

Практикум 4. Валидация

Задание 1.

В этом задании рассматривалась структура токсина, выделенного из яда китайского щитомордника (*Agkistrodon acutus*). Укус этой змеи вызывает сильное кровотечение и приводит к некрозам [1]. Вызвано это присутствием в яде акутогемолизина – фосфолипазы A2, гидролизующей sn-1 ацильную группу фосфолипидов. В большинстве фосфолипаз A2 каталитическая активность обусловлена связыванием 49 аспарагиновой кислоты иона Ca^{2+} , однако в акутогемолизинах эта аминокислота заменена на лизин, но активность при этом не теряется [2].

Представленный в структуре PDB ID 1MC2 акутогемолизин не обладает каталитической активностью, так как 102 фенилаланин препятствует попаданию субстрата в активный центр фермента. Все вышеописанные аминокислотные остатки хорошо покрываются электронной плотностью и в целом структура имеет высокое разрешение, которое составляет 0.85 \AA . $R_{\text{free}} = 0.121$, $R_{\text{obs}} = 0.095$ – это хорошие показатели, разница между ними небольшая, значит модель удачно оптимизирована. Но все-таки я бы выделила 10 триптофан, который чуть менее хорошо покрывается ЭП (рис. 1). N-концевая α -спираль, в которую входит этот остаток, участвует в образовании димеров, необходимых для осуществления гемолитической активности, при этом у других фосфолипаз A1 на 10 месте обычно находятся остатки Leu или Phe [2]. В этой же структуре триптофан¹⁰ имеет 2 альт-лока – радикал этого остатка расположен неоднозначно, при окраске структуры по B-фактору он окрашен в белый цвет, что дополнительно подтверждает подвижность ароматических колец (рис. 2). В PDB есть альтернативная структура для этого токсина, при выравнивании двух структур в PyMol видно, что триптофан у второй структуры (1MG6) соответствует альт-локу A в первой структуре (рис. 3), но так как у второй структуры разрешение ниже (и также отсутствуют файлы для визуализации ЭП), нельзя с уверенностью говорить, что такое расположение радикала наиболее стабильно. Таким образом, неактивность рассмотренного акутогемолизина может быть дополнительно обоснована неспособностью образовывать димеры из-за замены в N-концевой α -спирали и подвижностью замененного остатка. Структура 1MC2 подходит для изучения мономера белка, но каталитически активные димеры на ней изучать не получится.

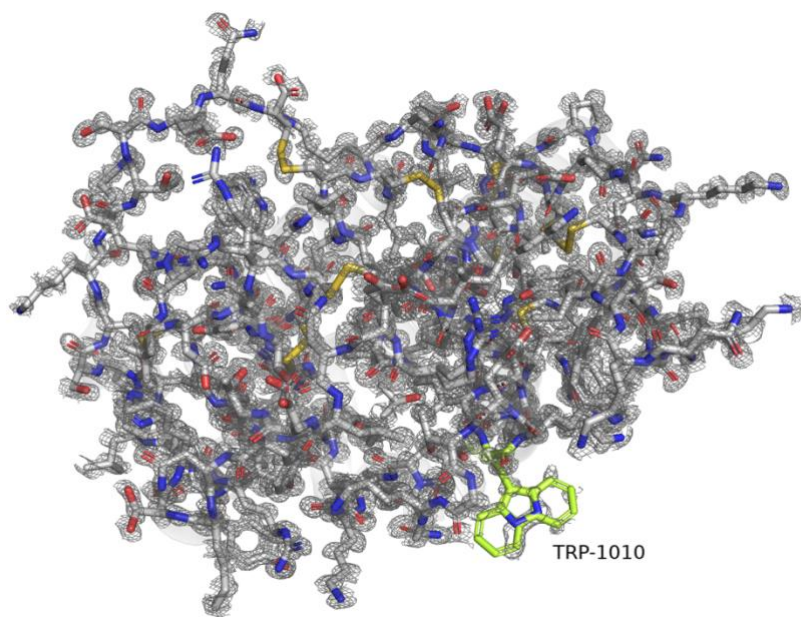


Рис. 1 Визуализация ЭП акутогемолизина 1MC2. Триптофан, который я посчитала не очень хорошо покрытым ЭП, окрашен в салатный цвет

