

УДК: 577.322

Связь формы белка со скоростью его сворачивания и разворачивания

Н.С. Богатырева *, Д.Н. Иванков **, М.Ю. Лобанов ***,
О.В. Галзитская ****

*Институт белка, Российская академия наук, Пуцино, Московская область, 142290,
Россия*

Аннотация. Проблема самоорганизации белковых глобул является на сегодняшний день одной из важных проблем молекулярной биологии. Несмотря на успех последних лет в понимании общих закономерностей процесса сворачивания, детали этого процесса до конца не ясны. Поэтому поиск факторов, влияющих на кинетику сворачивания, продолжается по сей день. В этой работе мы проверяем предположение о том, что более округлые белки сворачиваются медленнее вытянутых белков. С этой целью мы анализируем ряд структурных параметров, предназначенных для описания формы белка, и их связь со скоростью сворачивания и разворачивания. Мы показываем, что безразмерные структурные параметры, описывающие форму белка, плохо коррелируют со скоростями сворачивания и разворачивания белков. В то же время структурные параметры, учитывающие и форму, и размер белка, дают хорошее согласие с экспериментально наблюдаемыми скоростями сворачивания и разворачивания белка.

Ключевые слова: компактность, радиус инерции, радиус поперечного сечения, простой механизм сворачивания «все или ничего», сложный механизм сворачивания с накоплением промежуточного состояния.

ВВЕДЕНИЕ

Проблема сворачивания белка является одной из самых важных и актуальных проблем биофизики. Помимо своей фундаментальной значимости, понимание механизма сворачивания белка имеет огромное значение для решения многих практических задач, таких как разработка лекарств и создание искусственных белков с заданными свойствами. Нарушение правильного сворачивания белков *in vivo*, а также часто сопутствующий этому процесс агрегации, во многих случаях приводит к заболеваниям [1]. Поэтому предсказание скорости сворачивания также имеет ценность само по себе.

Время сворачивания различных белков различается на много порядков, от микросекунд до секунд и даже часов. Небольшие белки обычно сворачиваются быстро и без интермедиатов сворачивания (то есть, процесс проходит в одну стадию, наблюдается «одностадийная» кинетика), более крупные белки сворачиваются

* bogat@phys.protres.ru

** ivankov13@gmail.com

*** mlobanov@phys.protres.ru

**** ogalzit@vega.protres.ru

медленней, а при сворачивании часто наблюдаются метастабильные интермедиаты (то есть, процесс сворачивания проходит во много стадий, наблюдается «многостадийная» кинетика) [2].

Одной из первых аналитических теорий сворачивания однодоменных глобулярных белков была теория Финкельштейна–Бадретдинова [3]. В рамках этой теории, разработанной на основе нуклеационного механизма, было получено, что скорость сворачивания белка в точке перехода зависит от величины границы раздела двух фаз в переходном состоянии, а петли, «торчащие» из свернутой части белка, создают дополнительное поверхностное натяжение, замедляя процесс сворачивания белка (рис. 1). Отсюда следует, что, поскольку граница раздела двух фаз зависит для округлой глобулы от числа аминокислотных остатков L белковой цепи как $L^{2/3}$, то и скорость сворачивания белка должна зависеть от числа аминокислотных остатков в белке таким же образом:

$$\ln k_{mt} \sim -L^{2/3}, \quad (1)$$

где $\ln k_{mt}$ – логарифм константы скорости сворачивания в точке термодинамического равновесия двух состояний белка, нативного и денатурированного.

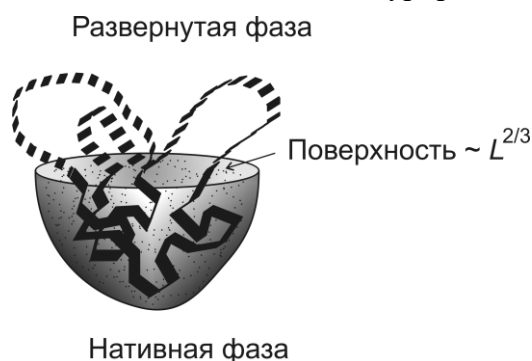


Рис. 1. Гипотетическое переходное состояние молекулы белка, соответствующее нуклеационному механизму. Аминокислотные остатки либо включены в нативную фазу (или в «ядро сворачивания», эти остатки показаны сплошной линией) и формируют нативные контакты с другими уже свернутыми остатками, либо находятся в развернутой фазе (эти остатки изображены пунктирной линией) — они полностью развернуты и не формируют никаких нативных контактов. Петли, выходящие из ядра, создают дополнительное поверхностное натяжение.

