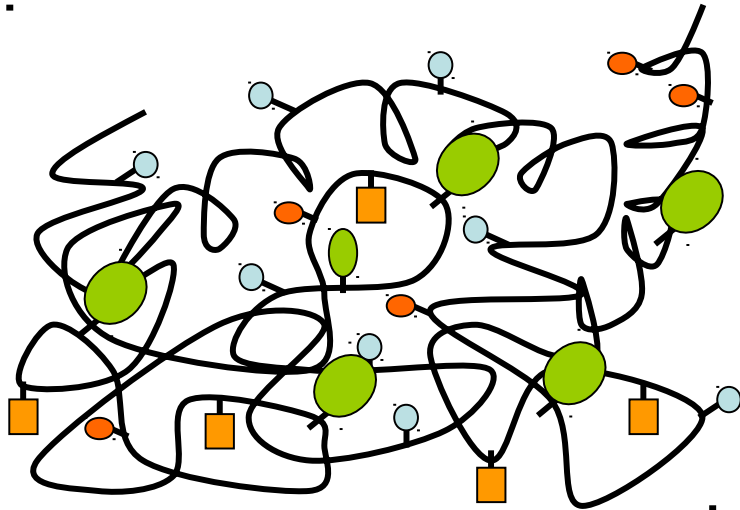


# O paradoxo do enovelamento de proteínas: um modelo simples



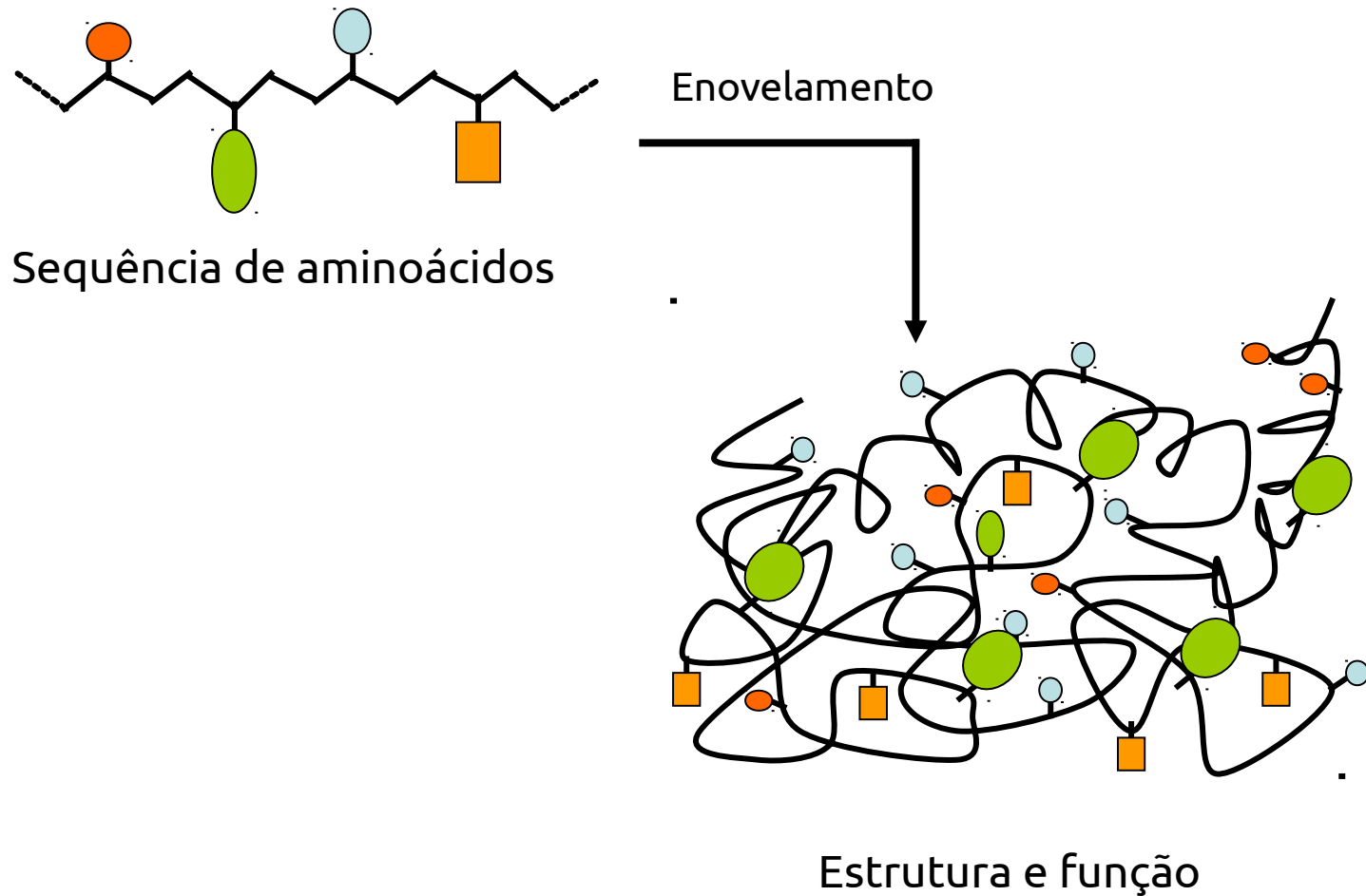
Leandro Martínez  
Instituto de Química  
UNICAMP

Seminários de Extensão do IMECC

26 de Outubro de 2016

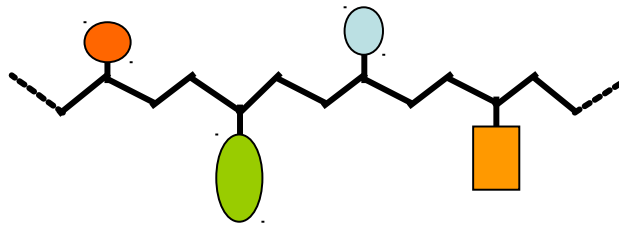
## O problema

Proteínas são sequências de aminoácidos que se enovelam tridimensionalmente formando estruturas específicas e funcionais



## O problema

### Número de conformações

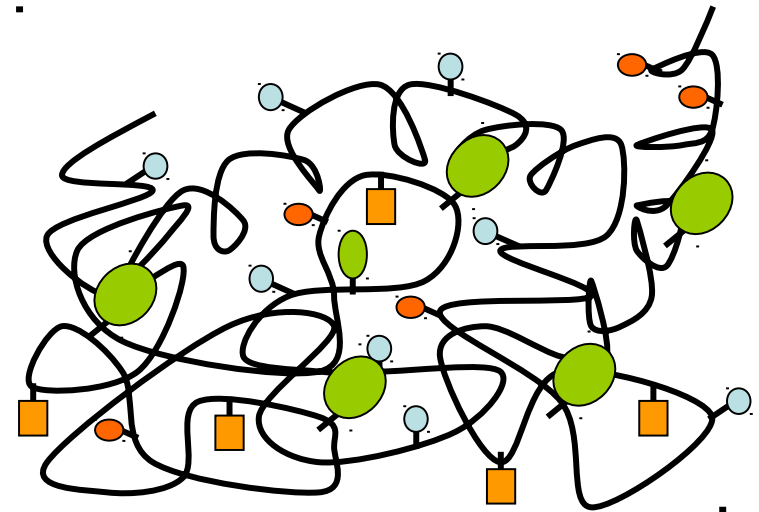


Sequência de aminoácidos

Se cada aminoácido pode assumir 3 conformações distintas, uma proteína com **300** aminoácidos tem  **$3^{300}$  estruturas possíveis.**

Se a busca fosse aleatória, à taxa de 1 estrutura a cada  $10^{-6}$  segundos

a proteína demoraria em torno de  
 $\sim 10^{137}$  s =  $10^{129}$  anos  
 para alcançar à estrutura nativa.



