Плазмин - Production #448

Project # 446 (New): Байндинг-активация плазмина

Стартовые системы

26.08.2021 11:23 - Alexander Zlobin

Status:	In Progress	**Start date:**	26.08.2021
Priority:	Normal	**Due date:**	01.11.2021
Assignee:	Egor Ilin	**% Done:**	0%
Category:			
Target version:			

Description

Необходимо
1) Собрать все кристаллы плазмина и плазминогена на данный момент
2) Отсортировать по тому, что это за форма и с чем связана

Далее подготовить системы для громакса для amber19sb
1) плазминоген-стрептокиназа
2) плазминоген
3) плазмин-стрептокиназа
4) плазмин

Внимательно отнестись к протонированиям, таутомерам, воде. Все эти фичи должны быть максимально унифицированными между системами.

History

#1 - 03.09.2021 10:57 - Alexander Zlobin

Комплекс плазмина со стрептокиназой: 1BML (NB: Это неактивный искусственный мутант по кат. серину, S>A, надо вернуть обратно)
Плазминоген: 1QRZ

Но это не отменяет задачи собрать и выровнять пространственно все, что есть. По сути готовое выравнивание уже есть на PDBeKB: https://www.ebi.ac.uk/pdbe/pdbe-kb/proteins/P00747

#2 - 07.09.2021 17:58 - Egor Ilin

Сделал стартовую систему. Лежит в папке home/domain/egor22366/plasmin/1_model_solv_ions.gro. В качетве mdp файла взял файл со страницы туториала по громаксу по лизоцину.

#3 - 09.09.2021 15:27 - Alexander Zlobin

- File strepto_555d2.result.zip added

Результаты альфафолда. Забавно, что он фолдит стрептокиназу прямо как в кристалле, при том, что плазмин отсутствует. Биас обучающей выборки или фича? Не понятно сходу.
Зато удобно, что накладывается хорошо. Не удобно, что одна петля очень хреново воссоздана, просто полукруг в пространстве. В связи с этим задача: надо прошарить литературу на предмет того, не разрезается ли сама стрептокиназа где-либо по ходу дела? Вдруг этой петли и не должно быть.

#4 - 18.09.2021 14:15 - Egor Ilin

/home/domain/egor22366/plasmin/plasmin_solv_ions готовый файл плазмина. Во время добавления ионов я взял параметр neutral и не указывал концентрацию. Добавились только хлоры. Это критично?

#5 - 20.09.2021 11:14 - Egor Ilin

Делаю плазминоген. Нашел структуры 4duu, разрешение паршивое, но что есть. Там есть вопрос про остатки в области 255/346, 437/455. Вгрузил электронную плотность, вроде ничего там нет, но есть сомнения. I NEED SOME HELP!

#6 - 20.09.2021 18:55 - Alexander Zlobin

По поводу ионов - ну, конечно depends. Если мы подозреваем, что конформация белка может меняться от ионной силы раствора, то такие вещи чекать обязательно. Плюс даже если мы не подозреваем, это могут подозревать рецензенты. Сильно проще взять 0.15M ибо это физраствор, т.е. плазма крови, а наши белки таки там и работают.

Плазминоген: Ну я же указал, есть 1QRZ. Super его на плазмин в твоей системе, удаляй плазмин, делай PDB, где вместо строчек атомов плазмина строчки атомов плазминогена. Можно все это сделать через PyMol, Например у тебя плазмин это chain A в некой системе plasmin (например это половина ячейки из 1BML). И ты еще загрузил 1QRZ. super (1qrz and chain A), (plasmin and chain A). create new_system, (plasmin and not chain A) or (1qrz and chain A). save plasminogen_system.pdb , new_system.

#7 - 30.09.2021 18:09 - Egor Ilin

- File Redmin.zip added

Прикрепляю архив и опись некоторых статей в списке. Последняя из описанных рассматривает 1BML, смотрели на поведение стрептокиназы. Достану не sci-hub версию, а то без картинок.
https://docs.google.com/document/d/1x_lpBKSHGpljWXHzQ7cPk2Gzvmd8zGTtrLbxO9ZhPfg/edit?usp=sharing

#8 - 04.10.2021 01:32 - Egor Ilin

/home/domain/egor22366/plasmin/plasmin_sk/plasmin_sk_2/pls_sk_out.gro
Путь к комплексу плазминоген и стрептокиназы. Брал киназу из модели, которую делал ты. Пока по раскручивание стрептокиназы ничего не нашел.

