

75-е БАХОВСКОЕ ЧТЕНИЕ
15 марта 2019 г.
Институт биохимии им. А.Н. Баха

60 лет самоорганизации белков

А. В. Финкельштейн

Института белка РАН
Пушино, Московская область

ЦЕПЬ ГЛОБУЛЯРНОГО БЕЛКА

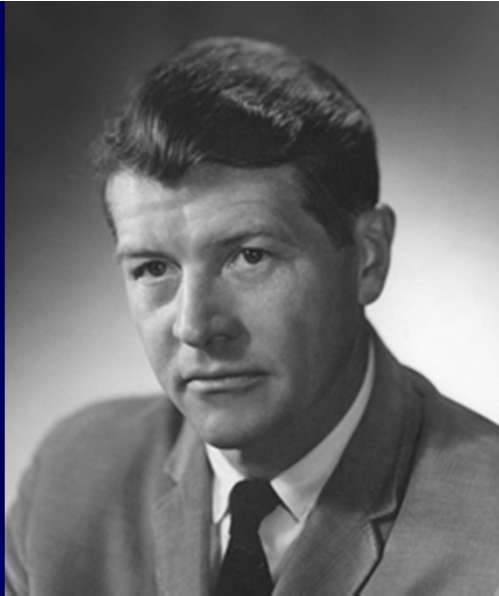


1-доменный
водорастворимый
глобулярный
белок
рибонуклеаза А
124 а.к.
 $\varnothing \approx 30\text{\AA}$

**МОЖЕТ СФОРМИРОВАТЬ СВОЮ УНИКАЛЬНУЮ
3-мерную СТРУКТУРУ СПОНТАННО *IN VITRO***

Anfinsen, 1961: Нобелевская премия, 1972

Merrifield, 1969: Нобелевская премия, 1984



Christian Boehmer
Anfinsen, Jr. (1916 – 1995)
Нобелевская премия 1972

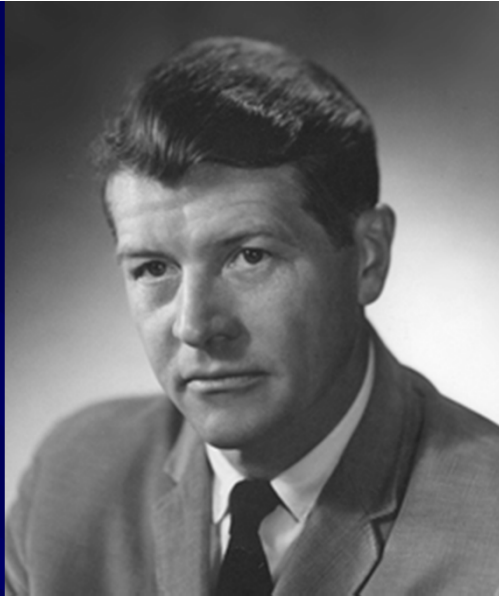
*Ренатурация природной рибонуклеазы А (1961)
“**термодинамическая гипотеза**”*



Robert Bruce
Merrifield (1921 – 2006)
Нобелевская премия 1984

*Сворачивание химически синтезированной
рибонуклеазы А (1969)*





Christian Boehmer
Anfinsen, Jr. (1916 –1995)
Нобелевская премия 1972

*Ренатурация природной рибонуклеазы А (1961)
“термодинамическая гипотеза”*



Robert Bruce
Merrifield (1921 – 2006)
Нобелевская премия 1984

*Сворачивание химически синтезированной
рибонуклеазы А (1969)*



*Ренатурация природной рибонуклеазы А (1956-59, PhD):
Образование правильных S-S связей, активность - 40%*

Lisa A. **Steiner** (1936)
MIT Professor of immunology (1967-)

(Сообщение: **D. Eisenberg**, *Cell*, 2018, 174, 8-11:
“Она ответила на важнейший вопрос, но его еще никто не задавал”)

Самоорганизация *in vitro*:

Цепь глобулярного белка сворачивается не только *in vivo* (на рибосоме), но и спонтанно *in vitro*

[Steiner, Anfinsen: 1959-1961; Merrifield: 1969]

Время самоорганизации: *in vitro*: от микросекунд до часов
in vivo: за минуты или менее

Ко-трансляционное сворачивание белка *in vivo* (на рибосоме) :
≈ как *in vitro* или принципиально другое ?

**Ко-трансляционное сворачивание белка in vivo (на рибосоме)
по крайней мере для небольших белков:**

≈ как *in vitro*

¹⁵N, ¹³C NMR:

Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. Eichmann C, Preissler S, Riek R, Deuerling E. *PNAS* 107, 9111 (2010):

«Полипептид [на рибосоме] остается неструктурированным во время элонгации, но сворачивается в компактную, нативоподобную структуру, когда сделана *вся цепь*.»

Тот же метод, тот же вывод: Han Y., ..., Qian S.-B., *PNAS* 109, 12467 (2012)

FRET (1 – 34):

Cotranslational protein folding on the ribosome monitored in real time

Holtkamp W., Kobic G., Jäger M., Joerg Mittelstaet J., Komar A.A., Rodnina M.V. *SCIENCE* 350, 1104 (2015)

«... котрансляционное сворачивание ... проходит через компактную ненативную конформацию, которая ... перестраивается в нативную структуру сразу после того, как *полная последовательность доменов* появилась из рибосомы.»

То же – in vitro:

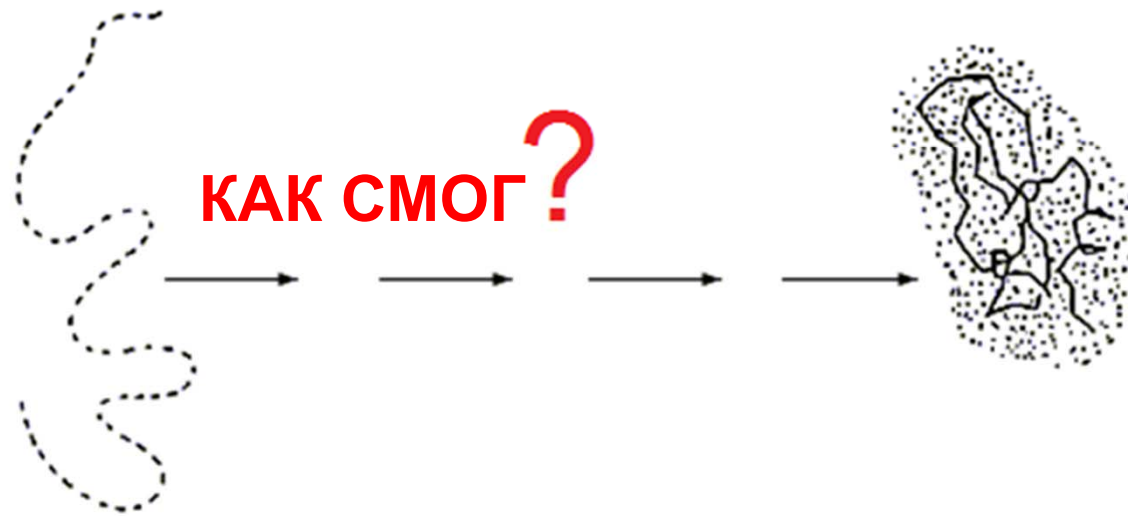
Flanagan J.M., Kataoka M., Shortle D., Engelman D.M. *Truncated [from C-end] staphylococcal nuclease is compact but disordered*. *PNAS* 89, 748 (1992).



Cyrus Levinthal
(1922 – 1990)

«Парадокс Левинтала» [1968-69] - парадокс спонтанной самоорганизации структуры белка

развернутая цепь: бездна конформаций ($>3^{100}$) нативная структура: 1 конформация



Как избежать полного перебора?? 3^{100} структур — 10^{26} лет??

ОСОБЫЕ ПУТИ??
ОСОБЫЕ ИНТЕРМЕДИАТЫ??

Как избежать
перебора всех 3^{100} структур??

Зачем знать, **КАК СМОГ** самоорганизоваться белок?

Во-первых, это просто интересно!

Для биологии — ??

Для дизайна искусственных белков —
надо знать, что строить:

шоссе (хороший путь к нативной структуре белка)

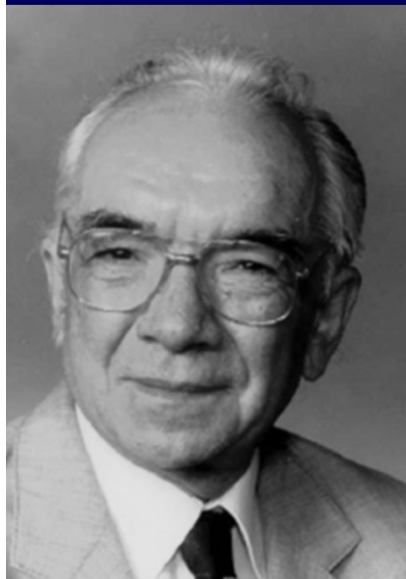
или

аэропорт (стабильную нативную структуру белка)?

Как избежать перебора??

НАЙТИ ОСОБЫЕ ИНТЕРМЕДИАТЫ??

“Рамочная модель” **стадийного** сворачивания белка (Птицын, 1973)



Олег
Борисович
ПТИЦЫН
(1929-1999)



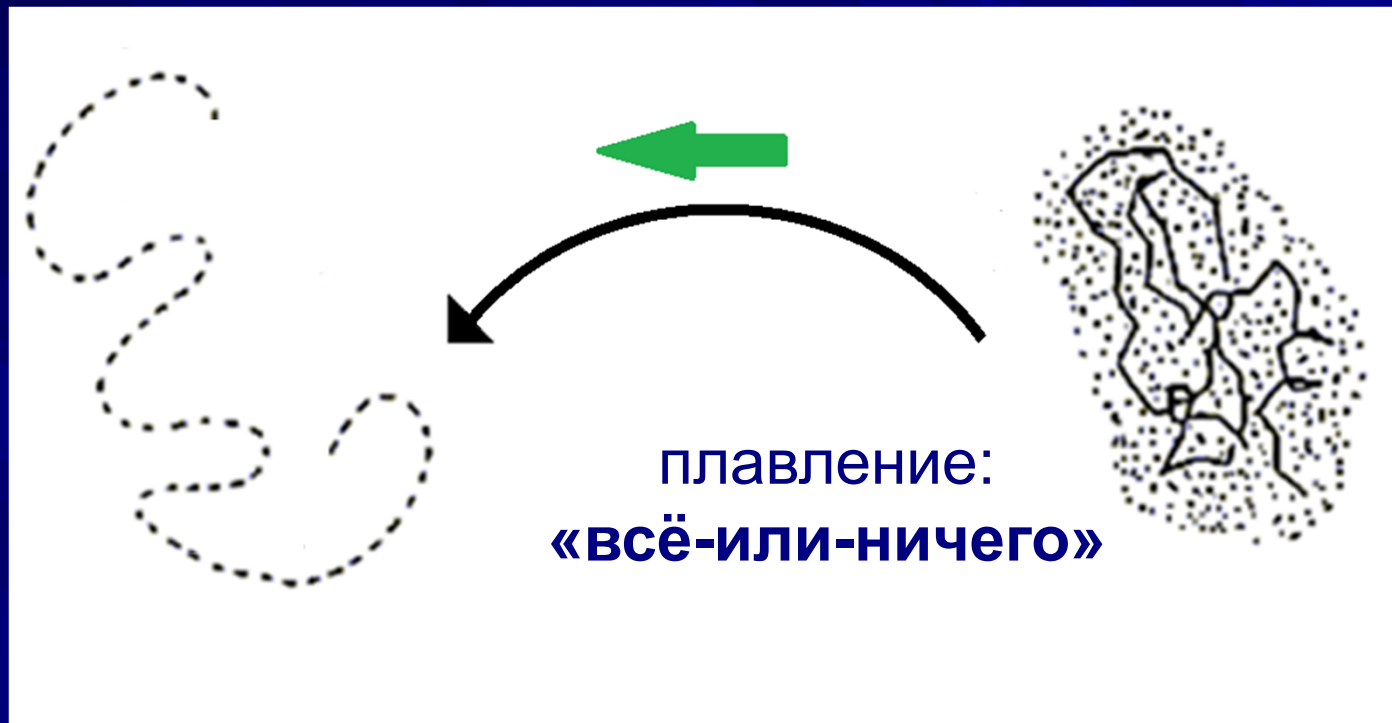
Перебор резко уменьшается, **НО** только если стабильность сильно растет от стадии к стадии

Однако: ЗАЧАСТУЮ НЕТ ИНТЕРМЕДИАТОВ
при денатурации белка!