**Como as proteínas atingem a estrutura nativa, na vida real, em frações de segundos ou minutos?**

**“Paradoxo de Levinthal”**

C. Levinthal, 1969



## Um modelo simples de enovelamento

CCIIICICICICCCCCIIIIIIIIIIICCIICICICIICIIIIICCCCCCCCCC...



$10^{90}$  estruturas possíveis.



CC...

Através de que MECANISMO de busca a proteína pode atingir o estado nativo em um tempo compatível com a realidade?

## Um modelo simples de enovelamento

Proposição de um MECANISMO de enovelamento:

1. A estrutura começa com uma conformação aleatória:

CCIIICICICICCCCIIIIIIIIIICCIICICICIICIIIIICCCCCCCCCC...

2. A cada  $\tau$  unidades de tempo ( $\tau \sim 1 \text{ ns}^*$ ) cada aminoácido sofre uma perturbação estrutural aleatória.

3. A perturbação leva a o estado C ou I de cada aminoácido com igual probabilidade.

CCIIICICICICCCCIIIIIIIIIICCIICICICIICIIIIICCCCCCCCCC...



ICCIICICICICCCCIICCIIIIIICCIIIIIICIIIIICCCCCCCCCC...



CCIIICICICICCCCIIIIIIIIIICCIICICICIICIIIIICCCCCCCCCC...

## Um modelo simples de enovelamento

### Probabilidade de enovelamento

t= 1 ns

CCCIICCCICIIIIIIICCIIIIIIIII...

1 ns

t= 2 ns

CCCIICCCIICCIICCCCCCIIICCC.....

1 ns

t= 3 ns

CIIIIIIIICCIICCIIIICCCCIICI.....

1 ns

t= 4 ns

CCCCCCCCCCCCCCCCCCCCCCCC.....

Quanto tempo leva a estrutura para passar pelo estado enovelado?

- N resíduos
- Perturbações independentes

Probabilidade de enovelamento instantânea:

$$P_{\text{enovelamento}} = 0,5^N$$

Probabilidade de permanecer desenovelada:

$$P_{\text{não-enovelamento}} = 1 - 0,5^N$$

Se  $t$  unidades de tempo se passaram, a estrutura tem probabilidade

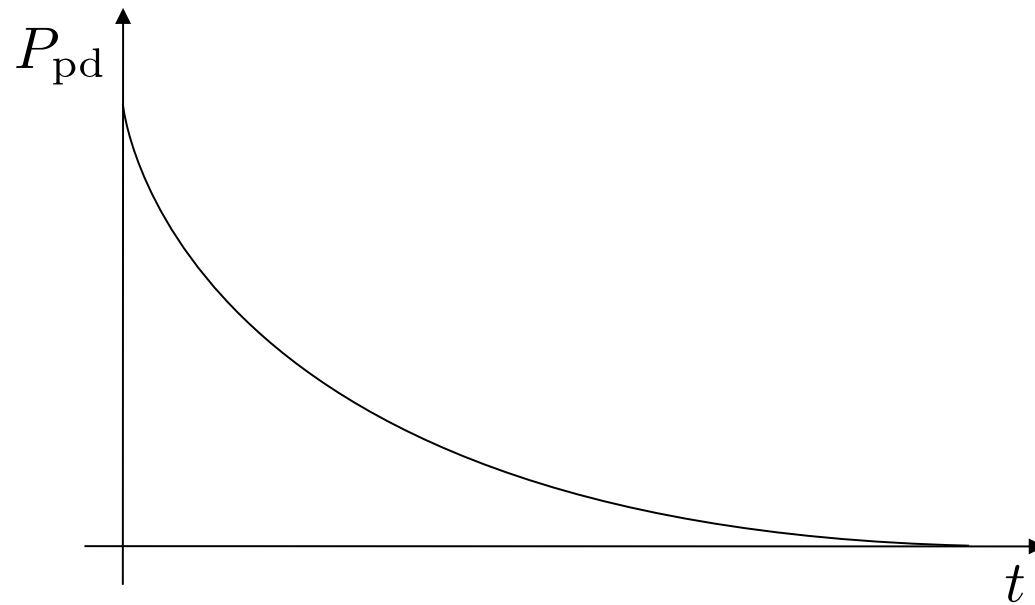
$$P_{\text{permanecer desenovelada}}(t) = (1 - 0,5^N)^{t/\tau}$$

de **não** ter passado pelo estado enovelado.

## Um modelo simples de envelhecimento

### Probabilidade de envelhecimento

$$P_{\text{permanecer desnovelada}}(t) = (1 - 0,5^N)^{t/\tau}$$



$$P_{\text{envelar}}(t) = 1 - P_{\text{permanecer desnovelada}}(t)$$

$$P_{\text{envelar}}(t) = 1 - (1 - 0,5^N)^{t/\tau}$$



## Um modelo simples de enovelamento

Qual o tempo necessário que para que 50% das proteínas atinjam a conformação correta?

