

Совмещение структур

Поиск структурных гомологов белка

С помощью сервиса PDBeFold был проведен поиск структурных гомологов белка репарации и рекомбинации RadA, идентификатор PDB 3ntu. В 2014 году я уже выполняла аналогичное задание: [ссылка](#). Тогда было обнаружено 26 структур, а подробно рассматривалась только одна находка. В этот раз найдено 43 структуры с настройками по умолчанию (**Lowest acceptable match 70%**) и 316 находок с **Lowest acceptable match 50%**. Возможно, в прошлый раз порог для поиска был строже. Я использовала список из 316 находок, так как среди 43 обнаруживаются только очень близкие структуры. Но в целом среди находок много структур 2015 и 2016 года, то есть можно сказать, что данный белок активно исследуют.

Я выбрала 4 структурных гомолога. Среди них есть и та структура, что была рассмотрена в работе 2014 года (4hyu:D). Они представлены в таблице 1, первая запись - это query.

Таблица 1. Выбранные структурные гомологи структуры 3ntu

##	RMSD	Nalgn	Nsse	Ngaps	Seq-%	Query	Target	Title
1	0	310	24	0	1	3ntu:A	3ntu:A	RADA RECOMBINASE D302K MUTANT IN COMPLEX WITH AMP-PNP
2	1,174	215	17	7	0,4605	3ntu:A	4hyu:D	FILAMENT OF OCTAMERIC RINGS OF DMC1 RECOMBINASE FROM HOMO SAPIENS
3	1,747	186	12	8	0,2419	3ntu:A	1xms:A	"E. COLI RECA IN COMPLEX WITH MNAMP-PNP"
4	1,776	276	20	9	0,3949	3ntu:A	3lda:A	YEAST RAD51 H352Y FILAMENT INTERFACE MUTANT
5	1,895	285	18	12	0,4035	3ntu:A	5jzc:C	HELICAL FILAMENT (Cryo-EM)

PDBeFold строит совмещение выбранных структур в Jmol или RasMol. Результат для выбранных находок показан на рисунке 1. Видно, что все структуры в целом похожи друг на друга по топологии. Больше всего выделяется структура 1xms:A белка RecA из E.coli. Известно, что белки архей RadA и эукариот Rad51 или DMC1 больше похожи друг на друга, чем RadA архей и RecA бактерий. В находках 4hyu:D и 1xms:A отсутствует N-концевой ДНК-связывающий домен, который есть в структуре 3ntu. Он представляет собой стандартный домен Helix-Hairpin-Helix. Возможно, этот домен не представляет особого интереса, поэтому его не включают в рекомбинантный белок. Также это может быть связано с удобством кристаллизации.

