

Task 1

- a) Take a closer look at the sequences arranged one below the other. Describe which codon they each begin with and explain the significance of this codon for translation.

- b) Now scroll slowly to the right and take a closer look at the alignment. Describe in bullet points what you notice and explain why this might be the case.

Carrying out the sequence alignment

The complete sequence alignment is carried out as follows. To do this, select the option „Align by MUSCLE (Codons)“ in the “Alignment” menu (see Figure 3).

Confirm the subsequent question "Nothing selected for alignment. Select all?" with "OK". Another window opens (see Figure 4) in which various parameters can be changed. The preset parameters are useful for our project, so you can confirm them with "OK". When the programme then asks "Would you like to remove gaps before alignment?", click on "Yes". Finally, select "w/o gaps" directly in the alignment window (see Figure 5).

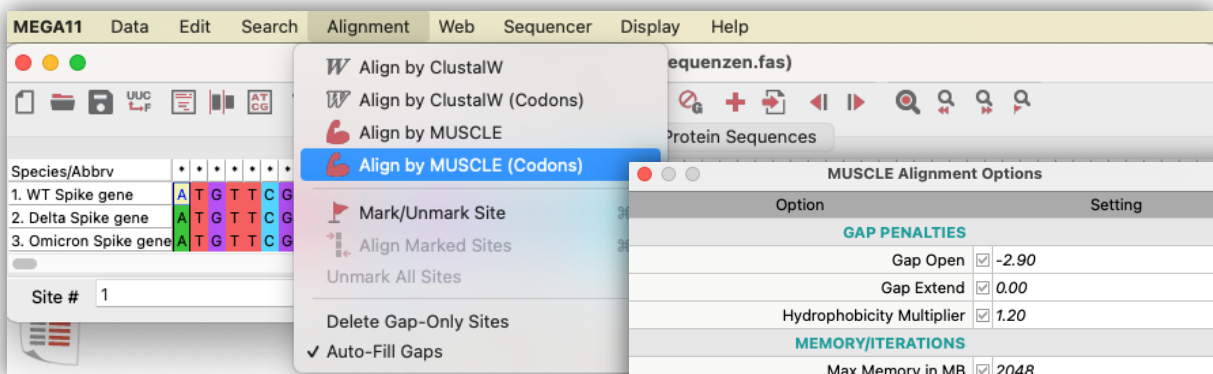


Figure 3

Figure 4

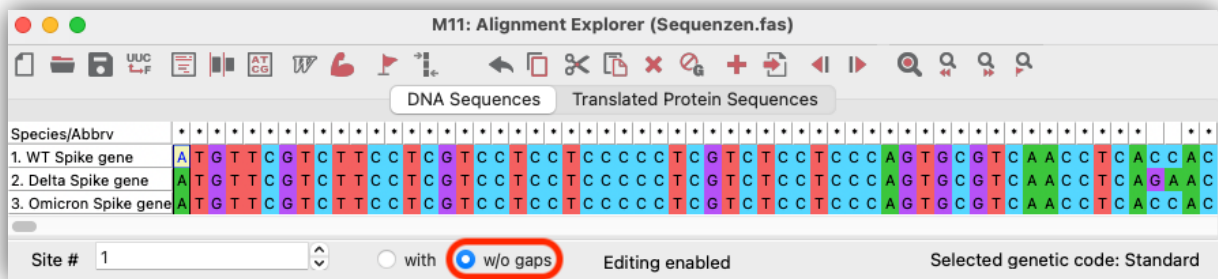


Figure 5

Task 2

- a) Now scroll slowly through the sequences again and look at the alignment. Explain what has changed.

- b) Give an example of each of the three different types of point mutation (substitution, deletion and insertion) that can be found in these sequences. Use the format C56G (cysteine at position 56 is changed to glycine).
NOTE: Specify the position in relation to the *wild type* sequence!

Variant	Mutation		
	Substitution	Deletion	Insertion
Delta			
Omikron			



Investigation of mutations

In the following, you will take a closer look at four different mutations. All numbers used in the following refer to the *wild type*.

Task 3

Examine the sequence alignment sites listed under a) - d) in more detail and name which variant is mutated and how. Describe the variant, the type of point mutation, the corresponding nucleotides and the resulting amino acids. Use the code column at the end of the worksheet for the latter.

HELP: Since 3 nucleotide bases together always make a codon, it is important to pay attention to which position you are looking at. If the number (site #) is dividable by three, this means that it is the third base in the codon. Site # $57 / 3 = 19$. In this case, position 57 is the third base. In order to translate the correct amino acid, the bases 55, 56 and 57 must be selected, i.e. ACC.

a) 205-210

b) 642/643

c) 1501-1503

d) 2041-2043

NOTE: You can navigate to a corresponding site by entering the corresponding number under "Site #" at the bottom left of the MEGA Alignment Explorer and pressing Enter. You may have to mark the current number and then overwrite it, if the delete key on the keyboard does not work.

Now compare your results regarding the amino acids with the programme by clicking on "Translated Protein Sequences" (see Figure 6). Answer the question "The current Genetic Code is: Standard. Is this correct?" with "Yes".

