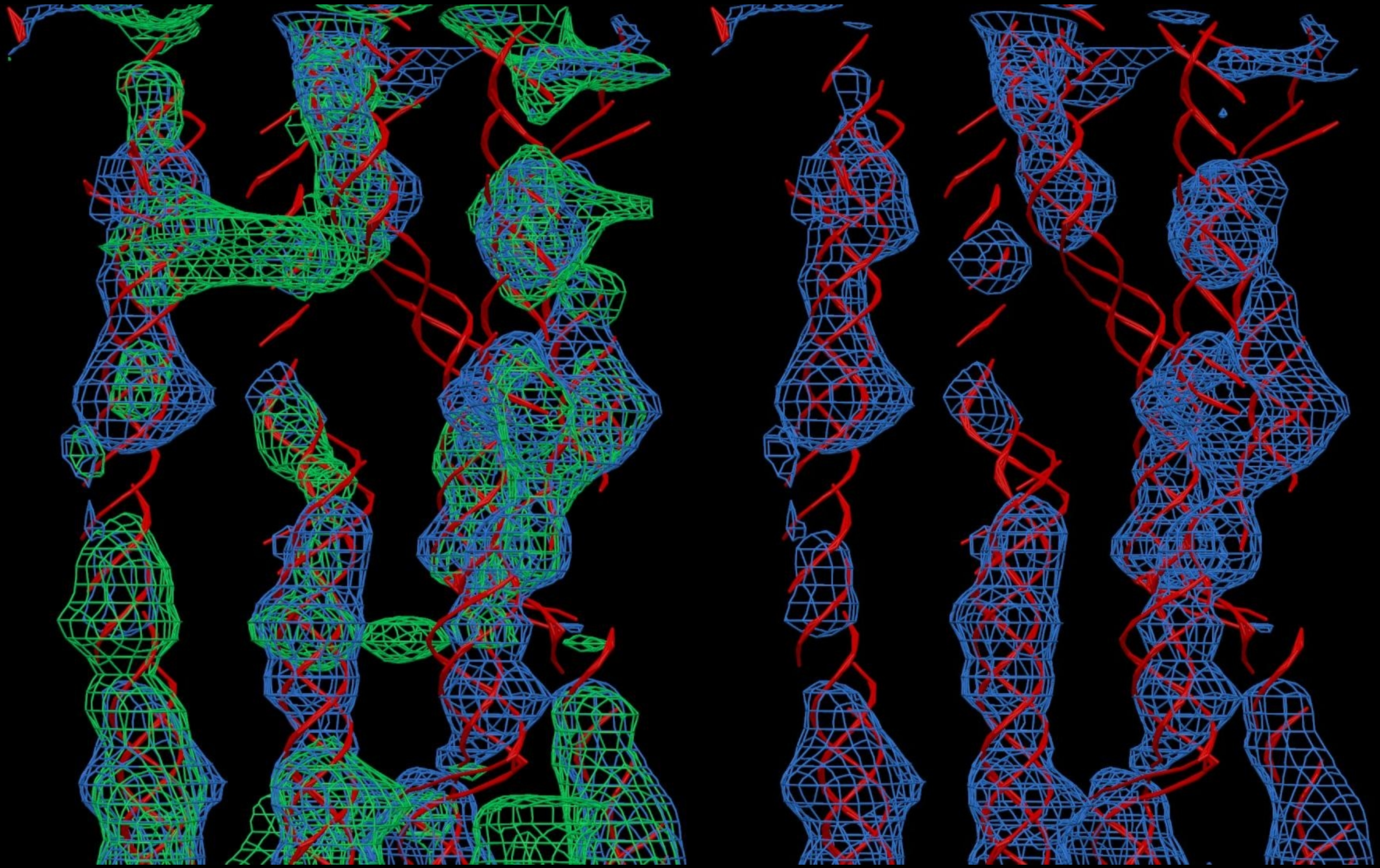
Final model path

Comparison of model to observed data



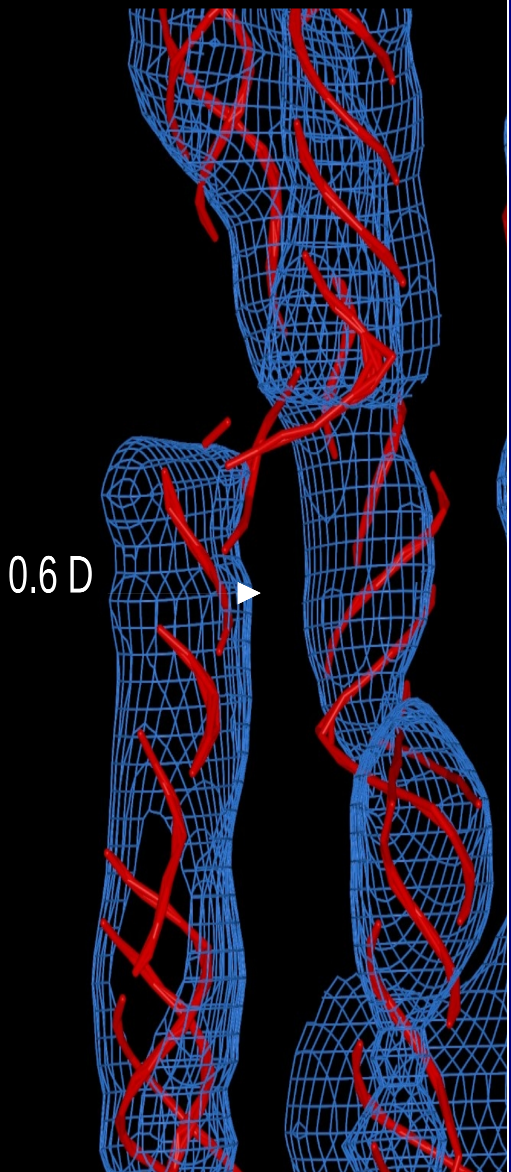
R-factor ~ 16.7% / 9.7 %

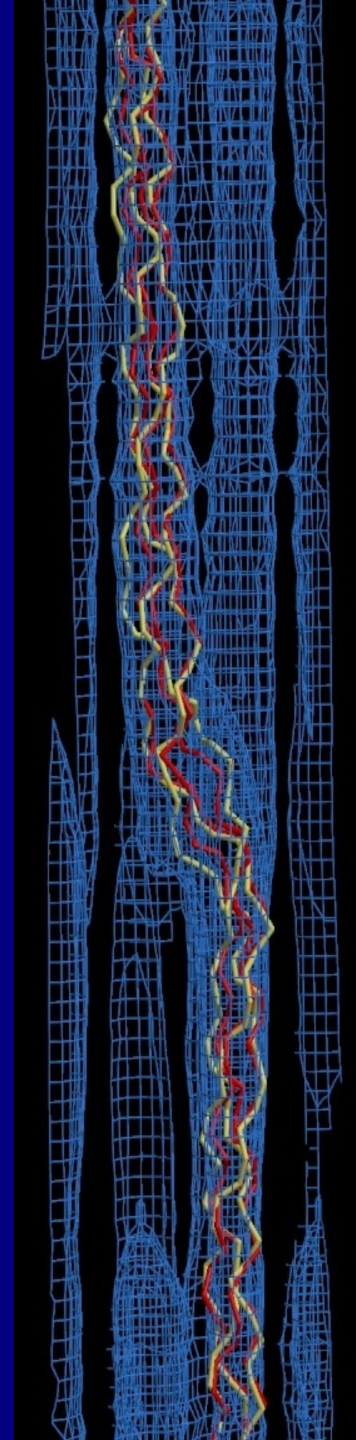

Fo
PhsObs
2Fo-Fc
PhsMixed
Sigma = 1
Gap
Overlap
N-Telopeptide
C-Telopeptide
N-Telopeptide



