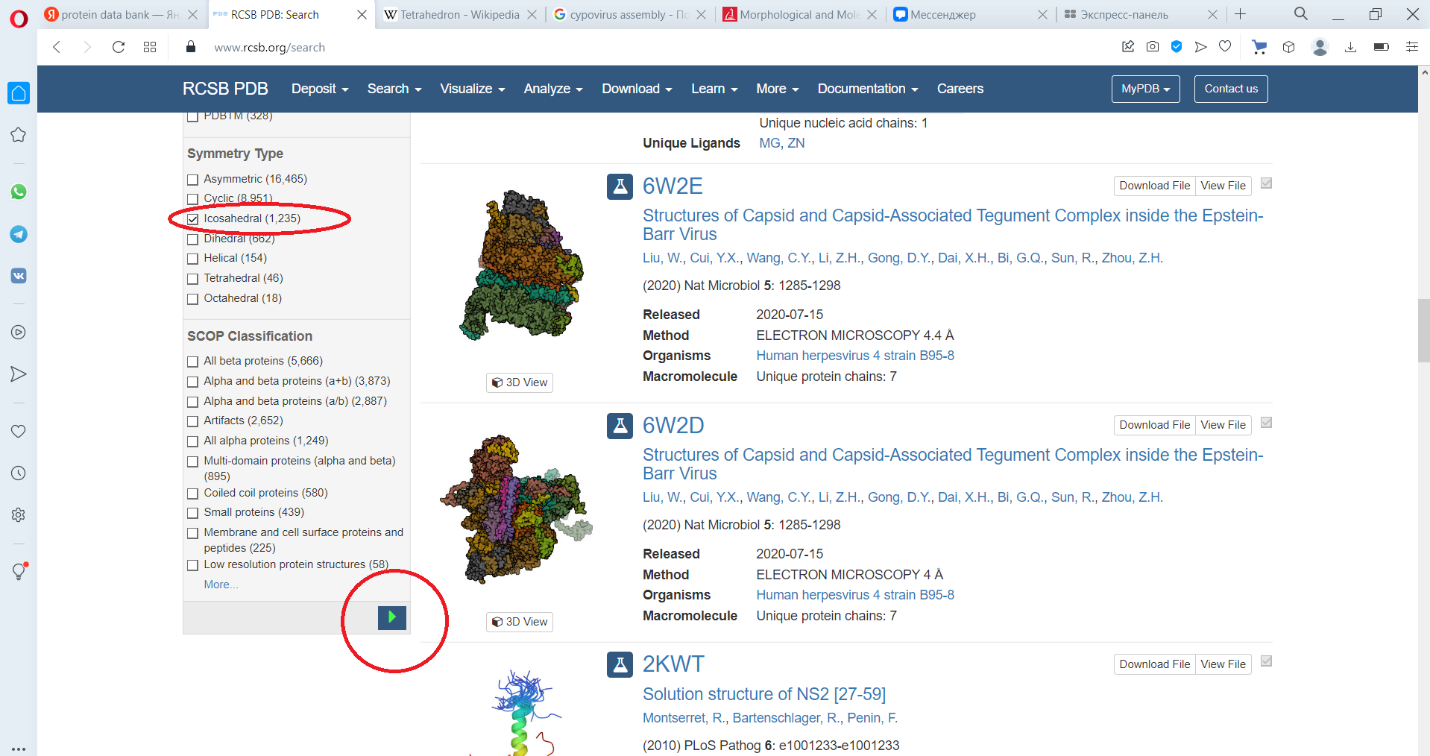
Ищем Protein data bank

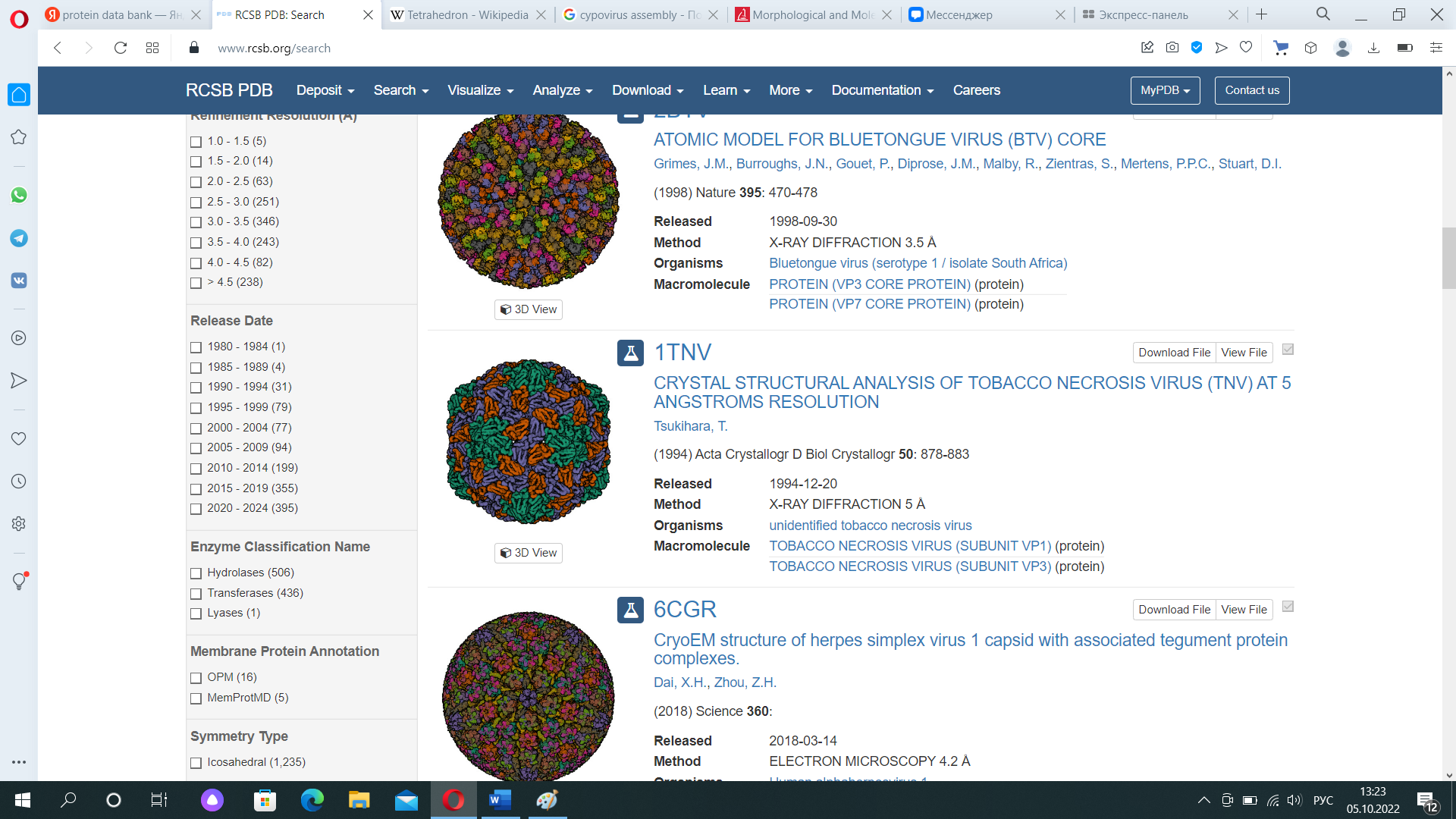
<https://www.rcsb.org>

Запускаем поиск по слову virus.

Далее в левом крайнем столбце в Symmetry type ставим галочку около icosahedral и жмем стрелочку в самом низу столбца.