С помощью этой теории был решен парадокс Левинтала [3, 4], а также получена одна из первых формул по оценке скорости сворачивания глобулярных белков. Согласно другим оценкам, скорость сворачивания глобулярных белков в отсутствие денатуранта должна зависеть от числа аминокислотных остатков в белке как $-\ln L$ [5], как $-L^{1/2}$ [6] или как $-L^{0.61}$ [7]. Однако с точки зрения предсказания скорости сворачивания все предложенные зависимости скорости сворачивания от числа остатков в белковой цепи примерно одинаково согласуются с наблюдаемыми скоростями сворачивания белков: коэффициент корреляции составляет около 65% [7–9]. Также было показано, что размер белка определяет скорость сворачивания только белков с многостадийной кинетикой сворачивания [10].

Помимо размера белка, есть еще факторы, которые могут дополнительно влиять на скорость сворачивания белка. Так, из нуклеационного механизма следует, что топология переходного состояния — ход белковой цепи в пространстве — должна быть такой же, как и у нативной структуры [11]. Это значит, что чем больше в белке контактов между остатками, далекими по цепи, тем более вероятно, что в переходном состоянии белок не сможет избежать петель, торчащих из нативоподобной части белка (рис. 1), и сворачивание будет более медленным. Именно это и наблюдается при сравнении с экспериментом: для белков примерно одного размера логарифм скорости сворачивания уменьшается с увеличением параметра «порядка контактов» (“contact order”, CO), равного среднему расстоянию (в аминокислотных остатках) по цепи между атомами, контактирующими в нативной структуре, нормированному на число

аминокислотных остатков в белковой цепи [12]. Однако сама по себе топология не может объяснить разницу в скоростях сворачивания для некоторых белков с одним и тем же ходом белковой цепи (SH3 домены, белки холодового шока, домены фибронектина, а также белки ферредоксиновой укладки) [13–17].

Из-за того, что CO не зависит от размера белка, он не может предсказывать скорости сворачивания всех белков. Объединение CO и числа аминокислотных остатков в белке в одну формулу дает «абсолютный порядок контактов» $AbsCO = CO * L$, включающий в себя влияние и топологии, и длины цепи белка. Этот параметр лучше предсказывает скорости сворачивания белков, чем CO и $L^{2/3}$ в отдельности [18].

Поскольку предсказание скорости сворачивания само по себе тоже имеет ценность, в последние несколько лет было предложено много параметров для предсказания скорости сворачивания [10, 18–26]. Некоторые из них являются модификациями CO [18, 21, 25], в других предсказание скорости сворачивания осуществляется по вторичной структуре [20, 22] и по числу контактов в нативной структуре [27]. Кроме того, поиск зависимостей также идет и чисто статистическими методами [24, 26, 28] (без рассмотрения какой-либо физической модели). Было обнаружено, что для белков с одностадийной и многостадийной кинетикой сворачивания в воде наблюдается различная зависимость скорости сворачивания от аминокислотного состава [26], а также от физико-химических свойств аминокислот белка [24, 28]. Однако статистические методы не могут дать физического объяснения найденных результатов.

Вышеперечисленные, довольно противоречивые, данные показывают, что теория самоорганизации белков нуждается в дальнейшем развитии, и необходим дальнейший поиск факторов, влияющих на скорость сворачивания.

В этой работе мы проверяем предположение о том, что более округлые белки должны сворачиваться медленнее вытянутых белков [29]. Это предположение следует из теории Финкельштейна–Бадретдинова, основанной на нуклеационном механизме. Действительно, при прочих равных условиях округлый, наиболее компактный белок, независимо от пути сворачивания, не может избежать большой площади границы раздела двух фаз в переходном состоянии. А продолговатый, менее компактный белок имеет возможность «выбрать» себе такой путь сворачивания, в котором сворачивание белка идет через малую площадь границы раздела фаз и, следовательно, через менее высокий барьер.

