

Нахождение гидрофобных кластеров в структуре белка

Для того чтобы оценить работу алгоритма CluD на разных типах структур был выбран набор структур по их размеру, количеству суединиц и составу (белок и ДНК-белковый комплекс).

Первая рассмотренная структура – 5o5s.

Это C терминальный домен RapZ из E.coli. Rap Z белок примечателен тем, что совсем недавно (2017-08-30) было обнаружено что он высококонсервативен среди бактерий поскольку он контролирует биогенез глюкоами фосфата.

Данная структура была выбрана поскольку она небольшая, имеет в составе альфа-спирали и бета-лист, и две 3/10 helix, а также изгиб между двумя альфа-спиралями.

Был проведен анализ CluD при изменении порога, в пределах которого атомы объединяются в один кластер (distance threshold), и минимального количества атомов в одном кластере (атомов, а не аминокислот).

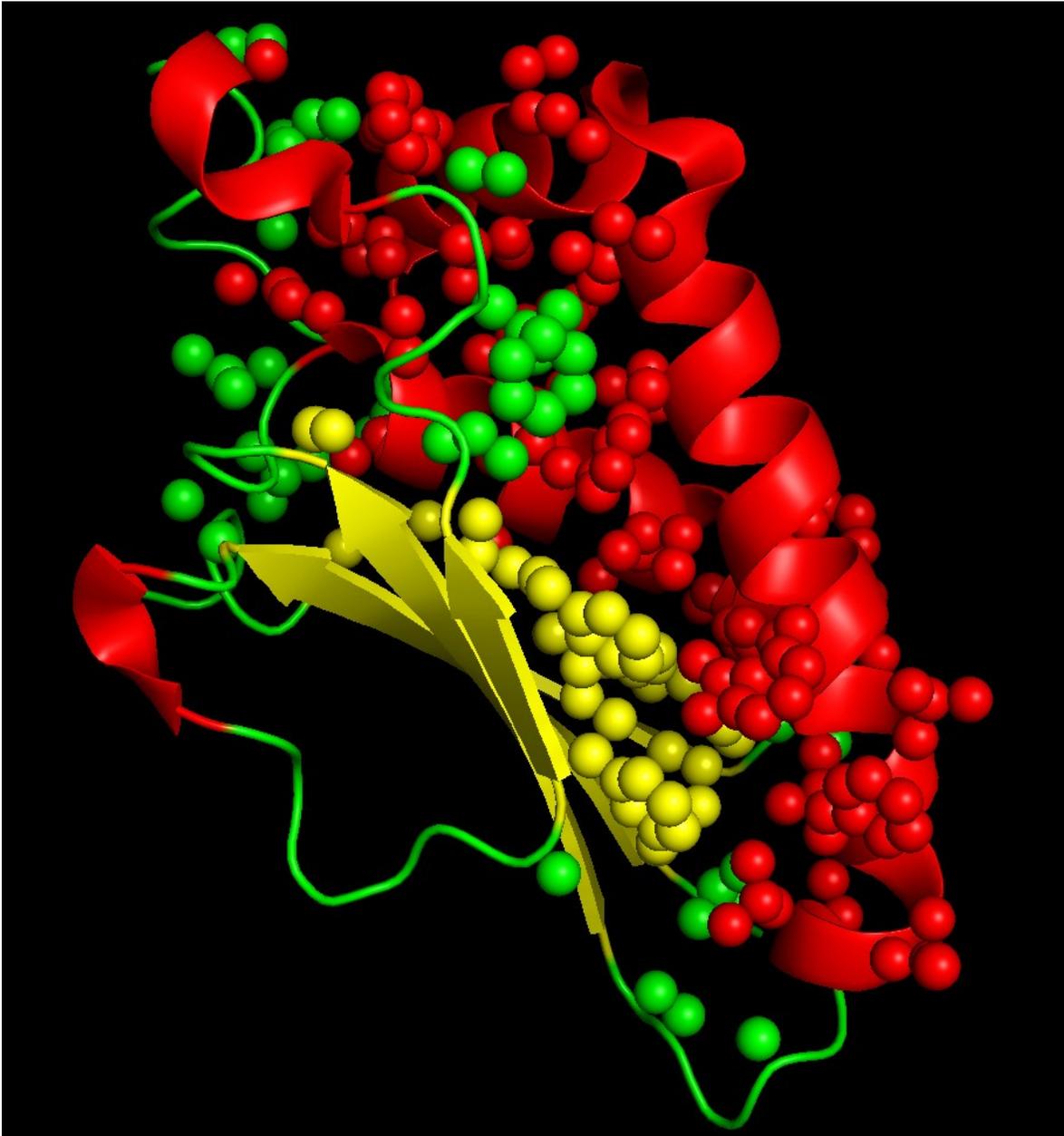
В целом для всех структур верно, что при уменьшении порога на расстояние крупные кластеры распадаются на более мелкие, а при увеличении порога на расстояние мелкие кластеры объединяются в более крупные (что ожидаемо). Однако часто при уменьшении порога на расстояние от большого “доминантного” кластера, образующего гидрофобное ядро молекулы, отпочковываются более мелкие кластеры из нескольких аминокислот, иногда не соответствующие какому-то определенному элементу вторичной структуры, а иногда покрывающие часть его (например часть альфа-спирали).