$$P_{\text{enovelar}}(t) = 1 - (1 - 0,5^N)^{t/\tau}$$

$$P_{\text{enovelar}}(t) = 0,5$$

$$1 - (1 - 0,5^N)^{t/\tau} = 0,5$$

$$t_{1/2} = \frac{\ln 0,5}{\ln(1 - 0,5^N)} \tau$$

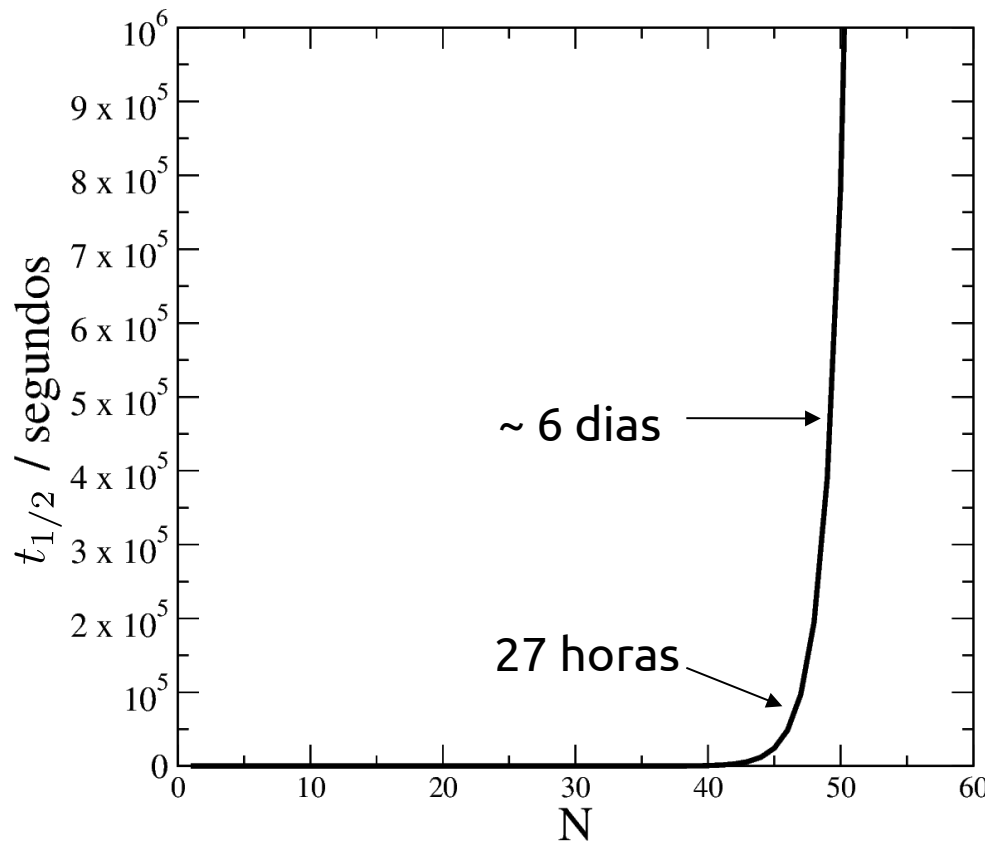
## Um modelo simples de enovelamento

Qual o tempo necessário que para que 50% das proteínas atinjam a conformação correta?

$$t_{1/2} = \frac{\ln 0,5}{\ln(1 - 0,5^N)} \tau$$

**Exemplo:**

$N = 30 ; \tau = 1 \text{ ns} : t_{1/2} = 1 \text{ segundo}$



Dependência com N:

Pequenas proteínas de ~40 aminoácidos podem se enovelar por este mecanismo.

Proteínas maiores (> 50 aminoácidos), **não**.

Mas isto era para probabilidades de 50% entre a correta e a incorreta. E se a correta for mais provável que a incorreta?

## Um modelo simples de enovelamento

### Probabilidade de enovelamento

t= 1 ns

CCCIICCCICIIIIIIICCIIIIIIII...

1 ns

t= 2 ns

CCCIICCCIICCIICCCCCCIIICCC.....

1 ns

t= 3 ns

CIIIIIIIICCIICCIIIICCCCIICI.....

1 ns

t= 4 ns

CCCCCCCCCCCCCCCCCCCCCCCC.....

Quanto tempo leva a estrutura para passar pelo estado enovelado?

- N resíduos
- Perturbações independentes

Probabilidade de enovelamento instantânea:

$$P_{\text{enovelamento}} = \alpha^N$$

Probabilidade de permanecer desenovelada:

$$P_{\text{não-enovelamento}} = 1 - \alpha^N$$

Se  $t$  unidades de tempo se passaram, a estrutura tem probabilidade

$$P_{\text{permanecer desenovelada}}(t) = (1 - \alpha^N)^{t/\tau}$$

de **não** ter passado pelo estado enovelado.

## Um modelo simples de enovelamento

Qual o tempo necessário que para que 50% das proteínas atinjam a conformação correta?

$$P_{\text{enovelar}}(t) = 1 - (1 - \alpha^N)^{t/\tau}$$

$$P_{\text{enovelar}}(t) = 0,5$$

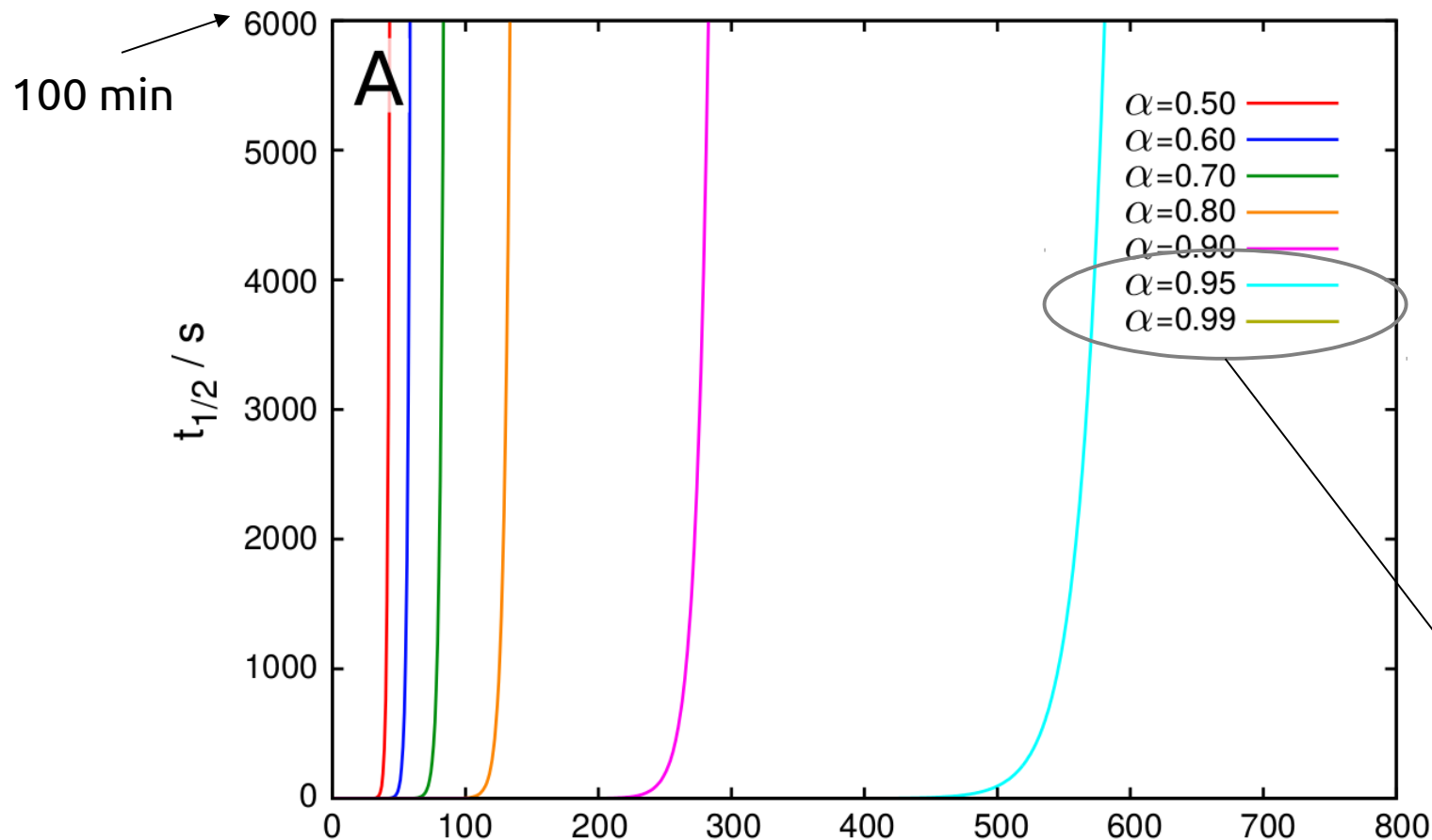
$$1 - (1 - \alpha^N)^{t/\tau} = 0,5$$

$$t_{1/2} = \frac{\ln 0,5}{\ln(1 - \alpha^N)} \tau$$

## Um modelo simples de enovelamento

Qual o tempo necessário que para que 50% das proteínas atinjam a conformação correta?