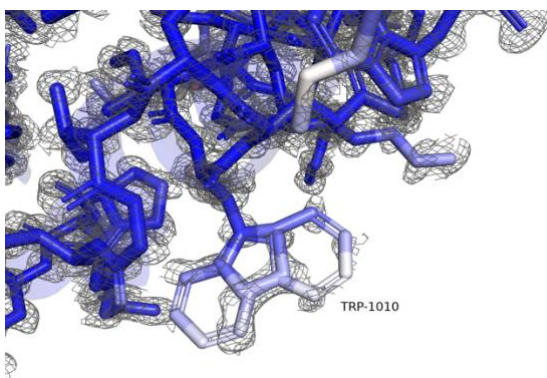


Рис. 2 Триптофан¹⁰, окрашенный по B-фактору. Белый цвет свидетельствует об относительной подвижности радикала

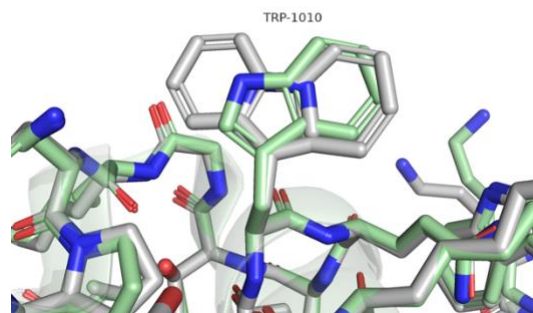


Рис. 3 Выравнивание триптофана¹⁰ структур 1MC2 (серый) и 1MG6 (зеленый). Положение триптофана в структуре 1MG6 ближе к аль-локу A в 1MC2
[PyMol сессия](#)

Задание 2.

В этом задании я рассматривала аминокислотные остатки, маргинальность которых была задокументирована в полном отчете по валидации структуры 1MC2 в PDB и также была замечена платформой для валидации 3D моделей MolProbity. Качество структуры хорошее, сильно маргинальных остатков по данным PDB нет (рис. 4), красной точкой на рисунке отмечены аминокислоты, которые не очень хорошо покрыты ЭП. Среди них и рассмотренный ранее триптофан. Для углубления в маргинальность остатков я выбрала триптофан¹⁰¹⁰, цистеин¹¹³⁴, аргинин¹⁰⁴³, аспарагин¹⁰⁸⁸, а на основе выдачи MolProbity аргинин¹¹⁰⁷ (аутлаер по длине связи C-N) и глутамин¹¹⁰⁸ (аутлаер по углу между углеродами). Выбранные аминокислоты, их функции в белке и особенности расшифровки структуры объединены в табл. 1.

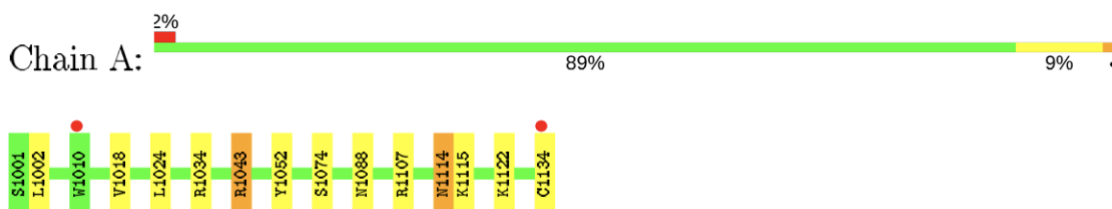


Рис. 4 Данные о качестве отдельных аминокислот структуры 1MC2. Красными точками указаны плохо покрытые ЭП остатки. Источник: PDB, 1MC2

Табл. 1 Выбранные аминокислотные остатки структуры 1MC2 и некоторые их характеристики

Аминокислота	Номер	Функционально важный остаток?[2]	В чем проблема (MolProbity)	Соответствие ЭП	Коммент
Триптофан	1010	Препятствует образованию кат. активных димеров	Угол CE2-CD2-CE3, Не очень хорошо покрыта ЭП	2 альт-лока, покрытие среднее	Рис. 5А
Аргинин	1043	-	Длина связи CZ--NH2: 5.2 σ , угол CD-NE-CZ: 33.0 σ	2 альт-лока, у одного плохо покрыт атом NH1	Рис. 5В
Аспарагин	1088	-	MolProbity: clear evidence for flip	Хорошее покрытие ЭП	Рис. 5С
Аргинин	1107	Солевой мостик удерживает димер в кристалле	Длина связи CZ--NH2: 4.3 σ	2 альт-лока	Рис. 5D
Глутамин	1108		Угол CB-CG-CD: 5.8 σ	2 альт-лока	Рис. 5D
Цистеин	1134	С-концевой, образует S-S связь с соседней альфа-спиралью	Длина связи C--O: 11.8 σ , Не очень хорошо покрыт ЭП	ОН – группа покрыта не очень хорошо, покрытие «ошибочной» C—O хорошее	Рис. 5Е

Триптофан 1010

Оба положения боковой цепи триптофана никак не конфликтуют с окружением, при этом углы CE2-CD2-CE3 в этих альт-локах не сильно отличаются от размера угла правильного шестиугольника - 120° (рис. 5). Скорее всего, маргинальность этого остатка связана с плохим покрытием ЭП, но в любом случае функциональная роль аминокислоты сохраняется.