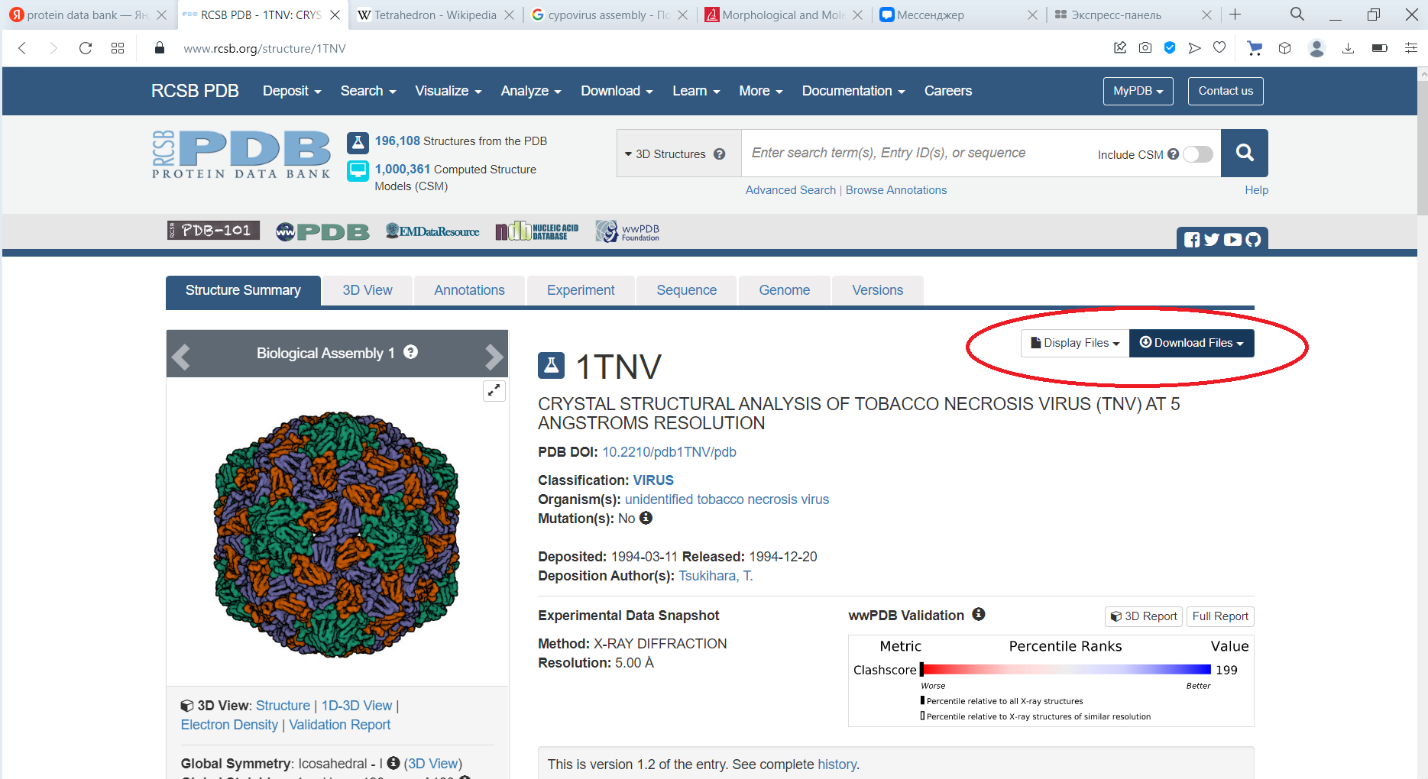


Ищем оболочку, состоящую из 3х цветов. Например, 1TNV.



Нажимаем на 1TNV.

В верхнем правом углу есть 2 кнопки: Display files and Download files. Нажимаем Download files, ищем PDB format, нажимаем. Скачивается файл, открыть его можно в блокноте.

