

Валидация

Задание 1.

Мне была дана PDB-структура [1YWR](#). Это модель структуры комплекса домена MAP киназы P38 с моноциклическим пиразолоном, выступающим в роли ингибитора. MAPK14 (Mitogen-activated protein kinase 14) – одна из четырех MAP киназ P38. Она играет роль в клеточном ответе, вызванном внеклеточным стимулом (провоспалительными цитокинами или клеточным стрессом) и ведущим к прямой активации транскрипционных факторов [1, 2].

Авторы статьи [2] заинтересованы в эффективности ингибиторов на всем каскаде реакций, ведущем к синтезу цитокина TNF- α . Домен MAP киназы P38 – одна из критических мишеней ингибирования. Целью расшифровки трёхмерной структуры домена в исследовании было выяснить особенности взаимодействия MAP киназы P38 с ингибитором.

Не помешает проанализировать данные о качестве модели из [отчёта](#) о валидации с PDB.

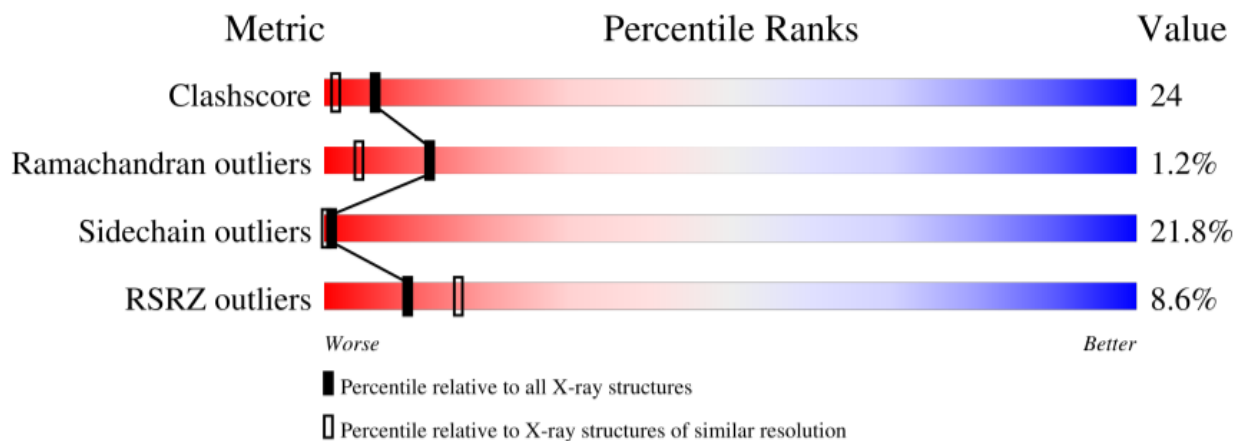


Рис. 1. Графический отчет с wwPDB X-ray Structure Validation Report

По Рис. 1 видно, что значения метрик для данной структуры в целом ниже, чем у большинства структур, что говорит о ее плохом качестве.

Разрешение 1.95 Å, полнота 98.7% (данные депозиторов). Разрешение не супер хорошее, зато полнота довольно высока.

R-фактор 0.221, R-free отсутствует (данные депозиторов), что уже начинает беспокоить, потому что нельзя понять, есть ли переоптимизация (вообще, переоптимизации не наблюдается, если разница между R-фактором и R-free меньше 10%).

Остатков, RSRZ которых оказался > 2, в модели 8.6%.

Clashscore 24, что лучше, чем у ~5% всех PCA-моделей в PDB и у ~1% моделей с похожим разрешением.

Из 338 остатков 4 (1.2%) – аутлайеры по карте Рамачандрана, это больше, чем у 96% структур с похожим разрешением.

65 остатков (22%) имеют отличные от ротамеров значения углов вращения вокруг боковых цепей (χ_1 - χ_4): больше чем у всех структур со схожим разрешением.

Аминокислотные остатки белка были покрашены по количеству геометрических критериев, по которым они являются аутлайерами (зеленым - 0, желтым - 1, оранжевым - 2, красным ≥ 3 , серым - остатки, не вошедшие в модель, красные точки над остатками и красная полоса над общей линией - остатки, не вписавшиеся в электронную плотность, у которых $RSRZ > 2$).

