

УДК 577.322.2:004.94

КОМПЬЮТЕРНЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ ТЕРМОСТАБИЛЬНОСТИ БЕЛКОВ И ИХ ПРИМЕНЕНИЕ В БИОЛОГИИ

© 2012 г. Н.А. Алемасов, Э.С. Фомин

Федеральное государственное бюджетное учреждение науки Институт цитологии и генетики
Сибирского отделения Российской академии наук, Новосибирск, Россия,
e-mail: alemasov@bionet.nsc.ru

Поступила в редакцию 15 июля 2012 г. Принята к публикации 31 августа 2012 г.

Проведена классификация существующих теоретических подходов к исследованию термостабильности белков. Методы компьютерного моделирования динамики белков позволяют наиболее полно оценить микро- и макропараметры исследуемых молекул, но без уточнений (например, в части учета распределения заряда) они ограничены в точности получаемых результатов. Перспективными ввиду низкой требовательности к вычислительным ресурсам являются также методы выделения жестких фрагментов белков. Они позволяют напрямую определить изменения в гибкости структуры при совершении аминокислотных замен. Но их недостатком является сложность оценки влияния окружающего растворителя и его термодинамических параметров.

Ключевые слова: термостабильность белков, свободная энергия, молекулярная динамика, «жесткие» фрагменты, электростатические и гидрофобные взаимодействия, термофилы, мезофилы.

ВВЕДЕНИЕ

Термостабильность белков отражает способность сохранять уникальную пространственную структуру полипептидной цепи под действием высокой температуры. Численно термостабильность определяется через разности энтальпии и энтропии нативного и денатурированного состояния белков (Matthews *et al.*, 1987; Talluri, 2011) и может рассматриваться с точки зрения термодинамических свойств составляющих белки аминокислот (Tanford, 1962). Существуют экспериментальные методы улучшения термостабильности белка: путем стабилизации его нативной формы, дестабилизации его денатурированной формы или комбинацией обоих подходов (Matthews, 1993).

Понимание факторов стабилизации белков, прежде всего, термофильных организмов, важно не только для выявления физико-химических принципов белкового фолдинга или механизмов стабилизации структуры и взаимодействия

белков. Это также является необходимым условием для эффективного проектирования новых ферментов, способных работать при высоких температурах, что полезно в промышленности (Vogt *et al.*, 1997).

В фундаментальных исследованиях термостабильные белки часто используются в процессе биологического анализа, например в качестве биосенсоров. Кроме того, большое значение представляет оптимизация ДНК-полимераз (Talluri, 2011), которые используются в ПЦР-реакциях (реакциях полимеразной цепной реакции) (Saiki *et al.*, 1985), в направлении их большей термостабильности, специфичности и процессивности. Существуют и эволюционные аспекты, обуславливающие повышенную термостабильность у термофильных организмов (Afonnikov *et al.*, 2001; Афонников и др., 2011).

Экспериментальные методы исследования термостабильности белков могут предоставить весьма подробную информацию о связи

структурных модификаций молекулы и соответствующих ей изменений термостабильности. Но и здесь существуют ограничения (Tidor *et al.*, 1991). Рассматривая кристаллические структуры как таковые, в эксперименте сложно измерить возможный энтропийный вклад в свободную энергию, обусловленный внутренними движениями и влиянием аминокислотных замен на денатурированное состояние. Кроме того, изменения в стабильности могут быть обусловлены различными взаимодействиями, дающими как положительный, так и отрицательный вклад в свободную энергию, который также сложно отделить экспериментально.

Для преодоления данных ограничений применяют теоретические методы, которые позволяют более детально рассмотреть термодинамику белков, а также влияние замен на термостабильность. Настоящая работа направлена на классификацию существующих теоретических подходов к исследованию термостабильности белков. По каждому классу методов сделаны выводы о применимости, ограничениях и получаемых с помощью методов этого класса результатах.

НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ

В работе А. Mozo-Villiaras и Е. Querol (2006) предлагается рассматривать три (возможно, пересекающихся) направления исследований термостабильности белков: сопоставление белков термофильных и мезофильных организмов; *ab initio* подходы; «доминирующие» взаимодействия. В настоящей работе предлагается расширить этот список еще одним классом методов.

Ab initio подходы (например метод молекулярной динамики (Alder, Wainwright, 1957)), а также менее трудоемкие методы, основанные на учете только «доминирующих» взаимодействий, требуют высоких вычислительных затрат. Поэтому в последние годы развиваются методы, основанные на выделении жестких фрагментов молекулы и изучении связи «жесткость/гибкость → термостабильность». Поскольку последние предполагают лишь статический анализ, они серьезно выигрывают по производительности у методов молекулярного моделирования.