(Привалов и др., 1973-74)

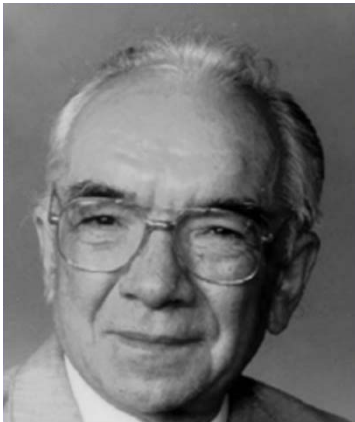


Петр
Леонидович
Привалов
1932



(по термодинамическому
критерию Ван'т Гоффа)

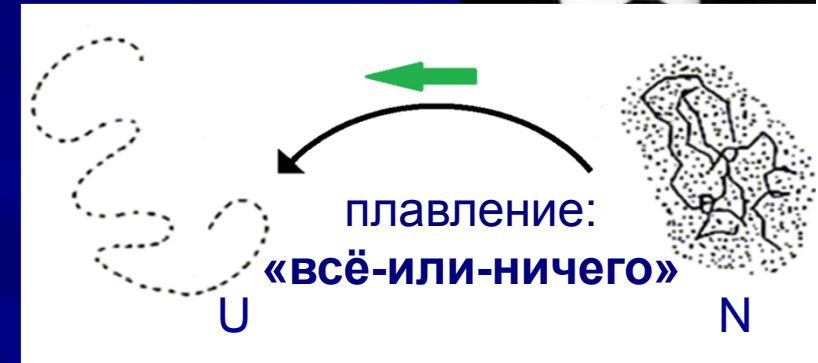
Два взгляда на сворачивание/разворачивание белка



ПТИЦЫН

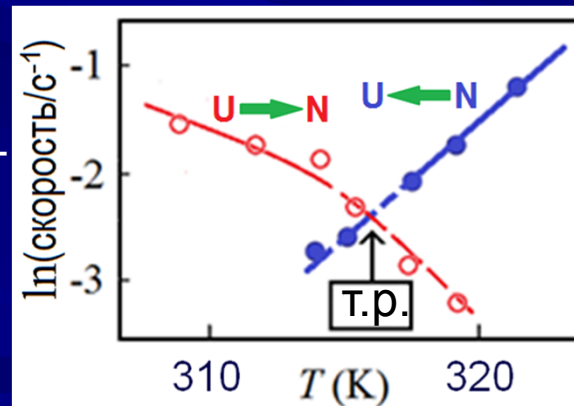


ПРИВАЛОВ



НО:

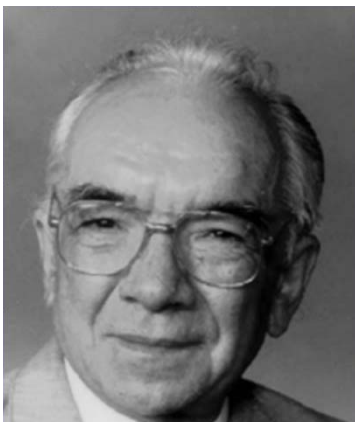
обратимость перехода – (Segawa, Sugihara, 1984)



– свидетельство одинаковости путей сворачивания и разворачивания в точке равновесия (т.р.)



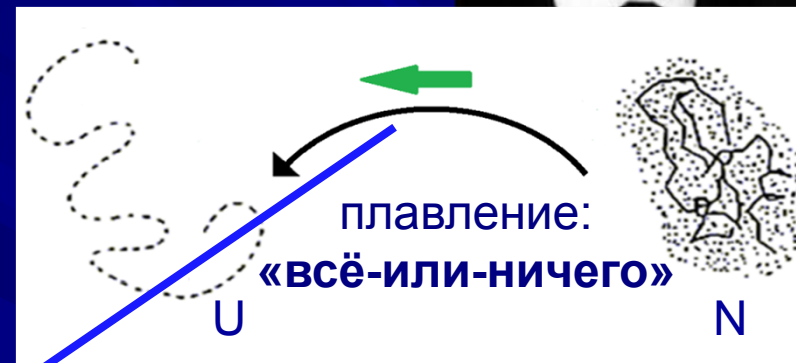
Два взгляда на сворачивание/разворачивание белка



ПТИЦЫН

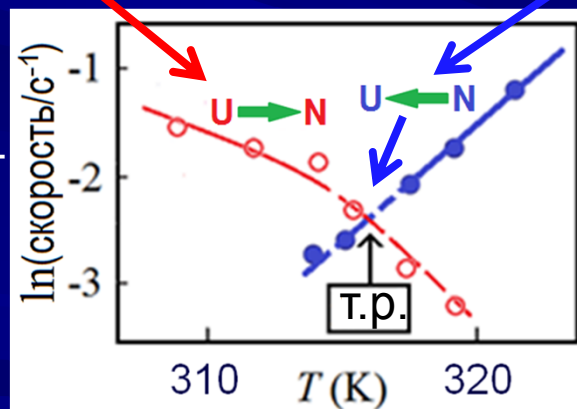


ПРИВАЛОВ



НО:

обратимость перехода – (Segawa, Sugihara, 1984)



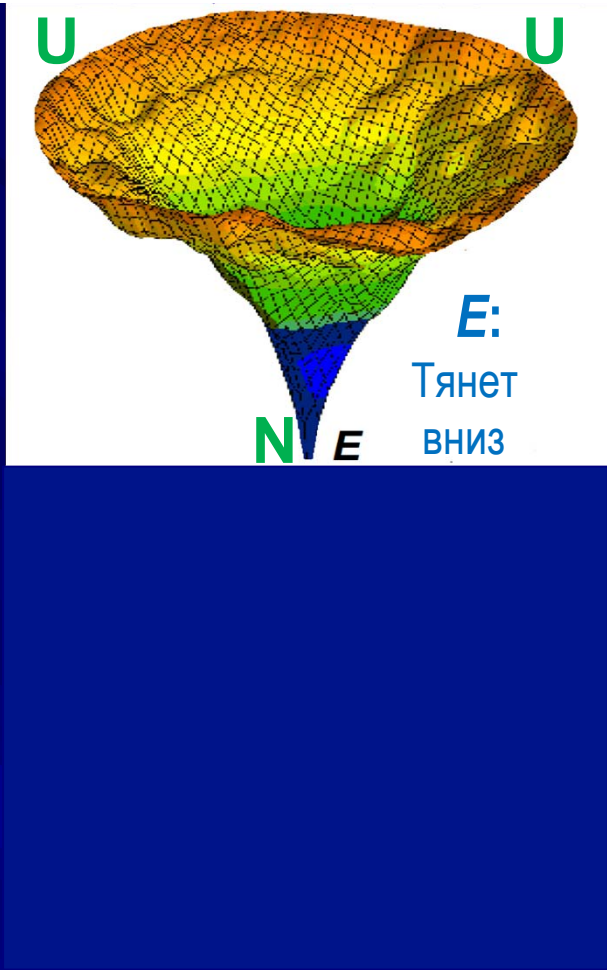
– свидетельство одинаковости путей сворачивания и разворачивания в точке равновесия (т.р.)

Вопрос – почему белок сворачивается не за 10^{26} лет, а за минуты – **относится и к точке равновесия**, где интермедиатов сворачивания нет!

Как избежать перебора??

Энергетическая «воронка»

Onuchic, Zwanzig, Szabo, Wolynes, Dill, etc., 1992 --

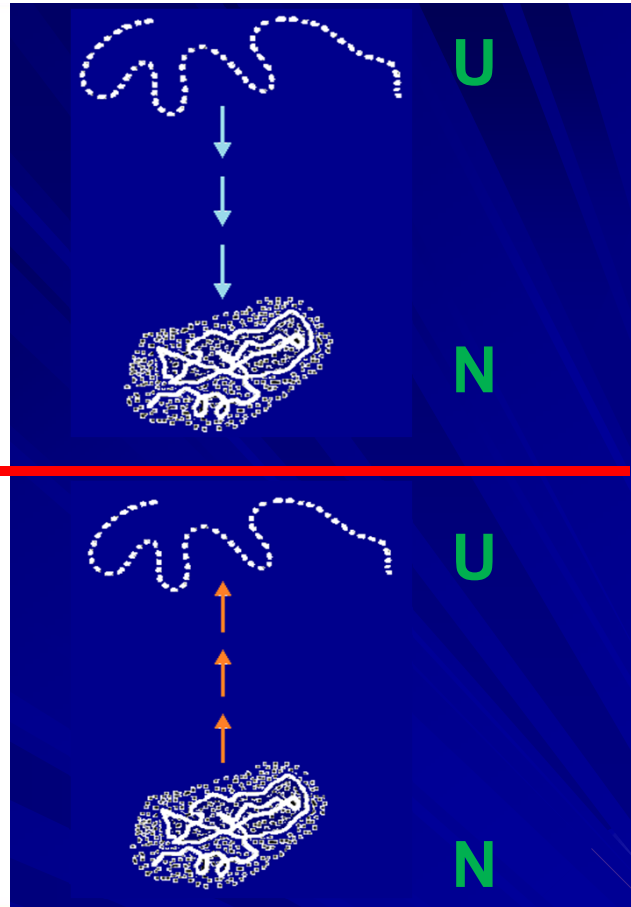
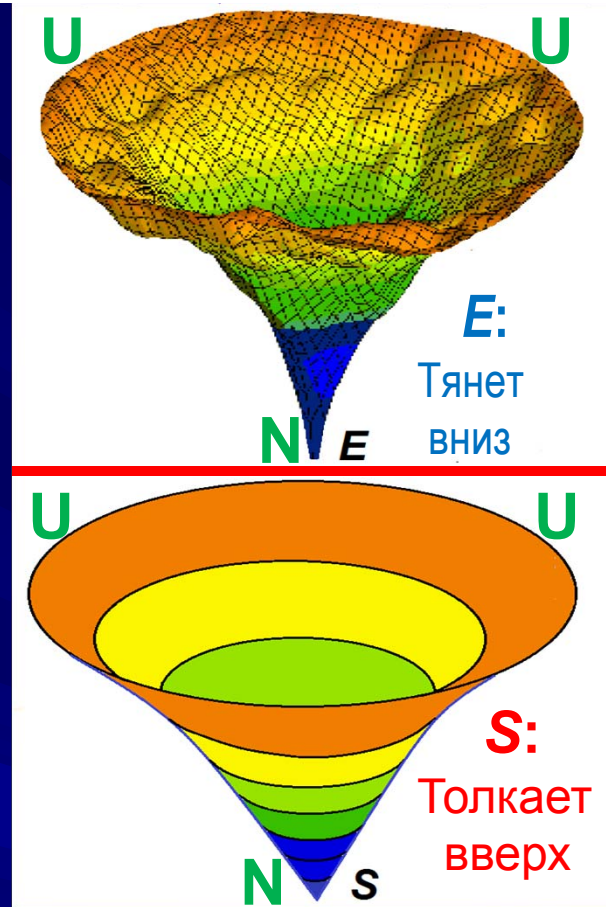


Как избежать перебора??

Энергетическая «воронка»

Onuchic, Zwanzig, Szabo, Wolynes, Dill, etc., 1992 --

НО: Энтропия: «анти-воронка»

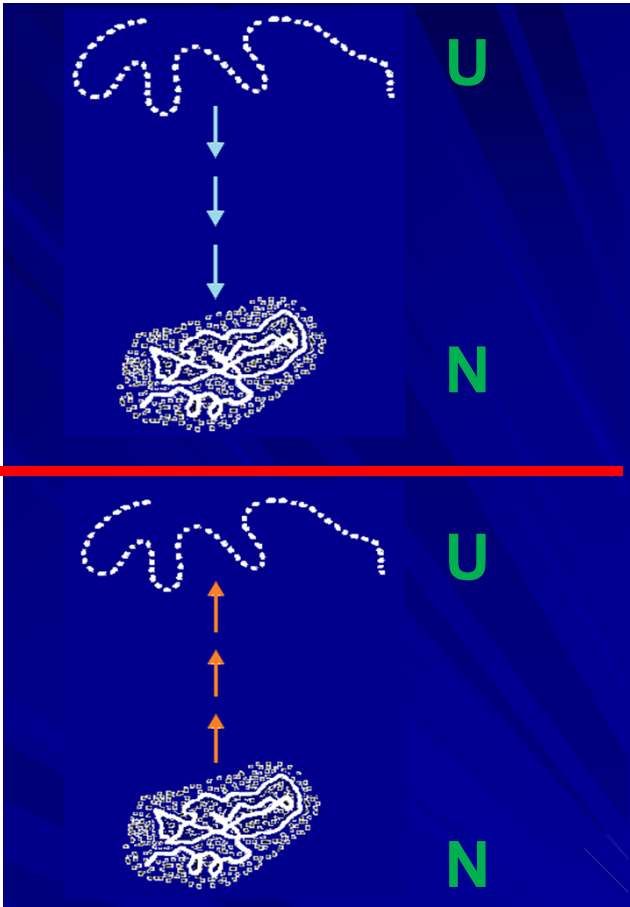
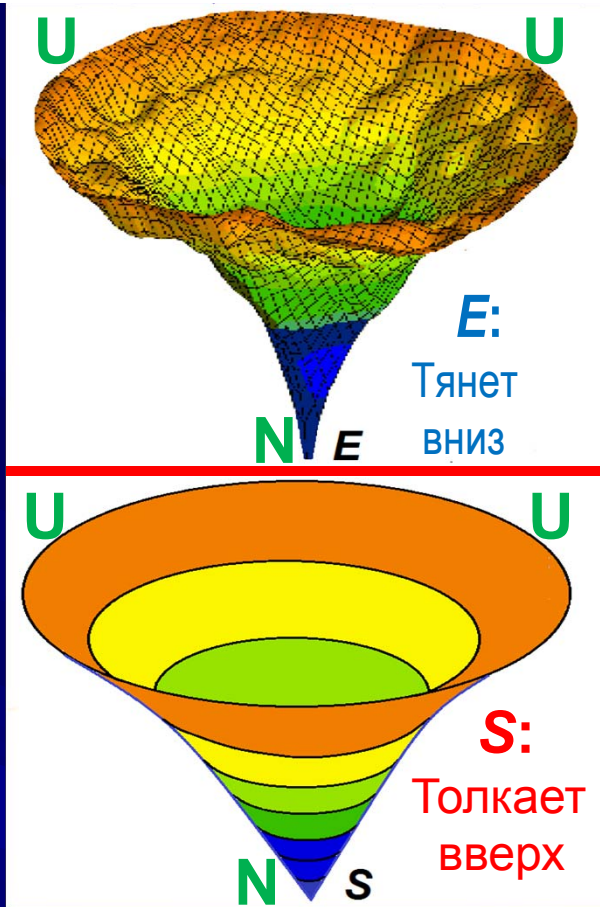


Как избежать перебора??

