

УДК 577.322.2:004.94

ИССЛЕДОВАНИЕ ТЕРМОСТАБИЛЬНОСТИ МУТАНТНЫХ ФОРМ БЕЛКА БАРНАЗЫ С ИСПОЛЬЗОВАНИЕМ ПРОГРАММНОГО КОМПЛЕКСА MOLKERN

© 2012 г. Э.С. Фомин, Н.А. Алемасов

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

Поступила в редакцию 14 мая 2012 г. Принята к публикации 30 мая 2012 г.

Выполнено исследование термостабильности ряда мутантных форм белка барназы методом λ -динамики. Метод реализован в рамках программного комплекса MOLKERN. Для исследования выбраны мутации, включающие аминокислоты с ненулевым зарядом, точность расчетов которых в методе λ -динамики почти в 1/4 случаев существенно отличается от экспериментальных данных (> 10 кДж/моль). Предложены оригинальные модификации метода λ -динамики, связанные с построением λ -потенциалов и учетом эффектов изменения заряда. Результаты выполненных расчетов (для мутации R72G) показали лучшее согласие с экспериментальными значениями, чем результаты других авторов.

Ключевые слова: барназа, молекулярная динамика, λ -динамика, разность свободных энергий, термостабильность белков.

Введение

Среди задач, возникающих в приложениях биотехнологии и связанных с рациональным проектированием новых белков (обеспечение большей специфичности к заданному субстрату, изменение специфичности, растворимости или проницаемости и пр.), получение термостабильных (т. е. способных работать с высокой активностью при температурах до 100 °С и выше) белков-ферментов имеет огромное практическое значение. Данная проблема решается внесением мутаций в исходный белок-фермент дикого типа, затрудняющих процесс денатурации.

Для этого в экспериментальных исследованиях используются метод случайного мутагенеза и сравнительный анализ ферментов с более термостабильными аналогами (Lehmann, 2001). Метод случайного мутагенеза ограничен трудоемкостью получения большого числа мутантов, а методы, основанные на сравнении, не могут выйти за пределы тех наборов аминокислотных последовательностей, которые присущи тер-

мофильным организмам. По причине наличия подобных ограничений в экспериментальных подходах большой интерес вызывают альтернативные теоретические методы, основанные на изучении пространственной структуры белков и нахождении точечных аминокислотных замен, которые оптимизируют внутрибелковые взаимодействия в направлении повышения термостабильности, что включает:

1. Исследования структуры полипептидной цепи и повышение ее жесткости путем замены входящих в нее аминокислот на пролин, обладающий низкой конформационной энтропией (Серов и др., 2002), что приводит к понижению конформационной энтропии процесса денатурации.

2. Исследования структуры полипептидной цепи и удаление напряжений в ней, обусловленных сильным отталкиванием между боковыми цепями аминокислотных остатков из-за их неоптимального взаимного расположения (например, путем замены остатков на глицин (Серов и др., 2005), что делает цепь более устойчивой к тепловым воздействиям.

3. Анализ ионных взаимодействий и «конструирование» белков с повышенным содержанием заряженных остатков на их поверхности, что стабилизирует белок за счет снижения диэлектрической проницаемости воды при увеличении температуры (Полянский и др., 2004).

4. «Реконфигурацию» структуры белков либо для увеличения ароматических стэкинг-взаимодействий (Puchkaev *et al.*, 2003), либо для возможности введения дополнительных дисульфидных связей (Pikkemaat *et al.*, 2002).

5. Поиск термолabileльных областей в белках методами молекулярной динамики (Полянский и др., 2004; Meharena *et al.*, 2010) и локализацию мутаций в данных областях белка.

Рассмотренные выше методы позволяют дать достаточно надежное качественное предсказание того, как может повлиять точечная мутация на термостабильность белка, поскольку учитывают реальные физические внутрибелковые взаимодействия. Однако подобные методы не могут дать точное численное значение оцениваемой величины, особенно в ситуации разнонаправленного действия физических эффектов. Задача точного численного расчета эффекта влияния точечной мутации белка на его термостабильность является в настоящее время «вызовом» для биоинформатики.

Задача расчета термостабильности сводится к вычислению разности свободных энергий между различными состояниями моделируемой системы: 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 нативной и денатурированной формой белка дикого типа, а ΔG_2 – разность свободных энергий между нативной и денатурированной формой мутанта (см. рис. 1). Расчет разностей ΔG_3 и ΔG_2 формально возможен, но в реальности он сталкивается с непреодолимыми вычислительными проблемами, обусловленными большим количеством конформаций, которые может принимать белок в процессе сворачивания в нативную структуру. Решение задачи по такому пути практически не отличается от решения задачи фолдинга белка, и вряд ли может быть получено в ближайшие годы.

Подход, позволяющий обойти данную проблему, основывается на том факте, что свобод-

ная энергия является функцией состояния и не зависит от пути, который проходит система. Решение получается через построение термодинамического цикла, который связывает четыре состояния белка: белок дикого типа, денатурированный белок дикого типа, белок-мутант и денатурированный белок-мутант. Сумма разностей свободной энергии Гиббса между этими состояниями удовлетворяет соотношению $\Delta G_1 + \Delta G_2 - \Delta G_4 - \Delta G_3 = 0$, где ΔG_1 – разность свободных энергий между денатурированным белком дикого типа и его мутантной формой, а ΔG_4 – разность свободных энергий между белком дикого типа и его мутантной формой. Из условия равенства нулю изменения свободной энергии вдоль всего цикла, следует, что искомая величина $\Delta\Delta G$ может быть рассчитана также через разность $\Delta\Delta G = \Delta G_1 - \Delta G_4$, где величины ΔG_1 и ΔG_4 могут быть получены без моделирования процессов фолдинга белка, однако требуют моделирования нефизического (т. е. не реализуемого в природе) непрерывного процесса «превращения» одних аминокислотных остатков в другие. Моделирование процесса подобных «превращений», называемых «алхимической» динамикой или λ -динамикой, получило популярность в последние годы (Knight *et al.*, 2009; Seeliger *et al.*, 2010). Решение подобной задачи требует вычислительных ресурсов на уровне не менее чем $10^4 \times 10^3 \times 10^9 = 10^{16}$ операций (произведение числа атомов, числа взаимодействий на атом и количества точек траектории) на одну точку λ -траектории, что могут обеспечить суперкомпьютеры петафлопного и выше уровня производительности.

Данная работа посвящена исследованиям термостабильности некоторых мутаций белка барназы методом λ -динамики, т. е. методом, ко-

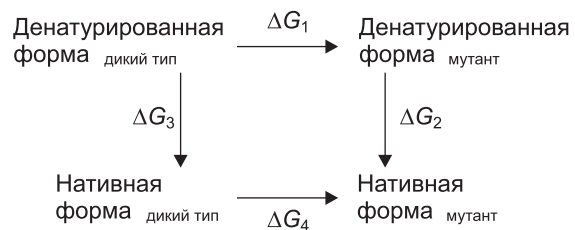


Рис. 1. Термодинамический цикл, который описывает процесс расчета термостабильности белка мутантной формы относительно его дикого типа.