Для проверки этого предположения нами был исследован ряд параметров, характеризующих «компактность» белковой глобулы с целью изучить, как они связаны со скоростями сворачивания и разворачивания в воде и скоростью сворачивания в точке термодинамического равновесия. Оказалось, что не связанные с размером белка параметры компактности, описывающие форму белковой глобулы, плохи для предсказания скоростей сворачивания белков, в то время как зависящие от размера белка параметры позволяют хорошо предсказывать скорость сворачивания белка.

МАТЕРИАЛЫ И МЕТОДЫ

Расчет площади доступной для растворителя поверхности S_{ASA} и молекулярной поверхности S_{mol} производился с помощью программы YASARA [30]. Площадью, доступной для растворителя поверхности S_{ASA} , считаем площадь, образованную центром модельной молекулы растворителя при обкатке ван-дер-ваальсовой поверхности белка этой молекулой. Модельная молекула растворителя представляет шар радиусом 1.4 Å.

Расчет недоступного растворителю объема белковой глобулы V_{ASA} (и V_{mol}) производился с помощью программы YASARA [30]. Недоступный растворителю объем белковой глобулы представляет собой объем, заключенный внутри доступной растворителю поверхности белка (молекулярной поверхности).

Набор исследуемых белков и белковых доменов был создан на основе базы структурной классификации белков SCOP [31] версии 1.65, с парной идентичностью аминокислотных последовательностей не выше 25%. В данном исследовании было проанализировано 3769 структур белковых доменов.

Расчетный радиус инерции белка в вакууме R_g — параметр, описывающий равновесное состояние конформации всей системы, — может быть рассчитан в два шага. Сначала находят координаты центра масс R_C из формулы:

$$\sum m_i(\mathbf{r}_i - \mathbf{R}_C) = 0, \quad (2)$$

где m_i — масса i -го неводородного атома, \mathbf{r}_i — его радиус вектор.

Если мы будем представлять атомы точками в трехмерном пространстве, то радиус инерции вычисляется по формуле:

$$R_g^2 = \sum m_i(\mathbf{r}_i - \mathbf{R}_C)^2 / M, \quad (3)$$

где M — масса всей системы.

Радиус инерции шара заданного объема V_{mol} рассчитывали по формуле

$$R_g^2 = \frac{3}{5} \left(\frac{3V_{mol}}{4\pi} \right)^{2/3}. \quad (4)$$

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для характеристики компактности белковых глобул нами были рассмотрены и рассчитаны следующие величины:

1) Параметры компактности, связанные с размером поперечного сечения белковой глобулы. Ранее в качестве меры округлости мы использовали отношение объема белка к площади его доступной поверхности, V_{ASA}/S_{ASA} [29]. Однако поскольку логарифм скорости сворачивания должен зависеть от величины площади поверхности раздела фаз в переходном состоянии, то вместо V_{ASA}/S_{ASA} более правильно, с физической точки зрения, использовать величину $(V_{ASA}/S_{ASA})^2$, которая имеет размерность поперечного сечения белка. Кроме этой величины, мы также проверили еще две величины, имеющие размерность поперечного сечения — $L^{2/3}$ и V_{ASA}/R_g . Надо сказать, что параметр $(V_{ASA}/S_{ASA})^2$ примерно пропорционален площади наименьшего сечения белковой глобулы в ее середине; эта величина пропорциональна $L^{2/3}$ для сферических глобул. Еще одной исследованной структурной характеристикой белка был радиус инерции белка R_g , рассчитанный по формуле (3).

2) Не связанные с размером белка параметры компактности, описывающие форму белковой глобулы. Это «коэффициент компактности», равный отношению площади доступной растворителю поверхности белка к площади поверхности шара того же объема, что и белок. Коэффициент компактности впервые появился в работе [32] как один из возможных способов нахождения белковых доменов. В дальнейшем эта величина с небольшими вариациями фигурировала в литературе как «индекс шероховатости» [33], «индекс глобулярности» [34] и как «компактность» S/S_{ASA}^* [35,36]. Кроме того, мы рассмотрели «порядок контактов» CO [12], радиус инерции белка, нормированный на радиус инерции шара того же объема, что и белок, R_g/R_g^* , а также отношение $V_{ASA}/S_{ASA}/R_g$.