Также чем больше размер структуры, тем больше кластеров (что логично), и тем больше кластеров небольшой величины (несколько аминокислот).

Для трех белковых структур – 5o5s, 2bnq и 3uob, варьировались параметры порога расстояния (в пределах 5.0 – 10.0) и минимального количества атомов (от 3 до 10).

После этого за неимением особого опыта с данным инструментом были взяты дефолтные настройки (3 атома и 5.4 Å).

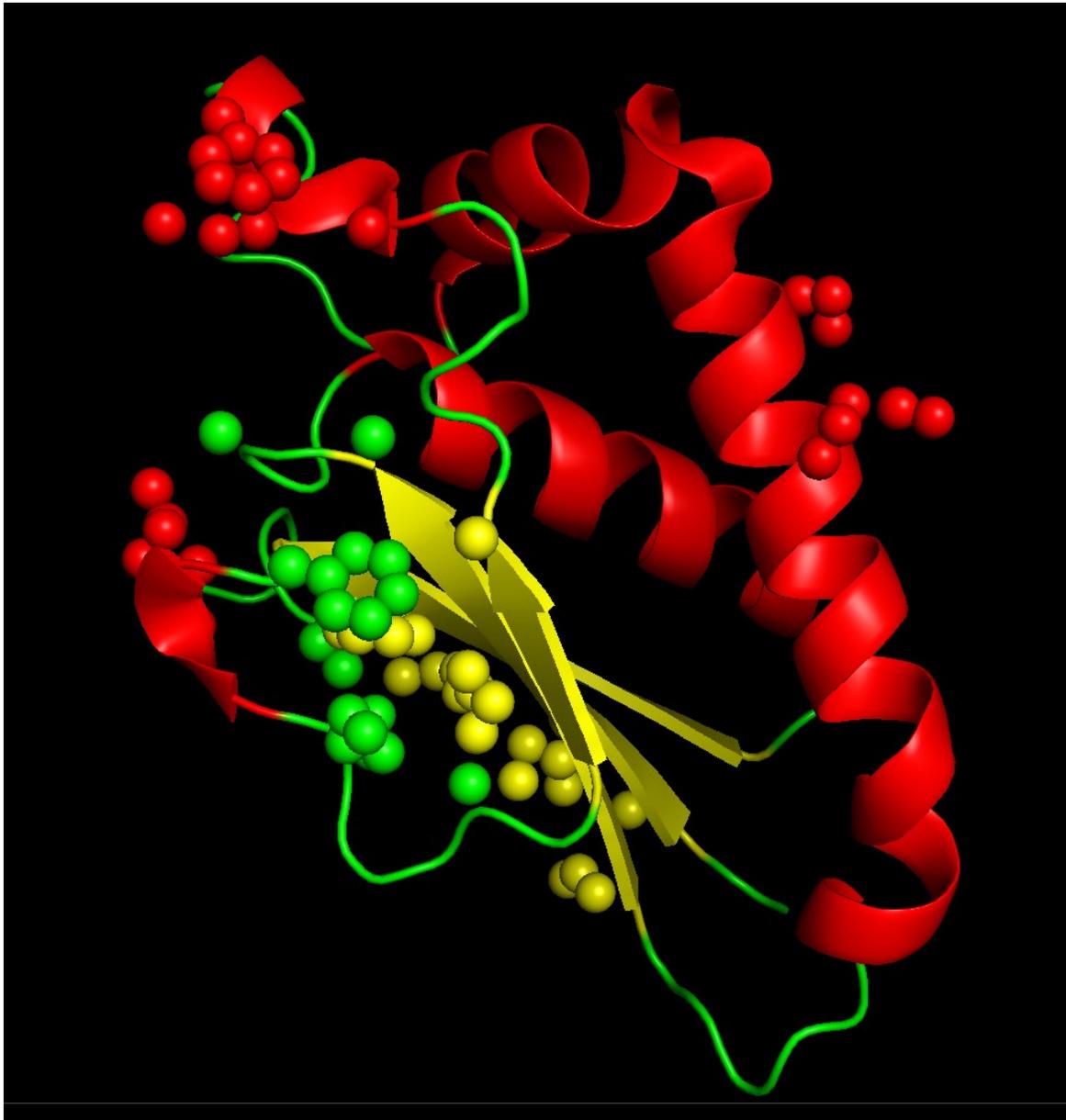
Рассмотрим выдачу CluD для структуры 5o5s.



