

УДК: 548.737

Использование программы кристаллографического уточнения Refmac в процессе выявления альтернативных конформаций в биологических макромолекулах

©2012 Соболев О.В.^{*}, Лунин В.Ю.^{**}

*Институт математических проблем биологии, Российская академия наук, Пущино,
Московская область, 142290, Россия*

Аннотация. Изучение смещений центров атомов при кристаллографическом уточнении макромолекулярной структуры без стереохимических ограничений позволяет обнаруживать присутствие альтернативных конформаций фрагментов полипептидной цепи. Процедуры автоматического принятия решений о наличии или отсутствии альтернативной конформации для конкретного фрагмента опираются на статистический анализ смещений атомов, полученных при тестовом уточнении большого количества белковых структур. При этом распределения вероятностей смещений оказываются зависимыми от используемой программы уточнения. В данной работе описаны построение баз смещений атомов, их анализ и разработка автоматических процедур идентификации альтернативных конформаций при использовании для уточнения структуры одного из наиболее популярных инструментов – программы Refmac.

Ключевые слова: структура биологических макромолекул, кристаллографическое уточнение, альтернативные конформации

ВВЕДЕНИЕ

Одной из особенностей используемых для рентгеноструктурного исследования кристаллов белка является то, что существенная часть (в среднем 50%) объема кристаллической ячейки занята растворителем. Боковые цепи аминокислотных остатков, находящиеся на поверхности белка и обращенные в растворитель, обладают повышенной подвижностью. Это приводит к тому, что даже в кристаллическом состоянии такие боковые цепи могут иметь различные (альтернативные) конформации для разных экземпляров молекулы в кристалле. Для кристаллов высокого качества такая неопределенность сводится к небольшому числу альтернатив, которые включаются в модель белка с соответствующими весами (коэффициентами заселенности), отражающими частоту встречаемости тех или иных конформаций в исследуемом кристалле. На данный момент изучаются случаи двух, реже – трех, альтернативных конформаций. Альтернативные конформации могут вводиться не только для боковых групп, но также для участков главной цепи [1–3]. Введение в модель альтернативных конформаций увеличивает точность расчета фаз структурных факторов, что повышает точность карт распределения электронной плотности во всем объеме элементарной ячейки. Более того, в ряде случаев наличие в кристалле

^{*} oleg@impb.psn.ru

^{**} lunin@impb.psn.ru

альтернативных конформаций связано с функцией белка [2,4–6]. Ранее было замечено, что количество остатков, для которых удается идентифицировать и ввести в модель альтернативные конформации, увеличивается с ростом разрешения набора экспериментальных данных [1,3,7]. Выявление альтернативных конформаций осуществляется, как правило, посредством трудоемкой процедуры визуального анализа карт распределения электронной плотности в ячейке кристалла и носит, во многом, субъективный характер. Нами были предложены алгоритмы автоматического принятия решений, позволяющие упростить этот процесс.

Процедура кристаллографического уточнения является, с математической точки зрения, процедурой оптимизации параметров модели, обеспечивающей как наилучшее соответствие модели экспериментальным данным, так и соответствие модели некоторым априорным (стереохимическим) требованиям. Обычно критерий, минимизируемый в процессе кристаллографического уточнения, состоит из двух частей. Первая часть критерия требует соответствия модели экспериментальным данным. Вторая часть критерия накладывает ограничения на геометрию модели. Необходимость присутствия второй части обусловлена относительно низким отношением количества экспериментальных данных к количеству уточняемых параметров. Она необходима для стабилизации процесса уточнения. Уточнение без стереохимических ограничений (свободное уточнение) нестабильно при низком и среднем разрешении экспериментальных данных. Тем не менее, хорошо упорядоченные участки модели могут уточняться без стереохимических ограничений на последних этапах работы над структурой при наличии экспериментальных данных высокого разрешения [8–10]. В некоторых случаях удается уточнить всю модель без стереохимических ограничений [5]. Однако было замечено, что даже при атомном разрешении плохо упорядоченные остатки сильно искажаются в процессе свободного уточнения, в то время как хорошо упорядоченные остатки стабильны [8].

Основой предложенного метода идентификации альтернативных конформаций явилась гипотеза, что существенное искажение фрагмента полипептидной цепи в ходе свободного уточнения является признаком наличия альтернативных конформаций [11]. Поэтому свободное уточнение изучаемого объекта, даже если оно не приводит к стереохимически состоятельной модели, может быть использовано как источник информации о степени неупорядоченности структуры.

