*6. Выберите вирусный белок с известной пространственной структурой; например, гомологичный белку, последовательность которого вы смотрели ранее. Поиском по сходству структур найдите еще два-три белка с похожей структурой. Используйте сервер PDBeFOLD (<http://www.ebi.ac.uk/msd-srv/ssm/>)*

*Поиск PDB код, выберите и цепочку, – против Whole PDB*

*Как выбрать находки. Смотрите на три параметра: (1) длину выравнивания входной структуры с находкой; лучше, чтобы она была побольше, и не сильно меньше, чем длина вашего белка (2) RMSD; выбирайте находки с RMSD 1-2.5 ангстрем; если таких недостаточно – расширьте эти границы (3) название находки; лучше, чтобы находки были из других вирусов (или организмов), или другими белкам, чем входной белок*

*Если не нашлось подходящих находок, то повторите, уменьшив на входной странице параметры Lowest acceptable match (%) во входной структуре и в находке. В презентации написано, что они значат.*

*Отметьте выбранные структуры в чекбоксах.*

*Внесите в протокол pdb коды выбранных белков и входного белка, названия белков и из какого организма ( вируса).*

*Не выходя из сервиса PDBeFOLD совместите выбранные структуры (*Submit for Multiple Alignment ); см. Скриншоты в презентации.

*Сохраните в совмещенные структуры в одном файле: Viewer Rasmol, Superpose whole entries – можно поставить галочку (если все находки – одно доменные белки), а можно и не ставить (если PDB файлы – большие, содержат комплексы белков и т.п. Хотя иногда и совмещение комплексов интересно). Нажмите view superposed и сохраните файл, дав ему понятное имя и расширение .pdb). См также в презентации.*

*Откройте скачанный файл в текстовом редакторе Notepad++ и удалите из него первые строки со скриптом – до строки exit включительно.*

*Откройте файл в Jmol. Имейте ввиду, что сервер переименовывает совмещаемые цепочки в A, B, C и так далее.*

*Создайте изображение совмещенных структур в модели backbone 80 (или trace, или cartoon); каждую из структур покрасьте в свой цвет. В протоколе укажите число совмещаемых C-alpha атомов (number of aligned residues) и Overall RMSD*

*Сохраните выравнивание последовательностей совмещенных структур, такое выравнивание предоставляет сервис PDBeFoldeFOLD. Откройте его в jalview, раскраска ClustalX, процент Identity – 100%. В протоколе укажите число и процент консервативных позиций. Проверьте и опишите, совпадают ли они сопоставленным С-альфа атомами.*

*8 (Не обязательное, но полезное). В белке с известной 3D структурой сравните «структурные домены» и консервативные домены согласно БД Pfam. Заполните табличку:*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *В структуре, на глаз* | *В БД структурных доменов CATH* | *Названия доменов* | *В БД консервативных доменов Pfam* | *Названия доменов* |
| *Число доменов* |  |  |  |  |  |
| *Границы доменов 1й* |  |  |  |  |  |
| *2й* |  |  |  |  |  |
| *3й* |  |  |  |  |  |

*Напишите в протоколе совпадают ли границы доменов CATH и Pfam.  
Приведите рисунок структуры, на котором выделены цветом домены.*

*Выберите свой белок или возьмите* любой из: 1C1D, 1dq3(2), 2ijd, 1WER (2), 1TPP(2), 1ZZZ(2).

Визуально определите границы структурных доменов. Для этого изобразите остов одной цепи в виде backbone 80. Раскрасьте по радуге (color group) от N-конца к C-концу. В первом приближении, в том месте остова, в котором разрезание линии приведет к распадению частей в пространстве без существенных зацеплений и находится граница доменов. Для некоторых доменов одного разреза мало, придется сделать два мысленных разрезе цепи. Выделите и раскрасьте домены командами select <фрагмент> (например, select 1-115:A) и color <ваш любимый цвет> (например, color green).

Сегодня есть две основные структурные классификации - SCOPe (structural classification of proteins, extended) и CATH. Адреса веб-ресурсов - www.google.ru. Будем использовать CATH.

Используйте БД PDBsum, в которой есть информация и из CATH, и из Pfam

Если в белке более трех доменов – можно игнорировать четвертый и последующие при выполнении этого задания.