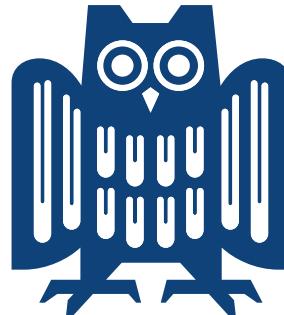


Structural Bioinformatics

Lecture 2

Experimental techniques for determination of
protein 3D structure.
The Protein Data Bank



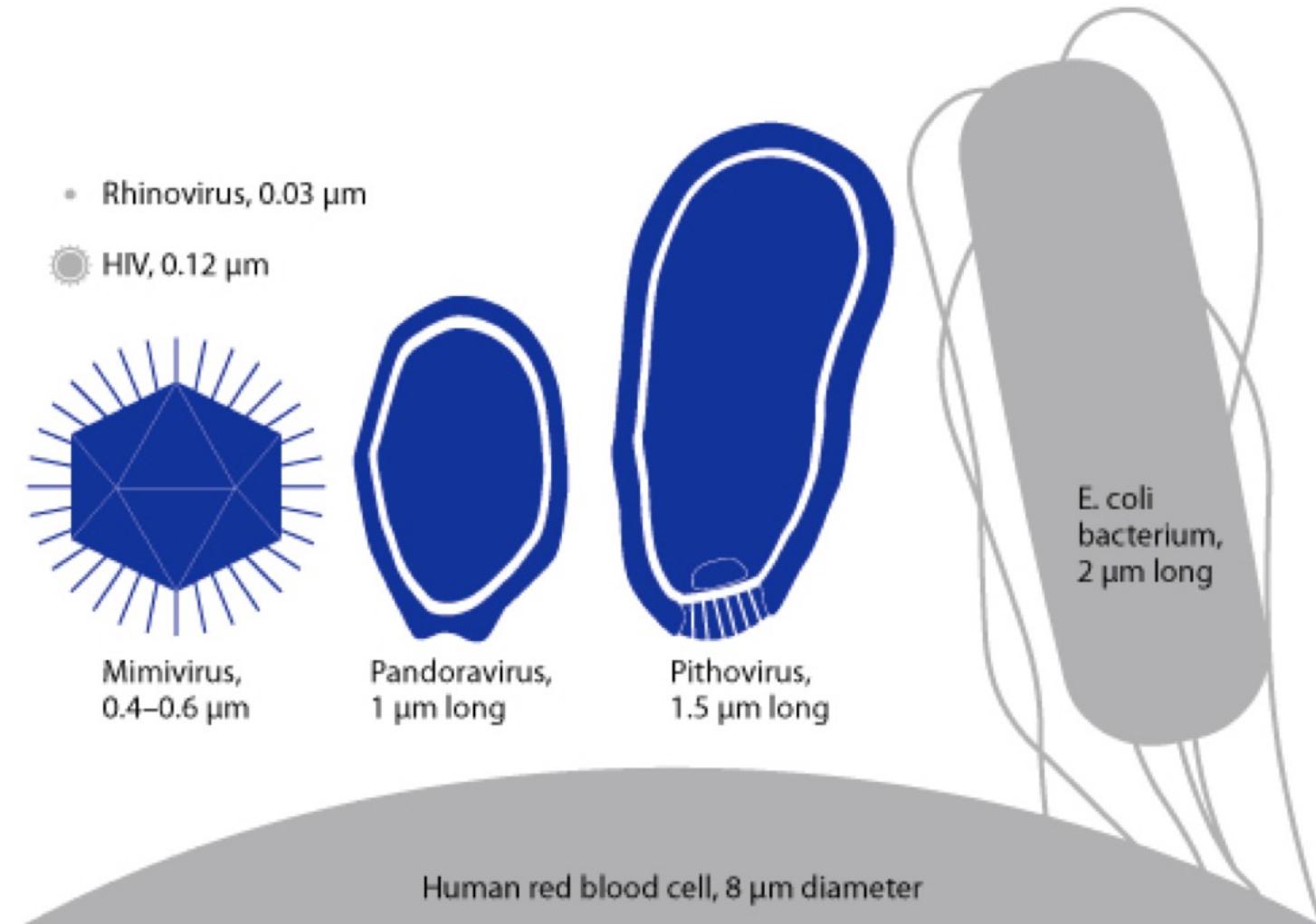
UNIVERSITÄT
DES
SAARLANDES



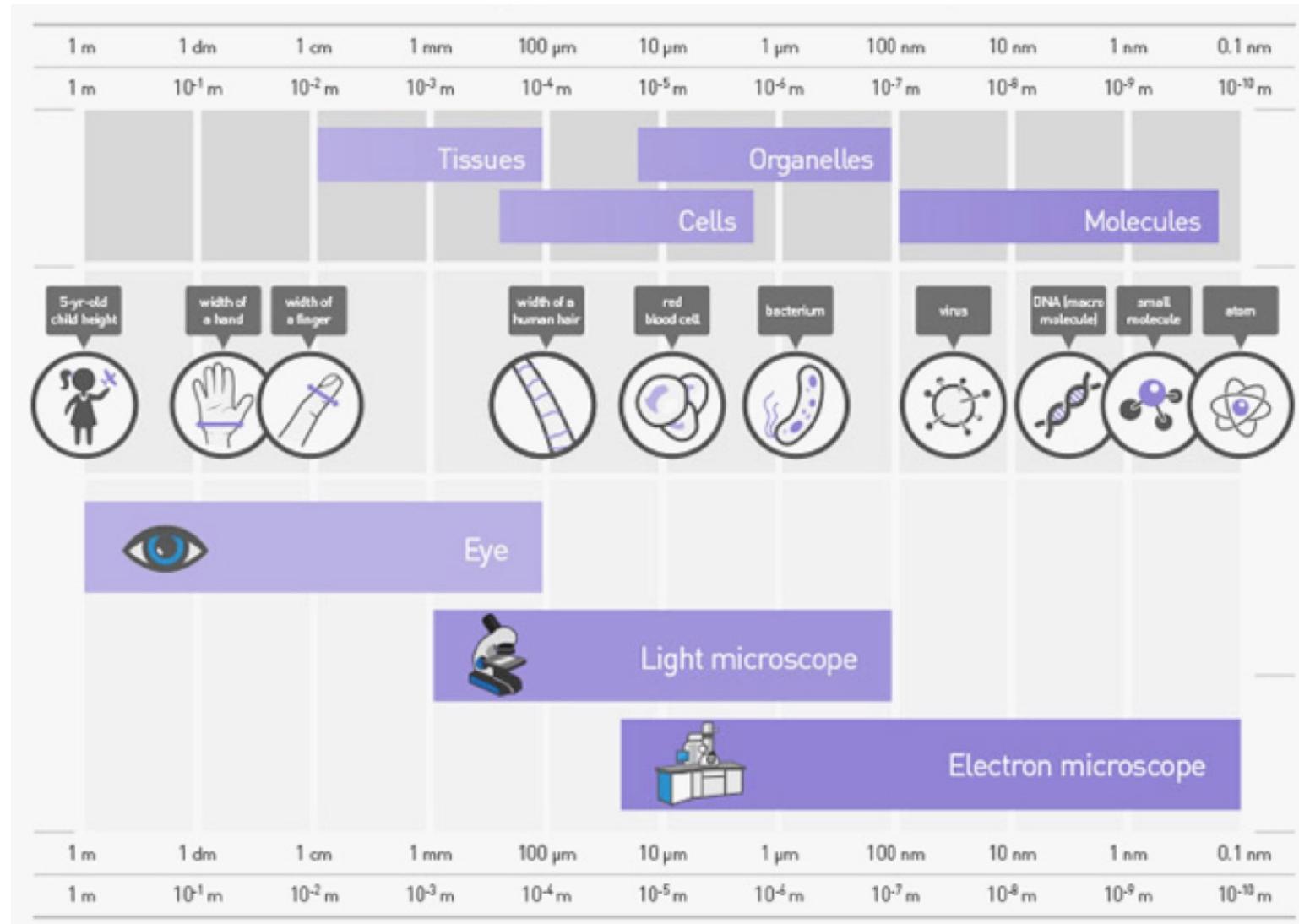
Outline of this lecture

- Experimental techniques for resolving 3D structures of individual biomolecules
- Protein identification via mass spectrometry
- The Protein Data Bank
- Added-value repositories

Size of microscopic objects



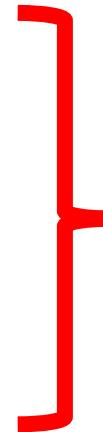
Limits of resolution vs. size of microscopic objects



<https://www.thermofisher.com/sぐ/en/home/life-science/cell-analysis/cell-analysis-learning-center/molecular-probes-school-of-fluorescence/imaging-basics/fundamentals-of-fluorescence-microscopy/epifluorescence-microscope-basics.html>

Major experimental approaches in structural biology

- X-ray crystallography
- NMR spectroscopy
- Electron microscopy
- Mass spectrometry



**Individual molecules
& complexes**

- Hydrogen/deuterium exchange
- Cross-linking
- Yeast two hybrid
- Tandem affinity purification



Not a part of this course

Determination of 3D structure of biomolecules

- X-ray crystallography
- NMR spectroscopy
- Electron microscopy

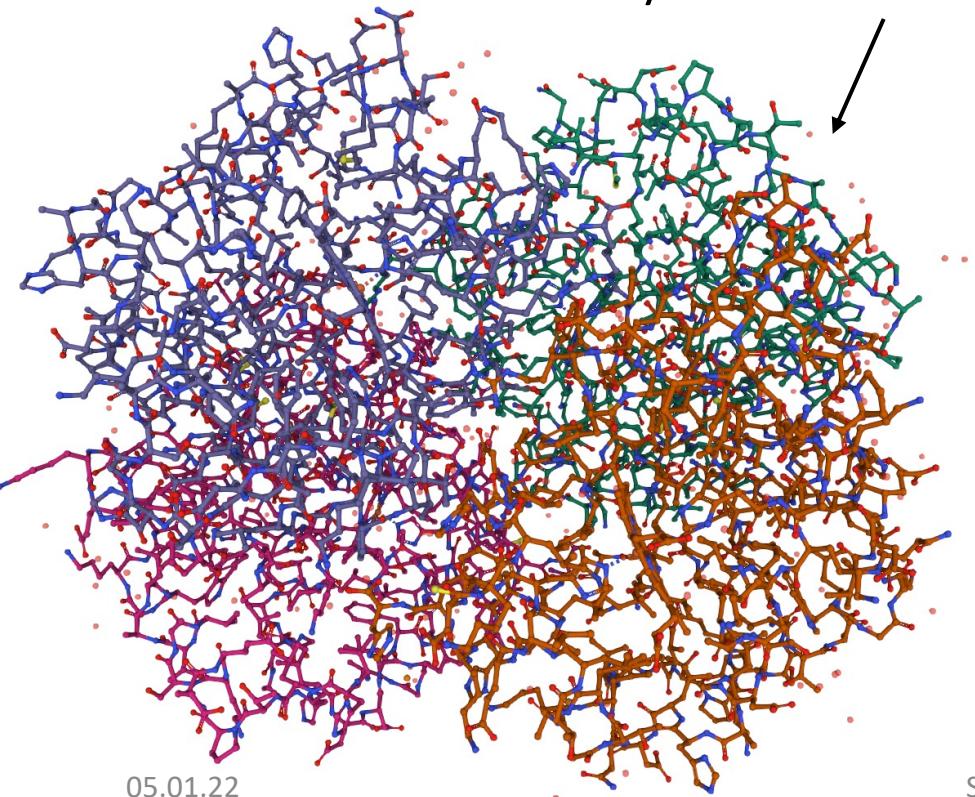
Experimental protein identification

- Mass spectrometry

Looking at protein 3D structures

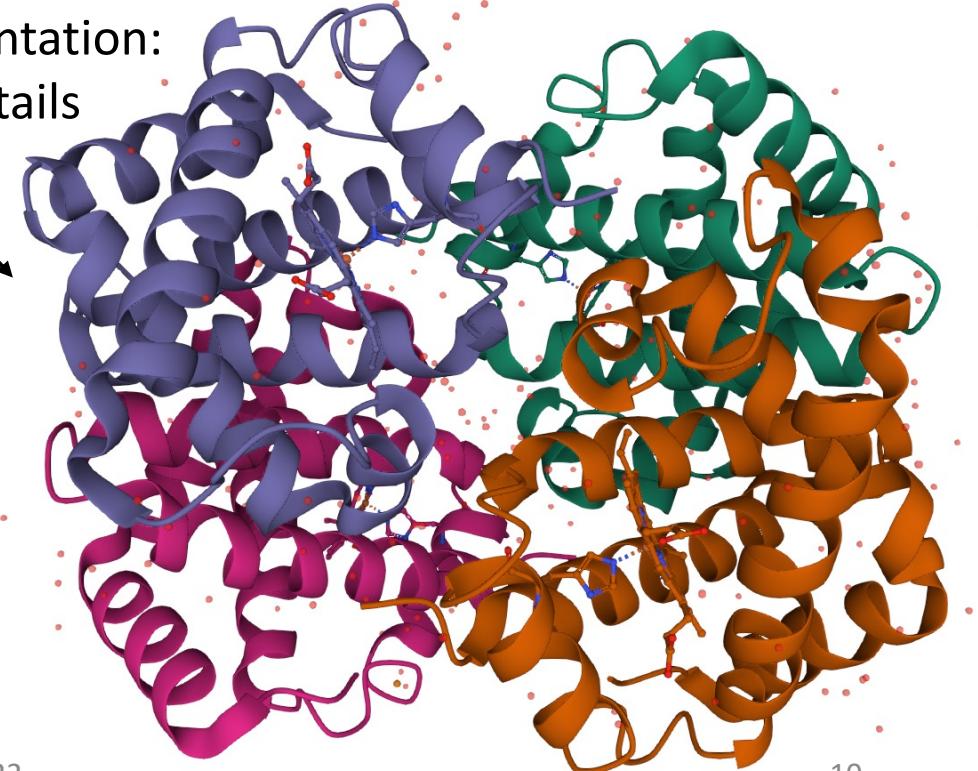
How to represent 1000s atoms?

- Example: haemoglobin, 4779 non-hydrogen atoms

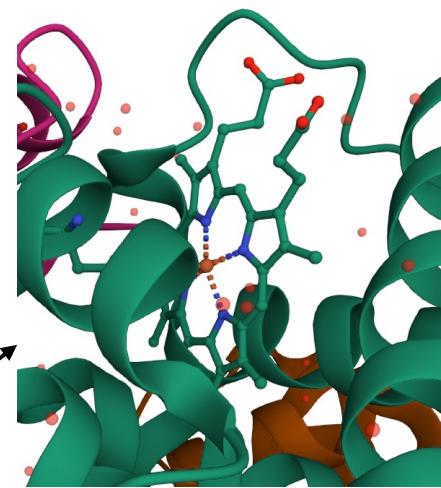


Ball-and-stick representation:
you don't see much

Cartoon representation:
no molecular details

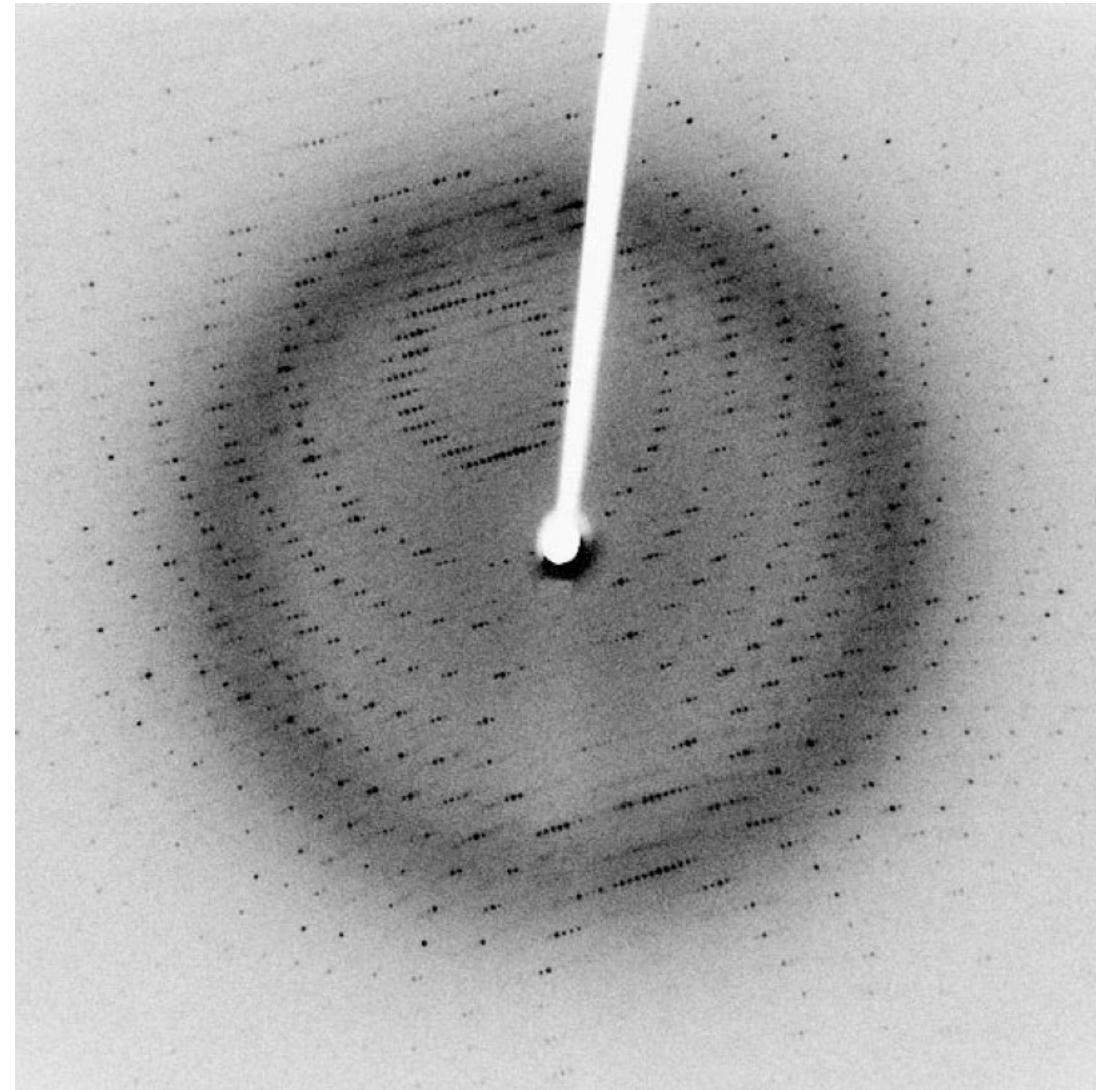
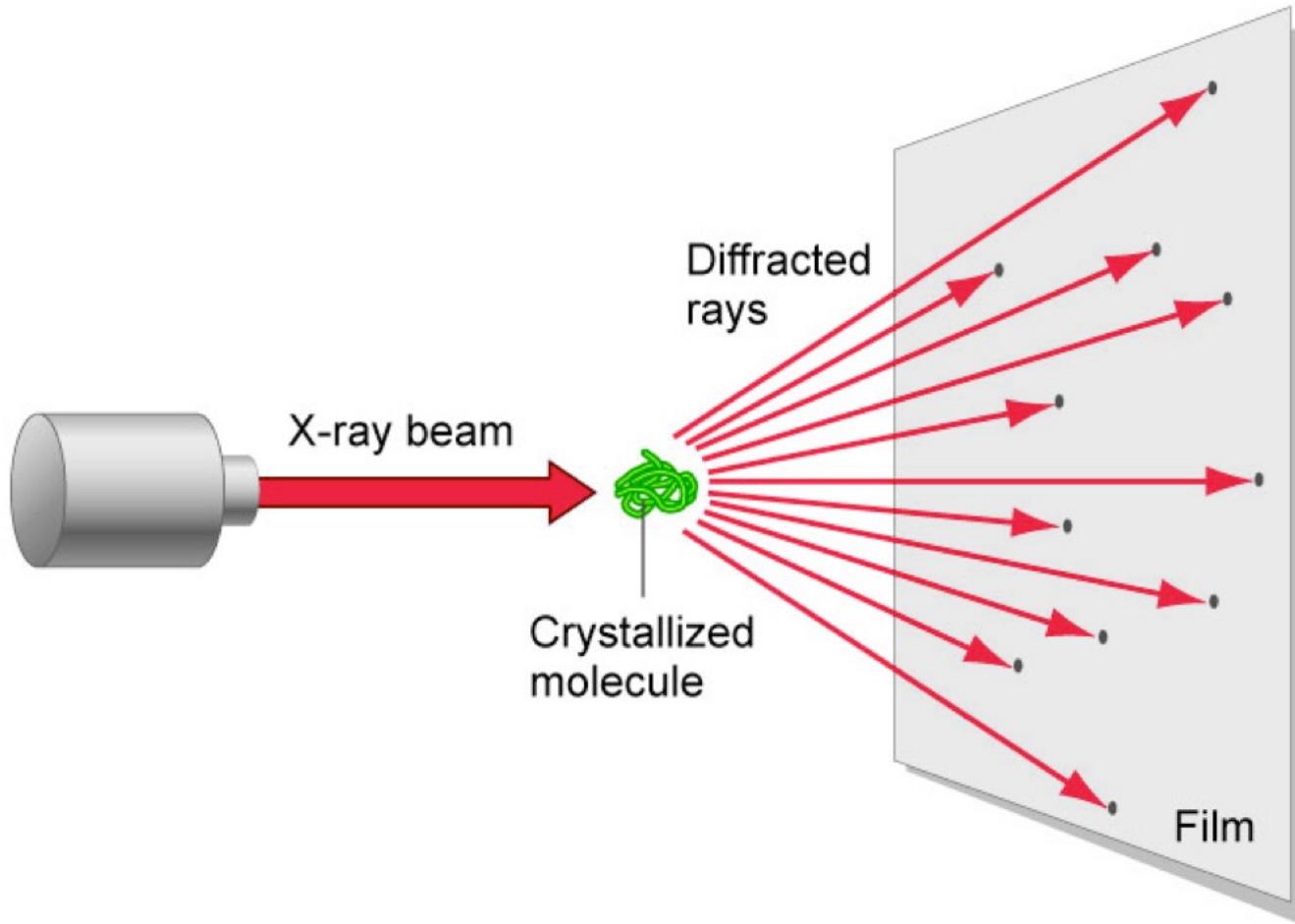