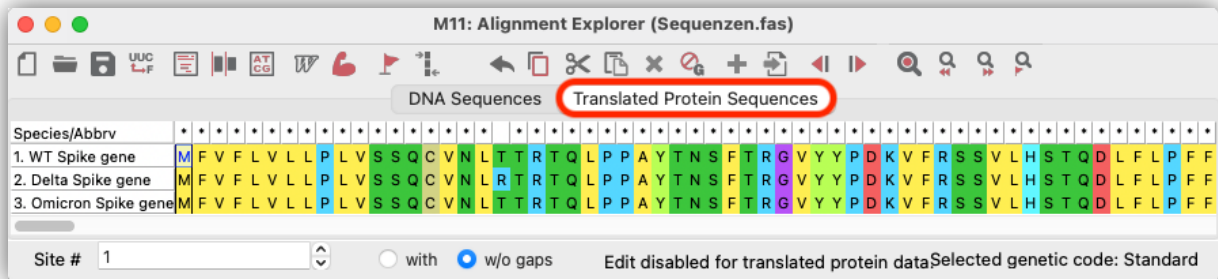


Figure 6

Now you can check for yourself whether you have translated the code correctly according to your results from task 3 a) - d).

Task 4

- a) In frequently used medical PCR tests, several sites in the genetic material of the virus are compared. The mutation Δ69-70 (in the protein sequence) falls within one of the binding sites for PCR primers. Explain how the PCR results of someone infected with the omicron variant differ from someone else infected with the delta variant.

- b) Analyse the side chains of the two amino acids in the mutation N501Y. Describe the chemical properties and differences.



- c) Examine the side chains of the amino acids of P681R at *Delta* and of P681H at *Omikron*. This site is close to a cleavage site that is important when the virus enters the host cell. Describe the chemical properties and differences. Discuss whether, and if so, what effects are to be expected for the infection

Additional task

Compared to viruses, higher animals have a much larger genome. Some genes are present more than once (through duplication or diploidy).

- a) Hypothesise what the multiple presence of genes may mean for the survival chances of an individual with a mutation.

- b) Explain whether a difference in the frequency of mutations (surviving to the next generation) is to be expected between viruses and higher animals such as humans.



Homework

Task 1

- a) Find out about the topic of artificially mutated viruses and draw up a list of pros and cons for a discussion group. Make a note of the sources you use.

- b) Find out about the advantages and disadvantages of experimenting with artificially modified and thus mutated viruses. Choose a side of the argument and prepare to present and defend it in a discussion round. Make a note of the sources you use.

Task 2

Find out more on the internet about PCR tests and their use to detect a corona infection.

- a) Briefly summarise the procedure of a PCR test.

- b) Explain why several genes are taken into account in the PCR test.

Coding circle

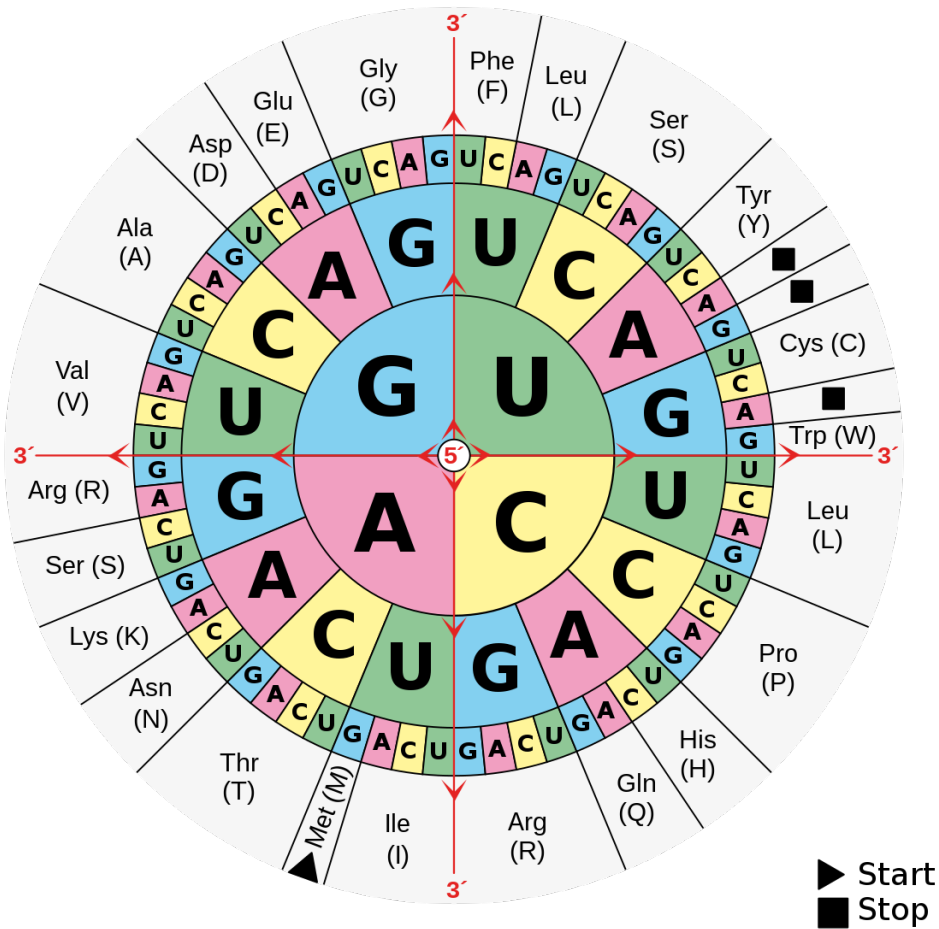


Figure 7: Coding circle
 Image: Mouagip/[Wikimedia Common](#); [public domain](#)