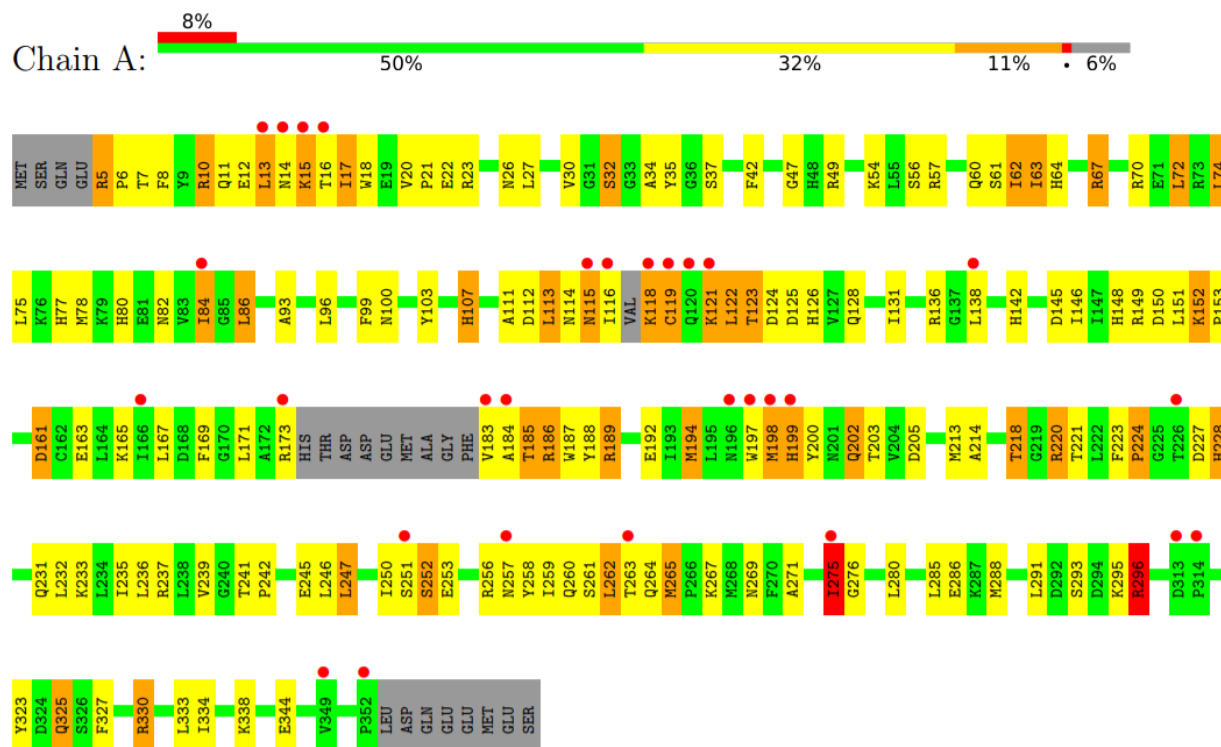


Рис. 2. Количество аутлайеров по критериям

В модели не имеется (хотя, скорее всего, не приведено) атомов, для которых нет подтверждающей их положение ЭП.

Однако в отчете оговорено три расхождения по остаткам между смоделированной и референсной (UniProt ID P47811) последовательностями. Два из них – специально введенные мутации, одно является конфликтом (в положении 263 вместо аланина нашли треонин). На Рис. 3 приведено изображение «конфликтного» остатка Thr-263 с электронной плотностью. Данный остаток находится близко к поверхности белка на петле, что скорее всего, обуславливает высокую подвижность участка белка вблизи остатка и плохую его расшифровку.

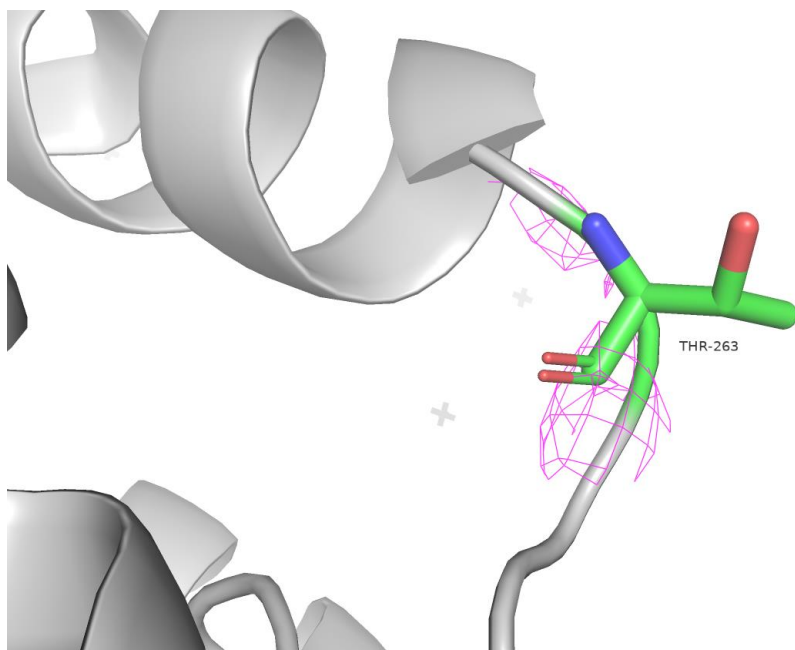


Рис. 3. Электронная плотность Thr-263 на 1.5 уровне подрезки (carve=2)

Задание 2.

Рассмотрела пять примеров остатков, маргинальных по разным критериям. Для анализа были выбраны данные MolProbity и PDB.

В качестве первого маргинального остатка рассмотрела Arg-296. Судя по отчету с PDB, он является маргиналом по ротамерам. При пороге $|Z|=5$ один из его углов (CD-NE-CZ) равен 141.51° вместо идеального 123.60° . Кроме этого, второй из его углов (CG-CD-NE) равен 126.70° вместо идеального 111.80° . По Рис. 4 видно, что атомы CG и CD даже на уровне подрезки не полностью вписаны в электронную плотность. Возможно, положение этих атомов немного неправильное, что привело к аутлайерам.

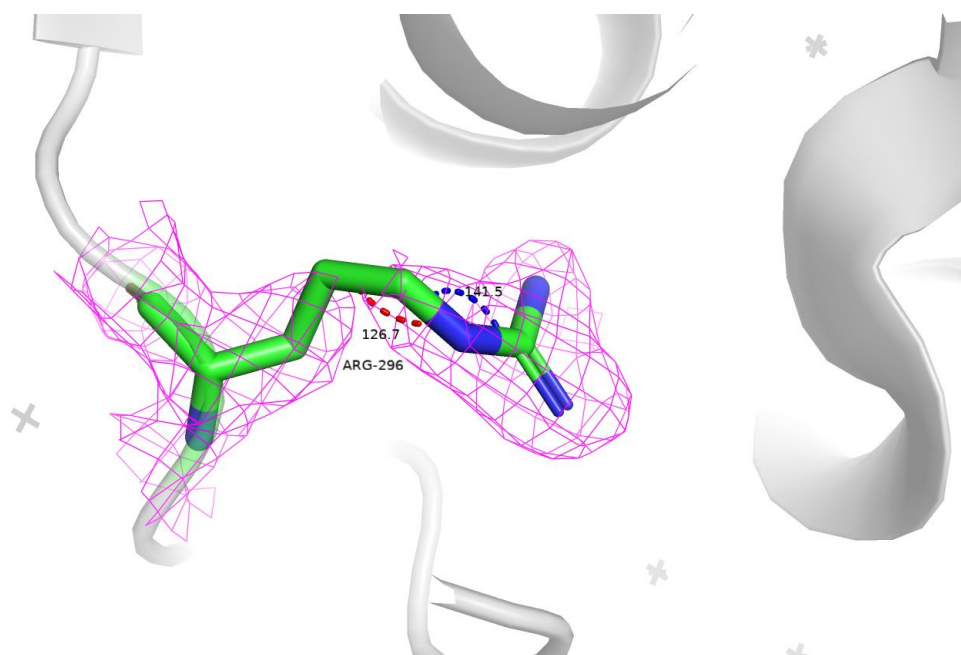


Рис. 4. Электронная плотность Arg-296 – маргинала по ротамерам - на уровне подрезки 2 (carve=2)

Кроме того, водород HG2 остатка Arg-296 слишком сильно (на 0.68 \AA , по данным MolProbity) перекрывается с водородом HB остатка Thr-203 (см. Рис. 5).