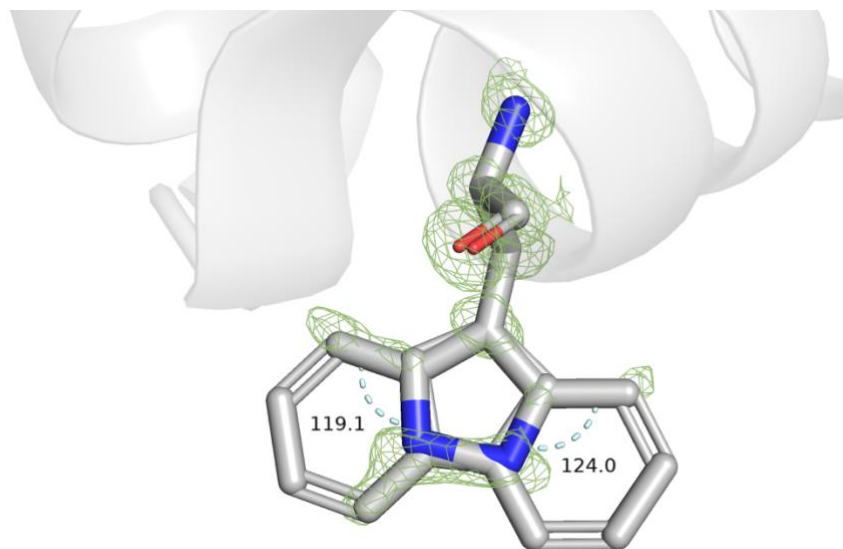


Рис. 5 Триптофан¹⁰¹⁰ и «неправильные» углы в кольце. ЭП, покрывающая остаток продемонстрирована в виде зеленого mesh визуализация на уровне подрезки 1

Аргинин 1043

Позиция альт-локов аргинина стабилизируется остатками окружения. В случае альт-лока А радикал стабилизирован лишь молекулами растворителя и судить о биологической правильности его расположения сложно (рис. 7, слева). Альт-лок В образует солевой мостик с глутамином соседней цепи, и также поддерживается связями с растворителем. Судя по всему, наиболее «правильная» и стабильная в природе конформация все-таки В, хотя покрытие ЭП тут не такое хорошее (рис. 7, справа). Рассмотрим размер угла CD-NE-CZ и длину связи CZ--NH2 этих конформеров: А – угол 128.7°, связь 1.3 Å; В – угол 170.6°, связь 1.4 Å. В целом, наиболее распространенный размер угла X4 аргинина приближен к 90° или 180°, таким образом размер угла в конформации В чуть более похож на правду, однако альт-лок А тоже имеет место быть – никакие элементы структуры этому не противоречат [3]. Длина связи, исходя из ЭП, определена верно.

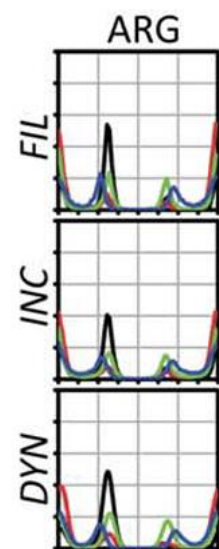


Рис. 6 Распределения значений двугранных углов для аргинина из датасета, проанализированного в статье, синим цветом показана кривая для угла X4 [3]