Экспериментальные данные по кинетике сворачивания мы взяли из базы данных KineticDB [37]. Они включают в себя: $\ln k_F$ — логарифм константы скорости сворачивания белковой цепи в воде, $\ln k_U$ — логарифм константы скорости разворачивания белковой цепи в воде, $\ln k_{mt}$ — логарифм константы скорости сворачивания белковой цепи в точке термодинамического равновесия двух состояний белка, которое достигается добавлением денатуранта или увеличением температуры. Индекс «two» у константы скорости означает, что кинетика сворачивания белка описывается кинетикой двух состояний, в этом случае наблюдается одна стадия, а

белок называется «одностадийным». Индекс «*multi*» означает, что для описания кинетики сворачивания необходимо три или более состояний белковой цепи, в этом случае наблюдается больше одной стадии, а белок называется «многостадийным». В случае, когда мы рассматриваем всю совокупность белков без разделения на типы сворачивания, мы будем опускать индекс стадийности у логарифма скорости сворачивания белка.

Таблица 1. Коэффициенты корреляции параметров компактности, связанных с размером поперечного сечения белка, с логарифмами скоростей сворачивания и разворачивания в воде и в точке термодинамического равновесия.

	$\ln k_F$	$\ln k_F^{multi}$	$\ln k_F^{two}$
$L^{2/3}$	-0.69±0.06	-0.80±0.07	-0.47±0.10
$(V_{ASA}/S_{ASA})^2$	-0.73±0.05	-0.76±0.08	-0.57±0.09
V_{ASA}/R_g	-0.72±0.05	-0.81±0.07	-0.53±0.09
R_g^2	-0.53±0.08	-0.72±0.09	-0.32±0.12
	$\ln k_U$	$\ln k_U^{multi}$	$\ln k_U^{two}$
$L^{2/3}$	-0.73±0.05	-0.57±0.13	-0.74±0.06
$(V_{ASA}/S_{ASA})^2$	-0.77±0.05	-0.53±0.14	-0.82±0.04
V_{ASA}/R_g	-0.76±0.05	-0.56±0.13	-0.82±0.04
R_g^2	-0.60±0.07	-0.54±0.14	-0.55±0.09
	$\ln k_{mt}$	$\ln k_{mt}^{multi}$	$\ln k_{mt}^{two}$
$L^{2/3}$	-0.69±0.06	-0.60±0.13	-0.67±0.07
$(V_{ASA}/S_{ASA})^2$	-0.73±0.05	-0.55±0.14	-0.75±0.06
V_{ASA}/R_g	-0.73±0.05	-0.60±0.13	-0.74±0.06
R_g^2	-0.55±0.08	-0.55±0.14	-0.48±0.10

Таблица 2. Коэффициенты корреляции не связанных с размером белка параметров компактности, описывающих форму белковой глобулы, с логарифмами скоростей сворачивания и разворачивания в воде и в точке термодинамического равновесия.

	$\ln k_F$	$\ln k_F^{multi}$	$\ln k_F^{two}$
<i>CO</i>	-0.01±0.11	0.25±0.18	-0.41±0.11
S_{ASA}/S_{ASA}^*	-0.33±0.10	-0.48±0.15	-0.19±0.13
R_g/R_g^*	0.23±0.10	-0.01±0.20	0.20±0.13
$V_{ASA}/S_{ASA}/R_g$	0.17±0.11	0.42±0.16	0.07±0.13
	$\ln k_U$	$\ln k_U^{multi}$	$\ln k_U^{two}$
<i>CO</i>	0.07±0.11	0.04±0.20	-0.09±0.13
S_{ASA}/S_{ASA}^*	-0.43±0.09	-0.34±0.17	-0.43±0.11
R_g/R_g^*	0.16±0.11	-0.15±0.19	0.15±0.13
$V_{ASA}/S_{ASA}/R_g$	0.29±0.10	0.39±0.17	0.26±0.12
	$\ln k_{mt}$	$\ln k_{mt}^{multi}$	$\ln k_{mt}^{two}$
<i>CO</i>	0.01±0.11	-0.01±0.20	-0.16±0.13
S_{ASA}/S_{ASA}^*	-0.39±0.09	-0.39±0.17	-0.37±0.11
R_g/R_g^*	0.18±0.11	-0.19±0.19	0.17±0.13
$V_{ASA}/S_{ASA}/R_g$	0.24±0.10	0.43±0.16	0.20±0.13

Чтобы показать, как рассматриваемые нами параметры влияют на скорость сворачивания и разворачивания белковой молекулы, рассчитывается коэффициент корреляции между этими величинами. Результаты приведены в табл. 1 для размерных параметров компактности и в табл. 2 для безразмерных параметров компактности белковых глобул.