Гидрофобное ядро структуры 5o5s. Сферами выделен самый большой кластер из выдачи CluD.

Было выделено 5 кластеров. Самый большой из них составляет гидрофобное ядро молекулы.

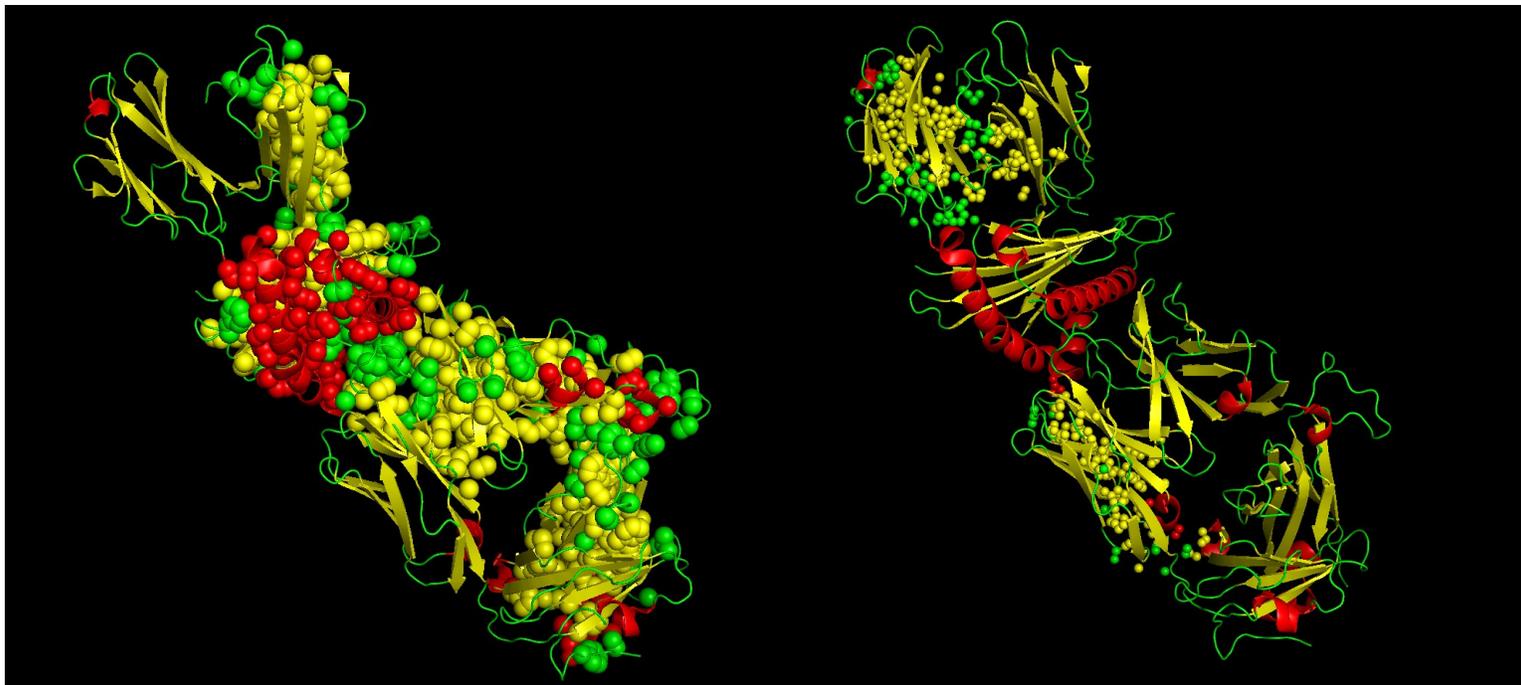
Второй по размеру обрамляет снаружи бета-лист. Остальные три кластера очень малы по размерам и не соответствуют элементам вторичной структуры. При увеличении минимального размера кластеров они объединяются с первым, самым большим кластером. Таким образом, можно заключить что на самом деле “функциональных” гидрофобных кластеров здесь два – гидрофобное ядро и кластер, обрамляющий бета-лист.



Кластеры меньшего размера в структуре 5o5s. При изображении всех кластеров одновременно невозможно установить их границы, поэтому самый большой кластер не показан.

Далее была рассмотрена структура гетеротетрамера 2bnq. Это структура T клеточного рецептора из *E. coli*.

Для нее CluD нашел 30 гидрофобных кластеров, наибольший из них покрывает всю молекулу кроме двух бета-листов. Остальные кластеры состоят из нескольких аминокислот на поверхности белка.



Самый большой кластер в структуре 2bnq (слева), и два кластера, соответствующие бета-листам (справа).