2Fo-Fc PhsObs 2Fo-Fc PhsMixed

Sigma = 2/3

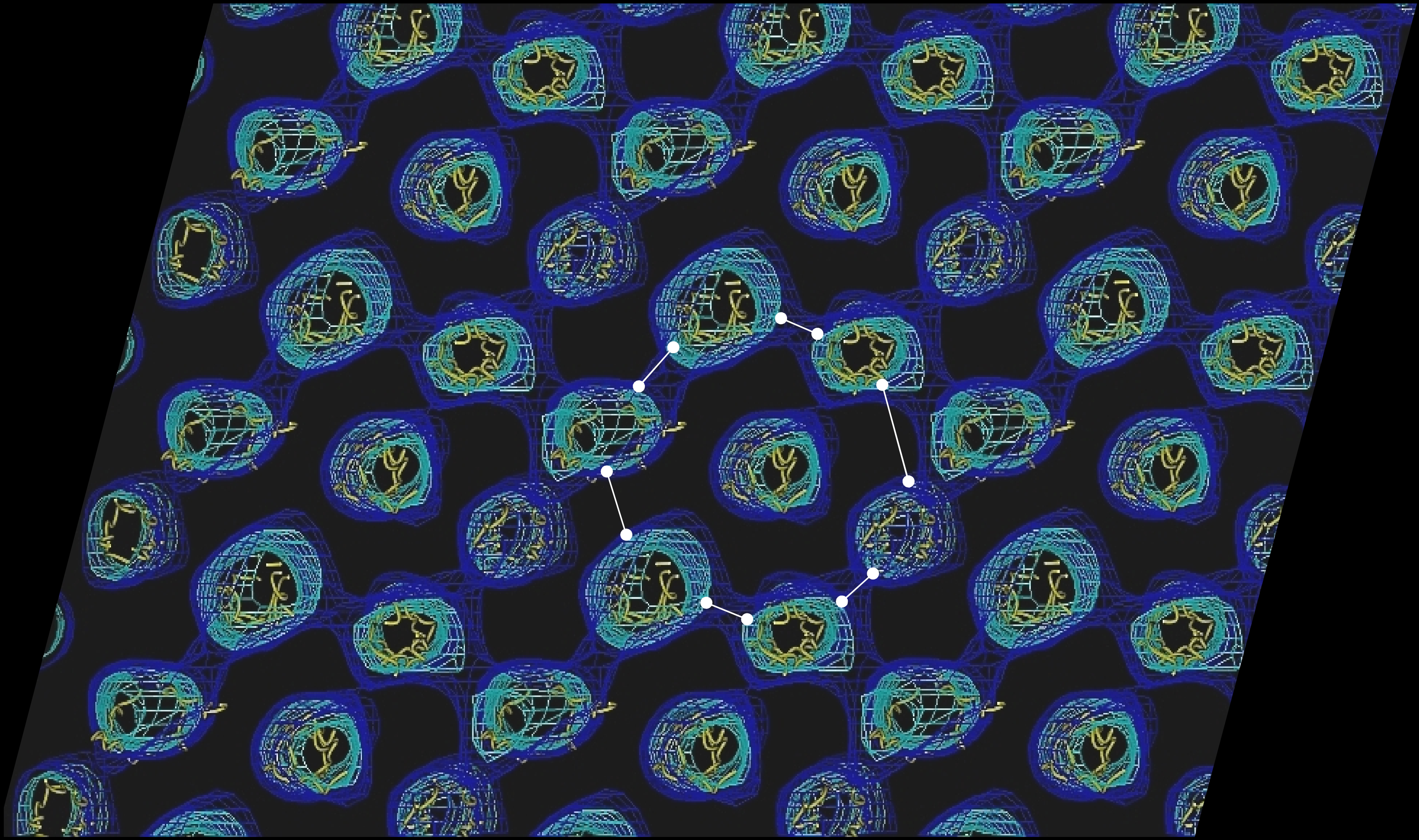




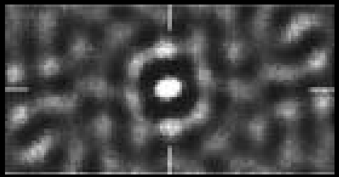
Sigma = 1/2

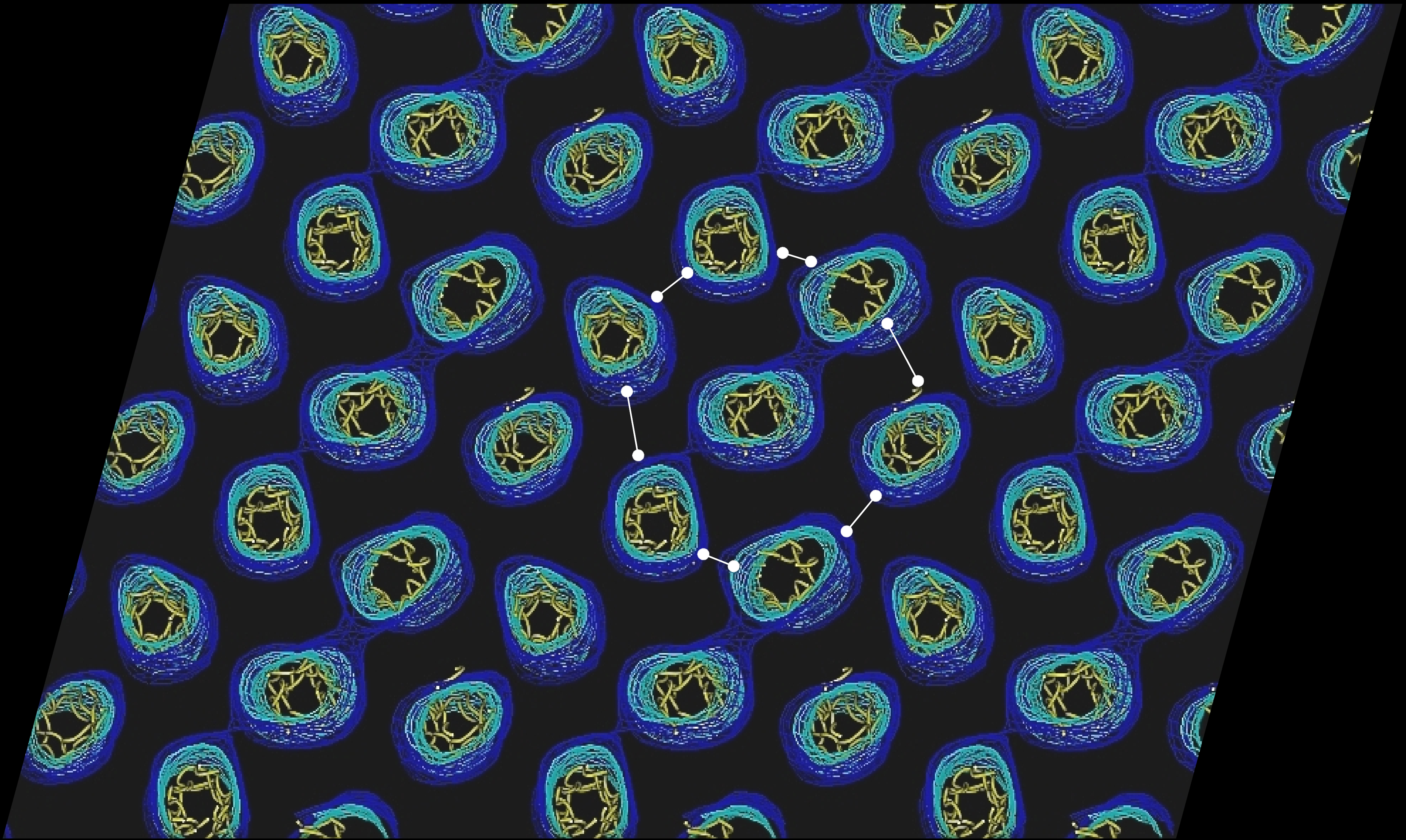
Quasi-hexagonal molecular packing

a common feature of homo and heterotypic fibrils?