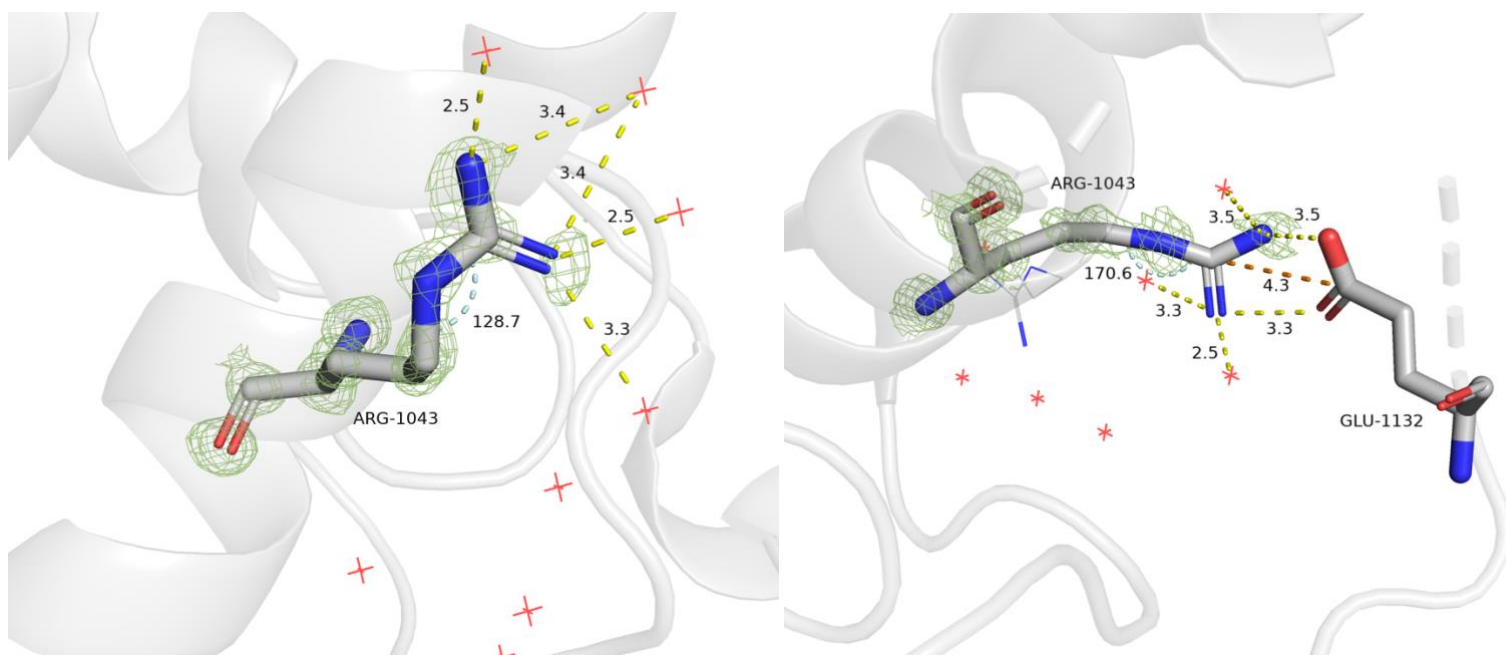


Рис. 7 Анализ взаимодействий, которые стабилизируют боковой радикал аргинина¹⁰⁴³ в случае конформера А (слева) и конформера В (справа). Желтым пунктиром обозначены водородные взаимодействия, солевой мостик – оранжевым
PyMol сессия: <https://kodom.fbb.msu.ru/~nr.burmistrova/term7/pr4/1043> Burmistrova.pse

Аспарагин 1088

Для аспарагина хорошо разрешена ЭП, атомы остова поддерживаются водородными связями с растворителем и с глицином¹⁰⁸⁶, однако радикал не так хорошо взаимодействует с окружением: я обнаружила водородное взаимодействие аминокетильной группы с водой, но карбонильная группа в данной репрезентации структуры никак не поддерживается. Я считаю, что водородное взаимодействие (рис. 8, синий пунктир) может поддерживать аминокетильную группу в случае перевернутого радикала (по рекомендациям MolProbity), при этом карбонильная группа так же будет поддерживаться водородным взаимодействием с растворителем.

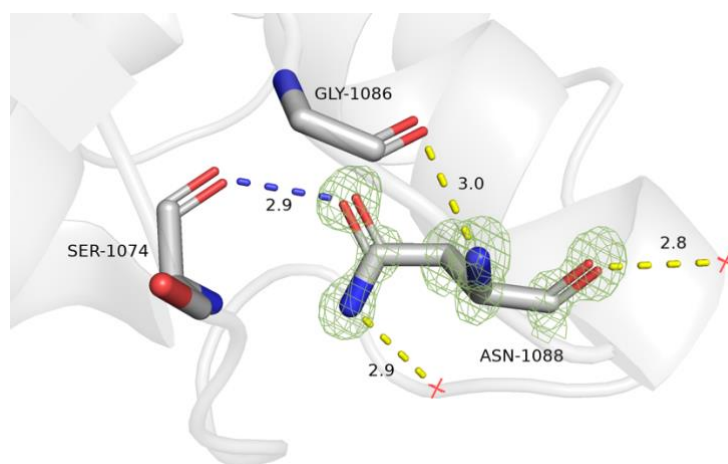


Рис. 8 Взаимодействия 1088 аспарагина с окружением, желтым пунктиром обозначены водородные взаимодействия, синим – водородная связь, которая стабилизирует аминокетильную группу в перевернутом аспарагине (flip)

Аргинин 1107 и глутамин 1108

В случае альт-лока А аминокислоты формируют прочные солевые мостики, удерживающие димер в кристалле. Эта конформация также стабилизируется водородными связями с окружением (рис. 9А). Длина связи CZ--NH₂ аргинина в этом случае составляет 1.3 Å. Угол СВ-СG-СD глутамина 122.4°. При рассмотрении структуры, содержащей альт-локи В обеих аминокислот заметно, что солевые мостики уже не образуются, но димер все-равно удерживается связями между аргинином и глутамином, но теперь уже менее прочными, водородными (рис. 9В). Радикалы дополнительно стабилизируются взаимодействиями с растворителем. Длина связи CZ--NH₂ аргинина также 1.3 Å. Угол СВ-СG-СD глутамина 105.6°.