Ранее нами были проведены исследования подвижности атомов при свободном уточнении отдельных структур, а также статистический анализ подвижности атомов при свободном уточнении 203 структур разрешения более 1.2Å [12]. Эти исследования показали, что анализ атомных сдвигов при свободном уточнении может быть использован для идентификации остатков, для которых наиболее вероятно наличие альтернативных конформаций. Были предложены методы визуального анализа атомных сдвигов с помощью диаграмм и автоматические процедуры для анализа атомных сдвигов. Предложенные методы анализа были реализованы в виде компьютерных программ Shift_plot и AC_prediction [13], обрабатывающих результаты свободного уточнения, осуществленного с использованием одной из популярных программ кристаллографического уточнения phenix.refine [14].

Целью данной работы является исследование подвижности атомов в свободном уточнении при использовании для уточнения другой популярной программы кристаллографического уточнения – Refmac [15]. В ходе работы получены параметры для процедур принятия решения при анализе результатов свободного уточнения, осуществленного с помощью этой программы. Полученные данные позволили расширить возможности программы AC_prediction, сделав возможным ее применение при проведении кристаллографического уточнения программой Refmac.

1. МАТЕРИАЛЫ И МЕТОДЫ

1.1. Выбор и подготовка моделей для исследования

Для исследования подвижности атомов при свободном уточнении и для получения параметров для процедур принятия решений была проведена выборка структур из Банка белковых структур PDB [16]. Они были поделены на 2 набора, соответствующих разным зонам разрешения экспериментальных данных. В первый набор (база 1) вошли модели с разрешением от 1.1 до 1.2Å, а во второй (база 2) – модели с разрешением лучше 1.1Å. Дополнительные ограничения, использованные при выборе моделей из PDB, представлены в таблице 1. Мы использовали в качестве одного из таких ограничений кристаллографический *R*-фактор, чтобы выбрать модели, в которые альтернативные конформации были внесены наиболее достоверно.

Процесс подготовки моделей состоял из трех этапов, а именно, подготовка модели, тестовое свободное уточнение и анализ величин атомных сдвигов. Целью первого этапа было «восстановить» модель в том виде, в котором она могла находиться до введения альтернативных конформаций. Для этого из модели удалялись введенные авторами модели альтернативные конформации, обладающие меньшим коэффициентом заселенности, а оставшимся атомам задавался единичный коэффициент заселенности. После этого проводилось стандартное уточнение с помощью программы Refmac. Уточнение длилось 10 циклов, использовались анизотропные температурные факторы, жесткая привязка атомов водорода, атомы воды оставались в модели. На втором этапе проводилось тестовое свободное уточнение, состоявшее из 10 циклов Refmac. На последней стадии были рассчитаны атомные сдвиги для всех неводородных атомов структуры и собраны в базы данных вместе с дополнительной информацией, такой как PDB-код структуры, номера остатка и атома, идентификатор альтернативной конформации, величина температурного фактора и т. д. В связи с большим количеством обрабатываемых моделей все процедуры осуществлялись в автоматическом режиме с использованием специально разработанных сценариев. С некоторыми моделями не удалось провести все запланированные манипуляции. В основном это было связано с невозможностью автоматической конвертации экспериментальных данных, невозможностью автоматически распознать лиганды и т. п. Некоторые уточнения закончились с высокими значениями *R*-факторов, поэтому структуры с *R*-факторами больше 0.15 были исключены из обеих баз. Тем не менее, количество моделей, для которых удалось получить данные об атомных сдвигах, достаточно для проведения статистического исследования.

Таблица 1. Выборка структур из PDB

	База 1	База 2
Параметры выборки		
Разрешение	$1.1 \leq d \leq 1.2\text{\AA}$	$d < 1.1\text{\AA}$
<i>R</i> -фактор		
<i>R</i> _{work}	0.13	0.12
<i>R</i> _{free}	0.16	0.15
Статистика		
Структур выбрано из PDB	189	102
Структур, над которыми проведены тесты	127	59
Количество неводородных атомов	243079	122566
Количество атомов в АК	14851	12649

1.2. Процедуры принятия решений

Ранее нами было разработано несколько процедур принятия решений по результатам свободного уточнения исследуемой структуры для автоматической классификации остатков на остатки, имеющие единственную конформацию (ЕК) и остатки, имеющие альтернативные конформации (АК). Каждая процедура использует результаты тестового свободного уточнения (ТСУ) для генерации списка остатков, для которых вероятно наличие АК. В данном исследовании ТСУ – это свободное уточнение программой Refmac со следующими параметрами: 10 циклов, анизотропные температурные факторы, атомы воды включены в модель. ТСУ должно проводиться для модели с хорошим разрешением экспериментальных данных (лучше 1.2\AA), которая уже довольно хорошо уточнена, ориентиром может служить достигнутый R -фактор около 0.15.