Таким образом, список рассматриваемых в настоящей работе классов методов выглядит следующим образом.

1. Сопоставление белков термофильных и мезофильных организмов напрямую, с тем чтобы выявить закономерности, ведущие к их температурной стабилизации (Perutz, Raidt, 1975; Daniel *et al.*, 1982; Vogt *et al.*, 1997; Kumar *et al.*, 2000, Szilagy, Zavodszky, 2000);

2. Теоретические методы, учитывающие «доминирующие» эффекты (электростатические и гидрофобные взаимодействия), как наиболее существенные среди всех взаимодействий, влияющих на термостабильность (Beveridge, DiCapua, 1989; Sharp, Honig, 1990; Zhu, Elcock, 2010; Nicholls *et al.*, 1991; Xiao, Honig, 1999);

3. Компьютерные алгоритмы, которые учитывают физические аспекты, влияющие на температурные характеристики белков (Tidor, Karplus, 1991; Guerois *et al.*, 2002; Gromiha, 2003; Khechinashvili *et al.*, 2006; Dehouck *et al.*, 2009; Potapov *et al.*, 2009; Seeliger, DeGroot, 2010);

4. Выделение «жестких фрагментов» для определения связи между гибкостью пространственной структуры белков и ее стабильностью (Radestock, Gohlke, 2011; Rader *et al.*, 2012).

Сравнение белков мезофилов и термофилов

Исторически сложилось так, что наиболее ранние работы для изучения термостабильности белков были направлены на сравнение белков мезофильных организмов и их термофильных аналогов. Одной из первых теоретических работ является работа М. Perutz и Н. Raidt (1975). Интересные результаты получены в работах R.M. Daniel с соавт. (1982), G. Vogt с соавт. (1997).

На наш взгляд, одной из наиболее показательных работ по направлению сравнения белков термофильных и мезофильных организмов является работа А. Szilagy и Р. Zavodszky (2000). В ней была исследована зависимость структурных характеристик белков от температуры, при которой существуют содержащие эти белки организмы. Авторы использовали наборы белков, представляющих 25 семейств, и рассчитали 13 структурных параметров, основывающихся на атомных координатах. Все параметры были объединены в следующие группы.

– «Вогнутость» белковой поверхности. Рассчитывалась с помощью программы Molecular Surface Package (Connolly, 1993).

– Водородные связи. Оценивались их число, количество непарных доноров и акцепторов. Расчет производился с помощью алгоритма Hb2 пакета молекулярного моделирования WHAT IF (Vriend *et al.*, 1990).

– Ионные пары. Рассчитывалось число ионных пар. Два противоположно заряженных остатка считались ионной парой, если кратчайшее расстояние между их противоположно заряженными атомами не превышало заранее установленного значения (4, 6 или 8 Å).

– Параметры вторичной структуры: доля α -спиралей, β -листов, нерегулярных областей. Рассчитывались с помощью программы DSSP (Kabsch, Sander, 1983).

– Полярные/неполярные, внешние/внутренние области белка. Рассматривалось отношение площадей поверхности полярных и неполярных областей для внешних и внутренних областей белка. Все параметры подсчитывались как функции температуры оптимального роста соответствующих организмов. Для расчета использовалась программа WHAT IF. Атомы N и O считались полярными, остальные – неполярными. Внутренние области отдельно рассчитывались для денатурированной и нативной конформации полипептидной цепи, полученные площади вычитались.

Исходя из рассчитанных значений для всего набора белков, А. Szilagyí и Р. Zavodszky (2000) сделали вывод о корреляции каждого из параметров с рядом характеристик. Среди них: оптимальная температура роста организма, из которого выделен исследуемый белок; изменение структурных параметров в белках умеренно термофильных организмов; изменение структурных параметров в белках крайне термофильных организмов.

Таким образом, были выявлены параметры, позволяющие отделить друг от друга умеренно и крайне термофильные организмы. Так, показано, что с ростом температуры число ионных пар увеличивается для обоих типов организмов. Кроме того, количество «вогнутостей» в белках крайне термофильных организмов значительно снижено, в то время как для умеренно термофильных организмов заметного изменения по

этому параметру не наблюдается в сравнении с белками мезофилов.

На наш взгляд, рассмотренный подход к выявлению связи термостабильности белков с изменениями их структурных характеристик при сравнении термофилов и мезофилов способен дать только качественную информацию, которая может быть использована (с невысокой точностью) для нахождения общих закономерностей при разработке более точных методов. Ввиду того что данный подход основан на изучении ограниченного числа белков семейств крайне и умеренно термофильных организмов, он не способен оценить количественно изменение термостабильности конкретных белков, т. е. заранее предсказать это изменение.

