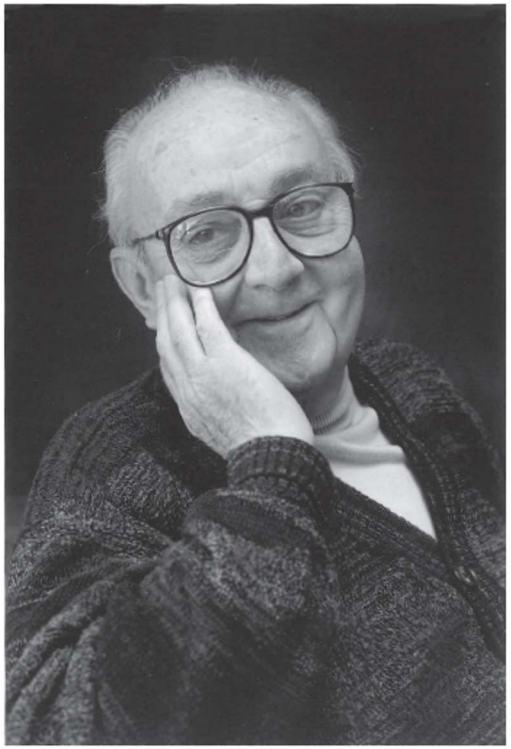
Getting the Right Answer for Wrong Reasons: An Example with the Macromolecular Rate Theory

Jinyun Tang and Bill Riley
Lawrence Berkeley National Laboratory

“All models are wrong but some are useful.”

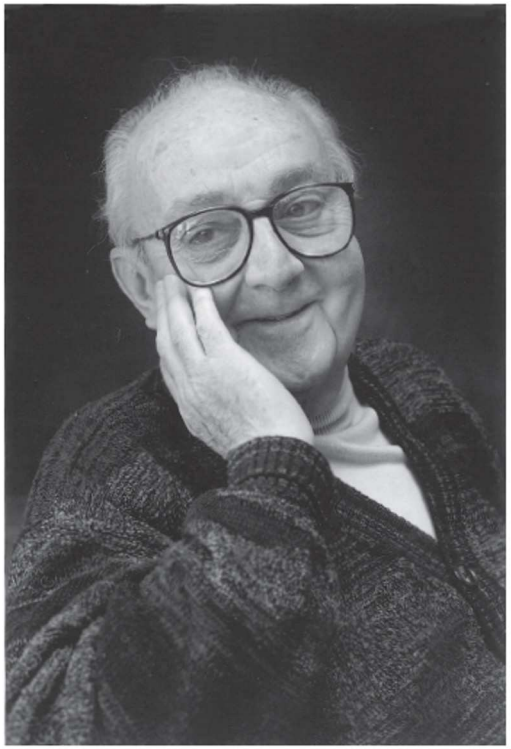
Who’s the first to say it?

Essentially all models are wrong, but some are useful.



George Box, 1919-2013

Essentially all models are wrong, but some are useful.



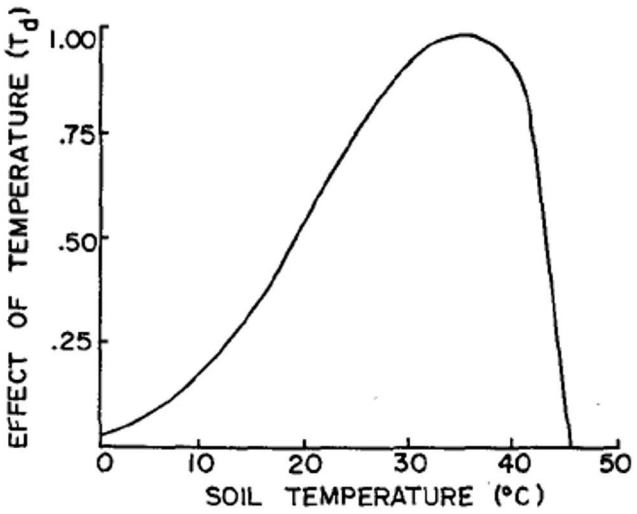
George Box, 1919-2013

Parsimony
Since all models are wrong the scientist cannot obtain a "correct" one by excessive elaboration.

Worrying selectively
Since all models are wrong the scientist must be alert to what is importantly wrong.

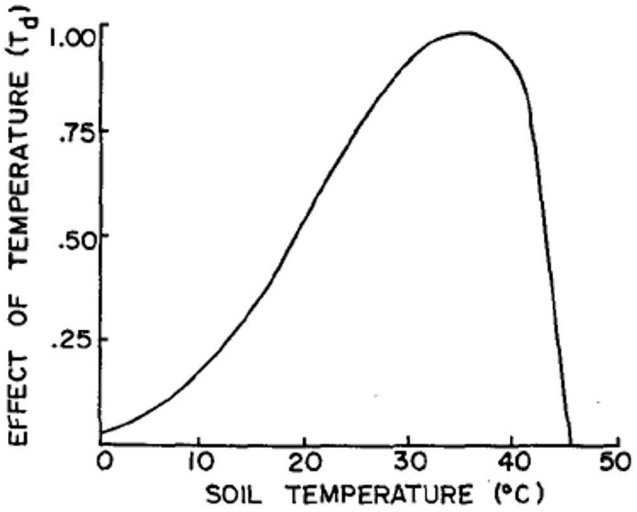
Box, 1976: Science and statistics.

The temperature sensitivity of biochemical rates in models

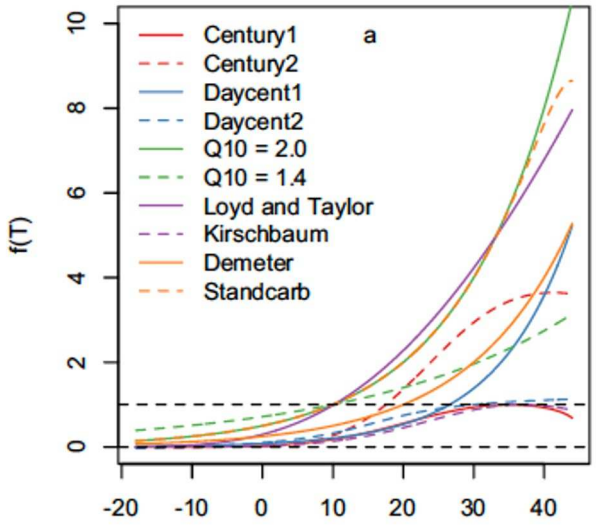


CENTURY (Parton et al., 1987)

The temperature sensitivity of biochemical rates in models



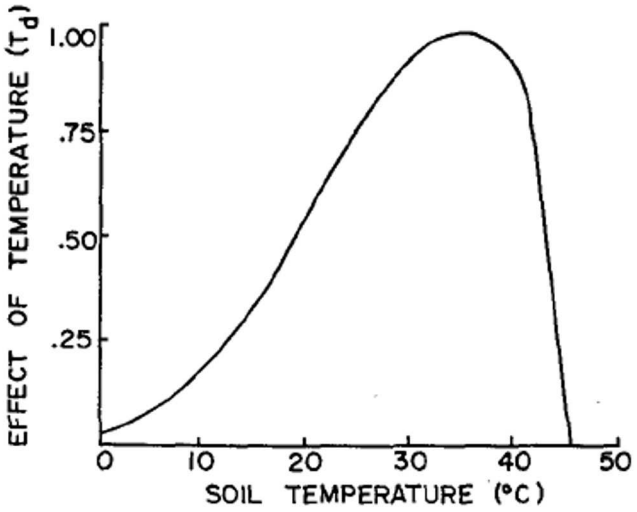
CENTURY (Parton et al., 1987)



