www.scienceinschool.org

On the trail of viruses: understanding virus research

Discussion answers

Activity 1: Short film about virus research

Learning objectives

After watching the film and discussing these questions, students should be able to:

- explain how physics techniques like X-ray crystallography are used to study virus structures;
- justify why international research institutions like the European XFEL are necessary;
- give examples of how structural biology has contributed to vaccines and antiviral drugs.

a) What surprises you about the connection between physics and biology?

Modern biology cannot operate without the tools of physics and engineering. The European XFEL generates X-ray flashes up to a billion times more intense than conventional X-ray sources, firing up to 27 000 pulses per second. These ultra-short, high-intensity pulses enable scientists to capture diffraction patterns of individual virus particles before radiation damage occurs — a principle called diffraction-before-destruction.[1,2]

Key points for class discussion:

- Physical instruments (particle accelerators, undulator magnets, superconducting technology) directly enable breakthroughs in structural biology.
- X-ray crystallography and single-particle imaging reveal the 3D architecture of viruses at atomic or near-atomic resolution.
- Without knowing the precise shape of a virus's surface proteins, it is impossible to design targeted vaccines or antiviral drugs.
- The COVID-19 pandemic demonstrated this directly: rapid structural elucidation of the SARS-CoV-2 spike protein enabled mRNA vaccine development within months.[3,4]

TEACHER NOTE

Students are often surprised that a 3.4 km underground tunnel filled with particle physics equipment is used to study something as biological as a virus. Emphasise that the size mismatch between the machine and the object it studies is precisely what makes it so powerful.

b) Why does modern biology need international research institutions?

Research facilities like the European XFEL are too large, too expensive, and too technically complex for any single university, company or country to build and operate alone.

The key arguments:

- Enormous investment required: The European XFEL cost approximately €1.25 billion to construct, spanning 3.4 km of underground tunnels across two German federal states. Such investment can only be justified through international cost-sharing.[5,2]
- 12 participating countries: The facility is operated by a non-profit GmbH with 12 shareholder countries. This distributes both financial burden and scientific benefit.
- Open access model: Any research group worldwide can apply for experimental beam time, democratising access to world-class infrastructure.
- Scientific critical mass: Modern structural biology requires simultaneous expertise in accelerator physics, X-ray optics, sample preparation, bioinformatics and medical translation – no single national group commands all these specialties.
- Pandemics do not respect borders: Research infrastructure that is internationally funded and accessible is better aligned with the global nature of viral outbreaks.

TEACHER NOTE

This question connects naturally to broader discussions about the European Research Area (ERA), the role of EU funding and why nations that might be geopolitical competitors still cooperate on science.

c) Could such methods also help with other viruses?

Yes. structural methods like those used at the European XFEL have already been applied across a wide range of medically important viruses. The techniques are broadly applicable because viruses share common structural features (capsids, envelope proteins, receptor-binding domains), that can be targeted once their 3D shape is known.

Examples:

- SARS-CoV-2: Cryo-EM structures of the spike protein were resolved within weeks of the start of the pandemic, directly enabling mRNA vaccine design. The antiviral drug Paxlovid targets the Mpro protease, the structure of which was determined by X-ray crystallography.[3,6,7]
- Influenza: Structural studies of hemagglutinin informed strategies for a universal flu vaccine. Oseltamivir (Tamiflu) targets neuraminidase, whose crystal structure guided its design.[8,9]
- HIV: Decades of X-ray crystallography and cryo-EM of the HIV capsid led to drugs like Lenacapivir (FDA-approved 2022).[7]

- Zika, Dengue, Ebola: Cryo-EM has been used to determine surface protein structures, which has guided vaccine candidate design.[10]
- Tick-borne encephalitis virus (TBEV): Single-particle imaging at the XFEL has been specifically simulated for TBEV — a virus endemic in Germany and Northern Europe, directly relevant to students.[11]

TEACHER NOTE

This is a strong opportunity to connect abstract physics-biology concepts to everyday public health. Students may have been vaccinated against influenza or tick-borne encephalitis — structural biology has contributed directly or indirectly to the development of those vaccines.

Extension idea: Ask students which virus they would most like scientists to solve next and why. This brings together structural biology, public health and personal relevance.