Доминирующие взаимодействия

Ряд исследователей полагают, что доминирующее влияние на стабильность пространственной структуры (т. е. на термостабильность) белков оказывают электростатические и гидрофобные силы (Beveridge, DiCarua, 1989; Sharp, Honig, 1990; Zhu, Elcock, 2010).

Электростатические взаимодействия. В работе L. Xiao и В. Honig (1999) подробно рассматриваются электростатические взаимодействия и их связь с термостабильностью. Авторы используют методы статистической механики для оценки множества состояний ионизации в нативном и денатурированном белке. Так, средний заряд каждой ионизируемой группы вычислен из среднего значения, полученного из всех состояний ионизации белка. Электростатическая энергия каждого ионизационного состояния, в свою очередь, определялась с помощью решения конечно-разностной формы уравнения Пуассона–Больцмана.

Кроме того, в рассматриваемой работе влияние электростатических взаимодействий складывалось из трех частей: $\Delta\Delta E(solv)$ – дестабилизирующий вклад в стабильность белка из-за десольватации ионизирующихся групп в процессе фолдинга; $\Delta\Delta E(hb)$ – сумма взаимодействий от водородных связей между заряженными и полярными группами; $\Delta\Delta E(cc)$ – кулоновские взаимодействия между заряженными группами. Следовательно, $\Delta\Delta E(elec) = \Delta\Delta E(solv) + \Delta\Delta E(hb) + \Delta\Delta E(cc)$.

Хотя диэлектрическая проницаемость воды при 20 °С равна 80 Ф/м, авторы проводили расчеты с $\epsilon = 4$ и $\epsilon = 20$. Под влиянием соли и при изменении диэлектрической константы раствора величина электростатической свободной энергии значительно изменяется. Результаты расчетов показывают, что при $\epsilon = 20$ полное изменение свободной энергии электростатических взаимодействий в процессе фолдинга для большинства случаев отрицательно. Кроме того, практически для всех белков термофильных организмов внутри одного семейства значение $\Delta\Delta G(elec)$ меньше, чем для мезофильных, что позволяет разделять соответствующие белки независимо от диэлектрической константы ϵ . Таким образом, действительно, электростатические взаимодействия вносят большой вклад в термостабильность белков гипертермофильных организмов при pH = 7 в каждом из четырех изученных семейств белков.

Другими словами, в гипертермофильных белках вклад электростатических взаимодействий в свободную энергию фолдинга более существенный, чем в их мезофильных аналогах. Этот вклад зависит от значения pH и, как считают L. Xiao и B. Honig (1999), может исчезнуть в результате изменения распределения заряда белка при определенных pH.

Гидрофобные взаимодействия. В соответствии с работой H. Dong с соавт. (2008) влияние гидрофобности на стабильность белков сводится к следующему: при замене гидрофобного остатка в гидрофобном ядре белка на полярный стабильность всей структуры падает, а при заполнении «вогнутостей» поверхности белка гидрофобной аминокислотой структура становится более стабильной. Особенно этот эффект характерен для замен в гидрофобном ядре белка. Изучение с разных сторон влияния гидрофобных взаимодействий на стабильность белковых структур проводилось в работах D. Eisenberg с соавт. (1986); S. Kumar, M. Meenatchi (2011).

Детальное изучение гидрофобных взаимодействий проводилось в работе S. Zhu, A. Elcock (2010), в которой с помощью метода молекулярной динамики сделана попытка напрямую выявить термодинамические особенности гидрофобных взаимодействий на примере ассоциации пар «ацетат–метиламмоний» и «метан–метан». Для моделирования использовалась программа

GROMACS 3.3 (van der Spoel *et al.*, 2005). Значения температуры устанавливались в ряд от –12,5 °С до 112,5 °С с шагом 12,5 °С. Использовался явный растворитель – водный раствор в рамках моделей TIP3P и TIP5P. Моделируемая область имела размеры $25 \times 25 \times 25 \text{ \AA}^3$. Рассчитывалась динамика системы в течение 500 нс для каждого значения температуры.

Из полученных «снимков» рассчитывались гистограммы: для пары «ацетат–метиламмоний» вычислялось расстояние между атомом углерода карбоксильной группы молекулы ацетата и атомом азота аминогруппы молекулы метиламмония; для пары «метан–метан» – расстояние между двумя углеродами. Кроме того, для пары ацетат–метиламмоний также рассчитывалась гистограмма по двум размерностям: по расстояниям между заряженными группами двух молекул (описывают электростатические взаимодействия) и между их метильными группами (описывают гидрофобные взаимодействия). Для 1D и 2D-гистограмм рассчитывались свободные энергии взаимодействия. Причем для расчетов было смоделировано два набора гистограмм с взаимодействующими между собой молекулами и с «выключенными» межмолекулярными взаимодействиями. Это было сделано для того, чтобы вычесть вклад конфигурационной энтропии в свободную энергию.

