Киназы антибиотикорезистентности - Exploration #406

Project # 419 (New): Второй металл и протоны в AmiN

Production # 420 (New): Динамики с ATP

Отбор систем-кандидатов (предреакционное состояние)

04.05.2021 14:56 - July Belyaeva

Status:	In Progress	**Start date:**	04.05.2021
Priority:	Normal	**Due date:**	01.11.2021
Assignee:	Alexander Zlobin	**% Done:**	50%
Category:			
Target version:			

Description

Была написана функция, которая обрабатывает pdb-файлы посчитанных 60 систем и считает для них ряд геометрических параметров. Соответствующие результаты далее собирает в датасет, с которым удобнее работать.
Есть некоторые предположения касательно параметров, которые стоит использовать для отбора систем-кандидатов предреакционного состояния - они кратко описаны следующем файле (структура датасета описана там же):
https://docs.google.com/document/d/1FTQt0dk6rON6bW0nhgrurNY7hheYIDP67nZlovwIlh8/edit?usp=sharing .

Хочется увидеть комментарии по поводу сделанных предположений, пожелания касательно рассчитываемых параметров (может быть, что-то стоит убрать, что-то наоборот - стоит добавить).

History

#1 - 20.05.2021 16:17 - Alexander Zlobin

- Tracker changed from Написать to Exploration

- Parent task set to #420

#2 - 16.07.2021 17:04 - Alexander Zlobin

У нас есть 120 систем по 5 реплик на каждую. Это очень много данных, и теперь нам предстоит непростая задача - найти во всем этом смысл, опираясь на визуальный анализ.

Ранжировки заносим в гугл-таблицу ami-systems-2021
https://docs.google.com/spreadsheets/d/105uMLk6Plu7TLwoJwsgkCO-TKR0d-2yIP82xXNVNpyo/edit?usp=sharing

На текущий момент у меня есть заказ на 2 вкладки:
1) RMSD только по тяжелым атомам хвоста трифосфата
2) RMSD только по тяжелым атомам хвоста амикумацина

Также в качестве инструмента для более адекватного запечатления нашей визуальной интуиции в цифрах был предложен dRMSD

#3 - 16.07.2021 19:59 - July Belyaeva

1) RMSD только по тяжелым атомам хвоста трифосфата (AMP-PNP): вкладка amp-pnp-3p-tail,
2) RMSD только по тяжелым атомам хвоста амикумацина (AMP-PNP): вкладка amp-pnp-ami-tail,
3) RMSD только по тяжелым атомам хвоста трифосфата (ATP): вкладка atp-3p-tail,
4) RMSD только по тяжелым атомам хвоста амикумацина (ATP): вкладка atp-ami-tail.

Жирным шрифтом отмечены те системы, которые преобладают в начале/конце отсортированного списка.

#4 - 22.07.2021 18:59 - July Belyaeva

Интересное наблюдение:
Для систем с протонированным аспартатом 222 наблюдается слабая, но значимая положительная корреляция **расстояния** от кислорода глутамата 36 до ближайшего к нему магния и **rms амикумацина** (мне показалось при визуальном просмотре, что это может быть связано). Утверждение не претендует на абсолютную точность, скорее интересное замечание.

#5 - 24.07.2021 14:23 - July Belyaeva

- File index.png added

Был произведен пересчет RMSD для систем в комплексе с AMP-PNP. Вычисление производилось по следующим группам:
1) хвост амикумацина
2) трифосфатный хвост

Как я потом выяснила, изначально неправильно считалось по двум причинам:
- вычислялось среднее значение RMSD по всей траектории, а не только по последней реплике
- некоторые системы в моей рабочей директории содержали устаревшие траектории, скопированные до пересчета реплик-импостеров

Ссылка на обновленную таблицу:
https://docs.google.com/spreadsheets/d/1gPTak0XOyOs2dYXvMV8ZMG5gi8nq8Tj0nRUEgXGrI8o/edit#gid=398140377

Примечательно, что системы, связанные с 1mg располагаются преимущественно в конце рейтинга, построенного по значениям RMSD как для amp-pnp, так и для амикумацина.