Видно, что, в соответствии с нашим предположением, параметры компактности, связанные с размером поперечного сечения белковой глобулы, в среднем, хорошо предсказывают скорости сворачивания и разворачивания белков, а также скорости сворачивания в точке термодинамического равновесия. В то же время, не связанные с размером белка параметры компактности, описывающие форму белковой глобулы, плохи для предсказания скоростей сворачивания и разворачивания белков.

Стоит заметить, что не всякий размерный параметр хорошо коррелирует со скоростью сворачивания и разворачивания белка. Из табл. 1 видно, что параметр $(V_{ASA}/S_{ASA})^2$, пропорциональный площади наименьшего поперечного сечения белковой глобулы, значительно лучше подходит для предсказания скорости сворачивания и разворачивания белка, чем R_g^2 , пропорциональный площади наибольшего сечения, что полностью подтверждает нашу гипотезу о роли наименьшего поперечного сечения белковой глобулы.

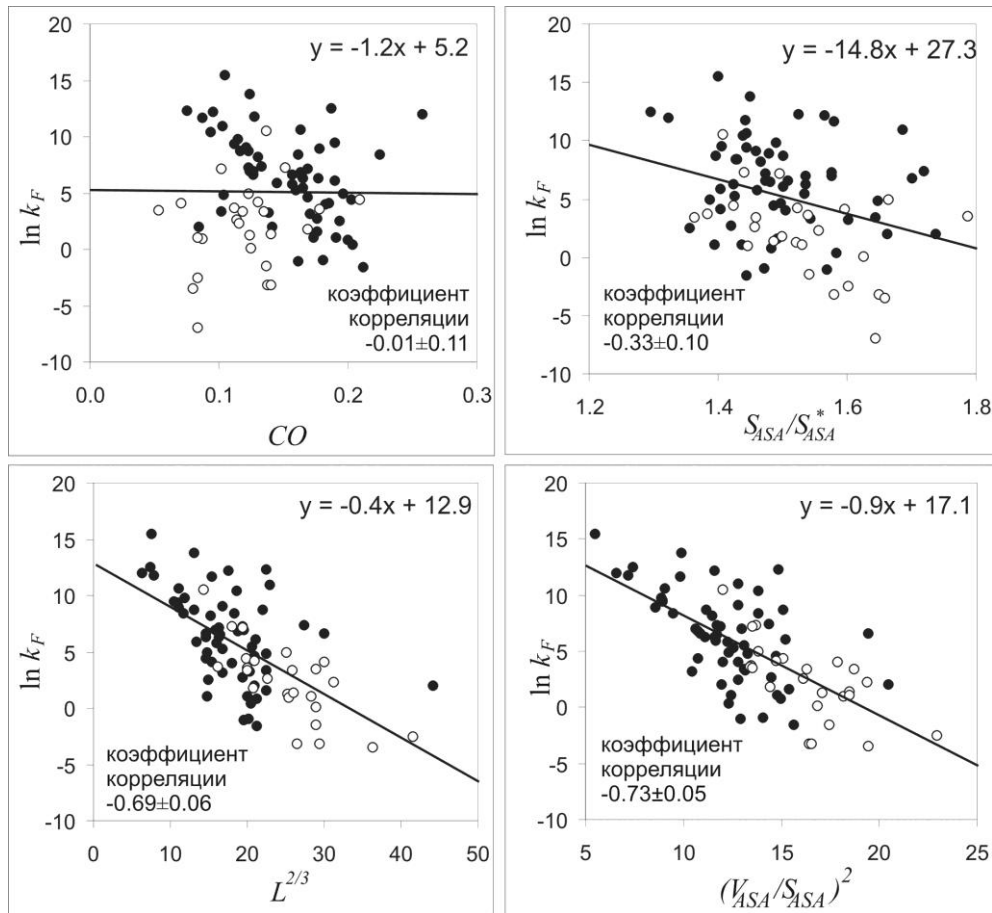


Рис. 2. Зависимости скоростей сворачивания/разворачивания белков $\ln k_F$ от нескольких из исследованных нами параметров компактности. Закрашенными кружками изображены белки, сворачивающиеся по одностадийному механизму, пустыми кружками — по многостадийному.

