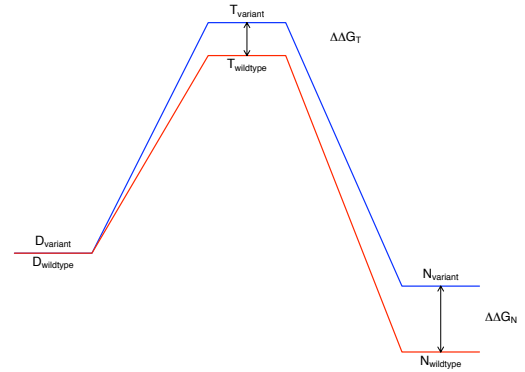


# On the Precision of Experimentally Determined Protein Folding Rates and $\Phi$ Values

Ingo Ruczinski

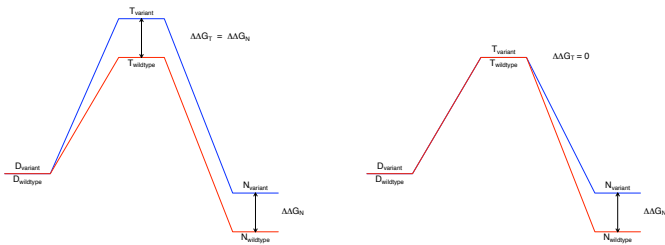
Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health

## Energy Profile



→ The  $\Phi$ -value is defined as the ratio  $\Delta\Delta G_T / \Delta\Delta G_N$ .

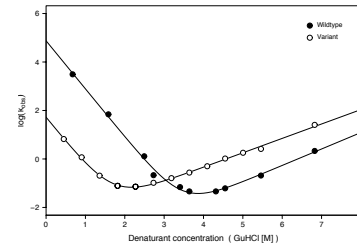
## Energy Profile



- If the part of the protein that contains the variant amino acid is fully structured in the transition state, we have  $\Delta\Delta G_T \approx \Delta\Delta G_N$ , and hence  $\Phi \approx 1$ .
- If the part of the protein that contains the variant amino acid is equal in denatured and the transition state, we have  $\Delta\Delta G_T \approx 0$ , and hence  $\Phi \approx 0$ .

At least this is the idea ...

## Phi-Value Estimation



$$\log(k_{\text{obs}}) = \log \left( \exp \left[ \log(k_f) + m_f \times \frac{C_{\text{GuHCl}}}{RT} \right] + \exp \left[ \log(k_u) + m_u \times \frac{C_{\text{GuHCl}}}{RT} \right] \right)$$

$$\Delta\Delta G_T = RT \times \left[ \log(k_f^{\text{wildtype}}) - \log(k_f^{\text{variant}}) \right]$$

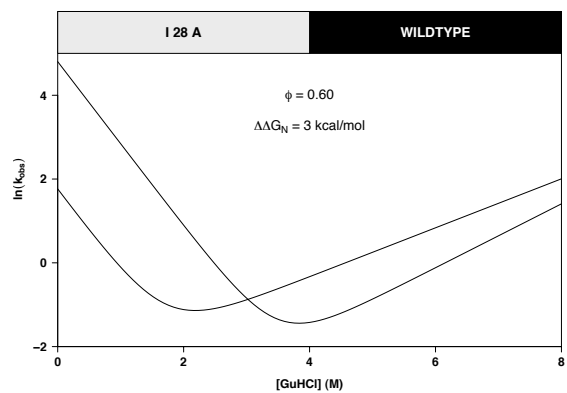
$$\Delta\Delta G_N = RT \times \left[ \log(k_f^{\text{wildtype}}) - \log(k_u^{\text{wildtype}}) - \log(k_f^{\text{variant}}) + \log(k_u^{\text{variant}}) \right]$$

## Questions

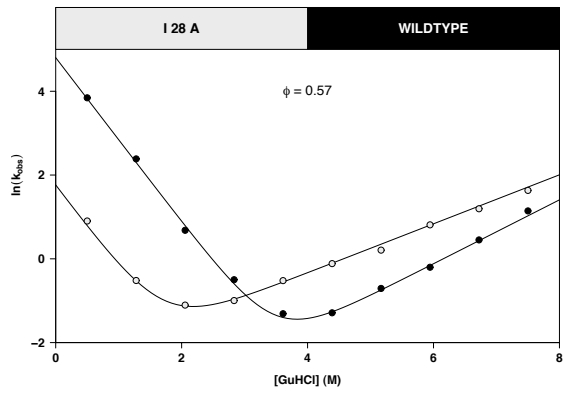
- What is the threshold for the difference in stability ( $\Delta\Delta G_N$ ) between two variants to assure reliable estimates of  $\Phi$ ?

It depends.