Отклонения как в размере угла глутамина, так и в длине С—N связи аргинина не такие большие, перекрытий с окружением не наблюдается и в целом эти значения находятся в пределах допустимого, так что маргинальность этих остатков, вероятнее всего, связана с особенностями структуры. Однако СВ атом глутамина в обоих альт-локах не покрыт ЭП, поэтому отклонения в размере угла также могут быть связаны с неправильным расположением этого атома, но на функциональность остатков это не влияет.

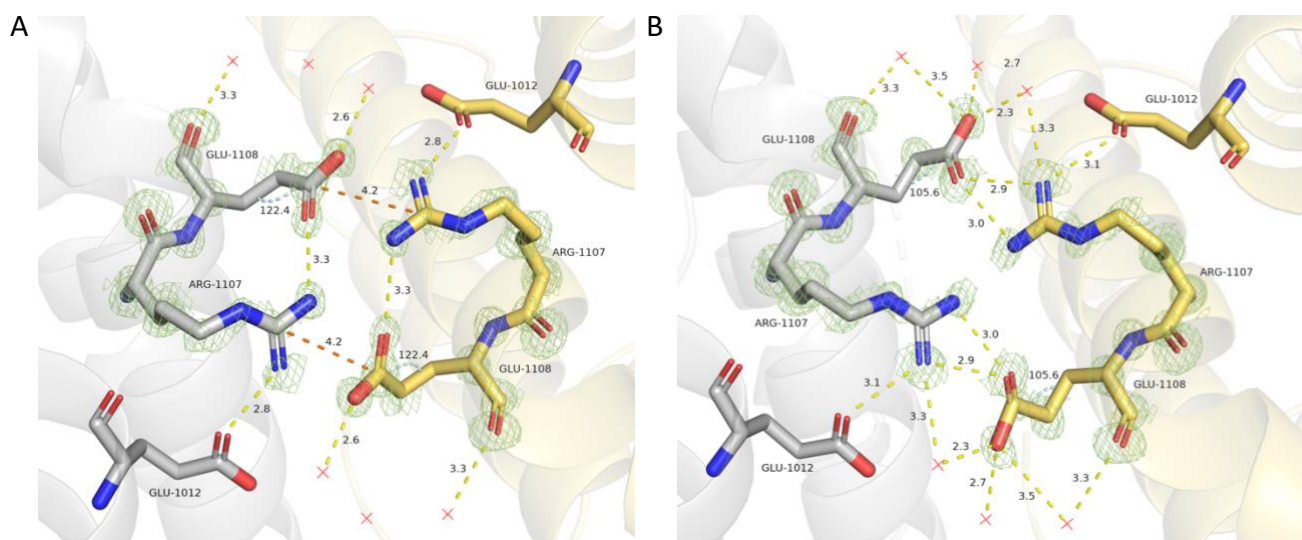


Рис. 9 Взаимодействия аргинина¹¹⁰⁷ и глутамин¹¹⁰⁸ с окружением. Димер поддерживается солевыми мостиками (оранжевый пунктир) и водородными взаимодействиями (желтый). А – обе аминокислоты находятся в конформации А, В – в конформации В

Цистеин 1134

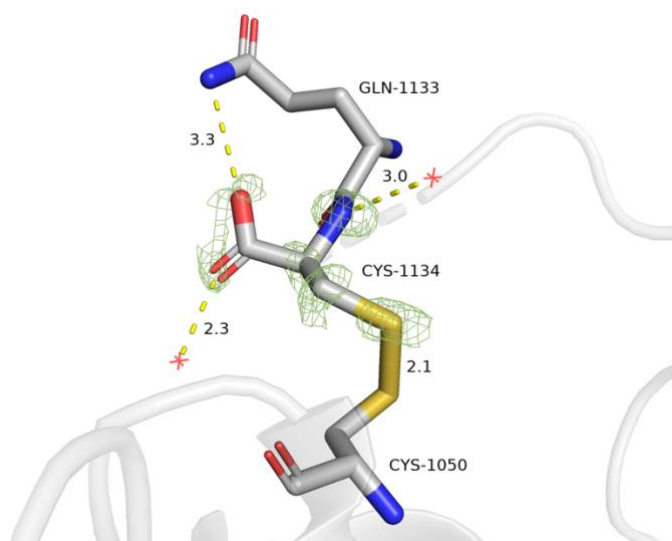


Рис. 10 Взаимодействия, стабилизирующие структуру цистеина¹¹³⁴. Водородные связи обозначены желтым пунктиром