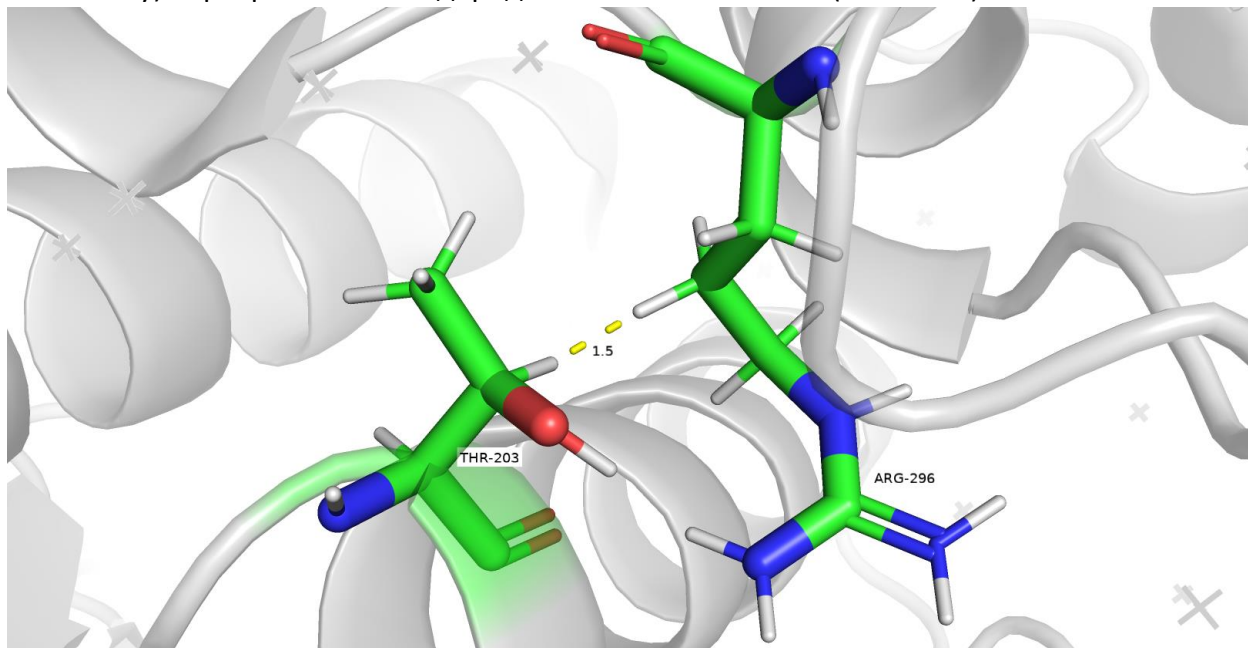


Рис. 5. Маргиналы по критерию clashcore

Еще один маргинальный остаток – Ile-275 – я выбрала по данным MolProbity. Данный остаток находится на петле на поверхности белка (см. Рис. 6). Остаток маргинален по торсионным углам. Если взглянуть на облако электронной плотности, соответствующее этому остатку, можно увидеть, что электронная плотность практически полностью не покрывает остаток Ile-275, что может говорить о причине появления аутлайеров. Торсионные углы ϕ и ψ равны 16.0 и 83.0 , соответственно (см. Рис. 6).

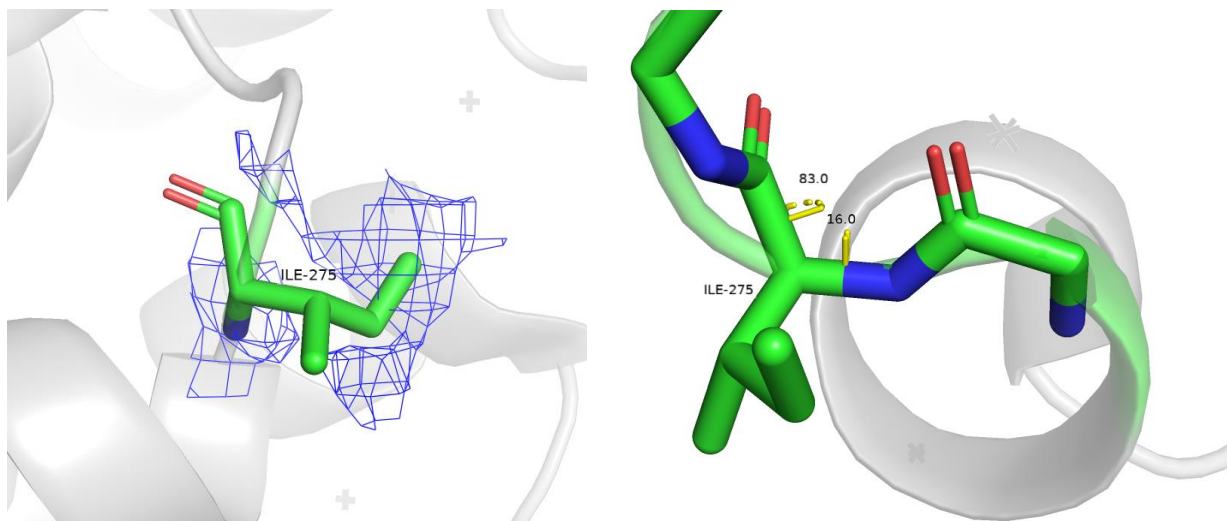


Рис. 6. слева: Электронная плотность остатка Ile-275, уровень подрезки 1, carve=2; справа: Торсионные углы ϕ и ψ остатка Ile-275

Остаток по ротамерам Leu-74 выбирала по данным отчета PDB (см. Рис. 7). Угол между атомами CA-CB-CG равен 129.5° вместо идеального 115.3° . Если посмотреть на электронную плотность остатка, можно заметить, что CG-атом не очень хорошо в нее вписался. Возможно, если его чуть-чуть «подвинуть», мы можем добиться лучшего угла CA-CB-CG.