Был построен scatter plot, иллюстрирующий распределение систем по значениям RMSD для amp-pnp и амикумацина - прикрепленное изображение ('1mg' : 'фиолетовый', '1mg1li' : 'синий', '2mg' : 'зеленый'). Наиболее плотное распределение, располагающееся ближе к левому нижнему углу, соответствующему меньшим значениям RMSD, характерно для зеленого кластера (системы с 2mg).

Предполагается следующий план (по обсуждению):
- сделать топ-10 систем из списка, ранжированного по минимальному среднему RMSD
- топ-10 систем по минимальному минимальному RMSD (т.е. берем не среднее, а минимальное значение из 5 реплик)
- собрать выборку без повторений из этих двух групп
- пересчитать RMSD для систем с ATP

Топ-10 систем (ранжирование по минимальному среднему RMSD, по хвосту амикумацина):
1) am0-neg_in_2mg 0.073879 '***'
2) am0-ash202_up_2mg 0.077588 '***'
3) ami-neg_up_1mg1li 0.078314 '***'
4) am0-glh36_up_2mg 0.080967 '***'
5) ami-neg_up_2mg 0.082281
6) am0-glh36_in_2mg 0.084450
7) am0-neg_up_2mg 0.085948 '***'
8) am0-glh36_out_1mg1li 0.086256 '***'
9) ami-neg_in_1mg1li 0.086817 '***'
10)am0-glh36_no_2mg 0.089919

Топ-10 систем (ранжирование по минимальному минимальному RMSD, по хвосту амикумацина):
1) am0-neg_up_2mg 0.044662 '***'
2) am0-ash202_up_2mg 0.045050 '***'
3) ami-neg_in_1mg1li 0.045632 '***'
4) am0-glh36_up_2mg 0.045646 '***'

5) am0-glh36_out_1mg 0.051157
6) am0-glh36_out_1mg1li 0.051628 '***'
7) am0-glh36_up_1mg1li 0.054862
8) ami-neg_out_1mg 0.055930
9) ami-neg_up_1mg1li 0.056051 '***'
10)am0-neg_in_2mg 0.056627 '***'

Вопрос: стоит ли в выборку для дальнейшего рассмотрения включать системы, которые встречаются только в группе 2 (не отмечены знаком '***' в группе 2)? Не было ли случайностью то, что одна из пяти реплик таких систем вошла в топ по RMSD?

#6 - 11.10.2021 21:40 - July Belyaeva

Основные технические моменты последних действий достаточно хорошо расписаны в сообщении выше.

Моя рабочая директория по этому проекту: /home/domain/data/julybel/amin-2021/
Рабочая директория Александра по этому проекту: /home/domain/data/zlobin/2021/kinase/

Небольшая литературно-идеологическая подводка:
_В настоящее время все большую актуальность приобретает проблема антибиотикорезистентности. Одна из стратегий, приводящая к возникновению данного явления, — киназо-опосредованное фосфорилирование. В качестве примера фосфотрансферазы с высокой каталитической эффективностью можно привести киназу AmiN, представителя нового подсемейства AmiN-подобных киназ. В недавней работе были сделаны первые шаги по изучению механизма работы комплекса киназы AmiN с субстратом - антибиотиком амикумацином. В ней производилось моделирование поведения одного типа молекулярной системы, состоящей из депротонированных кислотных остатков активного центра, протонированного положительно заряженного амикумацина, депротонированного гамма-фосфата ATP и 1 катиона магния (ami-neg-no-1mg). Также авторами было сделано предположение о том, что расположение ATP в сайте связывания фермента соответствует таковому для негидролизуемого аналога — AMP-PNP. Предположение о присутствии в активном центре одного катиона металла было сделано на основе кристаллографических данных, свидетельствующих о наличии слабой электронной плотности в области типичной локализации второго катиона магния. Такая ЭП может соответствовать ~20% occupancy для Mg2+, Li+, присутствующему в буфере для кристаллизации и просто молекуле воды.

