#### Московский государственный университет имени М.В. Ломоносова

Факультет биоинженерии и биоинформатики



**Отчет о качестве расшифровки структуры 1WPV методом рентгеноструктурного анализа**

Работа студентки 4-го курса

Прохоровой Евгении Александровны

Москва 2014

## Содержание

[Содержание 2](#_Toc403897702)

[Список сокращений 3](#_Toc403897703)

[Аннотация 4](#_Toc403897704)

[Введение 4](#_Toc403897706)

[Результаты и обсуждение 7](#_Toc403897708)

[1. Общая информация о модели 7](#_Toc403897709)

[2. Оценка качества модели 8](#_Toc403897711)

[3. Изучение маргинальных остатков структуры 14](#_Toc403897712)

[4. Сравнение модели из PDB с моделью из PDB\_redo 14](#_Toc403897712)

[Заключение 20](#_Toc403897713)

[Список литературы 22](#_Toc403897714)

## Список сокращений

Матричная рибонуклеиновая кислота

Electron Density Server

Histidine β-Naphthylamide

Hut operon positive regulatory Protein

Protein Data Bank

Real-Space R factor

мРНК

EDS

HBN

HutP

PDB

RSR

## Аннотация

## Данная работа систематизирует информацию, полученную при расшифровке структуры белка 1WPV. Проведен анализ показателей качества структуры, остатков, важных для выполнения белком своей функции, и маргинальных остатков. На основании полученных данных сделан вывод о качестве расшифровки структуры 1WPV методом рентгеноструктурного анализа.

## Введение

## Модель 1WPV (рис. 1, с) с разрешением 1,70 Å белка HutP сенной палочки (Bacillus subtilis) в комплексе с остатками гистидина и ионами магния была определена в 2004 году методом РСА (Рентгеноструктурный анализ) в Японии T. Kumarevel, H. Mizuno, P.K. Kumar [1].

## HutP (hut operon positive regulatory protein) – позитивный регулятор hut оперона сенной палочки. HutP регулирует экспрессию структурных генов hut сенной палочки, препятствуя терминации транскрипции. Антитерминация – один из нескольких механизмов, используемых бактериями для регуляции транскрипции оперонов. Однако последовательность HutP не несет сходства ни с последовательностями известных белков, работающих по данному механизму, ни с последовательностями РНК-связывающих белков.

## Ранее было показано, что для работы HutP необходимо наличие L-гистидина в растворе [2, 3]. Чтобы понять, какая именно группа данного остатка важна для функционирования белка, авторами структуры 1WPV были проанализированы 15 аналогов L-гистидина. В результате, было выяснено, что имидазольная группа и атомы остова наиболее важны для работы HutP [4, 5]. Кроме того, была получена структура белка в комплексе с гистидин-β-нафтиламидом (HBN, histidine β-naphthylamide) 1VEA (рис. 1, b) [4]. Данная структура позволила сделать вывод, что активная форма HutP представлена гексамером, и продемонстрировала существование нового РНК-связывающего мотива. Данный мотив взаимодействует с последовательностью мРНК из 21 нуклеотида [4], идентифицированной c помощью отбора in vitro и направленного мутагенеза. Эта последовательность: UAGNNNNUAGNNNNUAG, где N – любой нуклеотид, и каждый мономер HutP связывает UAG мотив. При этом как минимум 2 мотива необходимы для формирования комплекса [6].

## Т.к. для HutP были неизвестны близкие гомологи среди белков других бактерий, и для наиболее схожих по последовательности белков (60% сходства) тоже нет разрешенных структур без лиганда, авторы исследуемой мной структуры попытались определить для HutP кристаллографические структуры активного анти-терминационного комплекса, а также предшествующих его образованию форм белка.

## Так, наряду c 1WPV в статье были опубликованы еще 2 структуры: 1WPS и 1WMQ. 1WPS (рис. 1, а) – структура димера без лигандов разрешения 2,80 Å, 1WMQ (рис. 1, d) – структура с разрешением 1,60 Å активного комплекса из двух субъединиц, каждая из которых связана с молекулой РНК, остатком гистидина и ионом магния [1].