$$t_{1/2} = \frac{\ln 0,5}{\ln(1 - \alpha^N)} \tau$$



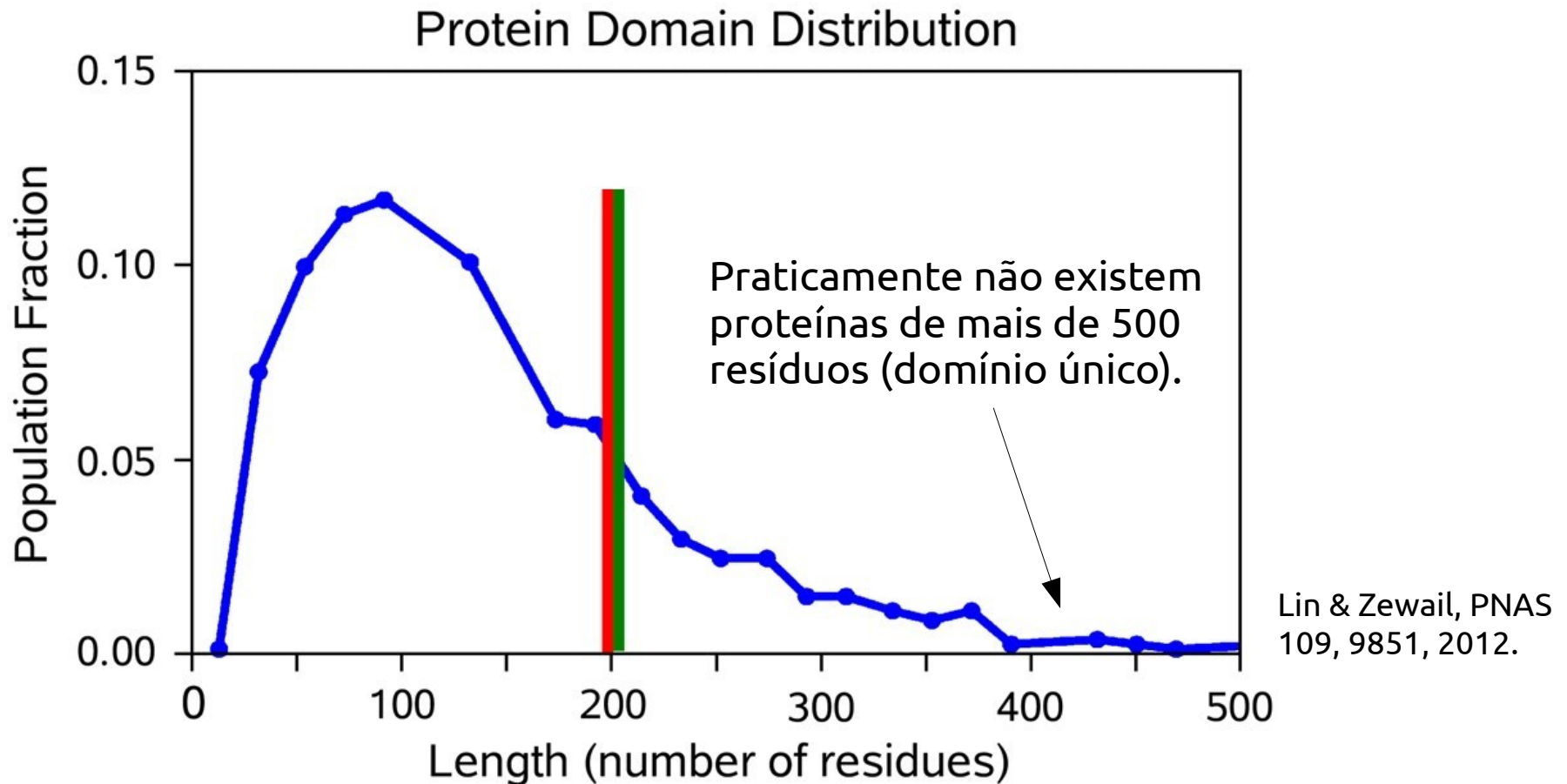
Preferência estatística pela conformação correta de cada resíduo.

Para  $\alpha \sim 0,95$ , proteínas de até 500 resíduos enovelam rápido.

## Um modelo simples de enovelamento

95% de preferência pela conformação correta permite o enovelamento de proteínas de ~500 resíduos.

$$t_{1/2} = \frac{\ln 0,5}{\ln(1 - \alpha^N)} \tau$$



**95% de preferência pelo estado correto fazem algum sentido?**

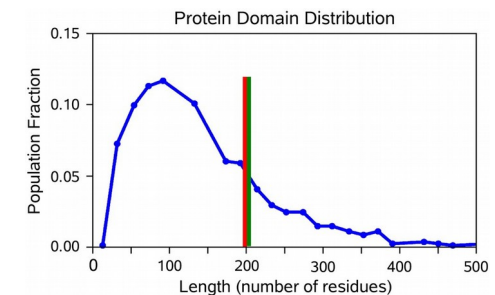
## Um modelo simples de enovelamento

### Resumo:

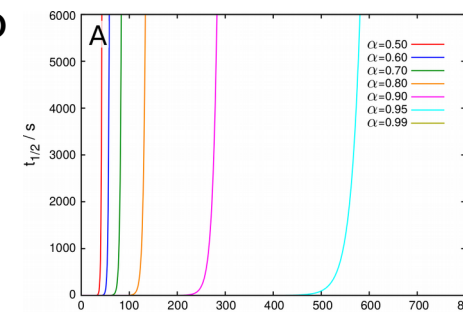
- Proteínas podem assumir múltiplas conformações.
- Definimos um mecanismo muito simples, de mudanças aleatórias na estrutura, para tentar explicar o enovelamento.

$$t_{1/2} = \frac{\ln 0,5}{\ln(1 - \alpha^N)} \tau$$

- Na natureza essencialmente só existem proteínas de menos de 500 aminoácidos.



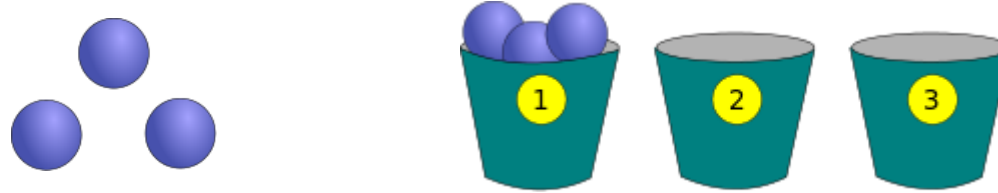
- Se as conformações corretas dos resíduos tiverem uma probabilidade de 95% em relação às incorretas, o mecanismo simples explica o enovelamento de proteínas reais.



95% de preferência pela estrutura correta fazem sentido?