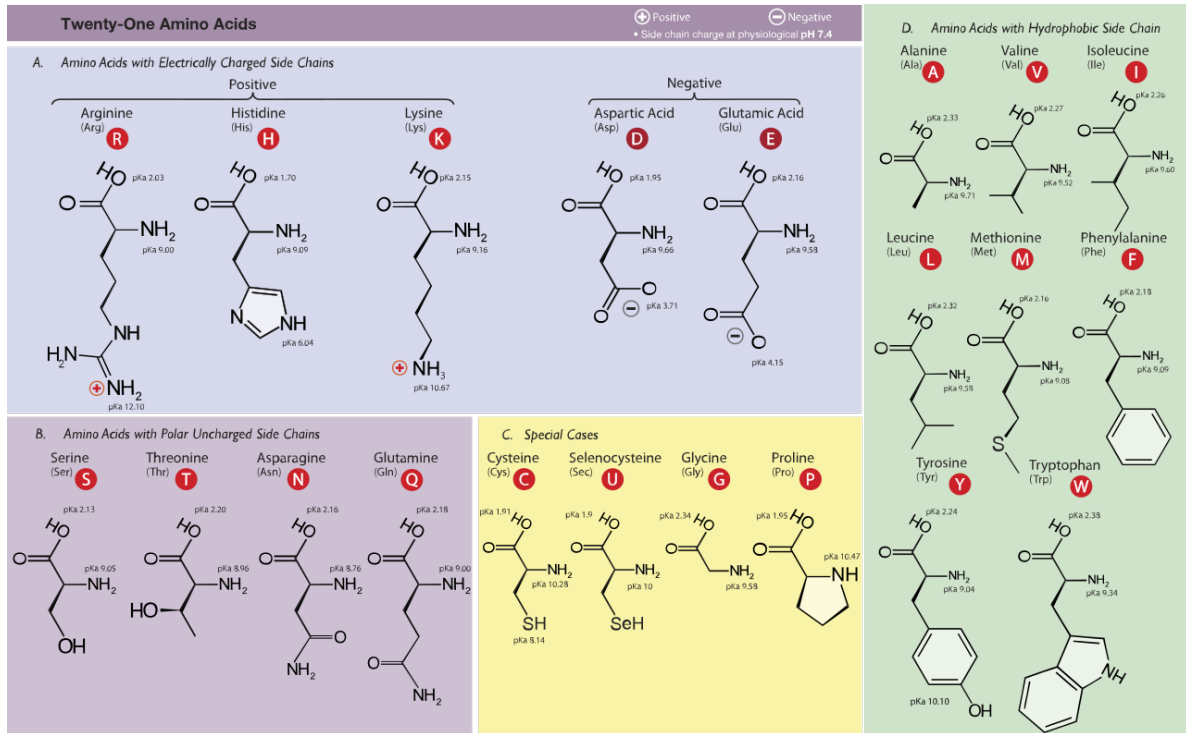


Figure 8: Overview of amino acids
Image: Bernd Huber/Wikimedia Commons; CC BY-SA 3.0

Sequences

The sequences must be copied and pasted into a text editor. On a Windows computer, you can use the Editor or Notepad++ programme for this. On an Apple computer, you can use the TextEdit programme. Now you should save the file as sequenzen.fas (figure 9). The file extension is **IMPORTANT** here. So make sure to delete the default ending *.txt and use *.fas instead.

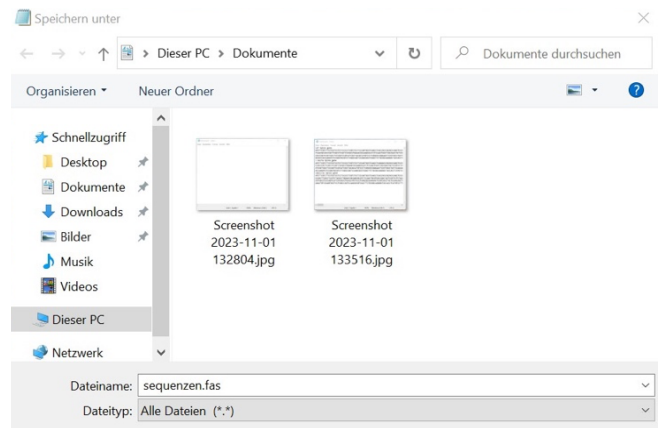


Figure 9

The rest of the document must be copied as described above:

>WT Spike gene

```
ATGTTTCGTCTTCCTCGTCCTCCTCCCCCTCGTCTCCTCCCAGTGGCTCAACCTCACCACCCGCACCCAG
CTCCCCCCCCGCCTACACCAACTCCTCACCCGCGGCTCTACTACCCCGACAAGGTCTCCGCTCCTC
CGTCTCCACTCCACCCAGGACCTCTTCTCCTCCCTTCTTCTCCAACGTACCTGGTCCACGCCATCCA
CGTCTCCGGCACCAACGGCACCAAGCGCTTCGACAACCCCGTCTCCTCCCCTTCAACGACGGCGTCTAC
```



TTCGCCTCCACCGAGAAGTCCAACATCATCCGCGGCTGGATCTTCGGCACCACCCTCGACTCCAAGA
 CCCAGTCCCTCCTCATCGTCAACAACGCCACCAACGTCGTCATCAAGGTCTGCGAGTTCAGTTCTGC
 AACGACCCCTTCTCGGGCTACTACCACAAGAACAACAAGTCTGGATGGAGTCCGAGTTCGCG
 TCTACTCCTCCGCCAACAACTGCACCTTCGAGTACGTCTCCCAGCCCTTCTCATGGACCTCGAGGGC
 AAGCAGGGCAACTTCAAGAACCTCCGCGAGTTCGTCTTCAAGAACATCGACGGTACTTCAAGATCTA
 CTCCAAGCACACCCCATCAACCTCGTCCGCGACCTCCCCAGGGCTTCTCCGCCCTCGAGCCCCTC
 GTCGACCTCCCATCGGCATCAACATCACCCGCTTCAGACCCTCCTCGCCCTCCACCGCTCCTACCT
 CACCCCGGCGACTCCTCCTCCGGTGGACCGCCGGCGCCGCCCTACTACGTGGTACCTCCA
 GCCCCGACCTTCTCCTCAAGTACAACGAGAACGGCACCATCACCGACGCGTCGACTGCGCCCT
 CGACCCCTTCTCCGAGACCAAGTGCACCCTCAAGTCTTACCGTCGAGAAGGGCATCTACCAGACC
 TCCAATTCCGCGTCCAGCCCACCGAGTCCATCGTCCGCTTCCCCAACATCACCAACCTCTGCCCTT
 CGGCGAGGTCTTCAACGCCACCCGCTTCGCTCCTGCTACGCCTGGAACCGCAAGCGCATCTCCAAC
 TGCGTCGCGGACTACTCGTCTCTACAACCTCCGCTCCTTCTCCACCTTCAAGTGCTACGGCGTCTCC
 CCCACCAAGCTCAACGACCTCTGCTTACCAACGTCTACGCCGACTCCTTCGTCATCCGCGGCGACG
 AGGTCCGCCAGATCGCCCCGGCCAGACCCGCAAGATCGCCGACTACAACCTACAAGCTCCCCGAC
 GACTTACCCGGTTCGTCATCGCCTGGAACCTCAACAACCTCGACTCCAAGGTTCGGCGGCAACTACA
 ACTACCTTACCCGCTTCTCCGCAAGTCCAACCTCAAGCCCTTCGAGCGGACATCTCCACCGAGATC
 TACCAGGCCGGTCCACCCCTGCAACGGCGTTCGAGGGCTTCAACTGCTACTTCCCCCTCCAGTCTT
 ACGGCTTCCAGCCCACCAACGGCGTTCGGTACCAGCCCTACCGCGTCGTCGCTCCTCTCCTTCGAGCT
 CCTCCACGCCCCCGCCACCGTCTGCGGCCCAAGAAGTCCACCAACCTCGTCAAGAACAAGTGCGT
 CAACTTCAACTTCAACGGCCTCACCGGCACCGGCTCCTCACCGAGTCCAACAAGAAGTTCCTCCCT