торый в принципе позволяет достичь точности вычисления, сравнимого с точностью экспериментальных данных. Предлагаемый подход включает в себя выполнение молекулярной динамики белков в водном окружении в полно-атомном приближении и накопление обширной статистики для расчета средних по ансамблю физических величин. Для исследования выбраны мутации, включающие аминокислоты с ненулевым зарядом, точность расчетов которых в методе λ -динамики существенно снижена относительно мутаций с нейтральными аминокислотами (Seeliger *et al.*, 2010). В исследовании используются оригинальные модификации метода λ -динамики, связанные с построением λ -потенциалов и учетом эффектов изменения заряда. Представлена реализация метода в рамках программного комплекса MOLKERN, приведены и обсуждены результаты расчетов, сделано сравнение с экспериментальными данными и расчетами других авторов.

Объект исследования и методы

В качестве объекта исследования в рамках метода λ -динамики был выбран белок барназы. Барназа (бактериальная рибонуклеаза) – небольшой белок (110 аминокислот), который синтезируется в бактерии *Bacillus amyloliquefaciens*. Вместе с белком, который называется барстар (89 аминокислотных остатков), эти две биомолекулы образуют комплекс; барстар является внутриклеточным ингибитором активности барназы. Белок плавится при температуре 50 °C при pH от 5 до 9 (Hartley *et al.*, 1989). Имеется пространственная структура барназы, наилучшее разрешение которой составляет 1,5 Å (Martin *et al.*, 1999). Кроме того, известно свыше 400 мутантных форм белка, образованных как точечными, так и множественными заменами аминокислот (Kumar *et al.*, 2006).

Наиболее часто барназа присутствует в качестве объекта в исследованиях фолдинга (Bucroft *et al.*, 1990; Serrano *et al.*, 1992; Fersht *et al.*, 2002), поскольку обладает обратимым сворачиванием. Множество работ (Clarke *et al.*, 1993; Johnson *et al.*, 1995; Pedersen *et al.*, 2002) посвящено исследованию факторов, влияющих на стабильность барназы, другие предусматривают комплексные исследования стабильности

мутантных форм этого белка. Причем среди последних встречаются как экспериментальные (Kellis *et al.*, 1989; Matouschek *et al.*, 1989; Horovitz *et al.*, 1992; Buckle *et al.*, 1993), так и теоретические (большая часть – компьютерные) исследования (Prevost *et al.*, 1991; Gilisa *et al.*, 1996).

Белок барназы является хорошо изученным объектом, в том числе и методом λ -динамики. Для большинства мутаций этого белка (более 70 %), изученных данным методом, расхождение рассчитанных данных с экспериментальными находится в пределах ± 4 кДж/моль, т. е. в пределах стандартной ошибки силового поля. Для мутаций, включающих заряженные аминокислотные остатки, точность рассчитанных данных существенно хуже: только 52 % находятся в пределах ± 4 кДж/моль от экспериментального значения, и для 24 % мутаций отклонение превышает 10 кДж/моль (Seeliger *et al.*, 2010).

Для исследования подобных мутаций в настоящей работе был использован метод λ -динамики, дополненный оригинальным методом учета эффектов изменения зарядового состояния аминокислотных остатков (Net-Q метод, см. Приложение 4). Метод реализован в рамках программного комплекса MOLKERN (Фомин *et al.*, 2006), предназначенного для эффективного выполнения основных задач молекулярного моделирования. MOLKERN использует силовые поля AMBER и GAFF. Он написан на языке C++ с использованием библиотек STL, BOOST, FFTW и др. Базовый код MOLKERN векторизован (через SSE и SSE2 расширения), использует многопоточность и выполняется в операционной системе Linux.

Выбор программного пакета MOLKERN для решения задачи исследования термостабильности оптимален также с точки зрения производительности. Например, на процессоре Intel(R) Core(TM)2 Quad CPU Q8400 2.66 GHz MOLKERN позволяет рассчитывать системы до 3 млн атомов, обеспечивая производительность до 0,1 млрд взаимодействий/с в 4-поточном исполнении. Эффективность реализации обусловлена как алгоритмической оптимизацией – все алгоритмы программы имеют вычислительную сложность, не превышающую $O(N \log N)$, так и использованием ряда оригинальных подходов

для обеспечения минимизации потока данных между глобальной памятью и CPU; а также автоматической подстройкой параметров в направлении предельной эффективности расчета при различных физических параметрах моделирования (Fomin, 2011).

При решении задачи термостабильности мутантов белка на вход MOLKERN подается файл формата PDB, содержащий координаты атомов белка дикого типа. В отличие от большинства существующих программных пакетов молекулярного моделирования MOLKERN позволяет выполнять точечные аминокислотные замены в белке, заданные в командной строке, автоматически; также автоматически строятся топология для λ -состояния комплекса и λ -потенциалы (см. Приложение 2). Далее происходят оптимизация структуры λ -состояния, заполнение заданной области моделирования молекулами воды, выполнение молекулярной динамики и накопление необходимой статистики. Дополнительный вызов MOLKERN с другими параметрами командной строки позволяет обработать полученные файлы статистики и получить термодинамические средние для физических величин, необходимых при расчете разностей свободных энергий ΔG методом термодинамического интегрирования (см. Приложение 1).

Модели белка барназы и методика расчетов

Модели белка. В работе использовались следующие модели белка барназы:

1. GXG – 3-аминокислотный пептид, который является моделью денатурированной формы белка (Seeliger *et al.*, 2010) и состоит из варьруемой между расчетами аминокислоты X, которая окружена остатками Gly.

2. BarnaseX – полноатомный белок барназы, являющийся моделью его нативной формы, в которой выполняются одиночные замены произвольной аминокислоты X.

Модель GXG разработана специально для расчета величины ΔG_1 , которая представляет собой разность свободных энергий между белком дикого типа и его мутантной формы в денатурированном состоянии (см. рис. 1). Данная модель позволяет обойти проблему, связанную с получением пространственной структуры

денатурированной формы белка и с расчетом ее конформационной энтропии. Применимость модели GXG для расчета величины ΔG_1 обусловлена тем, что в денатурированной форме все аминокислотные остатки белка в равной степени доступны растворителю, и в области любого остатка находится в среднем не более двух непосредственно связанных с ним соседей. При внесении мутации в белок пространственное окружение всех остатков, не связанных с мутацией, не меняется, и по этой причине они не дают вклада в величину ΔG_1 и могут быть выброшены из расчета. Применимость моделей, подобных GXG, создаваемых путем удаления всех незначимых для исследуемого явления атомов, обусловлена локальностью любых физических процессов, и использование подобных моделей является распространенным приемом (Messmer, 1979).