## Структуры конформаций HutP до формирования активной формы (1WPS и 1WPV) показывают, что ион магния координирует L-гистидин, способствуя необходимым для связывания определенной последовательности РНК изменениям структуры. Модель 1WMQ демонстрирует важность связывания гистидина и магния для взаимодействия HutP с консервативной последовательностью hut мРНК. Молекула РНК «оборачивается» вокруг белка (рис. 1, d), в результате чего перестраивается ее структура и дестабилизируется участок, отвечающий за терминацию [1].

## C:\Users\Jenny\Desktop\Безымянный.pngВ качестве обобщения полученных данных авторами структур были предложены модели активации комплекса терминации транскрипции (рис. 1, e) и связывания РНК (рис. 1, f).

## Рисунок 1. Модели HutP и предполагаемый механизм формирования комплекса, препятствующего терминации транскрипции. a-d. Поверхности димера HutP (а), HutP-HBN (b), комплекса HutP с гистидином и магнием (c), комплекса HutP с гистидином, магнием и РНК (d), окрашенные в соответствии с электростатическим потенциалом. е. Модель образования комплекса терминации транскрипции. f. Модель, предполагающая существование двух сайтов связывания мРНК (показаны в виде голубых прямоугольников) в области терминации. ГЦ-богатый участок выделен красным. РНК-связывающие остатки показаны розовым и зеленым. Взято из [1].

## Таким образом, благодаря данной работе, стали известны механизм образования активной формы HutP, регулирующей транскрипцию мРНК hut оперона сенной палочки, и взаимодействия, важные для его формирования.

## 

## Результаты и обсуждение

## 1. Общая информация о модели

### В модели 1WPV представлены три цепи (A, B, C) длиной по 147 остатков, каждая из которых связана с остатком гистидина и ионом магния, а также молекулы воды.

### Для решения фазовой проблемы при определении структуры использовался метод молекулярного замещения (MR, Molecular replacement) с использованием ранее разрешенной структуры 1VEA (HutP в комплексе с HBN) с разрешением 2,80 Å, для получения которой применялся метод многоволнового аномального рассеяния (MAD, Multi-wavelength anomalous dispersion) [4].

### Число измеренных рефлексов, представленных в файле структурных факторов cif [7] и указанных на сервере EDS [8] – 52035. Число использованных рефлексов, указанное в PDB-файле [7] – 50886. Полнота данных, указанная в PDB [7], равна 96,4%, а в EDS [8] – 94,5%. Число рефлексов с силой сигнала, превышающей стандартное отклонение более чем в три раза, равно 39770 (76,4%).

### Модель имеет вполне хорошее разрешение – 1,70 Å. Минимальное разрешение для используемых рефлексов, согласно данным PDB [7], равно 1,70 Å, тогда как сервер EDS [8] указывает минимальное разрешение 1,68 Å. Максимальное разрешение по данным PDB [7] равно 19,53 Å, а сервер EDS [8] указывает максимальное разрешение 24,69 Å.

### Параметры кристаллографической ячейки [2, 3]: a=77,79 Å, b=80,88 Å, c=75,17 Å; alpha=90,00, beta=90,00, gamma=90,00. Кристаллографическая группа структуры [2, 3] – P 21 21 2, в ячейке – 12 молекул.

## 2. Оценка качества модели

### Проведем оценку модели 1WPV в целом, проанализировав некоторые из известных индикаторов качества.

### R-фактор – величина, характеризующая, соответствие модели экспериментальным данным. Она минимизируется в процессе уточнения модели путем смещения координат атомов. Для структуры 1WPV R-фактор равен 23,3% – по данным PDB [7], и 23,5% – по данным EDS [8], что относится к хорошим значениям (менее 25%) и в то же время не вызывает подозрений о подгонке.

### Другой показатель качества модели в целом – R-free. Для рассматриваемой структуры по данным PDB он равен 25,9 [7]. R-free вычисляется по окончательной модели и по контрольным рефлексам (обычно берется 10% рефлексов) для контроля правильности модели. R-free данной модели вычислен по 5% рефлексов (2535). R-free считается хорошим, если он меньше 20%, и плохим – если он больше 40%. Т.е. в данном случае значение R-free достаточно хорошее и к тому же близко к значению R-фактора (немного больше, R-free – R-factor = 25,9 – 23,3 = 2,6%), что является показателями правильности модели. Значения R-free – R-factor меньше 10% говорят о возможной переоптимизации модели.