Энергетическая «воронка»

Onuchic, Zwanzig, Szabo, Wolynes, Dill, etc., 1992 --

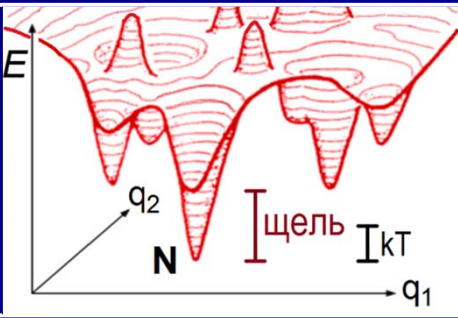
НО: Энтропия: «анти-воронка»



равно-стабильные **U** и **N** разделены

барьером (переход «всё-или-ничего»)

требует а. к. последовательности, обеспечивающей большую ($\gg kT$) «энергетическую щель» между наиболее стабильной структурой **N** и ее конкурентами



Shakhnovich, Gutin 1989-1993;
Šali, Shakhnovich, Karplus 1994

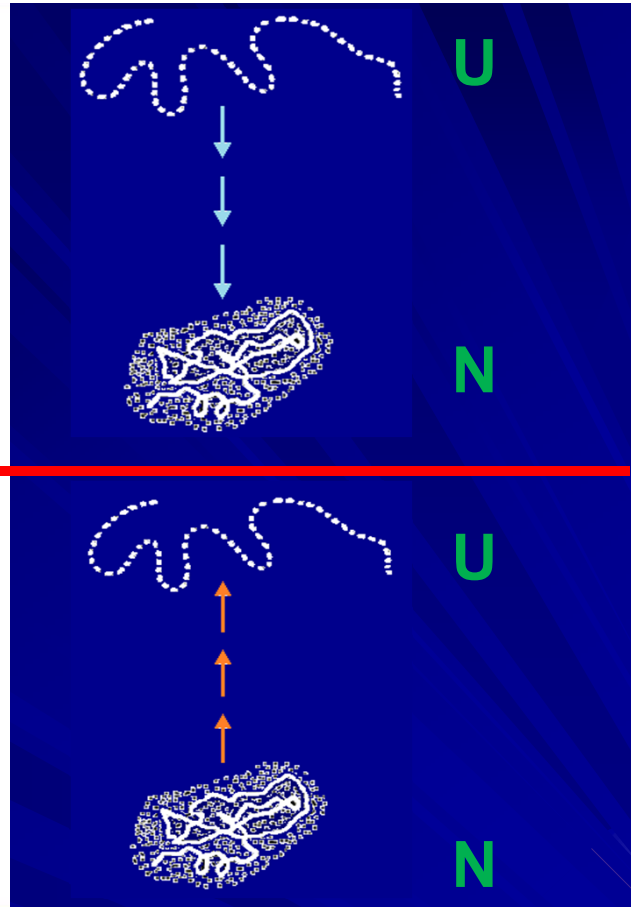
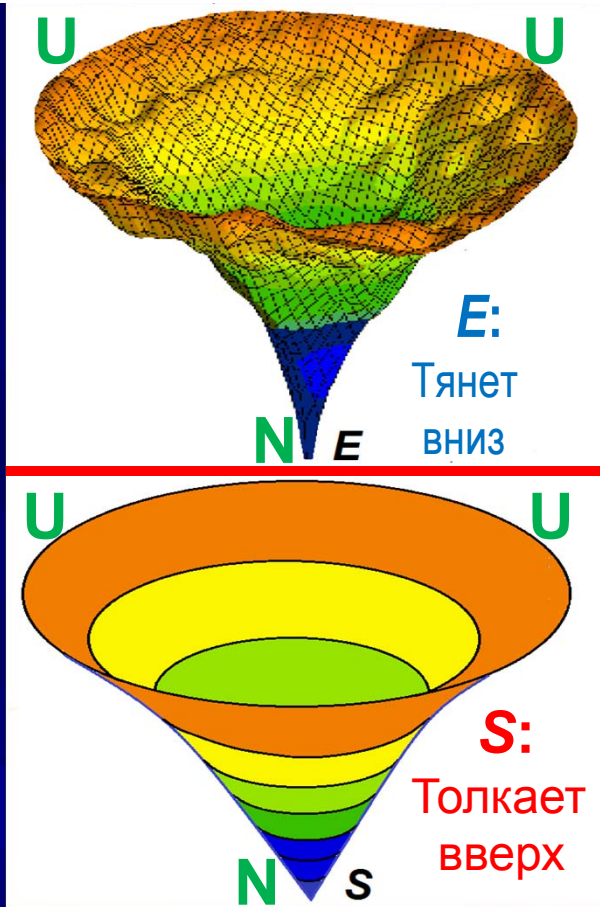


Как избежать перебора??

Энергетическая «воронка»

Onuchic, Zwanzig, Szabo, Wolynes, Dill, etc., 1992 --

НО: Энтропия: «анти-воронка»



Свободно-энергетический ландшафт: Если равно-стабильные **U** и **N** разделены барьером (переход «всё-или-ничего»): **вулкан**

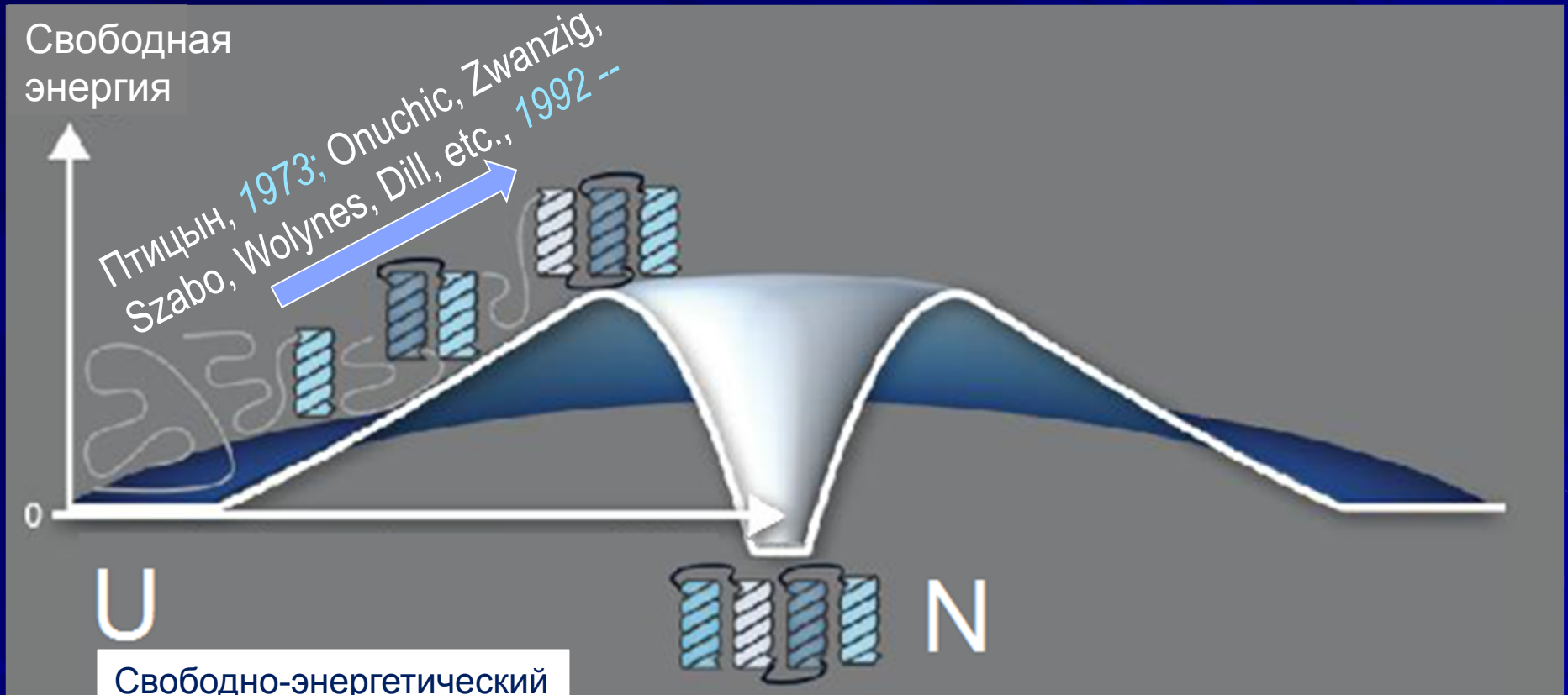


— и тогда E-воронка «тянет» ЛИШЬ близь **N**

Bogatyreva, Finkelstein, Prot. Eng. 2001
Dill, Rollins, JACS 2014

вулкан \approx парадокс Левинталя!

Как избежать перебора??

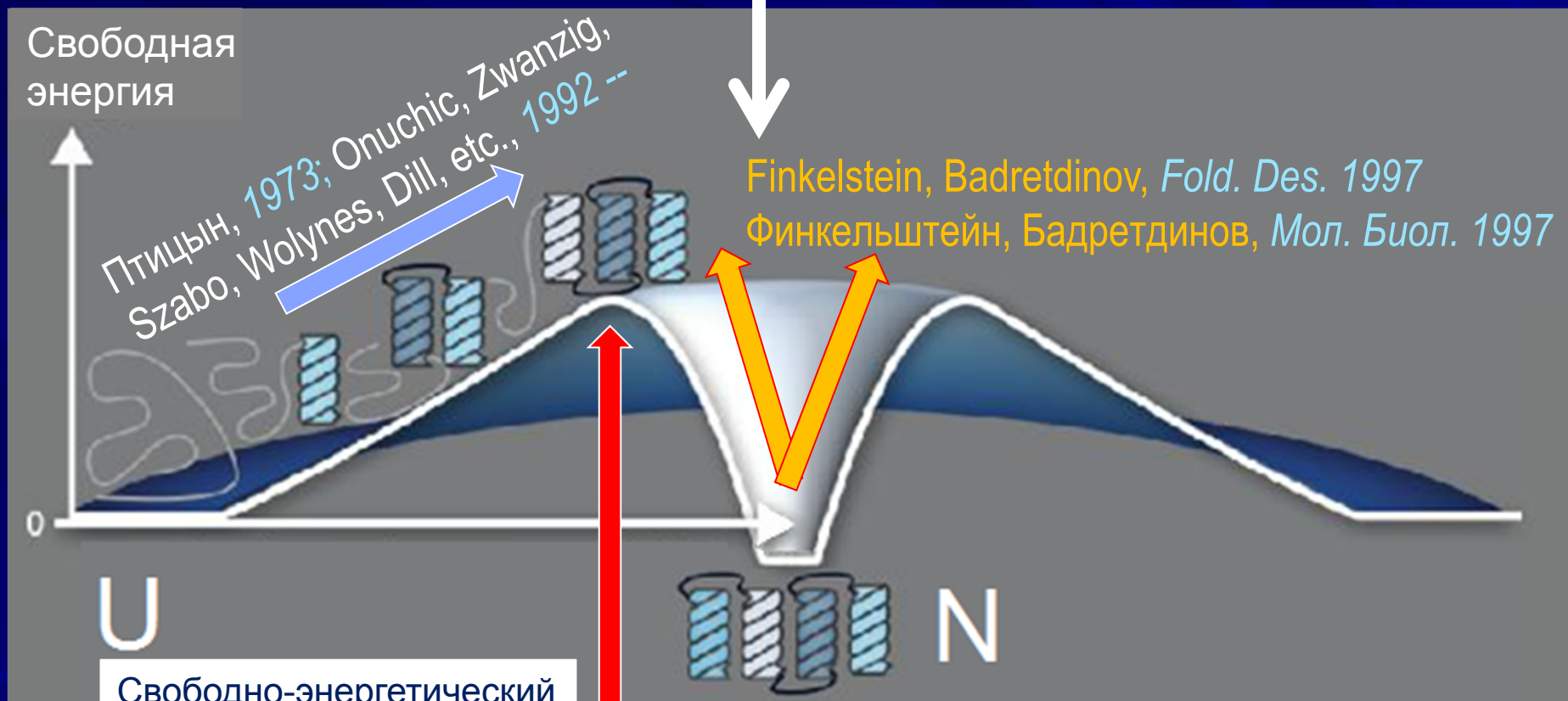


Свободно-энергетический «вулкан»

-TS↑: U → N
сторона

барьер \approx
перебор

Физика: «Принцип Детального Равновесия»:
Если **U** и **N** одинаково стабильны, то
скорости переходов **U** \rightarrow **N** и **N** \rightarrow **U** одинаковы



$-TS \uparrow$: **U** \rightarrow **N**
сторона

Барьер

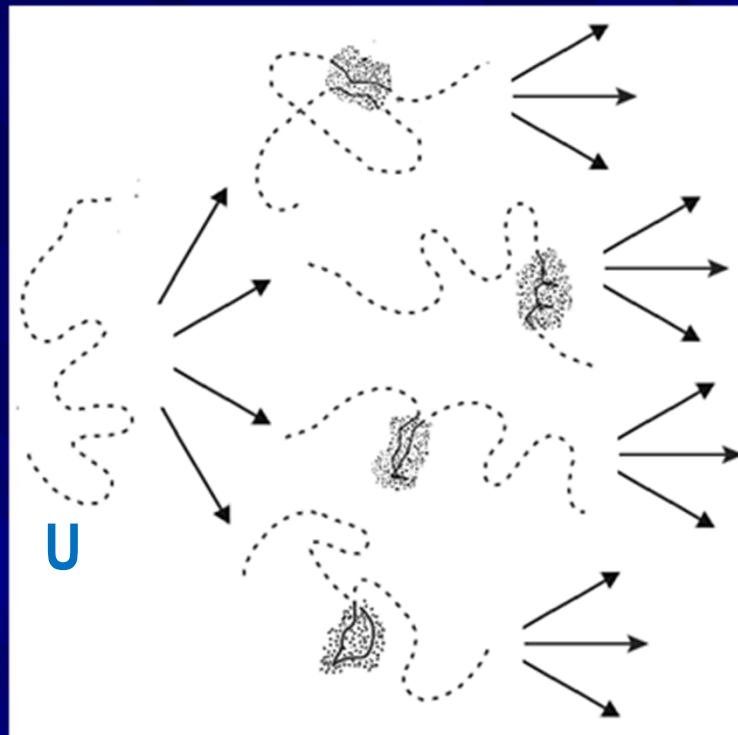
сторона
 $E \uparrow$: **N** \rightarrow **U**

«переходное состояние»

Чтобы оценить скорость сворачивания ($U \rightarrow N$) – проще рассмотреть переход $N \rightarrow U$, чем $U \rightarrow N$!