#9 - 05.10.2021 13:37 - Alexander Zlobin

Рекомендации по запуску

Запускать на xwing или gpu
Если на gpu, смотреть через nvidia-smi, какая gpu сейчас свободна. Давать одну из свободных по номеру любому запуску mdrun через флаг -gpu_id <id>

Все расчеты кроме минимизации сильно лучше считаются, если добавить флаг -update gpu при запуске mdrun

Обязательно задавать количество рангов и потоков. Флаги -ntmpi 1 -ntomp <N> при запуске mdrun. N можно варьировать от 4 до 8, смотреть, как счет идет быстрее.

Громакс:
source /home/domain/data/prog/gromacs-2021.1-pm-nompi/bin/GMXRC.bash
export LD_LIBRARY_PATH=/home/domain/data/prog/plumed-2.7.1-nompi/lib

После чего в этом окне терминала можно просто вбивать gmx

#10 - 10.10.2021 23:05 - Egor Ilin

/home/domain/egor22366/plasmin/plasmin_sk/plasmin_sk_2/npt.gro
/home/domain/egor22366/plasmin/plasmin_sk/plasmin_sk_2/md_0_1.xtc
Продинамил одну 1 ns. Получил 102 скана. Система плазминогена с стрептокиназой.

#11 - 18.10.2021 13:59 - Alexander Zlobin

- Due date changed from 04.10.2021 to 01.11.2021

Так, супер-важное изменение. В будущем нам надо будет прочекать, какая структура стабильнее -
а) где N-конец активаторной петли в плазмине протонирован, а аспартат, с которым он взаимодействует, анион и
б) нейтральный N-конец, аспартат тоже нейтральный

Нейтральные концы отсутствуют в полях Amber, но есть в поле charmm36m. Оно есть у меня, в /home/domain/data/zlobin/gmxlib. Это надо заэкспортить как GMXLIB. Тогда при pdb2gmx если не прописать выбор поля, то будет промпт командной строки, и выбрать там charmm36

Для плазминогена-то это не важно, конечно, но если мы будем в этом поле описывать плазмин, то чтобы все было четко и консистентно, плазминоген надо тоже в этом поле моделировать.

Если ты застрял с водой, то назначь окно в полчаса, когда мы сядем и я покажу свои действия и логику рассуждения.

#12 - 04.11.2021 16:29 - Egor Ilin

Сделал модели белков(плазмин, плазминоген и их комплексы со стрептокиназой). Пока ионы не добавлял, нужен опытный взгляд и можно в продакшн. Думаю сейчас написать скриптик для воды, а то реально отнимает немало времени, на крайняк, сделаб скриптик быстрого удаление найденой воды из файла громакса и топологии. Жду обратной связи(папка pls_sl-коплекс плазминогена со стрептокиназой).

#13 - 05.11.2021 13:48 - Alexander Zlobin

Что-то не так с complex_w.gro

gromacsplugin) Error reading atom 162228 from file, file does not match format
 ObjectMolecule: plugin 'gro' failed to read atoms.

Понял, что - ты не изменил число атомов. Оно значится как 162239, а весь файл - это 162230 строк. Атомов в нем 162227

Есть много проблемных залитых вод. Например 27881, 27827, 13880, ...

Есть проблемная "кристаллическая" вода - судя по всему это конфликт воды из кристалла и модели стрептокиназы из альфафолд. Например 642, 707, 657

Так же есть проблемы с этим в диких клэшах между плазмином и стрепток: место контакта остатков 317 и 505 имеет катастрофически малые шансы минимизироваться во что-то нормальное.

Вердикт: не стоило нам просто брать модель из АФ2 и просто подсовывать. Надо подработать локальные части, сделать химеру из кристалла и АФ2. Там, где лучше АФ2, взять его, там, где важная зона контакта, оставить кристалл. Это работа не простая, но и объект у нас не простой.