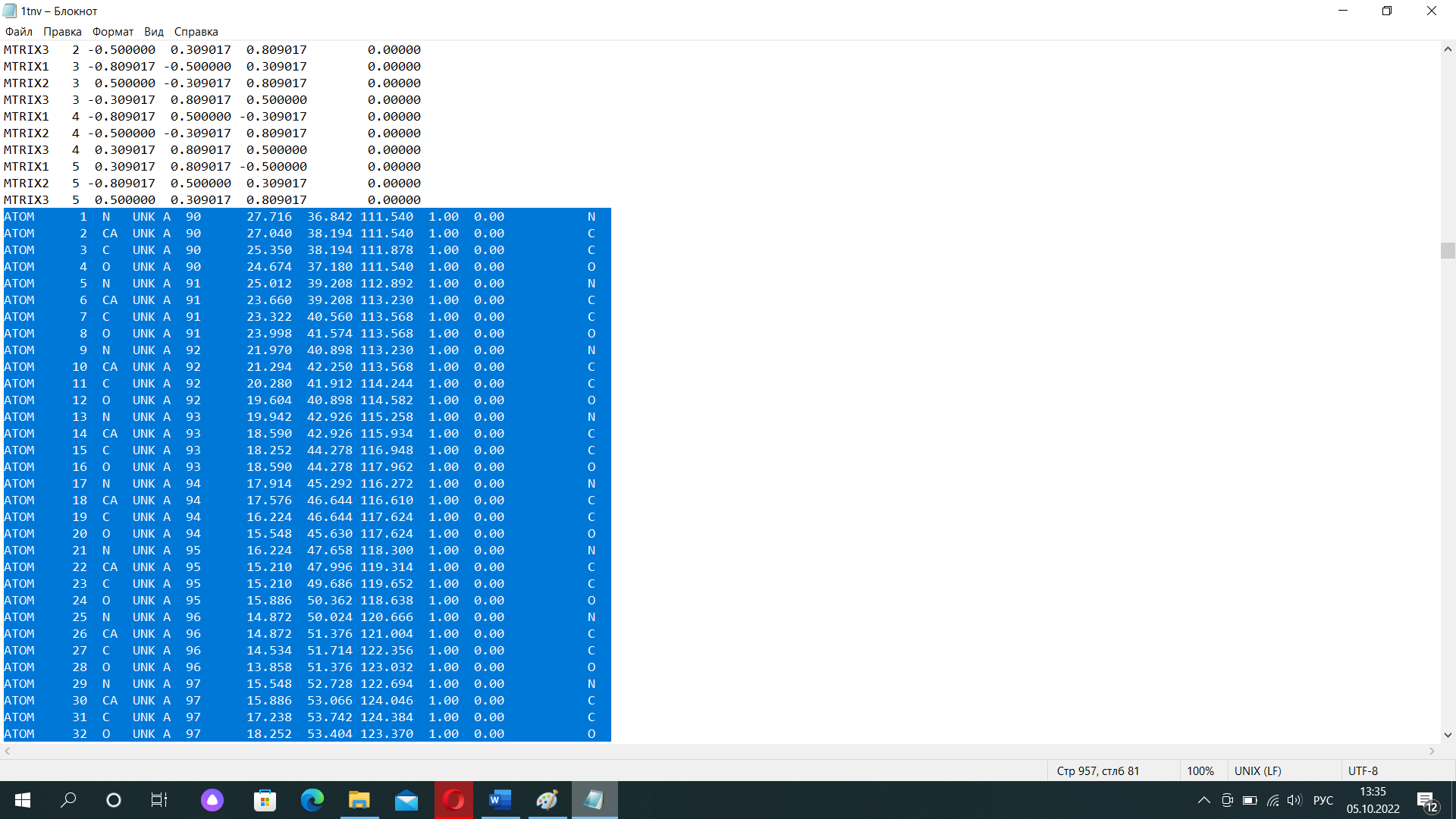


В самом начале файла есть информация о симметрийноэ-неэквивалентных белках в оболочке. Здесь это chain: A, B and C. Иногда вместо букв здесь цифры 1,2,3. По количеству цветов в оболочке.