$U \rightarrow N$:

Сверни то –
не знаю что

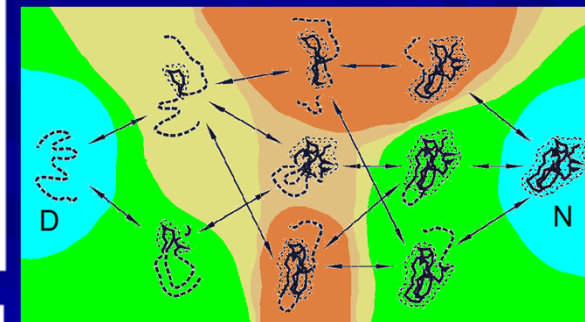
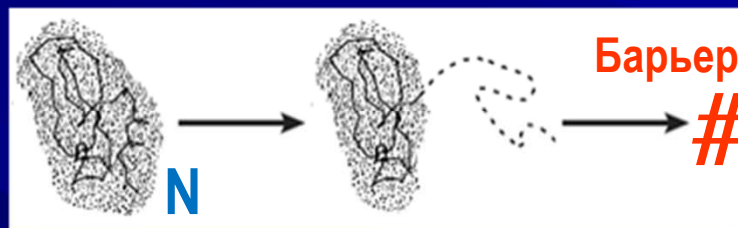


СЛОЖНО

$N \rightarrow U$:

Разверни ЭТО:

Оцени все



МНОГО ПРОЦЕ

Тоже несложно

быстрый переход идёт через *наименее нестабильные*, т.е. компактные промежуточные состояния

барьер
#

Finkelstein, Badretdinov, *Fold. Des.* 1997
Финкельштейн, Бадретдинов, *Мол. Биол.* 1997

Теория переходных состояний:

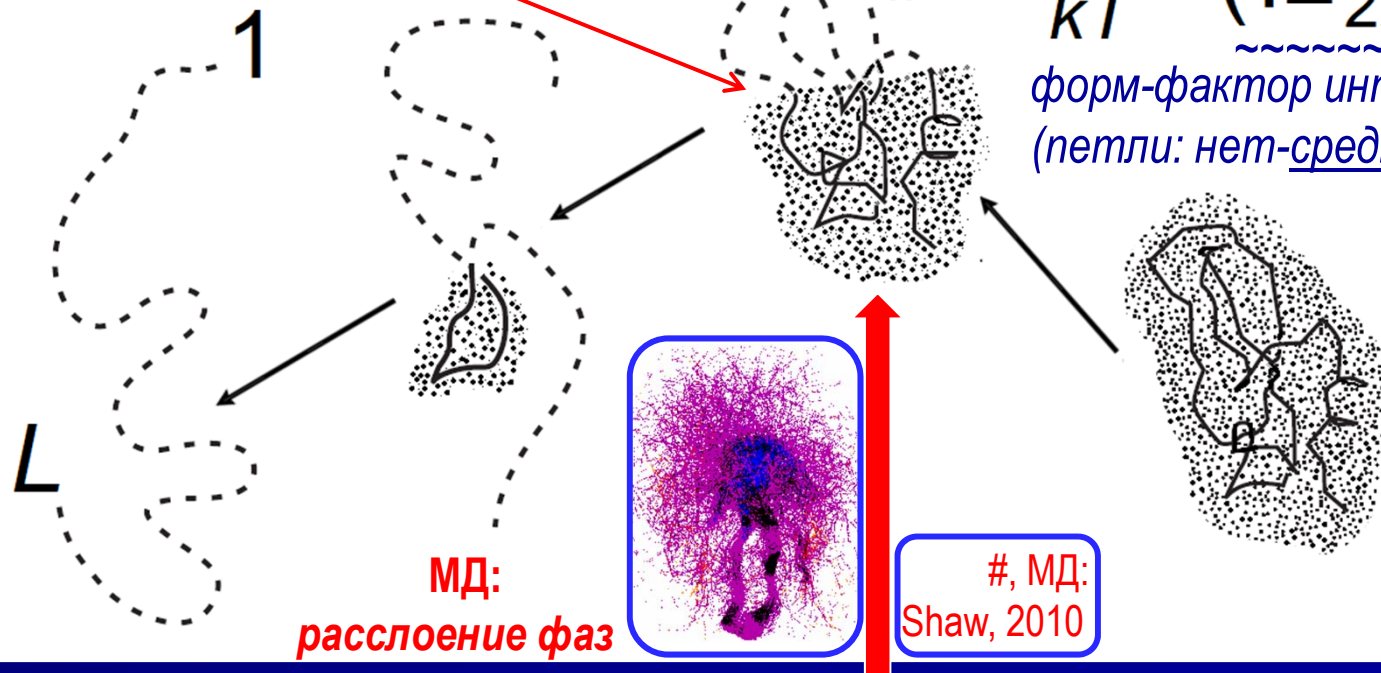
$$\text{скорость} = k_0 \cdot e^{-\frac{\Delta G^\#}{kT}}$$

расслоение фаз: интерфейс: $\sim L^{2/3}$ а.к.

В точке равновесия

$$\frac{\Delta G^\#}{kT} = (1 \pm \frac{1}{2}) L^{2/3}$$

форм-фактор интерфейса
(петли: нет-средне-много)



Узлы несущественны:
замедление – $v < \sim 2^{L/100}$

«переходное состояние»

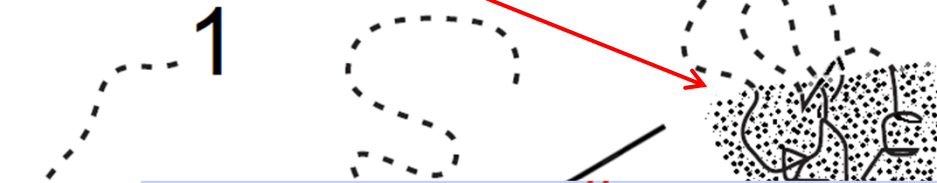
барьер
#

Finkelstein, Badretdinov, *Fold. Des.* 1997
Финкельштейн, Бадретдинов, *Мол. Биол.* 1997

Теория переходных состояний:

$$\text{скорость} = k_0 \cdot e^{-\frac{\Delta G^\ddagger}{kT}}$$

расслоение фаз: интерфейс: $\sim L^{2/3}$ а.к.



В точке равновесия:

$$\frac{\Delta G^\ddagger}{kT} = (1 \pm \frac{1}{2}) L^{2/3}$$

форм-фактор интерфейса
(петли: нет-средне-много)

Вклад в ΔG^\ddagger :

$\approx kT/2$ на 1 а.к. остаток интерфейса

L $\approx 6kT$ [$\approx kT (\frac{5}{2} \ln \frac{L/2}{L^{2/3}/6} - 1)$ при $L \sim 100$] на 1 петлю

число петель – от 0 до $L^{2/3}/6$

Узлы несущественны:
замедление – $v < \sim 2^{L/100}$

«переходное состояние»

барьер
#

Finkelstein, Badretdinov, *Fold. Des.* 1997

Финкельштейн, Бадретдинов, *Мол. Биол.* 1997

Теория переходных состояний:

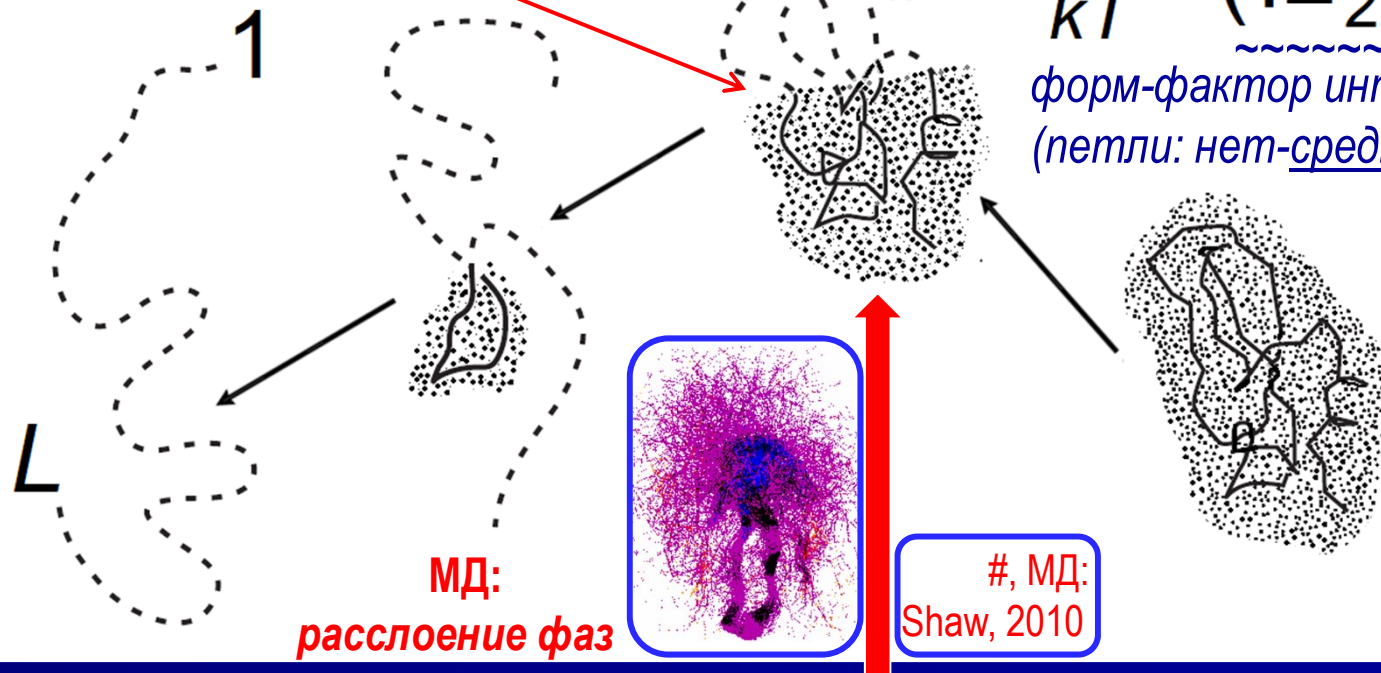
$$\text{скорость} = k_0 \cdot e^{-\frac{\Delta G^\#}{kT}}$$

расслоение фаз: интерфейс: $\sim L^{2/3}$ а.к.