## Precision

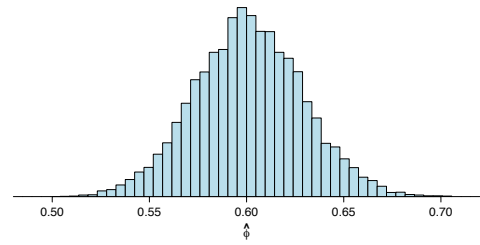


## Precision



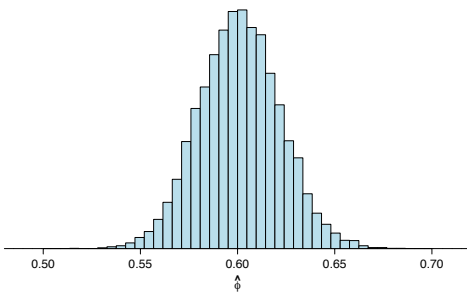
## Precision

10 data points / chevron



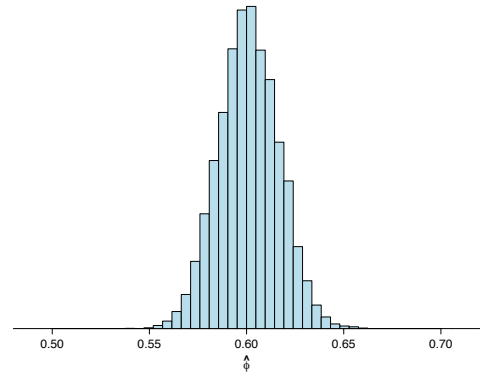
## Precision

20 data points / chevron

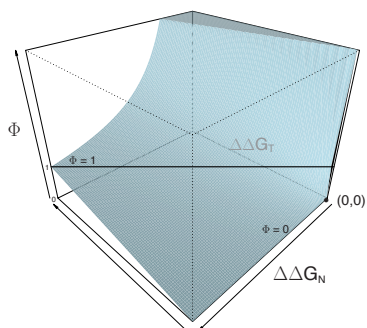


## Precision

40 data points / chevron



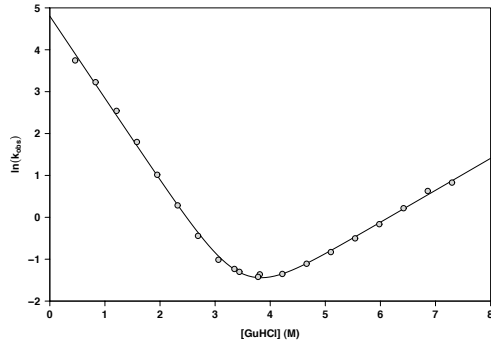
## Precision



## Questions

- What is the threshold for the difference in stability ( $\Delta\Delta G_N$ ) between two variants to assure reliable estimates of  $\hat{\phi}$ ?
- How can we construct valid standard errors for the estimates of  $\hat{\phi}$ ?

## Standard Error

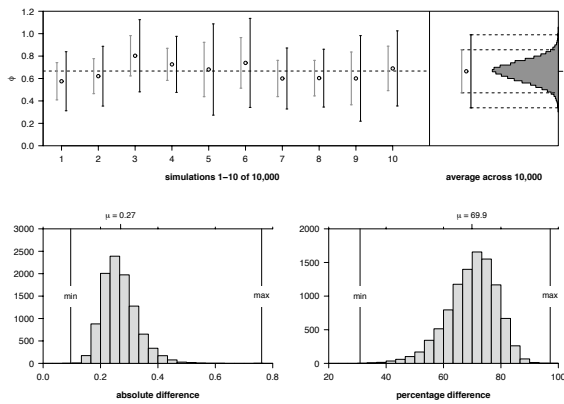


$$\ln(k_{\text{obs}}) = \log \left( \exp \left[ \log(k_f) + m_f \times \frac{[\text{GuHCl}]}{RT} \right] + \exp \left[ \log(k_u) + m_u \times \frac{[\text{GuHCl}]}{RT} \right] \right)$$

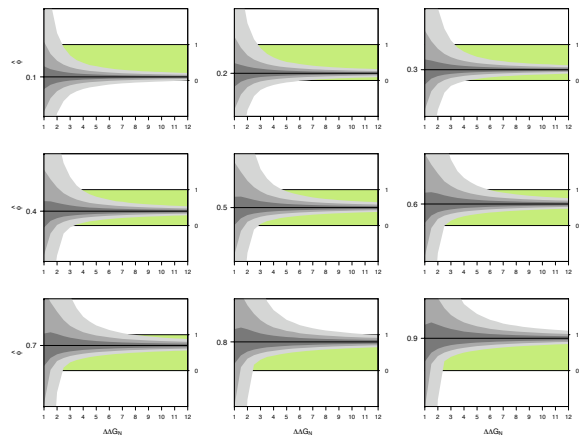
## Standard Error

$$\text{se}(\hat{\Phi}) = |\Phi| \times \sqrt{\left( \frac{\sigma_T}{\Delta\Delta G_T} \right)^2 - 2\rho_{\Delta\Delta G} \left( \frac{\sigma_T}{\Delta\Delta G_T} \right) \left( \frac{\sigma_N}{\Delta\Delta G_N} \right) + \left( \frac{\sigma_N}{\Delta\Delta G_N} \right)^2}$$