Для иллюстрации вышеизложенного на рис. 2 показаны характерные зависимости логарифма скорости сворачивания в воде: для не зависящего от длины цепи белка «порядка контактов» CO корреляция отсутствует вообще (коэффициент корреляции равен -0.01). Безразмерный параметр S_{ASA}/S_{ASA}^* плохо коррелирует со скоростью сворачивания (коэффициент корреляции равен -0.39), параметр длины цепи белка $L^{2/3}$

грубо предсказывает скорости сворачивания (коэффициент корреляции равен -0.69), а параметр $(V_{ASA}/S_{ASA})^2$, пропорциональный площади наименьшего поперечного сечения белковой глобулы, дает корреляцию -0.73 .

Таким образом, более округлые белки, действительно, сворачиваются и разворачиваются медленнее вытянутых белков. При прочих равных условиях округлый, наиболее компактный белок, независимо от пути сворачивания, не может избежать большой площади границы раздела двух фаз в переходном состоянии. А вытянутый, менее компактный белок имеет возможность «выбрать» себе такой путь сворачивания, в котором сворачивание белка идет через малую площадь границы раздела фаз и, следовательно, через менее высокий барьер. Очень характерный случай для иллюстрации этого тезиса показан на рис.3, где изображены структуры двух белков примерно одинакового размера. Из рис.3 видно, что площадь наименьшего поперечного сечения, проходящего через середину белка, для быстрее сворачивающегося, вытянутого белка VlsE меньше, чем для круглого белка, α -субъединицы триптофан-синтазы, сворачивающегося медленней.