В точке равновесия:

$$\frac{\Delta G^\#}{kT} = (1 \pm \frac{1}{2}) L^{2/3}$$

форм-фактор интерфейса
(петли: нет-средне-много)



Узлы несущественны:
замедление – $v < \sim 2^{L/100}$

«переходное состояние»

Опыт: $k_0 \sim 10^8 \text{c}^{-1}$
 скорость = $k_0 \cdot e^{-(1 \pm \frac{1}{2})L^{2/3}} \times e^{-0.4\Delta G/k_B T}$

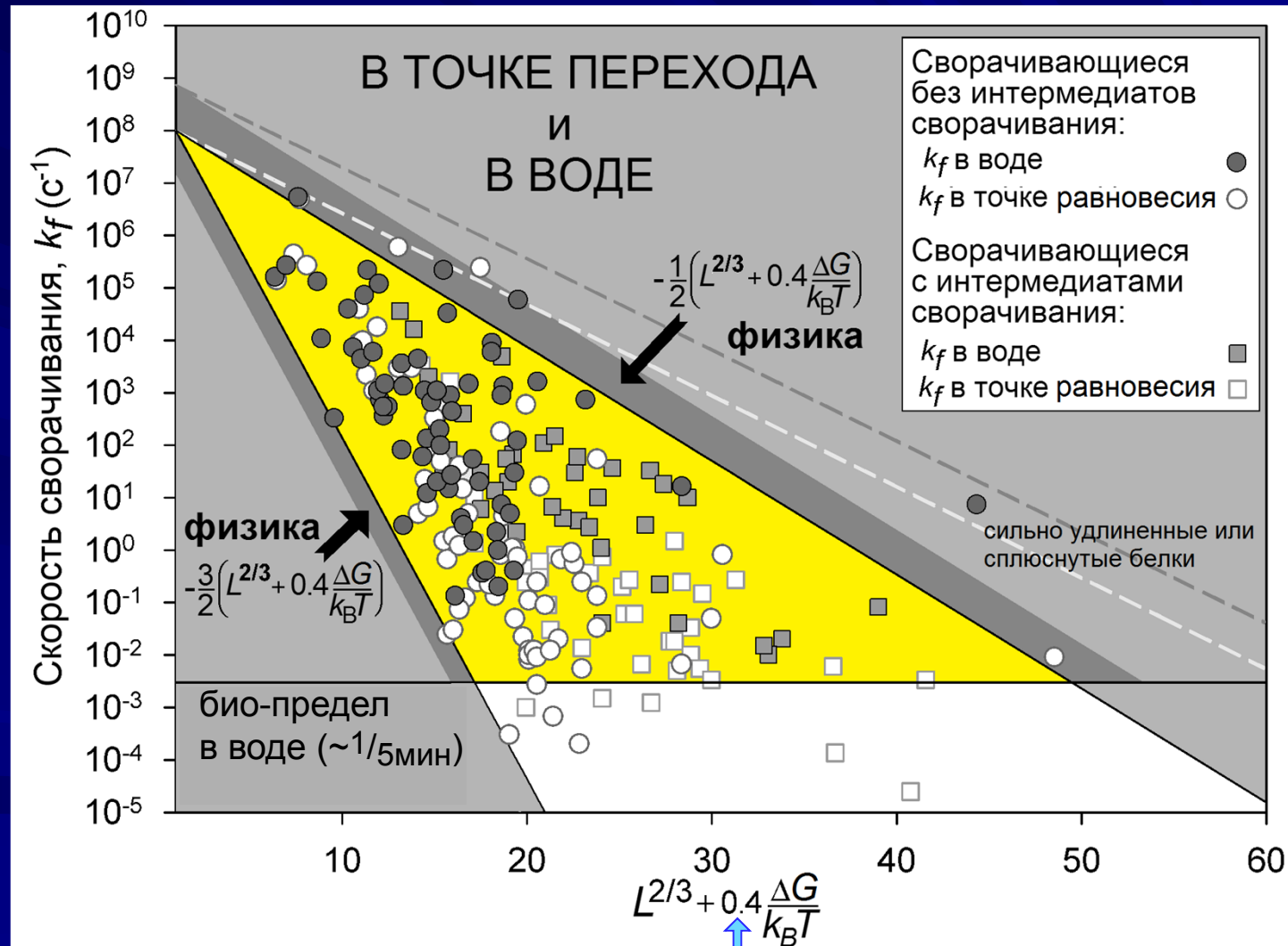
«В воде»: отклонение от точки равновесия

$$\Delta G = G_N - G_U$$

в точке равновесия

Garbuzinskiy, Ivankov, Bogatyreva, Finkelstein

“Golden triangle for folding rates of globular proteins.” *PNAS*, 2013



глобулярная часть переходного состояния охватывает чуть меньше половины белка

Однако наша теория еще не объяснила **парадокс Левинталя**
«в терминах Левинталя», т.е.

не оценила число подлежащих перебору конформаций
и не выяснила, почему перебор этих конформаций цепи
возможен хотя бы для небольших (~ 100 а.к. остатков) белков

Левинталь: полный перебор $\sim 3^L$ конформаций (на уровне а.к. остатков):
НЕВОЗМОЖЕН (даже для 50 а.к. остатков – $\sim 3^{50}$ структур ~ 10000 лет)

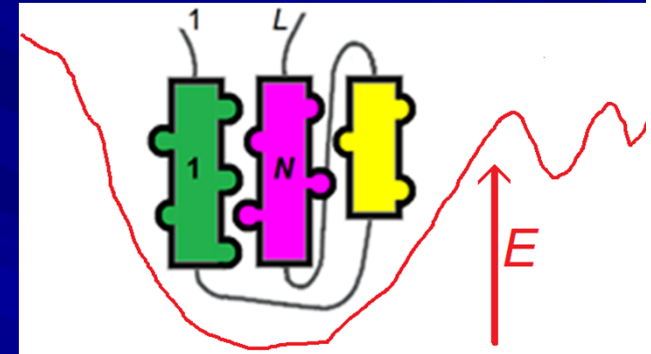
На уровне же формирования и упаковки
элементов вторичной структуры,

(аналог расслоения фаз)

полный перебор **ВОЗМОЖЕН** – перебору
подлежат лишь $\sim L^N$ конформаций

(при $L=70$ и $N=5$: перебор $\sim 10^9$ структур, что займет ~ 10 минут)

Основная причина уменьшения перебора (с $>3^L$ до $\sim L^N$): число N
структурных элементов в белке на порядок меньше числа а.к. остатков L

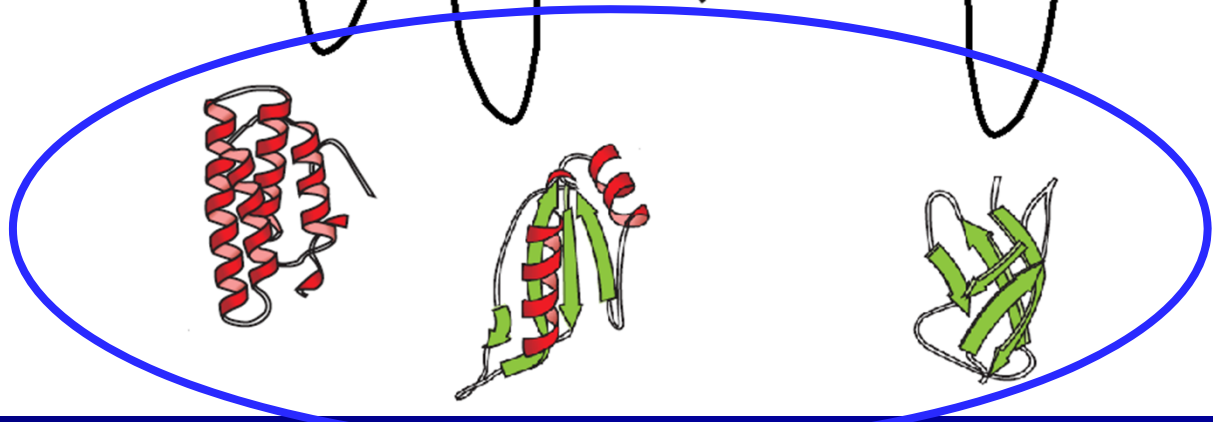
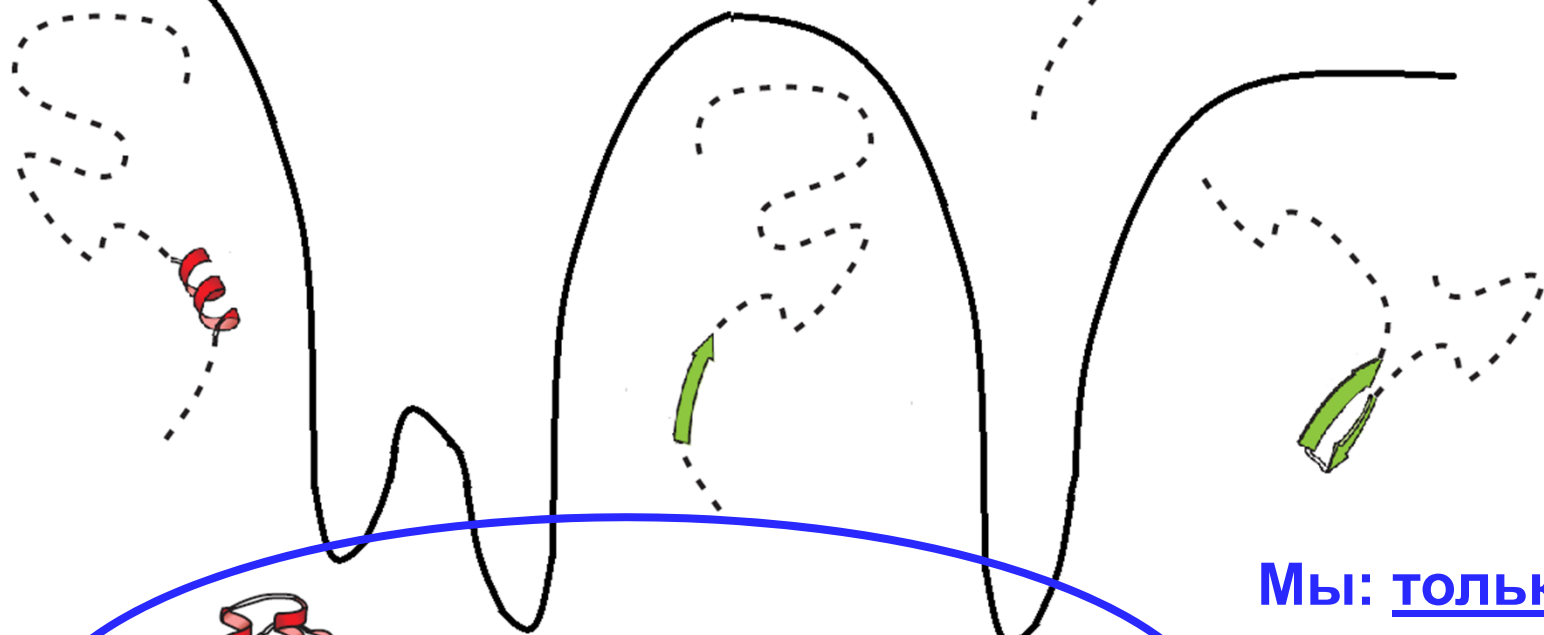


[Птицын, 1973]



Св. эн.

Левинталь: все структуры

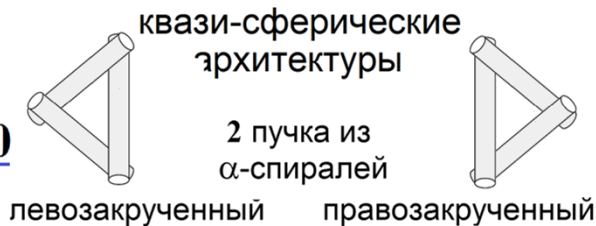


Мы: только
компактные
агрегаты
вторичных
структур

ПЕРЕБОР на уровне образования и укладки вторичных структур

(а) число наборов из N вторичных (α или β) структур: $\lesssim 2^N$

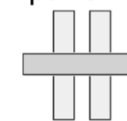
(б) архитектуры из N (тут: $N=3$) структурных элементов: немного, $\lesssim 10$



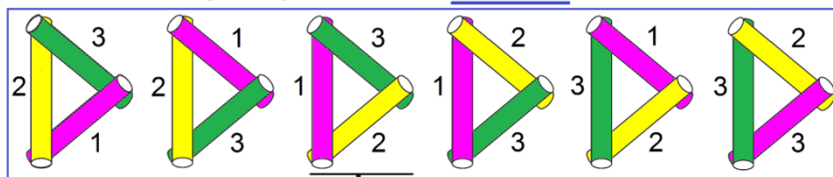
продольные архитектуры
2 продольные укладки из α -спиралей или β -тяжей



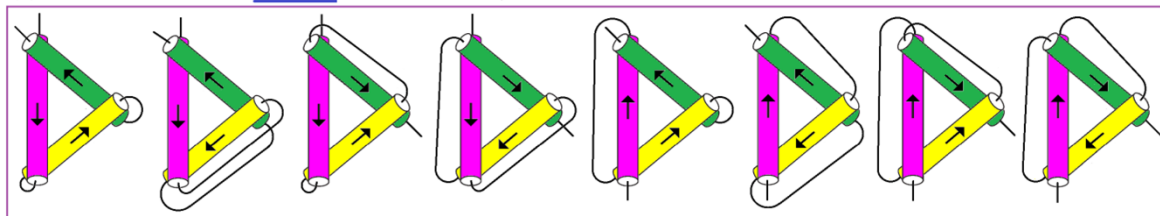
ортогональные архитектуры
1 ортогональная укладка из α -спиралей или β -тяжей



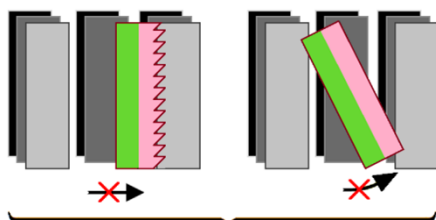
(в) упаковки: $N! = N \times (N-1) \times \dots \times 2 \times 1 \sim (N/e)^N$



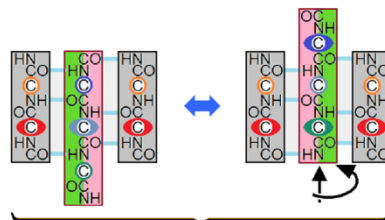
(г) топологии: $\lesssim 2^N$



Узлы несущественны:
из-за них рост числа топологий –
в $\lesssim 2^{L/100}$ раз



(д) поперечные сдвиги и наклоны: нет



(е) увязанные сдвиги и повороты

сдвиг/повороты:

L а.к. остатков, N втор. стр.: $\sim (L/N)^N$

ИТОГО: $\sim 2^N \times (N/e)^N \times 2^N \times (L/N)^N \sim L^N$

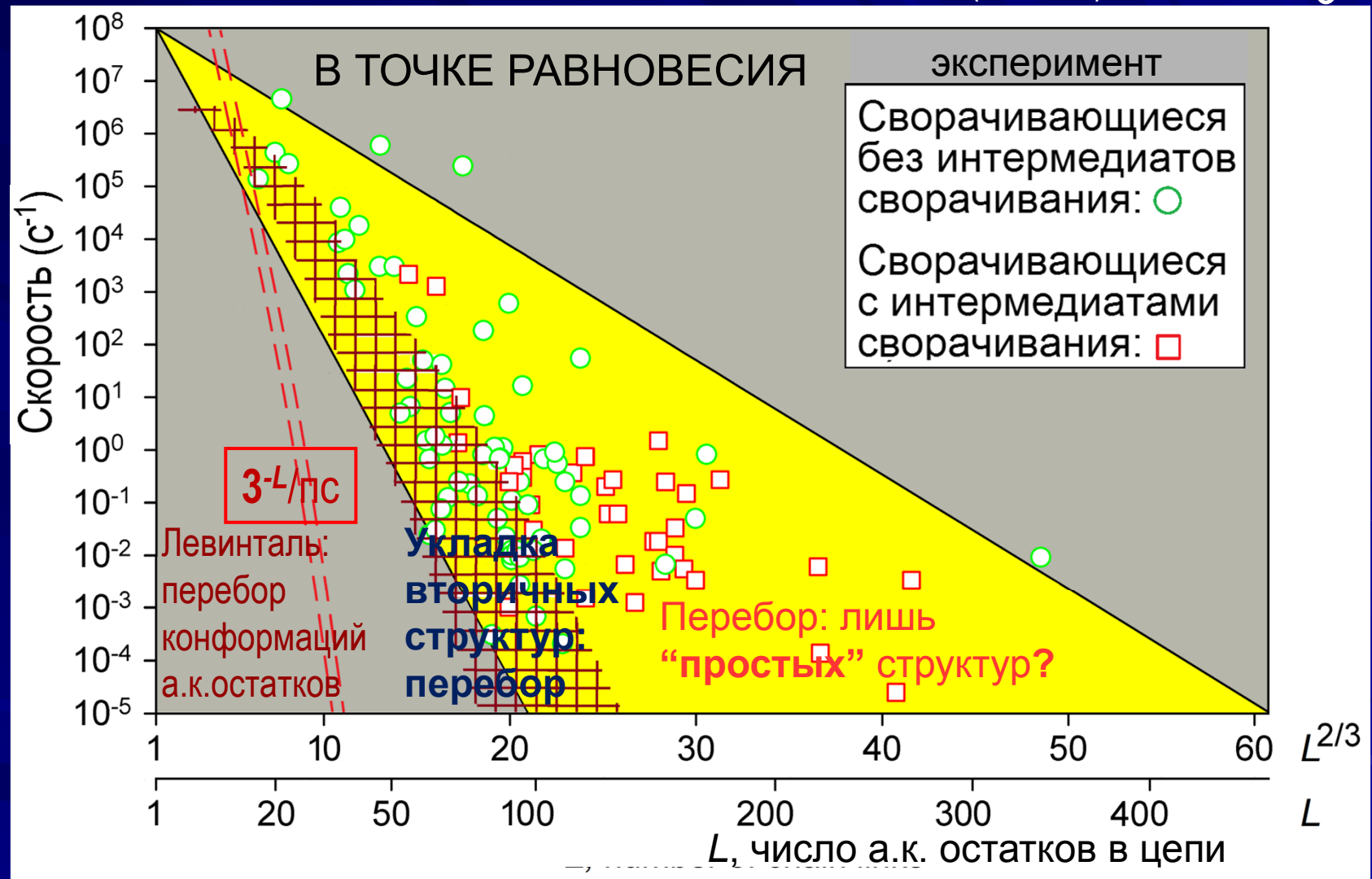
Перебор: $\sim L^N$: L остатков в цепи, $N \sim L / (10 \div 20)$ вторичных структур $\Rightarrow N \sim L / 3L^{1/3}$

статистика

теория

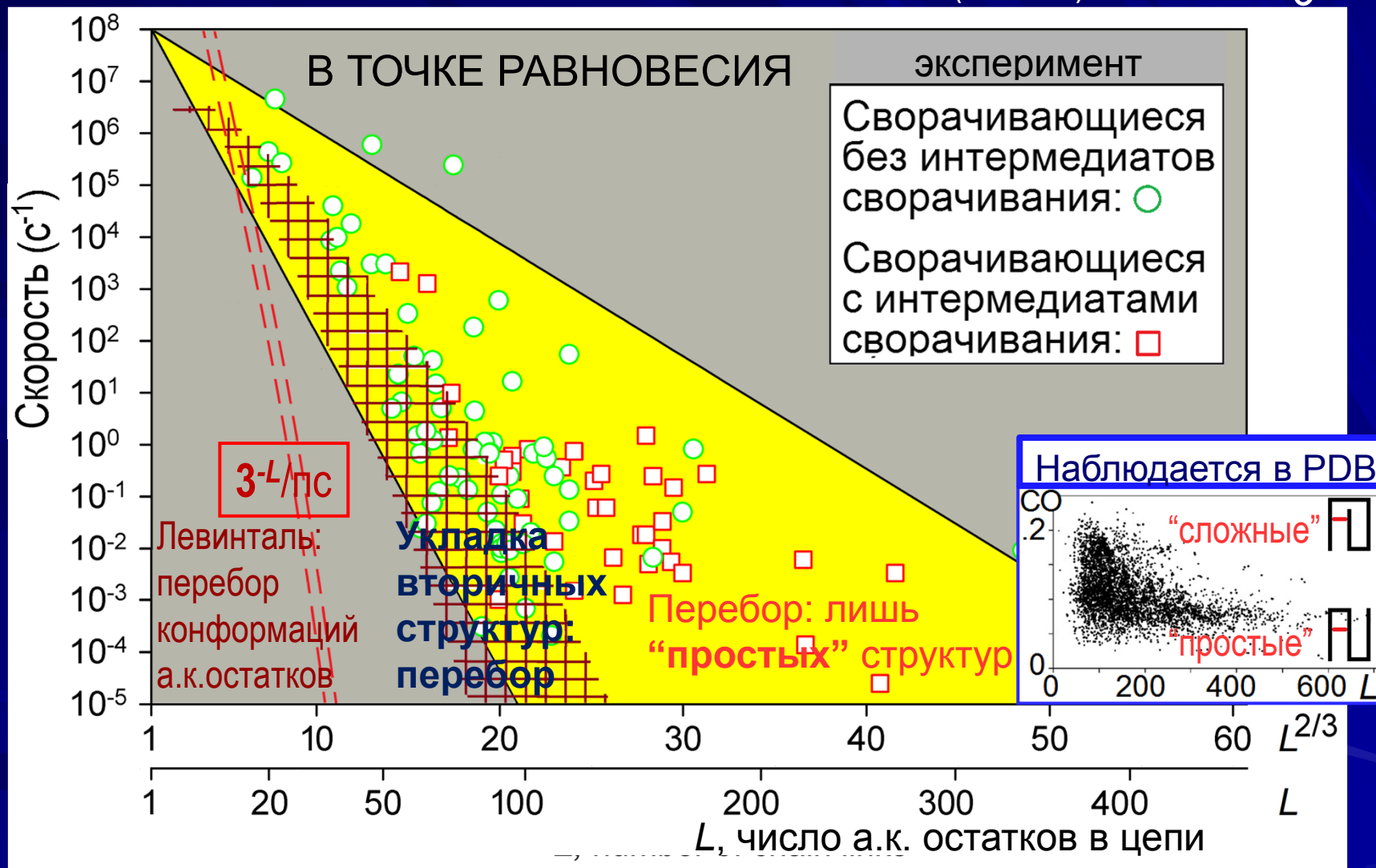
Образование и укладка вторичных структур: перебор:

время: $\sim 1\mu\text{s} \times L^N \Rightarrow$ скорость $\sim 1/(1\mu\text{s} \times L^N)$; $N \sim L/(10 \div 20) \Rightarrow \sim L^{2/3}/3$



Образование и укладка вторичных структур: перебор:

время: $\sim 1\mu\text{s} \times L^N \Rightarrow$ скорость $\sim 1/(1\mu\text{s} \times L^N)$; $N \sim L/(10 \div 20) \Rightarrow \sim L^{2/3}/3$



Заключение:

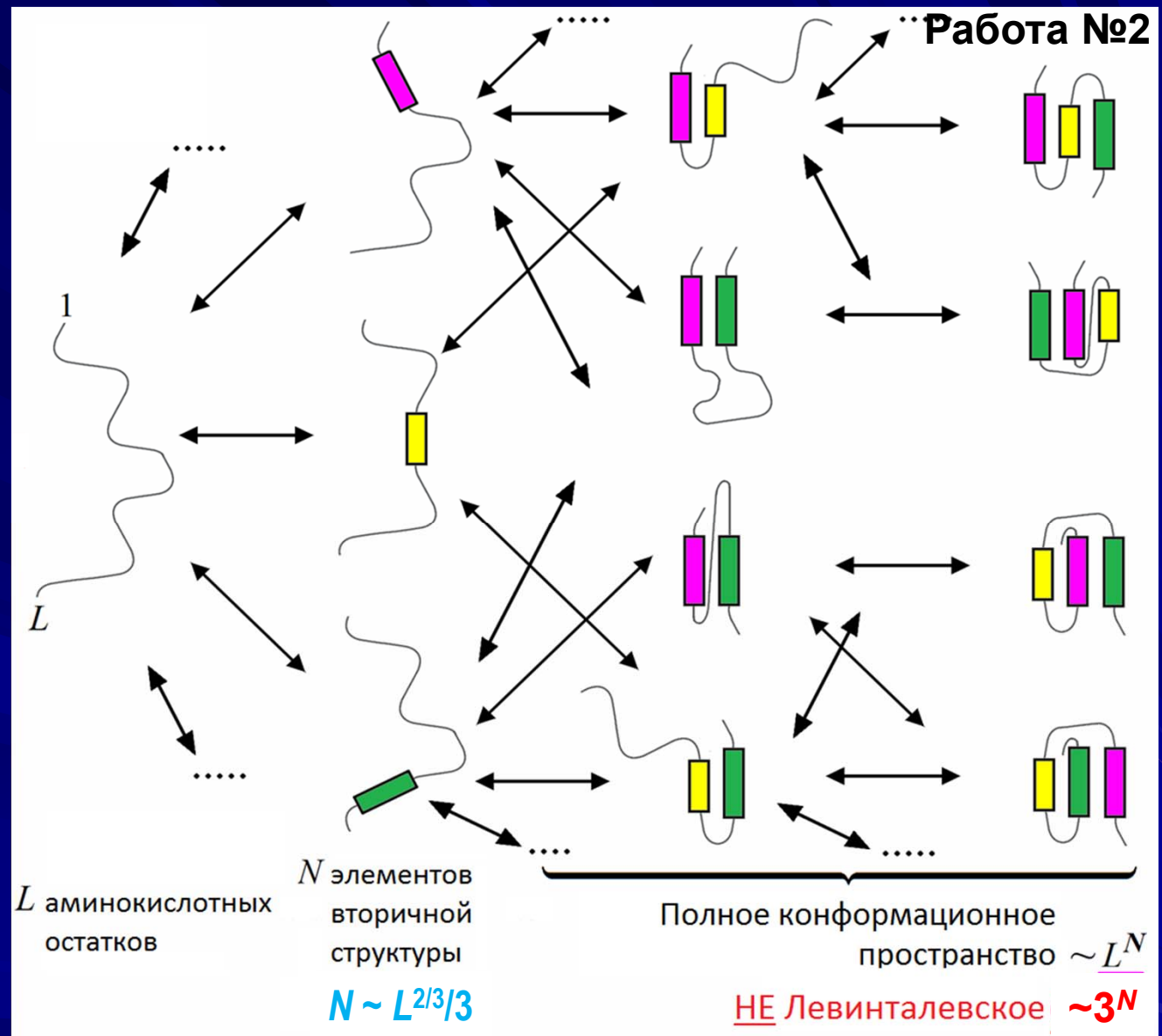
Полный перебор конформаций на уровне а.к. остатков: **НЕВОЗМОЖЕН**

Полный перебор конформаций на уровне элементов вторичной структуры: **ВОЗМОЖЕН** при $L < 100$

При $L > 100$: Перебор только «простых» структур

Парадокс Левинталя решён на уровне комбинаторики, даже без энергетики

Энергетические «воронки» и кинетическая отбраковка «слишком запутанных» структур, конечно, ещё более ускоряют сворачивание, - что особенно существенно для длинных белковых цепей



Заключение:

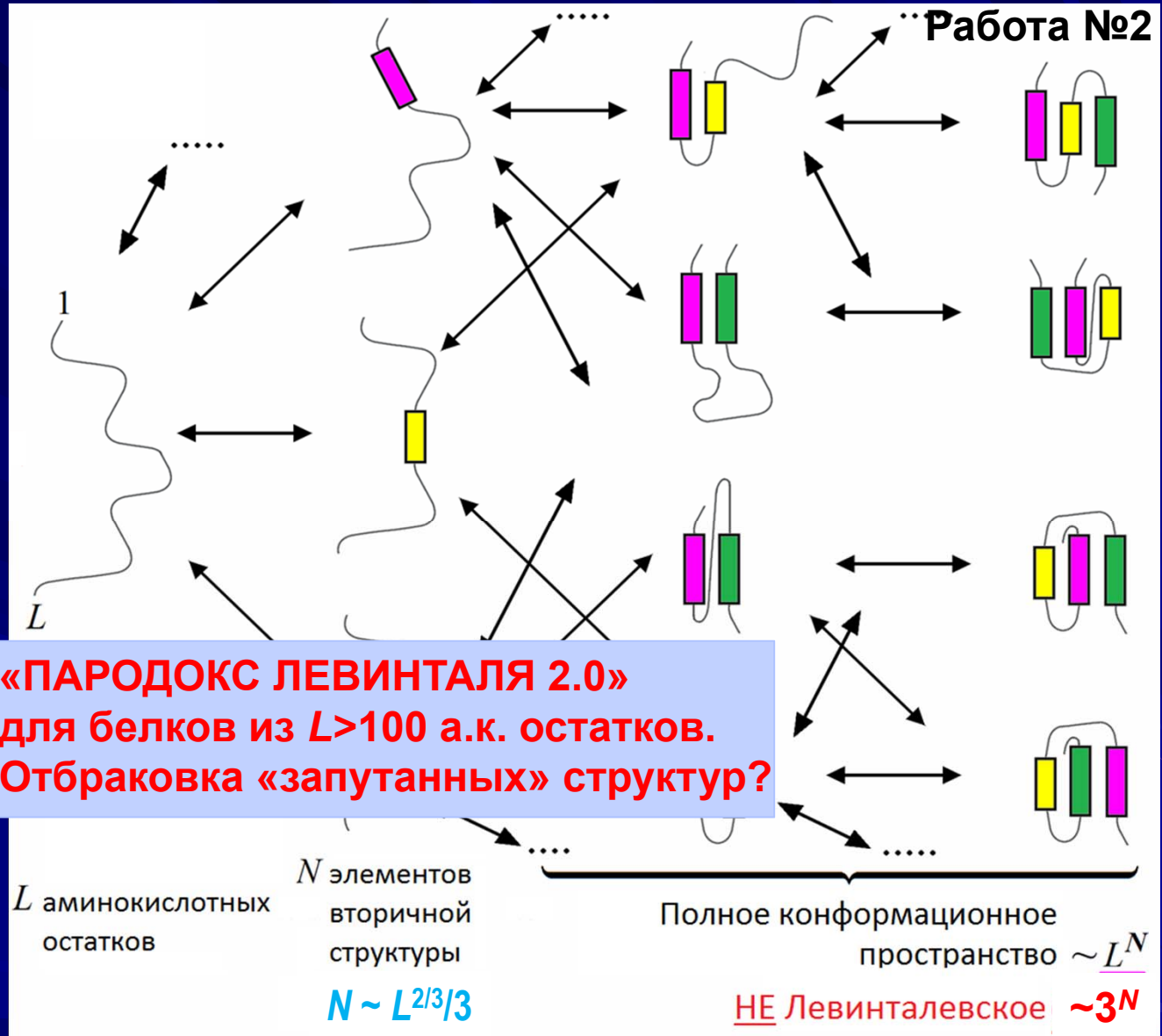
Полный перебор конформаций на уровне а.к. остатков: **НЕВОЗМОЖЕН**

Полный перебор конформаций на уровне элементов вторичной структуры: **ВОЗМОЖЕН** при $L < 100$

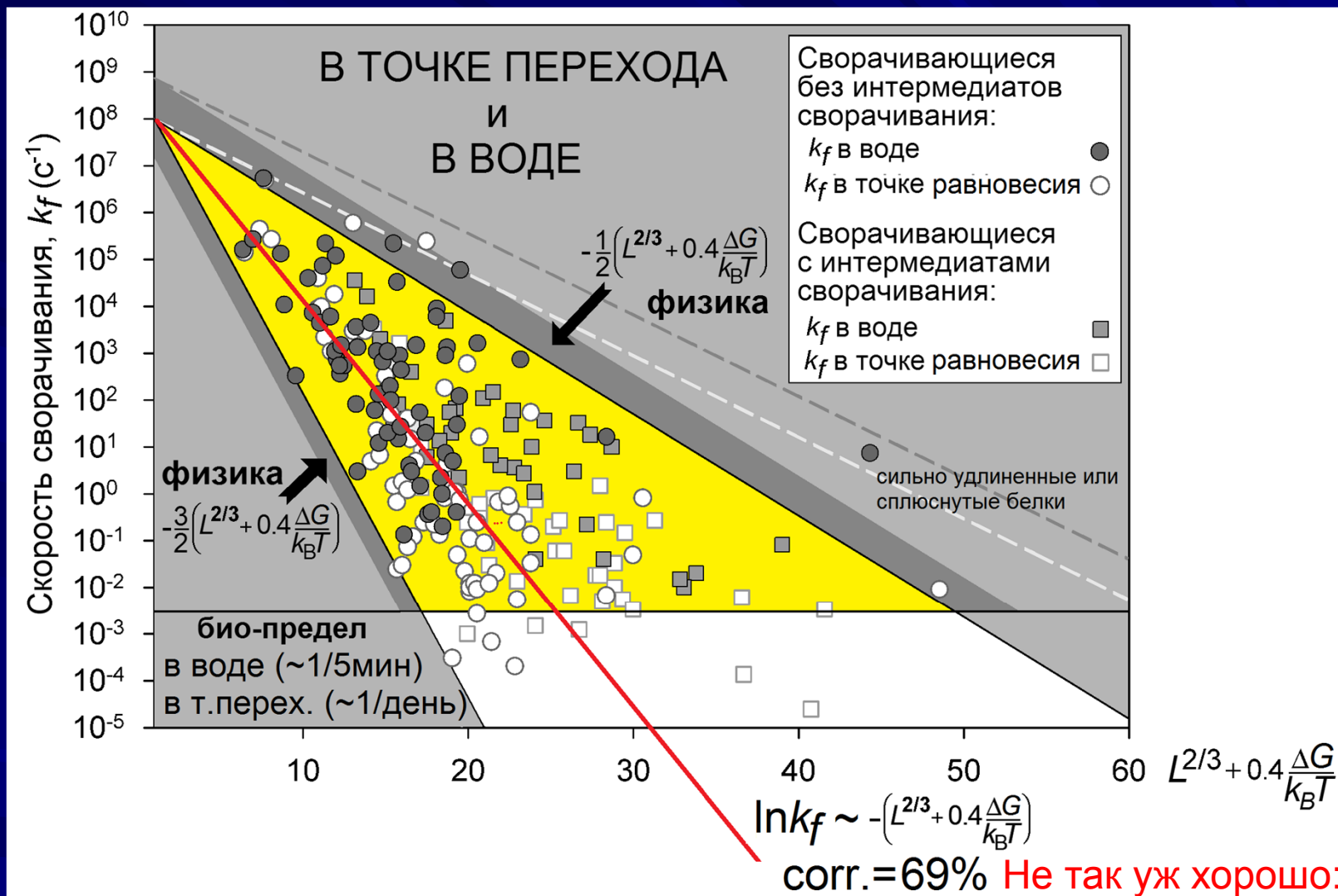
При $L > 100$: Перебор только «простых» структур

Парадокс Левинтала решён на уровне комбинаторики, даже без энергетики

Энергетические «воронки» и кинетическая отбраковка «слишком запутанных» структур, конечно, ещё более ускоряют сворачивание, - что особенно существенно для длинных белковых цепей



КОРРЕЛЯЦИЯ ПРЕДСКАЗАННЫХ (ФИЗИКА) СКОРОСТЕЙ С ОПЫТОМ



Finkelstein, Bogatyreva, Garbuzinskiy
FEBS Lett., 2013

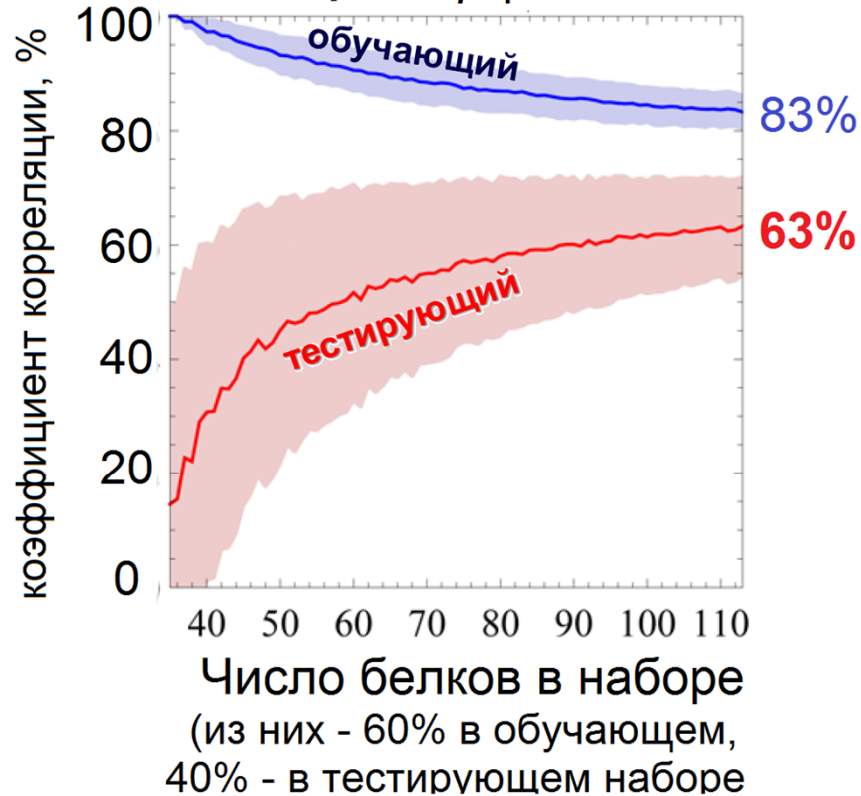
биоинформатический
 метод (напр., Huang & Tian
Proteins, 2006) дает 82%

СРАВНЕНИЕ:

БИОИНФОРМАТИКА

(20 оптимизируемых параметров Ω)

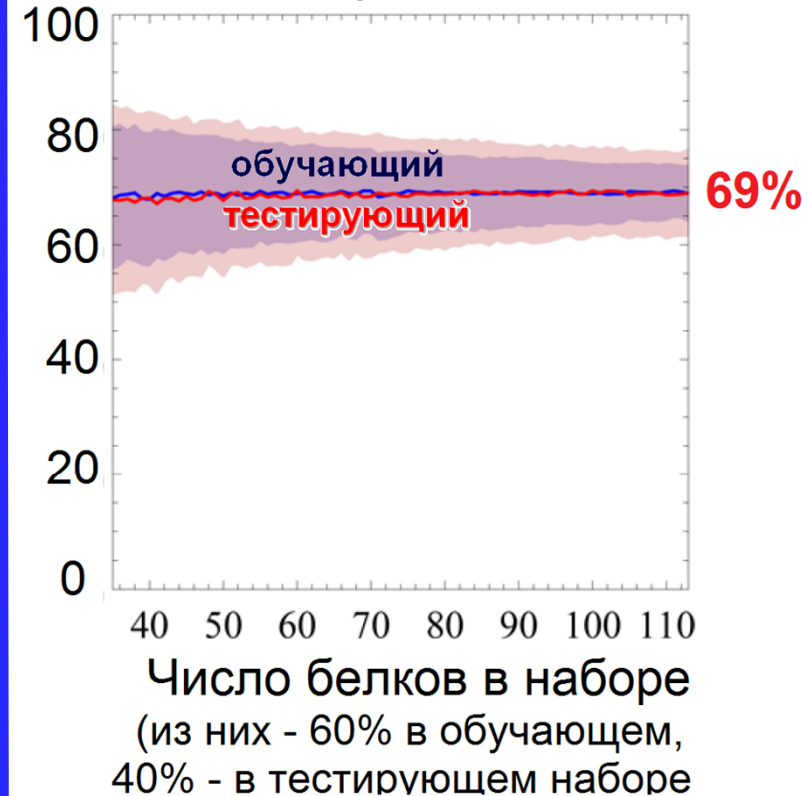
$$\ln k_f \sim \sum_{j=1}^{L_{PR}} \Omega(a_j)$$