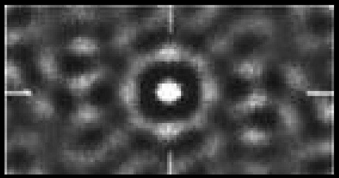


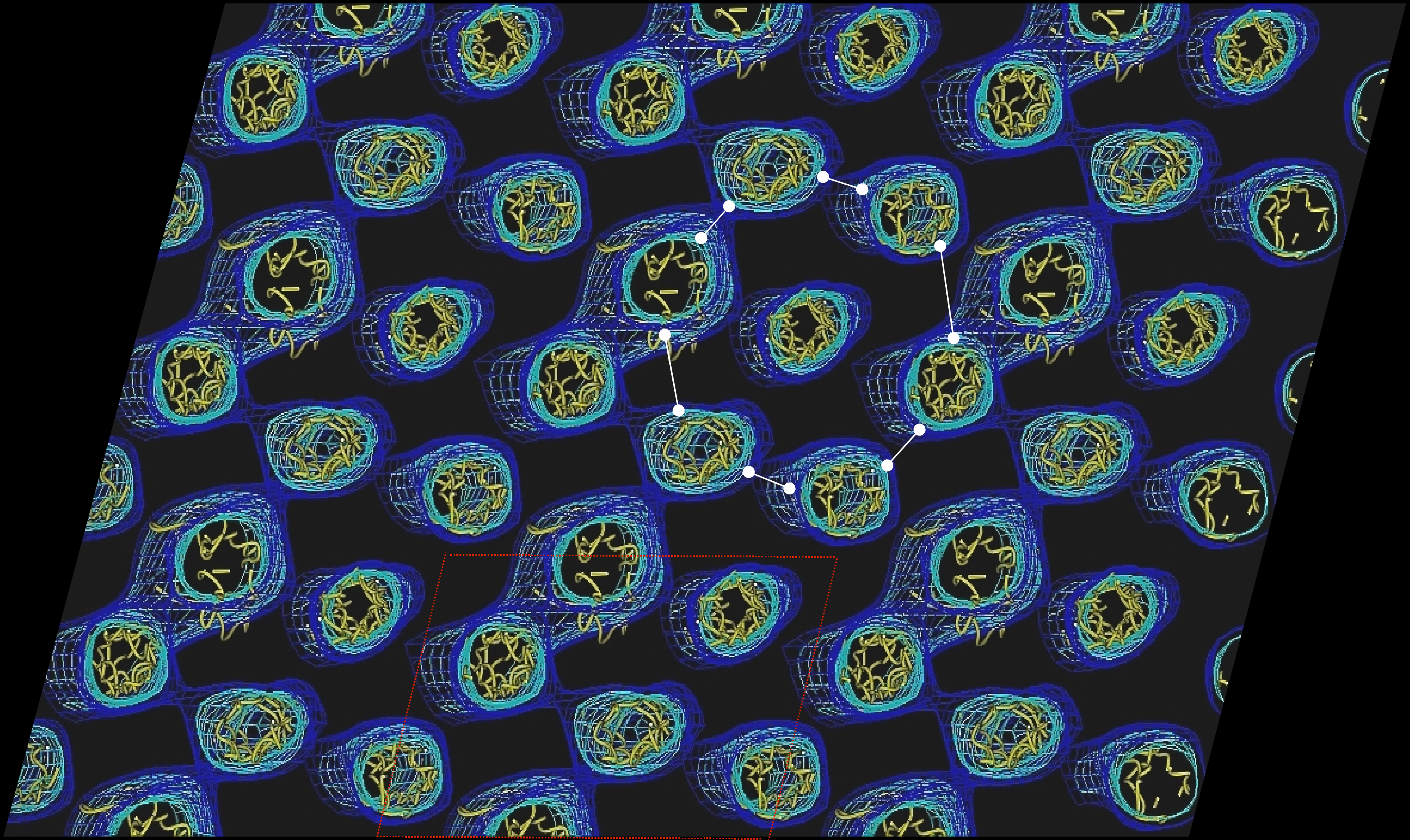
N-Term



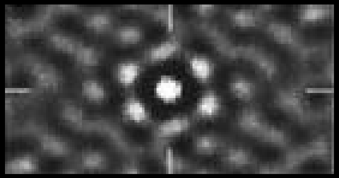


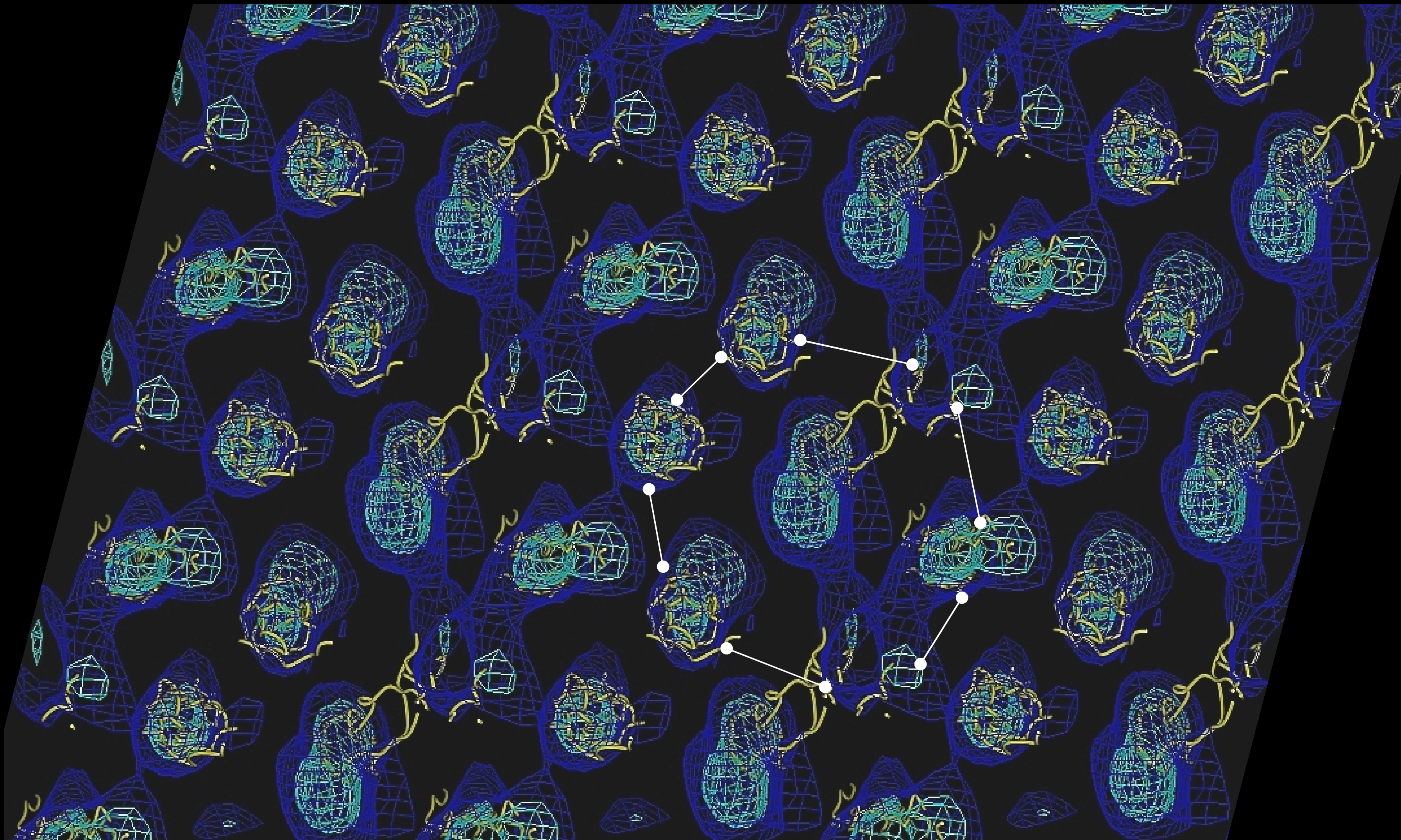
Overlap



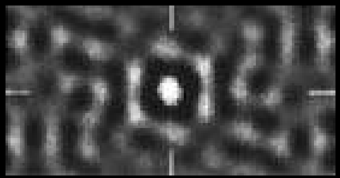