1.2.1. Интегральная мера искажения остатка

В ТСУ каждый атом получает свой сдвиг. Для рассмотрения вопроса об альтернативных конформациях в терминах аминокислотных остатков удобно ввести интегральную характеристику, которая будет показывать мобильность некоторой группы атомов (например, атомов боковой цепи конкретного остатка). В данной работе мы использовали две такие характеристики, а именно средний атомных сдвиг, рассчитываемый как среднее арифметическое атомных сдвигов группы атомов, и максимальный атомный сдвиг группы атомов.

1.2.2. Пороговые критерии

Первым применяемым способом формирования списка остатков, для которых необходимо введение в модель АК, является сравнение интегральной меры искажения каждого остатка, полученного в ТСУ, с некоторым пороговым значением. Если мера искажения больше порогового значения, то остаток классифицируется как АК, иначе как ЕК. Пороговое значение – это параметр процедуры принятия решения, и оно может быть подобрано в процессе предварительного анализа построенной базы данных таким образом, чтобы получить наилучшее качество предсказания. Необходимо заметить, что разные типы атомов, например, атомы главной (ГЦ) или боковой цепи (БЦ), могут иметь разную подвижность в свободном уточнении, и поэтому значения порогов должны быть подобраны индивидуально для разных типов атомов.

1.2.3. Критерии, основанные на правдоподобии

Другой тип процедур принятия решений основан на принципе максимального правдоподобия. В данном подходе используются два эмпирических распределения вероятности интегральной характеристики. Первое распределение соответствует остаткам в ЕК, второе – остаткам в АК. Эти распределения были получены эмпирически из баз данных, описанных в п. 1.1. Распределения могут быть разными для разных типов атомов. Для выработки решения относительно конкретного остатка рассчитываются интегральная характеристика и вероятность ее получения для остатков в ЕК и АК. Если вероятность больше для случая АК, то остаток классифицируется как АК, в противном случае, он классифицируется как ЕК.

1.3. Оценка эффективности процедур

Процедуры принятия решений, описанные выше, могут рассматриваться как бинарные классификаторы, и к ним могут быть применены стандартные статистические методы для оценки их качества. Для подбора параметров процедур и оценки их качества каждая база была поделена на две равные части. Первая часть

(обучающая выборка) использовалась для поиска порогов и расчета эмпирических распределений сдвигов. Вторая часть (тестовая выборка) использовалась для оценки качества процедур в предположении, что назначенные в PDB ЕК и АК являются «правильными ответами». Для каждого остатка мы считали результат процедуры «присутствует в АК» как положительный и результат «присутствует в ЕК» как отрицательный. Мы предполагали, что процедура дает правильный ответ, если он совпадает с приведенным в модели, депонированной в PDB, то есть либо наш критерий предсказывает присутствие АК, и в модели, депонированной в PDB, есть АК, либо процедура сигнализирует о ЕК, и в модели, депонированной в PDB, данный остаток присутствует в ЕК. Все остальные варианты рассматривались как неправильные предсказания. Предполагалось, что у остатка есть АК главной или боковой цепи, если хотя бы один атом в этой цепи находится в АК. Остатки аланина и глицина были исключены из расчета качества предсказания для боковых цепей.

1.3.1. Статистические критерии

Две главные характеристики, которые использовались для оценки качества критериев – это вероятности ошибок первого и второго рода (fp_rate , fn_rate):

$$fp_rate = \frac{FP}{TN + FP}, \quad (1)$$

$$fn_rate = \frac{FN}{TP + FN}, \quad (2)$$

где TP (true positive) – количество правильно предсказанных АК, FP (false positive) – количество неправильно предсказанных АК, TN (true negative) – количество правильно предсказанных ЕК, FN (false negative) – количество неправильно предсказанных ЕК.