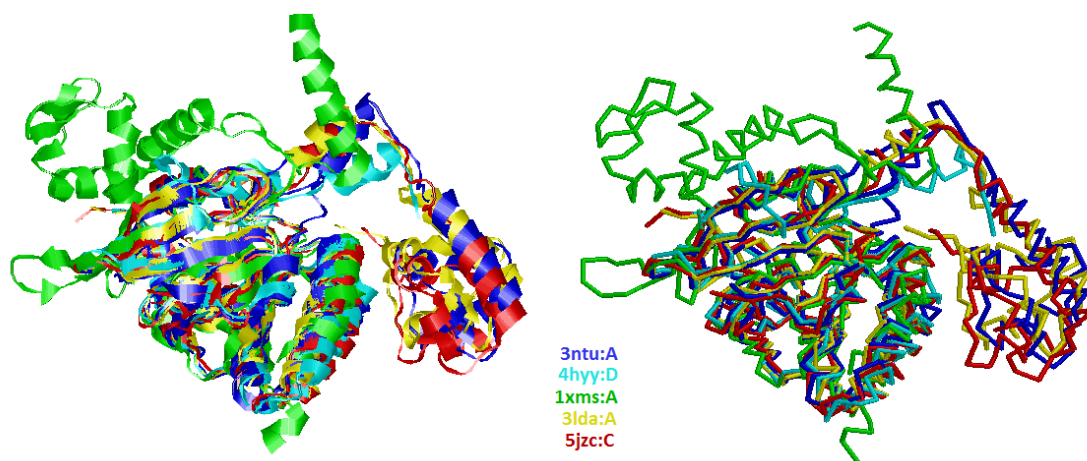


Рисунок 1. Совмещение выбранных структур в представлении cartoons (слева) и backbone (справа). Получено с помощью RasMol.

Сравнение множественных выравниваний

Рассмотрим множественные выравнивания:

- структурное, выданное PDBeFold;
- построенное по последовательностям с помощью Muscle.

В целом выравнивания очень близки, что можно увидеть на рисунке 2.

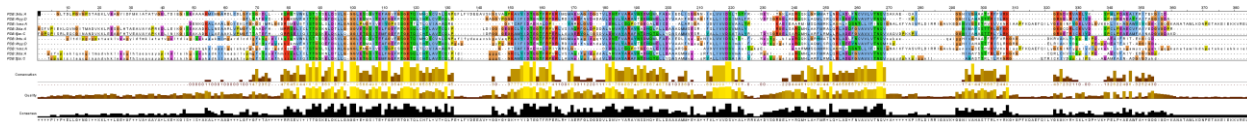


Рисунок 2. Общий вид двух множественных выравниваний. Сверху выравнивание пяти последовательностей по последовательности, снизу - по функции. Консервативные остатки покрашены в каждой группе Clustalx. Получено с помощью JalView.

Однако между двумя выравниваниями есть несколько небольших различий. Одно из них показано на рисунке 3. В последовательности 1xms идут подряд два остатка аланина. В выравнивании Muscle первый аланин попал в столбец с треонином и метионином, а второй аланин – в столбец, где есть еще три аланина и серин. В структурном выравнивании в столбец с аланинами попадает первый из аланинов последовательности 1xms, а перед ним – одиночный гэп. К тому же, после двух аланинов в 1xms есть лейцин, который попал в столбец с двумя лейцинами и треонином в версии Muscle, и в колонку с гэпами в версии PDBeFold.

Если посмотреть на совмещение структур, то становится понятно, почему так произошло. Отклонение всего участка цепи в структуре 1xms в этом месте привело к тому, что первый остаток аланина оказался ближе не к «своим» остаткам, а к следующим за ними по цепи. Таким образом, в данном случае выравнивание Muscle правильное, а структурное содержит ошибку.