### Пространственный R-фактор (RSR, real-space R factor) – величина, оценивающая, насколько электронная плотность, построенная по полученной модели, соответствует «экспериментальной» электронной плотности. В отличие от R-фактора и R-free, RSR (как и его Z-score, описываемый ниже) используется скорее для локальной оценки качества, однако его среднее значение для остатков структуры можно рассматривать как оценку качества модели в целом. Аналогично дело обстоит с картой Рамачандрана, ротамерами, температурным фактором и рядом других параметров, поэтому в данном разделе отчасти уделяется внимание и локальной оценке качества модели 1WPV.

### В данном случае средний RSR для модели составляет 11,4% (рис. 2) [8], что ненамного превышает 10%, поэтому является хорошим значением. Если же рассматривать RSR как характеристику локального качества модели, то маргиналы обычно имеют значения RSR более 20%. На рис. 2 видно, что, например, для цепи А белка маргиналами являются остатки 19-25 и 134. Вероятно, эти остатки плохо вписаны C:\Users\Jenny\Desktop\1wpv\rsr.pngв электронную плотность. Действительно, для аналогичных остатков 20-25 цепи С с RSR больше 20% это демонстрирует рис. 6.

### Рисунок 2. Графики пространственных R-факторов для остатков трех полипептидных цепей структуры 1WPV [8].

### C:\Users\Jenny\Desktop\1wpv\z.png Z-score RSR (относительная оценка RSR) – для вычисления Z-score остатка его RSR сравнивается со средним RSR того же типа остатков в выборке PDB-структур, имеющих такое же разрешение. Если RSR плохой, но с хорошей Z-score, то значит координаты атомов расшифрованы плохо, но не хуже, чем в других подобных структурах. Но в случае структуры 1WPV в среднем по субъединицам в модели имеется 5,41% остатков (рис. 3) [8], Z-score которых больше 2%, а это значит, что, по сравнению со структурами с таким же разрешением, 5,41% остатков из структуры плохо вписаны в электронную плотность, т.е. являются маргиналами.

### Рисунок 3. Z-score для остатков трех цепей структуры [8].

### Карты Рамачандрана (рис. 6) были построены с помощью сервера MolProbity [9]. Данный сервер обеспечивает построение карты Рамачандрана с лучшими на сегодняшний день границами предпочитаемых и допустимых областей. При построении карты к PDB структуре добавлялись атомы водорода (для выявления недопустимых наложений атомов).

### По окончанию построения помимо карты сервер выдает таблицу с параметрами, описывающими контакты между атомами и геометрию белковых молекул, рассмотрим и их.

### Первый из параметром – сlashscore – показатель, отображающий число недопустимых перекрываний атомов (более 0,4 Å) в расчете на 1000 – равен в данном случае 10,16, т.е. 1,016% атомов. Также указан показатель 66th percentile (N=819, 1,70Å ± 0,25Å) – 66% структур из выборки, включающей 819 структур с разрешением 1,70Å ± 0,25Å, имеют ClashScore хуже (больше), чем у данной, что больше половины.

### C:\Users\Jenny\Desktop\1wpv\b.pngЧисло остатков с маргинальными отклонениями боковых цепей от их нормальных ротамеров – 7 (2,13%) [9], что превышает предпочтительное значение – менее 1%. Скорее всего, этим ротамерам соответствуют большие значения температурных факторов (рис. 4). Температурный, или В-фактор, характеризует меру динамической неупорядоченности кристалла, обусловленной тепловым движением. С повышением температуры неупорядоченность повышается, увеличиваются тепловые колебания атомов. Как видно из рис. 4, структура содержит остатки, В-фактор которых больше 30 (больше значения, считаемого оптимальным), определение положения боковых цепей которых, вероятно, было затруднено.