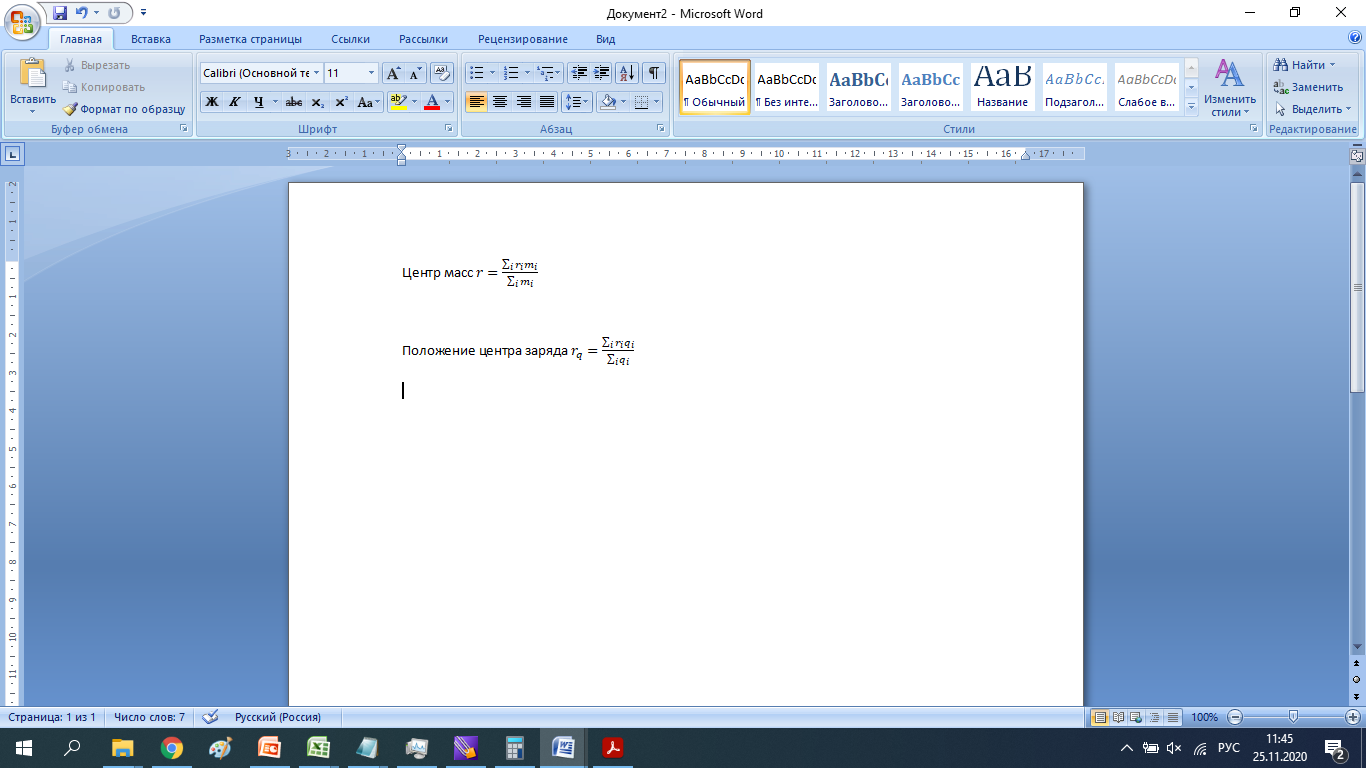
Изображение выглядит как текст

Автоматически созданное описание

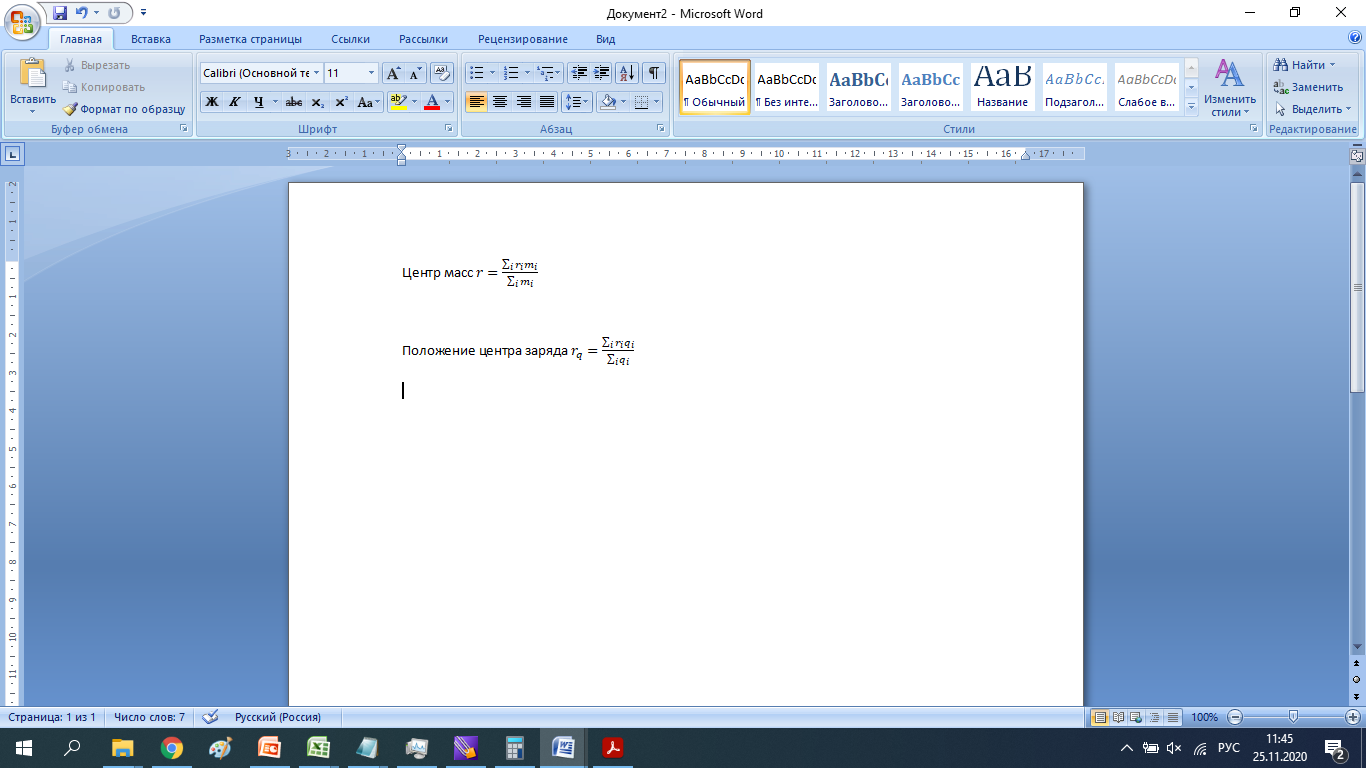
Листаем файл, находим строки, которые начинаются со слова ATOM. В 4-м столбце обычно информация об аминокислоте. В данном файле аминокислоты не определены. Стоит UNK, а должно быть: Asp, Glu, Tyr и другие. В 5-м столбце каждой такой строки содержится информация о белке (то есть указано название цепочки A, B или C). 9-й, 10-й, 11-й столбец – это координаты. Последний столбец – это название химического элемента (N, C, O).



Очевидно, что каждый белок состоит не из одного атома. Например, в этом файле цепочка A состоит из 744 атомов. Поэтому вам нужно найти центр масс каждого симметрийно-неэквивалентного белка.