Авторы показали, что результаты предсказаний изменений свободной энергии в рамках моделей TIP3P и TIP5P хорошо согласуются между собой для температур от 25 до 100 °С. Кроме того, в рамках модели воды TIP5P относительная прочность гидрофобных взаимодействий и солевых мостиков по большей части не изменяется в промежутке от 0 до 40 °С. При более высоких температурах солевые мостики дают более весомый вклад в стабильность.

На наш взгляд, основным недостатком примененного подхода является то, что для расчета только доминирующей части свободной энергии используется весьма ресурсоемкий метод молекулярной динамики. В то же время поправки, связанные с вкладами гидрофобных и электростатических взаимодействий, могут быть введены в методы компьютерного моделирования, рассмотренные ниже, с целью учета всех вкладов в свободную энергию.

Компьютерное моделирование

Задача расчета термостабильности в данном подходе сводится к вычислению разности свободных энергий между различными состояниями моделируемой системы. Так, например, расчет термостабильности мутантной формы белка требует вычисления величины двойной разности $\Delta\Delta G = \Delta G_3 - \Delta G_2$, где ΔG_3 – разность свободных энергий между глобулярной и денатурированной формами нативного белка, а ΔG_2 – разность свободных энергий между глобулярной и денатурированной формой мутанта (см. рис. 1). Данный подход успешно был применен в работе V. Tidor, M. Karplus (1991).

Свободная энергия является функцией состояния и не зависит от пути, который проходит система. Поэтому исходя из условия равенства нулю изменения свободной энергии вдоль всего цикла искомая величина $\Delta\Delta G$ может быть рассчитана также через разность $\Delta\Delta G = \Delta G_1 - \Delta G_4$. Величины ΔG_1 и ΔG_4 связаны с нефизическим, т. е. нереализуемым в природе, непрерывным процессом «превращения» одних аминокислотных остатков в другие (см. рис. 1). Ключевой особенностью этого «превращения» (или «алхимического» процесса) является использование метода λ -динамики, в котором в один и тот же момент существуют физические взаимодействия от заменяемой и заменяющей части молекулы (Kong, Brooks, 1996).

В работе D. Seeliger, B. De Groot (2010), которая посвящена «алхимическим» расчетам свободной энергии, с помощью распространенной программы GROMACS (Hess *et al.*, 2008) и силового поля amber99sb (Hornak *et al.*, 2006) было исследовано свыше 100 мутантных белков

барназы. Результаты расчетов показали лучшее согласие с экспериментальными данными, чем в работе R. Guerois с соавт. (2002). Корреляция рассчитанных и экспериментальных данных составила 0,86, а средняя абсолютная ошибка – 3,31 кДж/моль, что говорит о большем потенциале подходов «алхимических» расчетов.

Одним из подходов, реализующих расчет свободной энергии, является использование эмпирических энергетических функций (Guerois *et al.*, 2002). Но подобные функции часто дают неточные результаты при оценке энергии. Это происходит в случае расчетов свободной энергии белков, не входящих в обучающий набор или при оценке энергии замен, приводящих к значительному изменению конформации белка. Другой вариацией подходов к расчету свободной энергии является метод, основанный на упрощенных моделях, например, представляющий растворитель неявно (Vogobjev, 2011).

В нашей недавней работе (Фомин, Алемасов, 2012) были представлены результаты дальнейшего развития методов «алхимических» преобразований: в расчетную схему были включены эффекты поляризации. Установлено, что учет данных эффектов улучшает точность расчетов разностей свободных энергий для мутаций, включающих изменение заряда аминокислоты. Так, при расчете влияния замены R72G в барназе была существенно улучшена величина изменения свободной энергии в сравнении со значением, полученным в работе D. Seeliger, B. De Groot (2010). Полученное нами значение $-19,8$ кДж/моль находится заметно ближе к экспериментальному для этой мутации значению $-10,4$ кДж/моль, чем величина $-57,67$ кДж/моль, рассчитанная в работе D. Seeliger, B. De Groot (2010).

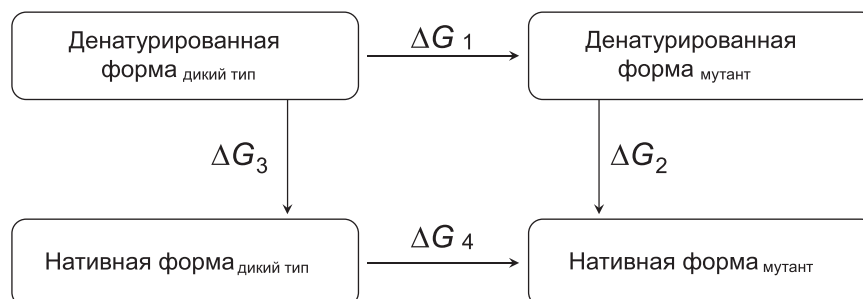


Рис. 1. Термодинамический цикл. По: D. Seeliger, B. De Groot (2010).