## Standard Error



## Standard Error



## Web Server

Rucznicki, Sonnick, and Plaxco

http://biostat.jhsph.edu/~rucznick/software/phi/

log's Post Estrez PubMed News Links Running Science Travel

Rucznick I, Sonnick TR, and Plaxco KW.  
Method for the accurate estimation of confidence intervals on experimental  $\Phi$ -values

Please read: important file format information for uploading your data. For options and output information, please read the help file.

- Upload your data:
- Specify the temperature:  $T =$   °C
- Specify the energy units:  kJ  kcal
- Specify the comparisons:  wildtype only  all comparisons
- Specify the type of fit:  individual fits  parallel arms
- Specify the denaturant concentrations: folding rate:  M; unfolding rate:  M
- Specify the coverage for the  $\Phi$ -value confidence intervals:  % coverage
- Specify the number of significant digits in the output:  significant digits
- And they are off:

Mail Ingo

## Web Server

Rucznicki, Sonnick, and Plaxco

http://biostat.jhsph.edu/~rucznick/software/phi/phi.cgi

log's Post Estrez PubMed News Links Running Science Travel

Results

Table 1: Estimates for the kinetic parameters. The energies are given in kJ.

	log(k)	$\Phi_{\text{GuHCl}}$	$m_f$	$m_u$	log(k)	$\Phi_{\text{GuHCl}}$	$m_f$	$m_u$	$\Delta\Delta G_f$	$\Delta\Delta G_u$	$\Phi_{\text{WT}}$	$m_{\text{WT}}$	$m_{\text{V55G}}$
WT	4.80	0.06	-4.86	0.08	-4.70	0.13	1.89	0.06	23.56	0.20	-6.75	0.07	
D8A	1.75	0.06	-4.97	0.16	-2.87	0.08	1.45	0.04	10.97	0.12	-4.42	0.15	
D8L	4.25	0.05	-5.05	0.08	-4.27	0.10	1.68	0.04	21.13	0.15	-6.74	0.07	
D8V	4.10	0.06	-4.89	0.09	-4.34	0.13	1.71	0.06	20.94	0.19	-6.81	0.08	
V55A	4.41	0.04	-4.26	0.10	-0.81	0.08	1.75	0.04	12.94	0.12	-6.01	0.09	
V55M	4.84	0.08	-4.80	0.12	-3.22	0.15	1.77	0.06	19.99	0.22	-6.87	0.10	
V55T	4.18	0.06	-4.33	0.14	-0.78	0.10	1.60	0.05	12.30	0.16	-6.13	0.13	
V55G	4.39	0.12	-4.45	0.72	2.18	0.18	1.41	0.14	5.47	0.27	-5.86	0.62	

A spreadsheet with the above data for download is here. The file is in csv format, which can directly be read, for example, into R and Excel.

Table 2: Estimates for the  $\Phi$ -values and the changes in free energy. The energies are given in kJ.

	$\Phi$	$\Phi_{\text{WT}}$	lower	upper	$\Delta\Delta G_f$	$\Delta\Delta G_u$	$\Phi_{\text{WT}}$	$\Phi_{\text{V55G}}$	$\Delta\Delta G_{\text{WT}}$	$\Delta\Delta G_{\text{V55G}}$
WT-D8A	0.60	0.02	0.96	0.64	7.56	0.20	12.99	0.37	0.20	
WT-D8L	0.66	0.11	0.30	0.80	3.96	0.20	2.43	0.40	0.13	
WT-D8V	0.66	0.13	0.40	0.92	1.74	0.21	2.62	0.43	0.13	
WT-V55A	0.09	0.02	0.06	0.12	0.96	0.18	10.62	0.37	0.12	
WT-V55M	-0.03	0.07	-0.16	0.11	-0.10	0.24	3.57	0.47	0.14	
WT-V55T	0.14	0.02	0.10	0.17	1.58	0.20	11.26	0.40	0.15	
WT-V55G	0.06	0.02	0.02	0.09	1.02	0.33	18.09	0.52	0.20	

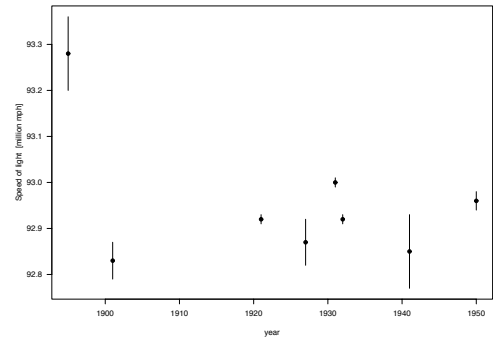
A spreadsheet with the above data for download is here. The file is also csv format.