В нашей работе мы стараемся перепроверить существующие данные и улучшить представления об особенностях работы киназы AmiN, реализуя более систематический подход к выбору гипотезы о структуре активного центра исследуемого фермента, а также используя более точную параметризацию молекулярной системы в процессах моделирования. Помимо всего прочего, существуют некоторые неопределенности касательно кристаллографических данных и причинах наличия слабой электронной плотности в области расположения второго катиона металла. Все эти факторы играют важную роль в определении структуры предреакционного состояния и, соответственно, результатах последующих симуляций._

Целью работы было определение того, какое состояние активного центра киназы AmiN наилучшим образом соответствует существующему кристаллу (6SUL), на основании которого делались предположения о стартовом состоянии молекулярной системы в процессе моделирования реакции. Полученные данные помогут сделать выводы о корректности трактовки существующих кристаллографических данных.

Основные результаты:
- Суммарно было собрано 120 систем, для каждой из которых была проведена MD-симуляция.
Почему систем 120: вариантов расположения протона в системе: 8 (E36/D202/D222/AmiA/ATP-up/ATP-in/ATP-out/no H+) * 3 варианта металла в системе (1 Mg2+/1 Li + / no metall), второй Mg2+ присутствовал в любом случае * 5 количество реплик.
- Отсутствие второго катиона противоречит данным рентгенструктурного анализа.
- Вероятно, ранее предложенное предреакционное состояние таковым не является (система ami-neg-no-1mg). Возможно, им является другая система - am0-neg0-in-2mg, эта структура находится на первом месте в топе рейтинга по минимальному среднему RMSD, рассчитанному по хвосту амикумацина для последнего фрейма.
- Наибольшее отклонение от координат кристалла имеют системы, имеющие в своем составе 1 катион магния, они характеризуются наименьшей стабильностью.

- Наибольшей стабильностью и наименьшей дисперсией значений RMSD обладают системы в комплексе с двумя катионами магния.

Дальнейшие планы:
- Как я понимаю, далее будет производиться раунд QM/MM-симуляций над топом-5 (топом-10 ?) систем, ранжированных по минимальному RMSD, рассчитанному для хвоста амикумацина в последнем фрейме.
- Стоит брать ранжирование по минимальным минимальным RMSD (RMSD рассчитывалось для последнего фрейма каждой из 5 реплик, соответствующих той или иной системе, выбиралось минимальное значение) или по минимальным средним RMSD (RMSD рассчитывалось для последнего фрейма каждой из 5 реплик, соответствующих той или иной системе, вычислялось среднее значение)? Мне кажется, второй вариант правильнее, так как полагаю, что минимальное значение, найденное в первом случае, может быть выбросом над общей совокупностью.
- *Мне нужна будет помощь в запуске QM/MM симуляций, либо ссылка на некоторое подобие туториала/мануала, так как ранее с такой задачкой я не сталкивалась*.

Прикрепляю ссылку на текст доклада конференции MCCMB'21: [[https://docs.google.com/document/d/1l3a3yZceix02GDWjRjv_pWD90_U0-L6jjk2yM_AIF38/edit?usp=sharing]]
Прикрепляю ссылку на постер с конференции MCCMB'21: [[https://drive.google.com/file/d/1x7Qycyk34Gd_Vg2_icyeeo84lxHmAKdQ/view?usp=sharing]]

#7 - 27.10.2021 10:48 - Alexander Zlobin

- Due date set to 01.11.2021
- Status changed from New to In Progress
- % Done changed from 0 to 50

1) РМСД по какой группе тебе кажется стоит использовать и почему?

2) Отбери, пожалуйста, топ-5 систем из АМП-ПНП, руководствуясь критерием из 1), и напиши их составы

3) Также вопрос про АТФ - по твоим наблюдениям, насколько лучшесть по критерию 1) совпадает с предреакционностью? Есть ли состояния, не оч по РСМД но неплохие по реакционности? Выпиши их в ответе на этот тикет.