Ball-and-stick
works well for ligands

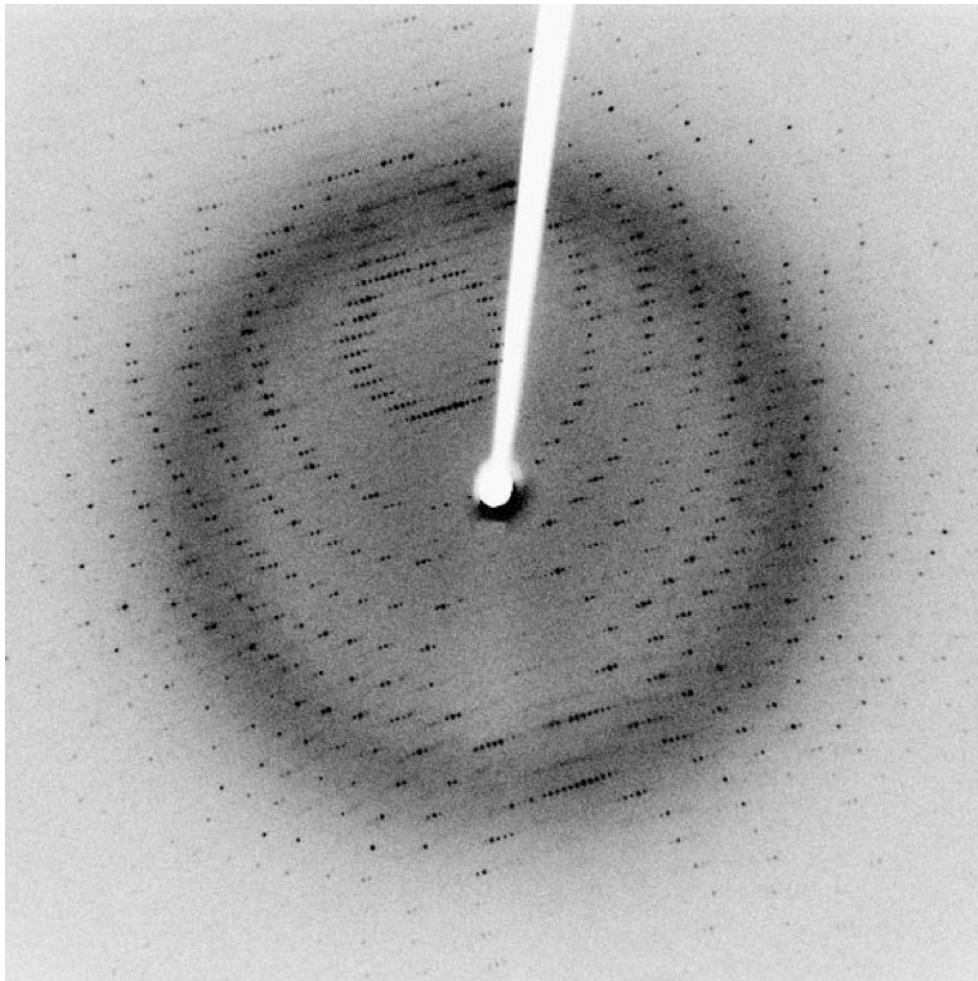


Experimental approaches to resolving 3D structure of biomolecules

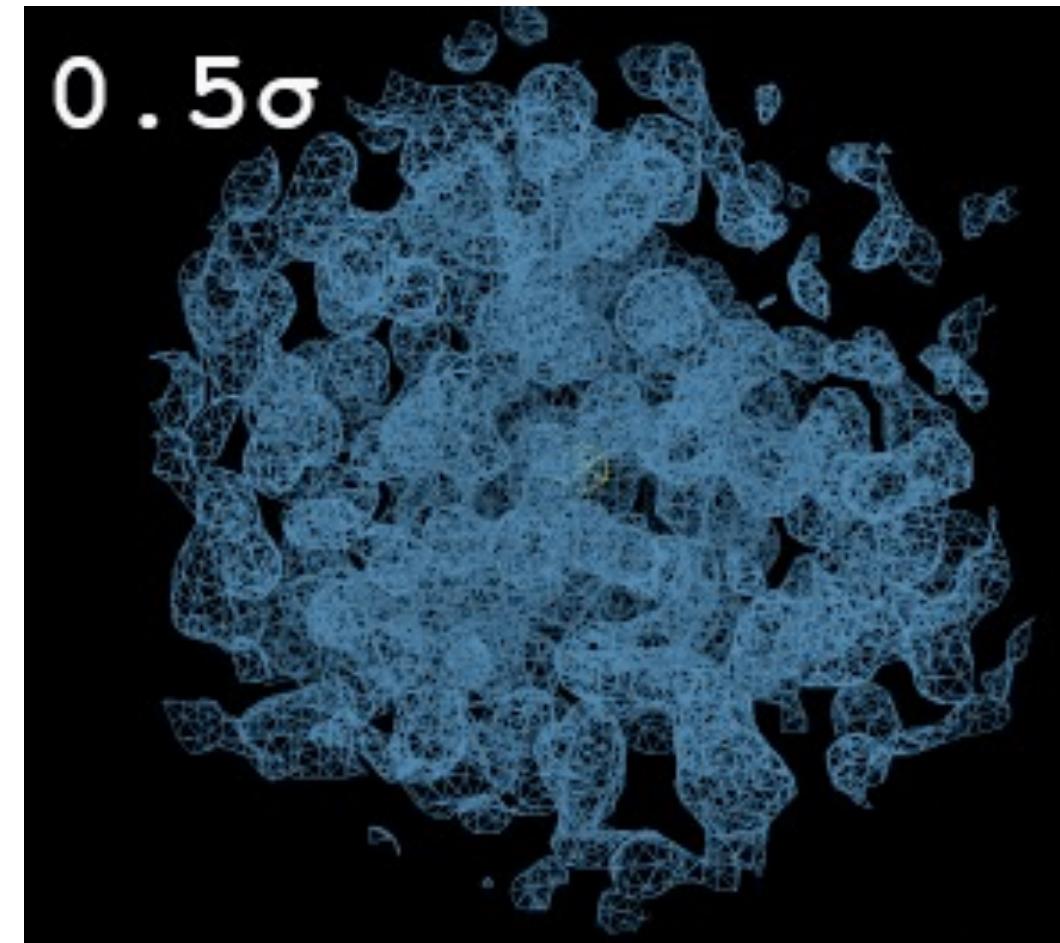
X-ray crystallography



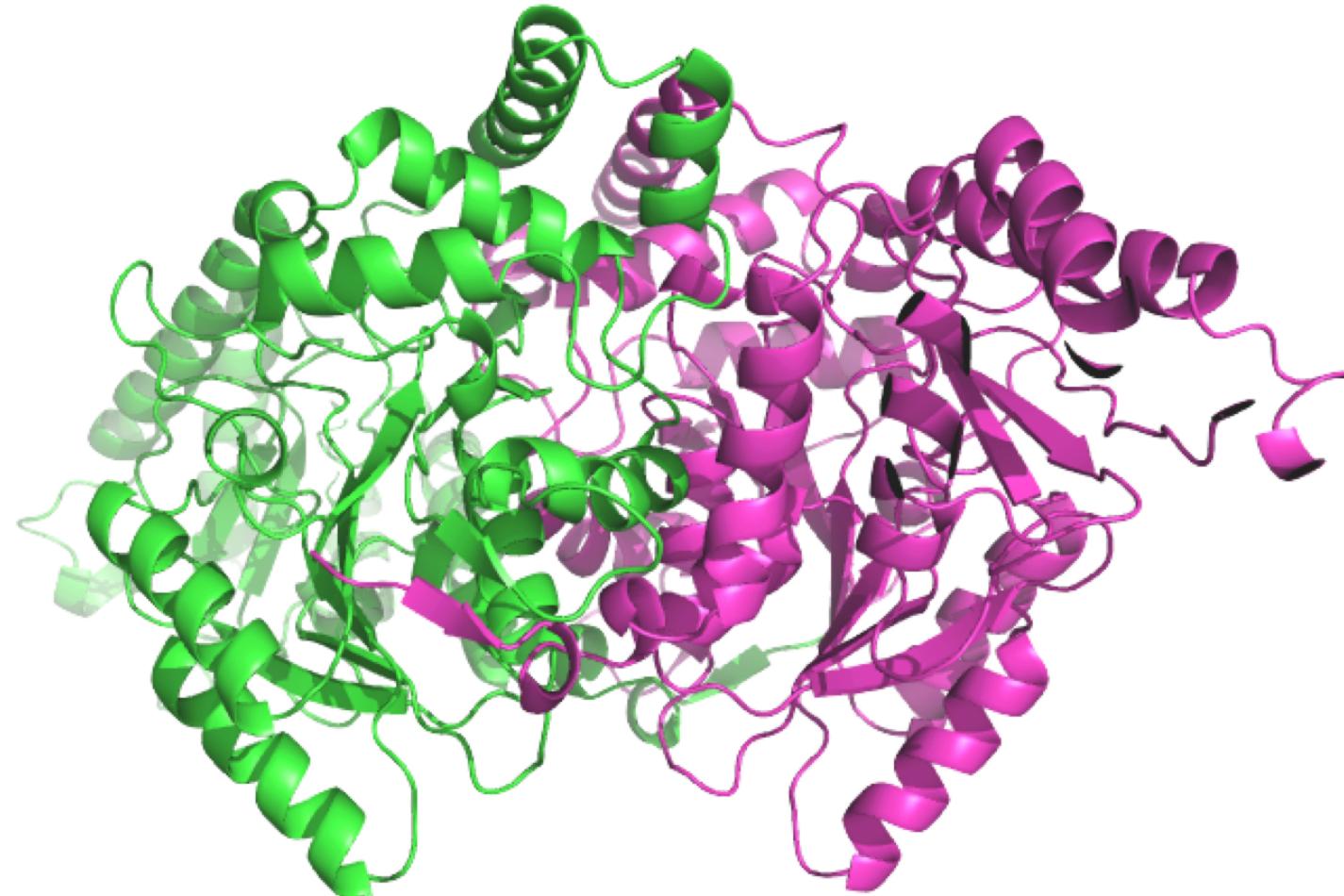
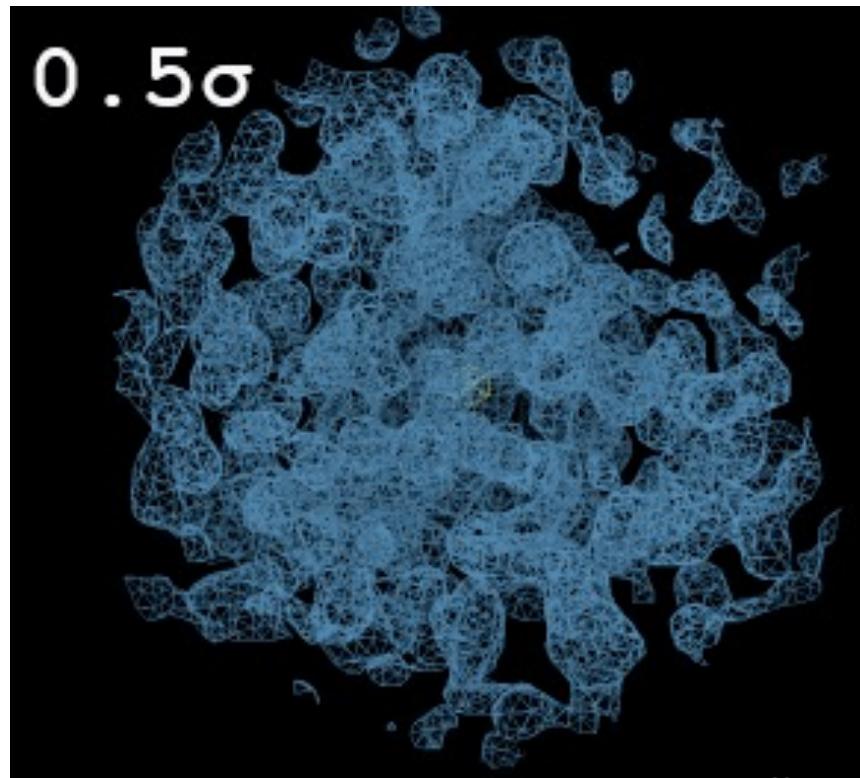
X-ray crystallography



Special software exists that turns diffraction patterns into electron density maps



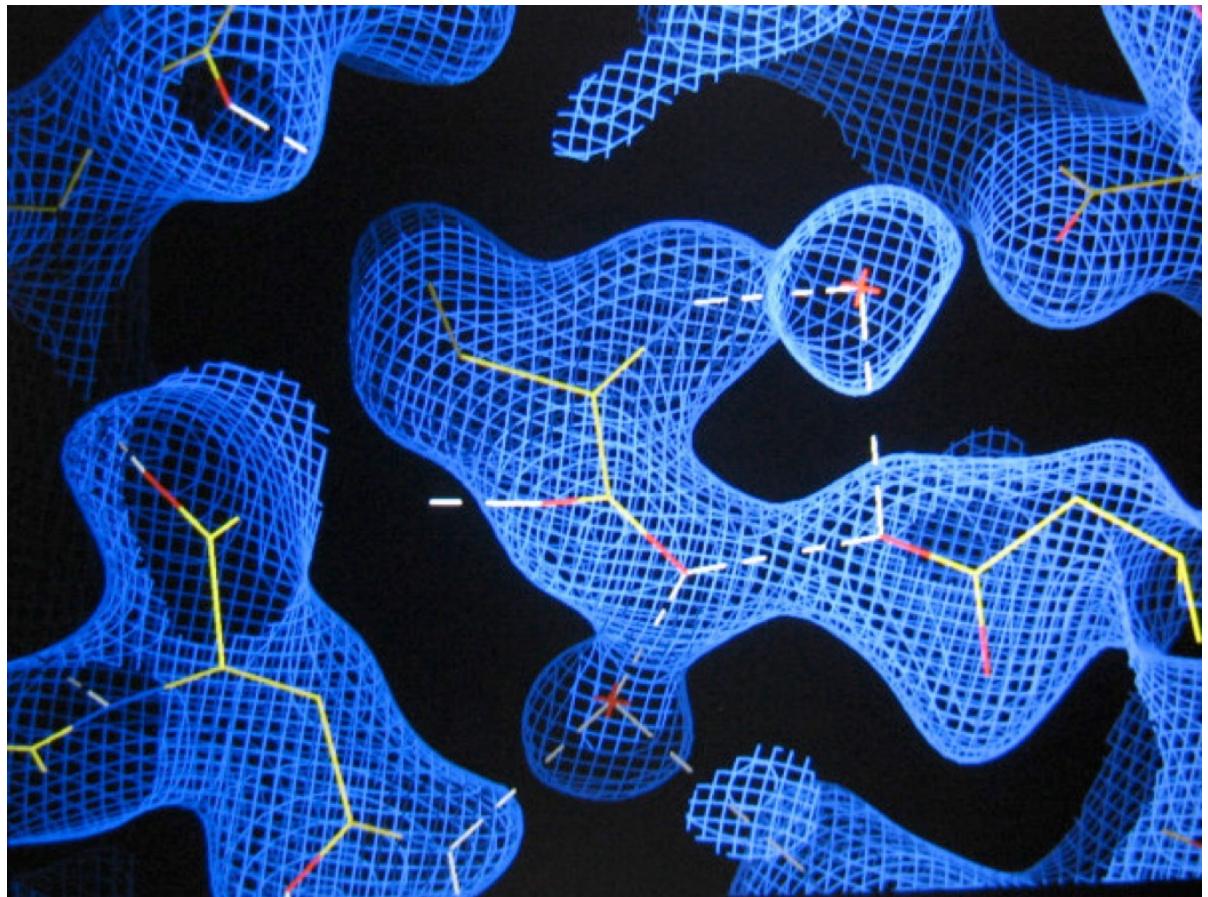
X-ray crystallography



**Here, there exist software
solutions as well**

Structure modelling and refinement

- Multidimensional data
- Take bond lengths/angles from idealized structure
- Cross-validation: use part of the data to build a model, rest of the data to test it => **R-value**



R-factor (value)

- **Structure factor $F(\mathbf{s})$:** $F(\mathbf{s}) = \int_{crystal} \rho(\mathbf{r}) e^{-2\pi i \mathbf{r} \cdot \mathbf{s}} d\mathbf{r}$,
 $\rho(\mathbf{r})$ is the electron density
 - describes the distribution of atoms in the crystal
- Electron density map can be reconstructed from given atom coordinates, and structure factor can be theoretically calculated
- **R-value:** $R = \frac{\sum ||F_{obs} - F_{calc}||}{\sum |F_{obs}|}$

R-free factor (value)

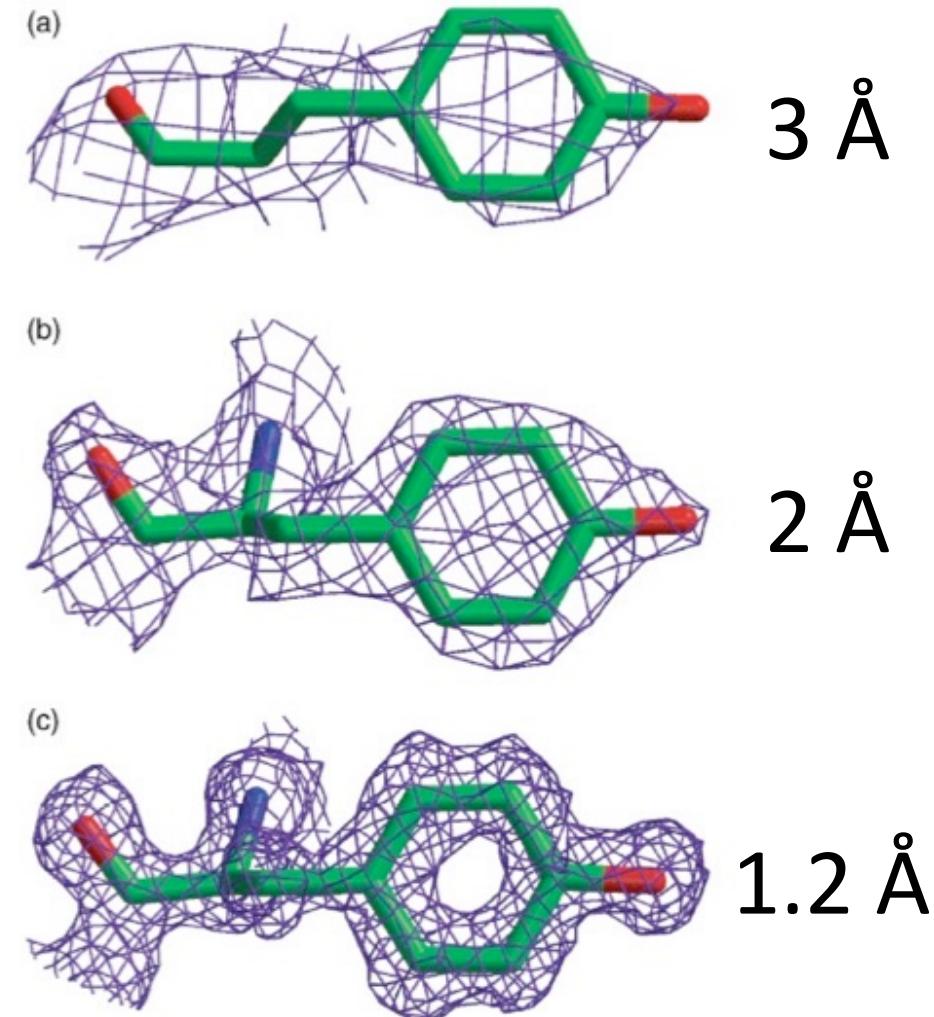
- **R-value:** $R = \frac{\sum ||F_{obs} - F_{calc}||}{\sum |F_{obs}|}$
- **R-free** is the same thing, but only 90% of the data is used for refinement