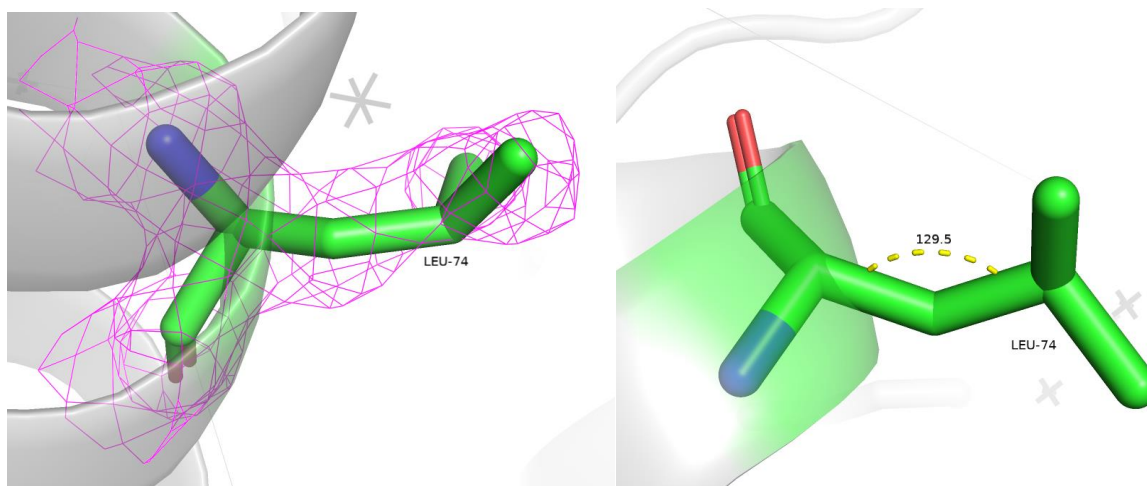


Рис. 7. слева: Электронная плотность остатка Leu-74, уровень подрезки 1.5, carve=2; справа: Угол CA-CB-CG остатка Leu-74

Последний выбранный мною маргинальный остаток по ротамерам – Arg-136 – был взят из отчета PDB. Угол NE-CZ-NH2 оказался равен 117.49° вместо идеального 120.3° (см. Рис. 8).

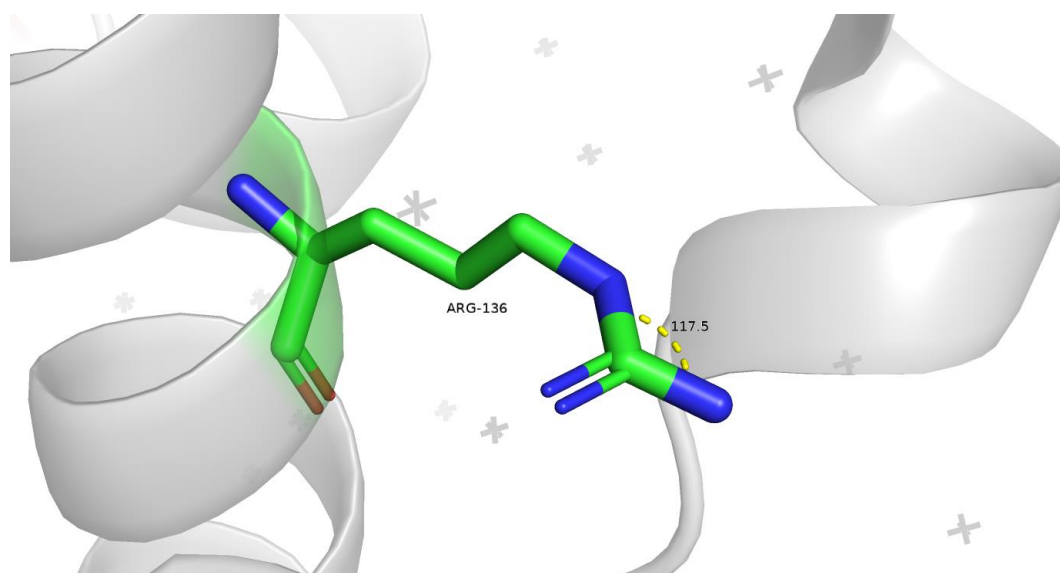


Рис. 8. Угол NE-CZ-NH2 маргинального остатка по ротамерам Arg-136

Ко всему прочему, проверила модель структуры на остатки, требующие разворота с помощью MolProbity. Таких остатков оказалось 11 штук (см. список ниже). Они находились как на поверхности белка, так и во внутренней его части, в том числе вблизи лиганда (см. Рис. 9). После регенерации модели с исправленными разворотами у структуры улучшились некоторые показатели: устранилось несколько клешей (в том числе с водой), немного увеличилось количество хороших ротамеров, чуть уменьшилось количество аутлайеров по ротамерам.

- A 60 GLN
- A 64 HIS
- A 80 HIS
- A 114 ASN
- A 115 ASN
- A 142 HIS
- A 148 HIS
- A 155 ASN
- A 231 GLN
- A 257 ASN
- A 264 GLN