Выделение жестких фрагментов

В работе А.Т. Rader с соавт. (2012) аналогично подходу, использованному в работе S. Raderstock, Н. Gohlke (2011), было сделано исследование связи аллостерии и жесткости структуры белка с его термостабильностью. Ключевым вопросом в этой работе было рассмотрение мутации A35V в GH12 эндоглюканазе *Trichoderma reesei*. Примечательно то, что эта замена делает белок способным выдерживать температуры на 8 °С выше, чем белок дикого типа. При этом разница в RMSD между этими двумя формами белка составляет всего порядка 0,4 Å.

Гибкость и стабильность биомолекулярных систем в этой работе рассчитывалась с помощью программы FIRST (Jacobs *et al.*, 2001). Модель, реализованная в FIRST, рассматривает физические взаимодействия между парами атомов в качестве ограничений (например ковалентные связи). Из этих ограничений формируется трехмерный граф, включающий полный набор взаимодействий, присутствующих в нативном состоянии белка. Жесткие фрагменты соответствуют тем наборам атомов, которые соединены так, что любое внутреннее движение нарушает хотя бы одно из ограниче-

ний. Остальные области молекулы рассматриваются как гибкие.

В FIRST каждая связь помечается как гибкая (способная вращаться) или жесткая (фиксированная относительно вращения). В соответствии с гибкостью связей определяется гибкость/жесткость аминокислотных остатков. С помощью данных авторы сравнили нативную структуру со структурой мутанта (на примере эндоглюканазы) и определили остатки, меняющие свою жесткость в мутантной структуре.

Так как в модели не учитывается температура, авторы ввели параметр E_{cut} – энергия разрыва водородной связи. Параметр рассчитывается для всех неявных водородных связей, основываясь на локальной геометрии конкретного донора, акцептора и атомов водорода. В зависимости от E_{cut} вводится условная температура как $\Theta = 300 - 20E_{cut}$.

Выбрав температуру денатурации нативного белка в качестве температуры для модели (346 К), Rader с соавт. (2012) рассчитали долю атомов в самом крупном жестком фрагменте (см. рис. 2). Очевидным недостатком данного подхода является невозможность оценить влияние растворителя на термостабильность белка, а также pH раствора и других физических характеристик (давление и т. п.).

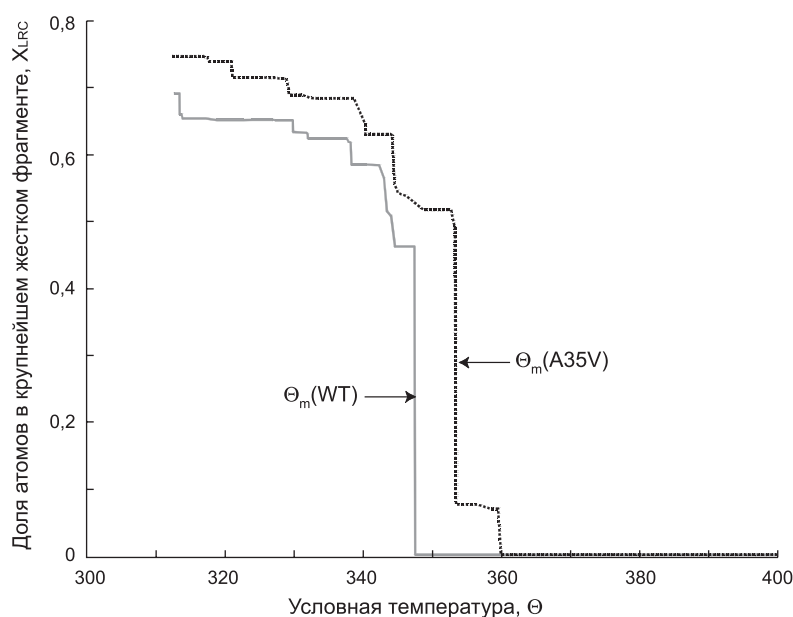


Рис. 2. Доля атомов в наиболее крупном жестком фрагменте в зависимости от условной температуры. По: Rader с соавт. (2012).