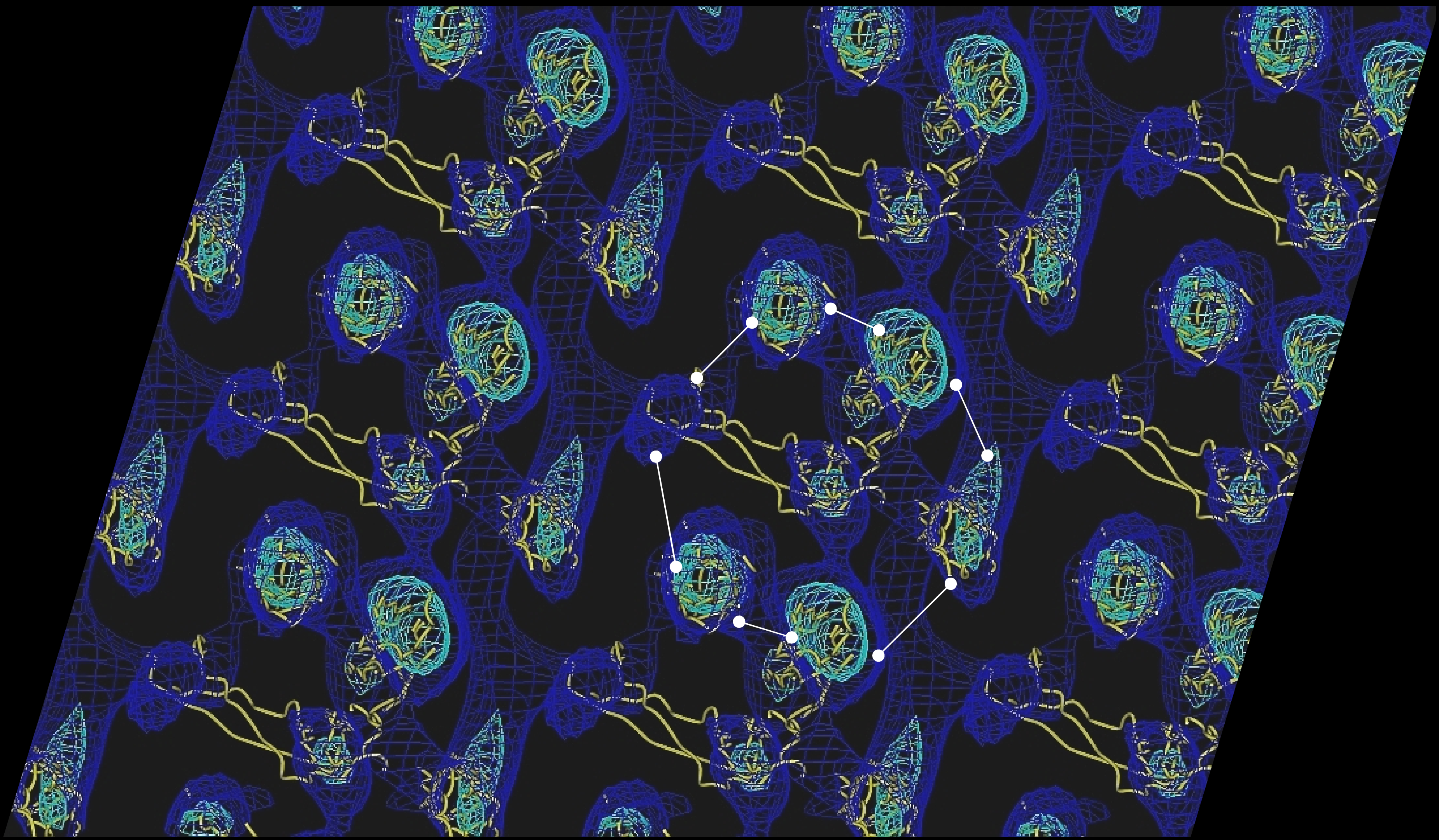
C-Term



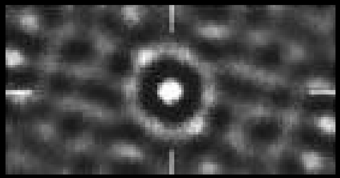


Gap 1

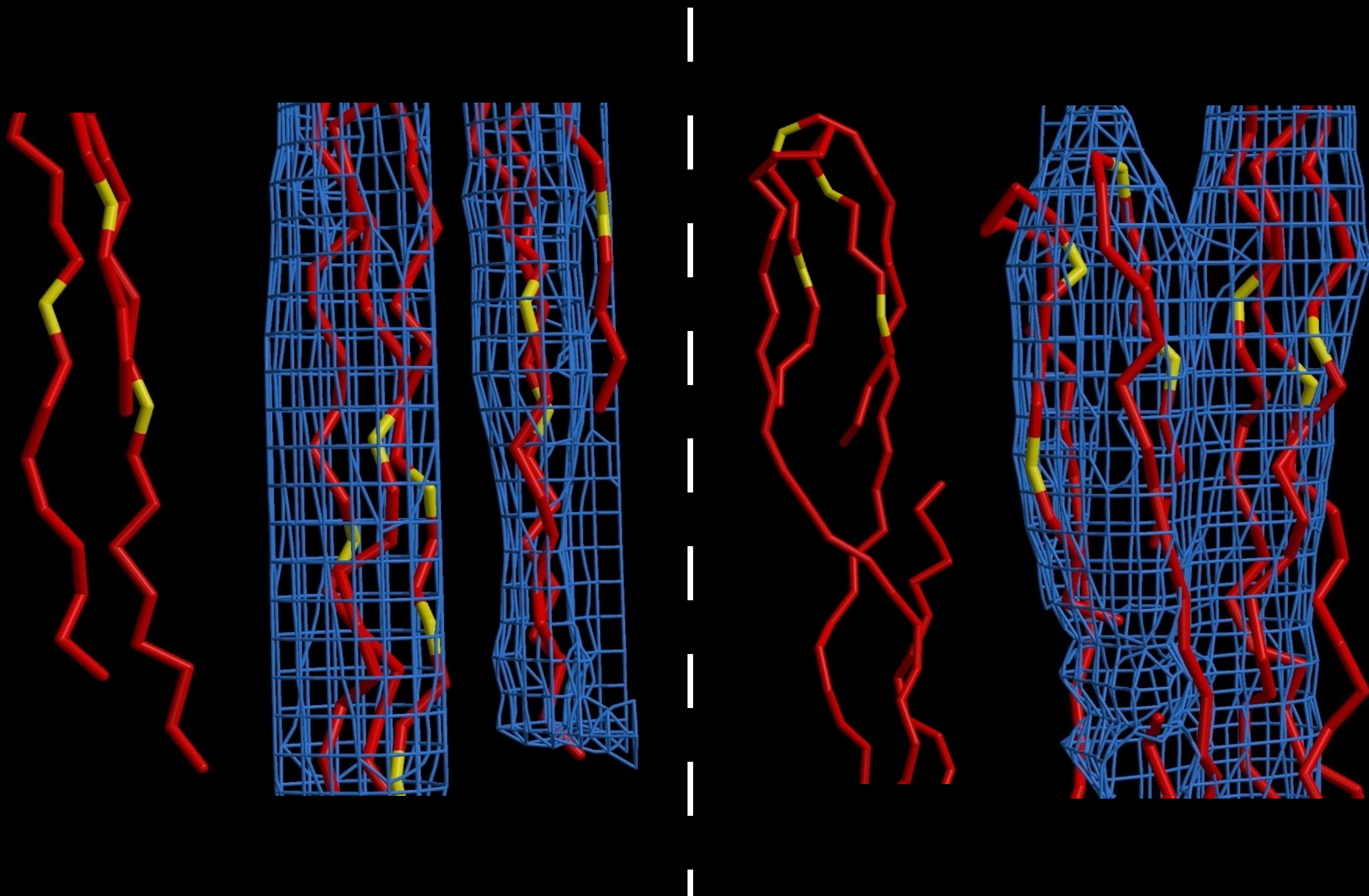




Gap 2



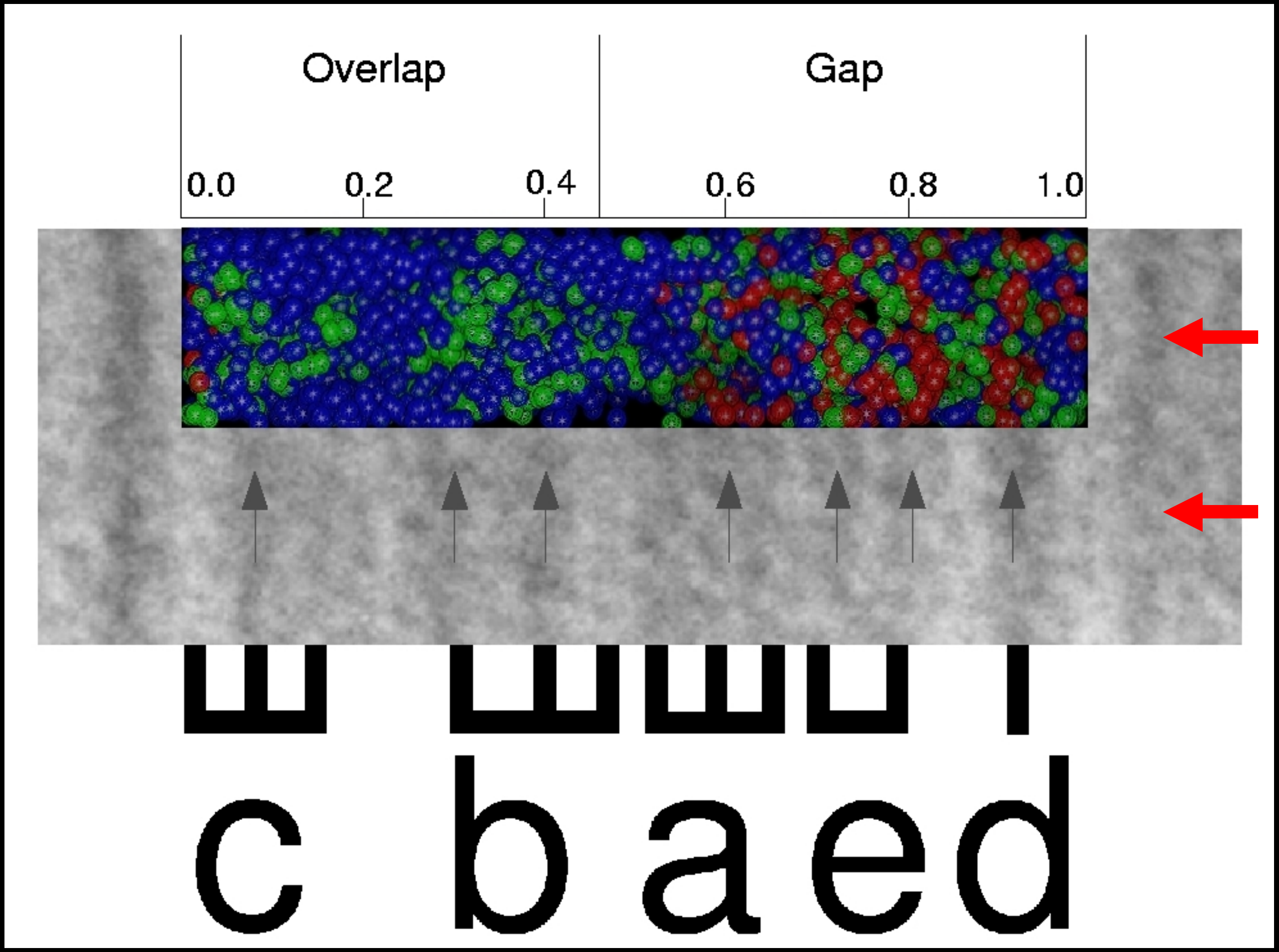
Telopeptides and the intermolecular cross-links