BarnaseX – полноатомный белок барназы, являющийся моделью его нативной формы, в которой производятся точечные замены аминокислоты X. Данная модель позволяет получить величину ΔG_4 (см. рис. 1), которая представляет разность свободных энергий между нативным белком дикого типа и его мутантной формой.

Методика расчетов. Все расчеты выполнялись со следующими параметрами:

1. Размер пространственной области моделирования – $30 \times 30 \times 30 \text{ \AA}^3$ (GXG модель) и $60 \times 60 \times 60 \text{ \AA}^3$ (BarnaseX модель); число молекул воды – 282 и 5577 соответственно.

2. Радиус обрезания потенциала – 10 \AA .

3. Число шагов оптимизации геометрии после построения системы – 100.

4. Используемый тип термодинамического ансамбля – NVT с температурой моделирования – 400 $^{\circ}\text{K}$.

5. Длина расчетных траекторий молекулярной динамики – 20 нс с шагом интегрирования – 2 фс.

6. Шаг интегрирования расширенной λ -траектории – от 1/8 до 1/10.

7. Период сохранения статистических величин – 1 пс.

8. Игнорируемый при расчетах ΔG_i диапазон траекторий молекулярной динамики – от 0 до 10 нс модельного времени; используемый для расчета ΔG_i диапазон траекторий – от 10 до 20 нс модельного времени.

Выбор параметров, таких, как шаги интегрирования МД траектории и λ -траектории, а также длина расчетных траекторий мотивировался необходимостью увеличения эффективности расчетов. За счет ограничения длины траекторий в расчетах происходит потеря точности, связанная с недооценкой энтропийного вклада (моделируемая система за заданное время не успевает обойти значимый объем конформационного пространства). Компенсация недооценки данного вклада осуществлялась повышением температуры моделирования до 400 °К. Увеличение температуры приводит к увеличению средней кинетической энергии молекул воды, чьи столкновения с атомами белка приводят к более частым поворотам ротамеров аминокислотных остатков, локализованных на поверхности белка. В результате система проходит большую часть пространства конфигураций за меньшее время моделирования, компенсируя негативный эффект ограничения длины расчетной траектории. Уравновешивание температуры выполнялось с помощью метода термостата Берендсена (Berendsen *et al.*, 1984).

Расчеты производились с использованием ресурсов ЦКП «Биоинформатика» СО РАН. Среди них вычислительный кластер, состоящий из: 64 двухпроцессорных серверов (Intel Xeon E5540), 48 двухпроцессорных серверов (Intel Xeon X5670) и параллельной файловой системы IBRIX для хранения данных: 4 шлюзовых сервера (HP DL380 G6), 4 дисковые полки (2 – HP MSA 2312sa DC, 2 – HP MSA 2000 DC) и сервер управления (HP DL360 G6). Расчеты оказались чрезвычайно затратны (до 400 ч на одну точку λ -траектории). Предположительно значительные затраты времени на расчеты обусловлены задержками файловой системы из-за необходимости сохранения большого объема статистических данных.

В силу существенных вычислительных затрат в работе решено было ограничиться только мутациями A32G и R72G. Данные мутации представлены по следующим их свойствам:

1. Мутация A32G затрагивает аминокислоты Ala и Gly, которые не обладают конформационной энтропией (остаток Ala – в силу симметрии группы CH_3 , а остаток Gly – в силу того, что он включает только атом H). Таким образом, в расчете ΔG отсутствует энтропийный вклад,

что в свою очередь позволяет ограничиться короткими траекториями динамики при наборе статистики. Расчет энергии данной мутации позволяет оценить корректность заложенных в MOLKERN базовых алгоритмов вычисления энтальпийного вклада в ΔG .

2. Мутация R72G включает аминокислоту Arg, которая обладает высокой конформационной энтропией, поскольку состоит из длинной цепи атомов $CH_2 - CH_2 - CH_2 - NH - C(NH_2)_2$. Расчет мутации R72G позволяет оценить корректность построения ансамбля конфигураций и расчета энтропийного вклада в ΔG . Дополнительным фактором, обусловившим выбор данной мутации, являлось то, что мутация R72G включает изменение зарядового состояния.

Результаты расчетов

Все расчеты термостабильности мутантных форм белка барназы относительно его дикого типа выполнены согласно формулам метода термодинамического интегрирования, описанного в Приложении 1. Основной величиной, которая необходима для расчетов, является величина $dH/d\lambda$. Совокупность этих величин собирается по всей траектории молекулярной динамики для каждого значения λ , характеризующего степень «превращения» белка дикого типа в мутантный. Среднее значение по ансамблю величины $\langle dH/d\lambda \rangle$ характеризует скорость изменения свободной энергии в точке λ согласно соотношению $\delta F = \langle dH/d\lambda \rangle \delta \lambda$.

На рис. 2 приведен график изменения величины $\langle dH/d\lambda \rangle$ вдоль λ -траектории в диапазоне $\lambda \in [0, 1)$ для перехода между диким типом и мутантом A32G для модели GXG денатурированного белка. Функция, отображенная на графике, построена с использованием ограниченного числа λ точек с шагом $d\lambda = 0,1$, и интерполяция между ними выполнена полиномами второго порядка. Для каждой точки показана теоретическая погрешность расчета, вычисленная согласно формулам Приложения 3.

Величина ΔG_1 вычислена путем интегрирования полученной функции на графике согласно формулам метода TI (см. Приложение 1), т. е.

$\Delta G_1 \approx \Delta F_1 = \int_0^1 \langle dH_1/d\lambda \rangle d\lambda$. Рассчитанное значение $\Delta G_1 = -37,6$ кДж/моль очень хорошо

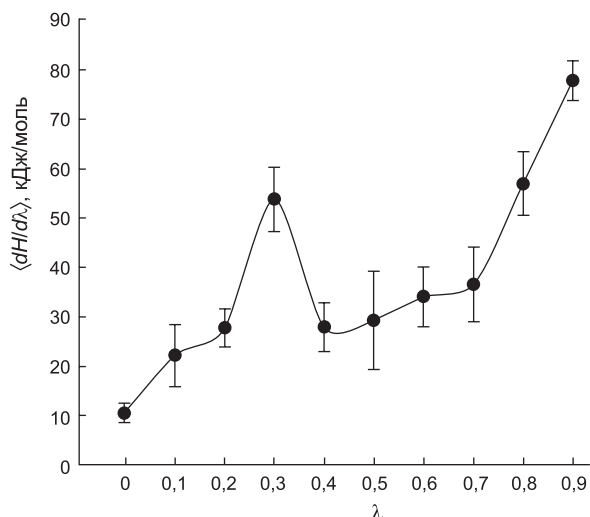


Рис. 2. Изменение величины $\langle dH/d\lambda \rangle$ вдоль λ -траектории для перехода между диким типом и мутантом A32G для модели GXG денатурированного белка.