R / R-free: your first criterion of structure quality

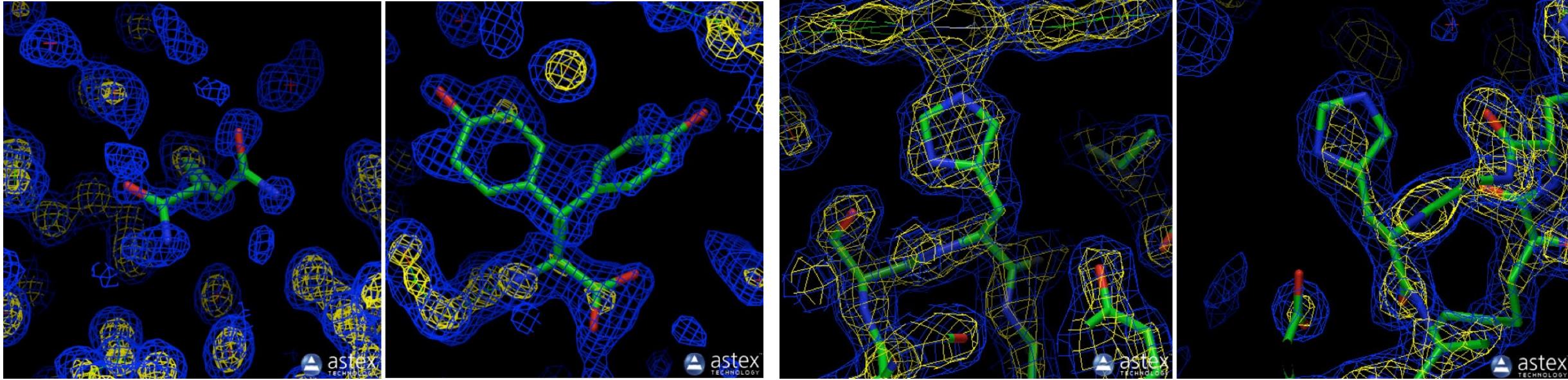
- **R-value:** $R = \frac{\sum ||F_{obs} - F_{calc}||}{\sum |F_{obs}|}$
- => R-value: the lower the better
 - ~0.40-0.60: totally random; ~0.20: good model
- R-value **positively correlates** with resolution
 - R-free should not exceed (resolution/10) by more than 0.05

Second thing to watch: structure resolution

- Measured in angstroms (ångströms), denoted Å
 - After Anders Jonas Ångström, Swedish physist
 - 10^{-10} m
- The lower the better
- For 3D structures of biomolecules, resolution $< 2 \text{ \AA}$ is good, $< 2.5 \text{ \AA}$ is OK
 - As low as $< 1 \text{ \AA}$ is possible



Occupancy & B-factor: measures of structural uncertainty



Occupancy: alternative conformations, e.g. of side chains

05.01.22

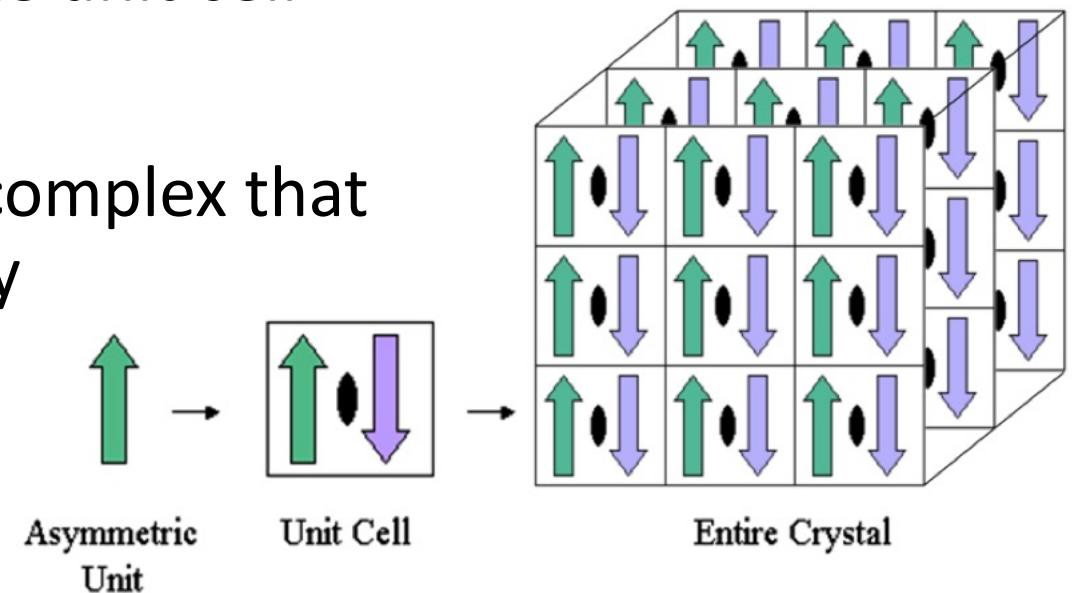
B-factor (temperature factor): reflects flexibility of atoms in their positions

Structural Bioinformatics WS 2021/22

20

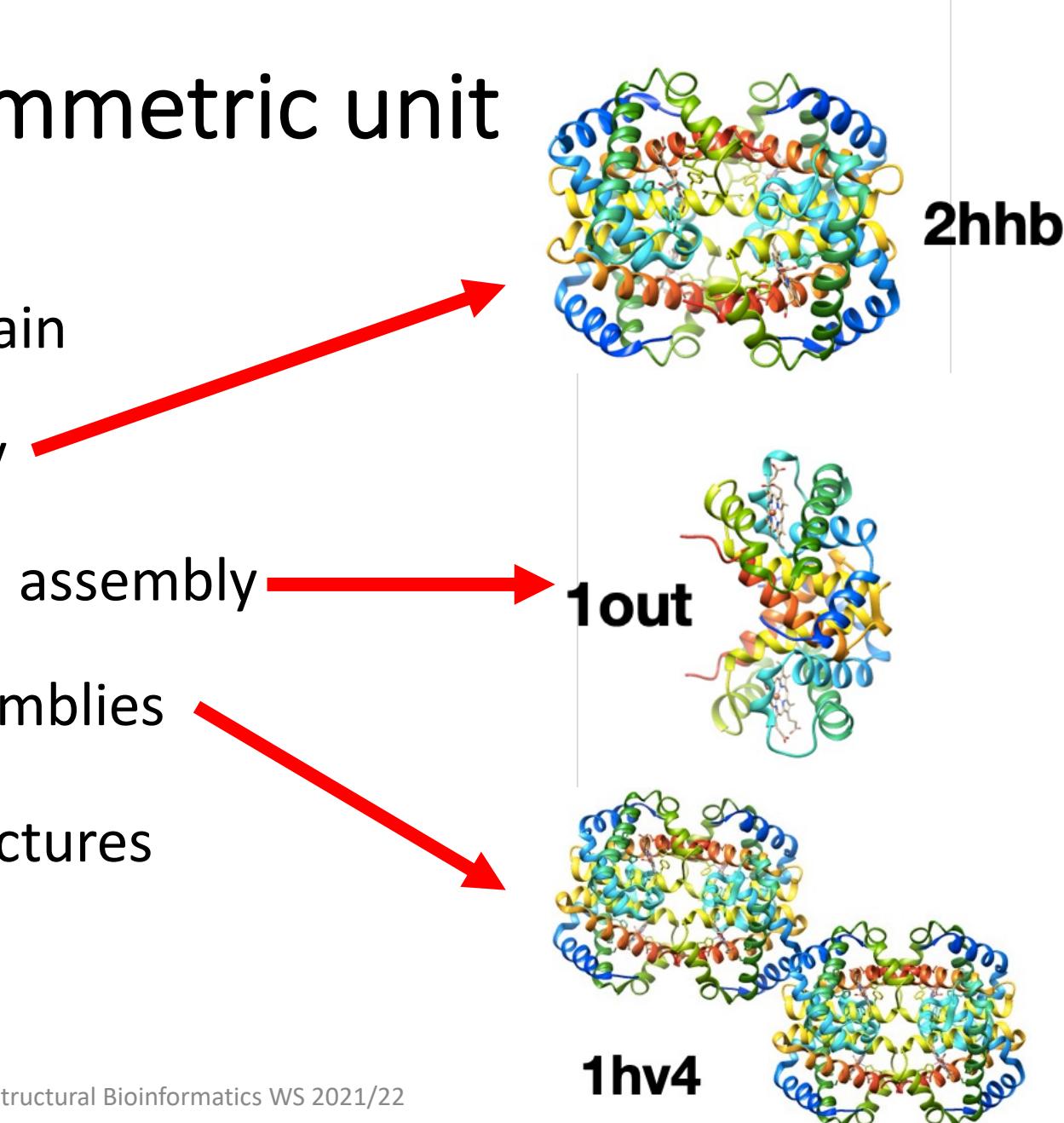
Biological and asymmetric unit

- **Asymmetric unit:** the smallest portion of a crystal structure to which symmetry operations can be applied in order to generate the complete unit cell
- **Unit cell:** the crystal repeating unit
- **Biological assembly:** protein or protein complex that represents a biologically functional entity
 - Experimentally proven or believed
 - Author provided or calculated (PISA)



Biological and asymmetric unit

- Asymmetric unit may contain
 - one biological assembly
 - a portion of a biological assembly
 - multiple biological assemblies
- **Example:** hemoglobin structures

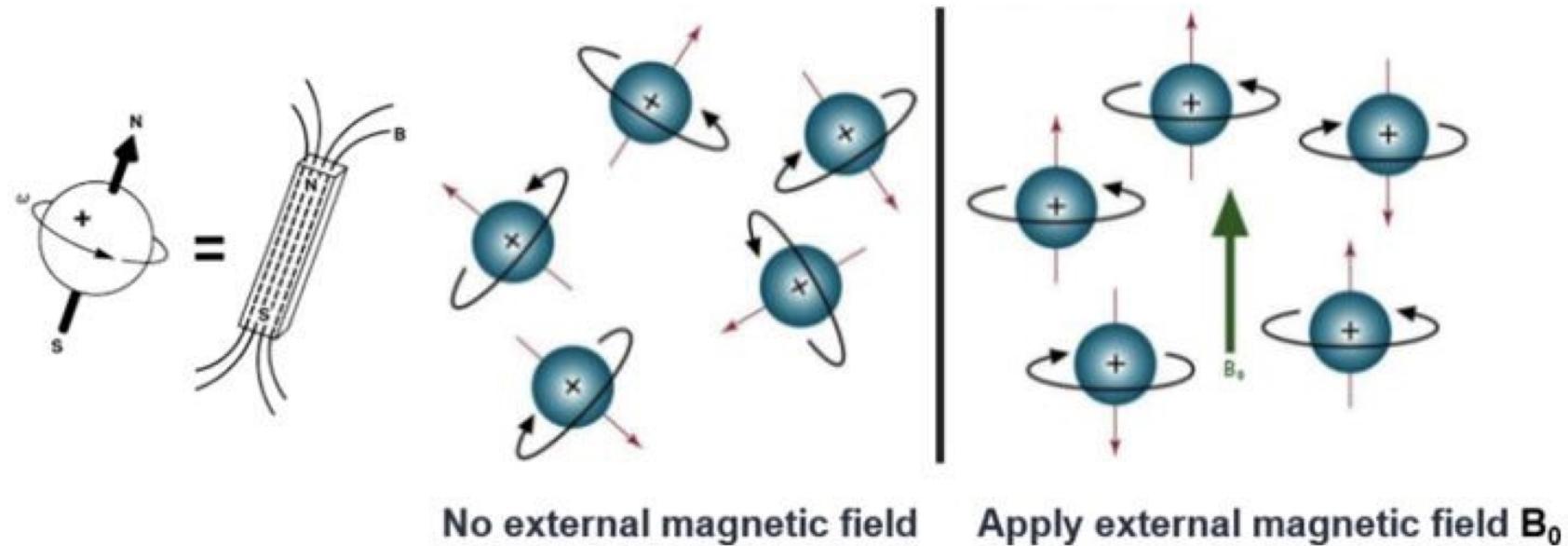


X-ray structure: important properties

- Resolution
- R-value and R-free value
- B-factor and alternative conformations

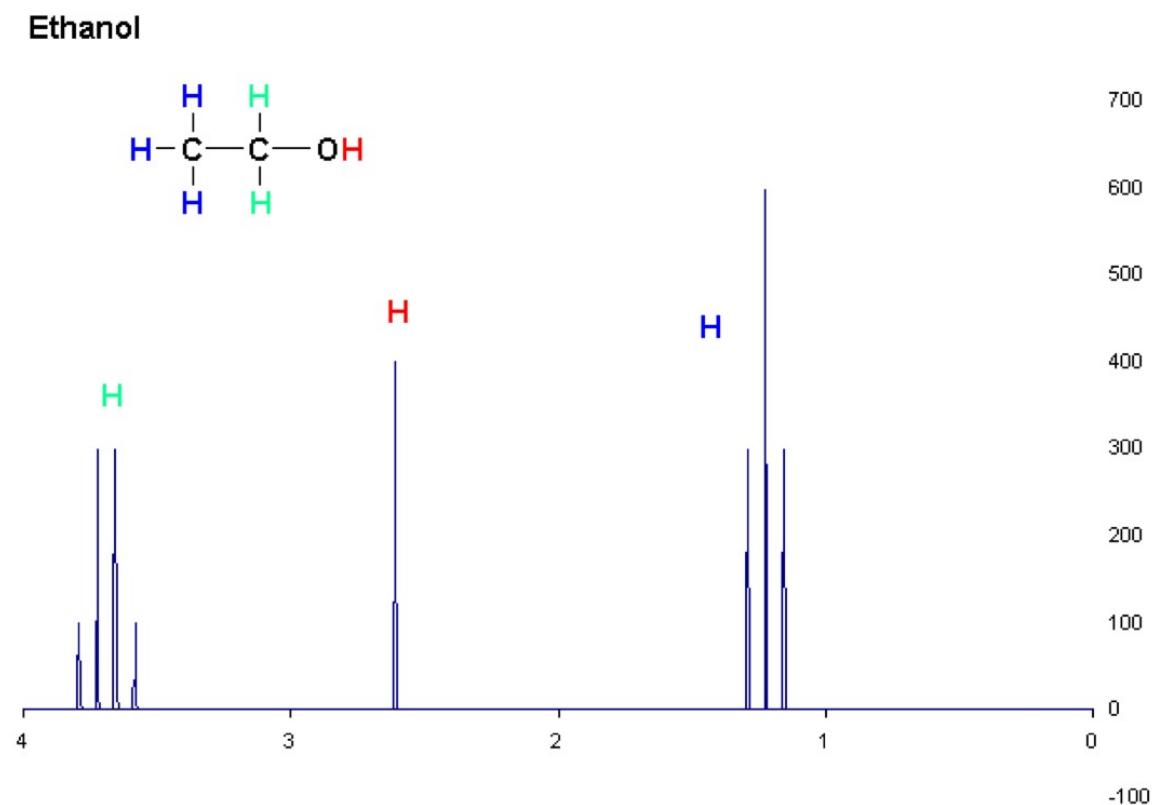
NMR (nuclear magnetic resonance) spectroscopy

- **Spin:** an intrinsic property of elementary particles; can be associated with nuclei with an odd number of protons and/or neutrons
- Spins respond to external magnetic field
- Resonance properties of different nuclei or same nuclei in different environment are different



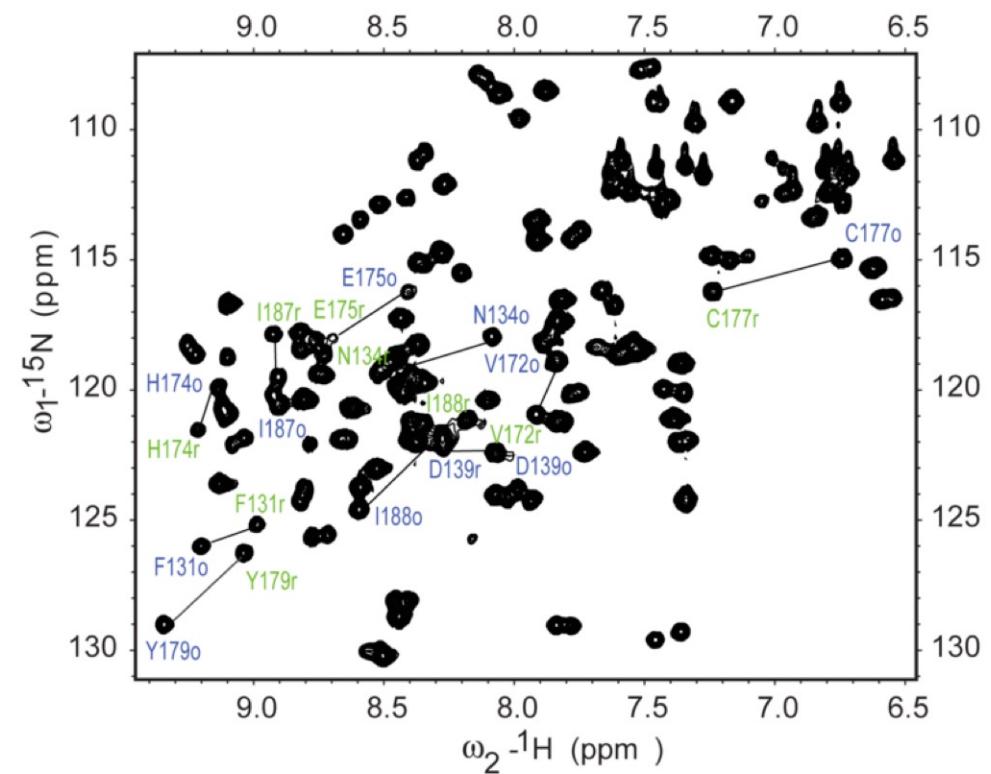
NMR spectroscopy

- **Idea:** record changes in atom spins in response to radio frequency (RF, 3 kHz ... 300 GHz) pulses: at some point the nuclei absorbs RF energy
- Nuclei in different environment absorb differently — **chemical shifts** => distance restrains
- Coupling between neighbouring active nuclei produces **multiplets**



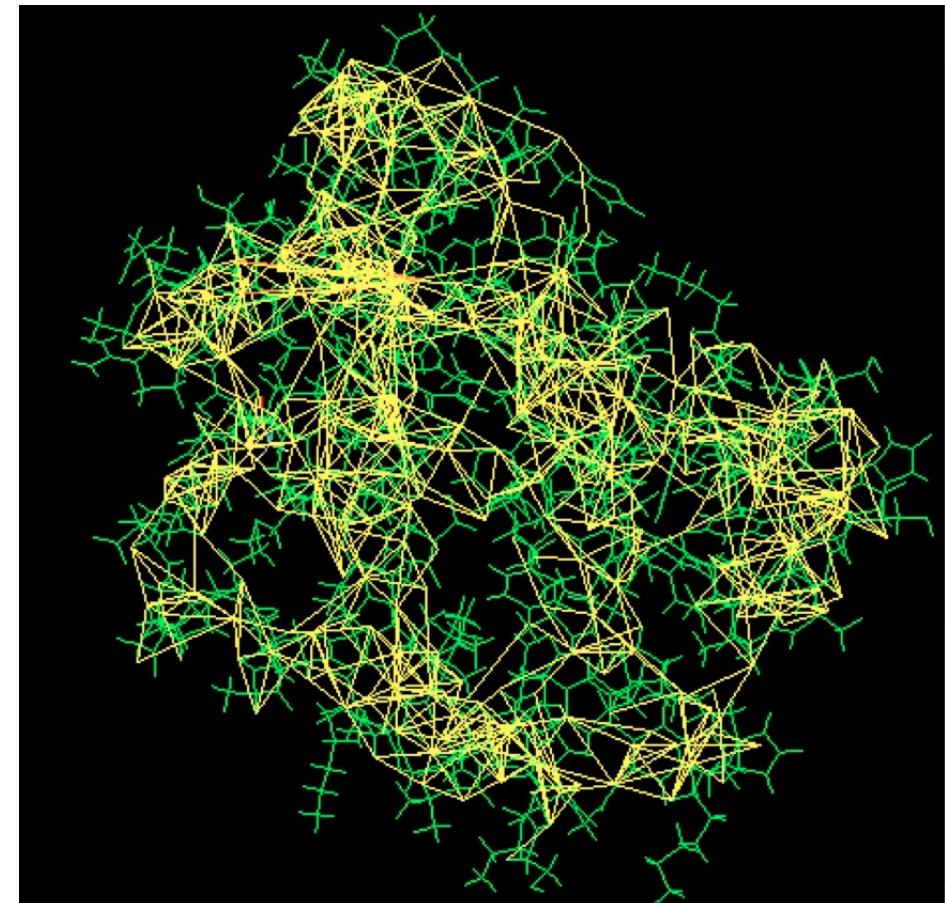
2D-NMR: more useful for protein structures

- Different frequencies allow to detect specific nuclei
- Use stable isotopes ^{15}N , ^{13}C , ^2H to increase signal-to-noise ratio
 - you detect only the labelled molecules
- Chemical shifts
- => 2D spectrum, in which intensity corresponds to a pair of chemical shifts for two proximal atoms of the corresponding types



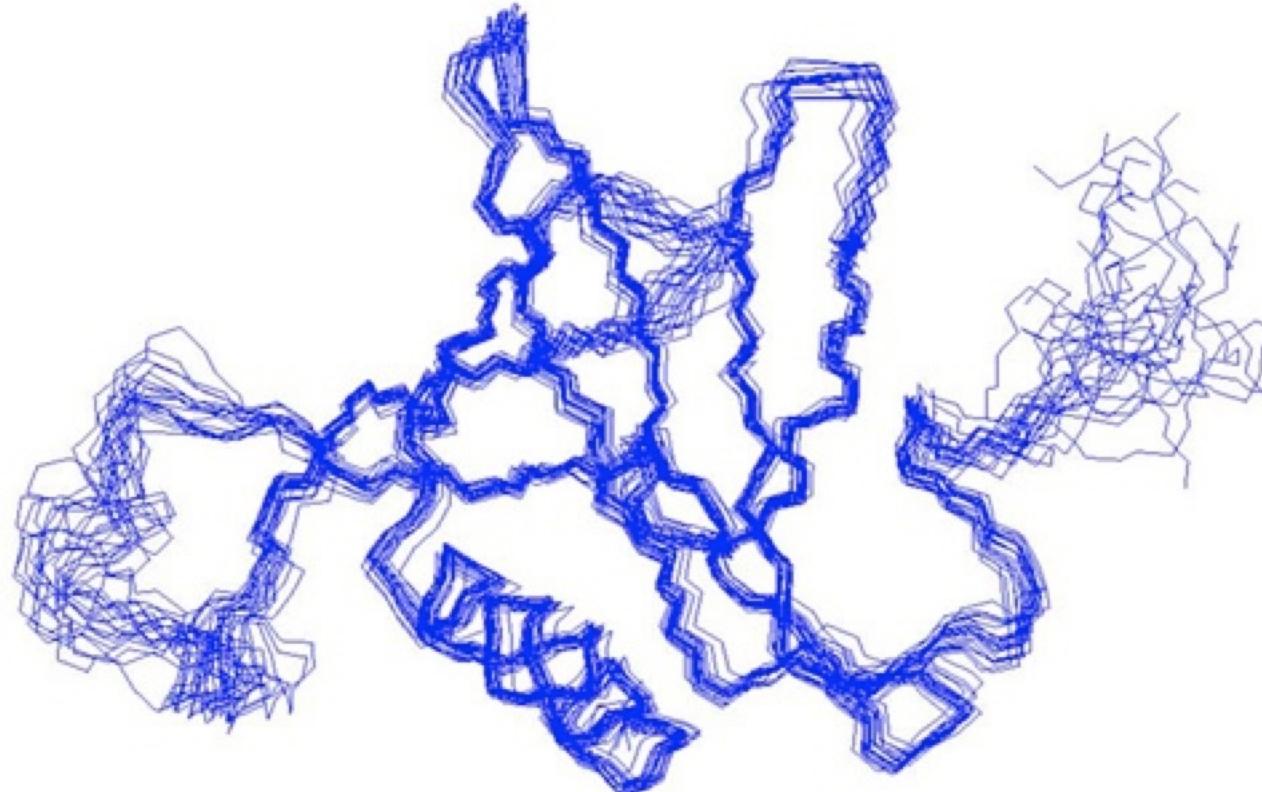
Resolving NMR spectra into 3D structures

