

Практикум № 10. Альфафолд2

Вариант: A7
последовательность:
GVLYVGSKTK
KGVVHGVATV
AEKTKEQVTN
VGGAVVTGVT
AVAQKTVEGA
GSIAAATGFV
KKD

1 Введение: Амилоиды

Я решила что все необходимое расскажу по ходу с картинками, поэтому перейдем сразу к

2 Выдача AlphaFold и ее валидация на структуре PDB

Я запустила белковый бласт, чтобы найти в PDB структуру Вот его выдача:

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	Structure of recombinantly assembled E46K alpha-synuclein fibrils [Homo sapiens]	Homo sapiens	116	116	100%	4e-35	100.00%	140	6UER_A
<input checked="" type="checkbox"/>	Human micelle-bound alpha-synuclein [Homo sapiens]	Homo sapiens	115	115	100%	1e-34	98.41%	140	1XQ8_A
<input checked="" type="checkbox"/>	Structure of alpha-synuclein fibrils [Homo sapiens]	Homo sapiens	114	114	100%	2e-34	98.41%	121	6FLT_A
<input checked="" type="checkbox"/>	cryo-em structure of alpha-synuclein fiber [Homo sapiens]	Homo sapiens	110	110	98%	5e-34	98.39%	63	6A6B_A
<input checked="" type="checkbox"/>	Cryo-EM structure of A53T alpha-synuclein amyloid fibril [Homo sapiens]	Homo sapiens	113	113	100%	7e-34	96.83%	140	6LRO_A
<input checked="" type="checkbox"/>	Chain A_ Alpha-synuclein [Homo sapiens]	Homo sapiens	112	112	100%	9e-34	96.83%	140	7EQF_A
<input checked="" type="checkbox"/>	Cryo-EM structure of alpha-synuclein H50Q Narrow Fibril [Homo sapiens]	Homo sapiens	112	112	100%	2e-33	96.83%	140	6PEO_A
<input checked="" type="checkbox"/>	Cryo-EM structure of phosphorylated Tyr39 a-synuclein amyloid fibril [Homo sapiens]	Homo sapiens	111	111	100%	4e-33	96.83%	140	6LIT_A
<input checked="" type="checkbox"/>	Chain A_ Alpha-synuclein [Homo sapiens]	Homo sapiens	105	105	92%	3e-31	98.28%	100	7LC9_A
<input checked="" type="checkbox"/>	Chain A_ Alpha-synuclein [Homo sapiens]	Homo sapiens	97.4	97.4	85%	1e-28	100.00%	55	6L4S_A
<input checked="" type="checkbox"/>	Crystal structure of human alpha-synuclein (32-57) fused to maltose binding protein (MBP) [Escherichia coli]	Escherichia coli	43.9	43.9	34%	2e-06	95.45%	397	3QZ7_A
<input checked="" type="checkbox"/>	Crystal structure of human alpha-synuclein (58-79) fused to maltose binding protein (MBP) [Escherichia coli]	Escherichia coli	43.5	43.5	34%	3e-06	100.00%	393	3QZ8_A
<input checked="" type="checkbox"/>	Structure of alpha-synuclein in complex with an engineered binding protein [Homo sapiens]	Homo sapiens	39.7	39.7	33%	4e-06	95.24%	22	4BXL_C

Рис. 1: Выдача protein Blast при поиске похожих структур

Итого мне попался кусочек длиной 63 аминокислоты человеческого белка alpha-synuclein, который имеет длину 140.

Все структуры из выдачи бласта, кроме 1XQ8, получены криоэлектронной микроскопией и это различные укладки из бэта-листов, которые полимеризуются в фибриллы.

2.1 Одна молекула

Однако Альфафолд предсказал одну длинную альфа спираль.

Это более или менее соотносится со структурой 1XQ8 - она была получена с помощью ЯМР, и считается, что именно так выглядит нативный белок, ассоциированный с мембранным пузырьком. Считается, что alpha-synuclein прилипает к липидной поверхности мембранного пузырька, который сливается с мембраной и выбрасывает в синаптическую щель нейротрансмиттеры. Также alpha-synuclein нужен как шаперон для сборки SNARE-комплекса, который отвечает за слияние мембран при экзоцитозе.