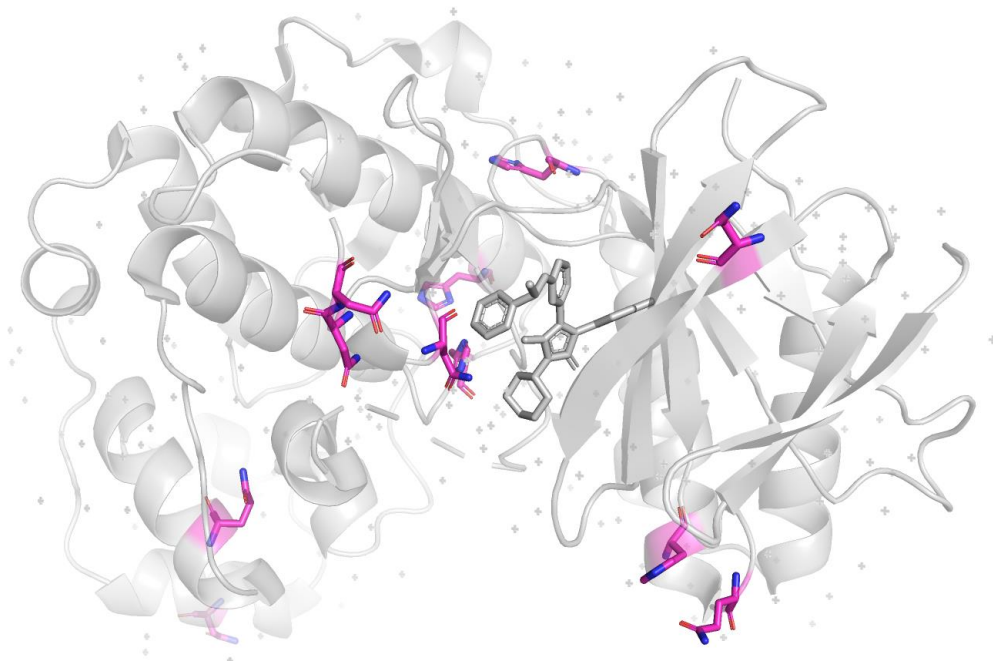


Рис. 9. Маргиналы, требующие разворота

Задание 3.

Судя по общим метрикам качества: разрешению, clashscore, аутлайерам по ротамерам и по карте Рамачандрана, данная мне структура далеко не самая хорошая. Вдобавок, в данной структуре имеются маргиналы в окружении лиганда (судя по сводной таблице MolProbity, это, в основном, клеши и маргиналы по ротамерам), что сильно затрудняет изучение взаимодействий белка с лигандом. Скорее всего, данную мне структуру не следует использовать для этой цели.

Задание 4.

В данном задании предлагалось улучшить модель белка с помощью сервиса PDB-Redo.

PDB-Redo улучшил R и R-free факторы. Изначально они равнялись 0,2227 и 0,2630, соответственно (замечу, что R-free исходной модели сервис предположил сам). После обработки в PDB-Redo они стали равны 0,1972 и 0,2499, соответственно. Значение R фактора стало меньше 20, что можно считать показателем хорошей модели.

По данным MolProbity, у переделанной модели clashscore упал с 24.12 до 6.34, количество аутлайеров по ротамерам снизилось с 63 до 13, «плохих углов» стало меньше (было 21, стало 2). Маргиналов по карте Рамачандрана немного прибавилось (с 4 в исходной модели до 7 в переделанной).

All-Atom Contacts	Clashscore, all atoms:	24.12	19 th percentile* (N=821, 1.95Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	63	21.14% Goal: <0.3%
	Favored rotamers	207	69.46% Goal: >98%
	Ramachandran outliers	4	1.20% Goal: <0.05%
	Ramachandran favored	314	94.58% Goal: >98%
	Rama distribution Z-score	-2.53 ± 0.41	Goal: abs(Z score) < 2
	MolProbity score [^]	3.25	4 th percentile* (N=13349, 1.95Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00% Goal: 0
	Bad bonds:	0 / 2791	0.00% Goal: 0%
Peptide Omegas	Bad angles:	21 / 3786	0.55% Goal: <0.1%
	Cis Prolines:	0 / 17	0.00% Expected: ≤1 per chain, or ≤5%
Additional validations	Chiral volume outliers	0/424	
	Waters with clashes	15/196	7.65% See UnDowser table for details

All-Atom Contacts	Clashscore, all atoms:	6.34	93 rd percentile* (N=773, 1.90Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	13	4.35% Goal: <0.3%
	Favored rotamers	267	89.30% Goal: >98%
	Ramachandran outliers	7	2.11% Goal: <0.05%
	Ramachandran favored	316	95.18% Goal: >98%
	Rama distribution Z-score	-0.57 ± 0.44	Goal: abs(Z score) < 2
	MolProbity score [^]	2.17	54 th percentile* (N=12147, 1.90Å ± 0.25Å)
	Cβ deviations >0.25Å	1	0.31% Goal: 0
	Bad bonds:	0 / 2793	0.00% Goal: 0%
Peptide Omegas	Bad angles:	2 / 3789	0.05% Goal: <0.1%
	Cis Prolines:	0 / 17	0.00% Expected: ≤1 per chain, or ≤5%
Additional validations	Twisted Peptides:	1 / 335	0.30% Goal: 0
	Tetrahedral geometry outliers	1	
	Waters with clashes	1/141	0.71% See UnDowser table for details

Рис. 10. Данные MolProbity по исходной модели 1uwr (сверху) и по модели, скорректированной PDB-Redo (снизу)

При совмещении исходной модели структуры 1uwr со скорректированной PDB-Redo (см. Рис. 11) видно, что изменения коснулись, в основном, некоторых петель и β-тяжей.

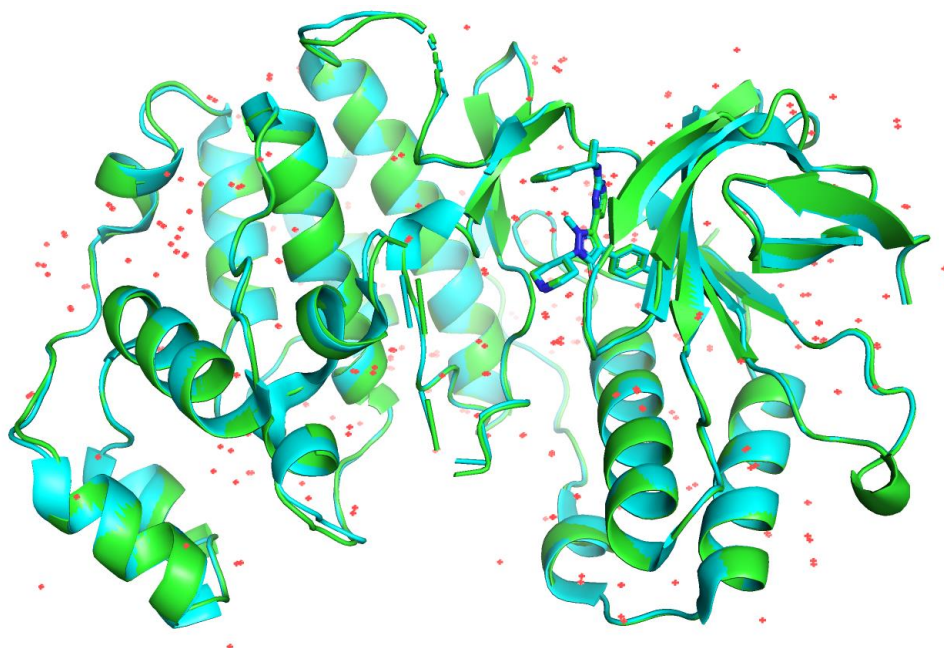


Рис. 11. Исходная структура белка (голубая), совмещенная с скорректированной PDB-Redo (зеленая)