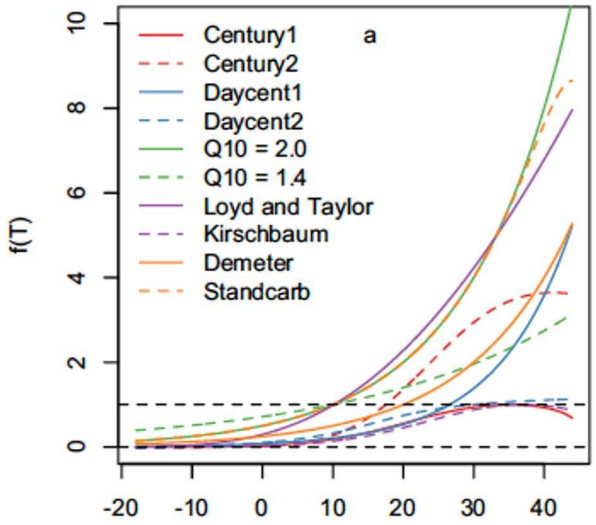
Sierra et al. (2015)

The temperature sensitivity of biochemical rates in models

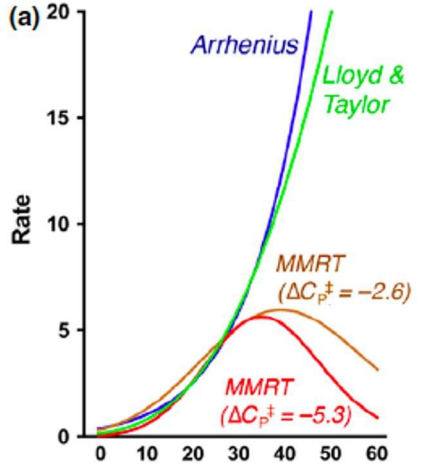
Organisms are dead/dying



CENTURY (Parton et al., 1987)

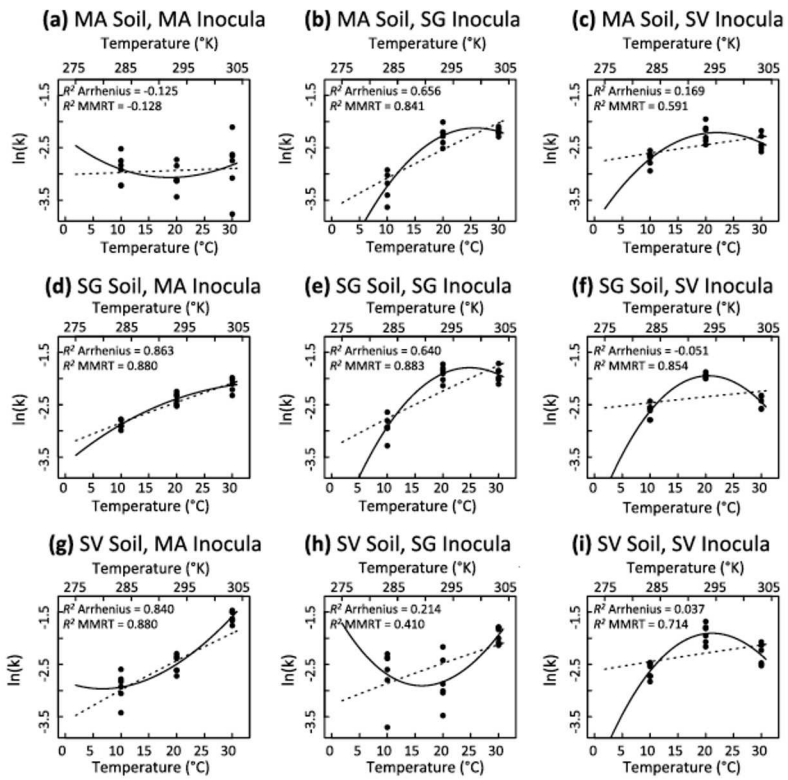


Sierra et al. (2015)

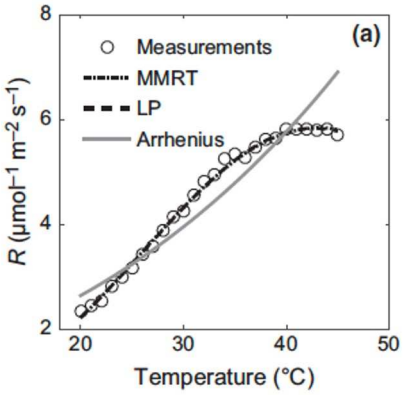


Schipper et al., 2014 Macromolecular Rate Theory (MMRT)

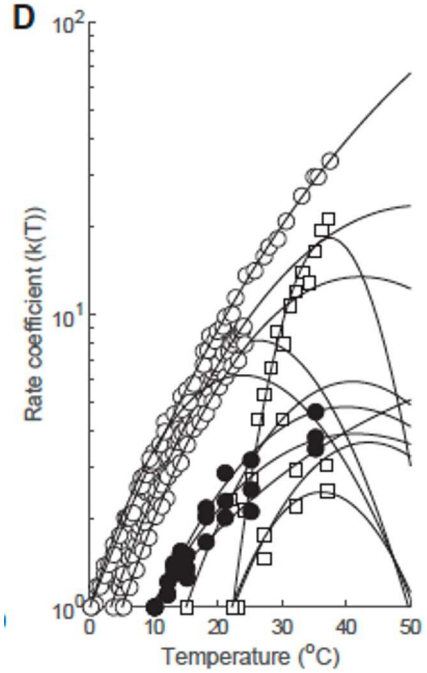
MMRT supremacy (empirical support)