N-telopeptide C-telopeptide

Sub-fibrillar structure

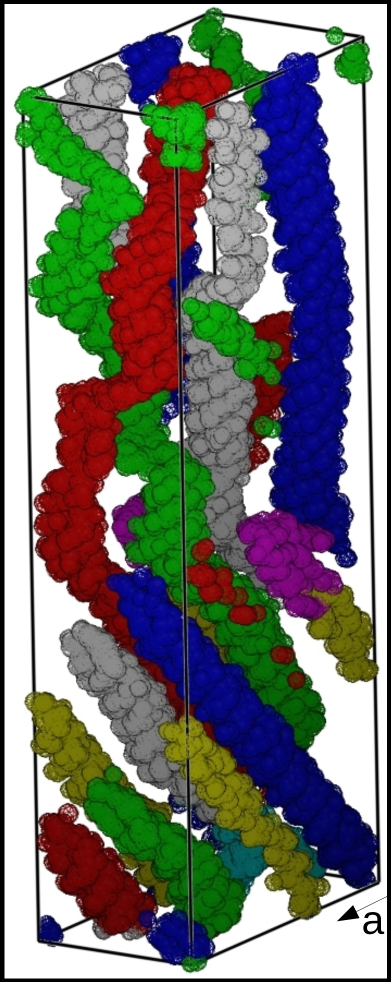
Q-factor value plot along fiber axis

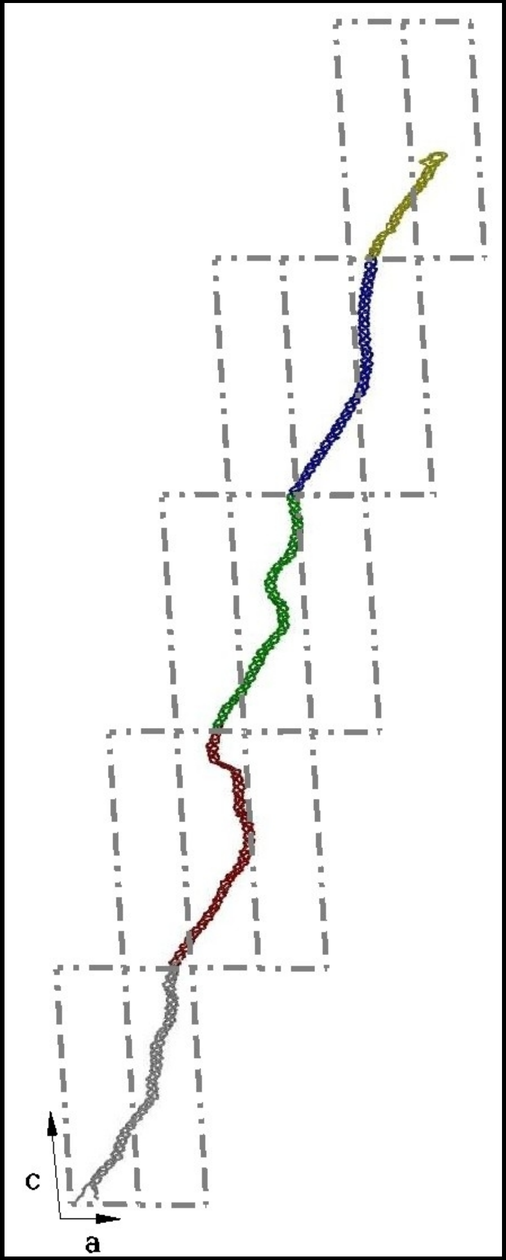


Q-factor (occupancy) plot

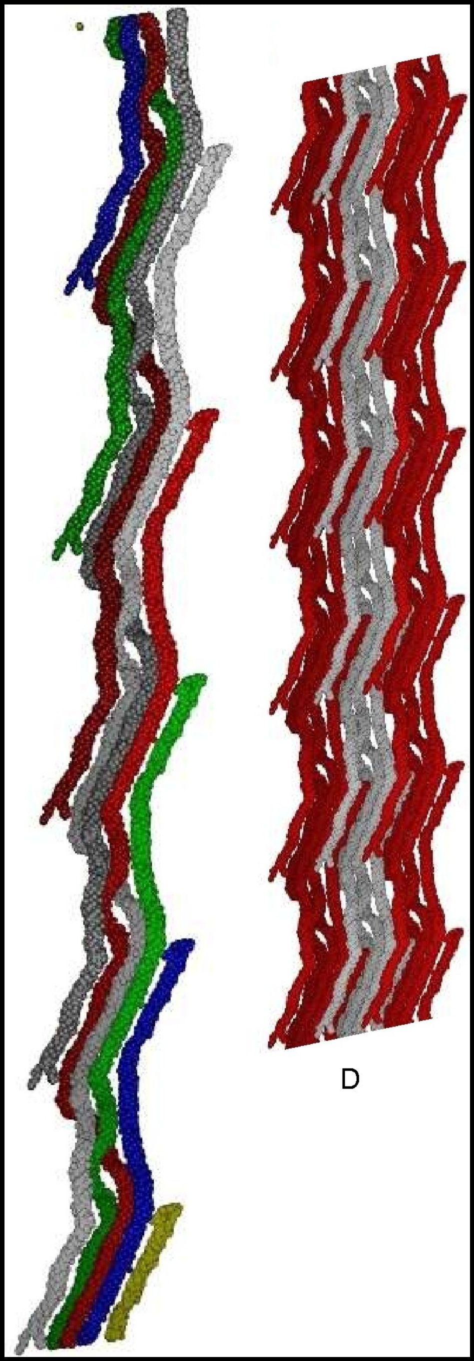
Positive stain EM of collagen fibrils

Collagen type I sub-fibrillar structure





C

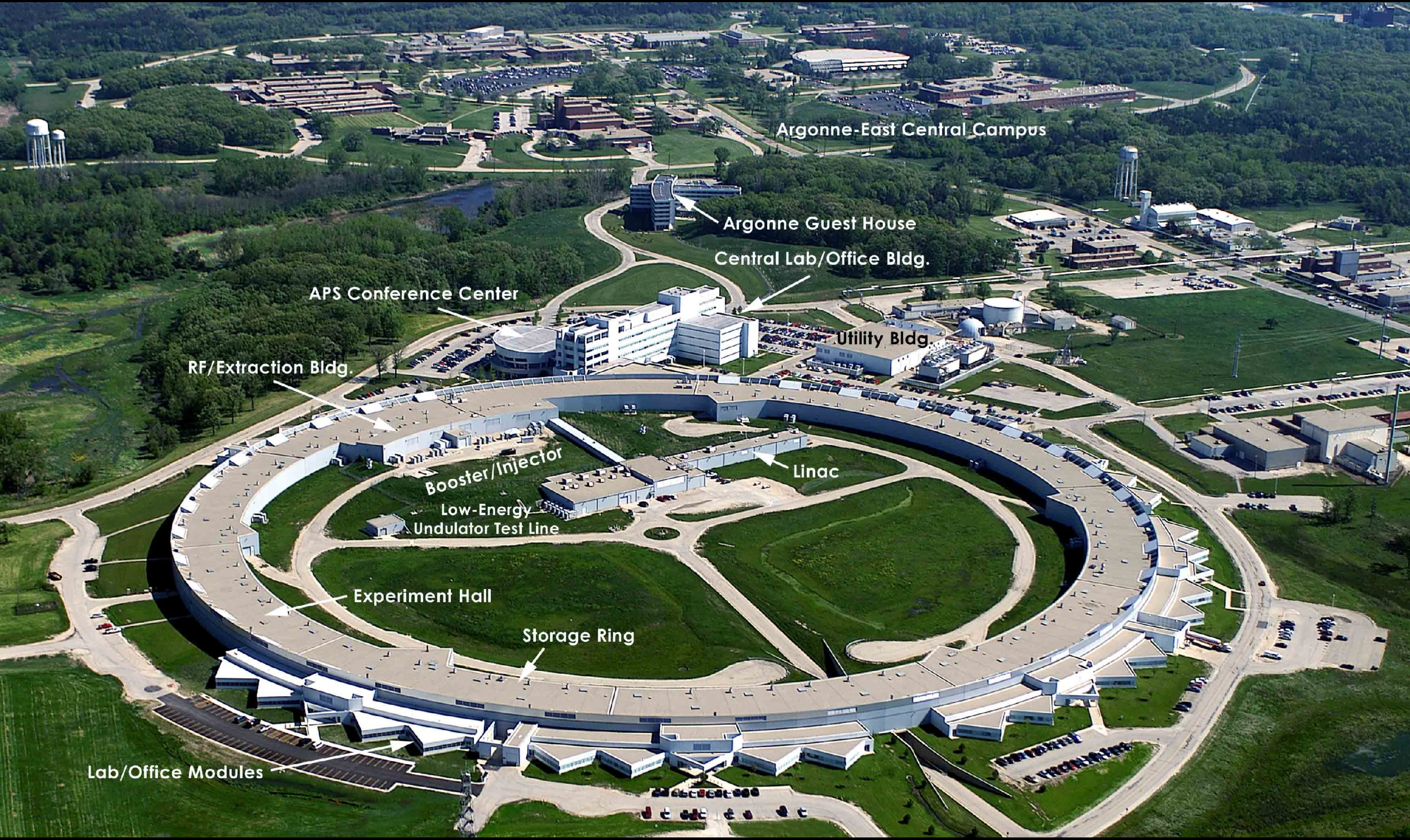


Resolution ~ 11 /5.16 Å

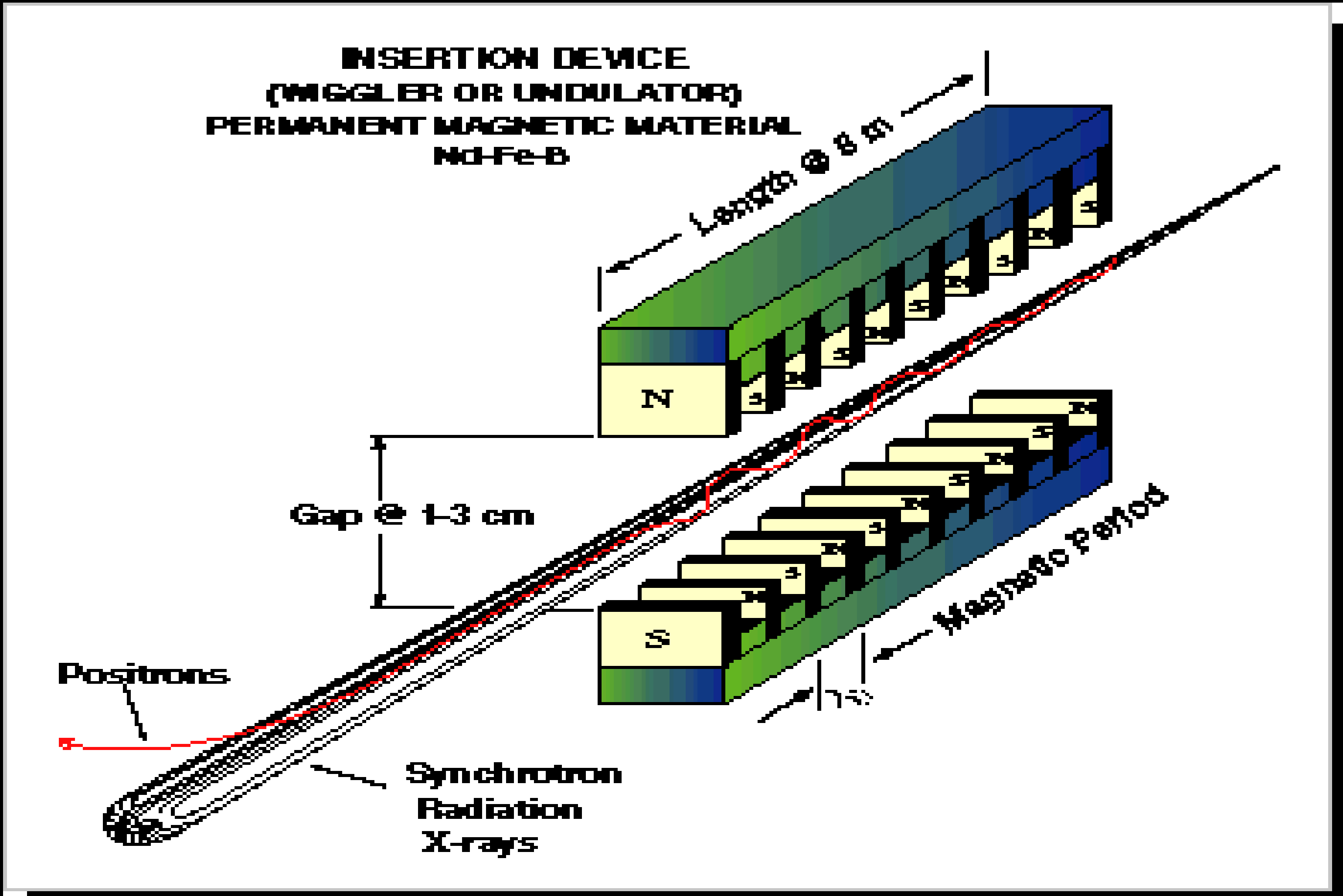
Synchrotrons and Fiber Diffraction

- Early work all done with conventional sources - why need synchrotron?
- Patterns weak, have high backgrounds, frequently have multiple closely spaced lattices
- Studies benefit from greatly increased beam quality
- Greatly increased flux permits time-resolved experiments