PDB-Redo улучшил вписанность в электронную плотность для 53 остатков, значительных ухудшений в соотношении, скорректированной модели с электронной плотностью, нет (разве что, для одной молекулы воды).

Если рассматривать остатки, выбранные мною в задании 2, то четыре из пяти остатков значительно лучше вписались в электронную плотность и перестали быть маргиналами по ротамерам и по клешам (за исключением остатка 275, который остался маргинальным по торсионным углам).

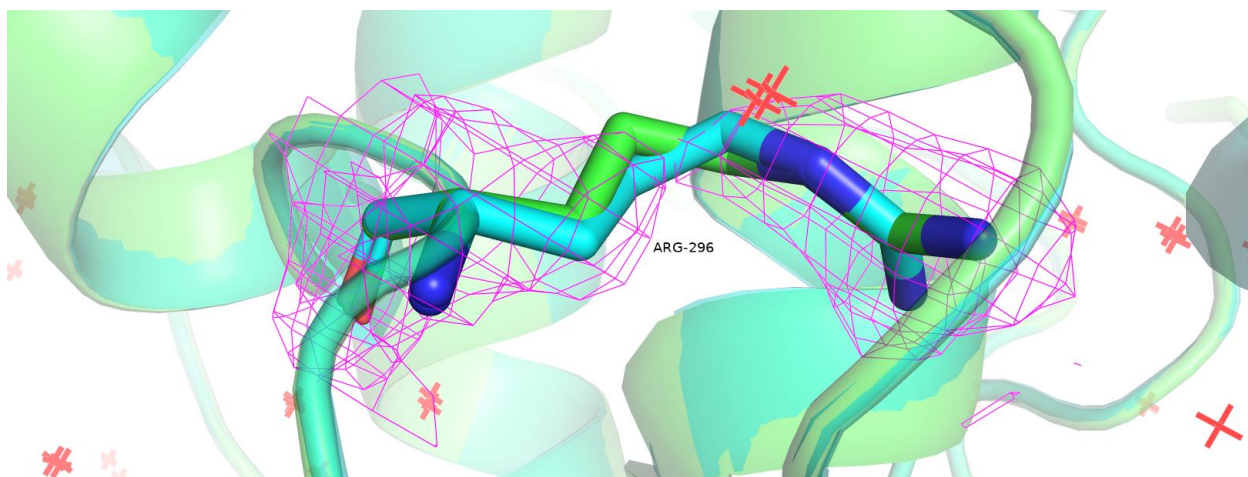


Рис. 12. Arg-296 оказался лучше вписан в электронную плотность и не считается плохим ротамером в скорректированной структуре (зеленой)

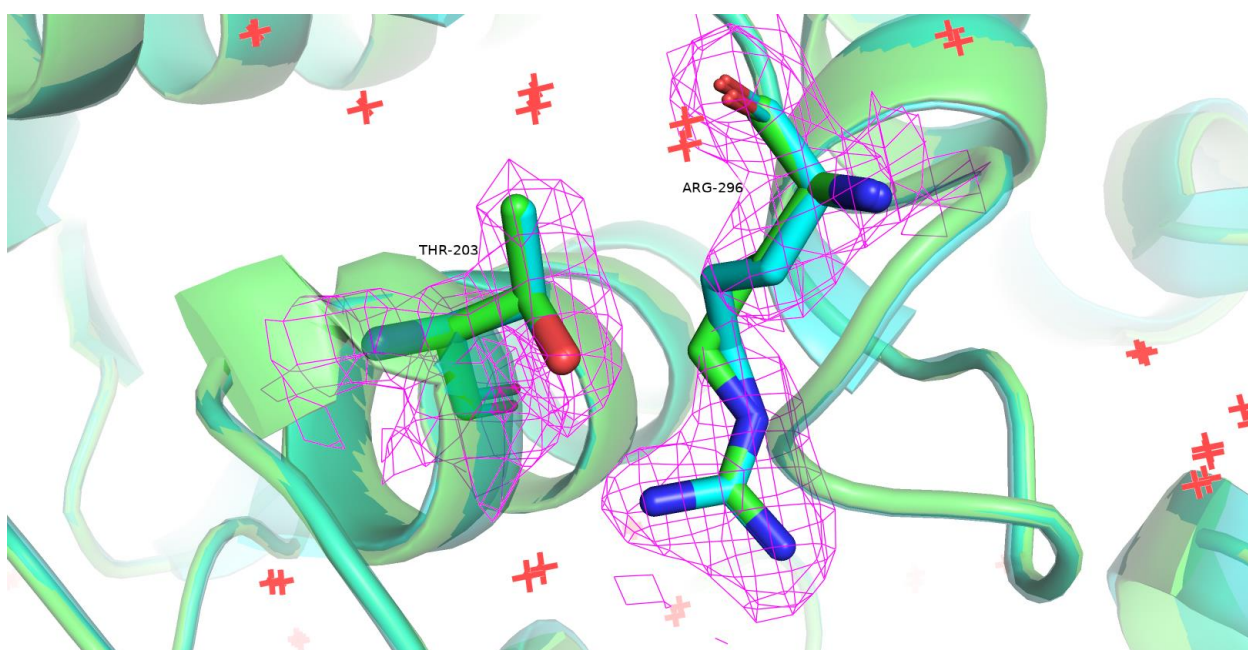


Рис. 13. Клеши между Arg-296 и Thr-203 больше нет в скорректированной структуре (зеленой)