(в пределах точности ± 4 кДж/моль силового поля AMBER) согласуется со значением $-36,1$ кДж/моль, полученным в работе Seeliger с соавт. (2010). Таким образом, для мутаций, в которых основной вклад в ΔG дает энтальпия, показано хорошее согласие получаемых MOLKERN результатов с расчетами других авторов.

На рис. 3 приведены графики изменения величины $\langle dH/d\lambda \rangle$ вдоль λ -траектории в диапазоне $\lambda \in [0, 1)$ для перехода между диким типом и мутантом R72G для модели GXG денатурированного белка (а) и модели нативного белка (б).

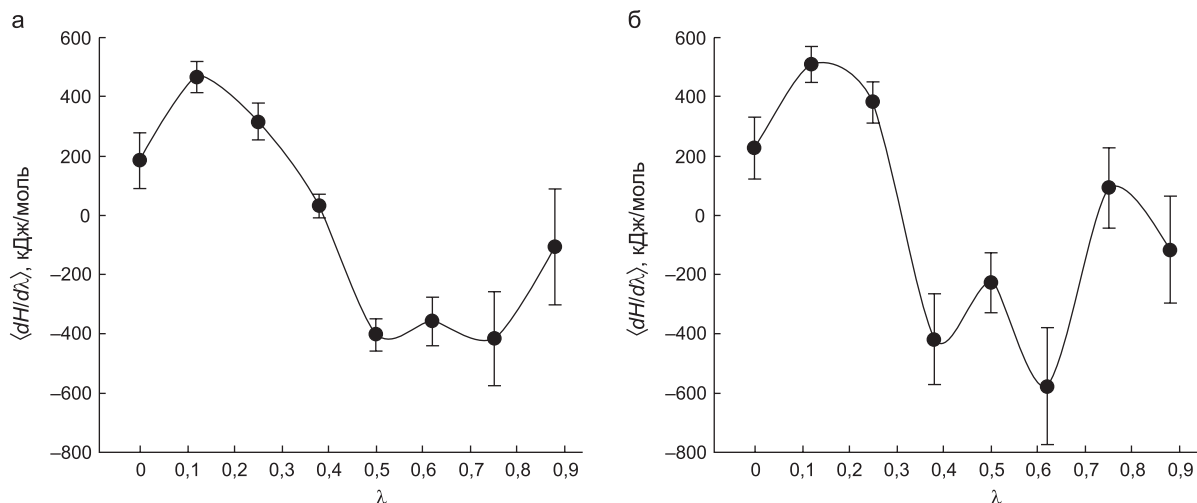


Рис. 3. Изменение величины $\langle dH/d\lambda \rangle$ вдоль λ -траектории для перехода между диким типом и мутантом R72G для модели GXG денатурированного белка (а) и для модели нативного белка BarnaseX (б).

Левый график позволяет рассчитать величину ΔG_1 , а правый график – ΔG_4 . Функции, отображенные на графике, построены с использованием ограниченного числа λ точек с шагом $d\lambda = 0,1$, и интерполяция между этими точками выполнена полиномами второго порядка. Для каждой точки показана теоретическая погрешность расчета, вычисленная согласно формулам Приложения 3.

Величины ΔG_1 и ΔG_4 , вычисленные согласно формулам метода ПИ, равны $\Delta G_1 = -36,6$ кДж/моль и $\Delta G_4 = -16,8$ кДж/моль, что в свою очередь дает значение для $\Delta\Delta G_{calc} = -19,8$ кДж/моль. Полученная величина $\Delta\Delta G_{calc}$ находится существенно ближе к экспериментальному значению $\Delta\Delta G_{exp} = -10,45$ кДж/моль (Kumar *et al.*, 2006), чем рассчитанная $\Delta\Delta G_{calc} = -57,67$ кДж/моль в работе Seeliger с соавт. (2010). Расхождение относительно экспериментального значения у Seeliger с соавт. (2010) составляет $-47,2$ кДж/моль против $-9,4$ кДж/моль в данной работе.

Значение величины $\Delta\Delta G_{exp} = -10,4$ кДж/моль для мутации R72G означает снижение термостабильности R72G мутантной формы белка барназы относительно его дикого типа. Поскольку аминокислота R72 локализована таким образом, что ее NH_2 -группы доступны воде и количество водородных связей в обоих случаях (денатурированный или нативный белок) одинаково для данной аминокислоты и не влияет на термостабильность, то основной эффект снижения термостабильности обусловлен энтропийным

вкладом. Действительно, аминокислотный остаток Arg представляет собой длинную цепочку, которая имеет большое число возможных конформаций в денатурированной форме белка. В нативной форме белка движение данной аминокислоты существенно ограничено. Это демонстрируется на рис. 4, где приведен фрагмент белка около аминокислотного остатка R72, причем остаток R72 на рисунке представлен в виде модели «шариков и проволочек». Существование белков с подобной мутацией, которая невыгодна с точки зрения термостабильности, должно быть обусловлено, с эволюционной точки зрения, и исследования механизмов функционирования белка с данной мутацией могут дать новую биологическую информацию.

Выводы

В работе выполнены исследования термостабильности мутаций A32G и R72G белка барназы методом λ -динамики. Исследование проведено с использованием программного комплекса MOLKERN, в котором реализован метод λ -динамики с оригинальными модификациями, связанными с построением λ -потенциалов и учетом эффектов изменения заряда. Результаты исследования для ограниченного числа мутаций

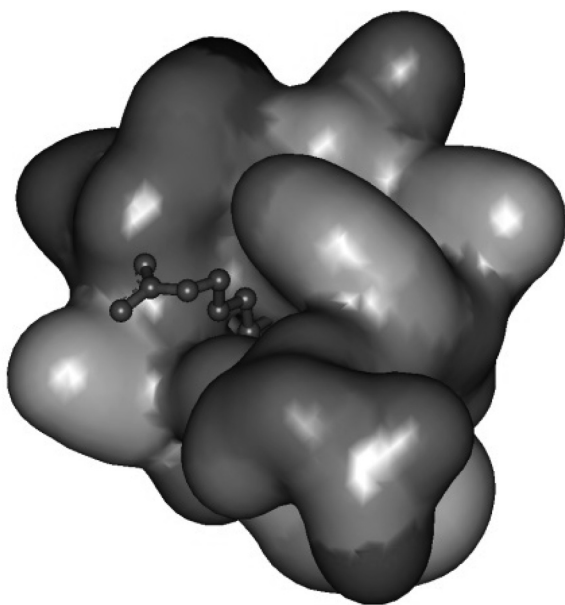


Рис. 4. Фрагмент белка барназы около аминокислотного остатка R72.

показали, что разработанный метод полностью согласуется с расчетами других авторов при расчете энтальпийного вклада в свободную энергию (мутация A32G) и дает более точные результаты при учете энтропийного вклада для мутаций с ненулевым зарядом (R72G), т. е. для случая, в котором отклонение от экспериментальных значений существенно (Seeliger *et al.*, 2010).