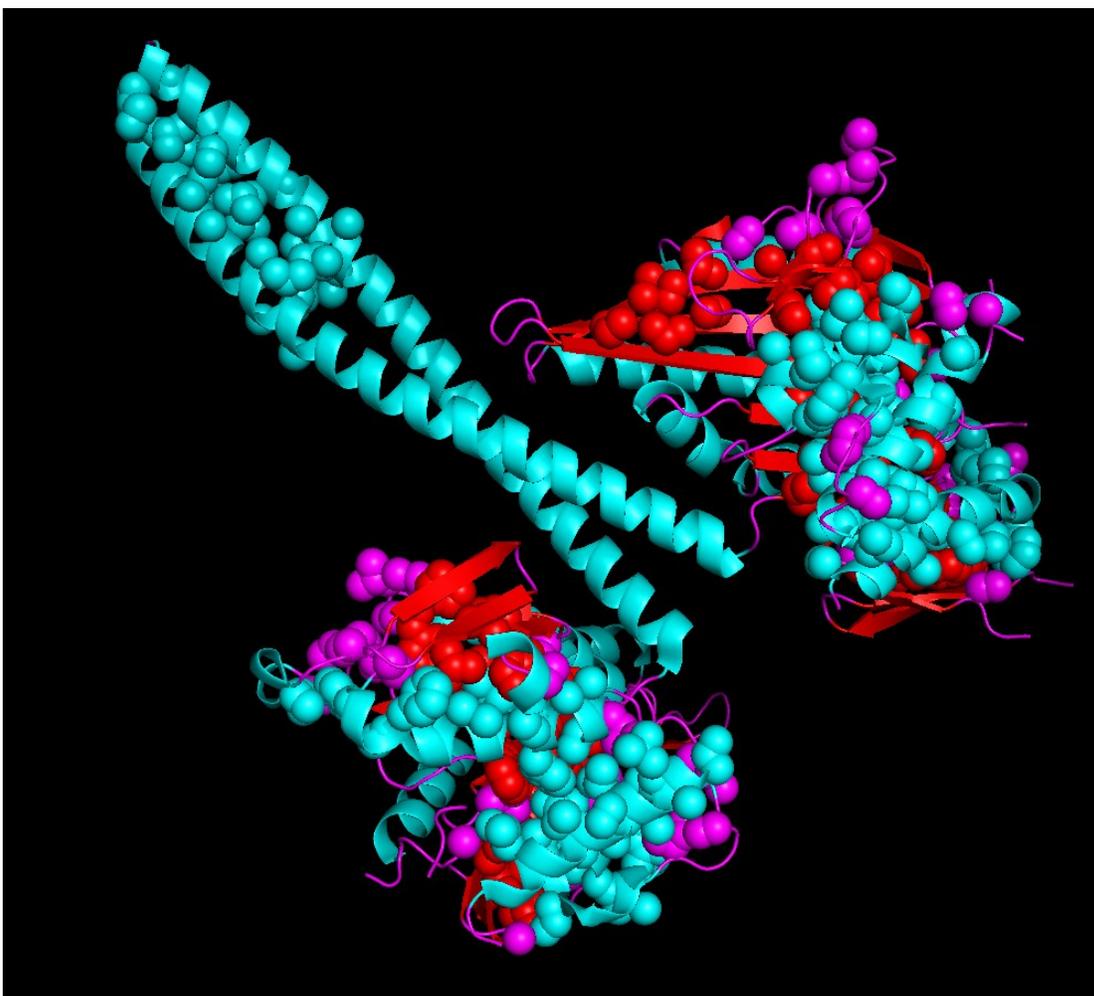
Далее была выбрана структура 3u06 (не путать с 3u0b). Структура была выбрана поскольку она является гомодимером с двумя альфа-цепями, между которыми ожидался гидрофобный кластер.

Структура принадлежит кинезину из *E. coli*. Каждый мономер состоит из А цепи и В цепи (альфа-спирали). Действительно, были найдены 32 гидрофобных кластера, один из которых соответствует взаимодействию двух α -спиралей боковыми поверхностями, богатыми гидрофобными остатками. Два остальных соответствуют А цепям мономеров.

Остальные кластеры соответствуют небольшим гидрофобным участкам на поверхности белка.

В целом можно сделать вывод, что для достижения более-менее правильного результата CluD необходимо уметь варьировать параметры порогового расстояния и минимального количества атомов в кластере в зависимости от таких параметров, как размер молекулы и количество гидрофобных остатков на поверхности молекулы и в гидрофобном ядре (например если мы рассматриваем фермент с гидрофобными остатками в каталитическом центре, то необходимо добиться того, чтобы они составляли различные кластеры).