# Um pouco de termodinâmica estatística



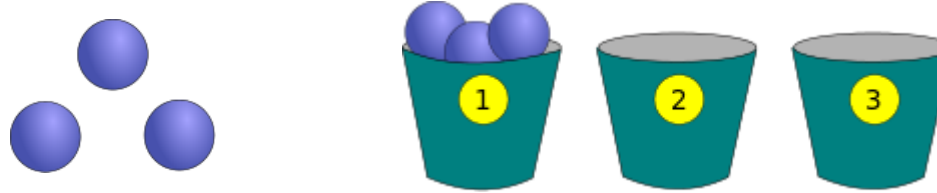
Bola 1	Bola 2	Bola 3
1	1	1
1	1	2
1	1	3
1	2	1
1	2	2
1	2	3
1	3	1
1	3	2
1	3	3
2	1	1
2	1	2
2	1	3
2	2	1
2	2	2
2	2	3
2	3	1
2	3	2
2	3	3
3	1	1
3	1	2
3	1	3
3	2	1
3	2	2
3	2	3
3	3	1
3	3	2
3	3	3

$3 \times 3 \times 3 = 27$  resultados possíveis

Todos os resultados são igualmente prováveis.

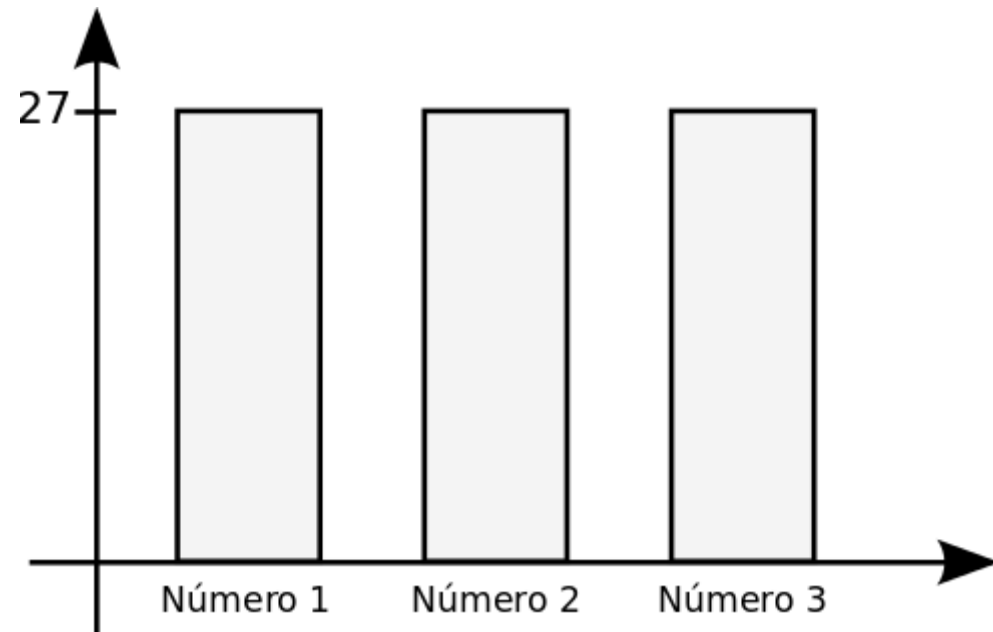
$$P = \left(\frac{1}{3}\right)^3$$

# Um pouco de termodinâmica estatística

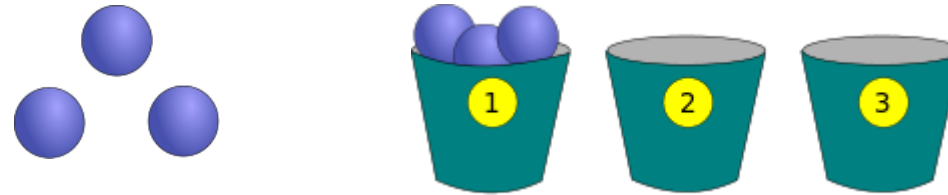


Bola 1	Bola 2	Bola 3
1	1	1
1	1	2
1	1	3
1	2	1
1	2	2
1	2	3
1	3	1
1	3	2
1	3	3
2	1	1
2	1	2
2	1	3
2	2	1
2	2	2
2	2	3
2	3	1
2	3	2
2	3	3
3	1	1
3	1	2
3	1	3
3	2	1
3	2	2
3	2	3
3	3	1
3	3	2
3	3	3

O número 1 aparece 27 vezes  
 O número 2 aparece 27 vezes  
 O número 3 aparece 27 vezes



# Um pouco de termodinâmica estatística



Mesmo sorteio, novas regras:

1. A cada balde corresponde uma “energia”

Balde 1: Energia = 1,00

Balde 2: Energia = 2,00

Balde 3: Energia = 3,00

2. Só são válidos os sorteios nos quais a energia total for 5,00

Bola 1      Bola 2      Bola 3

1	1	1
1	1	2
<b>1</b>	<b>1</b>	<b>3</b>
1	2	1
<b>1</b>	<b>2</b>	<b>2</b>
1	2	3
<b>1</b>	<b>3</b>	<b>1</b>
1	3	2
1	3	3
2	1	1
<b>2</b>	<b>1</b>	<b>2</b>
2	1	3
<b>2</b>	<b>2</b>	<b>1</b>
2	2	2
2	2	3
2	3	1
2	3	2
2	3	3
<b>3</b>	<b>1</b>	<b>1</b>
3	1	2
3	1	3
3	2	1
3	2	2
3	2	3
3	3	1
3	3	2
3	3	3

**Agora, só há 6 resultados admissíveis.**

# Um pouco de termodinâmica estatística



Bola 1      Bola 2      Bola 3

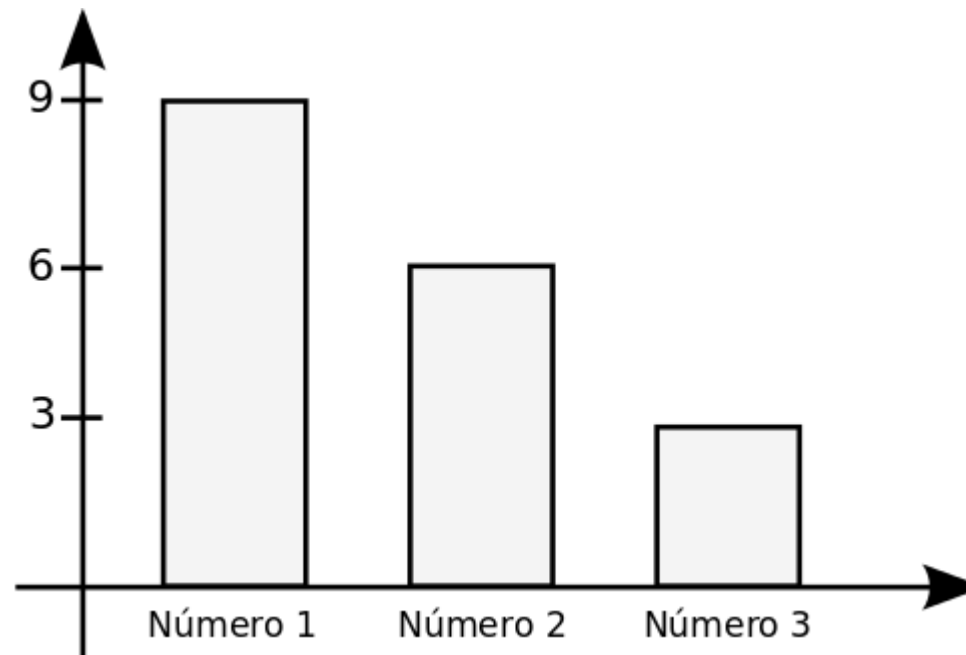
1	1	1
1	1	2
<b>1</b>	<b>1</b>	<b>3</b>
1	2	1
<b>1</b>	<b>2</b>	<b>2</b>
1	2	3
<b>1</b>	<b>3</b>	<b>1</b>
1	3	2
1	3	3
2	1	1
<b>2</b>	<b>1</b>	<b>2</b>
2	1	3
<b>2</b>	<b>2</b>	<b>1</b>
2	2	2
2	2	3
2	3	1
2	3	2
2	3	3
<b>3</b>	<b>1</b>	<b>1</b>
3	1	2
3	1	3
3	2	1
3	2	2
3	2	3
3	3	1
3	3	2
3	3	3