Благодарности

Данная работа поддержана грантами СО РАН (интеграционные проекты № 39, № 47, № 130), Министерством образования и науки РФ (Госконтракты № П857, № 07.514.11.4011), а также Программой Президиума РАН «Интеграция РАН 6.8 и НШ-5278.2012.4».

Приложения

1. Метод термодинамического интегрирования

Метод термодинамического интегрирования при расчете разностей свободных энергий используется не только в задачах определения термостабильности, но также в исследованиях связывания лигандов (Bash *et al.*, 1987; Kollman, 1993), при изучении стабильности белков (Dang *et al.*, 1989; Tidor *et al.*, 1991), ионного транспорта (Roux *et al.*, 1991), белкового фолдинга (Boczek *et al.*, 1995) и поверхностей свободной энергии (Bartels *et al.*, 1999; Simonson *et al.*, 2002).

Свободная энергия в каноническом (постоянные: N, V, T) ансамбле определяется термодинамическим потенциалом Гельмгольца: $F = -k_B T \ln Z$, где k_B – константа Больцмана, T – температура, Z – статистическая сумма. Для систем в жидкой фазе свободная энергия Гельмгольца F отличается от свободной энергии Гиббса G на величину pV , и их разности совпадают с хорошей точностью $\Delta G \approx \Delta F$, поскольку для жидкостей изменение $p\Delta V \approx 0$.

Статистическая сумма Z является функцией всех микросостояний $\{p, x\}$ системы, которые удовлетворяют условиям ансамбля, и выражается как:

$$Z = \int \exp \left[-\frac{H(p, x)}{k_B T} \right] dp dx, \quad (1)$$

где $H(p, x)$ – гамильтониан системы. При этом плотность вероятности $\rho(p, x)$ каждого микросостояния $\{p, x\}$ определяется как:

$$\rho(p, x) = \frac{1}{Z} \exp \left[-\frac{H(p, x)}{k_B T} \right], \quad (2)$$

что позволяет рассчитать значение среднего для любой физической величины $A(p, x)$:

$$\begin{aligned} \langle A \rangle &= \int \rho(p, x) A(p, x) dp dx = \\ &= \frac{1}{Z} \int \exp \left[-\frac{H(p, x)}{k_B T} \right] A(p, x) dp dx. \end{aligned} \quad (3)$$

Процессы перехода из одного состояния A в другое достаточно близкое состояние $B \approx A$ описываются значением относительной свободной энергии ΔF :

$\Delta F = F(B) - F(A) = -k_B T (\ln Z(B) - \ln Z(A))$, (4) где гамильтонианы $H(A)$ и $H(B)$ близких состояний A и B , входящие в статсуммы $Z(A)$ и $Z(B)$, непрерывным образом связываются друг с другом путем введения параметра λ , например, по такому соотношению:

$$H(A, B; \lambda) = (1 - \lambda)H(A) + \lambda H(B). \quad (5)$$

Введенный таким образом гамильтониан $H(\lambda) = H(A, B; \lambda)$ при $\lambda = 0$ совпадает с гамильтонианом системы A , а при $\lambda = 1$ совпадает с гамильтонианом системы B ; во всех промежуточных точках он соответствует некоторому физически нереализуемому (виртуальному) состоянию, являющемуся «смесью» состояний A и B .

Изменение свободной энергии $\delta F(\lambda)$ для «смешанной» системы при ее переходе из состояния λ в состояние $\lambda + \delta\lambda$ определяется соотношением:

$$\begin{aligned} \delta F &= \frac{dF}{d\lambda} \delta\lambda = -k_B T \delta\lambda \frac{d}{d\lambda} \ln Z = \\ &= \frac{\delta\lambda}{Z} \int \frac{dH}{d\lambda} \exp \left[-\frac{H}{k_B T} \right] = \left\langle \frac{dH}{d\lambda} \right\rangle \delta\lambda. \end{aligned} \quad (6)$$

Полное приращение свободной энергии ΔF между состояниями A и B выражается через сумму L приращений всех промежуточных переходов (поскольку свободная энергия является функцией состояния и не зависит от пути интегрирования):

$$\Delta F = \sum_{j=1}^{L-1} \Delta F(\lambda_j \rightarrow \lambda_{j+1}) = \int_0^1 \left\langle \frac{dH}{d\lambda} \right\rangle \delta\lambda. \quad (7)$$

Полученная формула является основной для метода термодинамического интегрирования (TI) (Kirkwood, 1935). Существуют различные

модификации рассмотренного подхода, например Free Energy Perturbation (Zwanzig, 1954) (FEP) и Bennett's Acceptance Ratio (Bennett, 1976) (BAR).

2. λ -потенциалы

В программном комплексе MOLKERN с каждой частицей i связано некоторое значение коэффициента λ_i как параметра силового поля, аналогичного таким параметрам, как масса m_i , заряд q_i , радиус ван-дер-ваальсового взаимодействия σ_i и силы этого взаимодействия ε_i , и определены правила комбинирования параметров $\lambda_i, \lambda_j, \dots, \lambda_n$ для взаимодействующих частиц i, j, \dots, n в виде произведения $\lambda_{ij\dots n} = \lambda_i \lambda_j \dots \lambda_n$. Для частиц, которые не являются виртуальными, т. е. которые не появляются и не исчезают в процессе динамики, значения параметра λ всегда равно 1.

Несвязующие парные (кулоновский и Ван-дер-Ваальса) потенциалы определены следующим образом:

$$V_{coul}(r_{ij}) = \frac{q_i \lambda_i q_j \lambda_j}{r_{ij}}; \quad (8)$$

$$V_{vdw}(r_{ij}) = 4\varepsilon_{ij} \lambda_i \lambda_j \left[\left(\frac{\sigma_{ij} \lambda_i \lambda_j}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij} \lambda_i \lambda_j}{r_{ij}} \right)^6 \right], \quad (9)$$

т. е. зависимость от параметров λ_i и λ_j для атомов i и j введена в параметры потенциалов как $q_i(\lambda) = q_i \lambda_i$, $\sigma_{ij}(\lambda) = \lambda_i \lambda_j \sigma_{ij}$ и $\varepsilon_{ij}(\lambda) = \lambda_i \lambda_j \varepsilon_{ij}$, где q_i и q_j – заряды атомов i и j , ε_{ij} и σ_{ij} – стандартные параметры силового поля, определяемые в соответствии с правилами комбинирования в используемом силовом поле, например, для AMBER по правилам $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$ и $\sigma_{ij} = (\sigma_i + \sigma_j)/2$.