Alster et al., 2016: soil incubations

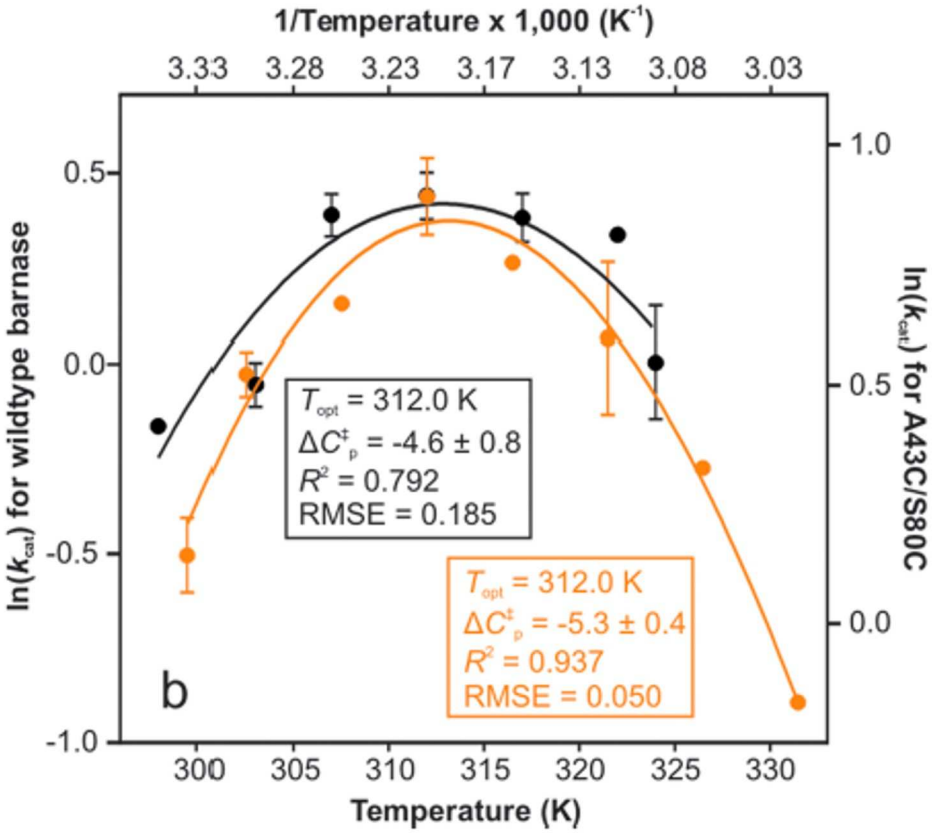


Liang et al. 2018: Leaf respiration



Pahlavan et al., 2022: Human neuron membrane conductance to cations.

MMRT supremacy: theoretical support



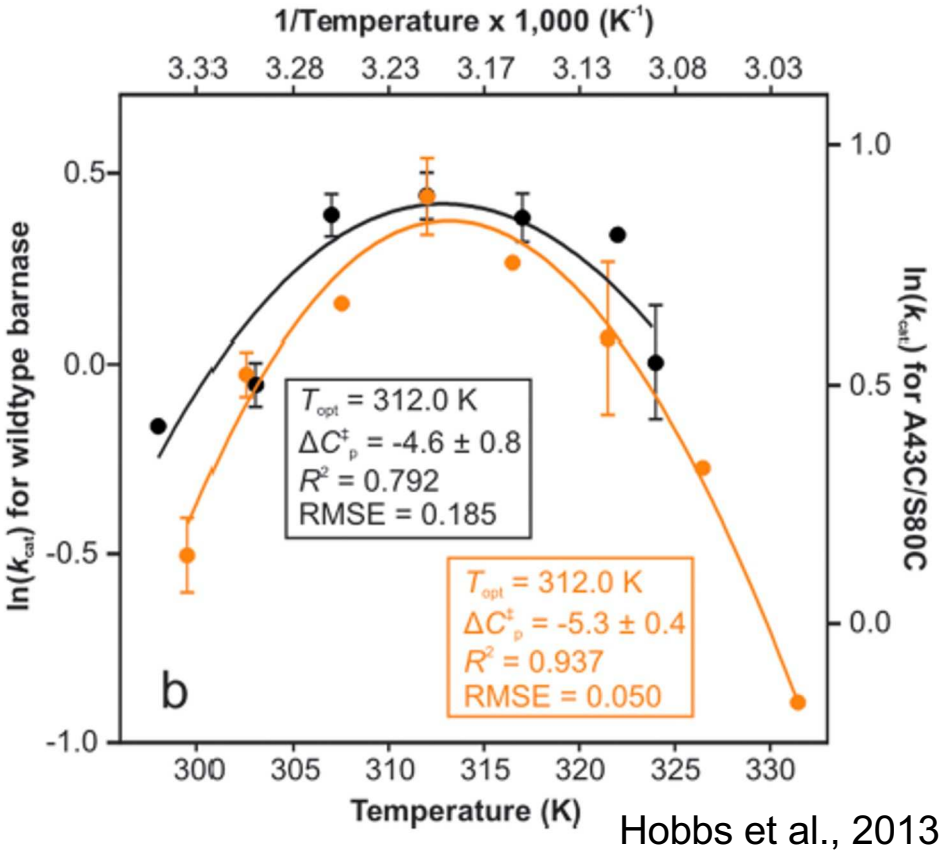
$$k = \frac{k_B T}{h} e^{-\Delta G^{\ddagger}/RT}$$ Transition state theory (Eyring, 1937)

$$\ln(k) = \ln\left(\frac{k_B T}{h}\right) - \frac{[\Delta G^{\ddagger}]}{RT}$$

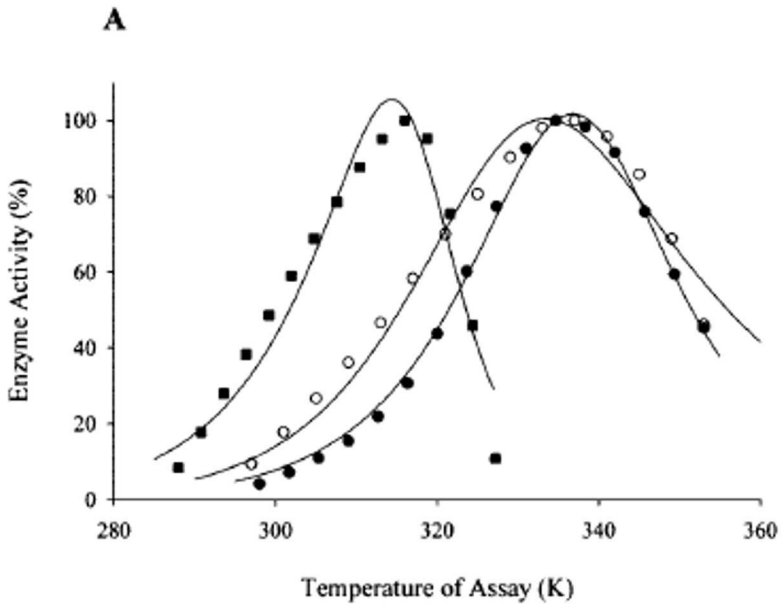
$$\ln(k) = \ln\left(\frac{k_B T}{h}\right) - \frac{[\Delta H^{\ddagger}_{T_0} + \Delta C^{\ddagger}_p (T - T_0)]}{RT} + \frac{[\Delta S^{\ddagger}_{T_0} + \Delta C^{\ddagger}_p \ln(T/T_0)]}{R}$$