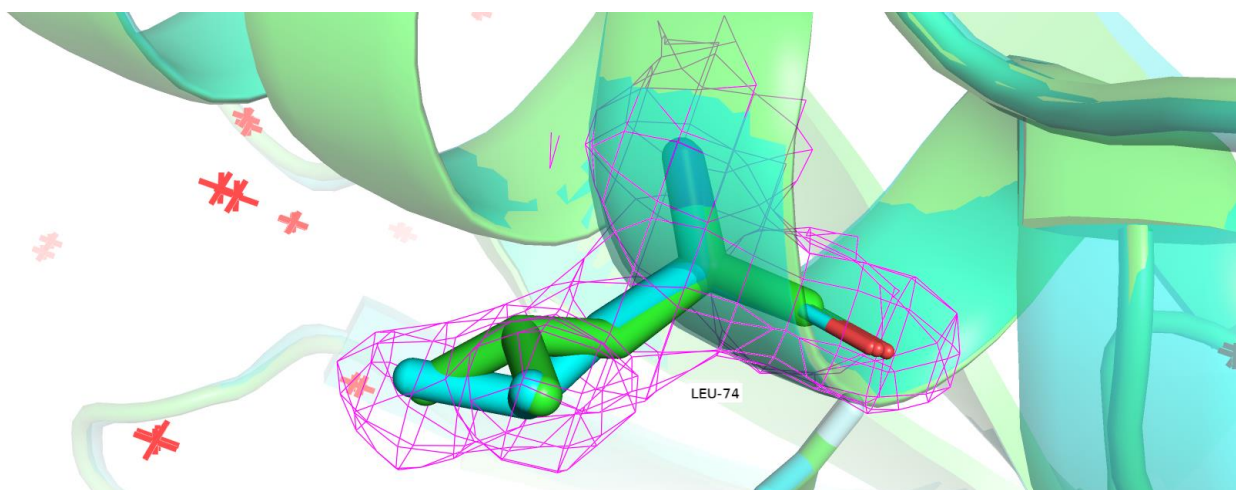


Рис. 14. Leu-74 оказался лучше вписан в электронную плотность и не считается плохим ротамером в скорректированной структуре (зеленой)

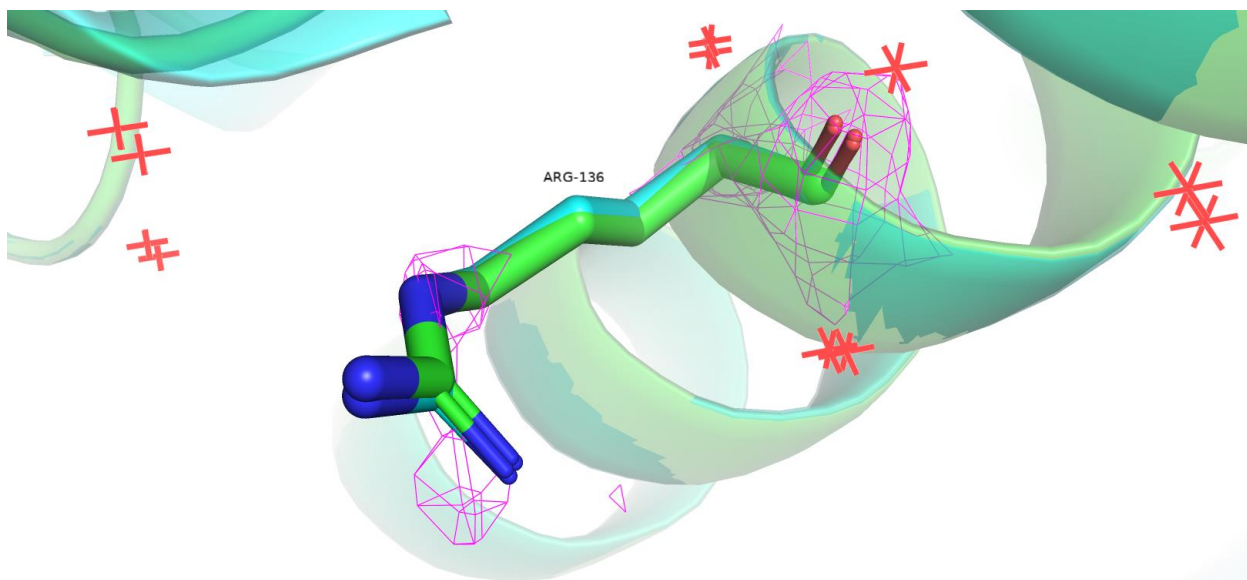


Рис. 15. Arg-136 оказался лучше вписан в электронную плотность (хоть и незначительно, судя по данным PDB-Redo) и не считается плохим ротамером в скорректированной структуре (зеленой)

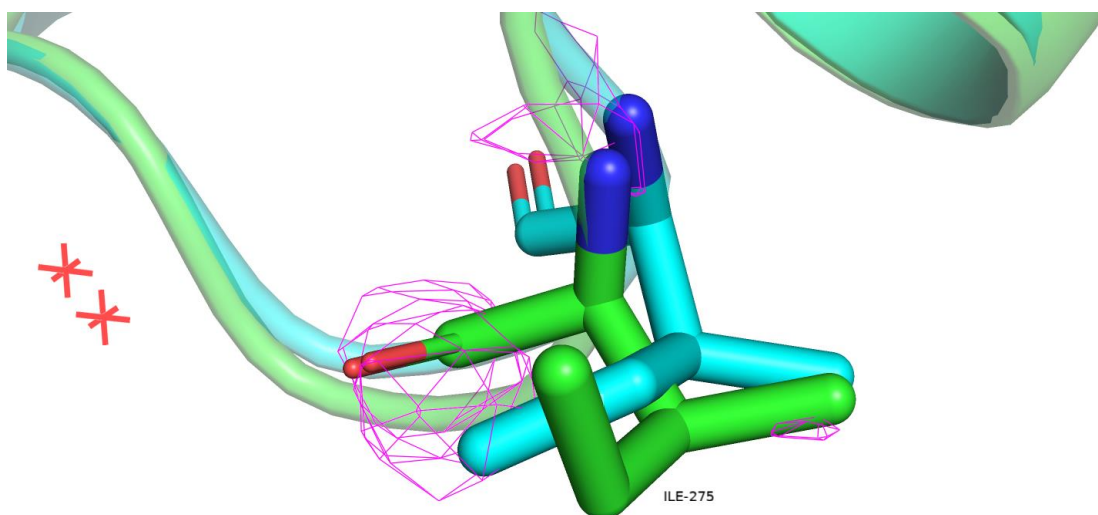


Рис. 16. Ile-275 оказался лучше вписан в электронную плотность, хотя не перестал быть маргиналом по карте Рамачандрана в скорректированной структуре (зеленой)

В качестве примеров остатков с наиболее значительными улучшениями вписанности в электронную плотность можно привести Glu-163 и Lys-121, которые после редактирования PDB-Redo перестали быть аутлайерами по клешам и по ротамерам, соответственно.