Связующие потенциалы валентных связей $V_b(r, \lambda)$, валентных $V_a(\alpha, \lambda)$ и двугранных $V_t(\varphi, \lambda)$ углов и зависимость кинетической энергии $K_i(\lambda)$ атома i от λ определены следующим образом:

$$V_b(r, \lambda) = \lambda_i \lambda_j k_{ij} (r - r_0)^2 / 2; \quad (10)$$

$$V_a(\alpha, \lambda) = \lambda_i \lambda_j \lambda_k k_{ijk} (\alpha_{ijk} - \alpha_0)^2 / 2; \quad (11)$$

$$V_t(\varphi, \lambda) = \lambda_i \lambda_j \lambda_k \lambda_l \sum_m v_m (1 + \cos(\varphi_{nm} - \varphi_0)); \quad (12)$$

$$K_i(\lambda) = \frac{\lambda_i m_i V_i^2}{2}, \quad (13)$$

где $\lambda_i, \lambda_j, \lambda_k, \lambda_l$ – параметры λ для атомов, формирующих соответствующую связь, валентный или двугранный угол.

3. Точность вычислений в λ -динамике

Точность вычислений разностей свободных энергий между состояниями А и В системы, главным образом, зависит от величины отбрасываемого члена второго порядка $d^2H(\lambda)/d\lambda^2$ в разложении гамильтониана $H(\lambda)$ и может быть оценена следующим образом.

Из формального определения свободной энергии $F(\lambda)$ для промежуточного состояния λ расширенного термодинамического ансамбля:

$$F(\lambda) = -kT \ln Z(\lambda), \quad (14)$$

где $Z(\lambda)$ – статистическая сумма заданной системы по всем ее состояниям, k – константа Больцмана, T – абсолютная температура, следует, что разность свободных энергий $\delta F(\lambda) = F(\lambda + \delta\lambda) - F(\lambda)$ между двумя близкими состояниями, задаваемыми гамильтонианами $H(\lambda)$ и $H(\lambda + \delta\lambda)$, равна:

$$\delta F(\lambda) = F(\lambda + \delta\lambda) - F(\lambda) = -kT \ln \frac{Z(\lambda + \delta\lambda)}{Z(\lambda)}. \quad (15)$$

Подставив в данную формулу определение статистической суммы

$$Z(\lambda) = \iint \exp\left[-\frac{H(\lambda)}{kT}\right] dpdx \quad (16)$$

и разложив гамильтониан H в точке $\lambda + \delta\lambda$ до членов второго порядка

$$H(\lambda + \delta\lambda) = H(\lambda) + \frac{dH}{d\lambda} \delta\lambda + \frac{1}{2} \frac{d^2H}{d\lambda^2} \delta\lambda^2 + o(\delta\lambda^3), \quad (17)$$

получим:

$$\begin{aligned} \delta F(\lambda) &= -kT \ln \left[\frac{\iint \exp\left[-\frac{H(\lambda + \delta\lambda)}{kT}\right] dpdx}{\iint \exp\left[-\frac{H(\lambda)}{kT}\right] dpdx} \right] \approx \\ &\approx -kT \ln \left[1 - \frac{\delta\lambda}{kT} \left\langle \frac{dH}{d\lambda} \right\rangle + \frac{1}{2} \left(\frac{\delta\lambda}{kT} \right)^2 \left[\left\langle \left(\frac{dH}{d\lambda} \right)^2 \right\rangle - kT \left\langle \frac{d^2H}{d\lambda^2} \right\rangle \right] \right] \approx \\ &\approx \delta\lambda \left\langle \frac{dH}{d\lambda} \right\rangle + \frac{\delta\lambda^2}{2} \left[\left\langle \frac{d^2H}{d\lambda^2} \right\rangle - \frac{1}{kT} \left[\left\langle \left(\frac{dH}{d\lambda} \right)^2 \right\rangle - \left\langle \frac{dH}{d\lambda} \right\rangle^2 \right] \right], \quad (18) \end{aligned}$$

где величина в угловых скобках $\langle \dots \rangle$ означает усреднение по ансамблю, а последний переход от $\ln(1 - \delta) \approx -\delta$ обусловлен малостью величины δ . Точность расчетов δF определяется членом при $\delta\lambda^2$ в полученном выражении.

4. Net-Q метод учета изменения заряда

Эффективные методы для динамической оценки зарядового перераспределения в зависимости от геометрии локального окружения базируются на принципе выравнивания электроотрицательности атомов при образовании молекулы (Sanderson, 1951), при этом атомные заряды трактуются как количество заряда, переместившегося от менее электроотрицательных атомов к более электроотрицательным. Данный принцип используется в методах БЕМ (electronegativity equalization method) (Mortier *et al.*, 1986), fluc-q (dynamical fluctuating charge model) (Rick *et al.*, 1994), CPE (chemical poten-

tial equalization) (York *et al.*, 1996), QEq (charge equilibration) (Rappe *et al.*, 1991) и ряде других (Zhu *et al.*, 1991; Bayly *et al.*, 1993; Chen *et al.*, 2000; Oliferenko *et al.*, 2001).

В программном комплексе MOLKERN реализован оригинальный метод расчетов зарядового распределения, называемый Net-Q методом, основанный на минимизации энергии электростатического поля при ограничениях, связанных с уравниванием электроотрицательностей атомов и сохранением полного заряда молекул. Движение заряда в методе разрешено только вдоль ребер графа, представляющего молекулярную систему, тем самым автоматически вводятся ограничения на дви-

жения заряда согласно требованиям модели QTRIE (Morales *et al.*, 2001) (коррекция асимптотического поведения распределения зарядов при удлинении расстояния между атомами) и обеспечивается линейная вычислительная сложность $O(N)$ вычислений. Для обеспечения динамики заряда в методе Net-Q используется подход расширенного ансамбля, аналогичный подходу flux-q.

Модельный гамильтониан зарядовой подсистемы Net-Q метода определен в виде:

$$H(\mathbf{p}, \mathbf{q}) = \frac{1}{2} \sum_k M_k f_k(r_{ij}) I_k^2 + \sum_i \left[\tilde{\chi}_i q_i + \frac{1}{2} \eta_i q_i^2 \right] + \sum_{ij} J_{ij}(r_{ij}) q_i q_j, \quad (19)$$

где I_k – ток, протекающий по связи k между атомами i и j ; $M_k(r_{ij})$ – фиктивная «масса» тока I_k ; $f_k(r_{ij})$ – функция, ограничивающая перенос заряда, $\tilde{\chi}_i = \sum_j f_k(r_{ij})(\chi_i^0 - \chi_j^0)$ – модифицированная согласно Chen с соавт. (2007) электроотрицательность атома.

Модель определена на графе $G = (V, E)$, представляющем молекулярную систему, где V – множество атомов системы, E – множество пар атомов, между которыми разрешен перенос заряда (например, все пары атомов, соединенных химической связью), и ограничена следующими требованиями:

1. Разрешены токи I_k только по ребрам графа G , что позволяет существенно сократить число переменных I_k . Число элементарных токов N_I связано с числом зарядов N_q соотношением: $N_I = N_q - 1 + N_{cycle}$, где N_{cycle} – число простых циклов (например, ароматических колец) в системе.

2. Используется правило Кирхгофа $\sum_{cycle} I_i = 0$ для замкнутых циклов тока в графе G , что обусловлено необходимостью учета в методе ароматических колец молекул.