- Each nucleus in a unique environment => unique chemical shifts
- Each cross-peak in 2D spectrum means proximity of two nuclei in question:
 $1.8 \dots 6 \text{ \AA}$, $\sim \text{intensity}^{-6}$
- Angle restraints (coupling of identical nuclei)
- Orientation restraints (external field)



Resolving NMR spectra into 3D structures

- Not all restraints can be satisfied and/or multiple solutions for the same set of restraints exist => multiple models



NMR spectroscopy vs. X-ray crystallography

- Physics completely different!
- Size limitations: NMR works best on structures < 25 kDa (~250 amino acids); X-ray works on anything you manage to crystallize
- NMR represents protein in solution
- NMR captures dynamics, X-ray is just a snapshot
- NMR allows other types of experiments (e.g. H/D exchange)

Microscopy: wavelength implies resolution

- **Abbe diffraction limit** (Ernst Abbe, 1873): spot separated by a distance smaller than wavelength are indistinguishable:

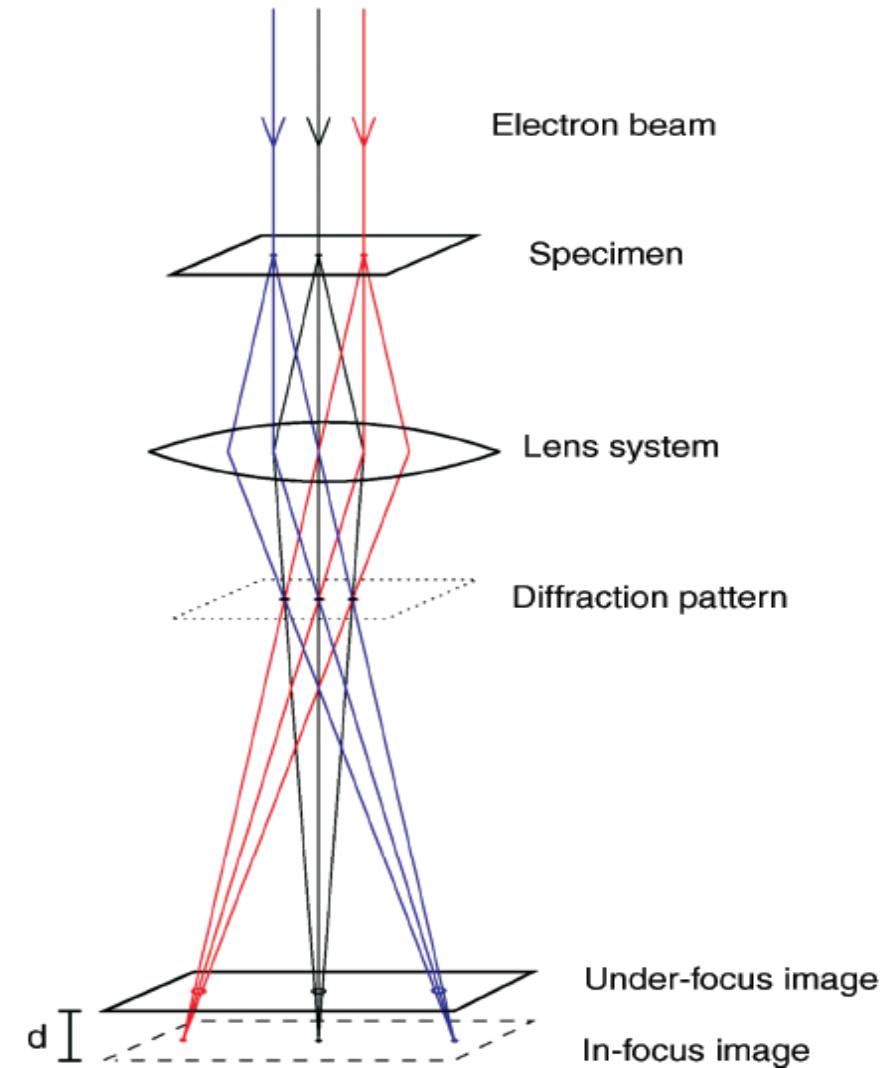
$$d = \frac{\lambda}{2n \sin \theta}$$

λ : wavelength; n : refractive index of the medium, ϑ : half-angle of the light beam
($n \sin \theta$: numerical aperture, NA, can reach 1.4-1.6 in modern devices)

- => for visible light $\lambda \sim 500$ nm => $d \sim 250$ nm (2500 Å)
- If high-energy electrons are used as the source of illumination,
 $\lambda \sim 100,000$ times smaller => $d \sim 50$ pm (0.5 Å)

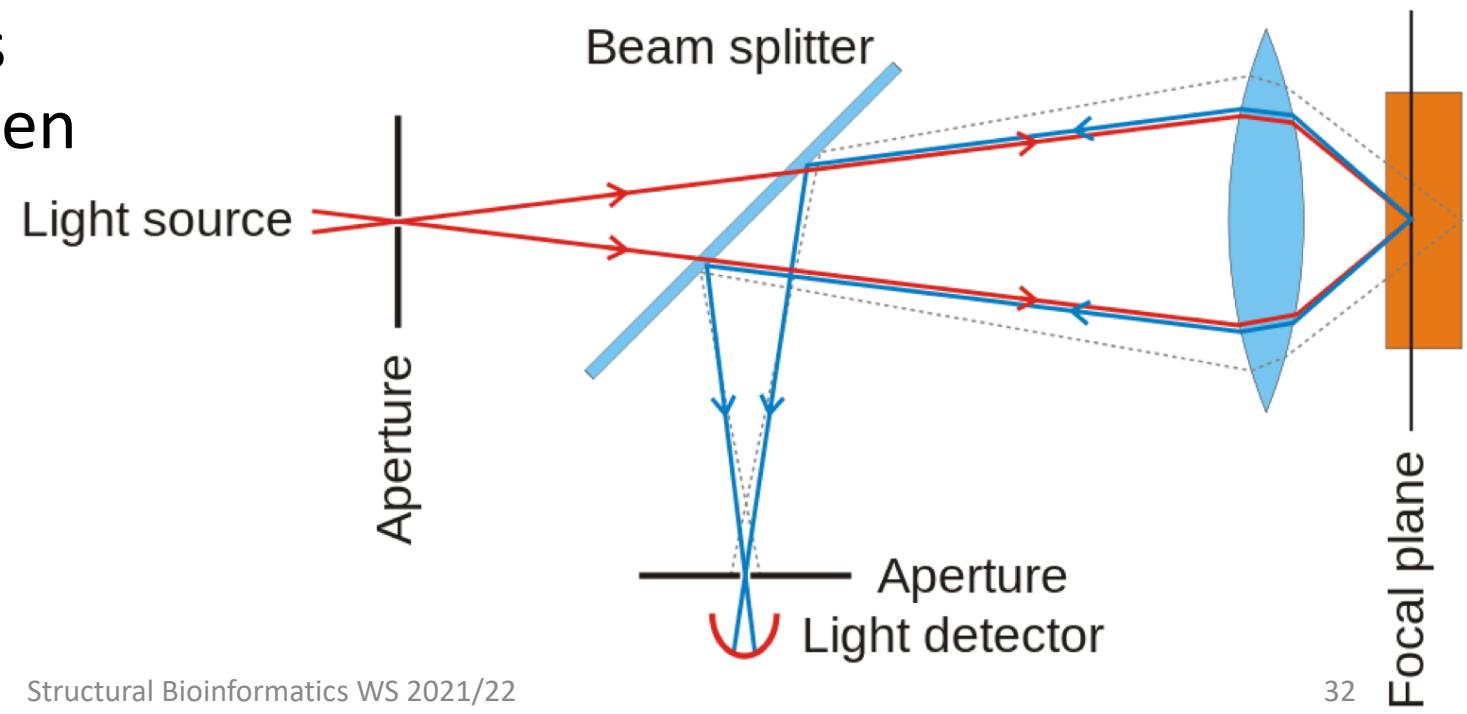
Electron microscopy

- Transmission electron microscopy (TEM)
- Resolution in the order of Å, in older (cheaper) devices $> 3\text{-}4 \text{ \AA}$ => much coarser than in X-ray or NMR
- Can deal with very large structures => amenable for large protein complexes or more
- Specimen has to be thin: $\sim 100 \text{ nm}$
- Scanning TEM (STEM): focus electron beam and scan the specimen in raster



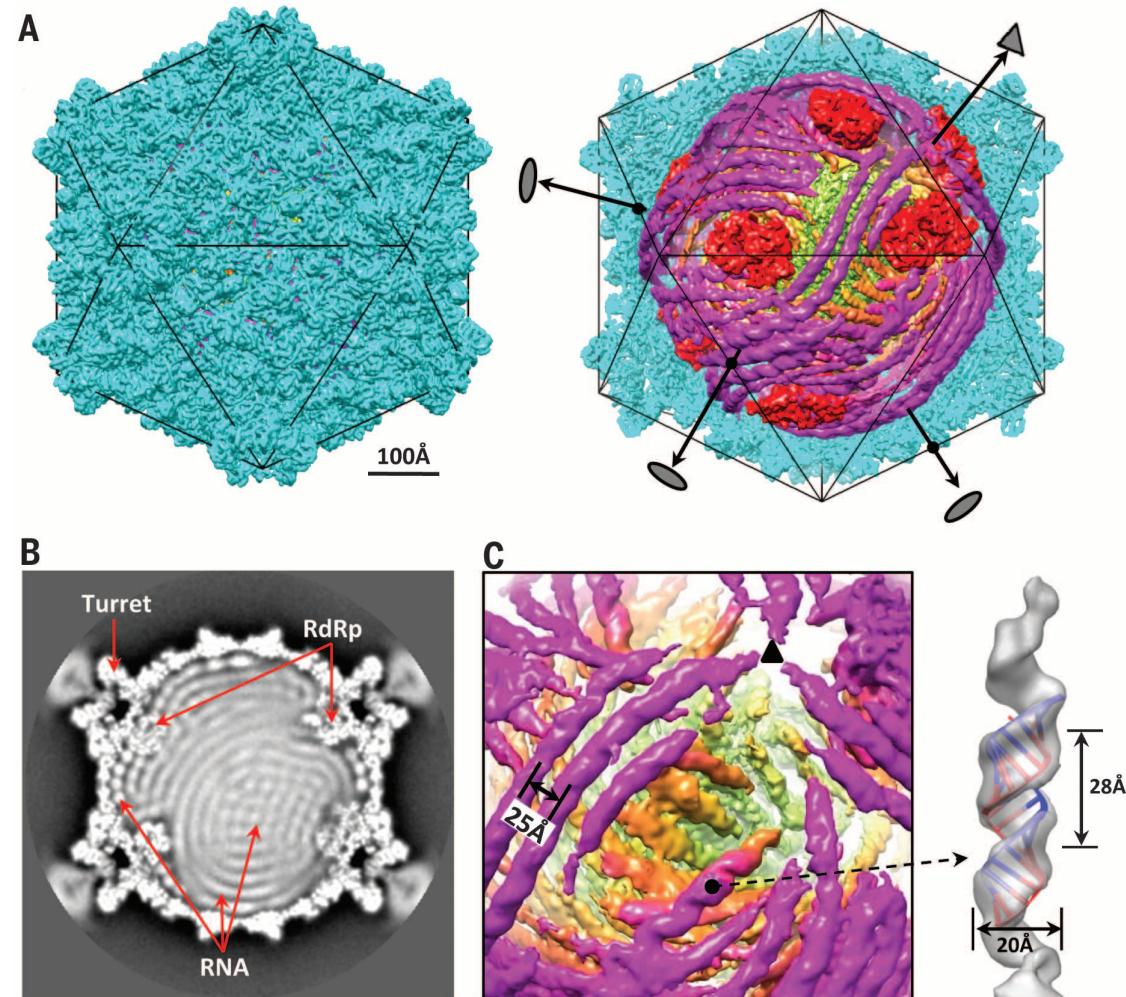
Confocal electron microscopy

- Removes out-of-focus signal
- Additional aperture in front of the detector
- Scanning confocal electron microscopy (CSEM) offers a 3D image of the specimen (theoretically, not made its way into practice yet)



Cryo-electron microscopy

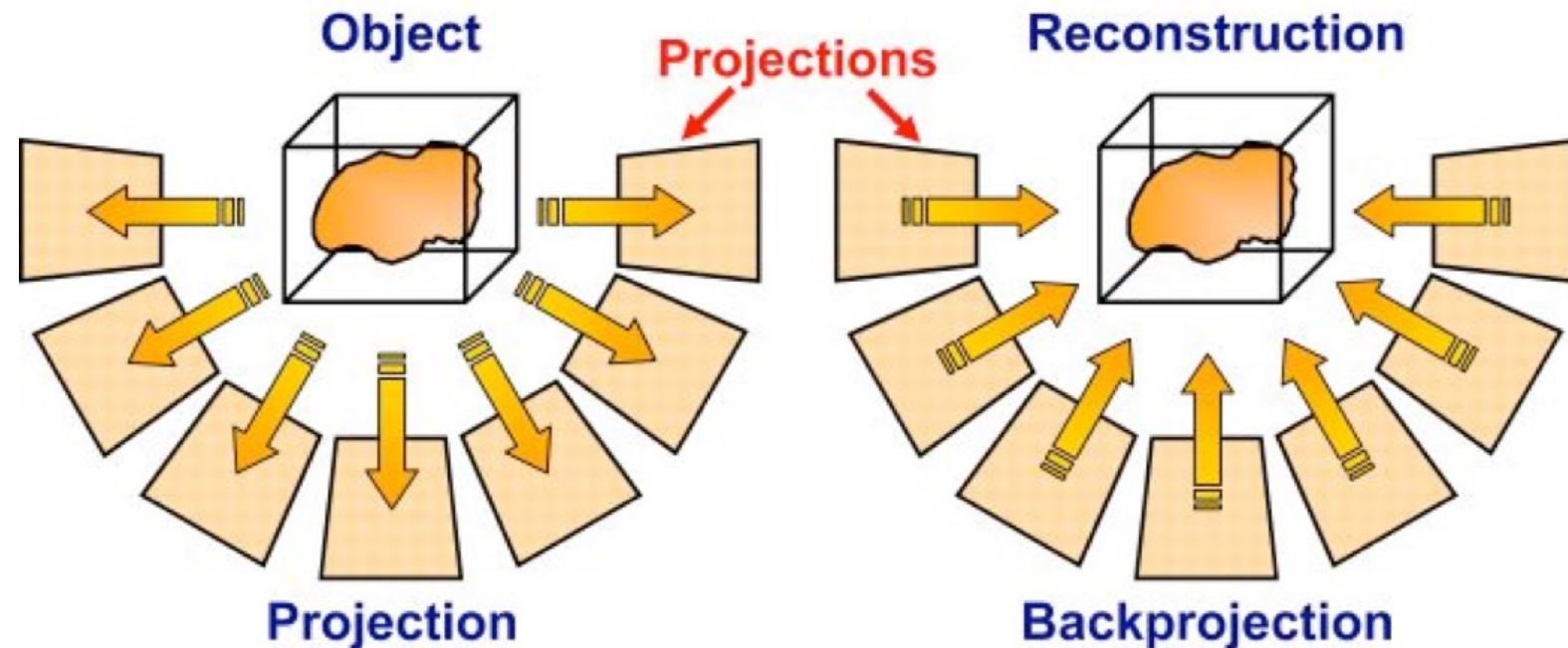
- A version of TEM when the sample is studied at very low temperature, e.g. frozen by liquid nitrogen
- Reduced radiation and vacuum damage
- Keeps sample hydrated



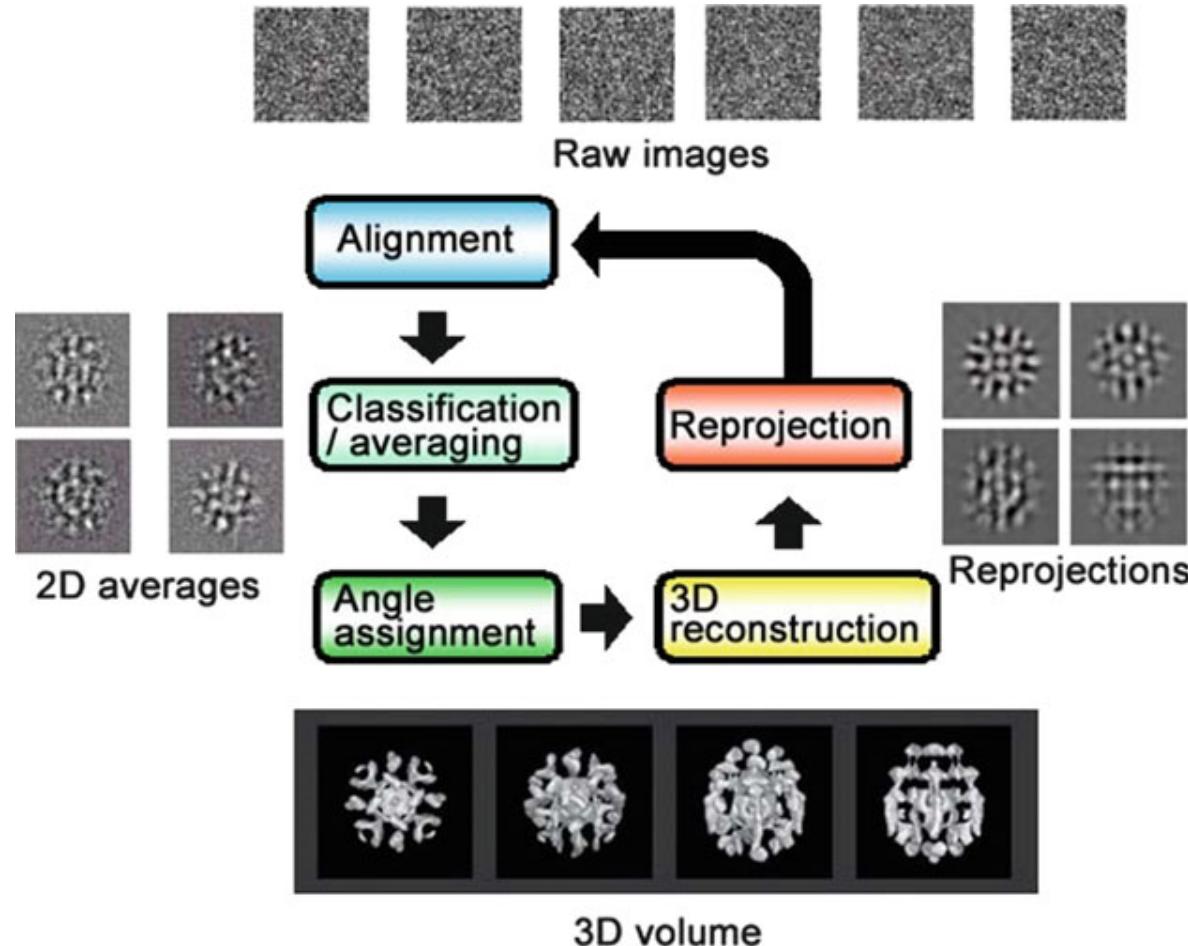
Cypovirus particle resolved with CryoEM,
Liu & Cheng, Science, 2015

Reconstruction of 3D image from 2D data (single particle analysis)