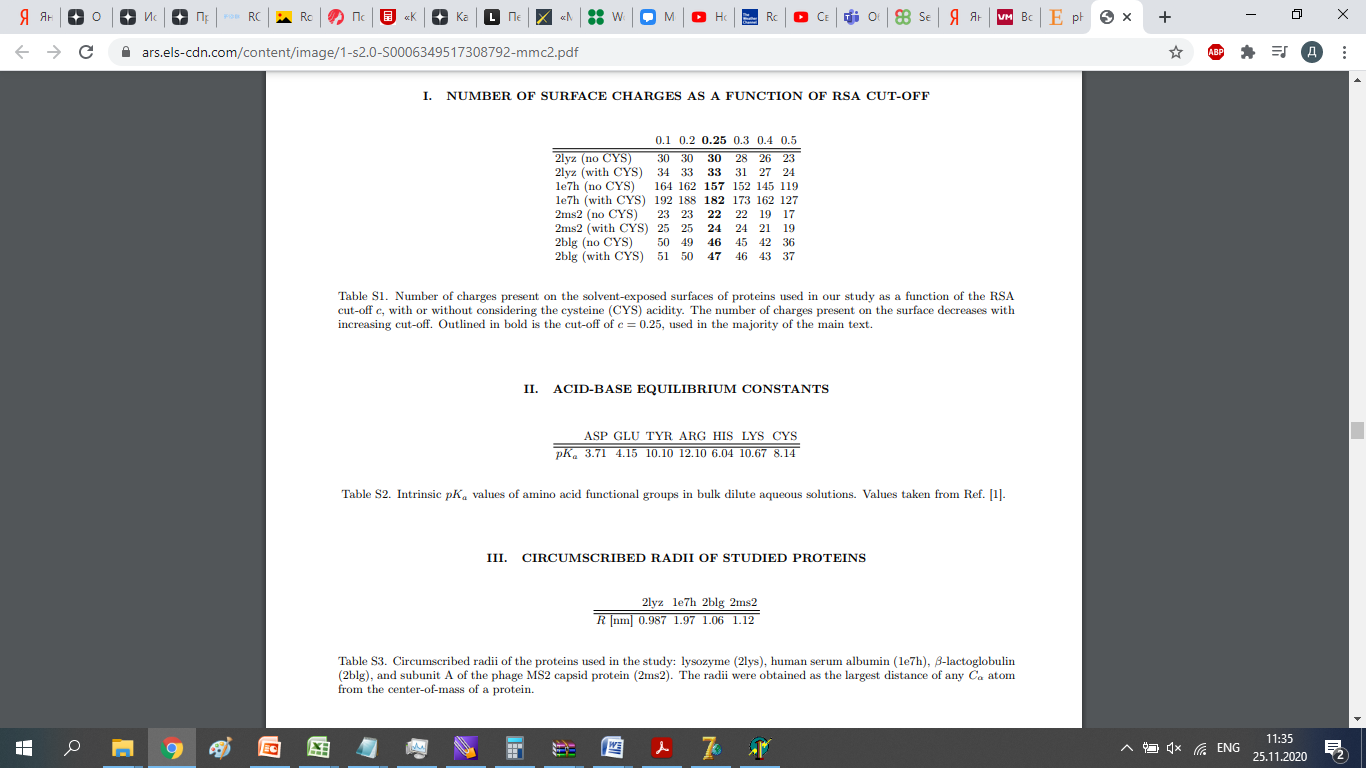
Кроме того, положения центров зарядов. Считаем при фиксированном значении pH=7.