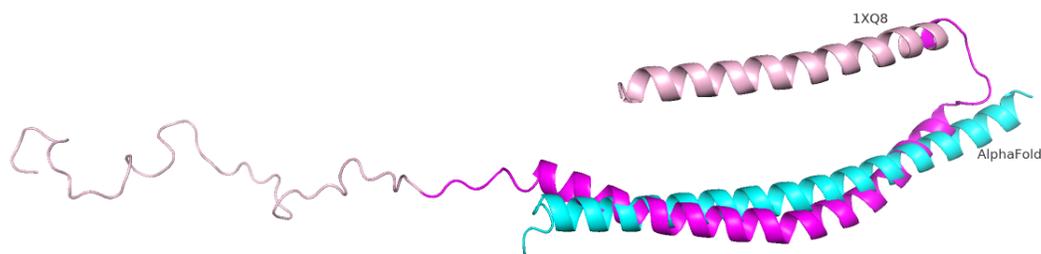


Рис. 2: Розовым цветом показана структура 1XQ8. Ярко розовым отмечен тот участок, который подавался на вход альфа-фолду. Голубым показана выдача альфа-фолда - модель ранга 1.

Как видно на рисунке 1, белок состоит из двух спиралей и подвижного С-конца. При этом соединительная петля между двумя спиральями, как ни странно, достаточно жесткая - она важна при контакте с мембраной и регулируется.

Для предсказания мне был дан ярко-розовый участок, включающий как длинную альфа-спираль, так и краевые не сложенные в структуру участки. Альфа-фолд предсказывает все участки как альфа-спиральные, что, пожалуй, ожидаемо - ведь на вход программе была дана не вся структура. Также видно, что изгиб предсказанной альфа-спирали гораздо менее явственный.

Когда я посмотрела на дополнительные модели, меня ждал приятный сюрприз.

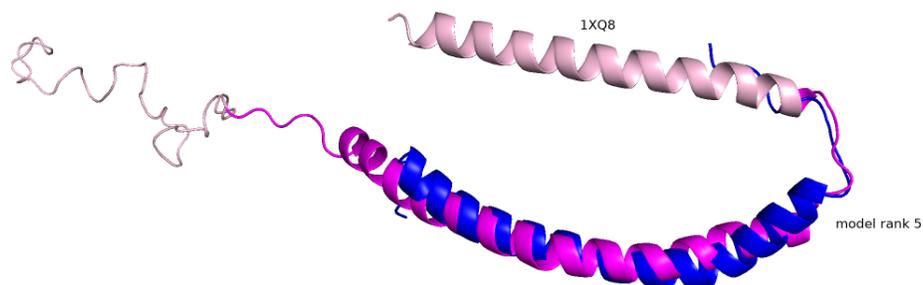


Рис. 3: Розовым цветом показана структура 1XQ8. Ярко розовым отмечен тот участок, который подавался на вход альфа-фолду. Темно синим показана выдача альфа-фолда - модель ранга 5.

Пятой по уверенности моделью оказалась гораздо более похожая на нативную спираль - и изогнутая, и ход петли был предсказан достаточно точно. Так что одну молекулу белка альфа-фолд предсказал неплохо, хотя и не очень уверенно.