#8 - 04.11.2021 18:19 - July Belyaeva

- File atp-top5-mean-rmsd-superimposition.pse added

Пересчитала rmsd для следующих групп (и для amp-pnp, и для atp):
- хвост ami
- полный ami
- трифосфатный хвост amp-pnp/atp.

Считалось два вида rmsd:
1) среднее rmsd по последнему фрейму (усреднение по пяти репликам системы)
2) минимальное rmsd в последнем фрейме (минимальное по пяти репликам системы).

Посчитанные результаты приведены в таблицах:
- amp-pnp: https://docs.google.com/spreadsheets/d/1195fLdSiI_CWkaMtjsEITyQSKLL0ZymIJl4LepOJ_jI/edit?usp=sharing
- atp: https://docs.google.com/spreadsheets/d/1IOgQMFVWbWq5abKtOHhUoOvZ18VhnSSTvYgajC93bg8/edit?usp=sharing .

Примечательно, что результаты пересчета совпадают с летними результатами (те, что приводились выше, для amp-pnp) - листы 'ami_tail_mean' и 'ami_tail_min' в таблице для amp-pnp, ссылка на которую прикреплена выше.

Вопрос 1

Грубо говоря, на старте реакции наибольший интерес для нас представляет пространственное расположение гамма-фосфата atp и хвоста амикумацина.
Если посмотреть на результаты расчета rmsd, для трех групп (перечислены выше), то мы увидим, что значения rmsd в целом больше для группы ami_full, как в системах с amp-pnp, так и в системах с atp.

На мой взгляд, ранжирование систем по результатам расчета rmsd для целого амикумацина дает достаточно грубую оценку ее стабильности - наибольшее значение имеет именно положение хвоста. В некоторых системах амикумацин будто бы "вылезает" из кармана связывания, но это можно понять и по значениям rmsd его хвоста.

Ранжирование систем по rmsd хвоста atp или amp-pnp также, полагаю, не очень хорошо использовать - нам ведь, по сути, не так важны отклонения альфа-фосфата.
Если посмотреть на рассчитанные значения rmsd для систем, то мы увидим, что трифосфатный хвост несколько более стабильный, нежели хвост амикумацина. Поэтому, стабильность системы можно оценивать только по хвосту амикумацина, так как отклонения его rmsd, фактически, означают нарушение одной из важных 'компонент' оптимального предреакционного состояния - положения хвоста ami.

Вопрос 2

Топ-5 систем с amp-pnp, ранжированных по значениям rmsd (среднее по последнему фрейму), рассчитанного для хвоста амикумацина:
1) am0-neg_in_2mg | 0.074
2) am0-ash202_up_2mg | 0.078
3) ami-neg_up_1mg1li | 0.078
4) am0-glh36_up_2mg | 0.081
5) ami-neg_up_2mg | 0.082

Топ-5 систем с amp-pnp, ранжированных по значениям rmsd (минимальное по последнему фрейму), рассчитанного для хвоста амикумацина:
1) am0-ash202_up_2mg | 0.045
2) am0-neg_up_2mg | 0.045
3) ami-neg_in_1mg1li | 0.046
4) am0-glh36_up_2mg | 0.046
5) am0-glh36_out_1mg | 0.051

В данном случае, полагаю, нам интересны значения среднего по последнему фрейму rmsd, так как это отражение картины между всеми репликами. То-есть, малое значение rmsd последнего фрейма может быть некоторым баесом по сравнению с трендом, который наблюдается в остальных траекториях.