Activity 2: DNA sequence analysis

a) Why do certain mutations appear in multiple variants?

Certain mutations appear independently in multiple variants due to a process called convergent evolution, whereby different viral lineages acquire the same beneficial mutations in response to identical selective pressures. During the first two years of the COVID-19 pandemic, recurrent mutations emerged at specific sites (K417, L452, E484, N501, and P681) across Alpha, Beta, Gamma and Delta variants.

These homoplasious mutations represent the virus's evolutionary response to natural selection. A mutation that significantly improves viral fitness, such as enhanced binding to human ACE2 or improved evasion of antibodies, will be selected for independently in different viral populations. This demonstrates that viral evolution is not random but it follows predictable pathways where the same genetic solutions solve the same survival challenges.[12,13]

TEACHER NOTE

Students may initially assume that mutations are random and unpredictable. Use this question to introduce the concept that natural selection filters out mutations, allowing only the beneficial ones to spread.

Advanced extension: Ask the students why not all mutations at these sites are identical across the variants (e.g., P681R in Delta vs P681H in Omicron), and what this tells us about the range of beneficial changes at a given site.

b) How does selection pressure drive mutation accumulation?

Selection pressure drives mutation accumulation through differential survival and replication. Primary selective pressures acting on SARS-CoV-2 are vaccine-elicited immunity, infection-induced antibodies, and population-level immunity. In this selective environment, mutations that provide advantages, such as enhanced transmissibility, improved receptor binding or antibody evasion, become enriched because viruses carrying them replicate more successfully.

During the pandemic, vaccination coverage was inversely correlated with mutation rates: higher vaccination coverage exerted stronger selective pressure, driving faster accumulation of immune-escape mutations. The Omicron variant is a prime example of this process, having acquired numerous mutations in the receptor-binding domain that enhance ACE2 binding and evading pre-existing antibody responses.[13-15]

TEACHER NOTE

This is an excellent opportunity to discuss the evolutionary arms race between viruses and the immune system. Students who understand this concept will better appreciate why annual flu vaccines are needed and why predicting future variants is scientifically challenging.

c) What role do amino acid properties play in protein function?

Amino acid properties determine how proteins interact with their environment and binding partners. Key properties are:

- Hydrophobic residues (e.g., carbohydrate side chains or aromatic residues): avoid water, form stable internal interactions and can create tight binding pockets
- Hydrophilic/polar residues: interact well with water and charged partners
- Charged (positively or negatively) residues: form salt bridges and ionic interactions

For example, the N501Y mutation in the spike protein replaces polar asparagine (hydrophilic) with tyrosine (aromatic, hydrophobic). The hydrophobic benzene ring of Y501 forms favourable interactions with hydrophobic residues (Y41 and K353) in the receptor ACE2, which mitigates virus transmission. The original polar N501 could not form efficiently these interactions and the improved binding increases transmissibility. However, the same change reduces affinity to the ACE2 receptor in cats due to a steric clash, demonstrating that a single amino acid change can be beneficial in one species and detrimental in another.[16,17]

TEACHER NOTE

If amino acid structures have not yet been covered, use the codon table to show how a single nucleotide change (A to T at position 501 of the spike gene) causes the N501Y substitution, connecting the DNA level to the protein level.

d) Why might some mutations increase viral transmission?

Mutations increase viral transmission through multiple mechanisms. Here are two critical examples illustrating these mechanisms:

N501Y – Enhanced receptor binding

As mentioned above the N501Y mutation increases transmissibility by enhancing binding between the spike protein's receptor-binding domain and the ACE2 receptor on the host cell. Tighter binding allows the virus to attach more efficiently to target cells, increasing the probability of successful cell entry. Fewer viral particles are needed to infect a cell, and infected individuals shed more efficiently transmissible virus.[16]

P681R/P681H – furin cleavage site

These mutations occur at the furin cleavage site (S1/S2 junction), where the spike protein must be cleaved by cellular proteases for optimal viral entry into the cell. P681R (Delta variant) increases furin protease substrate turnover, enhancing cleavage efficiency and allowing a more effective cell entry. P681H (Omicron variant) reduces cleavage efficiency by furin but creates a novel cleavage site for cathepsin G, representing an alternative entry pathway. These different cleavage properties affect cellular tropism: Delta with P681R shows enhanced replication in the lower respiratory tract, while Omicron with P681H preferentially enters cells through endocytosis and shows greater upper respiratory tract tropism.[18-20]

TEACHER NOTE

This helps students understand that the same site can be mutated in different ways with different consequences.