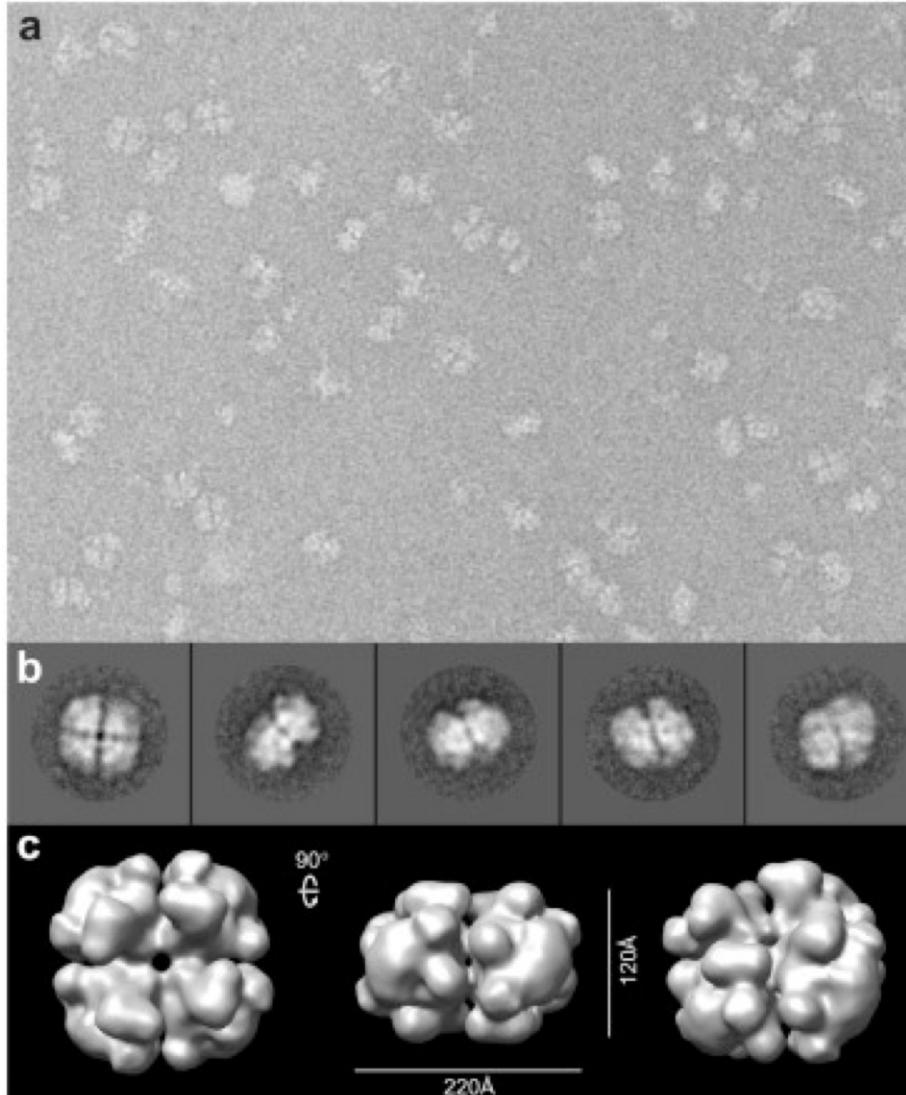
- You see your structure from different angles; or you rotate the sample
- This allows to reconstruct the 3D shape from 2D projections



Reconstruction of 3D image from 2D data (single particle analysis)

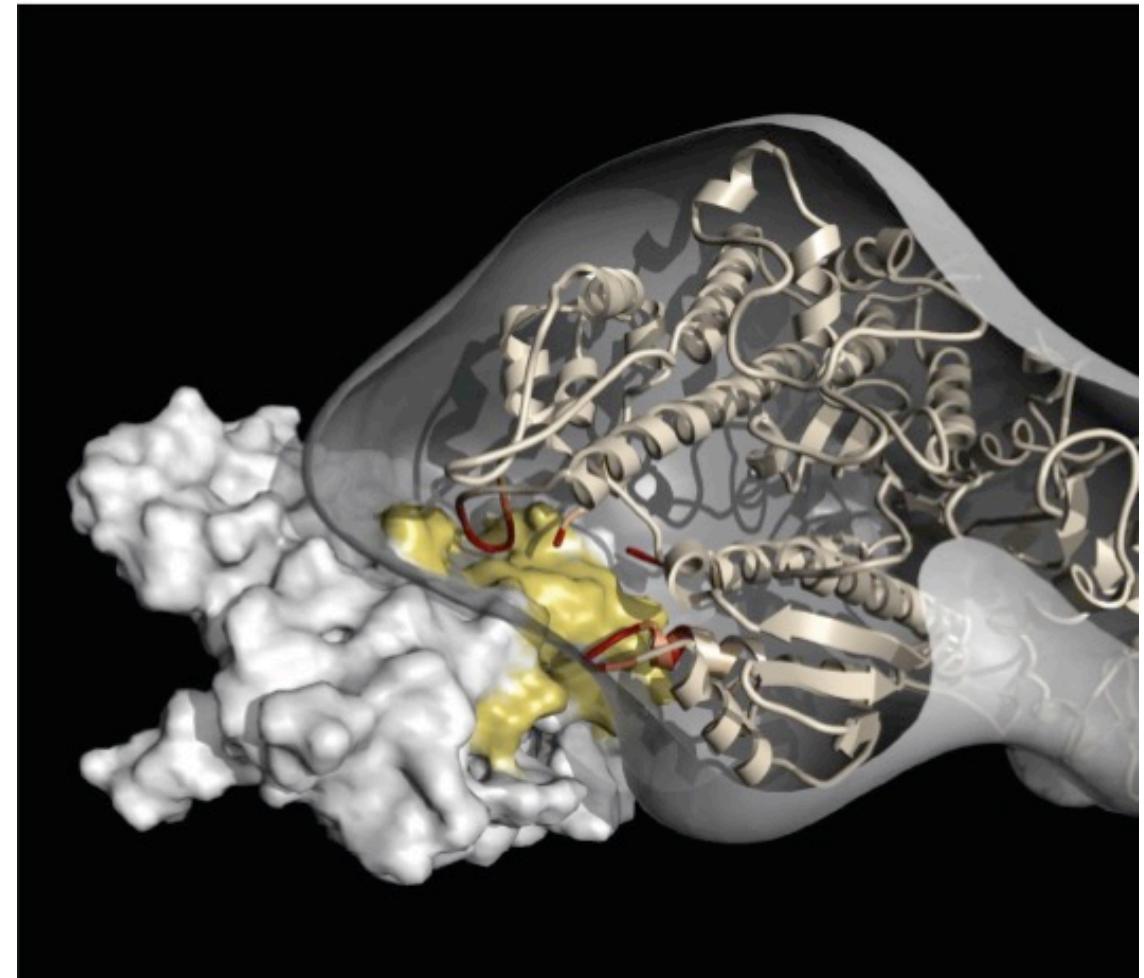


Reconstruction of 3D image from 2D data (single particle analysis)

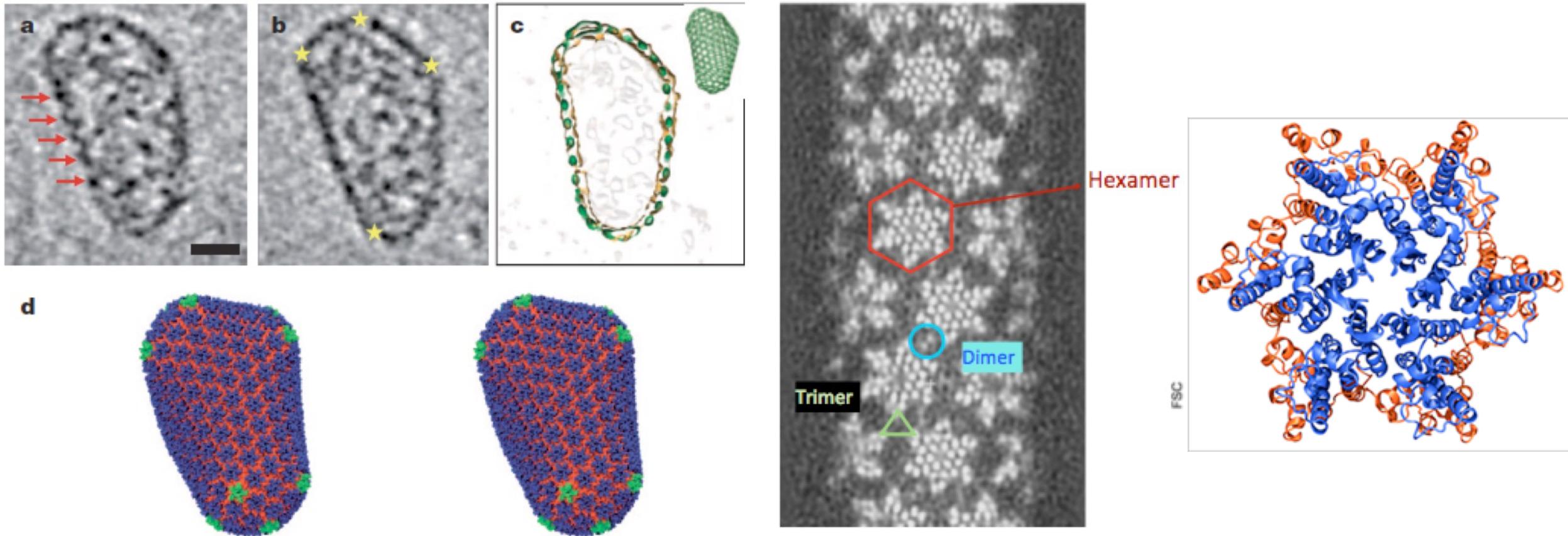


Hybrid approaches: electron microscopy + X-ray or NMR

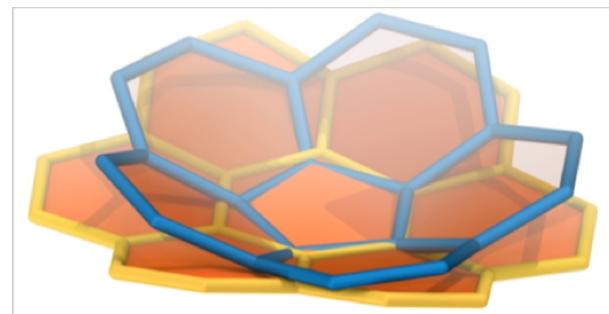
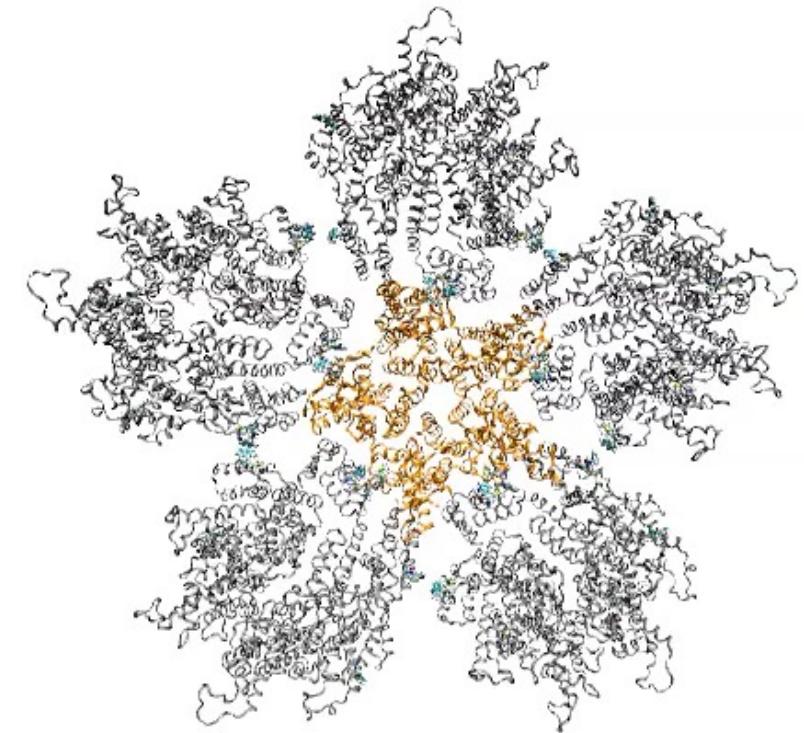
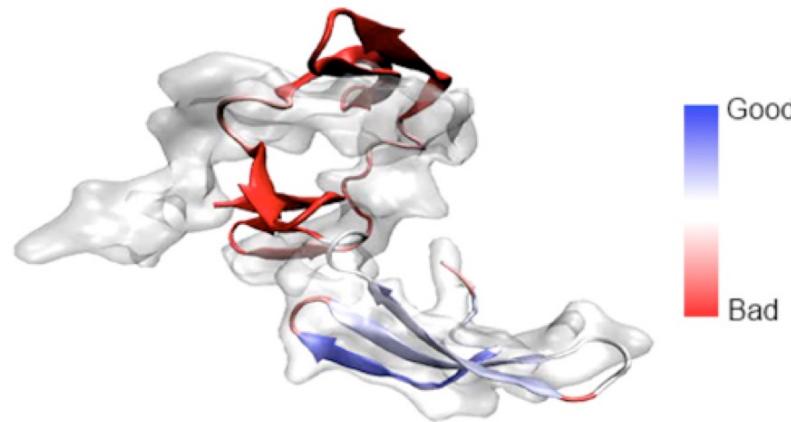
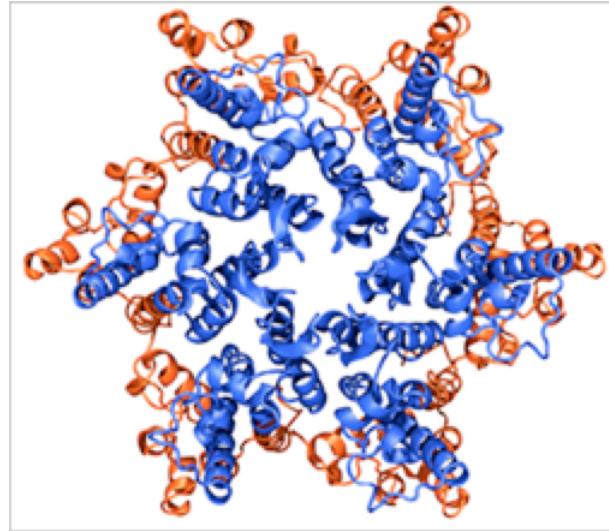
- EM gives you coarse image of a large structure; X-ray or NMR provide precise images of small parts of it



Bioinformatics-assisted 3D reconstruction: HIV-1 capsid



Bioinformatics-assisted 3D reconstruction: HIV-1 capsid



Flexible fitting: fitting an X-ray
structure into an EM map

The three techniques in comparison

X-ray crystallography	NMR spectroscopy	Cryo-electron microscopy
(+) Atomic resolution	(+) Atomic resolution	(+/-) Typically lower resolution (improving)
(+) Molecules and complexes of any size	(-) Limited to small to medium-sized molecules	(+) Applicable to very large molecules and complexes
(-) Sample needs to be crystallizable	(+) Sample in solution	(+) Close-to-native state
(-) Molecules need to be available in large quantity	(-) Molecules need to be available in large quantity	(+) Requires only small amounts of sample
(-) Static image	(+) Dynamics and interactions	(+) Dynamics and interactions

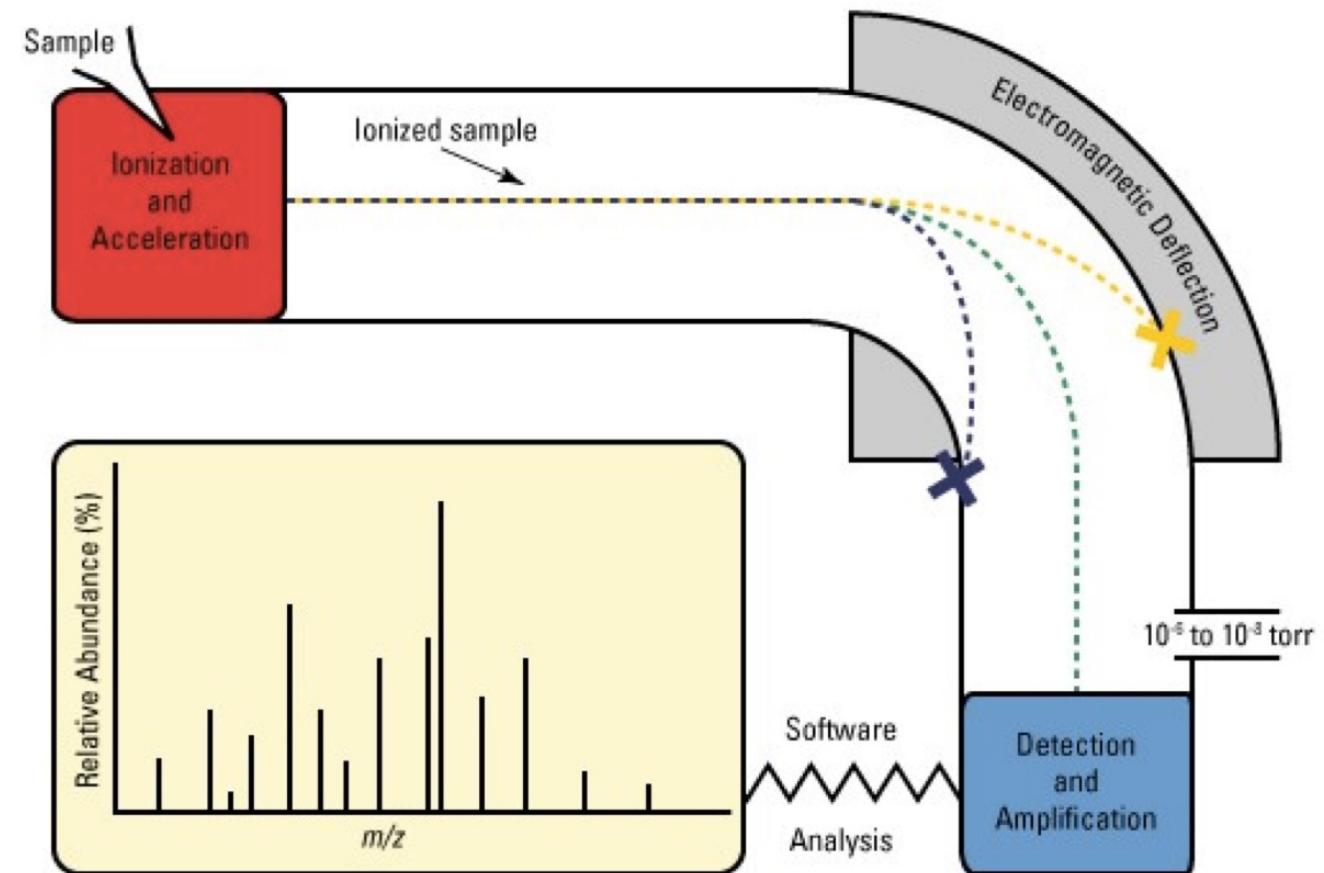
Protein detection with mass spectrometry

Mass spectrometry

- Identify proteins
- Identify other non-ribosomal peptides and ligands
- Deduce architecture of protein complexes
- DXMS to study protein dynamics
- Combine with other techniques to identify protein complexes

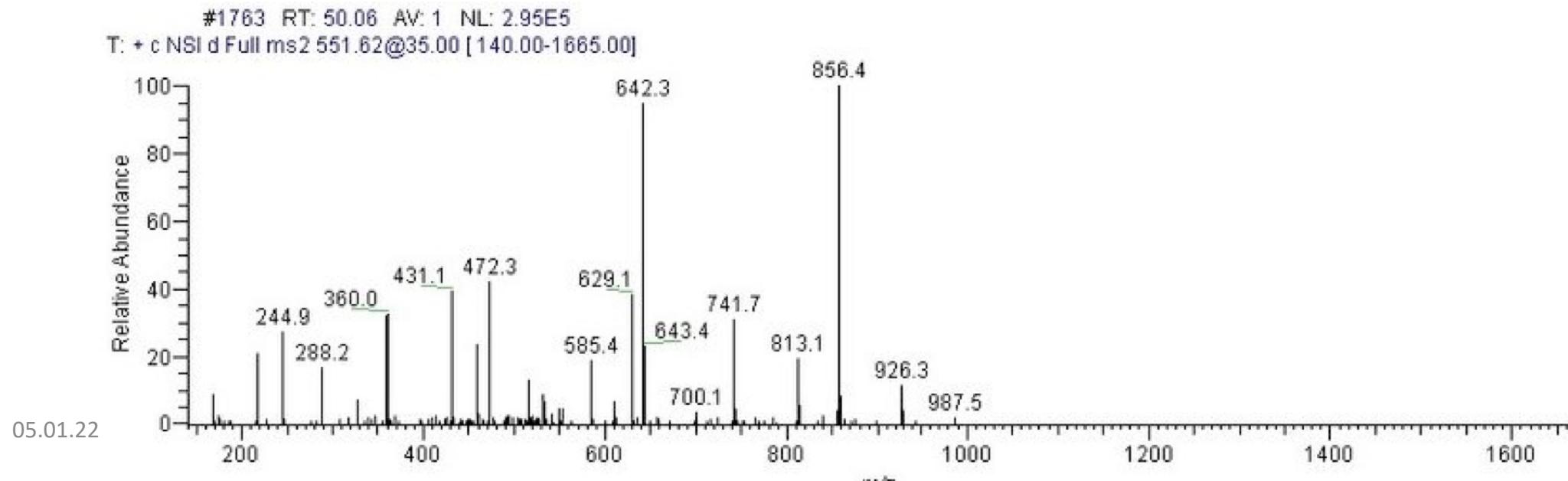
Mass spectrometry

- Sample preparation
(proteolytic digestion)
- Ionization and acceleration
- Detection of m/z ratio



Protein identification

- For a given peptide, m/z ratio can be computed theoretically
- => a limited number of peptides for each peak of the spectrum
- Known proteases used for digestion => known break points
- Database of all proteins (e.g. Uniprot)