### Рис. 4. График температурных факторов всех остатков структуры [8].

### Согласно же протоколу проверки качества WHAT\_CHECK (из пакета WHAT IF) [10], для белковой структуры, определенной при комнатной температуре также желательно, чтобы не более 1% атомов имели B-фактор меньше 5. Для структуры 1WPV данное значение – 0,9%, т.е. условие выполняется.

### C:\Users\Jenny\Desktop\Безымянный1.png Если же посмотреть на рис. 5 из выдачи MolProbity [9], можно заметить, что остатки с маргинальными отклонениями боковых цепей от их нормальных ротамеров расположены во основном на поверхности белка, где остатки всегда более подвижны.

### Рис. 5. Отклонения от нормы в структуре 1WPV, обнаруженные сервером MolProbity [9]. Остов белка выделен проволочной моделью и раскрашен по цепям (цепь А – черным, В – желтым, С – коричневым), розовым выделены области перекрывания атомов, зеленым обозначены маргиналы с карты Рамачандрана, оранжевым выделены боковые цепи маргинальных ротамеров.

### Карта Рамачандрана – хороший индикатор качества структуры, т.к. он, как правило, не зависит от процедуры оптимизации модели.

### 412 из 435 остатков модели 1WPV (94,7%) находятся в предпочтительных областях карты Рамачандрана, 434 из 435 (99,8%) – в разрешенных областях (рис. 6). По общепринятому критерию, более 90% остатков белка должно находиться в предпочтительных областях карты (хотя сервер MolProbity устанавливает более жесткие требования – более 98%), поэтому модель можно отнести к хорошей. Также модель практически соответствует требованию сервера на то, что более 99,8% остатков должно C:\Users\Jenny\Desktop\Безымянный1.pngC:\Users\Jenny\Desktop\Безымянный1.pngнаходиться в разрешенных областях.

### Рис. 6. Карты Рамачандрана, выданные программой MolProbity [9]. Фиолетовым кружком обозначен маргинальный остаток Glu23 цепи С.

### Число маргинальных остатков на карте Рамачандрана – 1 (0,23%). Это Glu23 (-60,2, 92,0) цепи C. Однако если посмотреть на рис. 6, General case, то можно сказать, что данный остаток пограничен разрешенной области, что не настолько критично, тем не менее он был внесен в табл. 1, №1 маргинальных остатков, т.к. является маргинальным и по некоторым другим критериям.

### Интегральная оценка структуры по данным сервиса (MolProbity score) [9] имеет значение 2,14. Данная оценка объединяет в себе такие показатели, как clashscore, наличие остатков с маргинальными по отклонению от ротамеров боковыми цепями и оценки на основе карты Рамачандрана, и нормирует их так, чтобы интегральная оценка получалась в той же шкале и единицах измерения, как разрешение кристаллографической структуры. То есть сервер оценил качество рассматриваемой структуры как соответствующее структуре разрешением в 2,14 Å, что на четверть больше, чем указанное в PDB [7] разрешение структуры 1,70 Å.

### Для MolProbity score кроме того указан показатель 43rd percentile (N=9248, 1,70Å ± 0,25Å) [9], т.е. 43% структур из определенной в 2006 году выборки, включающей 9248 структур (больше чем в выборке, рассматриваемой при оценке Clashscore, определенной в 2004 году) с разрешением 1,70Å ± 0,25Å, имеют MolProbity score хуже (больше), чем у данной.

### Число атомов Сβ с неприемлемым отклонением от ожидаемого положения – 0 (0%); число ковалентных связей, существенно отклоняющихся от теории – 0 из 3372 (0%); число валентных углов, существенно отклоняющихся от теории – 8/4562 (0,18%, а в идеале эта величина должна быть менее 0,10 %).

### В протоколе проверки качества WHAT\_CHECK [10] представлены результаты проверки модели и по другим критериям, не описанным в данной работе. Однако проверку по большинству из них модель прошла успешно, в принципе, как и практически по всем рассмотренным выше параметрам.

## 3. Изучение маргинальных остатков структуры

### Рассмотрим более детально некоторые из маргинальных остатков структуры 1WPV, о которых отчасти уже говорилось в предыдущем разделе. Для изучения выбраны остатки, маргинальные по значительному числу параметров (табл. 1).