## Questions

- What is the threshold for the difference in stability ( $\Delta\Delta G_N$ ) between two variants to assure reliable estimates of  $\Phi$ ?
- How can we construct valid standard errors for the estimates of  $\Phi$ ?
- How reproducible are  $\Phi$ -value measurements?

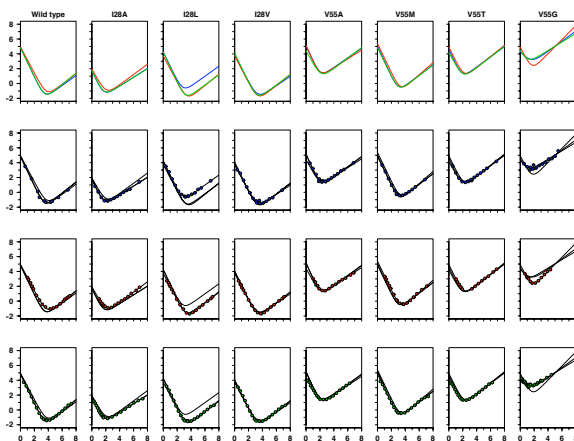
## Dang...

Estimates of the speed of light, with “confidence intervals” (1895 - 1950).



Youden (Technometrics, 1972).

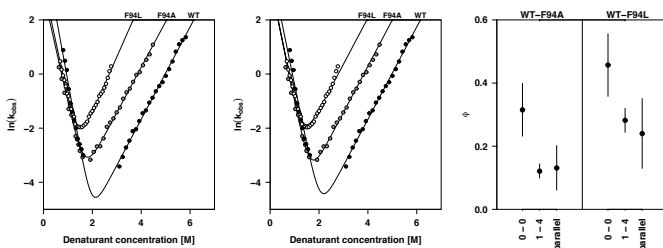
## Reproducibility?



## Questions

- What is the threshold for the difference in stability ( $\Delta\Delta G_N$ ) between two variants to assure reliable estimates of  $\Phi$ ?
- How can we construct valid standard errors for the estimates of  $\Phi$ ?
- How reproducible are  $\Phi$ -value measurements?
- What are the effects of other commonly employed techniques to calculate  $\Phi$  from kinetic data?

## Alternative Techniques



## References

- De Los Rios MA, Muralidhara BK, Wildes D, Sosnick TR, Marqusee S, Wittung-Stafshede P, Plaxco KW, Ruczinski I (2006). *On the Precision of Experimentally Determined Protein Folding Rates and Phi-Values.* *Protein Science*, 15(3): 553-63.
- Maxwell KL, Wildes D, Zarrine-Afsar A, De Los Rios MA, Brown AG, Friel CT, Hedberg L, Horng JC, Bona D, Miller EJ, Vallee-Belisle A, Main ER, Bemporad F, Qiu L, Teilum K, Vu ND, Edwards AM, Ruczinski I, Poulsen FM, Kragelund BB, Michnick SW, Chiti F, Bai Y, Hagen SJ, Serrano L, Oliveberg M, Raleigh DP, Wittung-Stafshede P, Radford SE, Jackson SE, Sosnick TR, Marqusee S, Davidson AR, Plaxco KW (2005). *Protein Folding: Defining a Standard Set of Experimental Conditions and a Preliminary Kinetic Data Set of Two-State Proteins.* *Protein Science*, 14(3): 602-16.
- Ruczinski I, Sosnick TR, Plaxco KW (2006). *Methods for the Accurate Estimation of Confidence Intervals on Protein Folding  $\Phi$  Values.* *Protein Science* (to appear).

## Acknowledgments

---

- [UC Santa Barbara](#) Department of Chemistry Biochemistry  
*Miguel de los Rios, Kevin Plaxco.*
- [University of Chicago](#) Department of Chemistry  
*Tobin Sosnick.*
- [Rice University](#) Department of Biochemistry and Cell Biology  
*BK Muralidhara, Pernilla Wittung-Stafshede.*
- [UC Berkeley](#) Department of Molecular and Cell Biology  
*David Wildes, Susan Marqusee.*

<http://biostat.jhsph.edu/~iruczins>