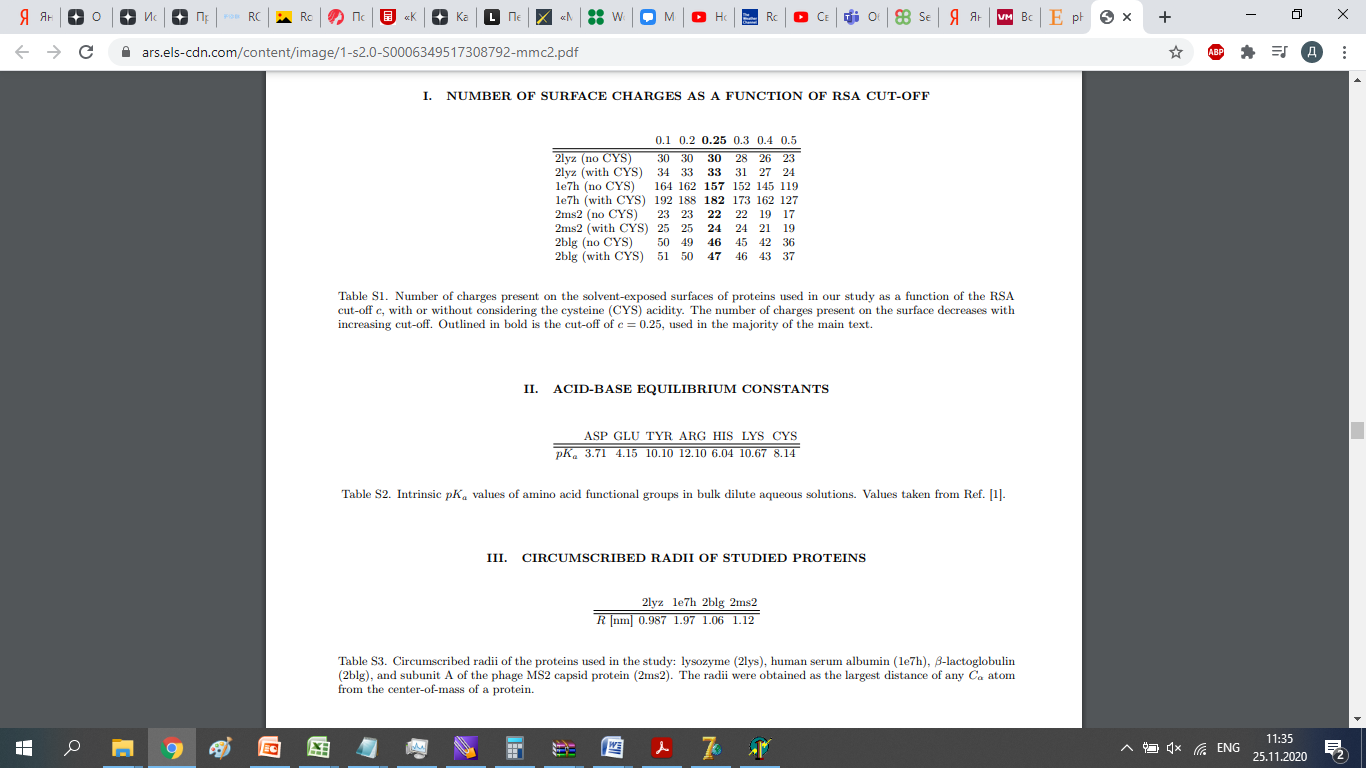


Partial charges of the ionizable residues were calculated using the Henderson-Hasselbach equation which follows the Langmuir isotherm55:

where is the dissociation constant of the protonated/deprotonated AAs (we used the values from ref. 63) and the signs depend on the type of amino acid: plus for basic AAs (Arg, His, Lys), minus for acidic AAs (Asp, Glu, Tyr). In the dissociable AAs breakdown, we did not include Cys which possesses a thiol functional end group that is a weak acid but is usually not considered to be one because the thiol group is often reactive and can form disulfide bonds.

1. W. M. Haynes, D. R. Lide and T. J. Bruno, Eds., *CRC Handbook of Chemistry and Physics*, CRC Press, 2016.





Про расчет заряда можно почитать дополнительно здесь:

<https://www.sciencedirect.com/science/article/pii/S0006349517308792>

Итак, задание.

* Выбрать вирус (каждому – свой), но так, чтобы были известны названия аминокислот и на картинке было 3 цвета, то есть состоящий из 3 симметрийно-неэквивалентных белков (цепочек).
* Найти центры масс белков. Это должны быть координаты 3-х точек.
* Найти центры зарядов белков.
* Прислать мне программы и файлы отчетов.