### Таблица 1. Маргинальные остатки структуры.

|  |  |  |
| --- | --- | --- |
| **№** | **Маргинальный остаток** | **Показатель отбора** |
| **1** | Glu23 цепи C **(рис. 7)** | Из запрещенной области карты Рамачандрана, (ϕ,ψ) = (-60,2, 92,0) [9];  RSR = 0,597; Z-score = 6,732; B-фактор = 68,170; валентный угол отклоняется от теории - N-CA-C: 4.02 σ [8] |
| **2** | Glu22 цепи C **(рис. 7)** | RSR = 0,355; Z-score = 4,953; B-фактор = 68,49 [8];  наложение 0,536 Å для кислорода данного остатка и углерода Ser24 [9] |
| **3** | Lys134 цепи С **(рис. 8)** | RSR = 0,301; B-фактор = 55,27 [8];  наложение 0,430 Å HD3 c 3131 HOH O цепи С [9] |
| **4** | Arg103 цепи В **(рис. 9)** | B-фактор = 41,57; отклонение боковой цепи от их нормального ротамера 0,1%, chi: 185,1, 152,6, 126,6, 59,2 [8];  наложение 0,595 Å HD2 c Lys41 HE2 цепи B; валентный угол отклоняется от теории - N-CA-C: 4,319 σ [9] |
| **5** | Молекула воды 3166 цепи B **(рис. 10)** | B-фактор = 52,27 [8];  наложение 0,882Å O с Glu20 HG3 цепи В [9] |

### Первый из рассматриваемых остатков – Glu23 цепи С (табл. 1, №1; рис. 7) – единственный остаток структуры, попавший в запрещенную область карты Рамачандрана [9], хотя 23-ие остатки цепей А и В очень схожи с остатком данной цепи и тоже находятся близко к границе на карте, но со стороны разрешенной области. Однако стоит отметить и, что только Glu23 цепи С характеризуется отклонением в значении валентного угла от теории – 1 из 8 случаев в модели. 23-ий остаток цепей В и С относится к группе остатков 20-25 и 19-25 в случае цепи А со значениями RSR больше 20% (рис. 2), Z-score больше 2 (рис. 3) и температурным фактором больше 40 (рис. 4) [8]. Поэтому не только Glu23 можно считать маргинальным, но и окружающие его остатки (однако в табл. 1 данной работы представлены только остатки, удовлетворяющие наибольшему числу критериев маргинальности из обсуждаемых примеров).

### C:\Users\Jenny\Documents\23.jpgЕсли обратиться к структуре белка (рис. 7), то можно увидеть, что данные остатки приходятся на поворот вторичной структуры и располагаются на поверхности белка, при этом практически не формируя связей даже с молекулами воды (по крайней мере, в случае Glu22 и Glu23). Соответственно, они довольно подвижны, и их электронная плотность описана плохо – на рис. 7 видно, что при уровне подрезки 1 (показана черным) ее поверхность для данных атомов не отображается. При уровне 2 (серым), несмотря на хорошее разрешение структуры в случае Glu22 и Glu23 поверхность электронной плотности не совпадает даже с остовами данных остатков.

### Рис. 7. Glu22 и Glu23 цепи С. Серым показана поверхность электронной плотности с уровнем подрезки 1,0, черным – 2,0. Молекулы воды представлены в виде красных сфер. Изображение получено с помощью The PyMOL Molecular Graphics System, Version 1.1, Schrödinger, LLC.

### Помимо описанных данных, свидетельствующих о маргинальности Glu22 цепи С, для данного остатка (только в цепи С) характерно еще и недопустимое наложение 0,536 Å кислорода и углерода Ser24, а также значения углов близкие к запрещенной области на карте Рамачандрана [9], поэтому было решено добавить его к таблице рассматриваемых маргинальных по многим параметрам остатков (табл. 1, №2). Серин же не был включен в таблицу, т.к. соответствующие ему значения на карте Рамачандрана располагаются в предпочтительной области.