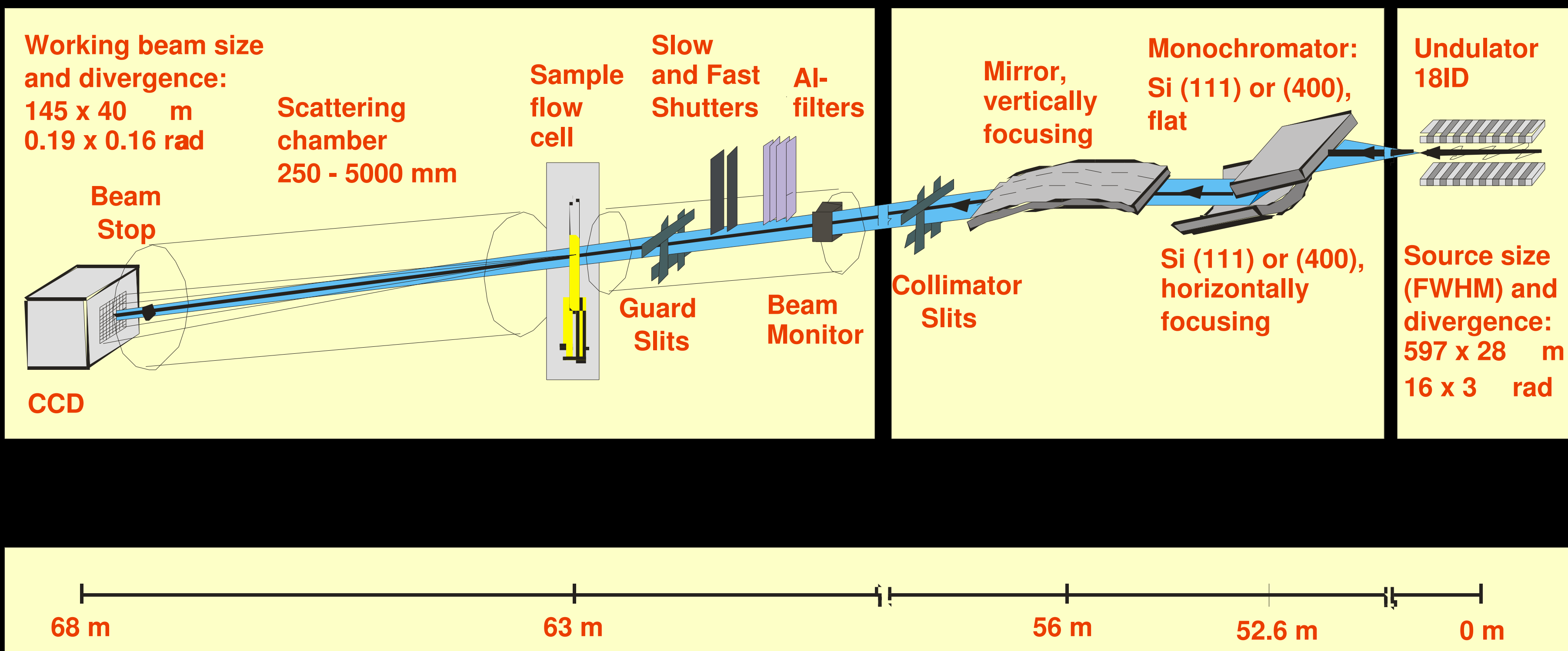
The Advanced Photon Source

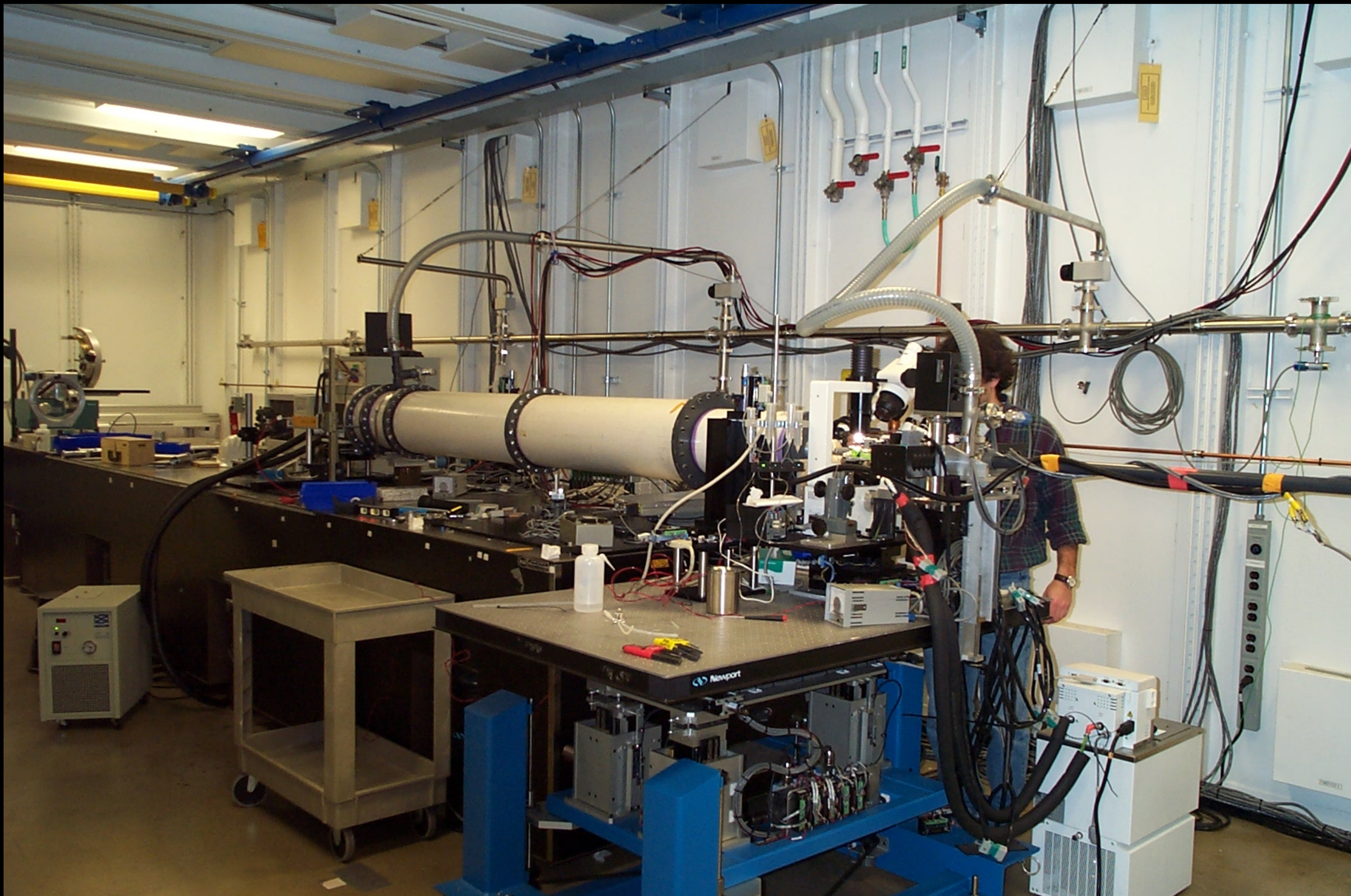


The APS is Optimized for Producing Undulator Radiation



Small-angle Instrument on the BioCAT 18ID - Undulator Beamline

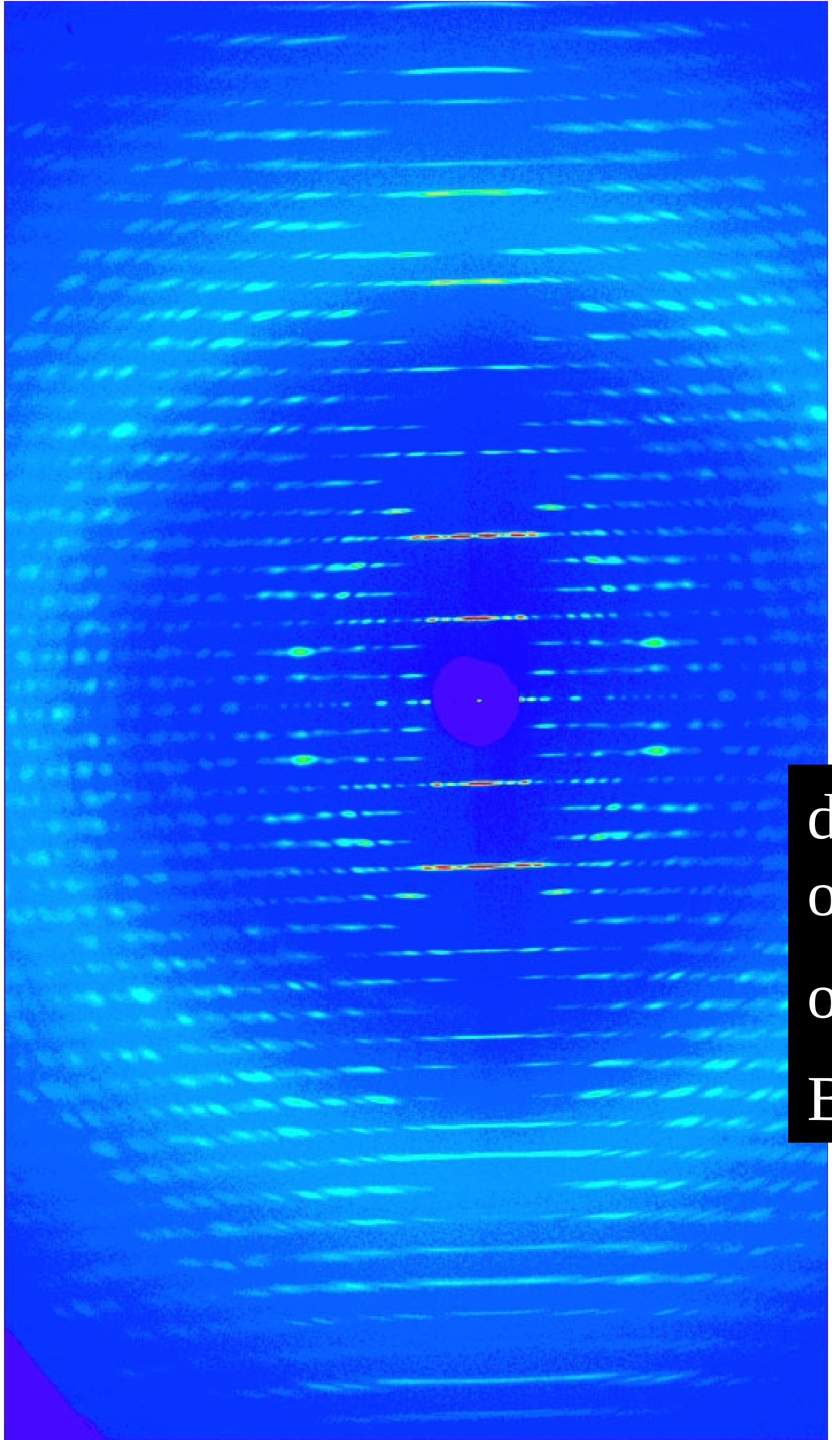




Newport

Undulator Beamline Performance

- Total X-ray flux 1-2.5 x10^{13} photons/s
- Focal spot size ranges from < 40 μm vertical and < 200 μm horizontal to ~3 x 1.5 mm
- First order resolution > 1500 Å
- Order to order resolution > 10000 Å

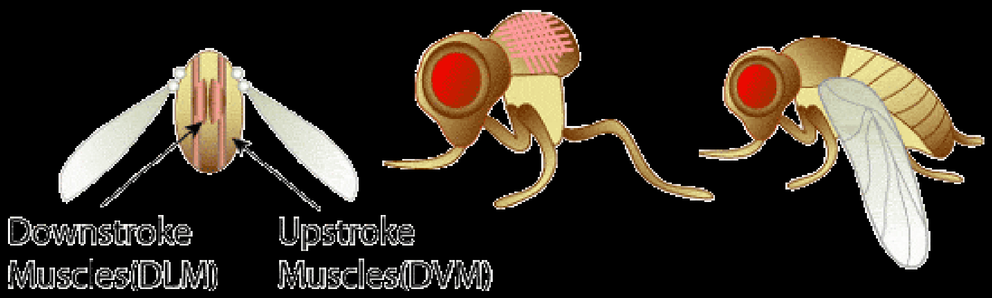


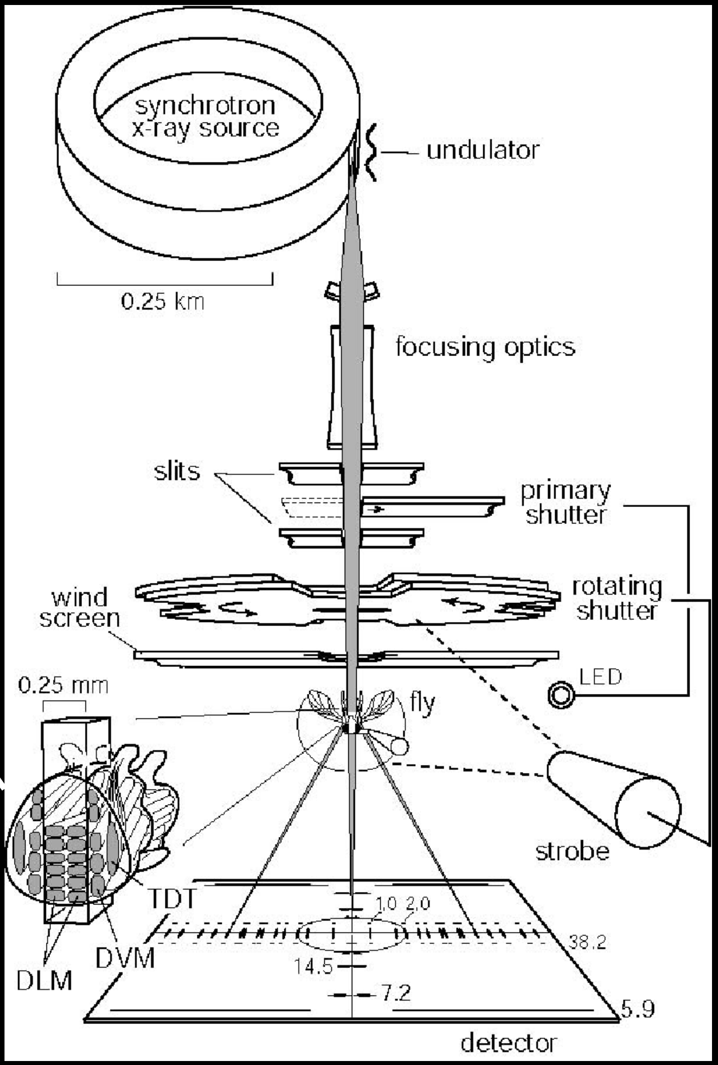
diffraction pattern from oriented sol:

odontoglossum ringspot virus

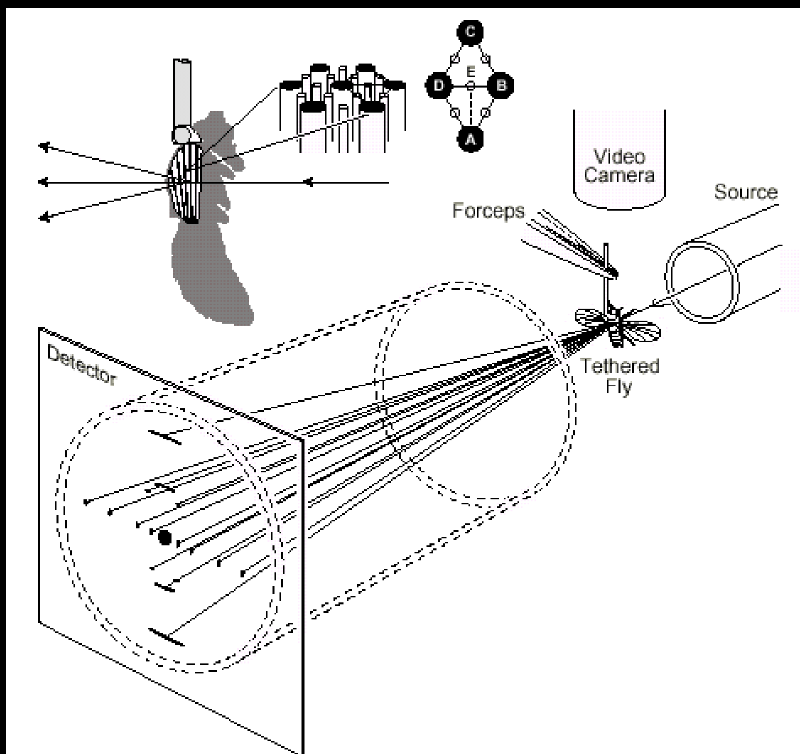
BioCAT 2001

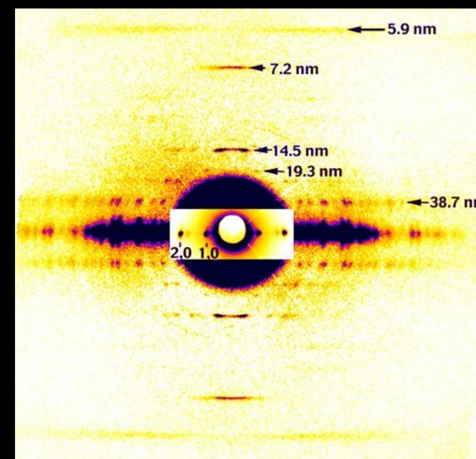
Functional integration
– time resolved

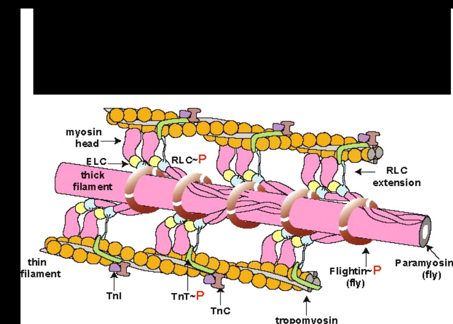




1st X-ray diffraction pattern from living fly



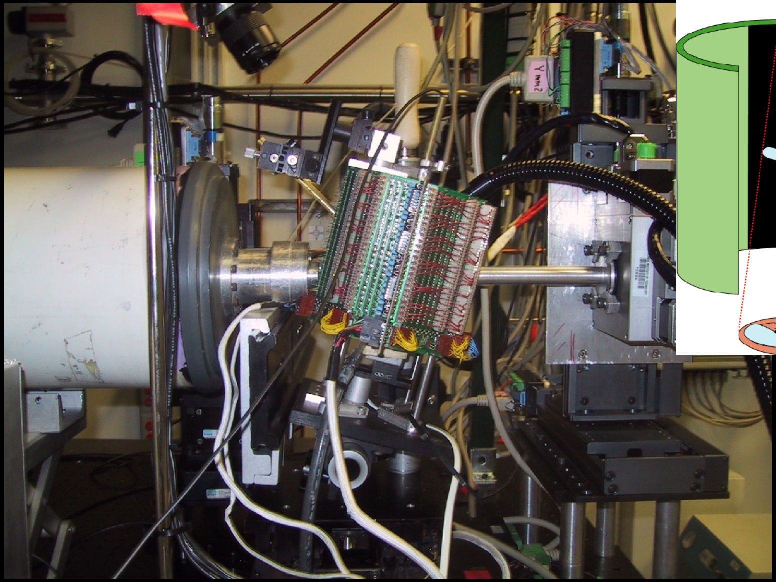


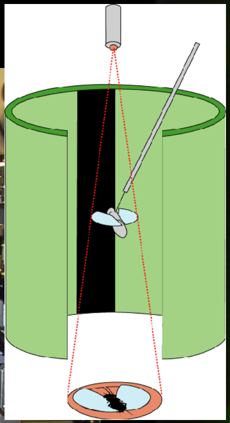


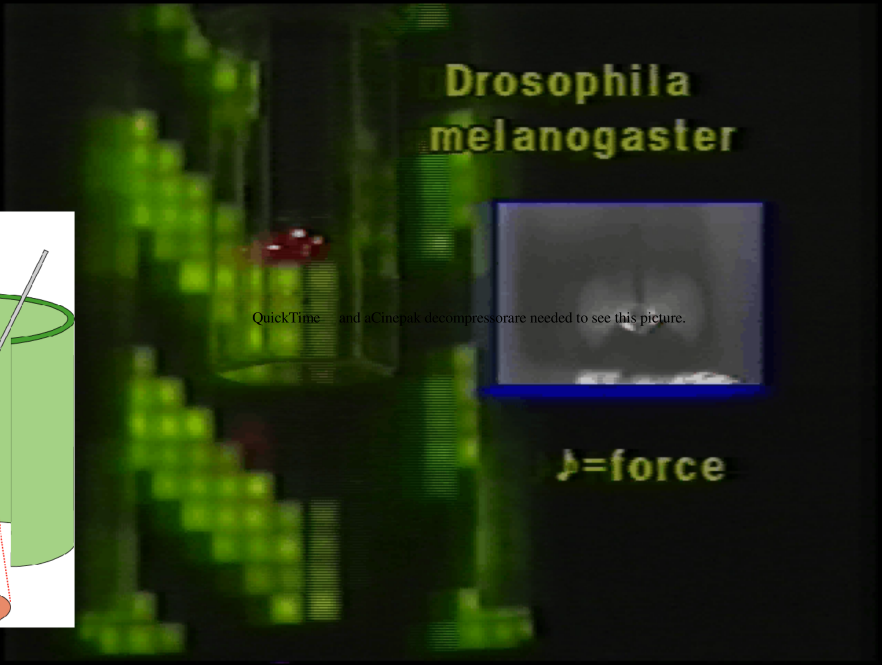
- Drosophila wing beat frequency ~ 200 Hz
- Obtained informative X-ray fiber patterns at the top and the bottom of the wing beat cycle (1 ms time resolution)
- Overturned the widely held ‘constant volume hypothesis’ of lattice spacing in intact muscle

“Fly Flight simulator”

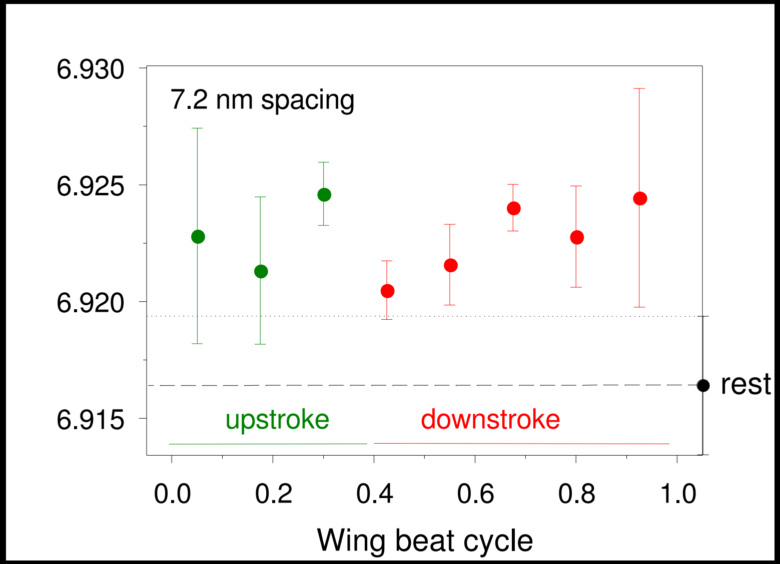
- Visual field simulator
- Wing-beat phase trigger
- Rapid shutter

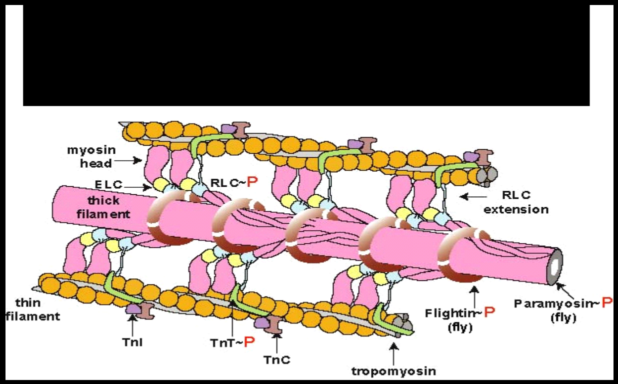






Time-resolved: 7.2 nm reflection





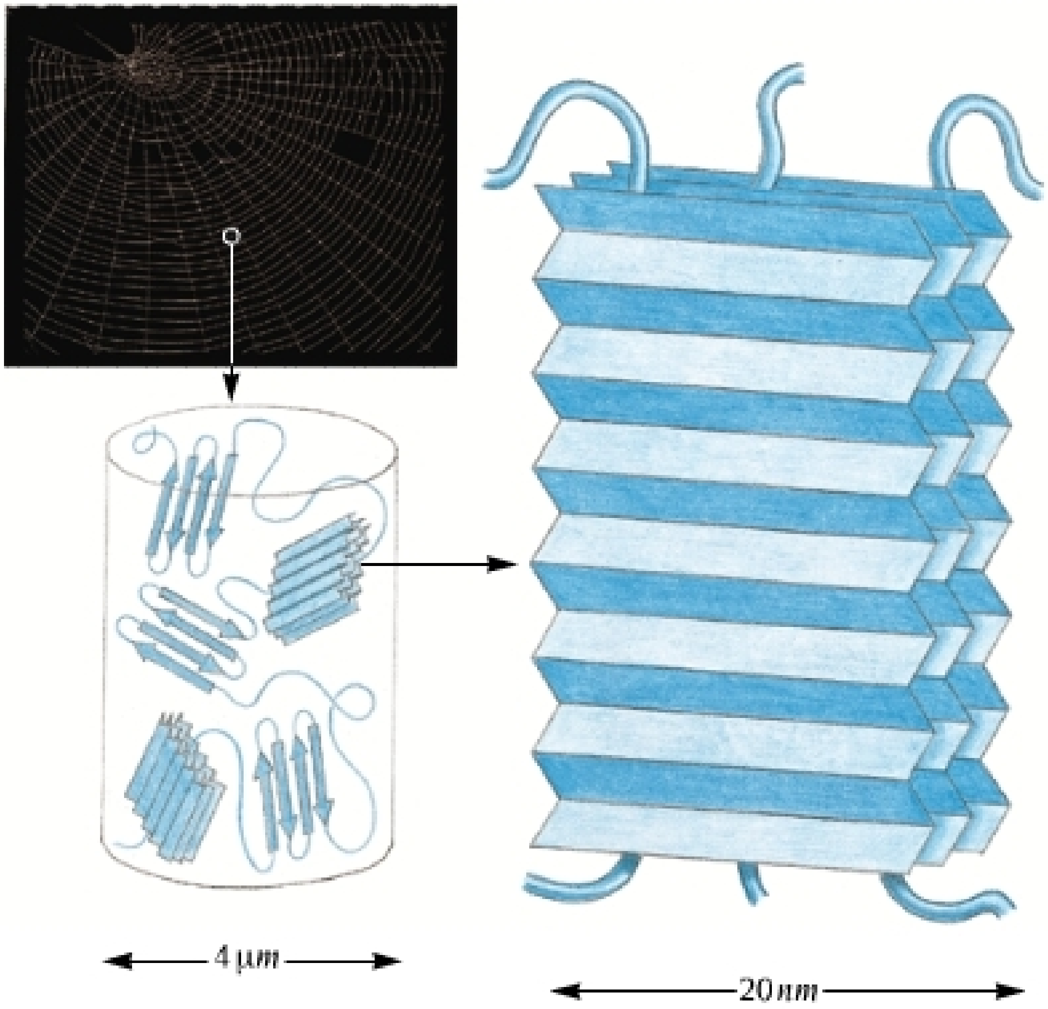
- Small changes in filament length index thick filament stiffness *in vivo*
- *Very* stiff (~0.09% change upon activation vs. 0.3% in frog muscle)
- Cyclical, same period as (and in phase with) wing beat