3. Обеспечено сохранение полного заряда в любом замкнутом подграфе (например, для любой молекулы, входящей в систему), т. е. $\sum_i q_i = 0$.

Гамильтониан зарядовой подсистемы Net-Q метода в канонической форме выглядит следующим образом:

$$H(\mathbf{P}, \mathbf{Q}) = \frac{1}{2} \mathbf{P} \mathbf{A} \tilde{\mathbf{M}}^{-1} \mathbf{A}^T \mathbf{P} + (\tilde{\boldsymbol{\chi}}, \mathbf{Q}) + \frac{1}{2} \mathbf{Q}^T \mathbf{J} \mathbf{Q}, \quad (20)$$

где $\mathbf{Q} = \{q_1, q_2, \dots, q_n\}$ и $\mathbf{P} = \{p_1, p_2, \dots, p_n\}$ – вектора зарядов и импульсов их движения, $\tilde{\mathbf{M}} = \mathbf{f}(\mathbf{r}_{ij}) \cdot \text{diag}\{m_1, m_2, \dots, m_n\}$ – диагональная матрица «масс» токов, \mathbf{A} – топологическая матрица графа молекулы, \mathbf{J} – матрица взаимодействий зарядов. Из гамильтониана следуют уравнения движения зарядовой подсистемы:

$$\dot{\mathbf{Q}} = \frac{\delta H}{\delta \mathbf{P}} = \mathbf{A} \tilde{\mathbf{M}}^{-1} \mathbf{A}^T \mathbf{P}, \quad (21)$$

$$\dot{\mathbf{P}} = \frac{\delta H}{\delta \mathbf{Q}} = -\{\tilde{\boldsymbol{\chi}} + \mathbf{J} \mathbf{Q}\}. \quad (22)$$

Поскольку полученные уравнения движения включают топологическую матрицу молекулярной системы \mathbf{A} , которая не требует обращения и является редкой, а матрица взаимодействий ограничена радиусом обрезания r_{cutoff} , то вычислительная сложность метода является линейной. Использование матрицы \mathbf{A} означает наличие принудительных запретов на перенос заряда между теми атомами, для которых соответствующий элемент матрицы \mathbf{A} равен нулю. Таким образом, матрица \mathbf{A} дает возможность компромисса между эффективностью расчетов и детальностью описания процессов переноса заряда. Ненулевые элементы матрицы \mathbf{A} разрешают перенос заряда по соответствующим связям, и ограничение переноса в этом случае, т. е. если длина связи превышает среднюю длину химической связи, выполняется функцией ограничения $f_k(r_{ij})$.

Литература

- Серов А.Е., Одинцева Е.Р., Упоров И.В., Тишков В.И. Использование карты Рамачандрана для повышения термостабильности бактериальной формиадегидрогеназы // Биохимия. 2005. Т. 70. Вып. 7. С. 974–979.
- Серов А.Е., Тишков В.И. Роль остатков пролина в стабильности прокариотических и эукариотических формиадегидрогенах // Вестник Моск. ун-та. Сер. 2. Химия. 2002. Т. 43. № 6. С. 345–349.
- Полянский А.А., Косинский Ю.А., Ефремов Р.Г. Локальные температурные изменения подвижности в молекулах тиоредоксинов влияют на их термостабильные свойства // Биооргани. химия. 2004. Т. 30. № 5. С. 470–480.
- Фомин Э.С., Алемасов Н.А., Чирцов А.С., Фомин А.Э. Библиотека программных компонент MOKKERN для построения программ молекулярного моделирования // Биофизика. 2006. Т. 51. Вып. 7. С. 110–113.
- Bartels C., Schaefer M., Karplus M. Determination of equilibrium properties of biomolecular systems using multidimensional adaptive umbrella sampling // J. Chem. Phys. 1999. V. 111. No. 17. P. 8048–8068.