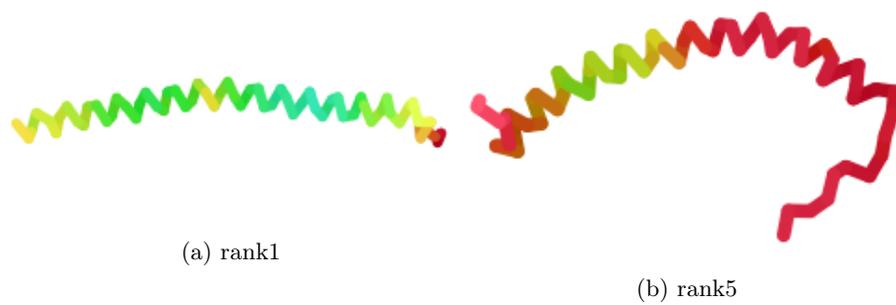


Рис. 4: Раскрашенные по pLDDT результаты выдачи Альфафолд

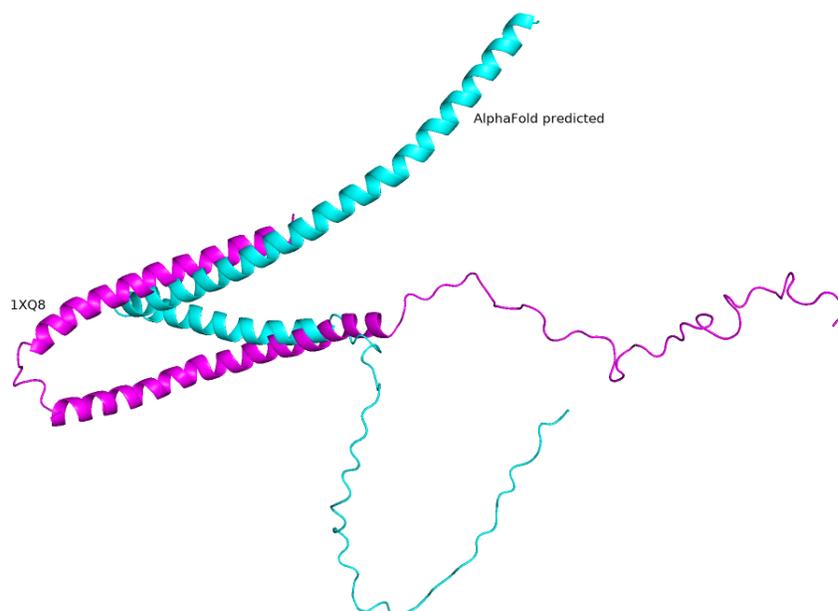


Рис. 5: Розовым цветом показана структура 1XQ8. Голубым показана выдача альфа-фолда - самая похожая на исходный белок модель.

Я попробовала подать Альфафолду всю последовательность из 140 остатков, но, к сожалению с такой задачей он тоже справился не очень - самая похожая структура показана на рисунке. Наверное, это тоже ожидаемо - во-первых, у белка есть длинный неупорядоченный хвост, во-вторых авторы упоминают, что возможно, две спирали могут иногда соединиться и становится одной длинной спиралью - то есть видимо, белок и в клетке достаточно изменчив.

Итого, кажется, что в программе есть bias в сторону предсказания как можно более устойчивых конформаций. Что, на самом деле, более чем оправдано - ведь сложно понять, это предсказание плохое, или просто белок в клетке очень подвижен.

Итого, я бы сказала, что использовать Альфафолд в целях предсказания небольших упорядоченных структур - можно и он здорово работает! Даже правильную альфа-спираль получилось предсказать.

2.2 Агрегация молекул

При полимеризации alpha-synuclein образует амилоидные тяжи. При этом жесткую структуру из бета листов образует как раз участок из 63 аминокислот который был дан мне для предсказания. Надо заметить, что моя последовательность отличалась на 1 аминокислоту от нативного белка, структура которого 1XQ8. Мой белок в точности соответствует однонуклеотидному

мутанту, полимеризованная структура которого - 6UFR. Два тяжа взаимодействуют друг с другом и формируют фибриллу именно вдвоем - важна зона контакта между двумя симметричными агрегатами.

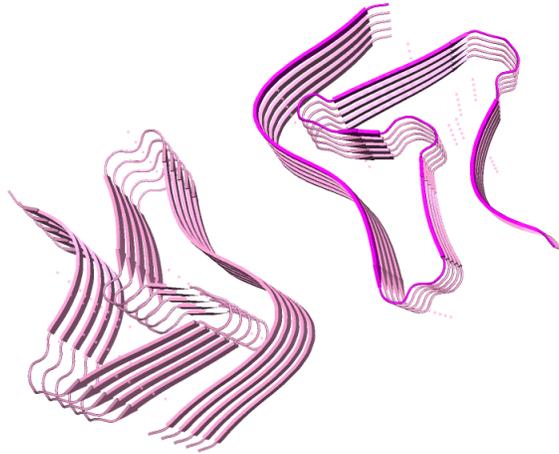


Рис. 6: Розовым цветом показана структура 6UFR, полученная с помощью криоэлектронной микроскопии. Ярко розовым отмечен тот участок, который подавался на вход альфа-фолду.