Non-ribosomal proteins and ligands

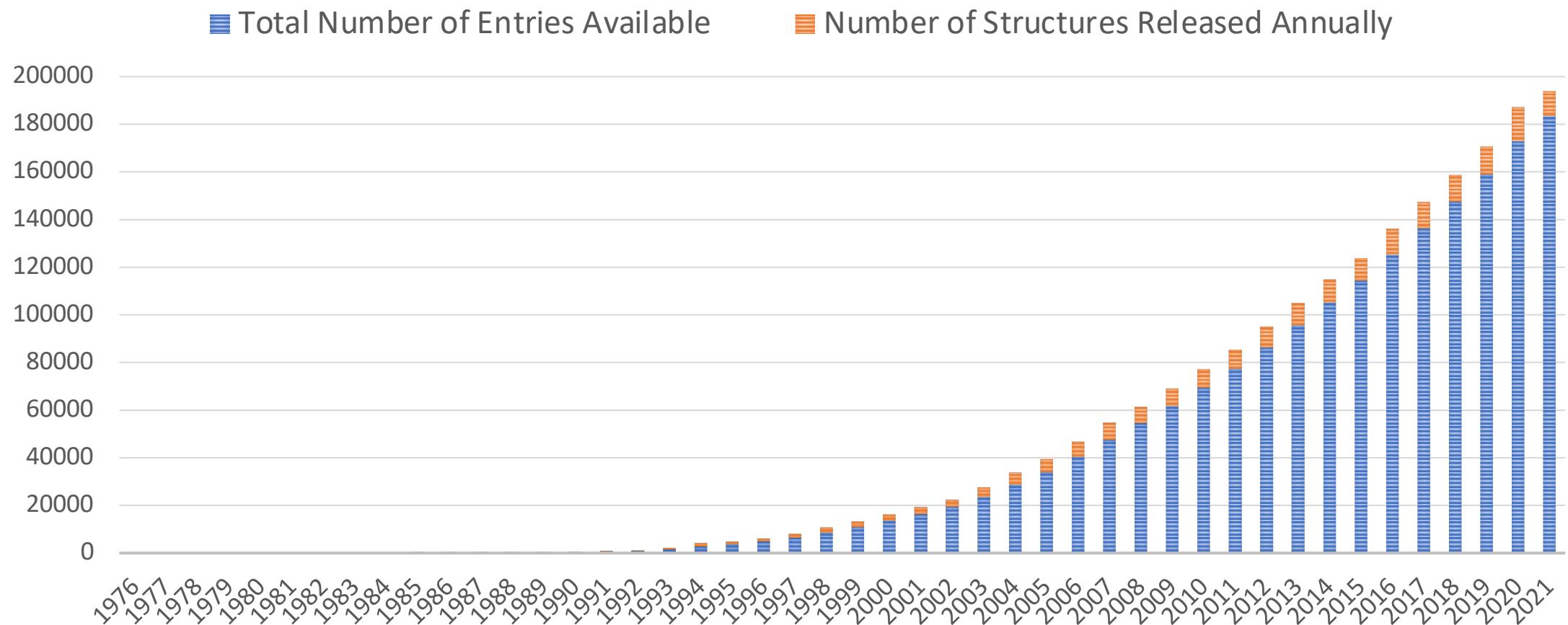
- Non-ribosomal peptides:
 - No standard amino acids
 - No database
 - A lot of them are cyclic
 - Can be analyzed with **tandem MS (MS/MS)**: two fragmentation steps (e.g. Ng et al., *Nature Methods* 2009)
- Ligands
 - Need a database of known molecules
 - Cannot be identified by MS alone

The Protein Data Bank

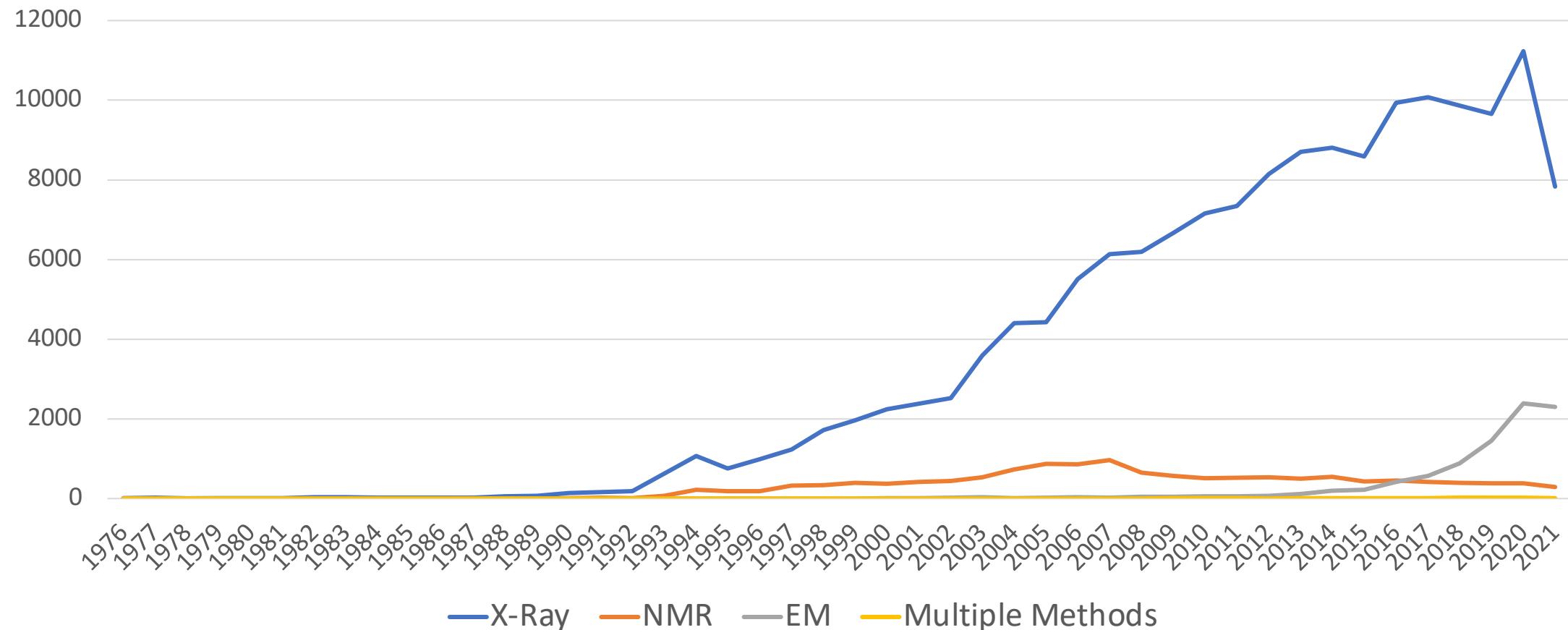
The Protein Data Bank (PDB)

- A repository to store all available 3D structures of biomolecules
 - Similar to GenBank for sequences
- <https://www.rcsb.org/>
 - <http://www.wwpdb.org/>, <https://www.ebi.ac.uk/pdbe/>,
<https://pdbj.org/>
 - No structural biology paper is accepted without depositing coordinated to the PDB

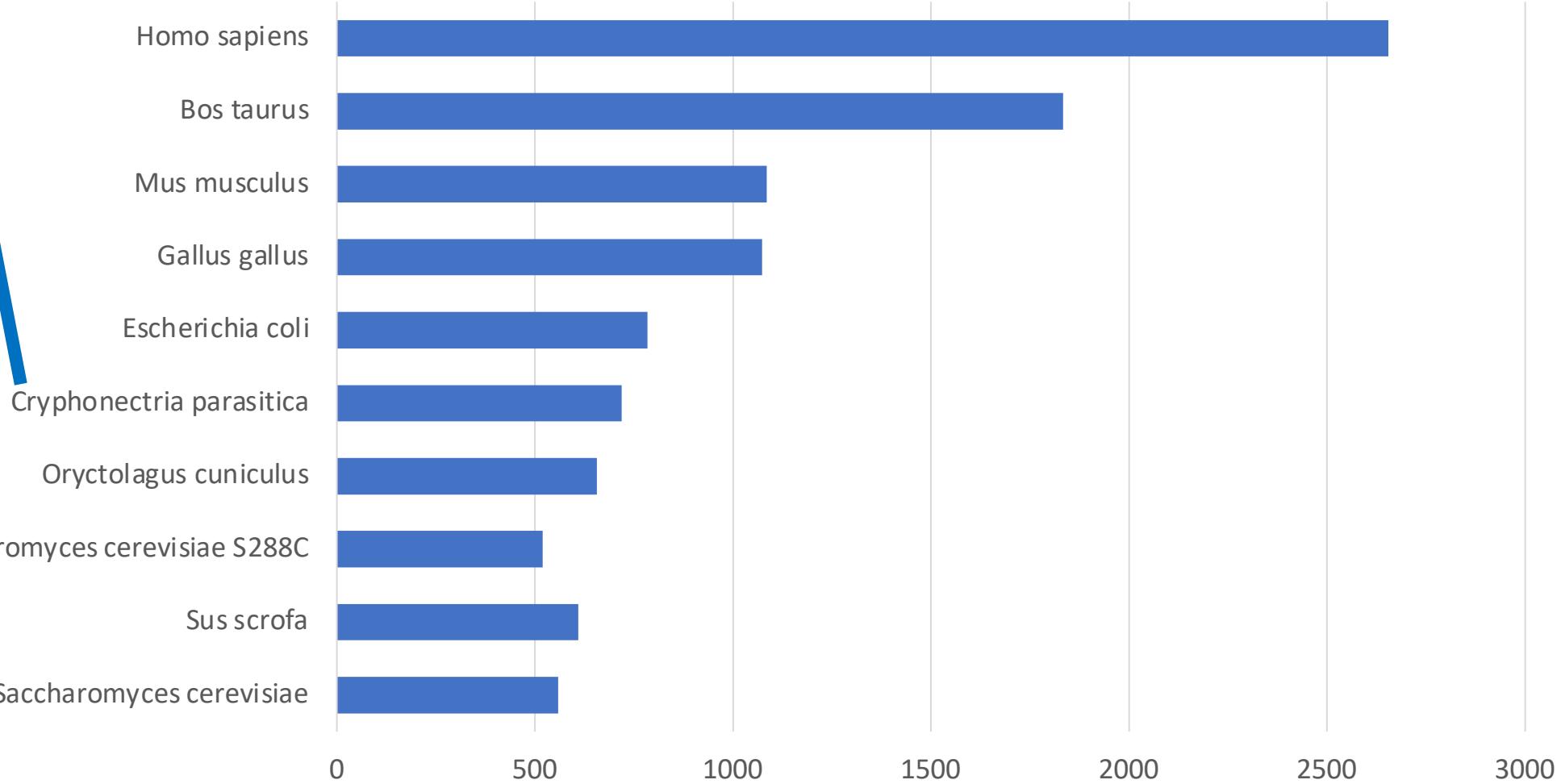
Annual data growth



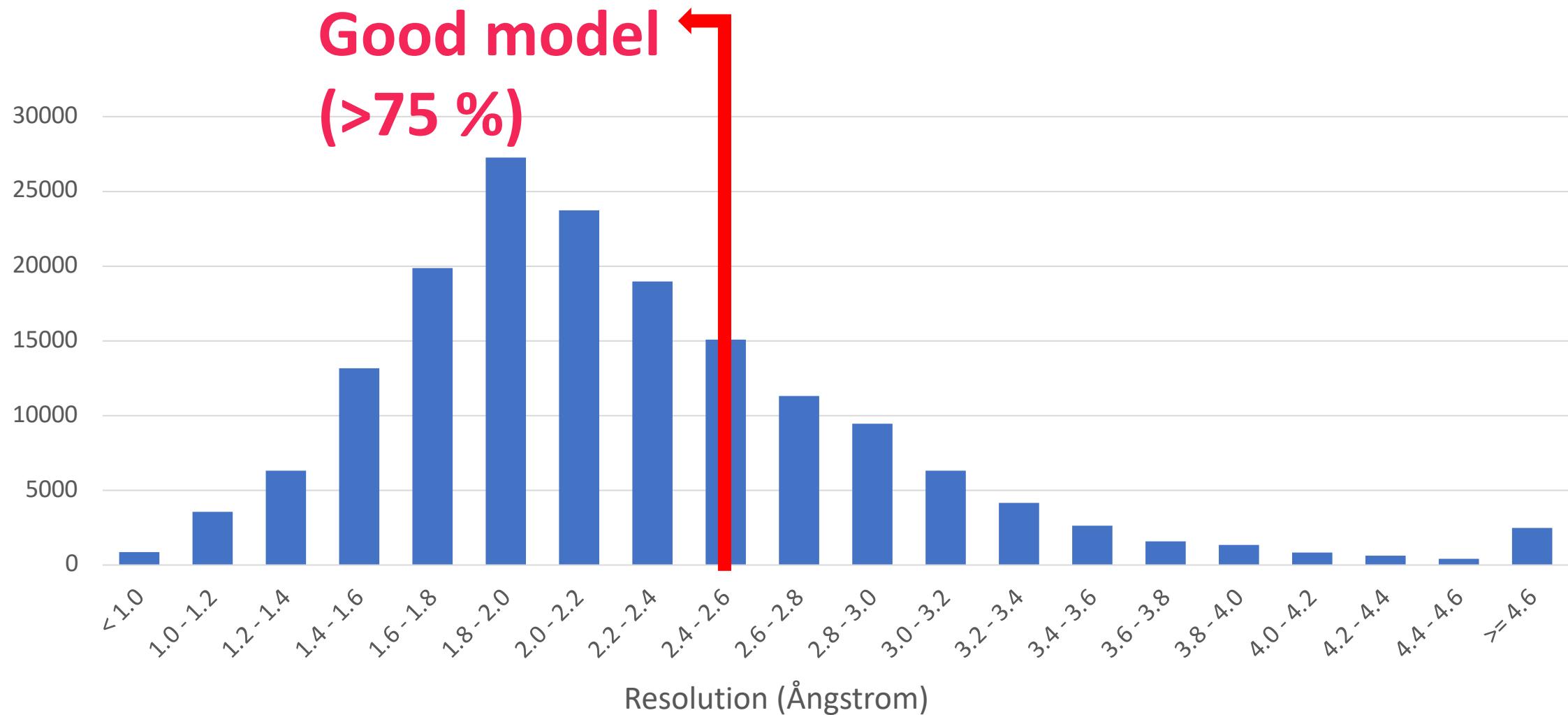
Data growth by experimental method per year



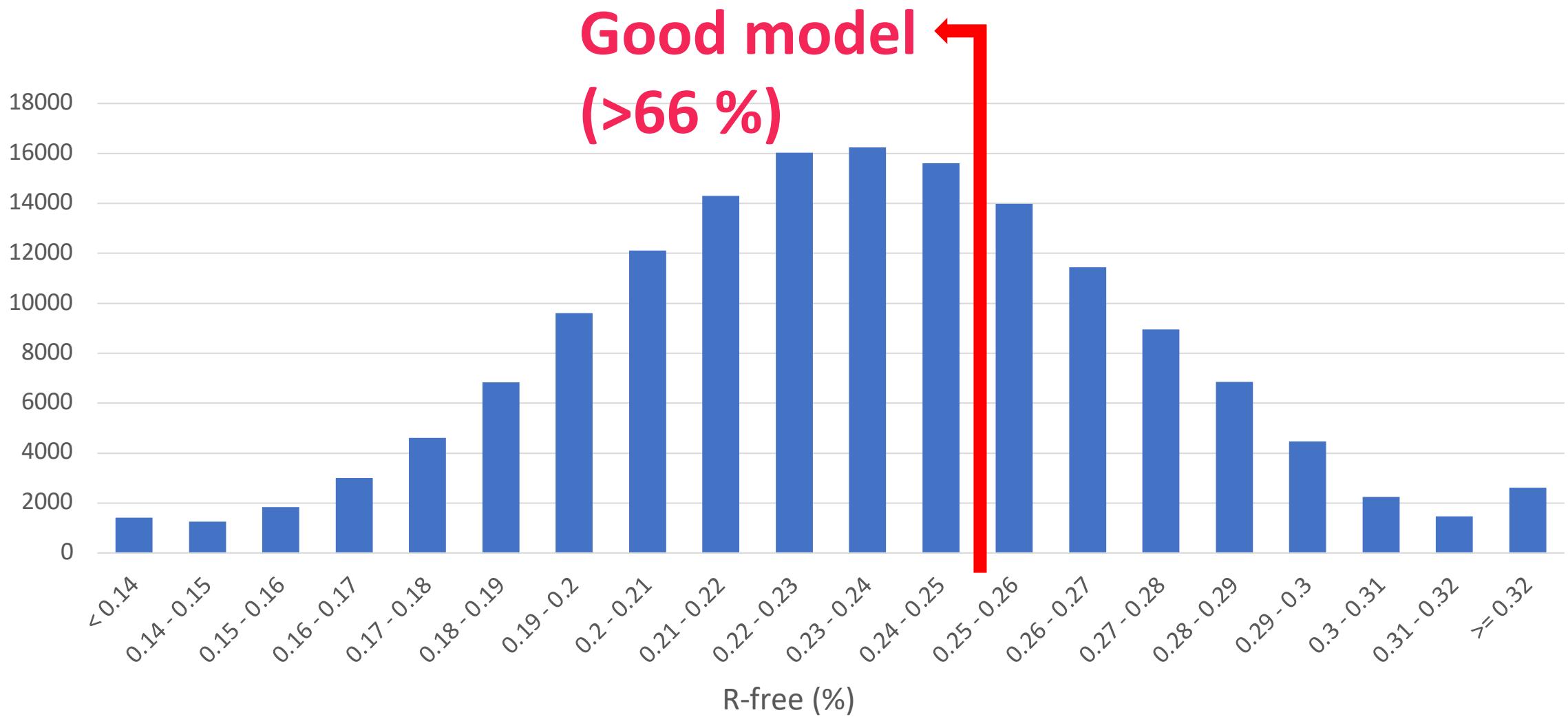
Top source organisms



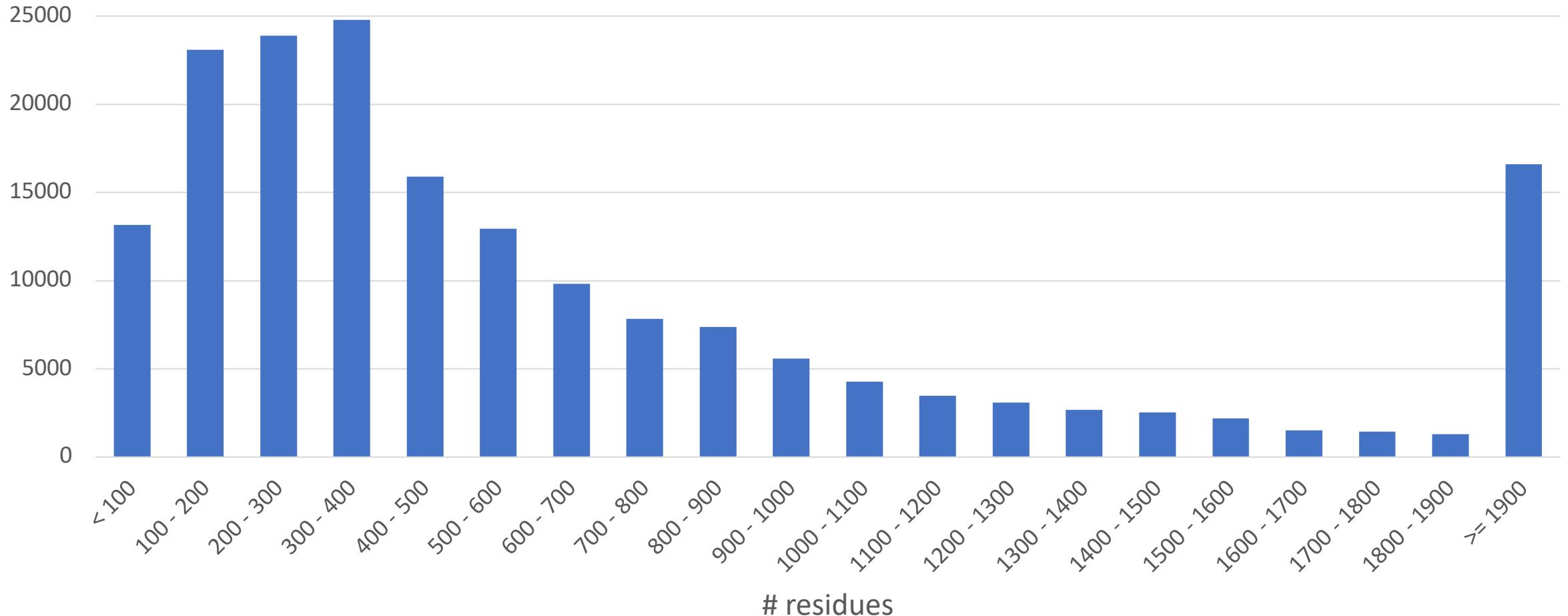
Structures' resolution



Structures' R-free value

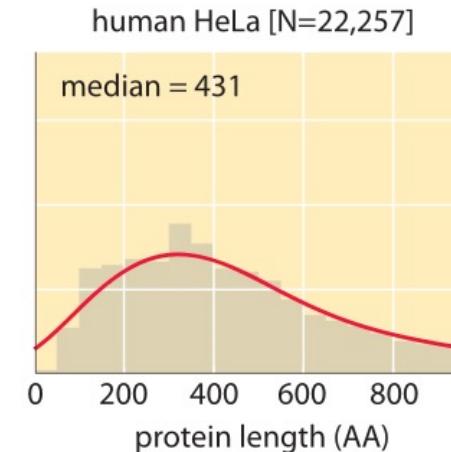
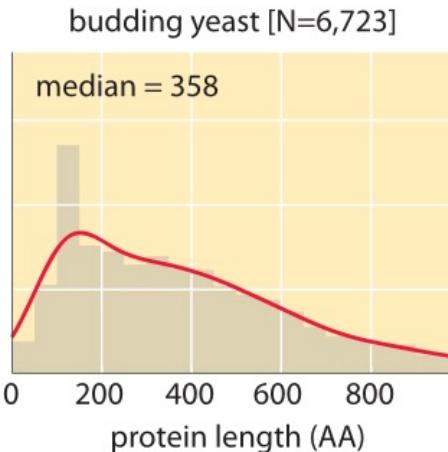
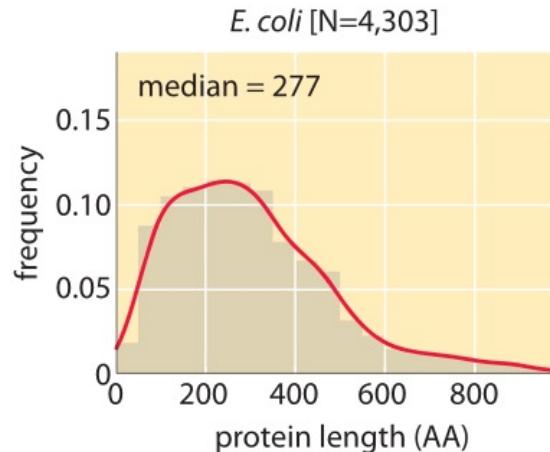