- Bash P.A., Singh U.C., Brown F.K. Calculation of the relative change in binding free energy of a protein inhibitor complex // *Science*. 1987. V. 235. No. 4788. P. 574–576.
- Bayly C.I., Cieplak P., Cornell W.D., Kollman P.A. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: The RESP model // *J. Phys. Chem.* 1993. V. 97. P. 10269–10280.
- Bennett C.H. Efficient estimation of free energy differences from Monte Carlo data // *J. Comput. Phys.* 1976. V. 22. No. 2. P. 245–268.
- Berendsen H.J.C., Postma J.P.M., DiNola A., Haak J.R. Molecular dynamics with coupling to an external bath // *J. Chem. Phys.* 1984. V. 81. P. 3684–3690.
- Beutler T.C., Mark A.E., van Schaik R.C. *et al.* Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations // *Chem. Phys. Lett.* 1994. V. 222. No. 6. P. 529–539.
- Boczko E.M., Brooks C.L. First-principles calculation of the folding free energy of a three-helix bundle protein // *Science*. 1995. V. 269. No. 5222. P. 393–396.
- Buckle A.M., Henrick K., Fersht A.R. Crystal structural analysis of mutations in the hydrophobic cores of barnase // *J. Mol. Biol.* 1993. V. 234. No. 3. P. 847–860.
- Bycroft M., Matouschek A., Kellis J.T. Jr. *et al.* Detection and characterization of a folding intermediate in barnase by NMR // *Nature*. 1990. V. 346. No. 6283. P. 488–490.
- Chen B., Potoff J.J., Siepmann J.I. Adiabatic nuclear and electronic sampling Monte Carlo simulations in the gibbs ensemble: Application to polarizable force fields for water // *J. Phys. Chem. B*. 2000. V. 104. P. 2378–2390.
- Chen J., Martinez T.J. QTPIE: Charge transfer with polarization current equalization. A fluctuation charge model with correct asymptotics // *Chem. Phys. Lett.* 2007. V. 463. P. 315–320.
- Clarke J., Fersht A.R. Engineered disulfide bonds as probes of the folding pathway of barnase: Increasing the stability of proteins against the rate of denaturation // *Biochemistry*. 1993. V. 32. No. 16. P. 4322–4329.
- Dang L.X., Merz K.M. Jr., Kollman P.A. Free energy calculations on protein stability: Thr-157 → Val-157 mutation of T4 lysozyme // *J. Amer. Chem. Soc.* 1989. V. 111. No. 22. P. 8505–8508.
- Fersht A.R., Daggett V. Protein folding and unfolding at atomic resolution // *Cell*. 2002. V. 108. No. 4. P. 573–582.
- Fomin E.S. Consideration of data load time on modern processors for the verlet table and linked cell algorithms // *J. Comput. Chem.* 2011. V. 32. No. 7. P. 1386–1399.
- Gilisa D., Rooman M. Stability changes upon mutation of solvent-accessible residues in proteins evaluated by database-derived potentials // *J. Mol. Biol.* 1996. V. 257. No. 5. P. 1112–1126.
- Hartley R.W. Barnase and barstar: two small proteins to fold and fit together // *Trends Biochem. Sci.* 1989. V. 14. No. 11. P. 450–454.
- Horovitz A., Matthews J.M., Fersht A.R. α -Helix stability in proteins. II. Factors that influence stability at an internal position // *J. Mol. Biol.* 1992. V. 227. No. 2. P. 560–568.
- Johnson C.M., Fersht A.R. Protein stability as a function of denaturant concentration: The thermal stability of barnase in the presence of urea // *Biochemistry*. 1995. V. 34. No. 20. P. 6795–6804.
- Kellis J.T. Jr., Nyberg K., Fersht A.R. Energetics of complementary side chain packing in a protein hydrophobic core // *Biochemistry*. 1989. V. 28. No. 11. P. 4914–4922.
- Kirkwood J.G. Statistical mechanics of fluid mixtures // *J. Chem. Phys.* 1935. V. 3/5. P. 300–314.
- Knight J.L., Brooks C.L. III. Lambda-dynamics free energy simulation methods // *J. Comput. Chem.* 2009. V. 30. P. 1692–1700.
- Kollman P. Free energy calculations: Applications to chemical and biochemical phenomena // *Chem. Rev.* 1993. V. 93. No. 7. P. 2395–2417.
- Kumar M.D.S., Bava K.A., Gromiha M.M. *et al.* ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions // *Nucl. Acids Res.* 2006. V. 34 (Database issue). P. D204–D206.
- Lehmann M., Wyss M. Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution // *Curr. Opin. Biotechnol.* 2001. V. 12. P. 371–375.
- Matouschek A., Kellis J.T. Jr., Serrano L., Fersht A.R. Mapping the transition state and pathway of protein folding by protein engineering // *Nature*. 1989. V. 340. No. 6229. P. 122–126.
- Martin C., Richard V., Salem M. *et al.* Refinement and structural analysis of barnase at 1.5 Å resolution // *Acta Crystallogr. D. Biol. Crystallogr.* 1999. V. 55. No. 2. P. 386–398.
- Mehareenna Y.T., Poulos T.L. Using molecular dynamics to probe the structural basis for enhanced stability in thermal stable cytochromes P450 // *Biochemistry*. 2010. V. 49. No. 31. P. 6680–6686.
- Messmer R.P. Cluster Model theory and its application to metal surface – adsorbate system // *The Nature of the Surface Chemical Bond* / Eds T.N. Rhodin, G. Ertl. Amsterdam, Noth-Holland Publ. Co., 1979. P. 51–111.
- Morales J., Martinez T.J. Classical fluctuating charge theories: The maximum entropy valence bond method and relationships to previous models // *J. Phys. Chem.* 2001. V. 105A. P. 2842–2850.
- Mortier W.J., Ghosh S.K., Shankar S.J. Electronegativity-equalization method for the calculation of atomic charges in molecules // *Am. Chem. Soc.* 1986. V. 108. P. 4315–4320.
- Oliferenko A.A., Palyulin V.A., Pisarev S.A. *et al.* Novel point charge models: reliable instruments for molecular electrostatics // *J. Phys. Org. Chem.* 2001. V. 14. P. 355–369.
- Pedersen J.S., Otzen D.E., Kristensen P. Directed evolution of barnase stability using proteolytic selection // *J. Mol. Biol.* 2002. V. 323. No. 1. P. 115–123.
- Pikkemaat M.G., Linssen A.B., Berendsen H.J., Janssen D.B. Molecular dynamics simulations as a tool for improving protein stability // *Protein Eng.* 2002. V. 15. No. 3. P. 185–192.
- Prevost M., Wodak S.J., Tidor B., Karplus M. Contribution of the hydrophobic effect to protein stability: analysis based on simulations of the Ile-96-Ala mutation in barnase // *Proc. Natl Acad. Sci. USA*. 1991. V. 88. No. 23. P. 10880–10884.
- Puchkaev A.V., Koo L.S., de Montellano P.R.O. Aromatic stacking as a determinant of the thermal stability of CYP119 from *Sulfolobus solfataricus* // *Arch. Biochem. Biophys.* 2003. V. 409. No. 1. P. 52–58.

- Rappe A.K., Goddard W.A. Charge equilibration for molecular dynamics simulations // *J. Phys. Chem.* 1991. V. 95. P. 3358–3363.
- Rick S.W., Stuart S.J., Berne B.J. Dynamical fluctuating charge force fields: Application to liquid water // *J. Chem. Phys.* 1994. V. 101. P. 6141–6156.
- Roux B., Karplus M. Ion transport in a model gramicidin channel. Structure and thermodynamics // *Biophys. J.* 1991. V. 59. No. 5. P. 961–981.
- Sanderson R.T. An interpretation of bond lengths and classification of bonds // *Science*. 1951. V. 114. P. 670–672.
- Seeliger D., de Groot B.L. Protein thermostability calculations using alchemical free energy simulations // *Biophys. J.* 2010. V. 98. No. 10. P. 2309–2316.
- Serrano L., Kellis J.T. Jr., Cann P. *et al.* The folding of an enzyme: II. Substructure of barnase and the contribution of different interactions to protein stability // *J. Mol. Biol.* 1992. V. 224. No. 3. P. 783–804.
- Simonson T., Archontis G., Karplus M. Free energy simulations come of age: Protein-ligand recognition // *Accounts Chem. Res.* 2002. V. 35. No. 6. P. 430–437.
- Tidor B., Karplus M. Simulation analysis of the stability mutant R96H of T4 lysozyme // *Biochemistry*. 1991. V. 30. No. 13. P. 3217–3228.
- York D.M., Yang W. A chemical potential equalization method for molecular simulations // *J. Chem. Phys.* 1996. V. 104. P. 159–172.
- Zhu S.B., Singh S., Robinson G.W. A new flexible/polarizable water model // *J. Chem. Phys.* 1991. V. 95. P. 2791–2799.
- Zwanzig R.W. High-temperature equation of state by a perturbation method. I. Nonpolar gases // *J. Chem. Phys.* 1954. V. 22. No. 8. P. 1420–1427.

STUDY OF THE THERMAL STABILITY OF BARNASE PROTEIN MUTANTS WITH MOLKERN SOFTWARE

E.S. Fomin, N.A. Alemasov

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

Summary

A thermal stability study of several barnase mutants has been carried out by the λ dynamics method. The method has been implemented in the MOLKERN software package. Mutations of amino acids with non-zero charge are chosen for the study, because in this case λ dynamics gives results differing dramatically (> 10 kJ/mol) from experimental data in nearly one-fourth of cases. A new modification of λ potentials is proposed, which takes into account charge changes, as well as the Net-Q method, in order to obtain reliable charge distributions for charged amino acids. The results obtained for the R72G mutation show a better agreement with experimental values than the results of other authors.

Keywords: barnase, molecular dynamics, λ dynamics, free energy difference, protein thermal stability.