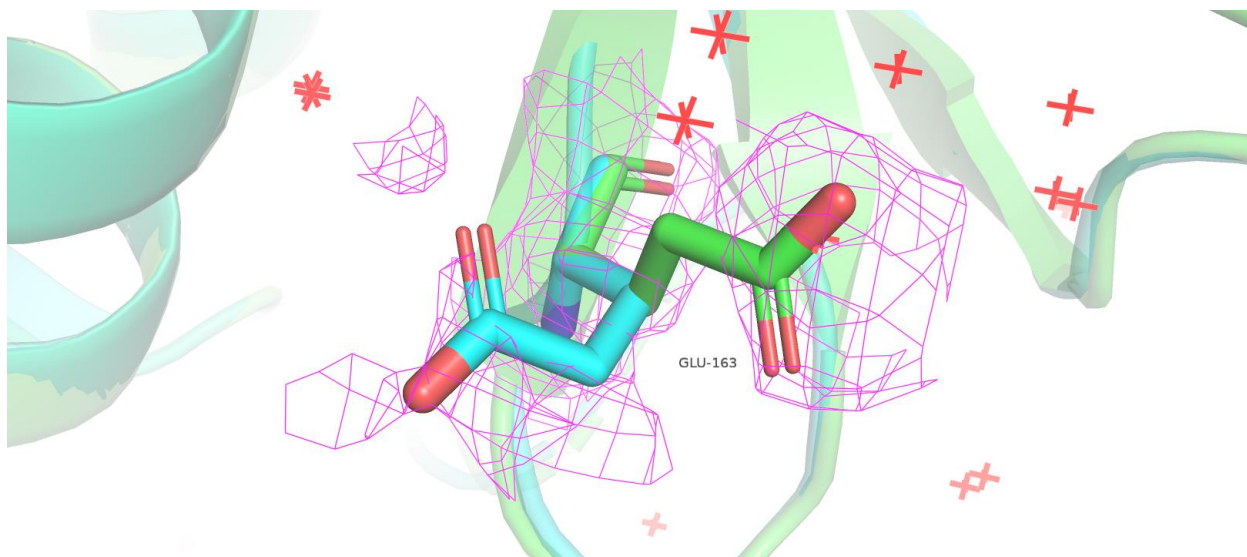


Рис. 17. Glu-163 оказался лучше вписан в электронную плотность и не считается аутлайером по клешам в скорректированной структуре (зеленой)

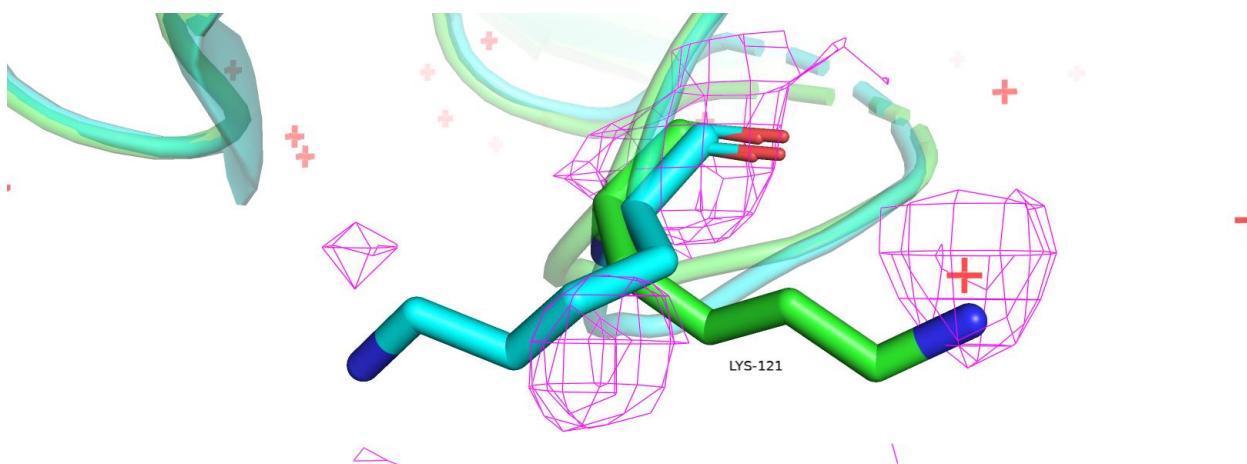


Рис. 18. Lys-121 оказался лучше вписан в электронную плотность и не считается плохим ротамером в скорректированной структуре (зеленой)

Аутлайеров вблизи лиганда стало меньше в скорректированной структуре.

Таким образом, структура, скорректированная PDB-Redo, скорее всего, может быть использована для изучения белка и его взаимодействия с лигандом.

Источники

1. <https://www.uniprot.org/uniprot/P47811>
2. Golebiowski, Adam & Townes, Jennifer & Laufersweiler, Matthew & Brugel, Todd & Clark, Michael & Clark, Cynthia & Djung, Jane & Laughlin, Steven & Sabat, Mark & Bookland, Roger & VanRens, John & De, Biswanath & Hsieh, Lily & Janusz, Michael & Walter, Richard & Webster, Mark & Mekel, Marlene. (2005). The development of monocyclic pyrazolone based cytokine synthesis inhibitors. *Bioorganic & medicinal chemistry letters*. 15. 2285-9. 10.1016/j.bmcl.2005.03.007.