Hobbs et al., 2013

MMRT supremacy: theoretical support



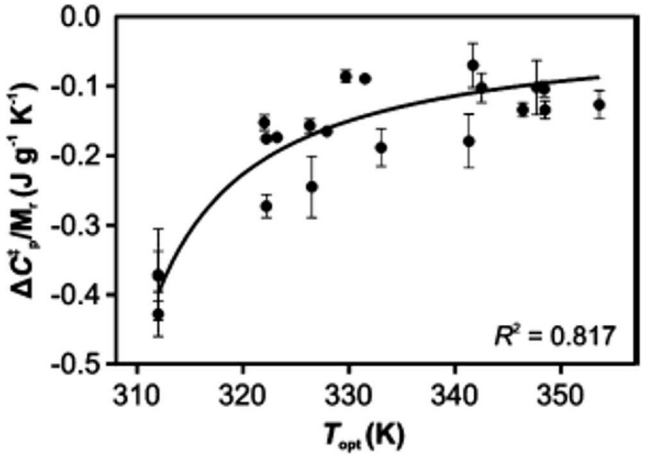
Hobbs et al., 2013



Peterson et al., 2004

MMRT predicts optimal temperature as a function of heat capacity

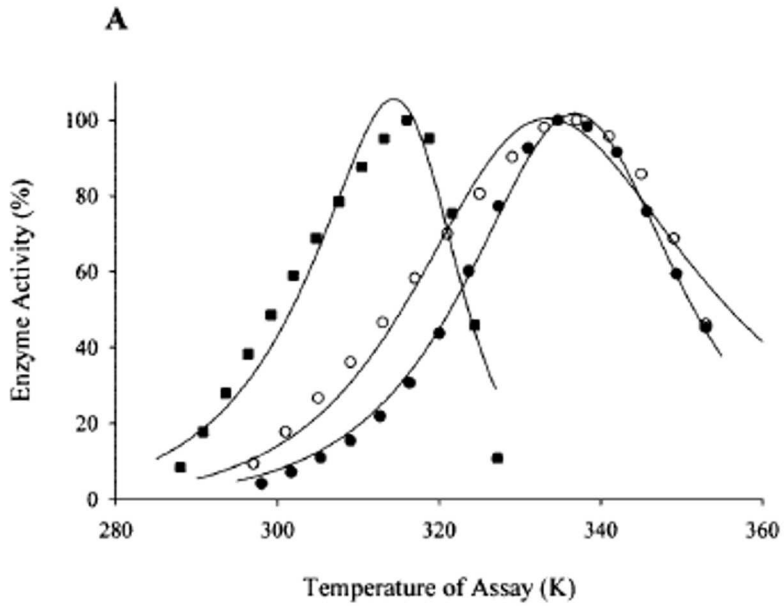
$$-\Delta C_p^{\ddagger} = \frac{\Delta H_{T_0}^{\ddagger} + RT_{\text{opt}}}{T_{\text{opt}} - T_0}$$



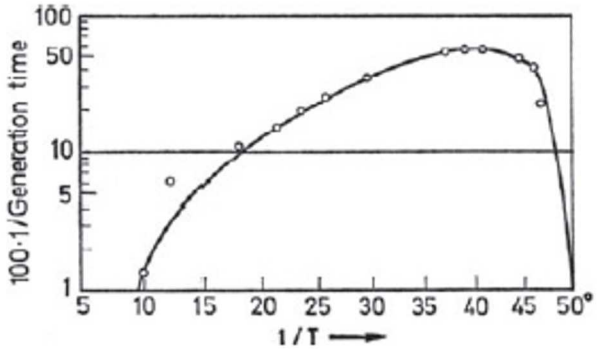
- For a given type of biochemical reaction, an optimal temperature (T_{opt}) is unically determined by heat capacity ($\Delta C_p^{=}$).
- T_{opt} or $\Delta C_p^{=}$ is important biochemical trait to characterize the thermal response of an organism.
- The change of T_{opt} or $\Delta C_p^{\Uparrow}$ reflects (evolutionary) thermal adaptation (e.g., Alster et al. 2022)

However, after re-analyzing the Hobbs et al. (2013) paper, we found something serious wrong with the MMRT.

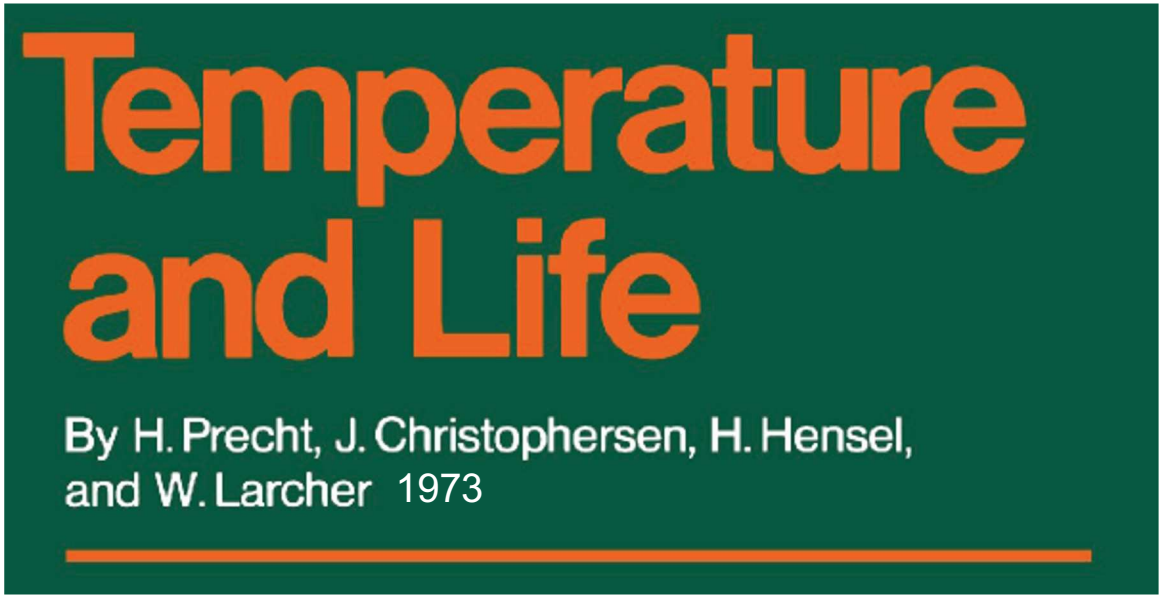
A history review of the non-monotonic temperature dependence of enzyme catalyzed reaction rates



Work in early days

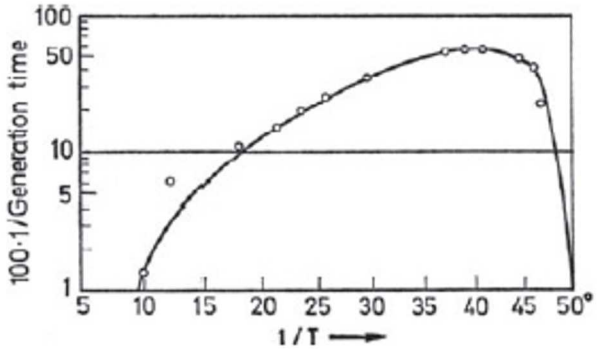


Microbial growth rate vs temperature, Buchanan and Fulmer (1930)