### C:\Users\Jenny\Documents\134.jpgДалее был рассмотрен остаток Lys134 цепи С (табл. 1, №3; рис. 8). Lys134 трех субъединиц белка характеризуются температурным фактором больше 40, однако только лизины цепей А и С характеризуются значениями RSR более 20%. При этом Z-score RSR менее двух, т.е., вероятно, по сравнению со структурами с таким же разрешением данные остатки хорошо вписаны в электронную плотность. Если же посмотреть на рис. 8 можно увидеть, что плотность Lys134 практически не определена. Скорее всего такое часто случается с подвижными поверхностными остатками лизина в структурах.

### Рис. 8. Lys134 цепи С. Серым показана поверхность электронной плотности с уровнем подрезки 1,0, черным – 2,0. Молекулы воды представлены в виде красных сфер. Изображение получено с помощью The PyMOL Molecular Graphics System, Version 1.1, Schrödinger, LLC.

### Другая причина выбора Lys134 цепи С для изучения – наложение 0,430 Å HD3 c водой 3131 цепи С [9]. Однако при исследования структуры оказалось, что данная молекула воды находится на расстоянии больше 20 Å от остатка и, видимо, сервер MolProbity выдал ошибку. Что касается лизинов других цепей – для остатка цепи В найдено наложение собственных водородов HA и HE2, для лизина цепи А наложения не характерны.

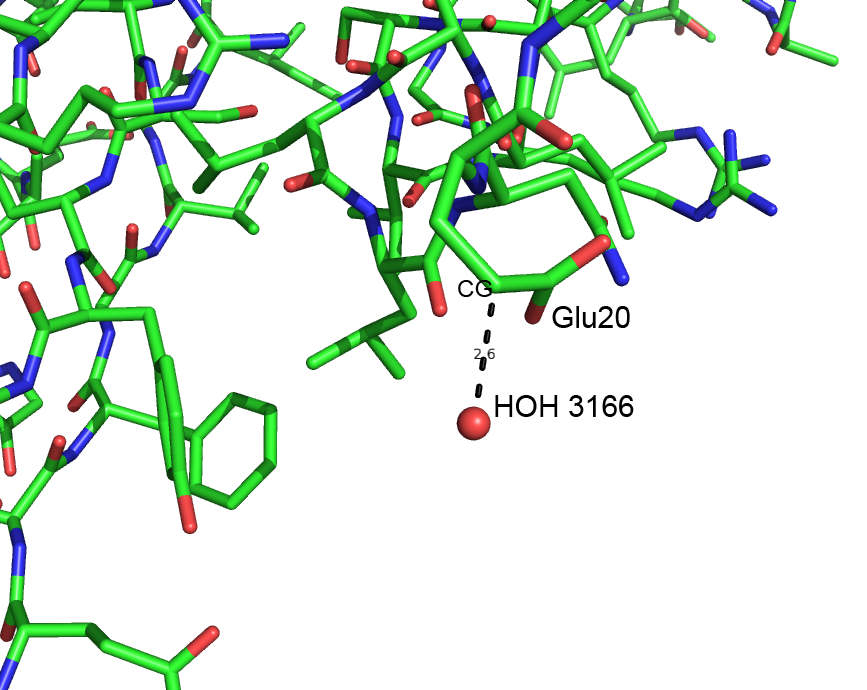
### Таким образом, Lys134 можно отнести к маргиналам, но по меньшему числу индикаторов, чем хотелось бы.

### Следующий кандидат – Arg103 цепи В (табл. 1, №4; рис. 9). Остаток имеет высокий B-фактор = 41,57 (и еще больше в остальных субъединицах), его боковая цепь характеризуется отклонением от нормального ротамера (0,1%, chi: 185,1, 152,6, 126,6, 59,2) [8]. Также наблюдается наложение 0,595 Å c Lys41 цепи B. В случае же цепи А наложений нет, С – наложение с молекулой воды. Кроме того, валентный угол Arg103 отклоняется от теории - N-CA-C: 4,319 σ и это характерно для остатков всех субъединиц C:\Users\Jenny\Documents\lys.jpgбелка [9].