 TCCAGCAGTTCGGCCGCGACATCGCCGACACCACCGACGCGTCCGCGACCCCCAGACCCTCGAG
 ATCCTCGACATCACCCCTGCTCCTTCGGCGGGCTTCCGTCATCACCCCGGCACCAACACCTCCAA
 CCAGGTTCGCGTCTCTACCAGGGTGTCAACTGCACCGAGGTCCCCGTCGCCATCCACGCCGACCA
 GCTCACCCCCACCTGGCGGTCTACTCCACCGGCTCCAACGTCTTCCAGACCCGCGCCGGTGCCTC
 ATCGGCGCCGAGCACGTCAACAACCTCCTACGAGTTCGACATCCCCATCGGCGCCGGCATCTGCGCC
 TCCTACCAGACCCAGACCAACTCCCCCGCCGCGCCCGTCCGTCGCTCCAGTCCATCATCGCT
 ACACCATGTCCCTCGGCGCCGAGAACTCCGTCGCTACTCCAACAACCTCCATCGCCATCCCCACCA
 CTTACCATCTCCGTACCACCGAGATCCTCCCCGTCTCATGACCAAGACCTCCGTTCGACTGCACCA
 TGACATCTGCGGCGACTCCACCGAGTGTCCAACCTCCTCCTCCAGTACGGTCTCTTCTGCACCCAG
 CTCAACCGCGCCCTCACCGGCATCGCCGTCGAGCAGGACAAGAACACCCAGGAGGTCTTCGCCCA
 GGTC AAGCAGATCTACAAGACCCCCCATCAAGGACTTCGGCGGCTTCAACTTCTCCCAGATCCTCC
 CCGACCCCTCCAAGCCCTCCAAGCGCTCCTTATCGAGGACCTCCTTCAACAAGGTACCCTCGCC
 GACGCCGGTTCATCAAGCAGTACGGCGACTGCCTCGGCGACATCGCCGCCGCGACCTCATCTGC
 GCCAGAAGTTCAACGGCCTCACCGTCTCCCCCTCCTCACCGACGAGATGATCGCCAGTACA
 CCTCCGCCCTCCTCGCCGGCACCATCACCTCCGGTGGACCTTCGGCGCCGGCGCCGCCCTCCAGA
 TCCCCTTCGCCATGCAGATGGCCTACCGCTTCAACGGCATCGGCGTCACCCAGAACGTCTCTACGA
 GAACCAGAAGTTCATCGCCAACCAGTTCAACTCCGCCATCGGCAAGATCCAGGACTCCCTCTCCTCC
 ACCGCCTCCGCCCTCGGCAAGCTCCAGGACGTGTC AACCAGAACGCCACGGCCCTCAACACCTC
 GTCAAGCAGTCTCCTCAACTTCGGCGCCATCTCCTCCGTCTCAACGACATCCTTCCCGCTCGA
 CAAGGTTCGAGGCCGAGGTCCAGATCGACCGCTCATCACCGGCCGCTCCAGTCCCTCCAGACCTA
 CGTACCCAGCAGTTCATCCGCGCCGCGAGATCCGCGCTCCGCCAACCTCGCCGCCACCAAGAT
 GTCCGAGTTCGCTCCTCGGCCAGTCCAAGCGGTCGACTTCTGCGGCAAGGGTACCACCTCATGTCC
 TCCCCAGTCCGCCCCCACGGCGTCTTCTCCTCACGTACCTACGTCCCCGCCAGGAGAAGA
 ACTTACCCACCGCCCCCGCCATCTGCCACGACGGCAAGGCCACTTCCCCCGGAGGGCGTCTTCG
 TCTCCAACGGCACCCACTGGTTCGTACCCAGCGCAACTTCTACGAGCCCCAGATCATCACCCCGA
 CAACACCTTCGTCTCCGGCAACTGCGACGTTCGTCATCGGCATCGTCAACAACACCGTCTACGACCC