Clinical connection: Omicron's upper airway tropism explains both its higher transmissibility (easier airborne spread) and its generally lower severity (less lung involvement) compared to Delta.

Extended Discussion: The Δ69-70 deletion

A vase study in mutation detection

Students may investigate the Δ69-70 deletion present in Alpha and Omicron variants. This deletion of two amino acids (H69 and V70) in the spike protein N-terminal domain disrupts the target site for one of the three PCR primer-probe sets used in the TaqPath SARS-CoV-2 real-time (RT)-PCR test, causing Spike-gene target failure (SGTF). This means that the spike gene cannot be detected by this test anymore, even though other virus genes can still be detected.

During the pandemic, SGTF was used as a rapid surveillance marker for variant identification with sensitivity and specificity exceeding 99% when validated by next-generation sequencing. A sensitivity >99% means that the test misses less than 1% of variants that have the mutation (false negatives), while >99% specificity means that the test incorrectly labels 1% of variant who do not have the mutation (false positives). This enabled presumptive variant identification within 24–48 hours, long before confirmatory sequencing results were available. This real-world example shows how understanding sequence mutations connects directly to clinical diagnostics and public health surveillance.[21]

TEACHER NOTE

This is one of the most powerful real-world examples in the entire activity series: a deletion of just two amino acids in the spike protein inadvertently broke one of the most widely used clinical PCR tests, and that broken test became a surveillance tool.

It illustrates that mutations can affect how we detect and track a virus in the population.

Activity 3: DNA sequence analysis

a) How does GC content affect primer stability and why does proper temperature selection prevent non-specific binding?

G-C base pairs are more stable as they form 3 hydrogen bonds while A-T pairs form only 2. Therefore, a higher G/C content increases the stability of the primer template duplex and thereby the melting temperature (T_m). Consequently, a G/C content within the range of 40–60% is beneficial because primers bind more tightly to the template DNA.[22]

A too low annealing temperature (T_a) allows the primer to bind non-specifically. At the perfect temperature the most stable formations (optimal base pair matching between primer and template) are predominant as the less stable formations with mismatches are broken up by thermal energy. A T_a that is too high (above T_m) prevents annealing altogether, yielding no PCR product. In practice, a gradient PCR is run (T_a ±5°C) to empirically verify the best annealing temperature.[23]

References

[1] Meents A, Wiedorn MO (2019) Virus Structures by X-Ray Free-Electron Lasers. *Annual Review of Virology* **6**: 161–176. doi: 10.1146/annurev-virology-092818-015724

[2] European XFEL: https://en.wikipedia.org/wiki/European_XFEL

[3] Kim H, Jung HS (2021). Cryo-EM as a powerful tool for drug discovery: SARS-CoV-2. *Applied Microscopy* **51**: 13. doi: 10.1186/s42649-021-00062-x

[4] European XFEL open for COVID-19 related research: https://www.xfel.eu/news_and_events/news/index_eng.html?openDirectAnchor=1777

[5] About the European XFEL: https://www.eiroforum.org/about-eiroforum/members/european-xfel/

[6] Asarnow D et al. (2024) Recent advances in infectious disease research using cryo-electron tomography. *Frontiers Molecular Bioscience* **10**: 1296941. doi: 10.3389/fmolb.2023.1296941

[7] Dutta M, Acharya P (2024) Cryo-electron microscopy in the study of virus entry. *Frontiers Molecular Bioscience* **11**: 1429180. doi: 10.3389/fmolb.2024.1429180

[8] Wu NC, Wilson IA (2020) Structural biology of influenza hemagglutinin: an amaranthine adventure. *Viruses* **12**: 1053. doi: 10.3390/v12091053

[9] Von Itzstein M (2007) The war against influenza: discovery and development of sialidase inhibitors. *Nature Reviews Drug Discovery* **6**: 967–974. doi: 10.1038/nrd2400