Entre os resultados admissíveis:

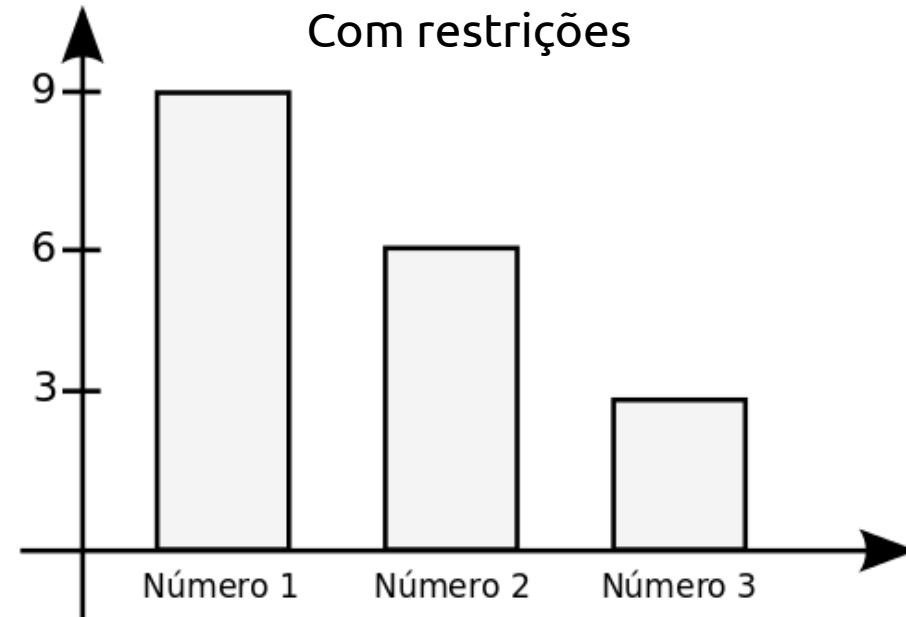
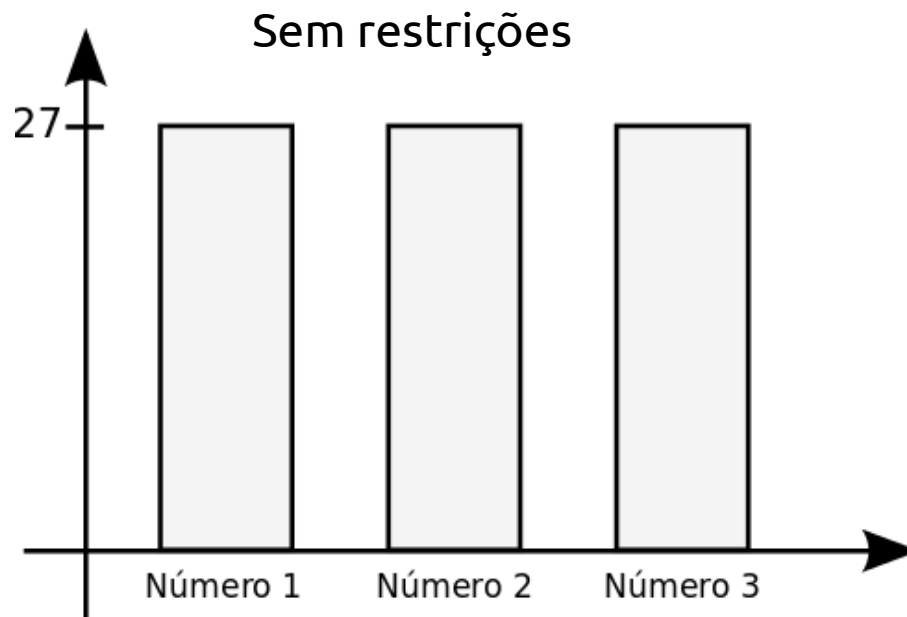
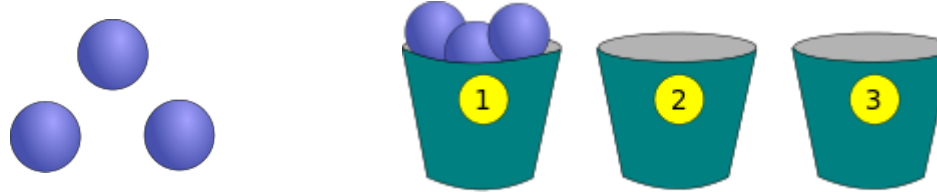
O número 1 aparece 9 vezes

O número 2 aparece 6 vezes

O número 3 aparece 3 vezes



## Um pouco de termodinâmica estatística



**Conclusão:** Em um conjunto de sorteios nos quais a soma dos resultados está restrita a um valor, **os valores menores tornam-se mais prováveis.**

## Um pouco de termodinâmica estatística

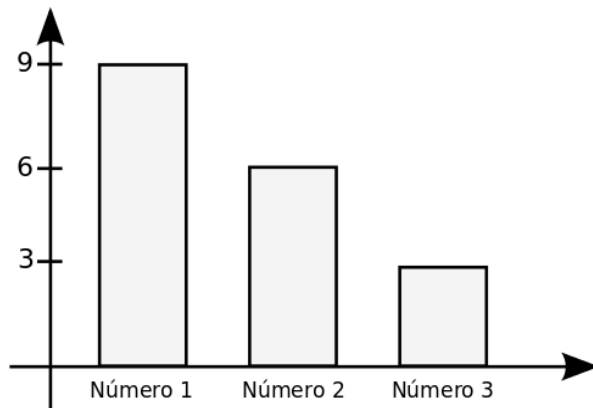


Imagine que nossa proteína está em uma garrafa térmica (isolada do ambiente).

A energia total do sistema é constante.

Cada aminoácido pode assumir conformações corretas e incorretas.