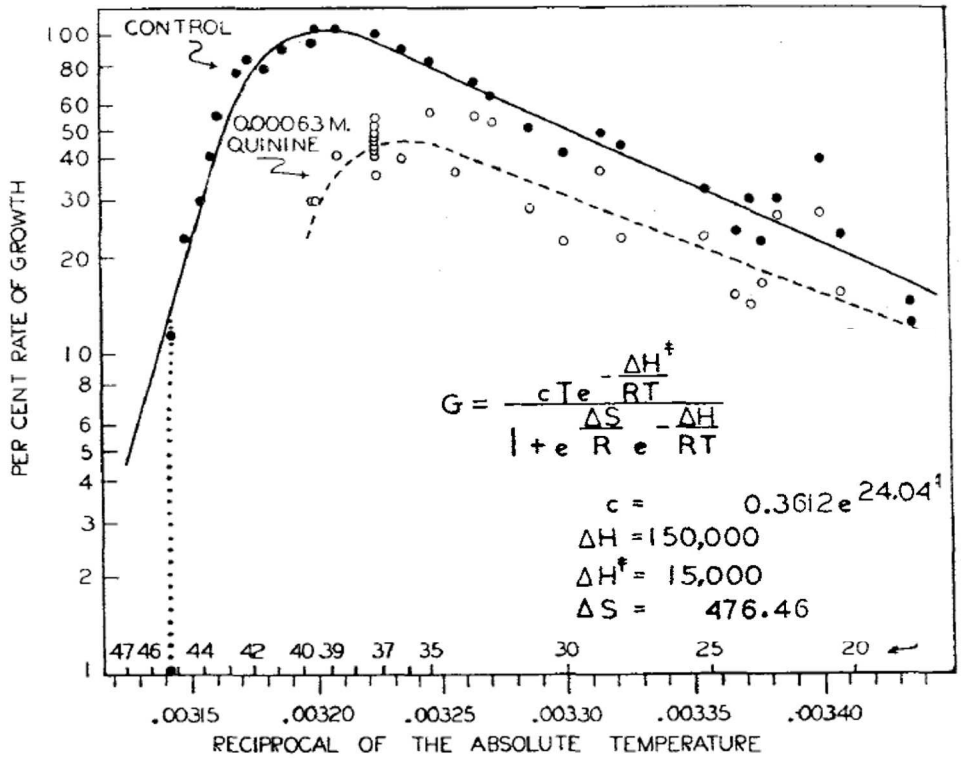
- Degree-day model, starting from Candolle (1855), Reibisch (1902)
- E. Coli growth rate, Johnson and Lewin (1946).

Work in early days

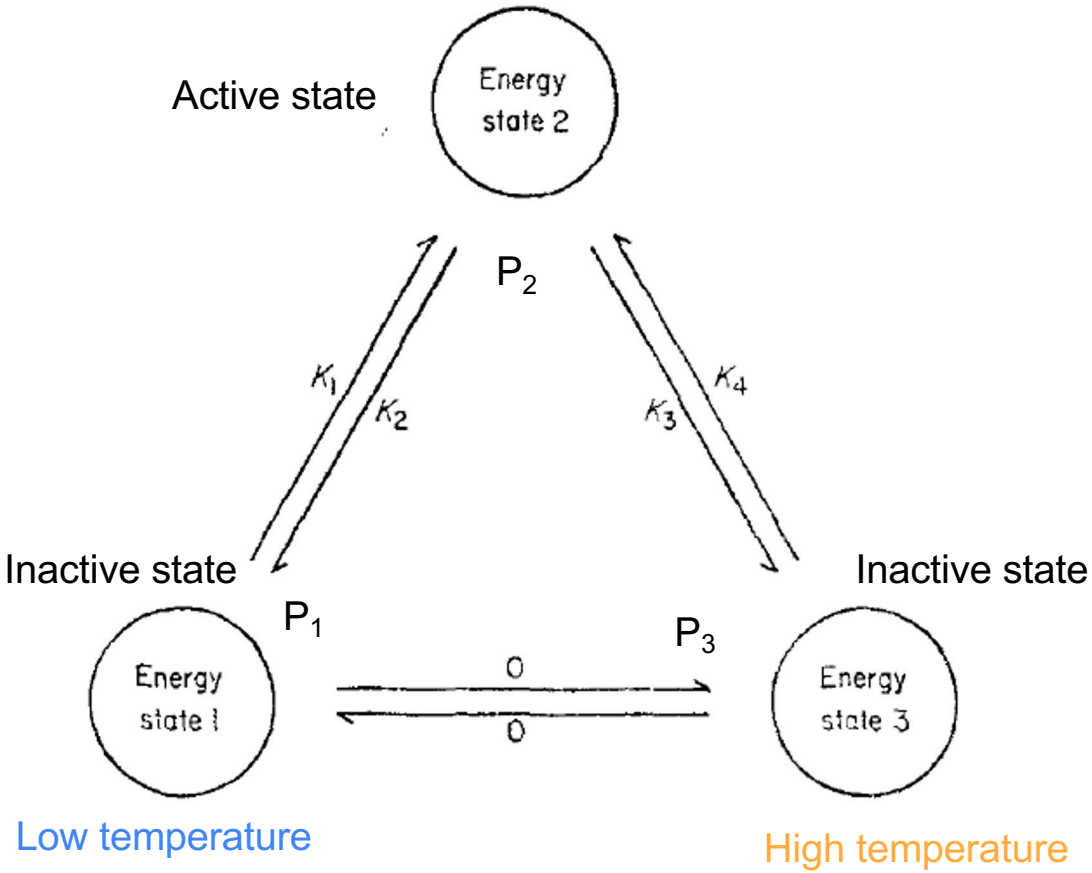


Microbial growth rate vs temperature, Buchanan and Fulmer (1930)

- Degree-day model, starting from Candolle (1855), Reibisch (1902)
- **E. Coli growth rate, Johnson and Lewin (1946).** (Control enzyme hypothesis: one enzyme bottlenecks the (growth) rate.)