Вероятности ошибок первого и второго рода являются конкурирующими значениями, а в случае пороговых критериев зависят от выбранного значения порога. На рис. 1а показаны вероятности ошибок первого и второго рода в зависимости от величины порога для критерия, использующего средние атомные сдвиги, для боковых внутренних цепей. Две этих кривых можно объединить в ROC-кривую, представленную на рис. 1б. Каждая точка на этой кривой соответствует некоторому пороговому значению, а координаты показывают вероятность ошибок первого рода и мощность ($1 - fn_rate$) критерия. Такая кривая показывает зависимость «выигрыша» (доля правильно определенных АК) от «стоимости» (доля неправильно предсказанных АК). Хорошая процедура принятия решений должна находить большую часть АК при

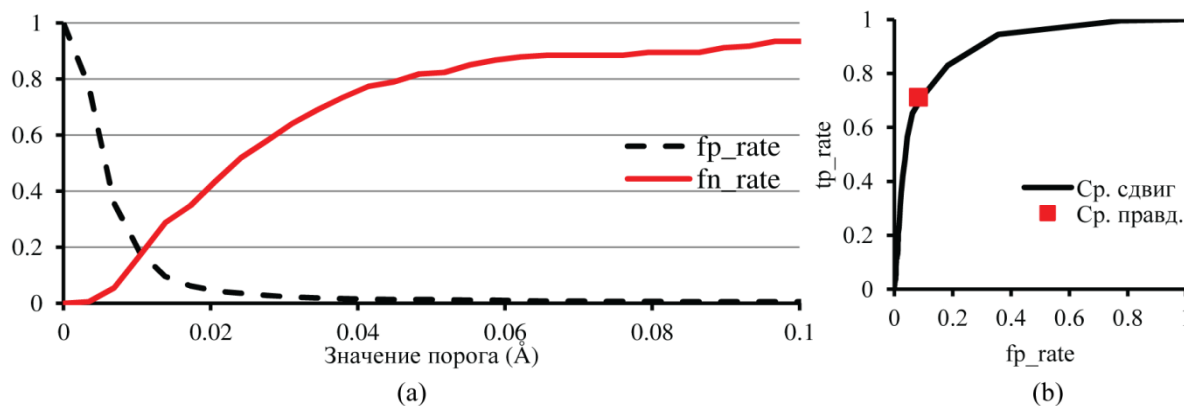


Рис. 1. Вероятности ошибок первого (черная пунктирная линия) и второго рода (красная сплошная линия) как функции величины порога (а), ROC-кривая и маркер для процедуры максимального правдоподобия с использованием средних атомных сдвигов (б). Использовался критерий на основе анализа величины среднего атомного сдвига, для боковых внутренних цепей (тестовая выборка, $0.099 < R_{work} < 0.129$).

малом количестве ложных срабатываний. Используемые в данной работе критерии, основанные на правдоподобии, не имеют параметров для оптимизации, поэтому могут быть показаны как одна точка в ROC-пространстве.

Для оценки критериев в данной работе также использовались положительное и отрицательное предсказательные значения (PPV , NPV), рассчитываемые как

$$PPV = \frac{TP}{TP + FP}, \quad (3)$$

$$NPV = \frac{TN}{TN + FN}. \quad (4)$$

В данном случае PPV показывает долю правильных предсказаний среди всех предсказанных АК, а NPV – долю правильных предсказаний среди всех предсказанных ЕК.

1.3.2. Интегральные характеристики

В пороговых критериях величины ошибок первого и второго рода являются конкурирующими и зависят от значения порога. То же справедливо и для PPV и NPV . Для удобства оценки эффективности процедур принятия решений мы использовали значение «сбалансированной точности» $bACC$,

$$bACC = \frac{(1 - fp_rate) + (1 - fn_rate)}{2} \quad (5)$$

которое было предложено ранее для оценки результатов в соревновании CASP в секции «поиск разупорядоченных участков в белках» [17]. При использовании пороговых критериев достигаемая точность $bACC$ зависит от выбранного значения порога. Эта величина использовалась нами для выбора оптимального значения порога. Порог выбирался так, чтобы максимизировать величину $bACC$, рассчитанную по обучающей выборке.

Другой интегральной характеристикой была выбрана площадь под ROC-кривой (AUC). Это значение может быть посчитано только для пороговых критериев. Большая площадь под ROC-кривой значит, что возможно выбрать порог таким образом, чтобы получить одновременно маленькие вероятности ошибок первого и второго рода. В нашем случае значение AUC эквивалентно вероятности получить большее значение интегрального сдвига для АК, чем для ЕК, если два таких остатка выбраны случайно [18].

2. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

2.1. Диаграммы атомных сдвигов

На рис. 2 представлена диаграмма атомных сдвигов для PDB структуры 1hg7 [19]. Маркерами обозначены остатки, для которых авторы структуры ввели АК. Цвет маркеров показывает уточненные значения коэффициентов заселенности для менее заселенной конформации. Легко видеть, что в ряде случаев АК были введены для остатков, для которых получены значительные величины атомных сдвигов в свободном уточнении. Более того, АК с коэффициентами заселенности, близкими к 0.5, имеют большие сдвиги, чем остатки, для которых коэффициент заселенности оказался около 0.3. В то же время не все остатки, для которых в модель были введены АК, получили большие атомные сдвиги. Такого типа диаграмма дает общее представление о степени упорядоченности различных частей структуры. Однако решение о том, какие величины сдвигов считать существенными, принимает исследователь. Описываемые ниже процедуры принятия решения помогают автоматизировать этот выбор.

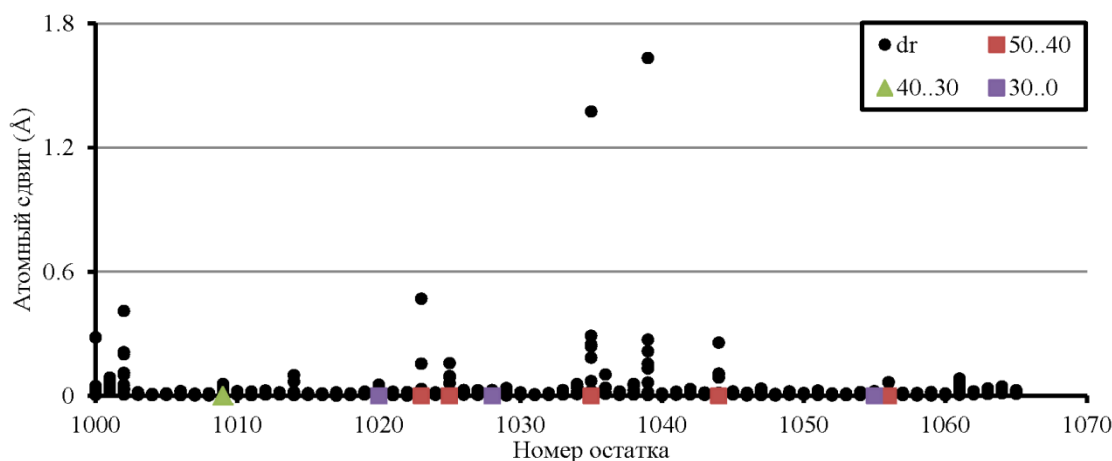


Рис. 2. Величины атомных сдвигов, полученные в ТСУ для структуры 1hg7. Каждый столбик точек представляет сдвиги неводородных атомов одного остатка. Остатки пронумерованы в соответствии с данными в PDB файле. Маркерами отмечены остатки, для которых были введены АК авторами структуры. Цвет и форма маркеров указывает на уточненные коэффициенты заселенности (в процентах) для наименее заселенной конформации.

2.2. Подвижность атомов при свободном уточнении

Все атомы в базах данных могут быть разделены на два типа, а именно атомы в ЕК и в АК. На рис. 3а показаны эмпирические распределения сдвигов для этих типов атомов для структур разрешения 1.1–1.2Å. Данные распределения значительно отличаются друг от друга, что создает возможность различать эти типы атомов по их атомным сдвигам в ТСУ. Распределение величин сдвигов атомов в ЕК имеет более выраженный пик, находящийся левее пика для АК. Распределение атомных сдвигов для АК более размыто и указывает на то, что для таких атомов более свойственны большие значения атомных сдвигов.

На рис. 3б показан средний сдвиг атомов разных типов при свободном уточнении в структурах разрешения 1.1–1.2Å. Видно, что для разных типов атомов средние значения атомных сдвигов отличаются. Кроме того, они растут с ростом R -фактора структуры. Эти результаты хорошо согласуются с полученными ранее при уточнении с помощью программы phenix.refine [12].