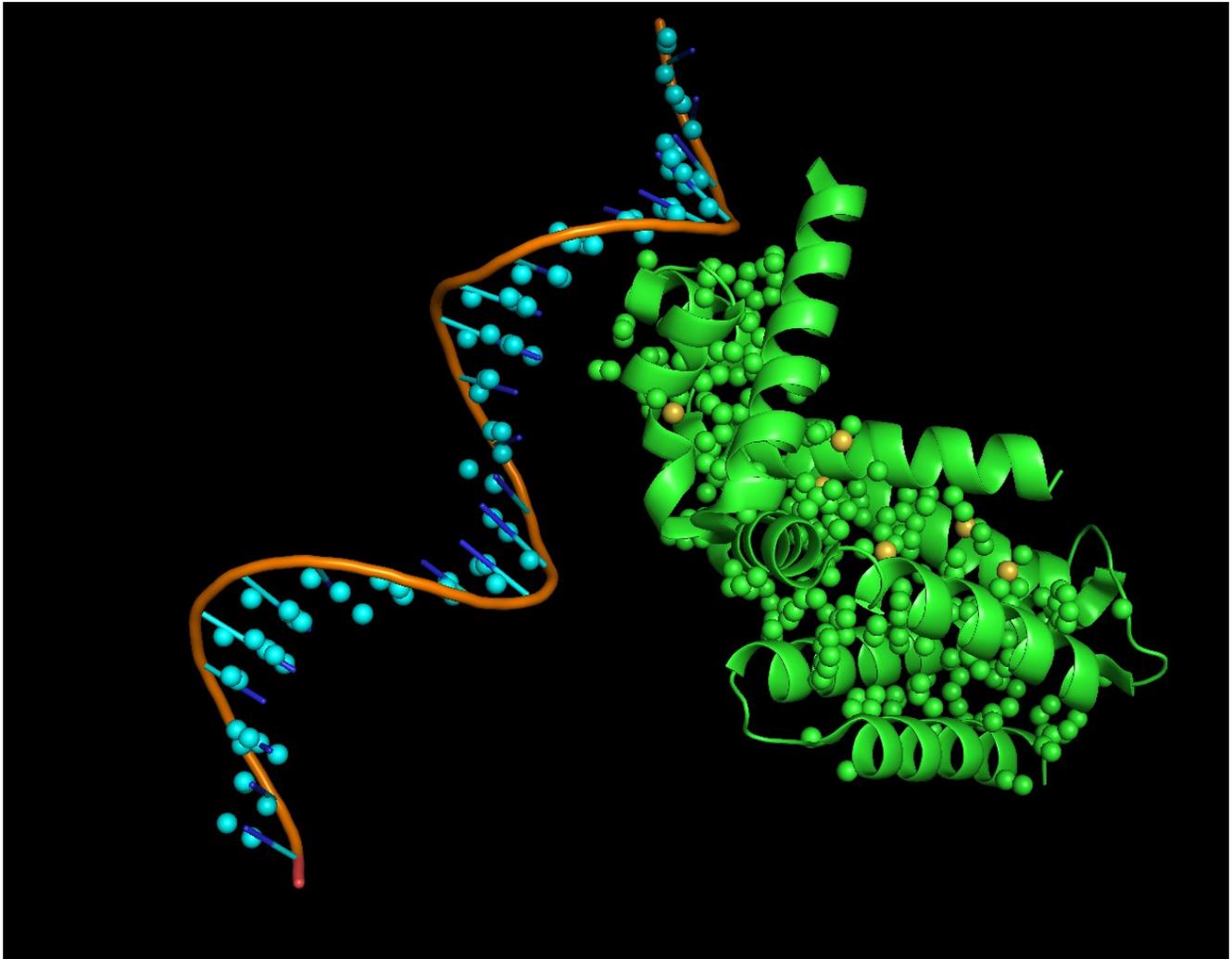
К сожалению, я не научилась правильно использовать данный инструмент.



Гидрофобные кластеры в структуре 3u06.

Далее была исследован ДНК-белковый комплекс 5ua2.

Было выделено 6 кластеров, три из которых принадлежат ДНК. Показана одна цепь ДНК и мономер белка для лучшей визуализации.