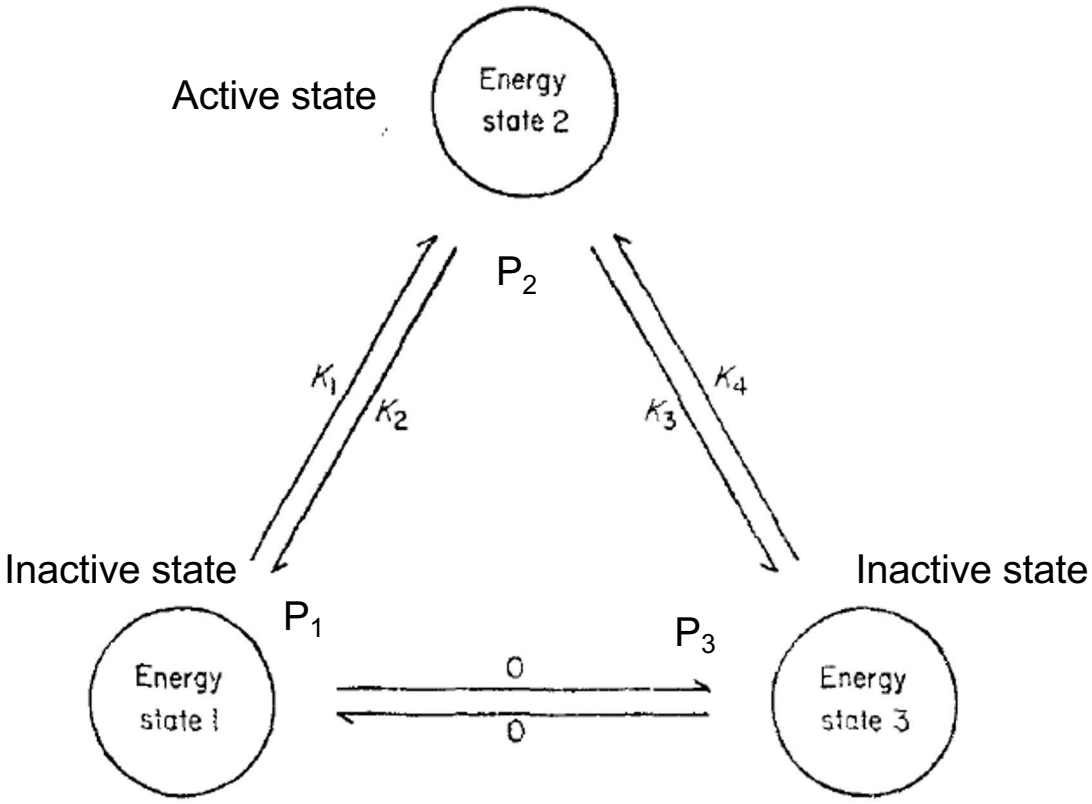


Milestone work by Peter Sharpe and Don deMichele (1977)



- Control enzyme hypothesis
- Enzymes exist in two types and three states.
- Rates proportional to the active enzyme amounts.

Milestone work by Peter Sharpe and Don deMichele (1977)



- Control enzyme hypothesis
- Enzymes exist in two types and three states.
- Rates proportional to the active enzyme amounts.

$$\frac{dP_1(t)}{dt} = -k_1 P_1(t) + k_2 P_2(t).$$

$$\frac{dP_2(t)}{dt} = k_1 P_1(t) - (k_2 + k_3) P_2(t) + k_4 P_3(t)$$

$$\frac{dP_3(t)}{dt} = k_3 P_2(t) - k_4 P_3(t).$$

k_is are by Eyring's transition state theory

$$k = \frac{k_B T}{h} e^{-\Delta G^{\ddagger}/RT}$$

Probability in active state

$$\frac{dP_1(t)}{dt} = -k_1 P_1(t) + k_2 P_2(t).$$

$$\frac{dP_2(t)}{dt} = k_1 P_1(t) - (k_2 + k_3) P_2(t) + k_4 P_3(t)$$

$$\frac{dP_3(t)}{dt} = k_3 P_2(t) - k_4 P_3(t).$$

Steady-state approximation: $dP_i/dt=0$

$$P_2 = 1/\left[1 + e^{(\Delta S_L - \Delta H_L/T)/R} + e^{(\Delta S_H - \Delta H_H/T)/R}\right]$$

Application to enzyme catalysis

Active state $$P_2 = 1/[1+e^{(\Delta S_L - \Delta H_L/T)/R} + e^{(\Delta S_H - \Delta H_H/T)/R}]$$

Enzyme kinetics

$$S + \varepsilon_f \underset{k'_{-1}}{\overset{k'_1}{\rightleftharpoons}} C \xrightarrow{k'_2} P + \varepsilon_f$$

$$\varepsilon_f + C = \varepsilon_a$$

Assuming no-substrate limitation

$$\frac{dP}{dt} = R_D = \frac{k'_1 S \varepsilon_c k'_2 P_2}{k'_1 S + k'_{-1} + k'_2}.$$

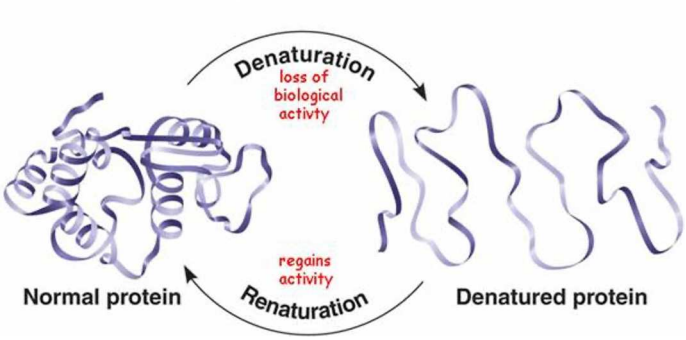
$$k'_1 S \gg k'_{-1} + k'_2$$

Combined with transition state theory

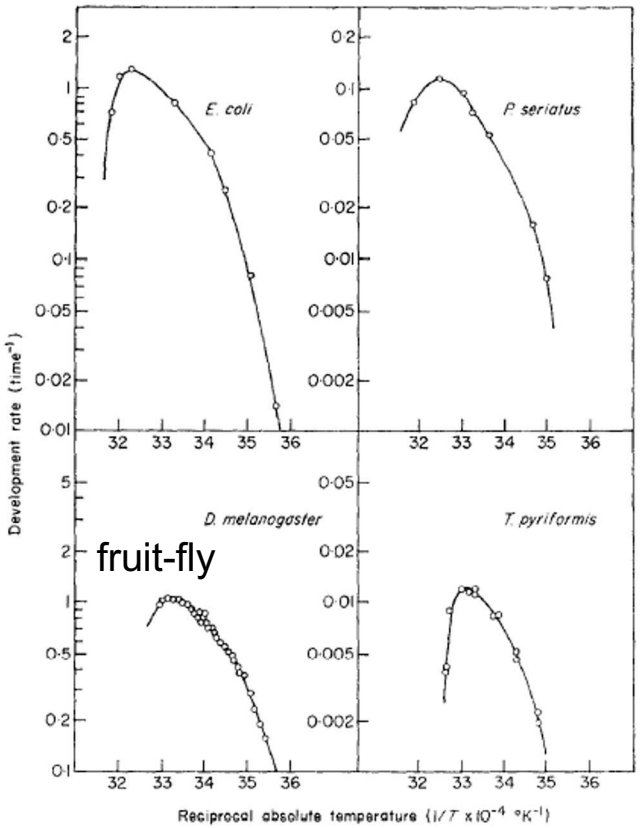
$$R_D = \frac{\varepsilon_c \dfrac{KT}{h} e^{(\Delta S_A^{\ddagger} - \Delta H_A^{\ddagger}/T)/R}}{1 + e^{(\Delta S_L - \Delta H_L/T)/R} + e^{(\Delta S_H - \Delta H_H/T)/R}}$$

Empirical observations supporting the Sharpe-deMichele theory's key assumption: thermally reversible denaturation