ФИЗИКА

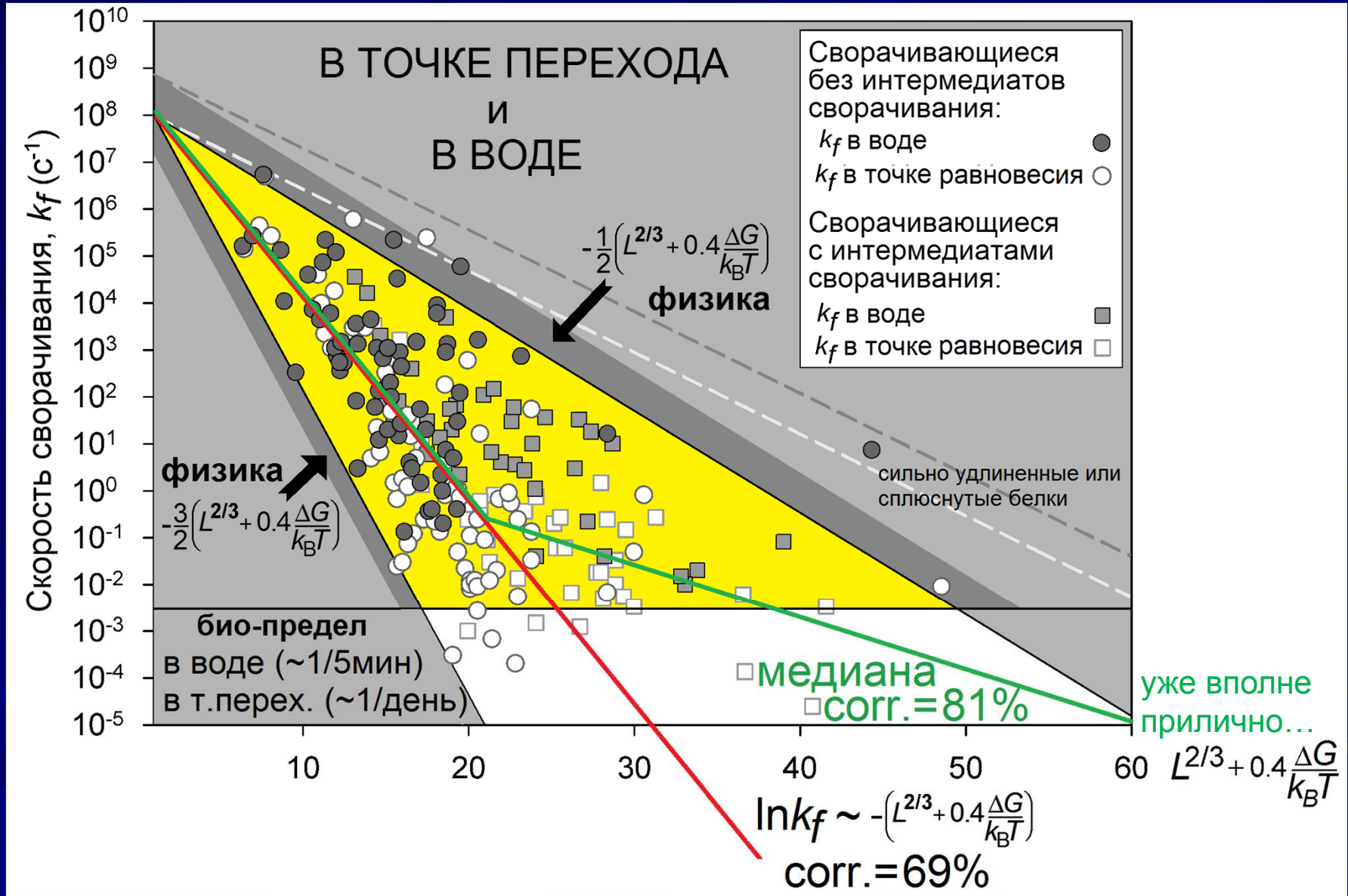
(НЕТ оптимизируемых параметров)

$$\ln k_f \sim -L^{2/3}$$



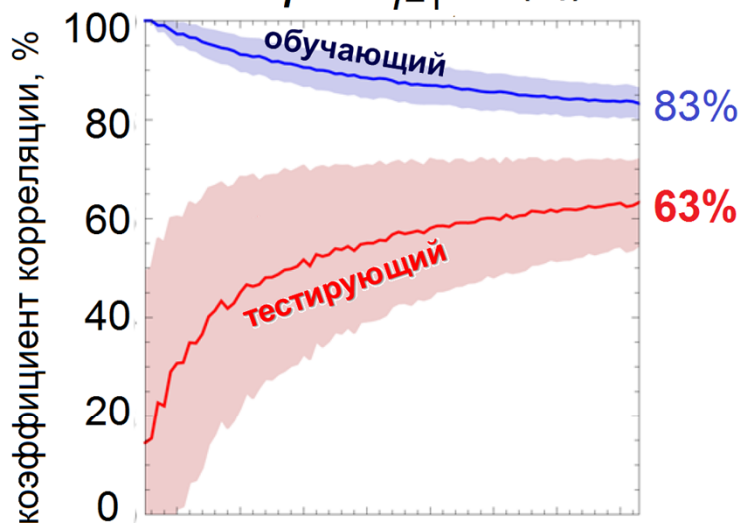
Существующие биоинформатические данные еще не позволяют получить на тестовой выборке то же качество предсказаний, что на обучающей

КОРРЕЛЯЦИЯ ПРЕДСКАЗАННЫХ (ФИЗИКА+БИО-ПРЕДЕЛ) СКОРОСТЕЙ С ОПЫТОМ



БИОИНФОРМАТИКА
(20 оптимизируемых параметров)

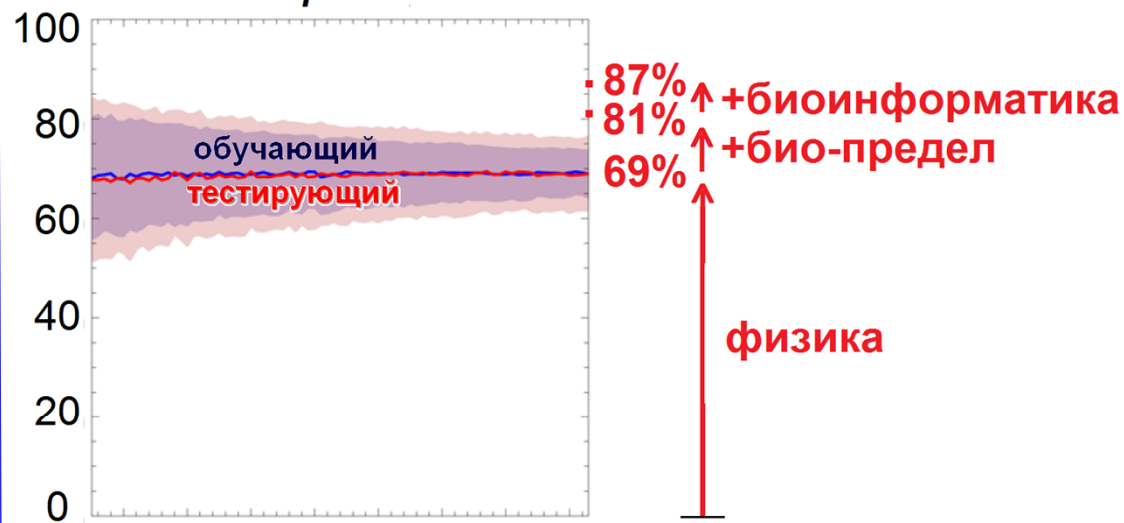
$$\ln k_f \sim \sum_{i=1}^{L_{PR}} \Omega(a_i)$$



Число белков в наборе
(из них - 60% в обучающем,
40% - в тестирующем наборе)

ФИЗИКА + БИО-ПРЕДЕЛ
(НЕТ оптимизируемых параметров)

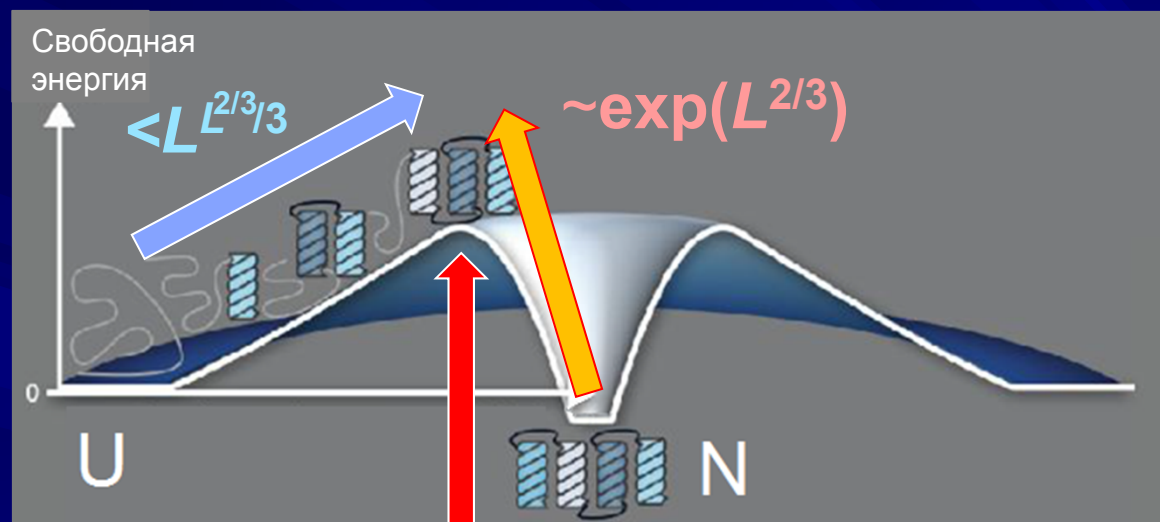
$$\ln k_f \sim -L^{2/3}$$



Число белков в наборе
(из них - 60% в обучающем,
40% - в тестирующем наборе)

Биоинформатика хороша для
уточнения базовой, основанной
на понимании сути дела оценки

Комплементарные точки зрения на проблему самоорганизации белка



Благодарю за внимание!

60 лет самоорганизации белков

А. В. Финкельштейн

Института белка РАН; Пущино, Московская область

Благодарности:

- О.Б. Птицыну, П.Л. Привалову, Е.И. Шахновичу, Ф.М. Гутину – за науку и многочисленные дискуссии
- Моим соавторам: А.Я. Бадретдинову, О.В. Галзитской, Д.Н. Иванкову, Н.С. Богатырёвой, С.А. Гарбузинскому
- Анонимным рецензентам наших статей
- МКБ, РФФИ, ННМИ – за поддержку 1-й части нашей работы
- Российскому Научному Фонду – за поддержку 2-й части нашей работы

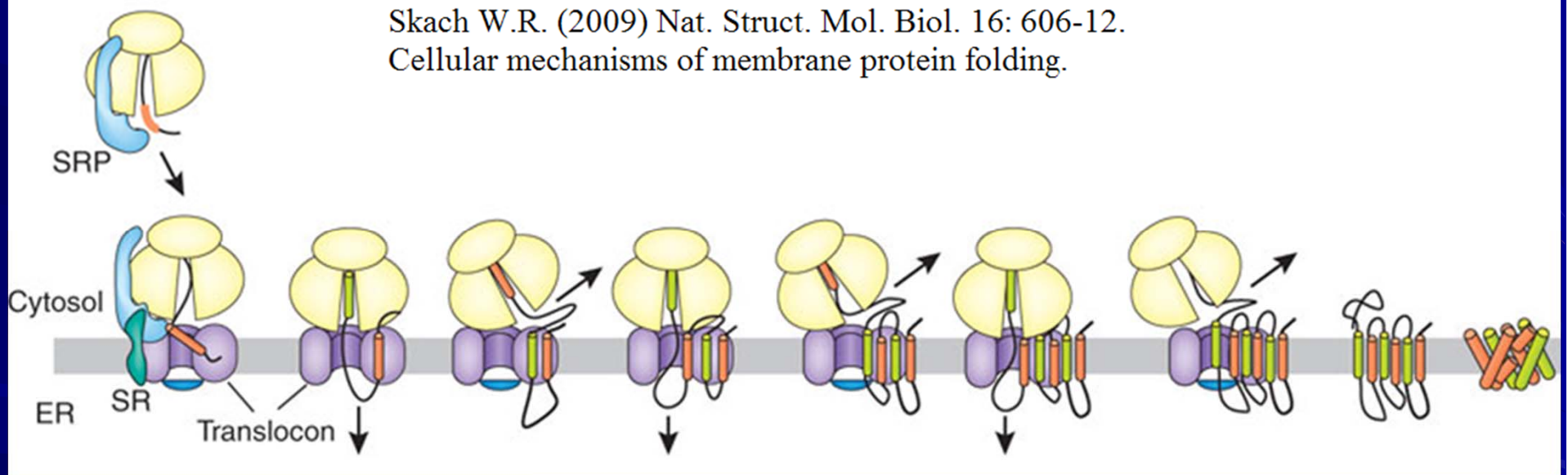
дополнения

Структурообразование

не-водорастворимых белков

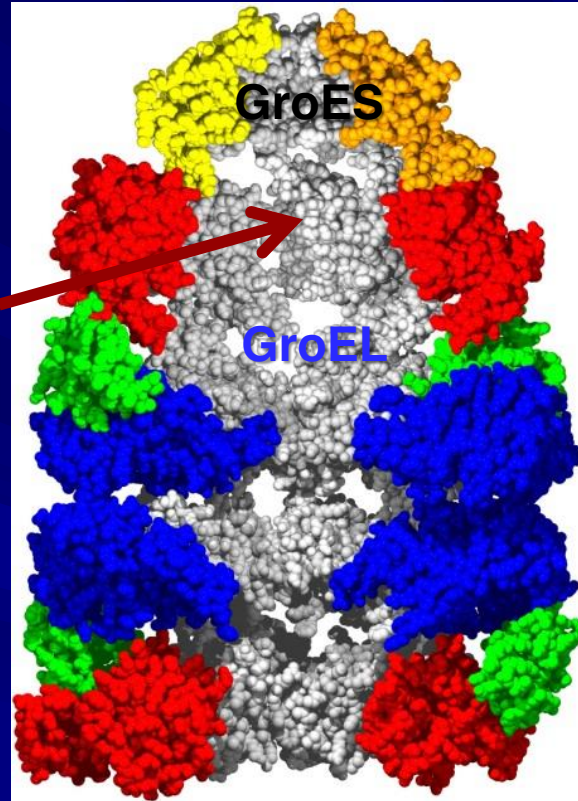
Мембранный белок *in vivo*

Сворачивание ассистируется “направляющими факторами” –
своеобразными шаперонами



Шаперон GroEL/ES

для водорастворимых белков



«Активная роль»?
– НЕТ

**Сворачивание: внутри
«ячейки Анфинсена»?»**

Ellis R.J. 2003
Curr. Biol. 13:R881-3

Неспецифичность:
GroEL/ES сворачивает
белок как из L-, так и из
D-аминокислот:

**“... двоякая активность
шаперона ...”**

Weinstock, Jacobsen, Kay, 2014,
PNAS 111, 11679-84

- ВНЕ!

**Пассивная и даже
сверх-пассивная роль
– GroEL/ES только
снижает концентрацию
еще свернутых белков
в растворе.**

Marchenkov & Semisotnov,
2009, *Int. J. Mol. Sci.*, 10:
2066-83