[10] Sirohi D et al. (2016) The 3.8 Å resolution cryo-EM structure of Zika virus. *Science* **352**: 467-470. doi: 10.1126/science.aaf5316

[11] Assalauova D, Vartanyants IA (2023) Structure of TBEV at X-ray free-electron lasers. *J Synchrotron Radiation* **30**: 24–34. doi: 10.1107/S1600577522011341

[12] Focosi D et al. (2023) Convergent evolution in SARS-CoV-2 Spike. *International Journal of Molecular Science* **23**: 2264. doi: 10.3390/ijms24032264

[13] Carabelli AM et al. (2023) SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nature Reviews Microbiology* **21**: 162–177. doi: 10.1038/s41579-022-00841-7

[14] Bloom JD, Neher RA (2023) Fitness effects of mutations on the spike protein of SARS-CoV-2 Omicron. *Virus Evolution* **9**: veae026. doi: 10.1093/ve/veae026

[15] Garg KM, Lamba V, Chattopadhya B (2023) Genomic Insights Into the Evolution and Demographic History of the SARS-CoV-2 Omicron Variant: Population Genomics Approach. *JMIR Bioinformatics and Biotechnology* **4**: 2563–3570. doi: 10.2196/40673

[16] Luan B, Wang H, Huynh T (2021) Enhanced binding of the N501Y-mutated SARS-CoV-2 spike protein to the human ACE2 receptor: insights from molecular dynamics simulations. *FBS Letters* **595**: 1454–1461. doi: 10.1002/1873-3468.14076

[17] Stojanov D (2023) Structural implications of SARS-CoV-2 Surface Glycoprotein N501Y mutation within receptor-binding domain [499-505] – computational analysis of the most frequent Asn501 polar uncharged amino acid mutations. *Biotechnology & Biotechnological Equipment* **37**. doi: 10.1080/13102818.2023.2206492

[18] Lubinski B et al. (2022) Spike Protein Cleavage-Activation in the Context of the SARS-CoV-2 P681R Mutation: an Analysis from Its First Appearance in Lineage A.23.1 Identified in Uganda. *Microbiology Spectrum* **10**: e01514-22. doi: 10.1128/spectrum.01514-22

[19] Mannar D et al. (2022). SARS-CoV-2 Omicron variant: Antibody evasion and cryo-EM structure of spike protein-ACE2 complex. *Science* **375**: 760-764. doi: 10.1126/science.abn7760

[20] Peacock TP et al. (2021) The furin cleavage site in SARS-CoV-2 spike protein. *Nature Microbiology* **6**: 899–909. doi: 10.1038/s41564-021-00908-w

[21] Kidd M et al. (2022) The spike gene target failure (SGTF) genomic signature is highly accurate for the identification of Alpha and Omicron SARS-CoV-2 variants. *Scientific Reports* **12**: 18641. doi: 10.1038/s41598-022-21564-y

[22] Marmur J, Doty P (1962) Determination of the base composition of DNA from its thermal denaturation temperature. *Journal of Molecular Biology* **5**:109–118. doi: 10.1016/S0022-2836(62)80066-7

[23] Guide to primer design for PCR: https://www.benchling.com/primer-design-for-pcr

[24] Keusch GT et al. (2022) Pandemic origins and a One Health approach to preparedness and prevention: Solutions based on SARS-CoV-2 and other RNA viruses. *PNAS* **119**: e2202871119. doi: 10.1073/pnas.2202871119

[25] Hagan MF, Zandi R, Uetrecht C (2024). Overview of the 2023 Physical Virology Gordon Research Conference—Viruses at Multiple Levels of Complexity. *Viruses* **16**: 895. doi: 10.3390/v16060895

[26] Fabius JM, Krogan NJ (2023). Lessons learned for pandemic preparedness: A collaborative network is imperative. *Cell Host & Microbe* **31**: 843-846. doi: 10.1016/j.chom.2023.05.008

[27] Moradian N et al. (2020) The urgent need for integrated science to fight COVID-19 pandemic and beyond. *Journal of Translational Medicine* **18**: 205. doi: 10.1186/s12967-020-02364-2