Таблица

Сведения о рассматриваемых подходах

Класс	Метод	Объект	Результаты
1	Сравнение последовательностей и пространственных структур гомологичных белков мезофилов и термофилов (Perutz, Raidt, 1975)	Бактериальный ферредоксин, гемоглобины А и А2 человека	В термостабильном ферредоксине найдены замены, формирующие водородные связи и солевые мостики. Гемоглобин А2 более стабилен, чем гемоглобин А за счет дополнительной связи между $\alpha 1$ и $\delta 1$ субъединицами
	Расчет доли полярных областей белка. Расчет числа водородных связей и солевых мостиков (Vogt <i>et al.</i> , 1997)	16 семейств белков мезофилов и термофилов	В 80 % случаев есть корреляция внутри семейства между термостабильностью и числом водородных связей. В 67 % случаев есть корреляция термостабильности с формированием ионных пар
	Расчет структурных параметров на основе атомных координат (Szilagyí, Zavodszky, 2000)	25 семейств белков умеренно и крайне термофильных организмов	Найдены совокупности параметров, позволяющие выделять термофильные организмы: например, с ростом температуры число ионных пар увеличивается
2	Статистическая механика для расчета изменения электростатической свободной энергии в процессе фолдинга (Xiao, Honig, 1999)	GDN, GAPDH, ферредоксин, CheY мезофилов и термофилов	Для термофильных белков внутри одного семейства свободная энергия меньше, чем для мезофильных. Опровергнута корреляция термостабильности и числа ионных пар
	Расчет площадей, доступных растворителю для аминокислот, расчет свободной энергии сольватации (Eisenberg <i>et al.</i> , 1986)	20 аминокислот	Корреляция рассчитанного изменения свободной энергии сольватации с экспериментом, равная 0,88
	Молекулярная динамика (GROMACS 3.3), потенциал средней силы (Zhu, Elcock, 2010)	Ацетат, метиламмоний, метан	Прочность гидрофобных взаимодействий и солевых мостиков не изменяется от 0 до 40 °С. При более высоких температурах солевые мостики доминируют в термостабильности
	Расчет кинетических и термодинамических величин (Kumar, Meenatchi, 2011)	20 аминокислот, цитохром b	Чем выше количество гидрофобных аминокислот, тем выше становится вклад в энергию Гиббса при кинетических и термодинамических расчетах
3	Молекулярная динамика (CHARMM), λ -динамика (Tidor, Karplus, 1991)	Лизоцим бактериофага T4, мутация R96H	Для исследуемой мутации получено значение $-7,9$ кДж/моль, что качественно соответствует эксперименту ($-13,4 \pm 5$ ккал/моль)
	Энергетическая функция (Guerois <i>et al.</i> , 2002)	1088 замен в различных белках	Наклон прямой линейной регрессии $-0,64$, коэффициент корреляции с экспериментом $-0,73$, стандартное отклонение $-4,27$ кДж/моль
	Молекулярная динамика (GROMACS 4), λ -динамика (Seeliger, De Groot, 2010)	109 замен барназы	Корреляция рассчитанных и экспериментальных данных $-0,86$, средняя абсолютная ошибка $-3,31$ кДж/моль
4	Анализ сети с ограничениями (CNA), расчет доли атомов в жесткой части сети (Radestock, Gohlke, 2011)	19 семейств белков мезофильных и термофильных организмов	Около 67 % семейств белков имели температуру денатурации выше для термофильных гомологов
	Анализ жесткости структуры (FIRST), нахождение наибольшего жесткого фрагмента (Rader <i>et al.</i> , 2012)	GH12 эндоглюканаза	Экспериментальное значение изменения температуры денатурации ($7,7$ °С) качественно подтверждается рассчитанным значением ($5,9$ °С)

ВЫВОДЫ

В настоящей работе представлен обзор наиболее распространенных методов теоретического исследования термостабильности. Уточнена классификация существующих подходов и представлены результаты, которые могут быть получены с их использованием. Основные сведения о рассмотренных подходах представлены в таблице. Так, наиболее распространенным методом исследования термостабильности белков является метод сравнения гомологов термофильных и мезофильных организмов. Этот подход естественным образом развился из экспериментальных методов и позволяет выявить самые существенные характеристики, которые влияют на способность белков функционировать при более высоких температурах.

В последнее время развиваются методы компьютерного моделирования динамики белков, которые позволяют наиболее полно оценить микро- и макропараметры исследуемых молекул, что обусловлено использованием базовых физических соотношений и уравнений движения. Методы молекулярного моделирования, однако, ограничены в точности и в будущем их прямое, без модификаций, использование вряд ли позволит получать более близкие к эксперименту результаты.

Необходимость модификаций методов моделирования также вытекает из существования доминирующего вклада в свободную энергию электростатических и гидрофобных взаимодействий. Уточнение параметризации электростатических и гидрофобных взаимодействий в потенциалах, определяющих движение атомов, является перспективным улучшением моделей динамики белков.