**As conformações tem diferentes energias.**



Se as conformações **corretas tem menor energia** que as incorretas, como a energia total é constante, as conformações corretas **serão mais prováveis!**

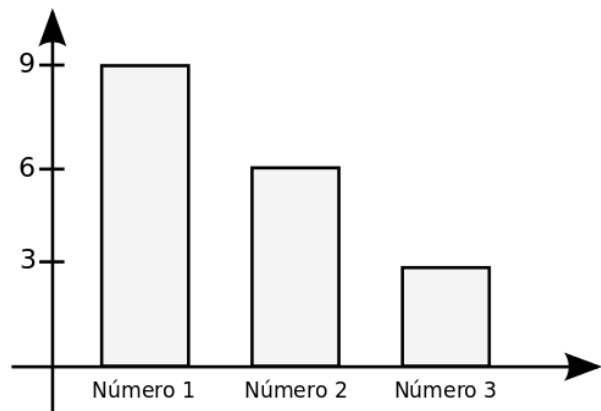
**Probabilidade**



**Energia**

“Distribuição de Boltzmann”:  $P \propto e^{-E/RT}$

## Um pouco de termodinâmica estatística



$$P \propto e^{-E/RT}$$

Probabilidade da conformação correta

$$\frac{P_C}{P_I} = \frac{\alpha}{1 - \alpha}$$

$$\frac{P_C}{P_I} = \frac{e^{-E_C/RT}}{e^{-E_I/RT}}$$

$$\frac{\alpha}{1 - \alpha} = \frac{e^{-E_C/RT}}{e^{-E_I/RT}}$$

$$E_C - E_I = -RT \ln \left( \frac{\alpha}{1 - \alpha} \right)$$

Essa é a diferença de energia que favorece uma ou outra conformação de acordo com a probabilidade relativa  $\alpha$ .

## Um pouco de termodinâmica estatística



Probabilidade



Energia

$$E_C - E_I = -RT \ln \left( \frac{\alpha}{1 - \alpha} \right)$$

95% de preferência pela estrutura correta fazem sentido?

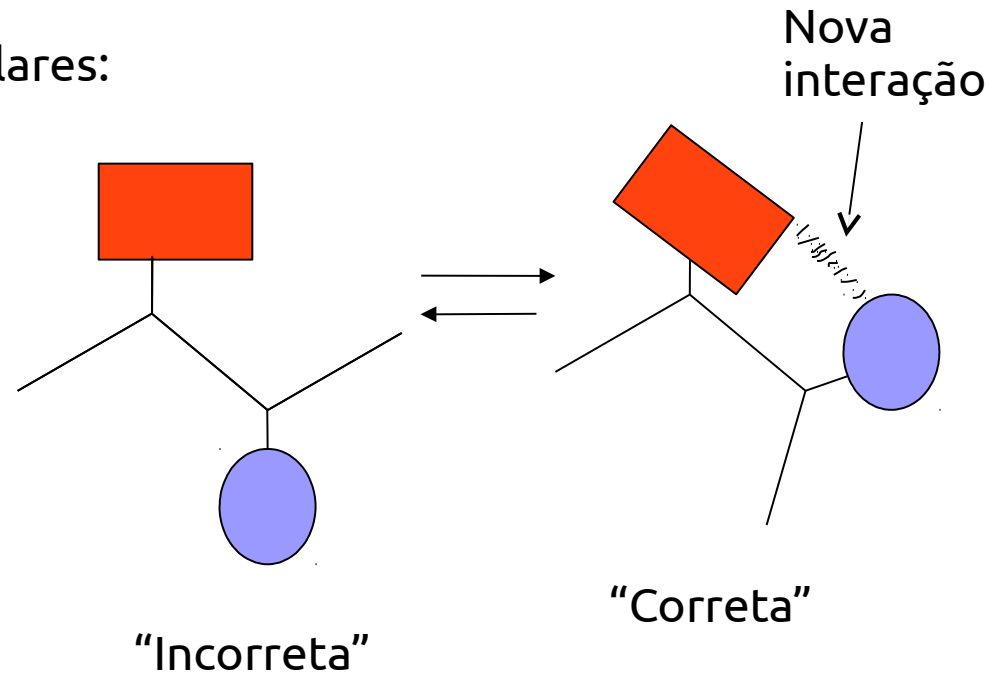
**Resposta:** Depende da diferença de energia que ela implica!

Se as **diferenças de energias** implicadas forem **razoáveis** do ponto de vista das **interações em nível molecular**, sim!



## Um pouco de termodinâmica estatística

Interações moleculares:



### Interações inter-moleculares:

Forças de London (van der Waals):  $\sim 0-1 \text{ kcal mol}^{-1}$

Dipolo-dipolo:  $\sim 0.5 \text{ a } 2 \text{ kcal mol}^{-1}$

Ligações de hidrogênio:  $\sim 10 - 16 \text{ kcal mol}^{-1}$

## Um pouco de termodinâmica estatística

$$E_C - E_I = -RT \ln \left( \frac{\alpha}{1 - \alpha} \right)$$

T = 36°C

$\alpha$	$E_C - E_I / \text{kcal mol}^{-1}$
0,5	0,0

## Um pouco de termodinâmica estatística

$$E_C - E_I = -RT \ln \left( \frac{\alpha}{1 - \alpha} \right)$$

T = 36°C

$\alpha$	$E_C - E_I / \text{kcal mol}^{-1}$
0,5	0,0
0,7	-0,5

## Um pouco de termodinâmica estatística

$$E_C - E_I = -RT \ln \left( \frac{\alpha}{1 - \alpha} \right)$$

T = 36°C

$\alpha$	$E_C - E_I / \text{kcal mol}^{-1}$
0,5	0,0
0,7	-0,5
0,9	-1.3

## Um pouco de termodinâmica estatística

$$E_C - E_I = -RT \ln \left( \frac{\alpha}{1 - \alpha} \right)$$

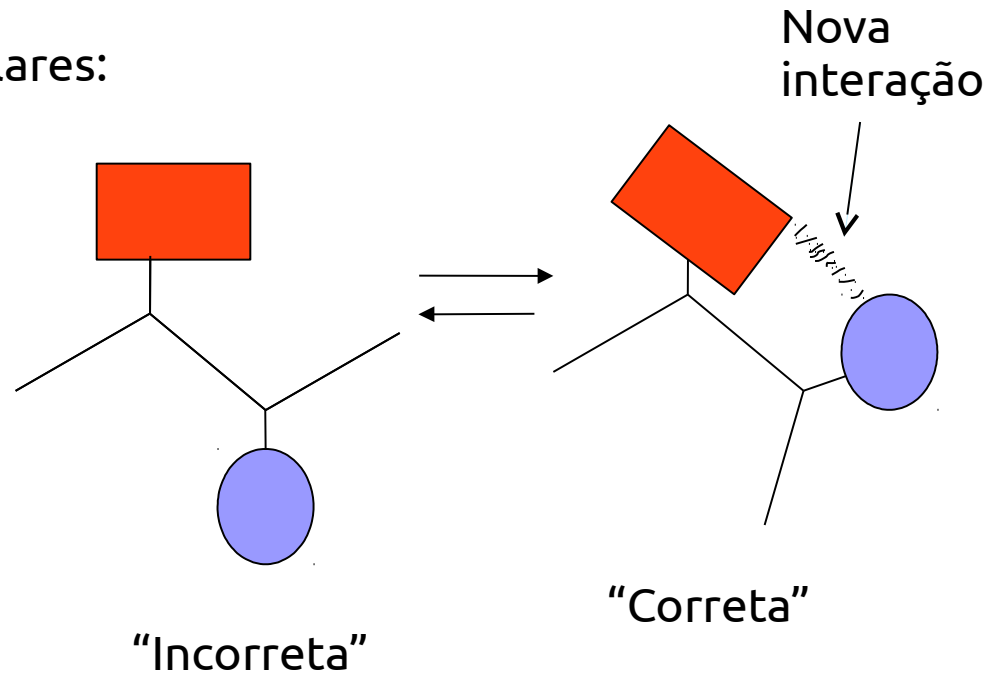
T = 36°C

$\alpha$	$E_C - E_I / \text{kcal mol}^{-1}$
0,5	0,0
0,7	-0,5
0,9	-1,3
0,95	-1,7

Se cada conformação correta for  $-1,7 \text{ kcal mol}^{-1}$  mais favorável que a incorreta, proteínas de 500 resíduos enovelam rápido!