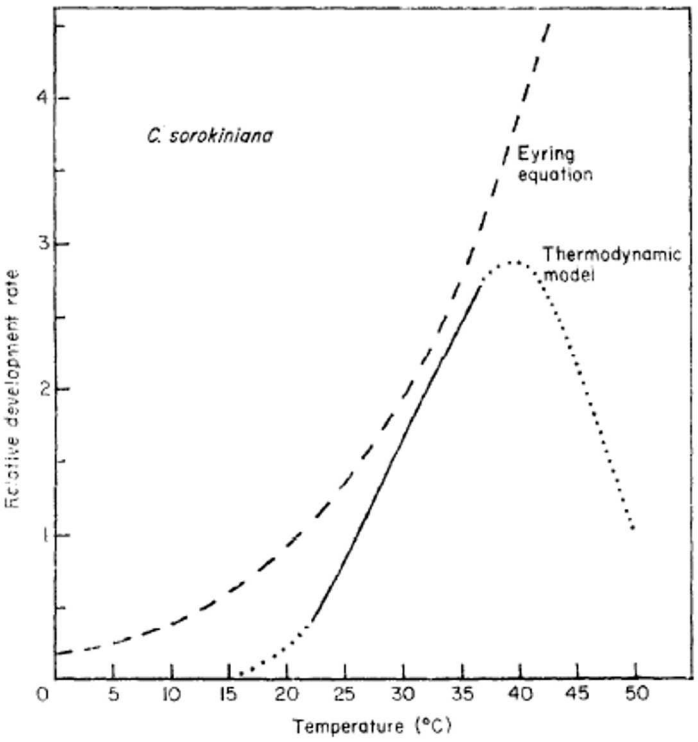
- Full range reversible inactivation, occurs independent of substrate presence: Sizer (1943), Northrup (1939).
- Cold inactivation, Scrutton and Utter (1965), Bergersen (1971), Huang and Cabib (173).
- Hot inactivation, Alexandrov (1964).



Performance of the Sharpe-deMichele theory



Development rate of Poikilotherm organisms.



Comparison to the Eyring model

Recap of the Sharpe-deMichele theory

Active state $$P_2 = 1/[1+e^{(\Delta S_L - \Delta H_L/T)/R} + e^{(\Delta S_H - \Delta H_H/T)/R}]$$

Enzyme kinetics

$$S + \varepsilon_f \underset{k'_{-1}}{\overset{k'_1}{\rightleftharpoons}} C \xrightarrow{k'_2} P + \varepsilon_f$$

$$\varepsilon_f + C = \varepsilon_a$$

Assuming no-substrate limitation

$$\frac{dP}{dt} = R_D = \frac{k'_1 S \varepsilon_c k'_2 P_2}{k'_1 S + k'_{-1} + k'_2}.$$

$$k'_1 S \gg k'_{-1} + k'_2$$

Combined with transition state theory

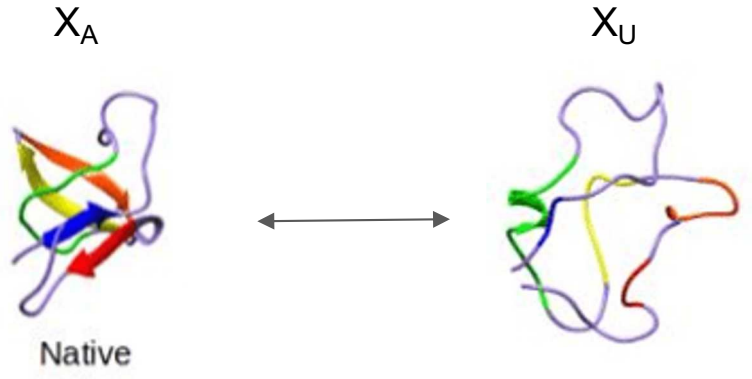
$$R_D = \frac{\varepsilon_c \dfrac{KT}{h} e^{(\Delta S_A^{\ddagger} - \Delta H_A^{\ddagger}/T)/R}}{1 + e^{(\Delta S_L - \Delta H_L/T)/R} + e^{(\Delta S_H - \Delta H_H/T)/R}}$$

Theory based on protein physics/ thermodynamics

Starting from 1930s:
e.g., Edsall, 1935.

Frank, 1945

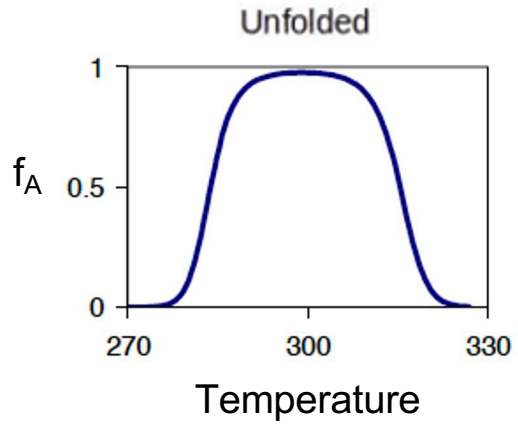
Murphy et al., 1990



Boltzmann law: $X_U=X_A\exp(-\Delta G/RT)$,
where $\Delta G=\Delta H-T\Delta S=C_P[(T-T_H)+T\ln(T_S/T)]$

Fraction of active proteins:
$f_A=X_A/(X_A+X_U)=1/(1+\exp(-\Delta G/RT))$

$$\Delta G = \Delta C_p \left[(T - T_H) - T ln \left(\frac{T}{T_S}\right)\right]$$



The chemical kinetics theory-1

$$E_a + S \underset{k_1^-}{\overset{k_1^+}{\leftrightarrow}} E_aS \xrightarrow{v_{max}} E_a + P,$$

E_a: active enzymes
S: substrates
E_aS: enzyme-substrate complex
P: products

$$F = v_{max}\frac{E_{aT}S}{K+S}.$$

$$v_{max} = v_{max,0}f_v(T) = v_{max,0}\left(\frac{T}{T_0}\right)exp\left(-\frac{\Delta G_V}{RT}\left(1-\frac{T}{T_0}\right)\right),$$

$K=(v_{max}+k_1^-)/k_1^+ \propto \exp(-\Delta H_K/RT)$

$k_1^+ \propto$ Diffusivity ~ $(T/T_0)\exp(-\Delta H^*/RT)$

k_1^- also follows the transition state theory, and often assumes $k_1^- << v_{max}$.

The chemical kinetics theory-2 & MMRT

$$F = \upsilon_{max,0} \frac{f_v(T) f_E(T) E_T S}{K_0 f_K(T) + S},$$

f_v(T): Eyring type function
f_K(T): Arrhenius type function
f_E(T): Partition function with positive heat capacity

The chemical kinetics theory-2 & MMRT

$$F = v_{max,0} \frac{f_v(T) f_E(T) E_T S}{K_0 f_K(T) + S},$$

f_v(T): Eyring type function
f_K(T): Arrhenius type function
f_E(T): Partition function with positive heat capacity