Радикал хорошо стабилизирован S-S связью. Довольно короткая длина C—OH связи (1 Å) может означать, что C-конец белка депротонирован за счет водородных связей и являться особенностью структуры [4]. Отсутствие протона также подтверждается более-менее равномерным распределением ЭП между кислородами COOH-группы.

Задание 3. PDB-REDO

В задании 2 мне удалось определить правильность расположения атомов в структуре на основе их связей с окружением, однако наука не стоит на месте и теперь кристаллографам не обязательно проверять структуры друг друга на особенности маргинальности остатков. Посмотрим на «исправленную» структуру в PDB-REDO.

Судя по R-факторам (рис. 11), модель в PDB-REDO чуть менее хорошо сходится с экспериментальными данными (R чуть выше, чем у оригинальной структуры), разница между R и R-free изменилась несущественно – модель не переоптимизирована.

Validation metrics from PDB-REDO		
	Original	PDB-REDO
Crystallographic refinement		
R	0.1050	0.1112
R-free	0.1192	0.1275

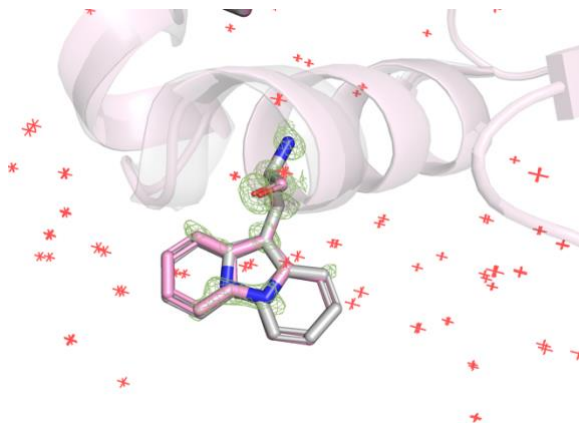
Рис. 11 Сравнение R-фактора и R-free в выдаче PDB-REDO

Significant structural changes	
Description	Count
Rotamers changed	4
Side chains flipped	0
Waters removed	21
Peptides flipped	0
Chiralities fixed	2
Residues fitting density better	47
Residues fitting density worse	34

Рис. 12 Результаты структурных изменений в PDB-REDO

В целом, структура 1mc2 изначально была отличной, поэтому в PDB-REDO не так много существенно измененных остатков. Интересно, что некоторые остатки стали хуже соответствовать ЭП (рис. 12). PDB-REDO все же переусердствовала?. Посмотрим, что у программы получилось на примере рассмотренных ранее маргиналов.

Триптофан 1010



Положение триптофана в PDB-REDO не изменилось (рис. 13). Видимо, углы альта-локов определены верно.

Рис. 13 Сравнение исходной структуры *Ims2* (серый) и исправленной в PDB-REDO (розовый)

Аргинин 1043

В случае аргинина некоторые изменения претерпел альта-лок В (рис. 14, слева). Угол стал меньше (был 170.6°), но при этом атомы NE и CD стали хуже вписаны в ЭП (рис. 14, справа). Скорее всего, такое положение радикала действительно является особенностью структуры, либо кристаллографы допустили ошибку при построении функции ЭП, что крайне маловероятно (я им полностью доверяю и мне нравится их структура).