## Um pouco de termodinâmica estatística

Interações moleculares:



### Interações inter-moleculares:

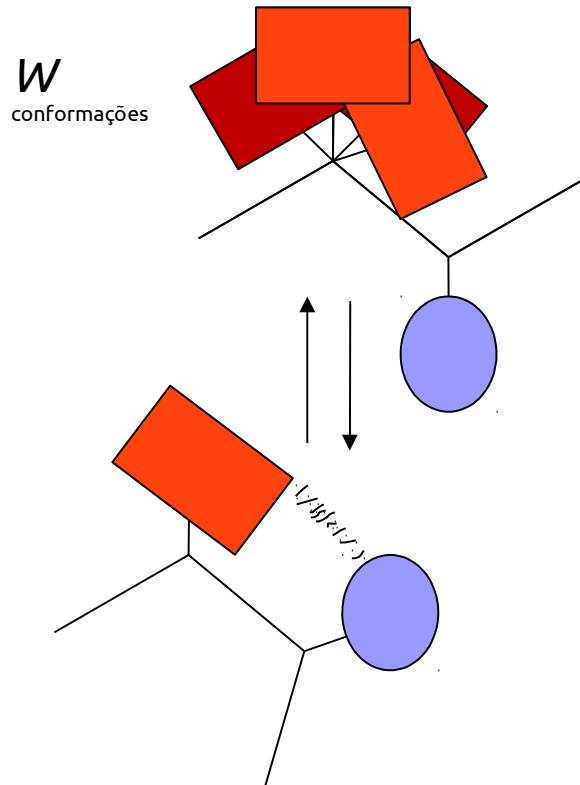
Forças de London (van der Waals):  $\sim 0-1 \text{ kcal mol}^{-1}$

Dipolo-dipolo:  $\sim 0.5 \text{ a } 2 \text{ kcal mol}^{-1}$

Ligações de hidrogênio:  $\sim 10 - 16 \text{ kcal mol}^{-1}$

## Um pouco de termodinâmica estatística

Mas deve haver mais conformações incorretas que corretas!



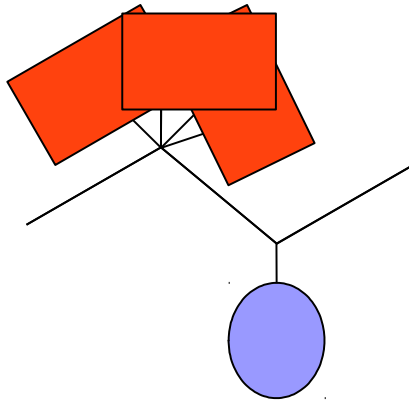
$$\frac{P_C}{P_I} = \frac{e^{-E_C/RT}}{W e^{-E_I/RT}}$$

↓  
Soma das probabilidades das conformações incorretas

$$E_C - E_I = -RT \ln \left( W \frac{\alpha}{1 - \alpha} \right)$$

## Um pouco de termodinâmica estatística

Mas deve haver mais conformações incorretas que corretas!



$$E_C - E_I = -RT \ln \left( W \frac{\alpha}{1 - \alpha} \right)$$

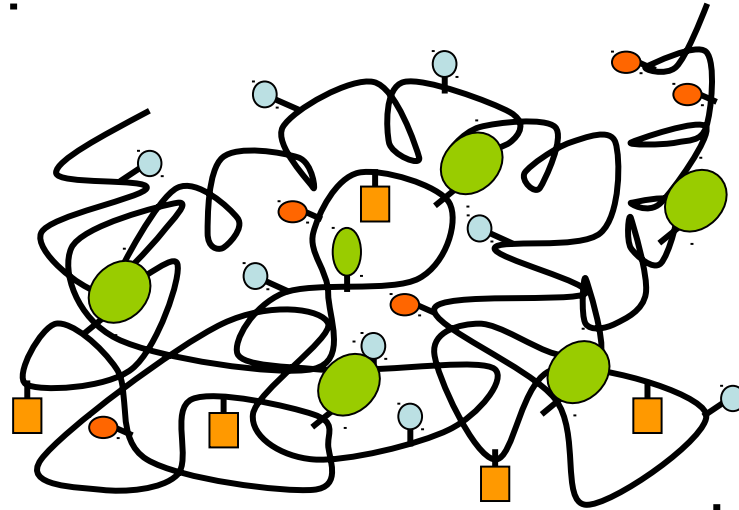
~ Número de confôrmeros  
dos aminoácidos  
nas proteínas.

Pessimista

$\alpha$	$W$	$E_C - E_I / \text{kcal mol}^{-1}$
0,95	1	-1,7
0,95	2	-2,4
0,95	10	-3,2
0,95	1000	-6,1
0,95	100.000	-8,9

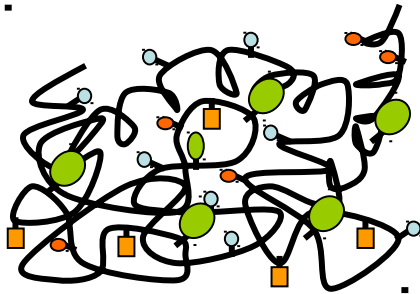


## Conclusões



- A principal contribuição para a solução do problema de enovelamento é a redução incremental da energia, localmente. Isto, somente, resolve o problema qualitativo, que passa a ser quantitativo.
- Os tamanhos das proteínas estão determinados em grande parte pela limitação do mecanismo acima.
- O mecanismo é válido porque as proteínas foram selecionadas pela evolução para apresentar superfícies de energia potencial pouco rugosas.

## Leituras recomendadas e agradecimentos



L. Martínez, “Introducing the Levinthal's protein folding paradox and its solution.” J. Chem. Educ. 91, 1818, 2014.

R. Zwanzig, A. Szabo, B. Bagchi, “Levinthal’s paradox”. PNAS 89, 20, 1992.

M. Karplus, “Behind the folding funnel diagram”, Nature Chem. Biol. 7, 401, 2011.

S. S. Plotkin, **J. N. Onuchic**, “Understanding protein folding with energy landscape theory. Part I: Basic concepts”. Quart. Rev. Biophys. 35, 111-167, 2002.

M. Lin, A. H. Zewail, “Simplicity in Complexity”. Annal. der Physik. 524, 379, 2012.

### Agradecimentos:

Lúcio / Laura

FAPESP (2010/16947-9) / CNPq

**Obrigado!**