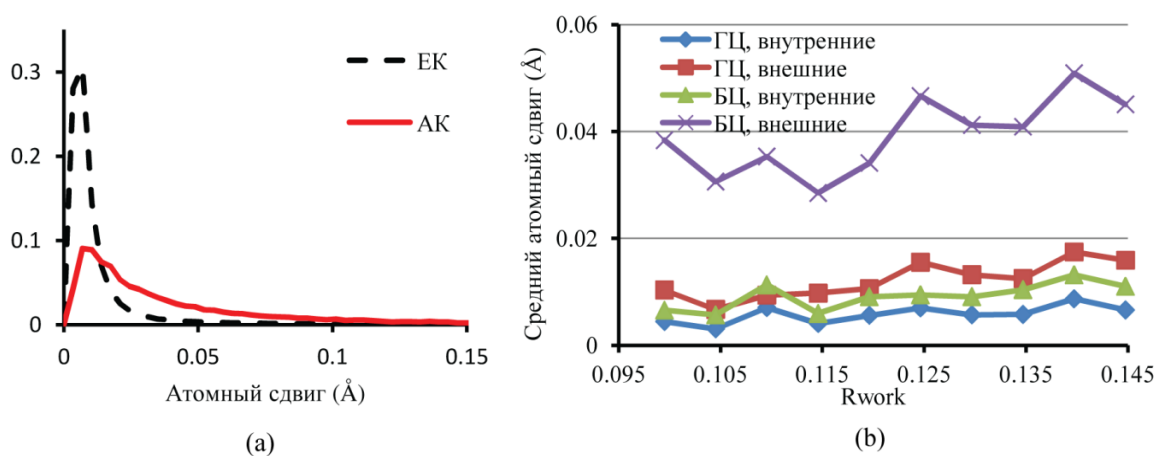


Рис. 3. Распределение атомных сдвигов для атомов в ЕК (черная пунктирная линия) и в АК (красная сплошная линия) для базы 1 (а). Средний атомный сдвиг разных типов атомов в ЕК в зависимости от величины R -фактора (база 1) (б).

2.3. Оценка процедур принятия решений

В ходе данной работы было протестировано пять процедур принятия решений – два пороговых: для среднего (ср_пор) либо максимального (макс_пор) атомного сдвига, и три, основанных на правдоподобии: вероятности среднего (ср_правд), максимального (макс_правд) либо индивидуальных (инд_правд) атомных сдвигов. Результаты тестирования представлены в таблицах 2 и 3. В таблицах 4 и 5 представлены результаты оценки порогового критерия на основе средних атомных сдвигов. На рис. 4 показаны ROC-кривые для четырех классов атомов. Величины ошибок, полученные для критериев, основанных на правдоподобии, показаны маркерами.

Тестирование процедур показало, что результаты, как и ожидалось, имеют ошибки обоих типов (ложные срабатывания и пропущенные АК). Качество предсказаний также отличается для разных типов атомов. Наилучшие предсказания были получены для боковых цепей, находящихся в центре белковой глобулы. Качество предсказаний для всех пяти процедур принятия решений оказалось близким, и не дает оснований для выбора какой-либо одной из них, как наилучшей.

Таблица 2. Сбалансированная точность (*bACC*) для различных критериев (база 1, тестовая выборка)

Вид критерия	Типы атомов					
	ГЦ, внутр.	ГЦ, внеш.	БЦ, внутр.	БЦ, внеш.	ГЦ	БЦ
ср_пор	0.697	0.664	0.817	0.730	0.684	0.787
макс_пор	0.687	0.670	0.797	0.715	0.693	0.773
ср_правд	0.650	0.619	0.792	0.687	0.640	0.743
макс_правд	0.690	0.649	0.781	0.663	0.695	0.721
инд_правд	0.650	0.640	0.760	0.688	0.650	0.735

Таблица 3. Площадь под ROC-кривой (*AUC*) для пороговых критериев (база 1, тестовая выборка, $0.099 < Rwork < 0.129$)