Также перспективными ввиду низкой требовательности к вычислительным ресурсам являются методы выделения жестких фрагментов белков. Они позволяют напрямую определить изменения в гибкости структуры при совершении аминокислотных замен. Их недостатком является сложность оценки влияния окружающего растворителя и его термодинамических параметров. Однако это не препятствует применению подобных методов для оценки термостабильности крупных систем, которые

сложно поддаются моделированию, например, методами молекулярной динамики.

БЛАГОДАРНОСТИ

Данная работа выполнена в рамках Госконтракта 07.514.11.4011 с Министерством образования и науки РФ.

Авторы выражают признательность С.А. Лашину за критическое рассмотрение рукописи данной статьи и за ценные советы по ее улучшению.

ЛИТЕРАТУРА

- Афонников Д.А., Медведев К.Е., Гунбин К.В., Колчанов Н.А. Важная роль гидрофобных взаимодействий при адаптации белков к высоким давлениям // Докл. АН. 2011. Т. 438. № 3. С. 412–415.
- Фомин Э.С., Алемасов Н.А. Программный комплекс l-mol Kern для расчетов разностей свободных энергий с учетом эффектов перераспределения заряда // Математ. биол. и биоинформатика. 2012. Т. 7. № 2. С. 398–409.
- Afonnikov D.A., Oshchepkov D.Y., Kolchanov N.A. Detection of conserved physico-chemical characteristics of proteins by analyzing clusters of positions with coadaptive substitutions // Bioinformatics. 2001. V. 17. No. 11. P. 1035–1046.
- Alder B., Wainwright T. Phase transition for a hard sphere system // J. Chemical Physics. 1957. V. 27. P. 1208.
- Beveridge D., DiCapua F. Free energy via molecular simulation: applications to chemical and biomolecular systems // Annu. Rev. Biophys. Biophys. Chem. 1989. V. 18. No. 1. P. 431–492.
- Connolly M. The molecular surface package // J. Mol. Graphics. 1993. V. 11. No. 2. P. 139–141.
- Daniel R.M., Cowan D.A., Morgan H.W., Curran M.P. A correlation between protein thermostability and resistance to proteolysis // Biochem. J. 1982. V. 207. No. 3. P. 641.
- Dehouck Y., Grosfils A., Folch B. *et al.* Fast and accurate predictions of protein stability changes upon mutations using statistical potentials and neural networks: Popmusic-2.0 // Bioinformatics. 2009. V. 25. No. 19. P. 2537–2543.
- Dong H., Mukaiyama A., Tadokoro T. *et al.* Hydrophobic effect on the stability and folding of a hyperthermophilic protein // J. Mol. Biol. 2008. V. 378. No. 1. P. 264–272.
- Eisenberg D., McLachlan A. Solvation energy in protein folding and binding // Nature. 1986. V. 319. No. 6050. P. 199–203.
- Gromiha M. Factors influencing the thermal stability of buried protein mutants // Polymer. 2003. V. 44. No. 14. P. 4061–4066.
- Guerois R., Nielsen J., Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations // J. Mol. Biol. 2002. V. 320. No. 2. P. 369–387.
- Hess B., Kutzner C., van der Spoel D., Lindahl E. Gromacs 4:

- Algorithms for highly efficient, load-balanced, and scalable molecular simulation // *J. Chem. Theory and Comput.* 2008. V. 4. No. 3. P. 435–447.
- Hornak V., Abel R., Okur A. *et al.* Comparison of multiple amber force fields and development of improved protein backbone parameters // *Proteins: Structure, Function, and Bioinformatics*. 2006. V. 65. No. 3. P. 712–725.
- Jacobs D.J., Rader A.J., Kuhn L.A., Thorpe M.F. Protein flexibility predictions using graph theory // *Proteins: Structure, Function, and Bioinformatics*. 2001. V. 44. No. 2. P. 150–165.
- Kabsch W., Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features // *Biopolymers*. 1983. V. 22. No. 12. P. 2577–2637.
- Khechinashvili N.N., Fedorov M.V., Kabanov A.V. *et al.* Side chain dynamics and alternative hydrogen bonding in the mechanism of protein thermostabilization // *J. Biomol. Struct. Dyn.* 2006. V. 24. No. 3. P. 255–262.
- Kong X., Brooks III C.L. λ -dynamics: A new approach to free energy calculations // *J. Chem. Phys.* 1996. V. 105. No. 6. P. 2414–2423.
- Kumar S., Meenatchi M. Virtual quantification of protein stability using applied kinetic and thermodynamic parameters // *IIOAB Lett.* 2011. V. 1. No. 1. P. 21–28.
- Kumar S., Tsai C., Nussinov R. Factors enhancing protein thermostability // *Protein Engineering*. 2000. V. 13. No. 3. P. 179–191.
- Matthews B. Structural and genetic analysis of protein stability // *Annu. Rev. Biochem.* 1993. V. 62. No. 1. P. 139–160.
- Matthews B., Nicholson H., Becktel W. Enhanced protein thermostability from sitedirected mutations that decrease the entropy of unfolding // *Proc. Natl Acad. Sci. USA*. 1987. V. 84. No. 19. P. 6663–6667.
- Mozo-Villiaris A., Querol E. Theoretical analysis and computational predictions of protein thermostability // *Curr. Bioinformatics*. 2006. V. 1. No. 1. P. 25–32.
- Nicholls A., Sharp K., Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons // *Proteins: Structure, Function, and Bioinformatics*. 1991. V. 11. No. 4. P. 281–296.
- Perutz M., Raidt H. Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2 // *Nature*. 1975. V. 255. P. 256–259.
- Potapov V., Cohen M., Schreiber G. Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details // *Protein Engineering Design and Selection*. 2009. V. 22. No. 9. P. 553–560.
- Rader A.J., Yennamalli R.M., Harter A.K., Sen T.Z. A rigid network of long-range contacts increases thermostability in a mutant endoglucanase // *J. Biomol. Struct. Dynam.* 2012. P. 1–10.
- Radestock S., Gohlke H. Protein rigidity and thermophilic adaptation // *Proteins: Structure, Function, and Bioinformatics*. 2011. V. 79. No. 4. P. 1089–1108.
- Saiki R.K., Scharf S., Faloona F. *et al.* Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia // *Science*. 1985. V. 230. No. 4732. P. 1350–1354.
- Seeliger D., De Groot B. Protein thermostability calculations using alchemical free energy simulations // *Biophys. J.* 2010. V. 98. No. 10. P. 2309–2316.
- Sharp K., Honig B. Electrostatic interactions in macromolecules: theory and applications // *Annu. Rev. Biophys. Biophys. Chem.* 1990. V. 19. No. 1. P. 301–332.
- Szilagy A., Zavodszky P. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey // *Structure*. 2000. V. 8. No. 5. P. 493–504.
- Talluri S. Advances in engineering of proteins for thermal stability // *Intern. J. Adv. Biotechnol. Res.* 2011. V. 2. No. 1. P. 190–200.
- Tanford C. Contribution of hydrophobic interactions to the stability of the globular conformation of proteins // *J. Amer. Chem. Soc.* 1962. V. 84. No. 22. P. 4240–4247.
- Tidor B., Karplus M. Simulation analysis of the stability mutant r96h of t4 lysozyme // *Biochemistry*. 1991. V. 30. No. 13. P. 3217–3228.
- van der Spoel D., Lindahl E., Hess B. *et al.* Gromacs: fast, flexible, and free // *J. Computat. Chem.* 2005. V. 26. No. 16. P. 1701–1718.
- Vogt G., Woell S., Argos P. Protein thermal stability, hydrogen bonds, and ion pairs I // *J. Mol. Biol.* 1997. V. 269. No. 4. P. 631–643.
- Vorobjev Y.N. Advances in implicit models of water solvent to compute conformational free energy and molecular dynamics of proteins at constant pH // *Computat. Chem. Methods Struct. Biol.* 2011. V. 85. P. 281–322.
- Vriend G. What if: a molecular modeling and drug design program // *J. Mol. Graphics*. 1990. V. 8. No. 1. P. 52.
- Xiao L., Honig B. Electrostatic contributions to the stability of hyperthermophilic proteins I // *J. Mol. Biol.* 1999. V. 289. No. 5. P. 1435–1444.
- Zhu S., Elcock A. A complete thermodynamic characterization of electrostatic and hydrophobic associations in the temperature range 0 to 100 °C from explicit solvent molecular dynamics simulations // *J. Chem. Theory Comput.* 2010. V. 6. No. 4. P. 1293–1306.

THEORETICAL METHODS FOR INVESTIGATING PROTEIN THERMOSTABILITY AND THEIR APPLICATIONS IN BIOLOGY

N.A. Alemasov, E.S. Fomin

Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia,
e-mail: alemasov@bionet.nsc.ru

Summary

Classification of existent theoretical approaches for investigating protein thermal stability is performed. Computer simulations allow to fully estimate micro- and macro properties of the molecules. But those approaches are limited in accuracy and therefore require certain improvements e.g. making them to allow for molecular charge distribution. Also promising methods are those dealing with rigid regions of proteins. They do not require a huge amount of computations and allow to directly determine changes in molecular structure flexibility caused by mutated amino acid residues. But using the only structure it is impossible to explicitly estimate an effect of solvent and its thermodynamical properties.

Key words: protein thermal stability, free energy, molecular dynamics, rigid regions, semi-empiric energy function, electrostatic and hydrophobic interactions, thermophiles, mesophiles.