Альфа фолд что для пяти, что для десяти молекул выдавал пучок из альфа спиралей, при этом они находили друг на друга - я кстати не поняла до конца почему так может происходить (почему в альфаволде из нескольких структур допустимо их накладывание друг на друга) и надо мне будет узнать. В целом вряд ли эта выдача имеет биологический смысл.

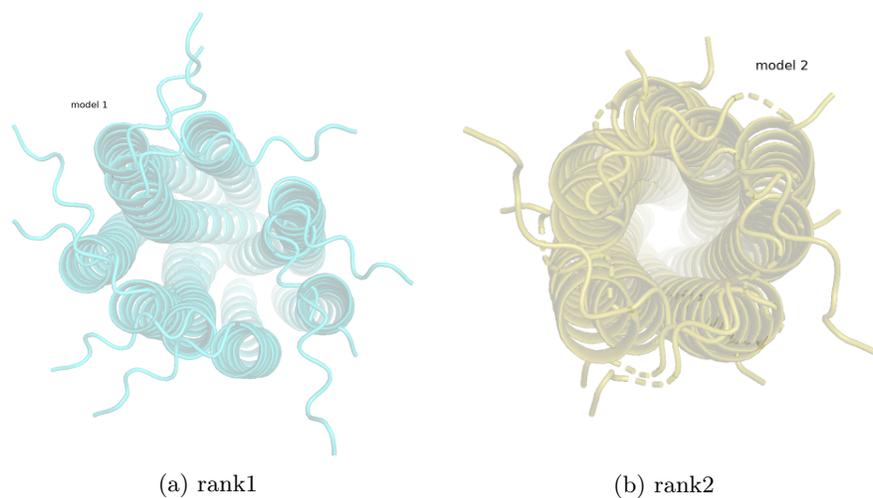


Рис. 7: Выдача AlphaFold при подаче 10 моделей.

Также с задачей предсказать агрегацию многих небольших белковых молекул в фибриллу альфа-фолд не справился.

Зато этому есть логичное, как мне кажется, объяснение.

2.3 Гидрофобность

Зная, что спираль нативного белка взаимодействует с липидной мембраной я выделила гидрофобные остатки. Оказалось, что все эти остатки в самой уверенной модели альфа-фолда для одной молекулы находятся на одной стороне спирали. Поэтому вполне логично предположить, что такие молекулы будут слипаться своими гидрофобными сторонами в пучки - что и предсказал альфа-фолд.

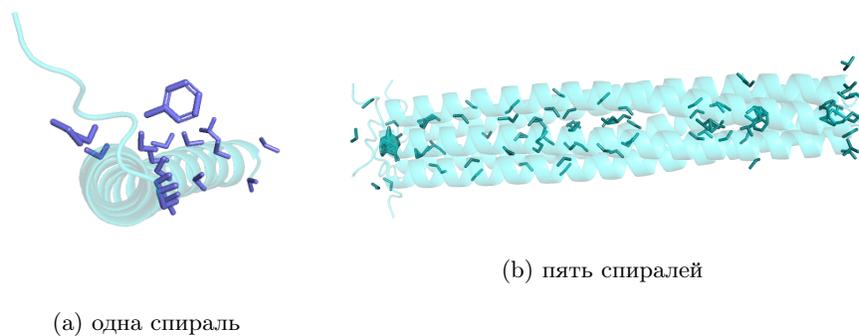


Рис. 8: Предсказания AlphaFold. Шарнирами показаны боковые цепи гидрофобных остатков.

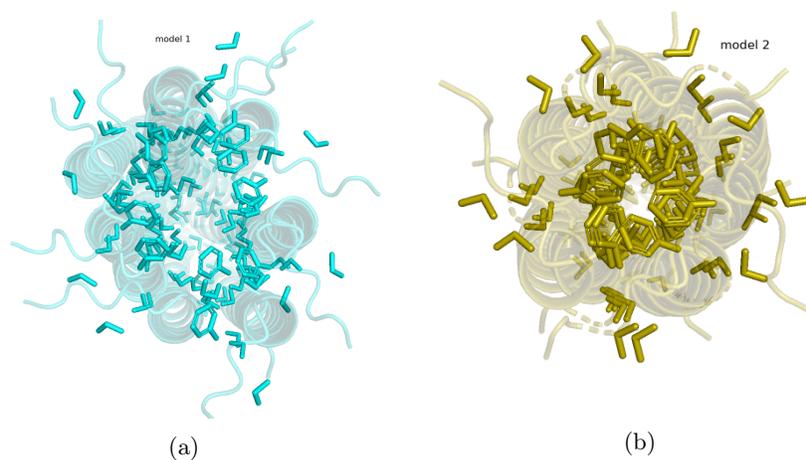


Рис. 9: Предсказания AlphaFold. Шарнирами показаны боковые цепи гидрофобных остатков. Вид пучков с "торца"

Кстати по поводу второй модели для десяти молекул - у меня есть ощущение, что это два наложенных друг на друга пучка по 5 спиралей.