CTCCAGCCCCGAGCTCGACTCCTTCAAGGAGGAGCTCGACAAGTACTTCAAGAACCACACCTCCCCCG
 ACGTCGACCTCGGGACATCTCCGGCATCAACGCCTCCGTCGTCAACATCCAGAAGGAGATCGACC
 GCCTCAACGAGGTGCGCAAGAACCTCAACGAGTCCCTCATCGACCTCCAGGAGCTCGGCAAGTACG
 AGCAGTACATCAAGTGGCCCTGGTACATCTGGCTCGGCTTCATCGCCGGCCTCATCGCCATCGTCATG
 GTCACCATCATGCTCTGCTGCATGACCTCCTGCTGCTCCTGCCTCAAGGGCTGCTGCTCCTGCGGCTC
 CTGCTGCAAGTTCGACGAGGACGACTCCGAGCCCCGTCTCAAGGGCGTCAAGTCCACTACACCTAA

> Delta Spike gene

ATGTTTCGTCTTCTCGTCTCCTCCCCCTCGTCTCCTCCCAGTGGTCAACCTCAGAACCCGACCCAG
 CTCCCCCCCCGCTACACCAACTCCTTACCCGCGGGCTACTACCCGACAAGGTCTTCCGCTCCTC
 CGTCTCCACTCCACCCAGGACCTTTCCTCCCCTTCTTCTCCAACGTACCTGGTTCACGCCATCCA
 CGTCTCCGGCACCAACGGCACCAAGCGCTTCGACAACCCCGTCTCCCCTTCAACGACGGCGTCTAC
 TTCGCCTCCACCGAGAAGTCCAACATCATCCGCGGTGGATCTTCGGCACCACCCTCGACTCCAAGA
 CCCAGTCCCTCCTCATCGTCAACAACGCCACCAACGTGTCATCAAGGTCTGCGAGTTCAGTTCTGC
 AACGACCCCTTCTCGGGCTTACTACCACAAGAACAACAAGTCTGGATGGAGTCCGGCGTCTACT
 CCTCCGCCAACAACTGCACCTTCGAGTACGTCTCCAGCCCTTCTCATGGACCTCGAGGGCAAGCA
 GGGCAACTTCAAGAACCTCCGCGAGTTCGTCTTCAAGAACATCGACGGTACTTCAAGATCTACTCCA
 AGCACACCCCATCAACCTCGTCCGCGACCTCCCCAGGGCTTCTCCGCCCTCGAGCCCCTCGTCGA
 CCTCCCCATCGGCATCAACATCACCCGCTTCAGACCCTCCTCGCCCTCCACCGTCTTACCTCACCC
 CCGGGACTCCTCCTCCGGTGGACCGCCGGCGCCGCCCTACTACGTGGTACCTCCAGCCCC
 GCACCTTCTCCTCAAGTACAACGAGAACGGCACCATCACCGACGCGTGCAGTGCGCCCTCGACC
 CCTCTCCGAGACCAAGTGCACCCTCAAGTCTTACCGTTCGAGAAGGGCATCTACCAGACCTCCAA
 CTTCCGCGTCCAGCCCACCGAGTCCATCGTCCGCTTCCCCAACATCACCAACCTCTGCCCTTCGGCG
 AGGTCTTCAACGCCACCCGCTTCGCCTCCGTCTACGCCTGGAACCGCAAGCGCATCTCCAAGTGCCT
 CGCCGACTACTCCGTCTCTACAACCTCCGCTCCTTCTCCACCTTCAAGTGTACGGCGTCTCCCCAC
 CAAGTCAACGACCTCTGCTTACCAACGTCTACGCCGACTCCTTCGTTCATCCGCGGGCAGCAGGTC
 CGCCAGATCGCCCCGGCCAGACCGGCAAGATCGCCGACTACAACACTACAAGTCCCCGACGACTTC
 ACCGGTGCATCGCTGGAACCTCAACAACCTCGACTCCAAGGTGGCGGCAACTACAACACTACC
 GCTACCGCTTCTCCGCAAGTCCAACCTCAAGCCCTTCGAGCGGACATCTCCACCGAGATCTACCA
 GGCCGGTCCAAGCCCTGCAACGGTGTGAGGGCTTCAACTGCTACTTCCCCCTCCAGTCTACGGC
 TTCCAGCCCACCAACGGCGTCCGCTACCAGCCCTACCGCGTTCGTCTCCTCTCCTTCGAGCTCCTCCA
 CGCCCCCGCCACCGTCTGCGGCCCAAGAAGTCCACCAACCTCGTCAAGAACAAGTGCCTCAACTTC
 AACTTCAACGGCCTCACCGGCACCGGCTCCTCACCGAGTCCAACAAGAAGTTCCTCCCCTTCCAGC
 AGTTCGGCCGCGACATCGCCGACACCACCGACGCGTCCGCGACCCCCAGACCCTCGAGATCTCG
 ACATACCCCCCTGCTCCTTCGGCGGGTCTCCGTTCATACCCCCGGCACCAACACCTCCAACCAGGT
 CGCCGTCTTACCAGGGTGTCAACTGCACCGAGGTCCCCGTGCCATCCACGCCGACCAGTCTCAC
 CCCACCTGGCGGTCTACTCCACCGGTCCAACGTCTTCCAGACCCGCGCCGGTGCCTCATCGGC
 GCGGAGCAGTCAACAACCTTACGAGTGCACATCCCCATCGGCGCCGGCATCTGCGCTCCTAC
 CAGACCCAGACCAACTCCAGACGCCGCGCCCGTCCGTTCGCTCCAGTCCATCATCGCTACACC
 ATGTCCCTCGGCGCCGAGAAGTCCGTCGCCTACTCCAACAACCTCCATCGCCATCCCCACCAACTTCAC
 CATCTCCGTACCACCGAGATCCTCCCCGTCTCCATGACCAAGACCTCCGTGACTGCACCATGTACA
 TCTGCGGGGACTCCACCGAGTGTCCAACCTCCTCCTCCAGTACGGTCTCTTCTGCACCCAGTCAAC
 CGCGCCCTCACCGGCATCGCCGTGAGCAGGACAAGAACACCCAGGAGGTCTTCGCCAGGTCAA
 GCAGATCTACAAGACCCCCCATCAAGGACTTCGGCGGCTTCAACTTCTCCAGATCCTCCCCGACC
 CCTCCAAGCCCTCCAAGCGTCTTTCATCGAGGACCTCCTTCAACAAGGTACCCCTCGCCGACGC
 CGGCTTCATCAAGCAGTACGGCGACTGCCTCGGCGACATCGCCGCCCGGACCTCATCTGCGCCCA
 GAAGTTCACGGCCTCACCGTCTCCCCCCCCCTCCTCACCGACGAGATGATCGCCAGTACACCTCC