Spider Silk

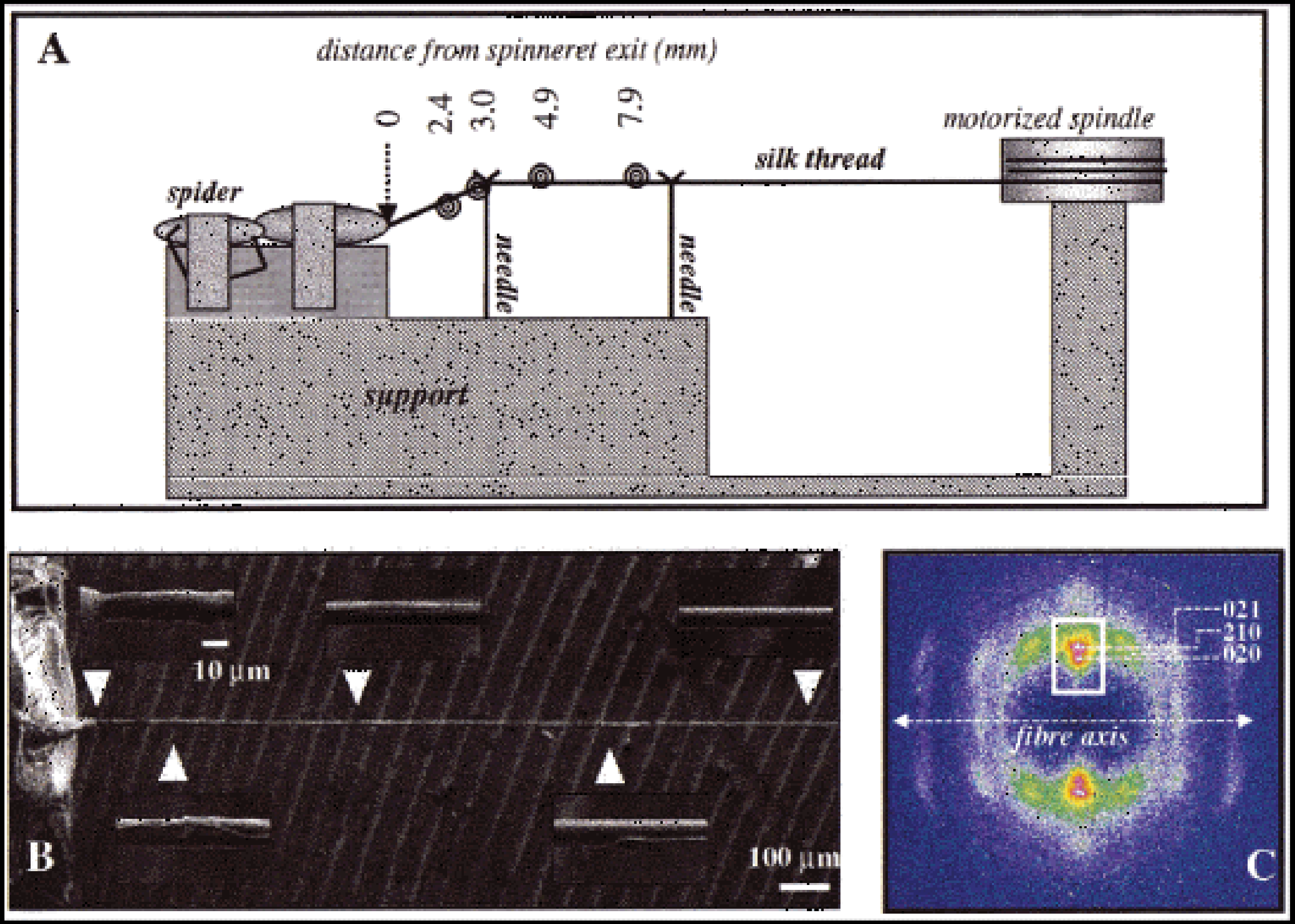
Arguably the strongest polymer known to humanity (tensile strength is in excess of that of steel of the same diameter)

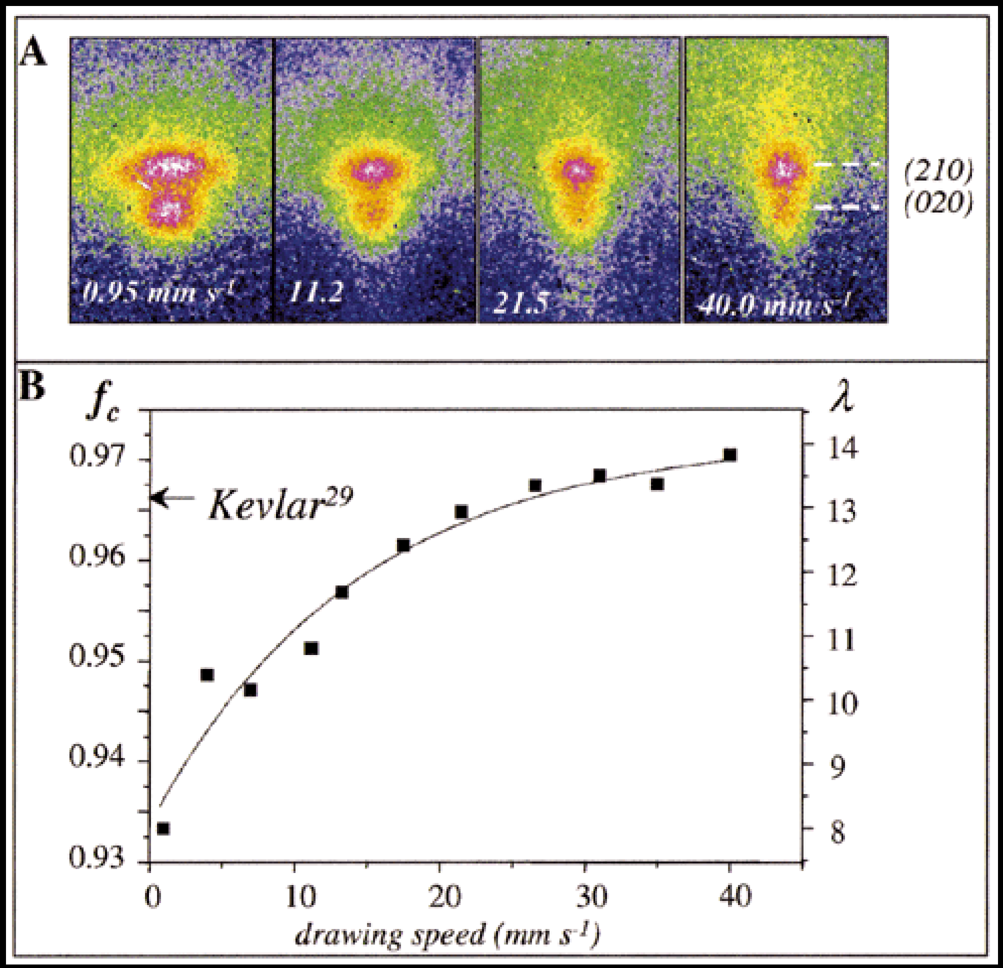
One protein many roles (polymorphic).
e.g. elastic high-tensile strength drag-line or spiders web

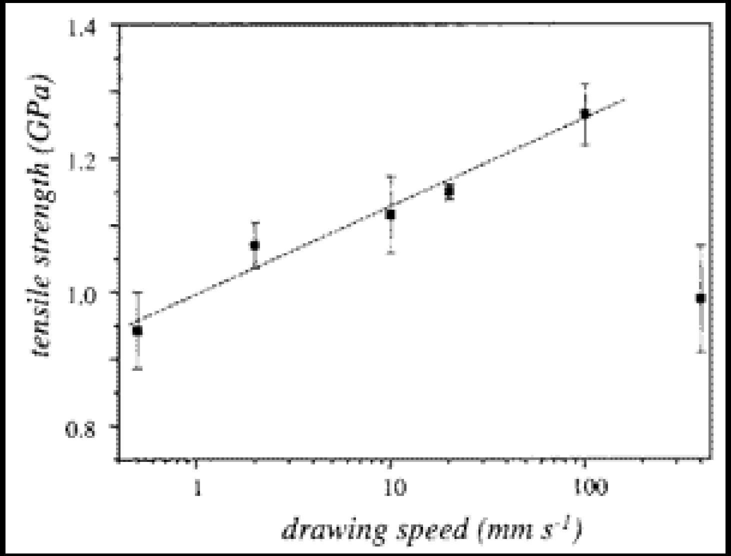
How can they be different if they are formed from the same polypeptide?


4 μm
20 nm

Polymorphoic form of spider silk dependent upon rate of syntheses




A
0.95 mm s-1
11.2
21.5
40.0 mm s-1
(210)
(020)
B
f_c
λ
← Kevlar[29]
drawing speed (mm s-1)


tensile strength (GPa)
drawing speed (mm s-1)

Analysis Software

- Rate limiting step is data analysis
- Long tradition of “rolling on your own”
- CCP 13 project http://www.ccp13.ac.uk
- Comprehensive data extraction suite
- Complementary NSF RCN Stubbs (Vanderbilt) PI will add angular deconvolution, other features to suite

Acknowledgements

BioCAT Staff:

Bob Fischetti, Ke Zhang, Shengke Wang, Sergey Stepanov, Rich Heurich, Darold Perry, Clareen Krolik, and Grant Bunker

Beamline Design:

Gerd Rosenbaum

National Center for Research Resources