Structures' size (in residues)

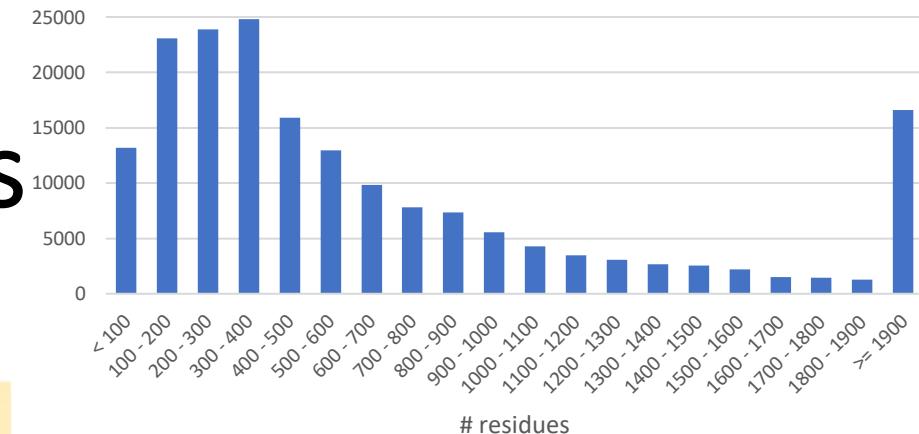
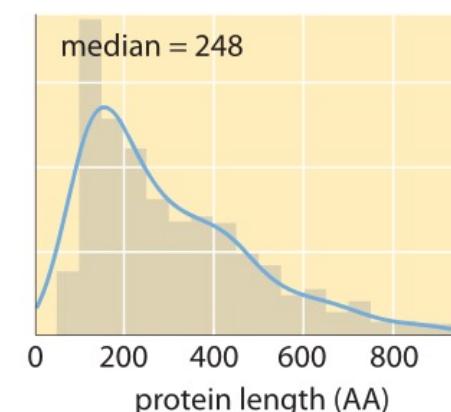
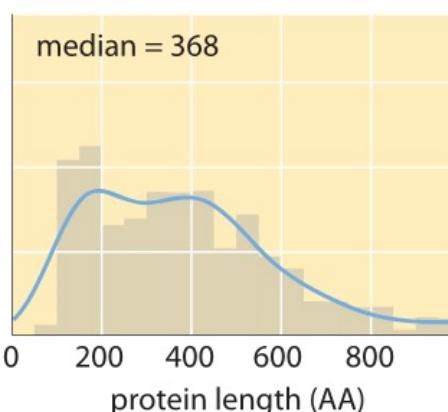
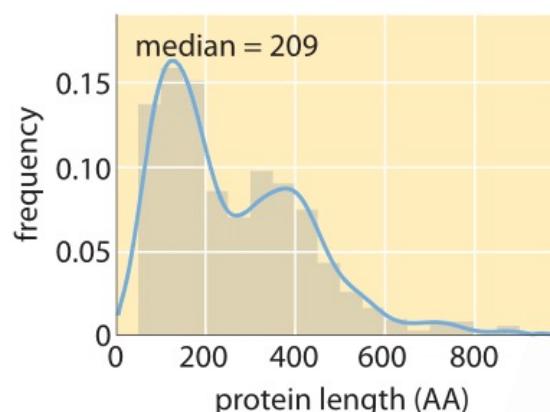


Median size/length of proteins

genomic length distribution



proteomic abundance weighted distribution

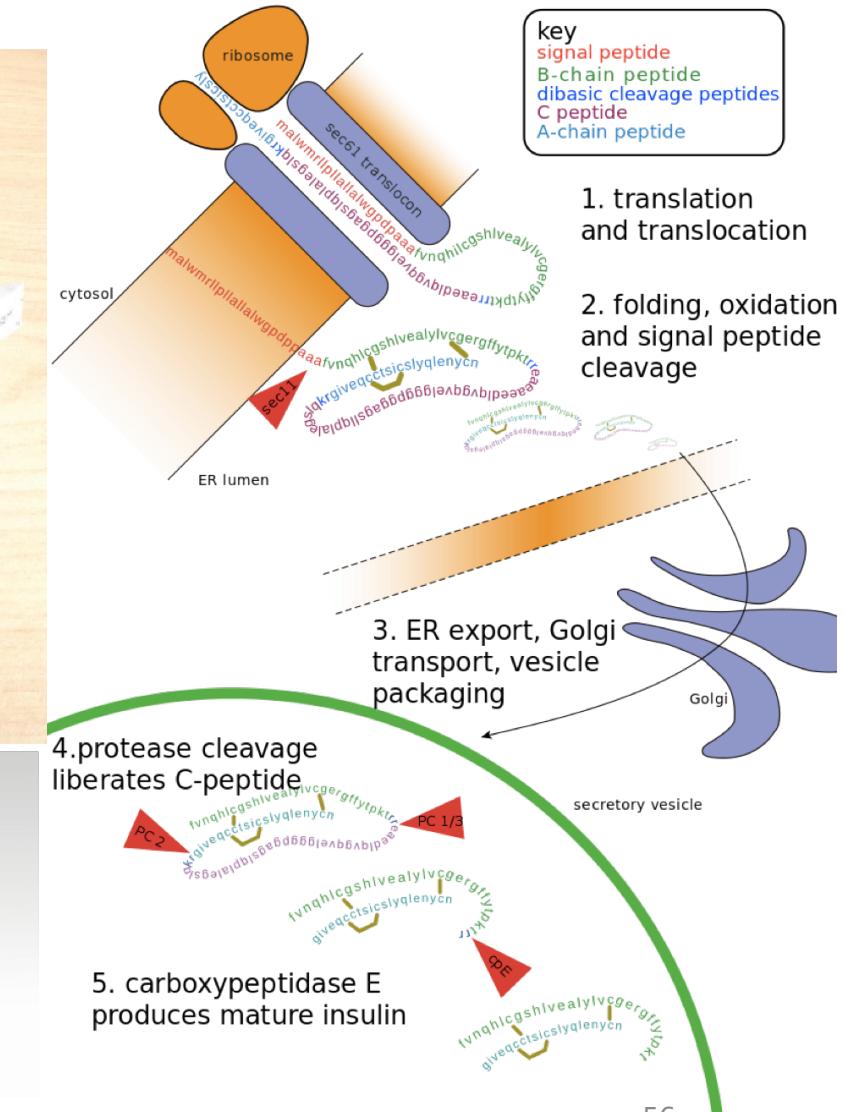


organism	median protein length (amino acids)
<i>H. sapiens</i>	375
<i>D. melanogaster</i>	373
<i>C. elegans</i>	344
<i>S. cerevisiae</i>	379
<i>A. thaliana</i>	356
5 eukaryotes (above)	361
67 bacteria	267
15 archaea	247

However, there are proteins longer than 1000, even 10000 aa!

PDB-101: an educational portal of PDB

- <http://pdb101.rcsb.org>
- Molecule of the Month
- Tutorials
- Videos
- Paper models, e.g.
<http://pdb101.rcsb.org/paper-models/insulin>



The anatomy of a PDB entry

Further exploring options

Structure Summary 3D View Annotations Experiment Sequence Genome Ligands Versions

Biological Assembly 1 ?

2E2I

RNA polymerase II elongation complex in 5 mM Mg+2 with 2'-dGTP

DOI: [10.22110/pdb2E2I/pdb](https://doi.org/10.22110/pdb2E2I/pdb) NDB: PH0029

Classification: TRANSCRIPTION,TRANSFERASE/DNA-RNA HYBRID

Organism(s): *Saccharomyces cerevisiae*

Mutation(s): No ⓘ

Deposited: 2006-11-14 Released: 2006-12-19

Deposition Author(s): Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., Kornberg, R.D.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 3.41 Å
R-Value Free: 0.316
R-Value Work: 0.266
R-Value Observed: 0.269

wwPDB Validation ⓘ

Metric	Percentile Ranks	Value
Rfree	21	0.300
Clashscore	4.2%	21
Ramachandran outliers	10.2%	4.2%
Sidechain outliers	4.0%	10.2%
RSRZ outliers	0.20	4.0%
RNA backbone	Worse	Better

3D View: Structure | Electron Density |
Ligand Interaction

Global Symmetry: Asymmetric - C₁
Global Stoichiometry: Hetero 10-Mer
- A1B1C1D1E1F1G1H1I1J1 ⓘ

Structure view

Display Files ▾ Download Files ▾

Download options

Method and quality

3D view

Sequence of 2E2I | RNA p... Chain 1: 5'-R(*AP*U... A [auth R]

1 6
AUCGAGAGGA

Structure

2E2I | RNA polymerase II elongatio...

Type Assembly

Asm Id 1: Author Defined Asse...

Dynamic Bonds X Off

Nothing Focused

Measurements

Structure Motif Search

Components 2E2I

Preset + Add

Polymer Cartoon

Ligand Ball & Stick

Ion Ball & Stick

Unit Cell C 1 2 1

Density

Assembly Symmetry

Export Animation

05.01.22

Sequence view

INSTANCE

D [auth A]

DNA-directed RNA polymerase II largest subunit - *Saccharomyces cerevisiae* [View Features in 3D](#)

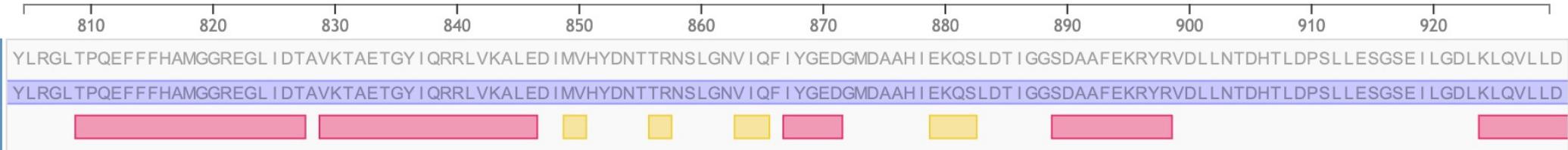
PDB INSTANCE **D[auth A]**

YLRGLTPQEFFFHAMGGREGI DTAVKTAETGY I QRRLVKALED IMVHYDNTTRNSLGNV I QF I YGEDGMDAAH I EKQLSLDT I GGSDAAFEKRYRVDLLNTDHTLDPSLLESGSE I LGDLKLQVLLD

UNIPROT **P04050**

YLRGLTPQEFFFHAMGGREGI DTAVKTAETGY I QRRLVKALED IMVHYDNTTRNSLGNV I QF I YGEDGMDAAH I EKQLSLDT I GGSDAAFEKRYRVDLLNTDHTLDPSLLESGSE I LGDLKLQVLLD

SECONDARY STRUCTURE



UNMODELED

RAMACHANDRAN OUTLIER

ANGLE OUTLIER

BOND OUTLIER

ROTAMER OUTLIER

RSRZ OUTLIER

RSCC OUTLIER

CIS-PEPTIDE

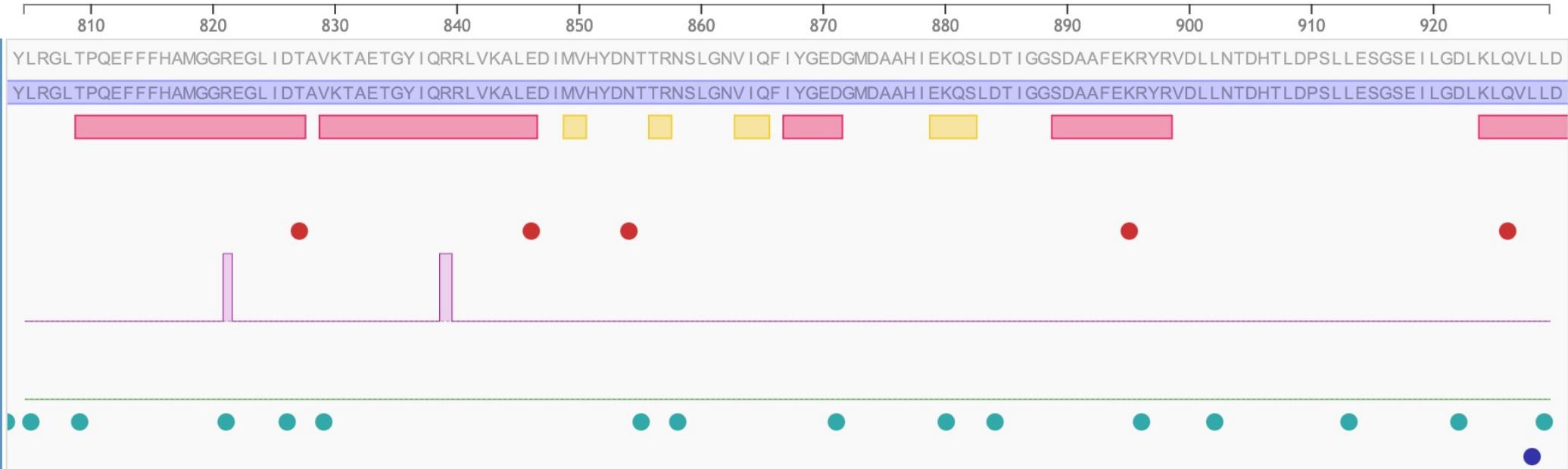
METAL COORDINATION

BINDING SITE **DGT**

BINDING SITE **MG**

BINDING SITE **ZN**

BINDING SITE **ZN**



Downloadable files

- Single structure or batch downloads
- rsync possible using an FTP site
- PDB or mmCIF format
 - PDB format is human-readable
 - Reference:
<http://www.wwpdb.org/documentation/format32/v3.2.html>

Resources for validation and analysis of protein structures

Overview of resources

- PDBsum (<http://www.ebi.ac.uk/pdbsum/>): next-choice resource after you have studied PDB
- Validation tools:
 - Procheck (part of PDBsum)
 - WHAT IF
 - PDB validation server
 - ...

PDBsum

Go to PDB code: 2e2i go  

Top page Protein DNA/RNA Ligands Metals Prot-prot Links

Transcription,transferase/DNA-RNA hybrid PDB id 2e2i

PDB id: 2e2i

Name: Transcription,transferase/DNA-RNA hybrid

Title: RNA polymerase ii elongation complex in 5 mm mg+2 with 2'- dgtp

Structure: 5'-r(Ap Up Cp Gp Ap Gp Ap Gp A)-3'. Chain: r. Engineered: yes. 28-mer DNA template strand. Chain: t. Engineered: yes. 5'-d(Cp Tp Gp Cp Tp Tp Ap Tp Cp Gp Gp Tp Ap G)-3'. Chain: n.

Source: Synthetic: yes. *Saccharomyces cerevisiae*. Baker's yeast. Organism_taxid: 4932. Strain: delta-rpb4. Strain: delta-rpb4

Resolution: 3.41Å **R-factor:** 0.269 **R-free:** 0.316

Authors: D.Wang,D.A.Bushnell,K.D.Westover,C.D.Kaplan,R.D.Kornberg

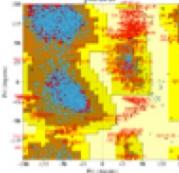
Key ref: D.Wang et al. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell*, 127, 941-954. PubMed id: [17129781](#)

DOI: [10.1016/j.cell.2006.11.023](#)

Date: 14-Nov-06 **Release date:** 19-Dec-06

Links

PROCHECK



Headers

References

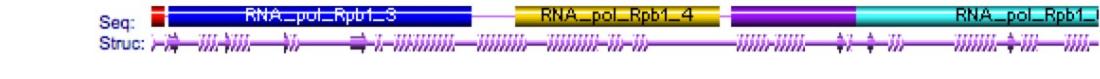
Protein chain A

P04050 (RPB1_YEAST) - DNA-directed RNA polymerase II subunit RPB1

Seq: RNA_pol_Rpb1_1 RNA_pol_Rpb1_2

Struc: 

Seq: RNA_pol_Rpb1_3 RNA_pol_Rpb1_4 RNA_pol_Rpb1_5

Struc: 

Seq: RNA_pol_Rpb1_6 RNA_pol_Rpb1_7

Struc: 

Seq: 1733 a.a.

Struc: 1411 a.a.

Protein chain B

PDBsum: protein tab

PDBsum

Go to PDB code: 2e2i go ?

Top page Protein DNA/RNA Ligands Metals Prot-prot Links PDB Id 2e2i

Protein chain A

Chain A (1411 residues)

UniProt code: P04050 (RPB1_YEAST) [Pfam]

Secondary structure:

The diagram illustrates the secondary structure of Protein chain A, which consists of 1411 residues. It features several alpha-helices (labeled H1 through H11) and beta-sheets. The structure is color-coded by residue type, with hydrophobic residues in purple and polar residues in yellow. The backbone of the protein is shown as a blue line, and the side chains of the amino acids are represented by colored sticks.

Protein chain A highlighted (click to view)

Jmol Snap Motifs

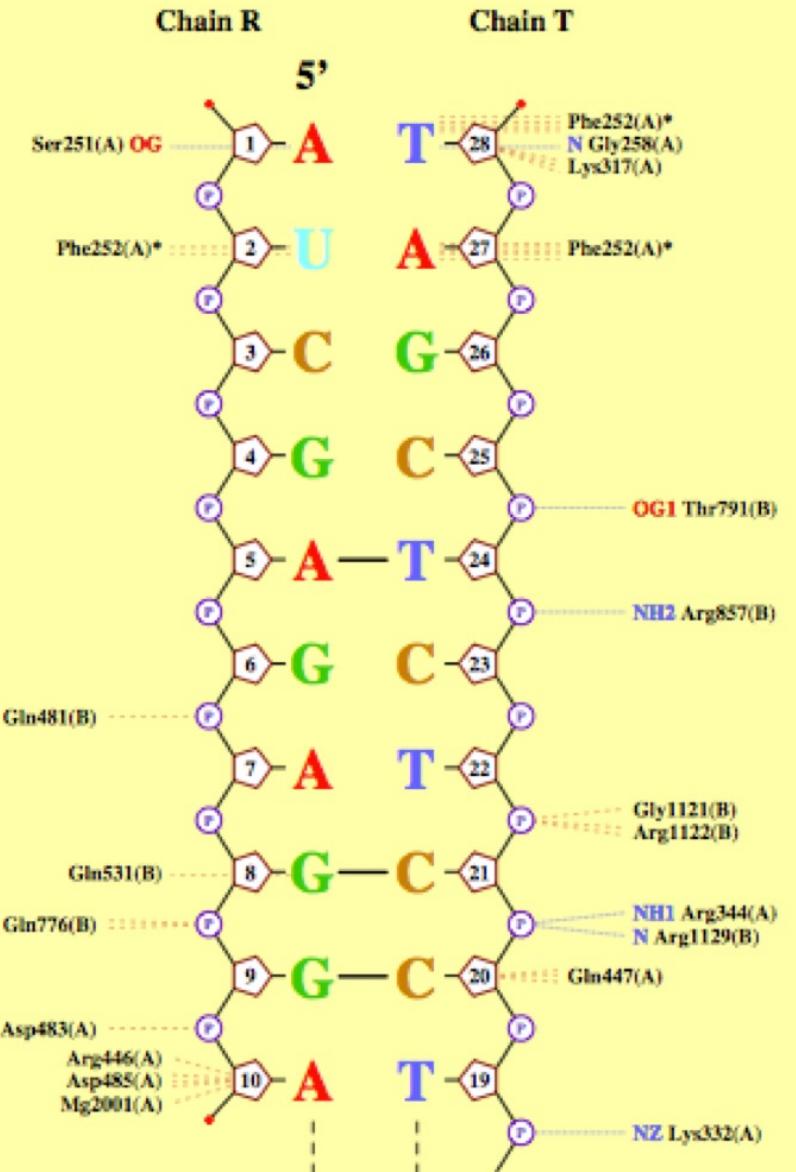
Secondary structure
Wiring diagram
Residue conservation
ProMotif
16 sheets
15 beta hairpins
3 psi loops
10 beta bulges
42 strands
56 helices