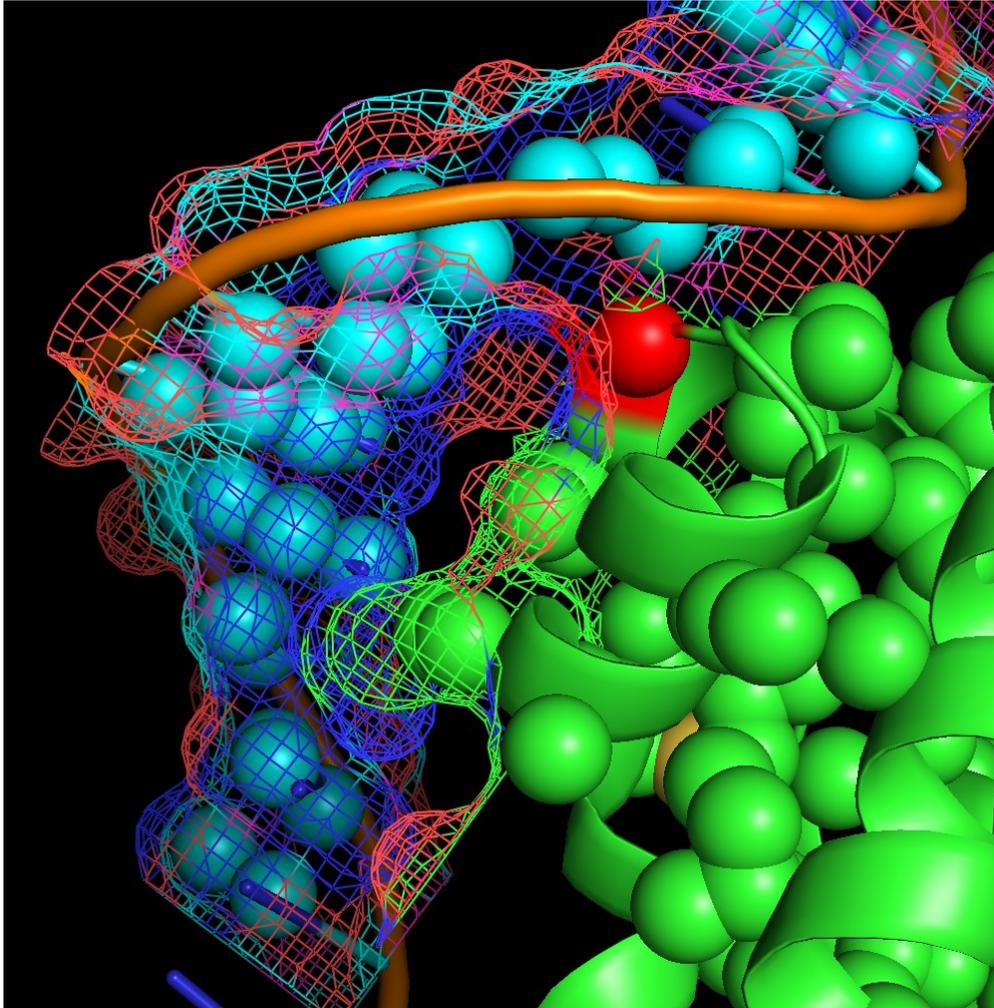


Структура ДНК-белкового комплекса 5ua2, отмечены три кластера для ДНК и один для мономера белка KstR.

Отдельно рассмотрим контакты между белком и цепью ДНК.

Был отдельно рассмотрен кластер в ДНК в непосредственной близости белка. Кластер ДНК (отмечен синим) включает аминокислоту из мономера белка KstR (отмечена красным).

Возможно в данном случае кроме гидрофобных взаимодействий наблюдаются также локальные полярные контакты.



Аминокислоты, обуславливающие гидрофобное взаимодействие между мономером белка и ДНК. Аминокислота, вошедшая в кластер вместе с нуклеотидами ДНК, отмечена красным.