Вопрос 3

Топ-5 систем с atp, ранжированных по значениям rmsd (среднее по последнему фрейму), рассчитанного для хвоста амикумацина:
1) am0-glh36_out_1mg | 0.070
2) ami-neg_up_1mg1li | 0.073
3) am0-neg_out_2mg | 0.076
4) ami-neg_in_2mg | 0.080
5) ami-neg_in_1mg1li | 0.092

Топ-5 систем с atp, ранжированных по значениям rmsd (минимальное по последнему фрейму), рассчитанного для хвоста амикумацина:
1) am0-neg_out_1mg | 0.040
2) am0-glh36_in_2mg | 0.044
3) am0-neg_no_1mg1li | 0.044

4) am0-neg_up_1mg | 0.044
5) am0-glh36_out_1mg | 0.048

Примечательно, что в этом топе все системы, кроме первой, характеризуются отсутствием протона на отобранных для рассмотрения аминокислотных остатках.
Прикрепляю PyMol-сессию **atp-top5-mean-rmsd-superimposition.pse**, в которой представлены 5 реплик каждой системы с atp из топа-5 с mean rmsd.

Реплики каждой системы, которые, как мне кажется, лучшим образом (из 5) могут соответствовать стартовому состоянию:
1) am0-glh36_out_1mg | am0-glh36_out_1mg_3 (угол O3B-PG-O5 = 156.8, расстояние PG-O5 (ami) = 3.7, расстояние O5 (ami) - OD2 (Asp202) = 3.4)
2) ami-neg_up_1mg1li | ami-neg_up_1mg1li_4 (угол O3B-PG-O5 = 144.6, расстояние PG-O5 (ami) = 3.7, расстояние O5 (ami) - OD2 (Asp202) = 3.2)
3) am0-neg_out_2mg | мне все не особо нравятся по геометрии
4) ami-neg_in_2mg | во всех системах протон на O1G залипает на кислород Glu36
5) ami-neg_in_1mg1li | мне все не особо нравятся по геометрии

На основании этого просмотра добавлю новый способ расчета rmsd - попробую сделать расчет rmsd по хвосту амикумацина + атомам O3B-PG-O1G-O2G-O3G, либо N3B-PG-O1G-O2G-O3G, составлю также рейтинг и посмотрю системы с atp.

#9 - 04.11.2021 22:36 - July Belyaeva

- File atp_top5_pg-ami_tail.pse added

Топ-5 систем по среднему rmsd (группа pg-ami_tail) с atp:
1) am0-glh36_in_2mg | 0.061
2) ami-neg_up_1mg1li | 0.061
3) am0-ash202_up_1mg1li | 0.062
4) am0-glh36_up_1mg1li | 0.064
5) am0-neg_no_1mg1li | 0.064

Результаты расчета rmsd по группе pg-ami_tail добавлены на соответствующие листы в таблицы, ссылки на которые в заметке выше.
Посмотрим, насколько эти системы похожи на предреакционное состояние:
1) am0-glh36_in_2mg | am0-ash202_up_1mg1li_4 (угол O3B-PG-O5 = 164.6, расстояние PG-O5 (ami) = 3.6, расстояние O5 (ami) - OD2 (Asp202) = 2.7)
2) ami-neg_up_1mg1li | am0-glh36_in_2mg_4 (угол O3B-PG-O5 = 151.2, расстояние PG-O5 (ami) = 3.7, расстояние O5 (ami) - OD2 (Asp202) = 2.7)
3) am0-ash202_up_1mg1li | мне все не особо нравятся по геометрии, элементы далеко друг от друга
4) *am0-glh36_up_1mg1li * | am0-neg_no_1mg1li_4 (угол O3B-PG-O5 = 173.1, расстояние PG-O5 (ami) = 4.0, расстояние O5 (ami) - OD2 (Asp202) = 3.3)
5) am0-neg_no_1mg1li | мне все не особо нравятся по геометрии

Прикрепляю соответствующую PyMol-сессию: **atp_top5_pg-ami_tail.pse**

_
Будет хорошо, если ты сможешь оценить рассуждения по степени адекватности.

Files

index.png	17.5 KB	24.07.2021	July Belyaeva
atp-top5-mean-rmsd-superimposition.pse	38.3 MB	04.11.2021	July Belyaeva
atp_top5_pg-ami_tail.pse	38.3 MB	04.11.2021	July Belyaeva