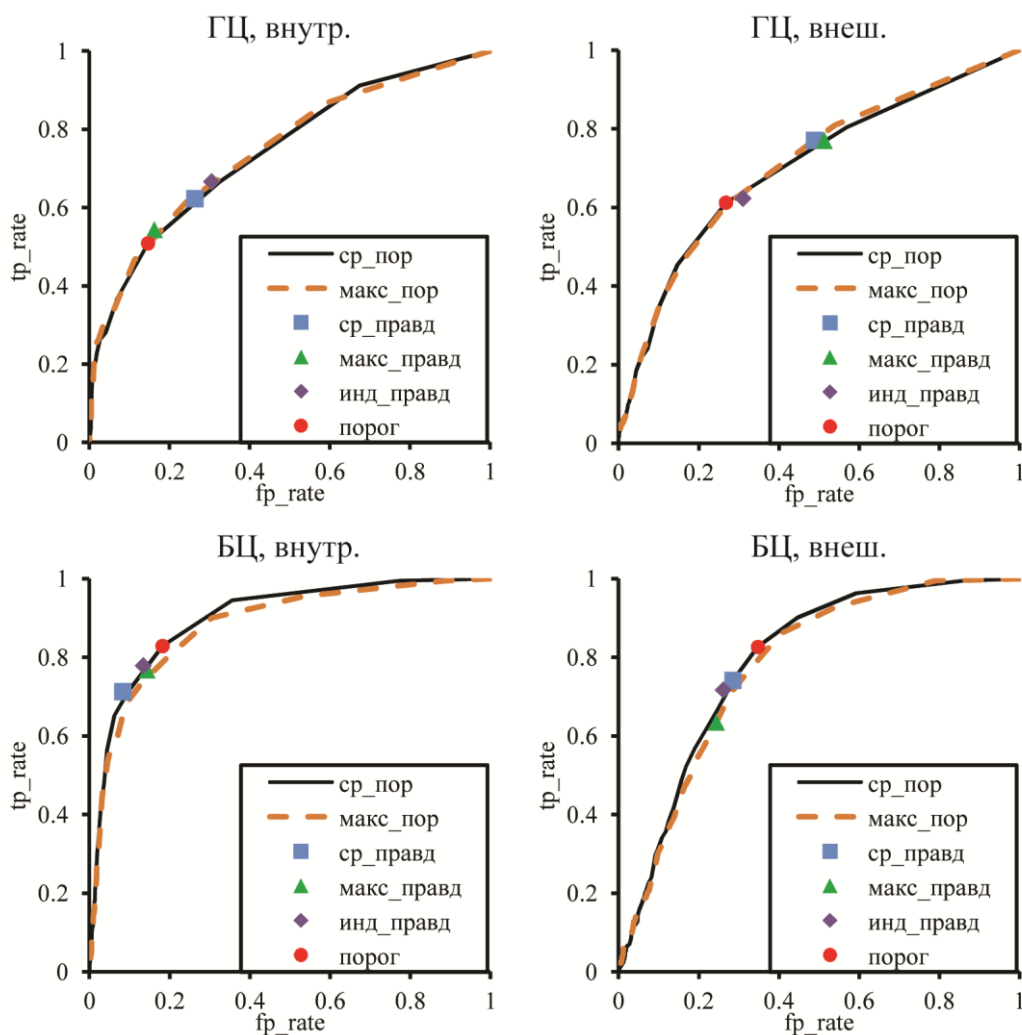
Вид критерия	Типы атомов			
	ГЦ, внутр.	ГЦ, внеш.	БЦ, внутр.	БЦ, внеш.
ср_пор	0.744	0.703	0.898	0.789
макс_пор	0.746	0.708	0.881	0.777

Таблица 4. Результаты тестирования порогового критерия, основанного на анализе средних величин атомных сдвигов для базы 1 (1.1–1.2Å), тестовая выборка

Типы атомов	<i>bACC</i>	fp_rate	fn_rate	<i>PPV</i>	<i>NPV</i>	Доля АК
ГЦ, внутр.	0.697	0.184	0.422	0.079	0.986	0.027
ГЦ, внеш.	0.664	0.235	0.437	0.144	0.962	0.066
БЦ, внутр.	0.817	0.190	0.176	0.207	0.987	0.057
БЦ, внеш.	0.730	0.386	0.155	0.280	0.957	0.151
ГЦ (итог)	0.684	0.202	0.431	0.107	0.978	0.041
БЦ (итог)	0.787	0.263	0.162	0.249	0.978	0.094

Таблица 5. Результаты тестирования порогового критерия, основанного на анализе средних величин атомных сдвигов для базы 2 ($d < 1.1 \text{ \AA}$), тестовая выборка

Типы атомов	$bACC$	fp_rate	fn_rate	PPV	NPV	Доля АК
ГЦ, внутр.	0.841	0.115	0.204	0.312	0.985	0.062
ГЦ, внеш.	0.778	0.271	0.173	0.297	0.968	0.122
БЦ, внутр.	0.868	0.114	0.149	0.455	0.981	0.101
БЦ, внеш.	0.753	0.350	0.144	0.460	0.928	0.259
ГЦ (итог)	0.822	0.169	0.188	0.304	0.980	0.083
БЦ (итог)	0.828	0.197	0.146	0.458	0.966	0.163

**Рис. 4.** Результаты тестирования критериев (база 1, тестовая выборка, $0.099 < Rwork < 0.129$) для разных типов атомов. Сплошные и пунктирные линии показывают ROC-кривые для пороговых критериев на основе анализа средних (sr_por) и максимальных (maks_por) атомных сдвигов соответственно. Прямоугольники, треугольники и ромбы показывают результативность критериев, основанных на максимизации правдоподобия средних (sr_prawd) и максимальных (maks_prawd) атомных сдвигов и индивидуальных атомных сдвигов (ind_prawd). Круглым маркером (porog) показан оптимальный порог для порогового критерия на основе анализа средних атомных сдвигов.