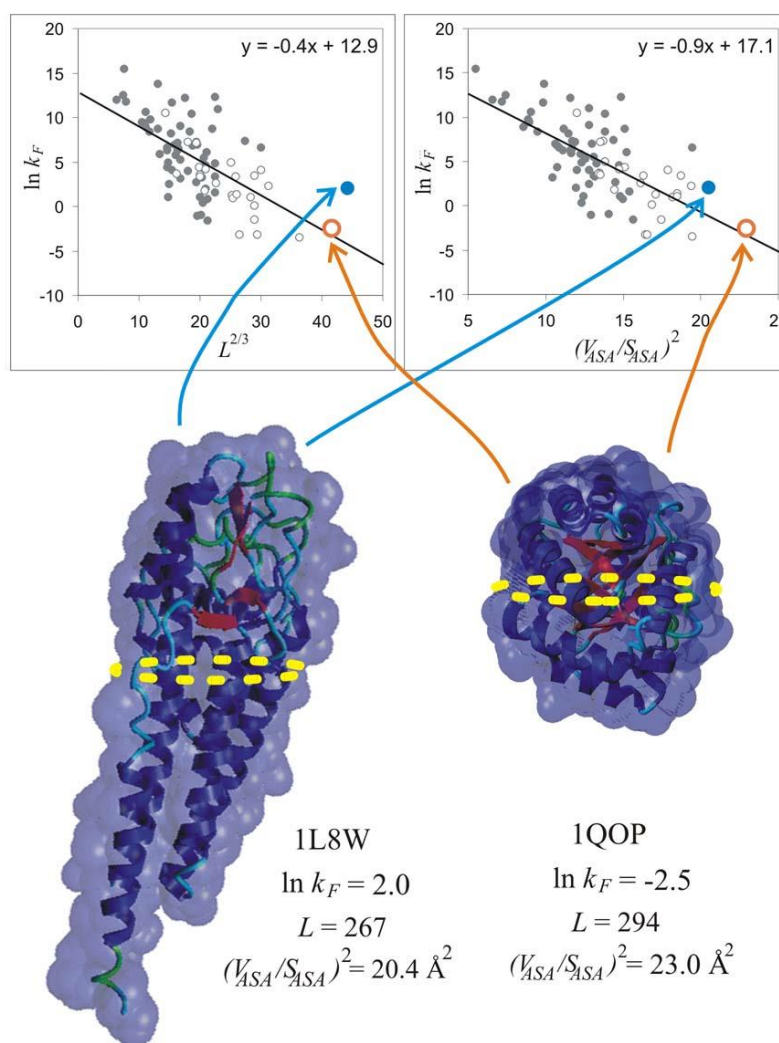


Рис. 3. Пример сравнения двух больших белков примерно одного размера. Менее сферичный одностадийный белок VlsE изображен слева (PDB код 1L8W), а более сферичный многостадийный белок α -субъединица триптофан-синтазы (PDB код 1QOP), изображен справа. Более сферичный белок α -субъединица триптофан-синтазы сворачивается медленней и имеет большее значение величины $(V_{ASA}/S_{ASA})^2$. Для удобства сравнения размеров и величин $(V_{ASA}/S_{ASA})^2$ сверху приведены графики зависимости $\ln k_F$ от $L^{2/3}$ и от $(V_{ASA}/S_{ASA})^2$, на них VlsE выделен сплошным синим кружком, α -субъединица триптофан-синтазы – пустым оранжевым кружком.

ЗАКЛЮЧЕНИЕ

Возможность успешного предсказания скорости сворачивания и разворачивания белков, исходя из физической идеи, является свидетельством правильного понимания механизма самоорганизации белков. Из рассмотренного анализа видно, что безразмерные параметры, описывающие форму белка, плохи для предсказания скоростей сворачивания белков. Также видно, что учет размера наименьшего поперечного сечения белковой глобулы улучшает качество предсказания скоростей сворачивания белков.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (№ 08-04-00561) и Фонда содействия отечественной науке.

Мы благодарим А.В. Финкельштейна за обсуждение результатов.

СПИСОК ЛИТЕРАТУРЫ

1. Chiti F., Dobson C.M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 2006. V. 75. P. 333–366.
2. Jackson S.E. How do small single-domain proteins fold? *Fold. Des.* 1998. V. 3. № 4. P. R81–R91.
3. Финкельштейн А.В., Бадретдинов А.Я. Физические причины быстрой самоорганизации стабильной пространственной структуры белков: решение парадокса Левинтала. *Мол. биология.* 1997. Т. 31. С. 469–477.
4. Levinthal C. Are there pathways for protein folding? *J. Chim. Phys. Chim. Biol.* 1968. V. 65. P. 44–45.
5. Gutin A.M., Abkevich V.I., Shakhnovich E.I. Chain length scaling of protein folding time. *Phys. Rev. Lett.* 1996. V. 77. № 27. P. 5433–5456.
6. Thirumalai D. From minimal models to real proteins: time scales for protein folding kinetics. *Journal de Physique Orsay Fr.* 1995. V. 5. P. 1457–1467.
7. Koga N., Takada S. Roles of native topology and chain-length scaling in protein folding: a simulation study with a Gō-like model. *J. Mol. Biol.* 2001. V. 313. № 1. P. 171–180.
8. Finkelstein A.V., Galzitskaya O.V. Physics of protein folding. *Phys. Life Rev.* 2004. V. 1. № 1. P. 23–56.
9. Galzitskaia O.V., Ivankov D.N., Finkelstein A.V. Folding nuclei in proteins. *FEBS Lett.* 2001. V. 489. № 2-3. P. 113–118.
10. Galzitskaya O.V., Garbuzynskiy S.O., Ivankov D.N., Finkelstein A.V. Chain length is the main determinant of the folding rate for proteins with three-state folding kinetics. *Proteins.* 2003. V. 51. № 2. P. 162–166.
11. Fersht A.R. Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.* 1997. V. 7. № 1. P. 3–9.
12. Plaxco K.W., Simons K.T., Baker D. Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* 1998. V. 277. № 4. P. 985–994.
13. Plaxco K.W., Guijarro J.I., Morton C.J., Pitkeathly M., Campbell I.D., Dobson C.M. The folding kinetics and thermodynamics of the Fyn-SH3 domain. *Biochemistry.* 1998. V. 37. № 8. P. 2529–2537.
14. Guijarro J.I., Morton C.J., Plaxco K.W., Campbell I.D., Dobson C.M. Folding kinetics of the SH3 domain of PI3 kinase by real-time NMR combined with optical spectroscopy. *J. Mol. Biol.* 1998. V. 276. № 3. P. 657–667.
15. Zerovnik E., Virden R., Jerala R., Turk V., Waltho J.P. On the mechanism of human stefin B folding: I. Comparison to homologous stefin A. Influence of pH and trifluoroethanol on the fast and slow folding phases. *Proteins.* 1998. V. 32. № 3. P. 296–303.