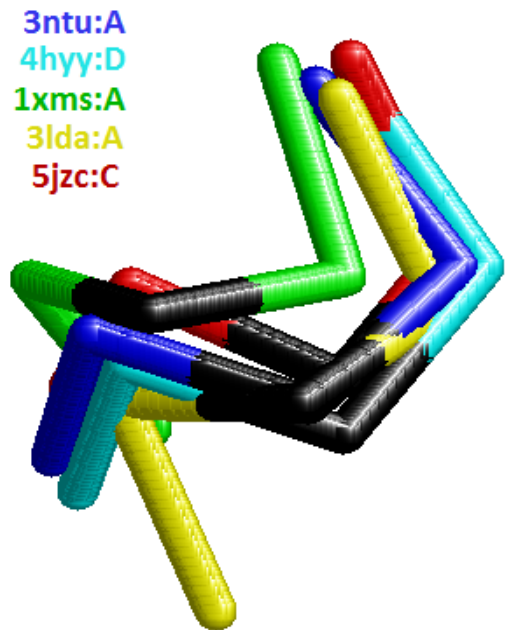
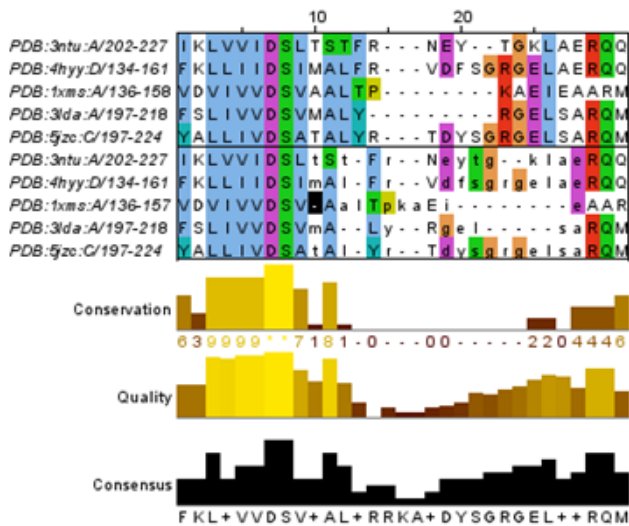
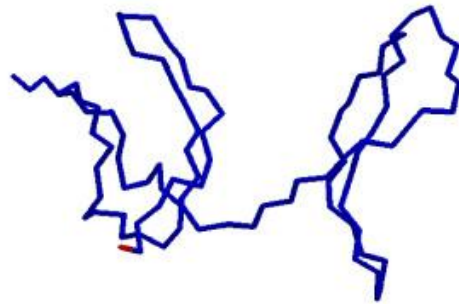


Рисунок 3. Различия в множественных выравниваниях. Слева: участок множественных выравниваний, сверху выравнивание пяти последовательностей по последовательности, снизу - по функции. Черным выделен одиночный гэп (описание в тексте). Справа: совмещение структур, черным выделены остатки из столбца выравнивания с одиночным гэпом.

Преимущества гибкого выравнивания относительно жесткого

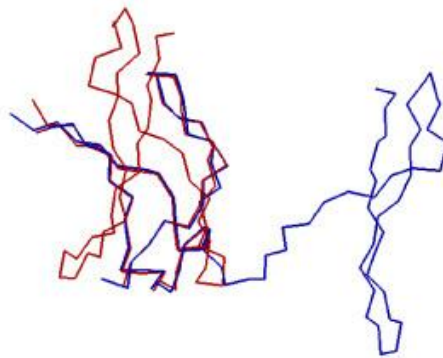
Ряд программ, в том числе PDBeFold, используют метод жесткого выравнивания, то есть стараются совместить структуры в целом. Другой подход – метод гибкого выравнивания – позволяет выравнивать между собой наиболее совместимые локальные участки структур. Иногда таким методом удастся получить существенно более удачные структурные выравнивания. Пример приведен на рисунке 4.



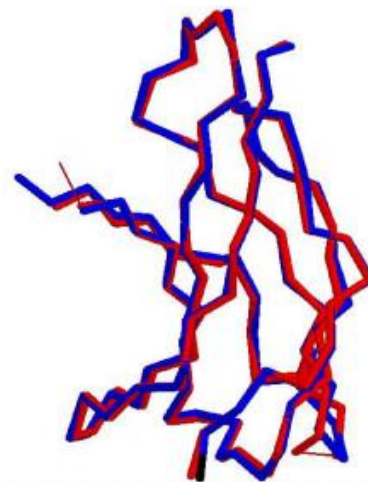
(a) pdbID:1u42 chain:A 101 residues



(b) pdbID:1u36 chain:A 100 residues



(c) DALI (N/*rmsd*): 55/0.8Å



(d) FlexSnap (N/*rmsd*/hinges): 100/0.89Å/1

Рисунок 4. Совмещение структур 1u42 (а) и 1u36 (b) методами жесткого (c) и гибкого (d) выравнивания. В первом случае совмещено 55% остатков, во втором – 100%. Источник: Salem et al: **Flexible Non-sequential Protein Structure Alignment. Algorithms for Molecular Biology** : AMB. 2010;5:12.