### Рис. 9. Arg103 цепи B. Серым показана поверхность электронной плотности с уровнем подрезки 2,0. Молекулы воды представлены в виде красных сфер. Изображение получено с помощью The PyMOL Molecular Graphics System, Version 1.1, Schrödinger, LLC.

### Полярные атомы Arg103 находятся у поверхности белка (не выходя из нее), но в принципе все за исключением одного образуют водородные связи. В гуанидиновой группе NE и NH2 атомы образуют связи с молекулами воды, однако NH2 атом не формирует таковых и соседствует с гидрофобным радикалом Leu39, поэтому это тоже можно отнести к признакам маргинальности остатка. Однако по некоторым индикаторам остаток не подходит: Arg104 трех цепей имеют хорошие значения RSR и его Z-score, и, как и в случае рассматриваемого выше Lys134, Arg103 цепей находится в предпочтительной области карты Рамачандрана [9].

### Также была рассмотрена молекула воды 3166 цепи В (табл. 1, №5; рис. 10). Молекула была взята как имеющая самое высокое значение наложения – 0,882 Å. Накладывается она с HG3 Glu20 цепи В, судя по данным MolProbity [9]. Действительно, на рис. 10 можно увидеть, что кислород данной молекулы воды располагается довольно близко к СG атому Glu20. Однако Glu20 – поверхностный остаток, и его электронная плотность разрешена плохо. При уровне подрезки 1,5 и более для его атомов она не визуализируется, при уровне подрезки 1,4 и менее видна только очень маленькая незначительная область правее атома CG, если ориентироваться по рис. 10. Видимо, данный остаток подвижен, о чем свидетельствуют и высокие значения В-фактора его атомов, увеличивающиеся по мере удаления атома от остова – В-фактор атома N = 50,65, а В-фактор OE2 = 58,89 [7].

****

### Рис. 10. Наложение кислорода молекулы воды на атом кислорода остатка Glu20 цепи В. Молекула воды показана в виде красной сферы. Изображение получено с помощью The PyMOL Molecular Graphics System, Version 1.1, Schrödinger, LLC.

### В-фактор данной молекулы вод также высок – 52,27 [8] – что, как и наложение, является признаком маргинальности. Таким образом, вариант взаимного расположения воды 3166 и Glu20 цепи В, приведенный в PDB файле, маловероятен.

### В статье, в которой получена структура 1WPV [1], не ведется обсуждение описанных выше остатков, но было бы интересно проанализировать и остатки, упоминаемые авторами. Наибольшее внимание в статье уделено остаткам сайта связывания L-гистидина и магния (рис. 11) и остаткам важным для связывания РНК (в основном для структуры 1WMQ комплекса HutP с гистидином, магнием и РНК). Однако данные остатки не являются маргинальными ни по одному из упоминаемых в работе индикаторов. Электронная плотность в данных сайтах связывания и вокруг также C:\Users\Jenny\Documents\act.jpgразрешена хорошо.

### Рис. 11. Остатки сайта связывания L-гистидина и магния субъединицы А структуры 1WPV. Серым показана поверхность электронной плотности с уровнем подрезки 2,0, черным – 3,5. Ион магния представлен в виде желтой сферы, молекулы воды – в виде красных сфер. Остаток L-гистидина не показан для более наглядной визуализации остатков сайта. Изображение получено с помощью The PyMOL Molecular Graphics System, Version 1.1, Schrödinger, LLC.

### Другие остатки, про которые упоминается в статье, также нельзя отнести к маргинальным ни по одному из рассматриваемых критериев.

## 4. Сравнение модели из PDB с моделью из PDB\_redo

### Сравним исходную PDB структуры со структурой оптимизированной стандартной проверенной программой в PDB\_redo [11], исходя из представленных авторами PDB файла экспериментальных данных.

Полученная структура, действительно, имеет более высокое качество по части параметров. Так улучшились R-фактор и R-free **(табл. 2)**.