GCCCTCCTCGCCGGCACCATCACCTCCGGCTGGACCTTCGGCGCCGGCGCCCTCCAGATCCCC
 TTCGCCATGCAGATGGCCTACCGCTTCAACGGCATCGGCGTCACCCAGAACGTCCTCTACGAGAACC
 AGAAGTCATCGCCAACCAGTTCAACTCCGCCATCGGCAAGATCCAGGACTCCCTCTCCTCCACCGC
 CTCCGCCCTCGGCAAGCTCCAGAACGTCTCAACCAGAACGCCAGGCCCTCAACACCCTCGTCAA
 GCAGCTCTCCTCCAACCTTCGGCGCCATCTCCTCCGTCTCAACGACATCCTCTCCCGCCTCGACAAGG
 TCGAGGCCGAGGTCCAGATCGACCGCCTCATCACCGGCCGCTCCAGTCCCTCCAGACCTACGTCA
 CCCAGCAGTCTATCCGCGCCGCCGAGATCCGCGCCTCCGCCAACCTCGCCGCCACCAAGATGTCCG
 AGTGGCTCTCGGCCAGTCCAAGCGCGTCTGACTTCTGCGGCAAGGGCTACCACCTCATGTCTTCCC
 CCAGTCCGCCCCCACGGCGTCTTCTCCTCCACGTACCTACGTCCCCGCCAGGAGAAGAACTTC
 ACCACCGCCCCGCCATCTGCCACGACGGCAAGGCCACTTCCCCCGGAGGGCGTCTTCGTCTCC
 AACGGCACCCACTGGTTCGTACCCAGCGCAACTTCTACGAGCCCCAGATCATCACCCACCGACAACA
 CCTTCGTCTCCGGCAACTGCGACGTCTCATCGGCATCGTCAACAACACCGTCTACGACCCCTCCA
 GCCCGAGCTCGACTCCTTCAAGGAGGAGCTCGACAAGTACTTCAAGAACCACACCTCCCCGACGTG
 GACCTCGGCCGACATCTCCGGCATCAACGCCTCCGTCTGTAACATCCAGAAGGAGATCGACCGCCTC
 AACGAGGTCGCCAAGAACCTCAACGAGTCCCTCATCGACCTCCAGGAGCTCGGCAAGTACGAGCAG
 TACATCAAGTGGCCCTGGTACATCTGGCTCGGCTTCATCGCCGGCCTCATCGCCATCGTCATGGTCA
 CATCATGCTCTGCTGCATGACCTCCTGCTGCTCCTGCCTCAAGGGTCTGCTGCTCCTGCGGCTCCTGCT
 GCAAGTTCGACGAGGACGACTCCGAGCCCGTCTCAAGGGCGTCAAGCTCCACTACACCTAA

>Omicron Spike gene

ATGTTTCTTCTCCTCGTCTCCTCCCCCTCGTCTCCTCCCAGTGGTCAACCTCACCCCGCACCCAG
 CTCCCCCCCCGCTACACCAACTCCTTACCCCGGGCGTCTACTACCCGACAAGGTCTTCCGCTCCTC
 CGTCTCCTCCACTCCACCCAGGACCTTCTCCTCCCCTTCTTCTCCAACGTACCTGGTTCACGTGATCTCC
 GGCACCAACGGCACCAAGCGCTTCGACAACCCCGTCTCCCCCTTCAACGACGGCGTCTACTTCGCT
 CCATCGAGAAGTCCAACATCATCCGCGGTGGATCTTCGGCACCACCCTCGACTCCAAGACCCAGTC
 CCTCCTCATCGTCAACAACGCCACCAACGTCTCATCAAGGTCTGCGAGTTCAGTTCTGCAACGACC
 CCTCCTCGACCACAAGAACAACAAGTCTGGATGGAGTCCGAGTTCGCGTCTACTCCTCCGCCAAC
 AACTGCACCTTCGAGTACGTCTCCAGCCCTTCTCATGGACCTCGAGGGCAAGCAGGGCAACTTCA
 AGAACCTCCGCGAGTTCGTCTTCAAGAACATCGACGGTACTTCAAGATCTACTCCAAGCACACCCCC
 ATCATCGTCCGCGAGCCCGAGGACCTCCCCAGGGCTTCTCCGCCCTCGAGCCCCTCGTCGACCTCC
 CCATCGGCATCAACATCACCCGCTTCCAGACCCTCCTCGCCCTCCACCGTCTACCTACCCCCGGC
 GACTCCTCCTCCGGTGGACCGCCGGCGCCGCCCTACTACGTGGTACCTCCAGCCCCGCACC
 TTCCTCCTCAAGTACAACGAGAACGGCACCATCACCGACGCCGTCTGACTGCGCCCTCGACCCCTCT
 CCGAGACCAAGTGCACCCTCAAGTCTTACCGTCTGAGAAGGGCATCTACCAGACCTCCAACCTCCG
 CGTCCAGCCACCGAGTCCATCGTCCGCTTCCCCAACATCACCAACCTCTGCCCCCTCGACGAGGTCT
 TCAACGCCACCCGCTTCCGCTCCGTCTACGCCTGGAACCGCAAGCGCATCTCCAACCTGCGTCCGCCA
 CTACTCCGTCTCTACAACCTCGCCCCCTTCTTACCTTCAAGTGTACGGCGTCTCCCCACCAAGCT
 CAACGACCTCTGCTTACCAACGTCTACGCCGACTCCTTCGTCTATCCGCGGGCAGGAGGTCCGCCAG
 ATCGCCCCCGGCCAGACCGGCAACATCGCCGACTACAACCTACAAGTCCCCGACGACTTACCCGGC
 TGCGTCTATCGCCTGGAACCTCAACAAGCTCGACTCCAAGGTGACGGCAACTACAACCTACTTACC
 GCCTTCTCCGCAAGTCCAACCTCAAGCCCTTCGAGCGCGACATCTCCACCGAGATCTACCAGGCGCG
 CAACAAACCCTGCAACGGCGTCCGCCGCTTCAACTGCTACTTCCCCCTCAGGTCTTACAGCTTACAG
 CCCACCTACGGCGTCCGCCACCAGCCCTACCGCGTCTGCTCCTCTCCTTCGAGCTCCTCCACGCC
 CCGCCACCGTCTGCGGCCCAAGAAGTCCACCAACCTCGTCAAGAACAAGTGGTCAACTTCAACTT
 CAACGGCCTCAAGGGCACCGGCGTCTCACCGAGTCCAACAAGAAGTCTCCTCCCCCTCCAGCAGTTC
 GGCCGCGACATCGCCGACACCACCGACGCCGTCCGCGACCCCCAGACCCTCGAGATCCTCGACAT
 CACCCCTGCTCCTTCGGCGGGTCTCCGTCTATACCCCGGCACCAACACCTCCAACCAGGTCCGC