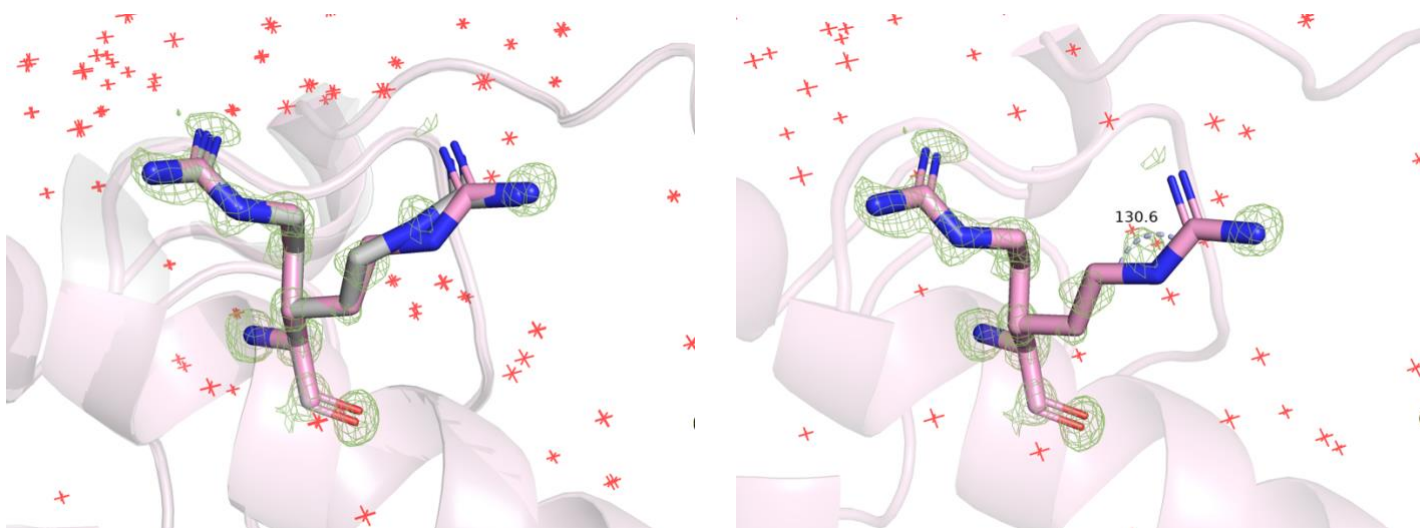


Рис. 14 Сравнение структуры бокового радикала Arg43. Справа – выравнивание *Ims2* (серый) со структурой из PDB-REDO (розовый). Слева показано, что исправленная аминокислота в случае альта-лока В плохо вписывается в ЭП

Аспарагин 1088

Аспарагин, как и рекомендовала MolProbity оказался перевернут в PDB-REDO. Это действительно более правильное положение радикала – теперь он способен образовывать несколько водородных связей, которые хорошо его стабилизируют (рис. 15). Помимо этого, PDB-REDO немного уменьшила угол $O-C-NH_2$: его значение в структуре из PDB составляло 125.8° . В целом, это изменение несущественно и атомы все равно вписываются в ЭП. Скорее всего, такое положение атомов уменьшает расстояние до доноров и акцепторов водородной связи, делая взаимодействия с окружением более прочными.

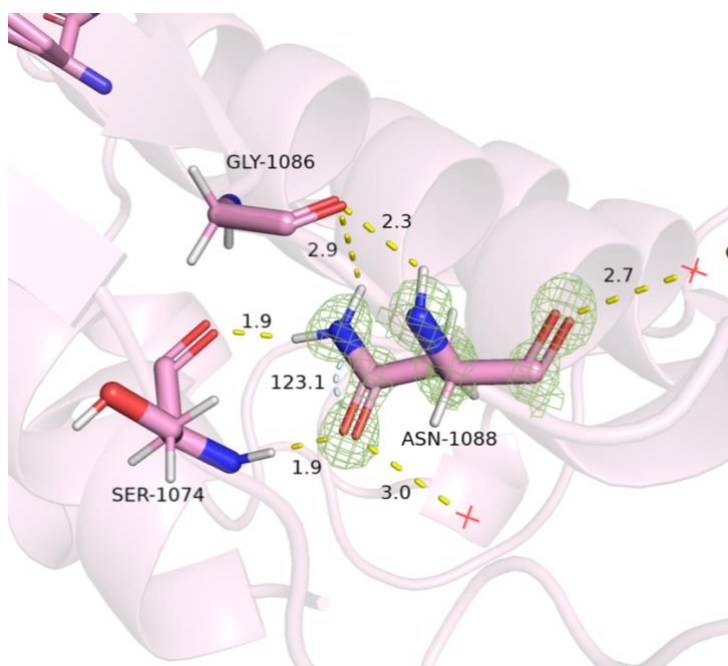


Рис. 15 Структура Asn88 из PDB-REDO и его взаимодействия с окружением. Желтым пунктиром показаны водородные взаимодействия.

[PyMol сессия: https://kodomofbb.msu.ru/~nr.burmistrova/term7/pr4/asn_1088_redo_Burmistrova.pse](https://kodomofbb.msu.ru/~nr.burmistrova/term7/pr4/asn_1088_redo_Burmistrova.pse)

Аргинин 1107 и глутамин 1108

Положение атомов тут существенно не изменилось, координаты определены верно (рис. 16).