Topology

Postscript viewer Hera Postscript viewer

05.01.22 64

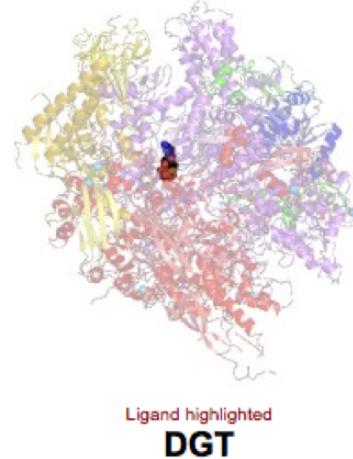
PDBsum: DNA/RNA tab



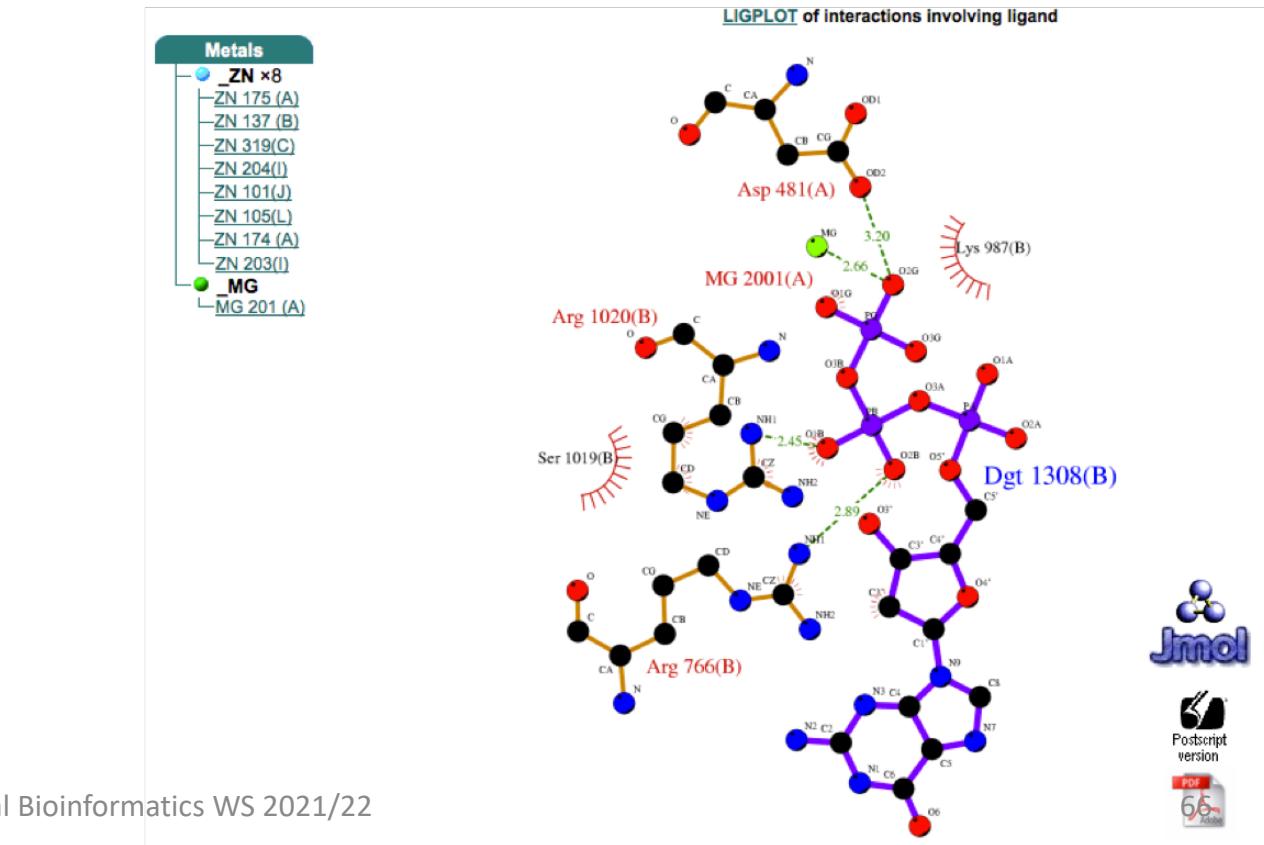
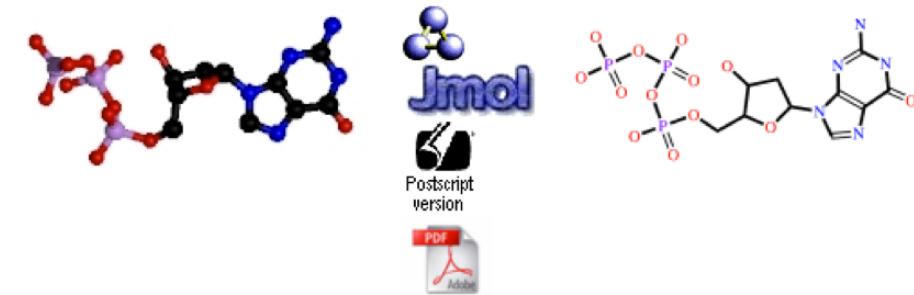
Key

- 3 Backbone sugar and base-number P Phosphate group * Residue/water on plot more than once
----- Hydrogen bond to DNA
..... Nonbonded contact to DNA (< 3.35Å)
88 W Water molecule and number

PDBsum: ligands tab



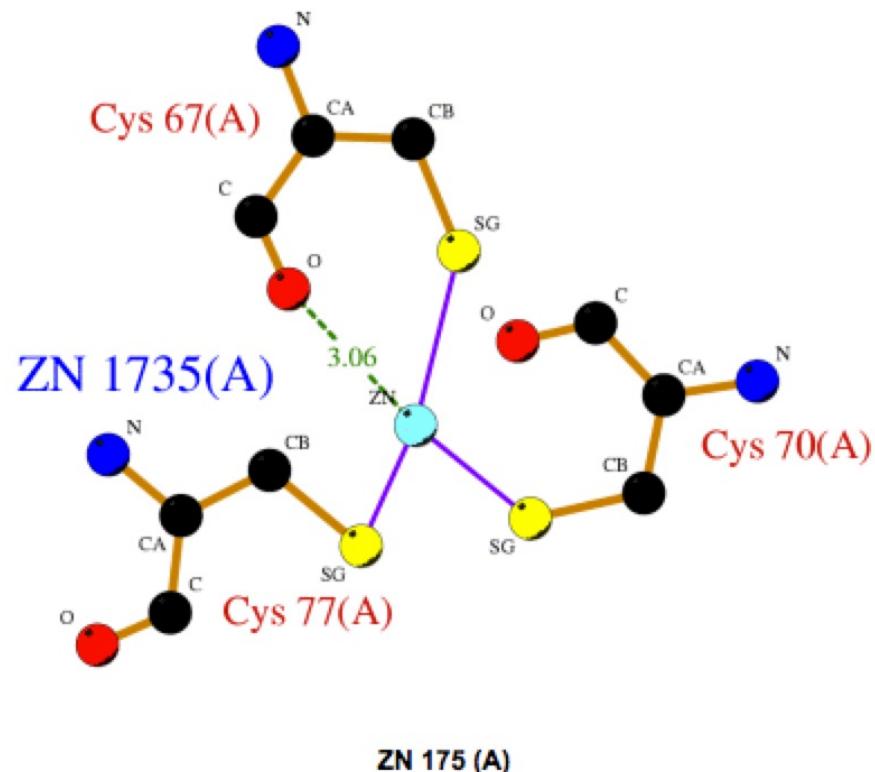
Ligand **DGT** - 2'-Deoxyguanosine-5'-Triphosphate
Formula: $C_{10}H_{16}N_5O_{13}P_3$



PDBsum: metals tab

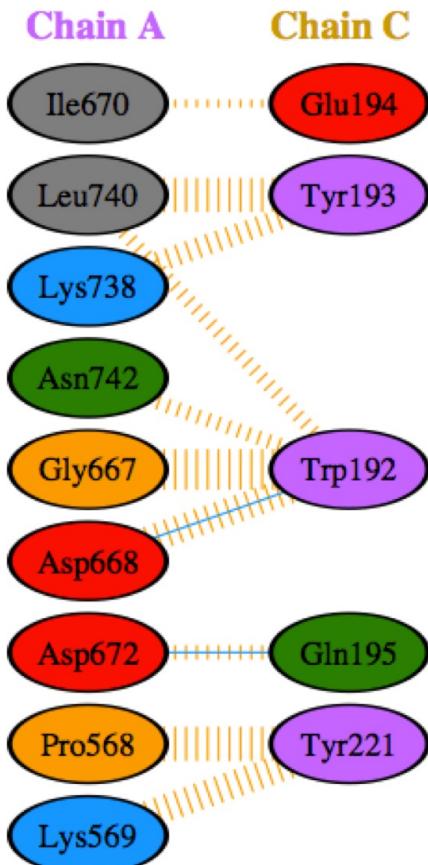
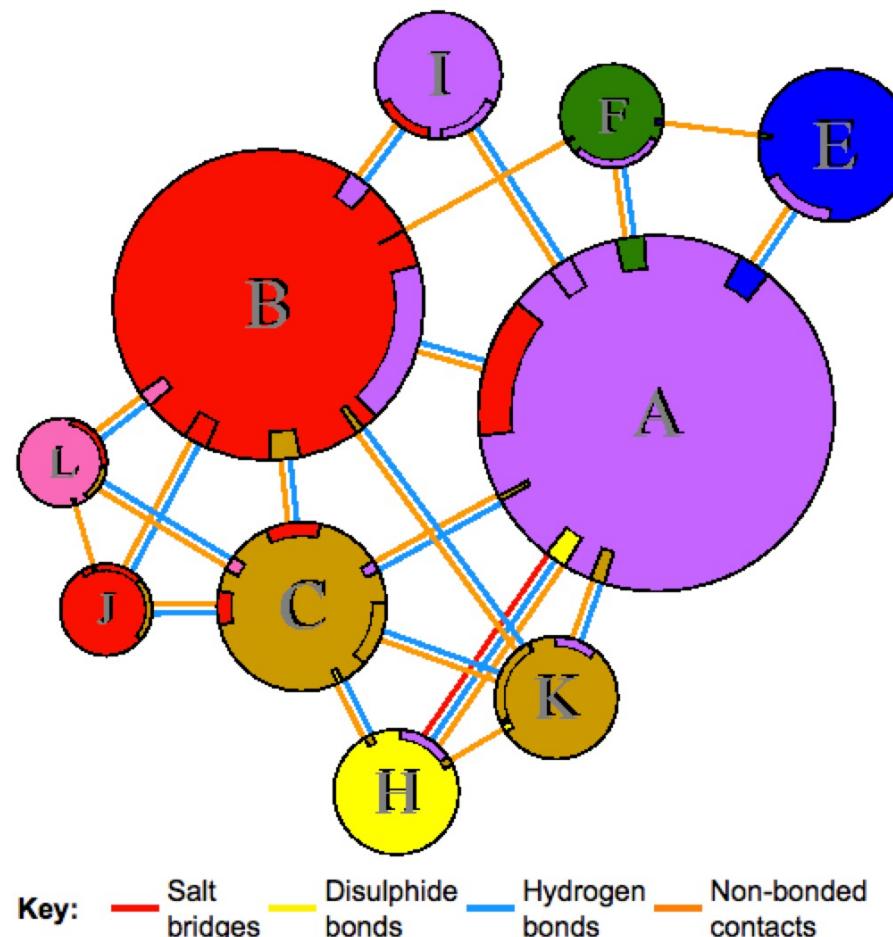
Metal ion _ZN - Zinc ion

LIGPLOT of interactions involving metal



PDBsum: protein-protein interactions tab

Interfaces summary for 2e2i



Interface statistics

Chains	No. of Interface residues	Interface area (Å ²)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
A _H B	181:171	9470:9615	-	-	77	1006
A _H C	9:5	379:385	-	-	2	34
A _H E	44:34	2189:2253	-	-	14	200
A _H F	25:32	1897:1855	-	-	8	122
A _H H	23:23	1492:1533	1	-	9	121
A _H I	27:27	1638:1586	-	-	13	131
A _H K	16:20	978:1020	-	-	4	96
B _H C	35:32	1746:1786	-	-	10	208
B _H F	1:1	44:49	-	-	-	1
B _H I	22:26	1493:1440	-	-	4	79
B _H J	38:24	1436:1689	-	-	7	182
B _H K	10:9	460:504	-	-	2	33
B _H L	22:18	948:1056	-	-	7	99
C _H H	3:2	178:198	-	-	1	8
C _H J	20:21	1104:1140	-	-	9	90
C _H K	28:33	2133:2105	-	-	6	114
C _H L	15:7	493:613	-	-	7	58
E _H F	1:2	130:114	-	-	-	9
H _H K	2:2	208:203	-	-	-	3
J _H L	1:1	57:52	-	-	-	2

Other repositories and added-value databases

Molecular Modeling Database (MMDB)

- Structure resource from NCBI
- [http://www.ncbi.nlm.nih.gov/
Structure/MMDB/mmdb.shtml](http://www.ncbi.nlm.nih.gov/Structure/MMDB/mmdb.shtml)
- Integrated into NCBI system of sequence, genome and literature information

The screenshot shows the NCBI Structure Home page for 3D Macromolecular Structures. At the top, there's a search bar with a dropdown set to "Structure" and a red box highlighting the search input field. To the right of the search bar is a "GO" button. Below the search bar, the page title "3D Macromolecular Structures" is displayed, along with a brief description of the database. A red box highlights the "Retrieve by MMDB ID or PDB ID:" input field and its "Go" button. On the left side, there's a section titled "Resources" with a link to "Molecular Modeling Database (MMDB)". A red box highlights the "CBLAST" link, which is described as a tool for comparing a query protein sequence against all protein sequences from resolved 3D structures. The word "CBLAST" is also highlighted with a red box. Red arrows point from the text labels "Freetext search", "Search using an ID", and "Sequence similarity search" to their respective highlighted areas on the page.

Freetext search

Search using an ID

Sequence similarity search

RNA Polymerase II Elongation Complex in 5 MM Mg⁺² With 2'- Dgtp

Citation: [?](#)

Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis.

Wang D, Bushnell DA, Westover KD, Kaplan CD, Kornberg RD

Cell(Cambridge, Mass.) (2006) 127 p.941

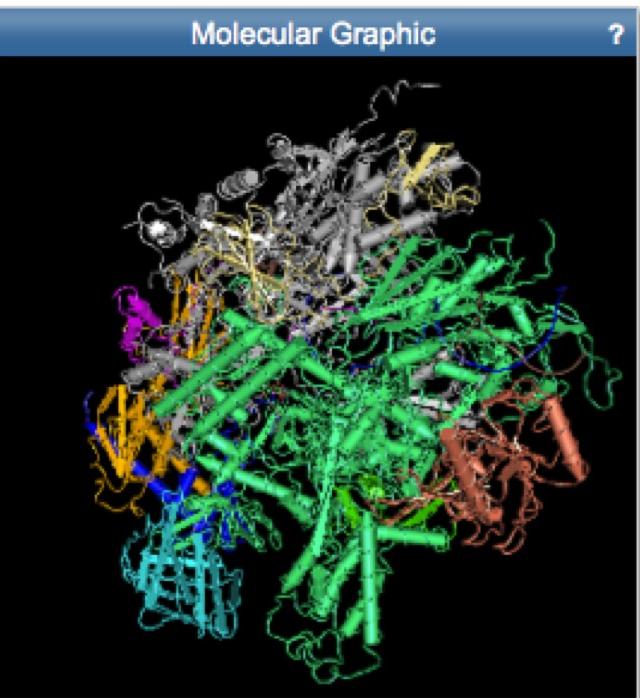
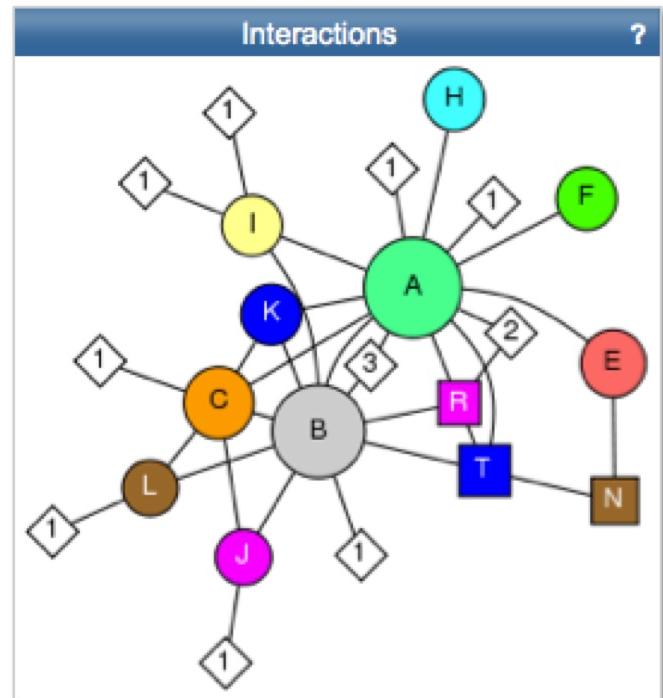
Default Biological Unit

All Biological Unit (1)

Asymmetric Unit

[?](#)

Biological Unit: tridecameric; determined by author [?](#)



View or Save 3D Structure [?](#)

File Format: Cn3D [▼](#)
 Display As: 3D structure [▼](#)
 Data Set: alpha-Carbons [▼](#)

NOTICE
 In order to view this biological unit properly, please upgrade to Cn3D 4.3.

Nucleic Acids Database (NDB)

- <http://ndbserver.rutgers.edu/>
- # released structures: 11047 (~10x less than in PDB)
- A subset of PDB entries that contain nucleic acids
- Special tools for analysis of these structures

Featured Tools

RNA 3D Motif Atlas, a representative collection of RNA 3D internal and hairpin loop motifs

Non-redundant Lists of RNA-containing 3D structures

RNA Base Triple Atlas, a collection of motifs consisting of two RNA basepairs

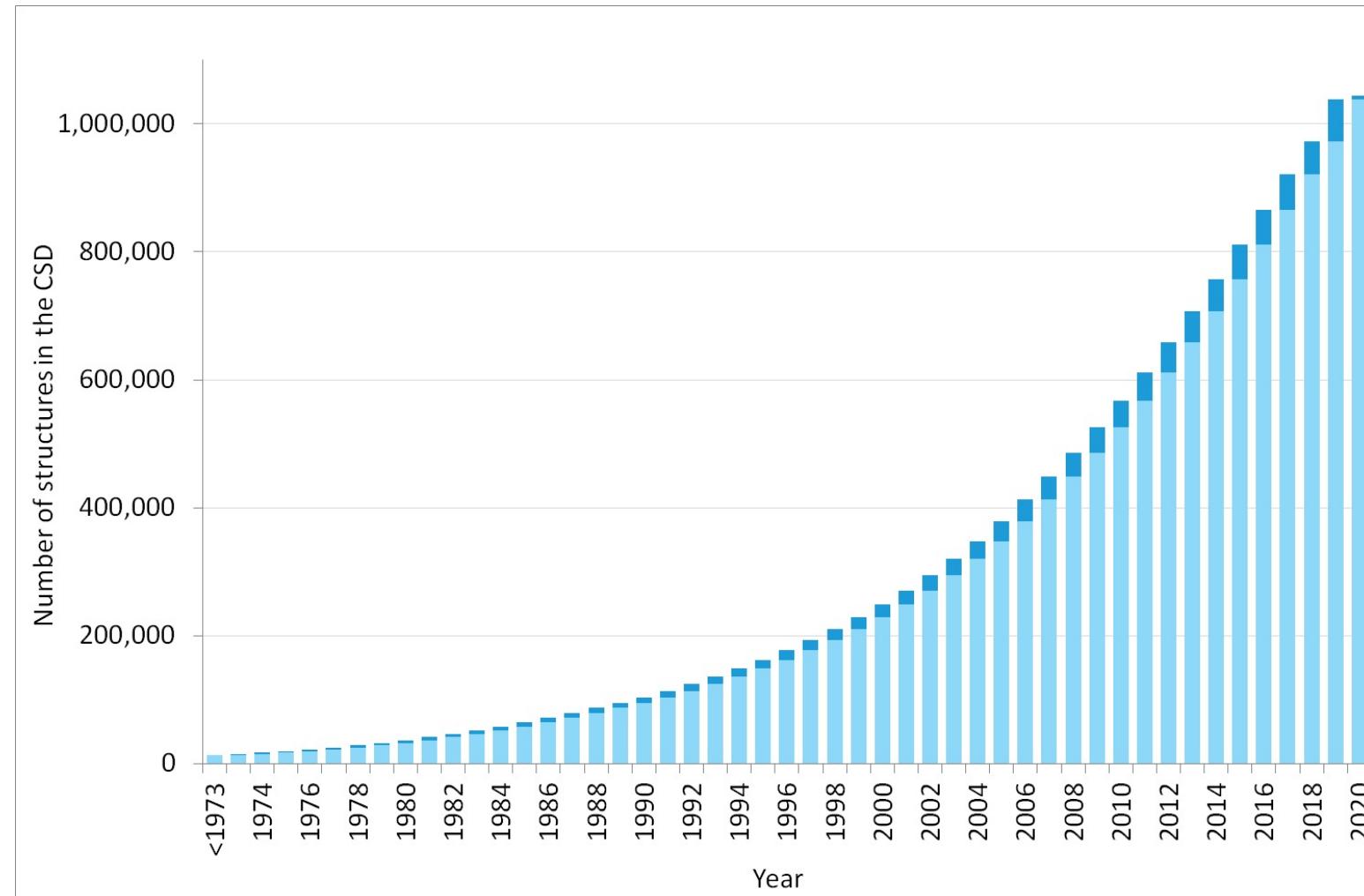
WebFR3D, a webserver for symbolic and geometric searching of RNA 3D structures

R3D Align, an application for detailed nucleotide to nucleotide alignments of RNA 3D structures



Cambridge Structural Database: small molecules

- Major repository for small-molecule crystal structures
- <https://www.ccdc.cam.ac.uk/solutions/csd>
 - [system/components/csd/](https://www.ccdc.cam.ac.uk/system/components/csd/)
- Unfortunately, proprietary



Summary and potential exam questions

- Three major techniques for determination of 3D structure of biomolecules
- Rule-of-thumb: what is a good resolution? how to detect good R-value and R-free value?
- Difference in the dynamics of the molecules in X-ray crystallography and NMR spectroscopy experiments
- Typical size of molecules amenable to X-ray, NMR, and EM
- Typical resolution range for X-ray, NMR, and EM
- How peptides are identified in MS?
- Different types of information available in PDB
- Other databases containing structural data