ЗАКЛЮЧЕНИЕ

Тестовое свободное уточнение и последующий анализ величин атомных сдвигов может выявить наиболее нестабильные участки структуры. Эти участки являются наиболее вероятными кандидатами для введения в атомную модель структуры альтернативных конформаций, и поэтому должны быть проверены (с использованием карт распределения электронной плотности) в первую очередь. Предложенные формальные процедуры хотя и не позволяют абсолютно точно определить остатки, для которых необходимо введение альтернативных конформаций, но позволяют существенно сократить количество остатков, которые необходимо обработать вручную.

В представленной работе были получены пороговые значения и эмпирические распределения, расширяющие область применения разработанных ранее подходов, делая возможным их использование при проведении кристаллографического уточнения популярной программой уточнения Refmac. Эти данные были добавлены в программы для анализа значений атомных сдвигов. Программы и документация доступны по адресу: www.impb.ru/lmc/programs/ac_prediction/

Работа выполнена при поддержке РФФИ, грант № 12-04-31096.

СПИСОК ЛИТЕРАТУРЫ

- Howard E.I., Sanishvili R., Cachau R.E., Mitschler A., Chevrier B., Barth P., Lamour V., Van Zandt M., Sibley E., Bon C., Moras D., Schneider T.R., Joachimiak A., Podjarny A. *Proteins*. 2004. V. 55. P. 792–804.
- Ševčík J., Dauter Z., Wilson K.S. *Acta Crystallographica D*. 2004. V. 60. P. 1198–1204.
- Wang J., Dauter M., Alkire R., Joachimiak A., Dauter Z. *Acta Crystallographica D*. 2007. V. 63. P. 1254–1268.
- Rypniewski W.R., Østergaard P., Nørregaard-Madsen M., Dauter M., Wilson K.S. *Acta Crystallographica D*. 2001. V. 57. P. 8–19.
- Getzoff E.D., Gutwin K.N., Genick U.K. *Nat. Struct. Biol.* 2003. V. 10. P. 663–668.
- Kursula I., Wierenga R.K. *J. Biol. Chem.* 2003. V. 278. P. 9544–9551.
- Addlagatta A., Krzywda S., Czapinska H., Otlewski J., Jaskolski M. *Acta Crystallographica D*. 2001. V. 57. P. 649–663.
- Dauter Z., Sierer L.C., Wilson K.S. *Acta Crystallographica D*. 1992. V. 48. P. 42–59.
- Dodd F.E., Hasnain S.S. *Acta Crystallographica D*. 1995. V. 51. P. 1052–1064.
- Koepke J., Scharff E.I., Lücke C., Rüterjans H., Fritsch G. *Acta Crystallographica D*. 2003. V. 59. P. 1744–1754.
- Соболев О.В., Лунин В.Ю. *Математическая биология и биоинформатика*. 2008. Т. 3. С. 50–59. URL: [http://www.matbio.org/downloads/Sobolev2008\(3_50\).pdf](http://www.matbio.org/downloads/Sobolev2008(3_50).pdf) (дата обращения 21.12.2012).
- Sobolev O.V., Lunin V.Y. *Acta Crystallographica D*. 2012. V. 68. P. 1118–1127.
- Sobolev O.V. *Computational Crystallography Newsletter*. 2012. V. 3. P. 32–34.
- Afonine P.V., Grosse-Kunstleve R.W., Echols N., Headd J.J., Moriarty N.W., Mustyakimov M., Terwilliger T.C., Urzhumtsev A., Zwart P.H., Adams P.D. *Acta Crystallographica D*. 2012. V. 68. P. 352–367.
- Murshudov G.N., Skubak P., Lebedev A.A., Pannu N.S., Steiner R.A., Nicholls R.A., Winn M.D., Long F., Vagin A.A. *Acta Crystallographica D*. 2011. V. 67. P. 355–367.
- Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. *Nucleic Acids Res.* 2000. V. 28. P. 235–242.
- Noivirt-Brik O., Prilusky J., Sussman J.L. *Proteins*. 2009. V. 77. Suppl. 9. P. 210–216.
- Fawcett T. *Pattern Recognition Letters*. 2006. V. 27. P. 861–874.

19. Antson A.A., Smith D.J., Roper D.I., Lewis S., Caves L.S., Verma C.S., Buckley S.L., Lillford P.J., Hubbard R.E. *J. Mol. Biol.* 2010. V. 305. P. 875–889.

Материал поступил в редакцию 18.12.2012, опубликован 25.12.2012.