16. Perl D., Welker C., Schindler T., Schroder K., Marahiel M.A., Jaenicke R., Schmid F.X. Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat. Struct. Biol.* 1998. V. 5. № 3. P. 229–235.
17. van Nuland N.A., Chiti F., Taddei N., Raugei G., Ramponi G., Dobson C.M. Slow folding of muscle acylphosphatase in the absence of intermediates. *J. Mol. Biol.* 1998. V. 283. № 4. P. 883–891.
18. Ivankov D.N., Garbuzynskiy S.O., Alm E., Plaxco K.W., Baker D., Finkelstein A.V. Contact order revisited: influence of protein size on the folding rate. *Protein Sci.* 2003. V. 12. № 9. P. 2057–2062.
19. Punta M., Rost B. Protein folding rates estimated from contact predictions. *J. Mol. Biol.* 2005. V. 348. № 3. P. 507–512.
20. Ivankov D.N., Finkelstein A.V. Prediction of protein folding rates from the amino acid sequence-predicted secondary structure. *Proc. Natl. Acad. Sci. USA.* 2004. V. 101. № 24. P. 8942–8944.
21. Zhou H., Zhou Y. Folding rate prediction using total contact distance. *Biophys. J.* 2002. V. 82. № 1. Pt 1. P. 458–463.
22. Gong H., Isom D.G., Srinivasan R., Rose G.D. Local secondary structure content predicts folding rates for simple, two-state proteins. *J. Mol. Biol.* 2003. V. 327. № 5. P. 1149–1154.
23. Capriotti E., Casadio R. K-Fold: a tool for the prediction of the protein folding kinetic order and rate. *Bioinformatics.* 2007. V. 23. № 3. P. 385–386.
24. Gromiha M.M., Thangakani A.M., Selvaraj S. FOLD-RATE: prediction of protein folding rates from amino acid sequence. *Nucleic Acids Res.* 2006. V. 34. P. W70–W74.
25. Gromiha M.M., Selvaraj S. Comparison between long-range interactions and contact order in determining the folding rate of two-state proteins: application of long-range order to folding rate prediction. *J. Mol. Biol.* 2001. V. 310. № 1. P. 27–32.
26. Ma B.G., Chen L.L., Zhang H.Y. What determines protein folding type? An investigation of intrinsic structural properties and its implications for understanding folding mechanisms. *J. Mol. Biol.* 2007. V. 370. № 3. P. 439–448.
27. Makarov D.E., Keller C.A., Plaxco K.W., Metiu H. How the folding rate constant of simple, single-domain proteins depends on the number of native contacts. *Proc. Natl. Acad. Sci. USA.* 2002. V. 99. № 6. P. 3535–3539.
28. Gromiha M.M. A statistical model for predicting protein folding rates from amino acid sequence with structural class information. *J. Chem. Inf. Model.* 2005. V. 45. № 2. P. 494–501.
29. Galzitskaya O.V., Bogatyreva N.S., Ivankov D.N. Compactness determines protein folding type. *J. Bioinform. Comput. Biol.* 2008. V. 6. № 4. P. 667–680.
30. Krieger E., Koraimann G., Vriend G. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins.* 2002. V. 47. № 3. P. 393–402.
31. Murzin A.G., Brenner S.E., Hubbard T., Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 1995. V. 247. № 4. P. 536–540.
32. Wetlaufer D.B. Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci. USA.* 1973. V. 70. № 3. P. 697–701.
33. Richards F.M. Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* 1977. V. 6. P. 151–76.
34. Wodak S.J., Janin J. Location of structural domains in protein. *Biochemistry.* 1981. V. 20. № 23. P. 6544–6552.
35. Zehfus M.H., Rose G.D. Compact units in proteins. *Biochemistry.* 1986. V. 25. № 19. P. 5759–5765.

36. Tsai C.J., Nussinov R. Hydrophobic folding units at protein-protein interfaces: implications to protein folding and to protein-protein association. *Protein Sci.* 1997. V. 6. № 7. P. 1426–1437.
37. Bogatyreva N.S., Osypov A.A., Ivankov D.N. KineticDB: a database of protein folding kinetics. *Nucleic Acids Res*, 2009, Database Issue, в печати

Материал поступил в редакцию 14.11.2008, опубликован 01.12.2008.