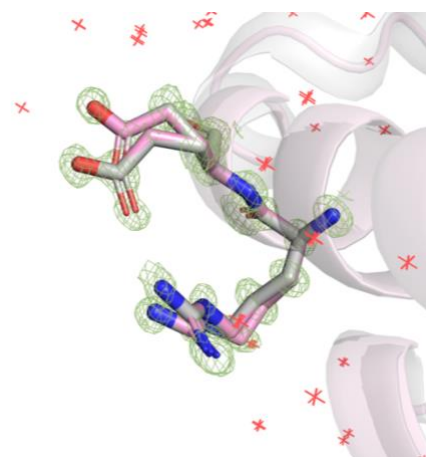


Рис. 16 Выравнивание Arg107 и Glu108 структуры *Irc2* (серый) и исправленной в PDB-REDO (розовый)

Цистеин 1134

Положение цистеина по данным PDB-REDO изменилось (рис. 17, слева): уменьшение торсионного угла пси привело к повороту свободного С-конца, при этом, немного сместился из своей ЭП атом СА. Длина связи С—ОН изменилась незначительно (на 0.1Å), также программа убрала ближайшую к -COOH – группе цистеина воду, из-за чего водородная связь стала немного длиннее (рис. 17, справа).

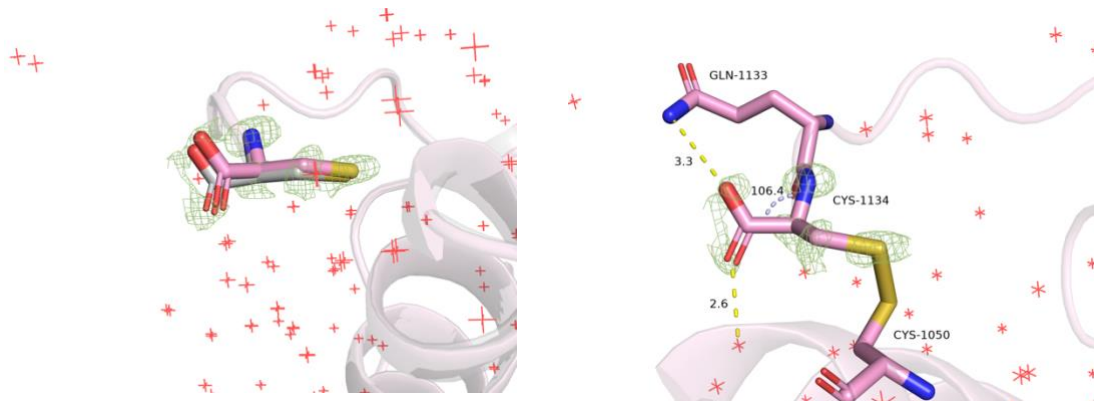


Рис. 17 Слева – выравнивание Cys134 структуры *Ims2* (серый) и исправленной в PDB-REDO (розовый).
Справа – анализ взаимодействий исправленного цистеина с окружением

Вывод: ~~не стоит давать PDB-REDO и так хорошую модель.~~ Видимо, из-за высокого качества структуры и усердной работы кристаллографов, улучшить структуру не удалось (а если и удалось, то не существенно). В исходной структуре атомы рассмотренных аминокислот лучше соответствовали своей электронной плотности, чем на выходе программы. Скорее всего, маргинальность остатков вызвана особенностями конкретной структуры.

PyMol сессия для 3 задания:

https://kodom.fbb.msu.ru/~nr.burmistrova/term7/pr4/PDB-REDO_Burmistrova.pse

Литература:

- 1) Wikipedia contributors. (2023, November 11). Deinagkistrodon. In Wikipedia, The Free Encyclopedia.
- 2) Liu Q, Huang Q, Teng M, Weeks CM, Jelsch C, Zhang R, Niu L. The crystal structure of a novel, inactive, lysine 49 PLA2 from *Agkistrodon acutus* venom: an ultrahigh resolution, AB initio structure determination. *J Biol Chem.* 2003 Oct 17;278(42):41400-8. doi: 10.1074/jbc.M305210200. Epub 2003 Jul 19. PMID: 12871974.
- 3) Scouras AD, Daggett V. The Dynameomics rotamer library: amino acid side chain conformations and dynamics from comprehensive molecular dynamics simulations in water. *Protein Sci.* 2011 Feb;20(2):341-52. doi: 10.1002/pro.565. PMID: 21280126; PMCID: PMC3048419.
- 4) Lin J, Pozharski E, Wilson MA. Short Carboxylic Acid-Carboxylate Hydrogen Bonds Can Have Fully Localized Protons. *Biochemistry.* 2017 Jan 17;56(2):391-402. doi: 10.1021/acs.biochem.6b00906. Epub 2016 Dec 30. PMID: 27989121; PMCID: PMC5805389.