Гидрофобные остатки в настоящей агрегированной структуре тоже расположены внутри фибриллы и в месте контакта фибрилл.

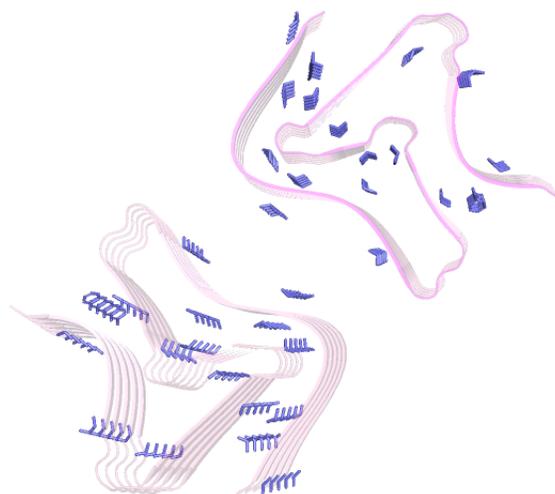


Рис. 10: Структура 6UFR. Синим показаны боковые цепи гидрофобных остатков.

Также фибрилла поддерживается ионными взаимодействиями К-Е как между петлями одного пептида, так и между двумя тяжами .

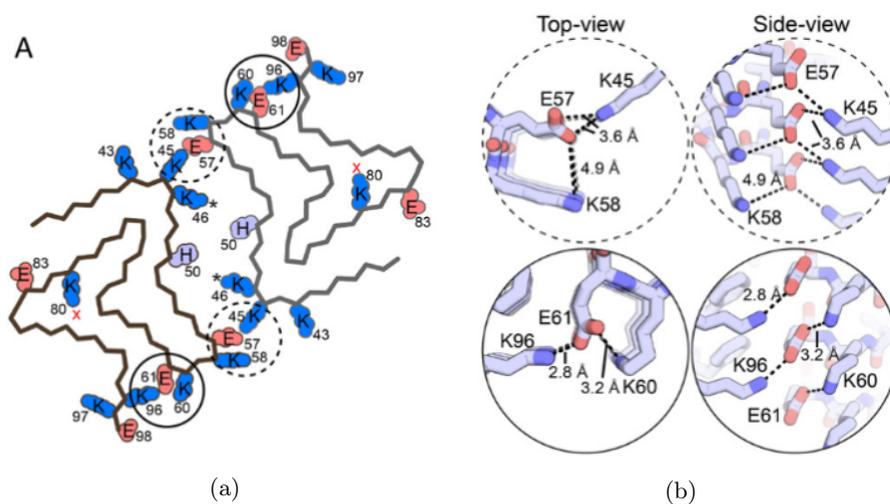


Рис. 11: Изображения из оригинальной статьи (doi:10.1073/pnas.1917914117). Показаны ионные взаимодействия, играющие роль в поддержании амилоидной фибриллы. Приведено изображение для мутанта, который был дан мне для предсказания.

Видимо, альфа-фолду все же тяжело предсказывать структуры, которые сильно стабилизируются именно межмолекулярными взаимодействиями - ведь особенность амилоидных тяжей именно в том, что водородные связи бета-листов - это водородные связи между остовами разных молекул. Также дополнительная стабилизация может происходить, как в случае alpha-synuclein, с помощью ионных взаимодействий между отдельными кластерами молекул.

Итого я думаю, что для таких задач (предсказание полимеризации белковых структур), альфа-фолд не очень подходит. С другой стороны эта задача специфичная - тут должен быть особый упор именно на межмолекулярные взаимодействия и их влияние на вторичную структуру, так что может быть, в обозримом будущем, появится отдельная программа для таких задач.