### Таблица 2. Сравнение значений R-фактора и R-free до и после оптимизации структуры 1WPV в PDB\_redo.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Значения в PDB [7]** | **Значения, подсчитанные по параметрам, используемым авторами [11]** | **Значения после полной оптимизации в PDB\_redo [11]** |
| **R-фактор** | 0,2330 | 0,2309 | 0,1762 |
| **R-free** | 0,2590 | 0,2245 | 0,2023 |
| **R-free – R-фактор** | 0,0260 | -0,0064 | 0,0261 |

### Настораживает, что подсчитанные PDB\_redo R-фактор и R-free по данным, используемым авторами структуры, довольно сильно отличаются от представленных в PDB (табл. 2), что наводит на мысль о подгонке авторами. Тем не менее разность R-фактора и R-free после оптимизации фактически равно значению, представленному авторами.

### Кроме того, число маргиналов на карте Рамачандрана уменьшилось с 1 до нуля, что вполне адекватно, т.к. данный маргинал располагался очень близко к допустимой области.

### C:\Users\Jenny\Documents\1цуц.jpgЕсли визуально сравнить полученную структуру с представленной в PDB можно увидеть, что в пределах вторичной структуры изменений практически нет (рис. 12, A). Однако на рис. 12, Б видно, что координаты части остатков, в основном поверхностных, изменились.

### Рис. 12. Сравнение PDB структуры 1WPV и оптимизированной в PDB\_redo версии [11]. А. Наложение вторичных структур тримеров. Б. Наложение остатков А субъединицы. Зеленым изображена исходная PDB структура, голубым - исправленная. Изображение получено с помощью The PyMOL Molecular Graphics System, Version 1.1, Schrödinger, LLC.

## 

## Заключение

### Подводя итог изучению показателей индикаторов качества модели, можно сделать вывод о высоком качестве структуры 1WPV. Электронная плотность большинства атомов белка хорошо описана, в частности отлично разрешена плотность остатков, участвующих в связывании лигандов, что необходимо для функционирования HutP. Фактически все атомы находятся в предпочтительной или разрешенной областях карты Рамачандрана, практически не встречается отклонений от теоритических значений связей и углов. Многие параметры оценки качества модели имеют значения не хуже, чем у других структур схожего разрешения.

### Конечно, в модели имеются и маргинальные остатки, но они, по мнению авторов, не формируют контактов с лигандами белка, а, следовательно, не столь важны при изучении его функции.

## Список литературы

1. Kumarevel T, Mizuno H, Kumar PK. 2005. Structural basis of HutP-mediated anti-termination and roles of the Mg2+ ion and l-histidine ligand. Nature 434:183–191
2. Wray, L. V. Jr & Fisher, S.H. Analysis of Bacillus subtilis hut operon expression indicates that histidinedependent induction is mediated primarily by transcriptional anti-termination and that amino acid repression is mediated by two mechanisms: regulation of transcription initiation and inhibition of histidine transport. J. Bacteriol. 176, 5466–5473 (1994)
3. Oda, M., Kobayashi, N., Ito, A., Kurusu, Y. & Taira, K. Cis-acting regulatory sequences for antitermination in the transcript of the Bacillus subtilis hut operon and histidine-dependent binding of HutP to the transcript containing the regulatory sequences. Mol. Microbiol. 35, 1244–1254 (2000)
4. Kumarevel, T. S. et al. Crystal structure of activated HutP: an RNA binding protein that regulates hut operon in Bacillus subtilis. Structure 12, 1269–1280 (2004)
5. Kumarevel, T. S., Mizuno, H. & Kumar, P. K. R. Allosteric activation of HutP protein, that regulates transcription of hut operon in Bacillus subtilis, mediated by various analogs of histidine. Nucleic Acids Res. Suppl. 3, 199–200 (2003)
6. Kumarevel, T. S., Gopinath, S. C. B., Mizuno, H. & Kumar, P. K. R. Identification of important chemical groups of the hut mRNA for HutP interactions that regulates the hut peron in Bacillus subtilis. Nucleic Acids Res. 32, 3904–3912 (2004)
7. <http://www.rcsb.org/pdb/explore.do?structureId=1WPV>
8. <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=1wpv>
9. <http://molprobity.biochem.duke.edu/index.php>
10. <http://swift.cmbi.ru.nl/servers/html/index.html>
11. <http://www.cmbi.ru.nl/pdb_redo/uw/1uwz/>