Метакомментарий: судя по pls_sk_box, получается очень большая ячейка. Думаю можно поджать

#14 - 06.11.2021 23:09 - Egor Ilin

Обновил модельку в pls_sl. Вроде должно быть ок. Возможно убил не всю ненужную воду.

#15 - 08.11.2021 13:10 - Egor Ilin

/home/domain/egor22366/pls_sl/iter_2/complex_w.gro

#16 - 10.11.2021 01:14 - Egor Ilin

Продинамил плазминоген со стептокиназой. Проучилось 13.7 нс. Ставил на 20, но отключился терминал (не знаю как это связано, но все же). /home/domain/egor22366/pls_sl/iter_2/md_0_1.tpr.
Собрал и остальные модели. /home/domain/egor22366/plasmin_sk,plasmin,plasminogen/complex_w.gro

#17 - 13.11.2021 10:46 - Alexander Zlobin

- Status changed from New to In Progress

Используй tmux, чтобы создавать набор терминалов на сервере, которые будут работать, когда ты уже закроешь все со своей стороны. Сейчас ты будешь генерить много данных - надо уже перейти работать в папки по пути /home/domain/data/...(что-либо).

Внутри npt увидел это: On 1 MPI rank, each using 64 OpenMP threads. Для любого расчета надо указывать ntomp, не только для продакшн. Ты пропустил флаг -update gpu, тесты числа потоков стоит переделать с ним

При этом тест md_with_flag_only смысла не имеет - мы не можем себе позволить 64 потока на расчет. Задача была потестить наличие или отсутствие флагов -nb gpu -bonded gpu -pme gpu -pmefft gpu (всех разом).

#18 - 13.11.2021 10:57 - Alexander Zlobin

update gpu дает больше всего прирост. Его надо использовать **всегда**, если есть гпу.

Остальные флаги включены. Риппер, 3080
-ntomp 4: 139 ns/day
-ntomp 6: 143 ns/day
-ntomp 8: 147 ns/day
-ntomp 10: 148 ns/day
-ntomp 12: 149 ns/day

Остальные флаги выключены.
-ntomp 4: 138 ns/day
-ntomp 6: 147 ns/day
-ntomp 8: 151 ns/day
-ntomp 10: 153 ns/day
-ntomp 12: 154 ns/day

Видно, что в обоих случаях быстро выходим на плато. Оптимально выходит 8 потоков, без дополнительных флагов. Поставь в таком сетапе наносекунд на 500, потом посмотрим, что получится, обсудим.

#19 - 27.11.2021 13:08 - Egor Ilin

Посмотрел первые 50 наносекунд комплекса плазминогена со стрептокиназой. Он разваливается((. Файл
/home/domain/data/ilin/pls_sl/traj0-50_fit.xtc

#20 - 14.12.2021 01:07 - Egor Ilin

Выпал на неделю, не трогал системы. Решил действовать по твоему алгоритму(выравнивал по боковым цепям). Телепортации не исчезли. Кста, что там с сириусной статьей.
Файлы лежат в
/home/domain/data/ilin/pls_sl/traj0-50_fit.xtc
/home/domain/data/ilin/pls_sl/traj0-50_2_fit.xtc
/home/domain/data/ilin/pls_sl/traj_end_fit.xtc
/home/domain/data/ilin/pls_sl/traj_end_2_fit.xtc

#21 - 14.12.2021 01:23 - Alexander Zlobin

Такими темпами мы никуда не уедем. Либо пересмотри свои приоритеты, либо я забираю проект, он с 2022 года становится важен для лаборатории.

Разобраться, как работать с траекториями - задача посильная и не требует каких-то сакральных знаний.

Для статьи по ДФПазе нужно полностью пересмотреть пространство связанных поз ДФП в ДФПазе, мы что-то сильно упускаем. Сейчас не до этого, гранты, под которые могла бы пойти эта работа, кончились.

Files

strepto_555d2.result.zip	1.16 MB	09.09.2021	Alexander Zlobin
Redmin.zip	34.7 MB	30.09.2021	Egor Ilin