GTCCTCTACCAGGGTGTCAACTGCACCGAGGTCCCCGTCGCCATCCACGCCGACCAGCTACCCCCA
 CCTGGCGCGTCTACTCCACCGGCTCCAACGTCTTCCAGACCCGCGCCGGTGCCTCATCGGGCCG
 AGTACGTCAACAACCTCCTACGAGTGCGACATCCCCATCGGGCCGGCATCTGCGCCTCCTACCAGAC
 CCAGACCAAGTCCCACCGCCGCGCCCGTCCGTCCGCTCCAGTCCATCATCGCTACACCATGTCC
 CTCGGCGCCGAGAACTCCGTGCGCTACTCCAACAACCTCCATCGCCATCCCCACCAACTTCACCATCTC
 CGTACCACCGAGATCCTCCCCGTCTCCATGACCAAGACCTCCGTGACTGCACCATGTACATCTGCG
 GCGACTCCACCGAGTGCTCCAACCTCCTCCTCCAGTACGGTCTTTCTGCACCCAGCTCAAGCGCGC
 CCTCACCGGCATCGCCGTGAGCAGGACAAGAACACCCAGGAGGTCTTCGCCAGGTCAAGCAGAT
 CTACAAGACCCCCCATCAAGTACTTCGGCGGCTTCAACTTCTCCAGATCCTCCCCGACCCCTCCA
 AGCCCTCAAGCGCTCCTTCATCGAGGACCTCCTCTTCAACAAGGTACCCCTCGCCGACGCCGGCTT
 CATCAAGCAGTACGGCGACTGCCTCGGGCAGATCGCCGCCGCGACCTCATCTGCGCCAGAAAGTT
 CAAGGGCCTACCGTCTCCCCCCCCCTCCTCACCGACGAGATGATCGCCCAGTACACCTCCGCCCTC
 CTCGCCGGCACCATCACCTCCGGTGGACCTTCGGCGCCGGCGCCGCCCTCCAGATCCCCTTCGCC
 ATGCAGATGGCCTACCGCTTCAACGGCATCGGGGTACCCAGAACGTCTCTACGAGAACCAGAAG
 CTCATCGCCAACCAGTTCAACTCCGCCATCGGCAAGATCCAGGACTCCCTCTCCTCCACCGCCTCCGC
 CCTCGGCAAGTCCAGGACGTGTCGAACCACAACGCCAGGCCCTAACACCCTCGTCAAGCAGCT
 CTCCTCAAGTTCCGGCGCCATCTCCTCCGTCTCAACGACATCTTCTCCCGCCTCGACAAGGTGAGG
 CCGAGGTCCAGATCGACCGCCTCATCACCGGCCGCTCCAGTCCCTCCAGACCTACGTACCCAGC
 AGCTCATCCGCGCCGCCGAGATCCGCGCCTCCGCCAACCTCGCCGCCACCAAGATGTCCGAGTGCG
 TCCTCGGCCAGTCCAAGCGCGTGCCTTCTGCGGCAAGGGTACCACCTCATGTCTTCCCCAGTC
 CGCCCCCACGGCGTCTTCTCCTCACGTACCTACGTCCCCGCCAGGAGAAGAAGTTCACCACC
 GCCCCCGCATCTGCCACGACGGCAAGGCCACTTCCCCCGGAGGGCGTCTTCTGCTCTCAAACGGC
 ACCACTGGTTCGTCACCCAGCGCAACTTCTACGAGCCCCAGATCATCACCCAGACAACACCTTCGT
 CTCGGCAACTGCGACGTGTCATCGGCATCGTCAACAACACCGTCTACGACCCCTCCAGCCCCGAG
 CTCGACTCCTTCAAGGAGGAGCTCGACAAGTACTTCAAGAACCACACCTCCCCGACGTGACCTCG
 GCGACATCTCCGGCATCAACGCTCCGTGTCGAACATCCAGAAGGAGATCGACCGCCTCAACGAGG
 TCGCAAGAACCTCAACGAGTCCCTCATCGACCTCCAGGAGCTCGGCAAGTACGAGCAGTACATCAA
 GTGGCCCTGGTACATCTGGCTCGGCTTCATCGCCGGCCTCATCGCCATCGTCATGGTACCATCATGC
 TCTGCTGCATGACCTCCTGCTGCTCCTGCCTCAAGGGTGTGCTCCTGCGGCTCCTGCTGCAAGTTC
 GACGAGGACGACTCCGAGCCCCGTCTCAAGGGCGTCAAGTCCACTACACCTAA

Acknowledgements

This resource has been sponsored by the Joachim Herz Stiftung and produced by the European XFEL.