$$k = \frac{k_B T}{h} e^{-\Delta G^{\ddagger}/RT}$$

The chemical kinetics theory-2 & MMRT

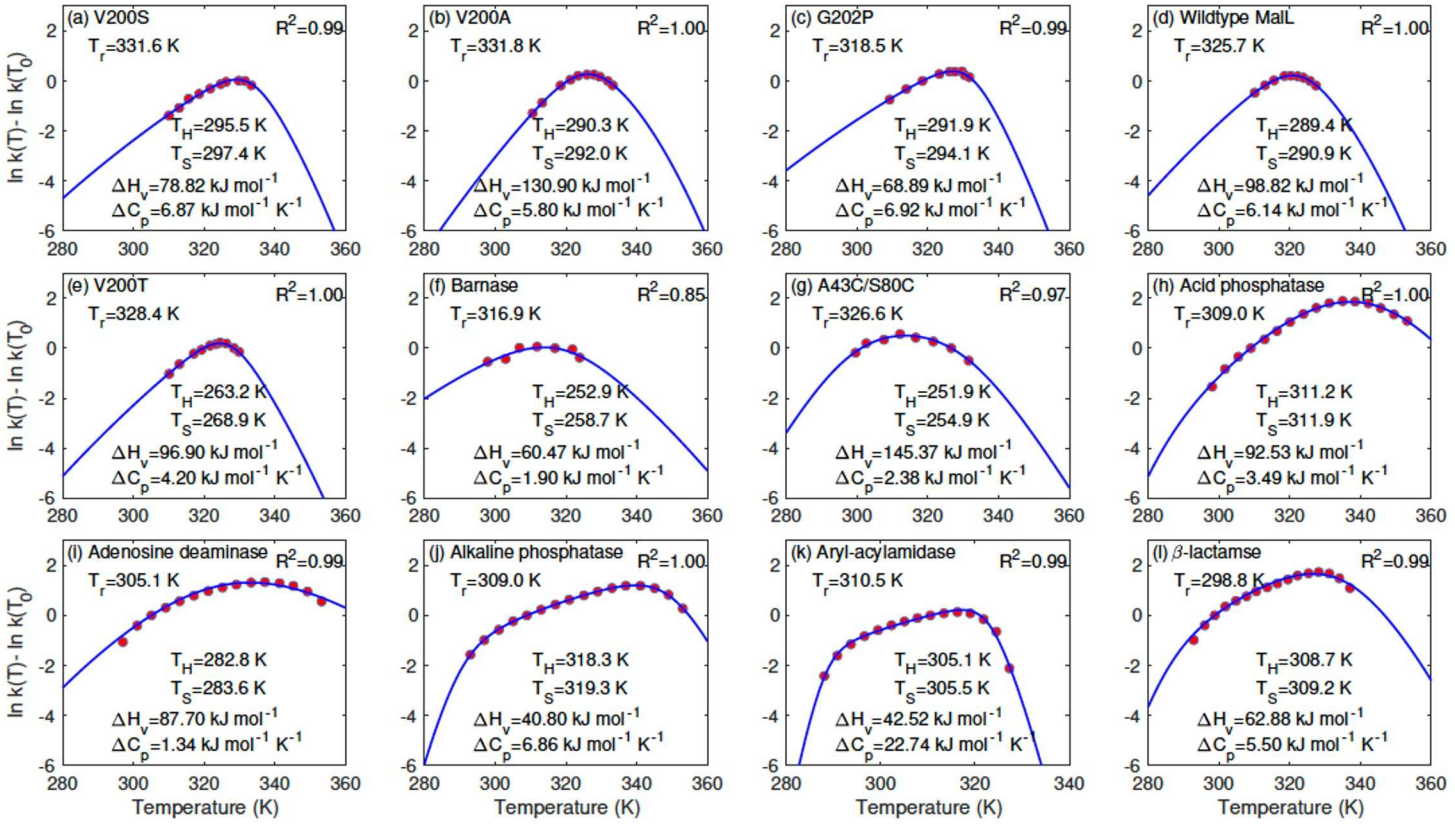
$$F = v_{max,0} \frac{f_v(T) f_E(T) E_T S}{K_0 f_K(T) + S},$$

Enzyme assay (Peterson et al., 2004, also Hobbs et al., 2013): **Substrate concentrations were maintained at 10 times K to minimize the effects of any possible increases in K with temperature.**

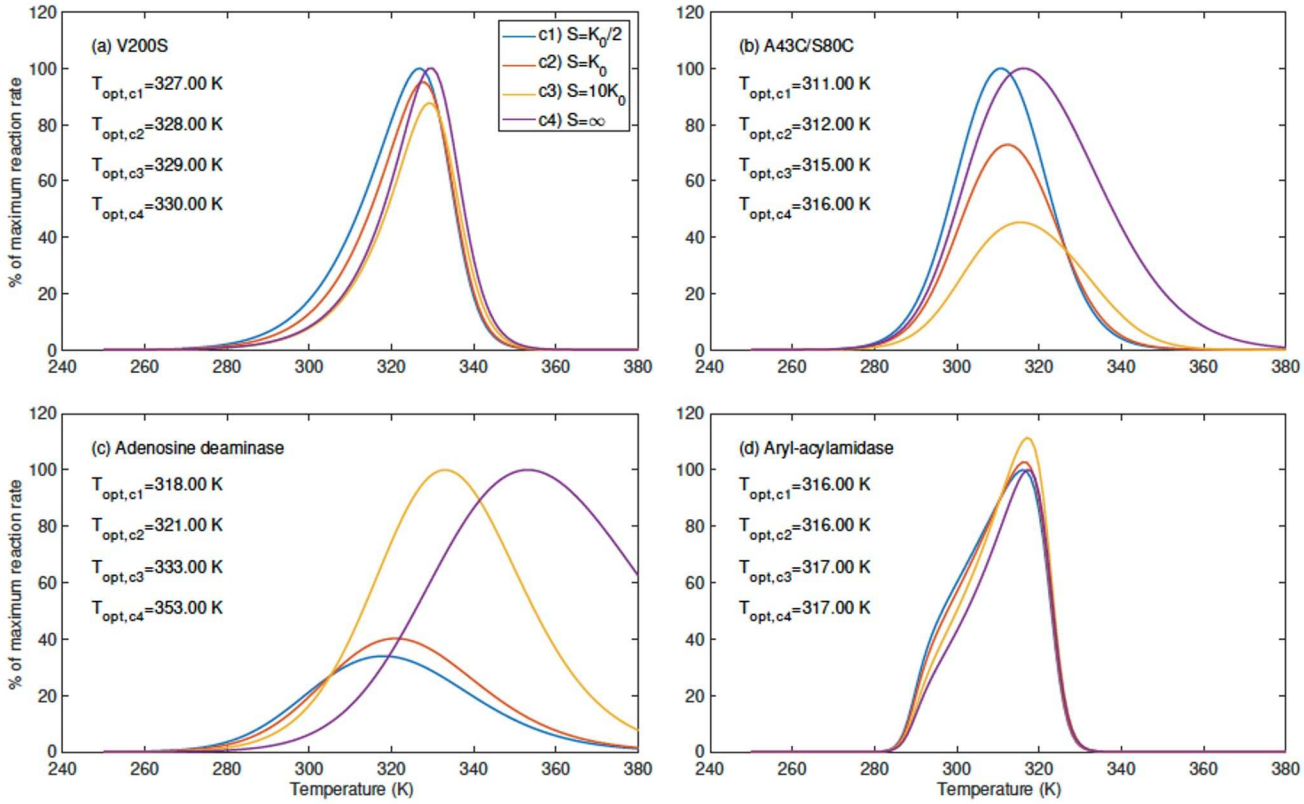
Hobbs et al (2013) also used substrate of ~2xK, but assume dK/dT=0.

$$\ln k(T) = \ln v_{max,0} + \ln(1 + S/K_0) + \ln\left(\frac{f_v(T) f_E(T)}{f_K(T) + S/K_0}\right).$$

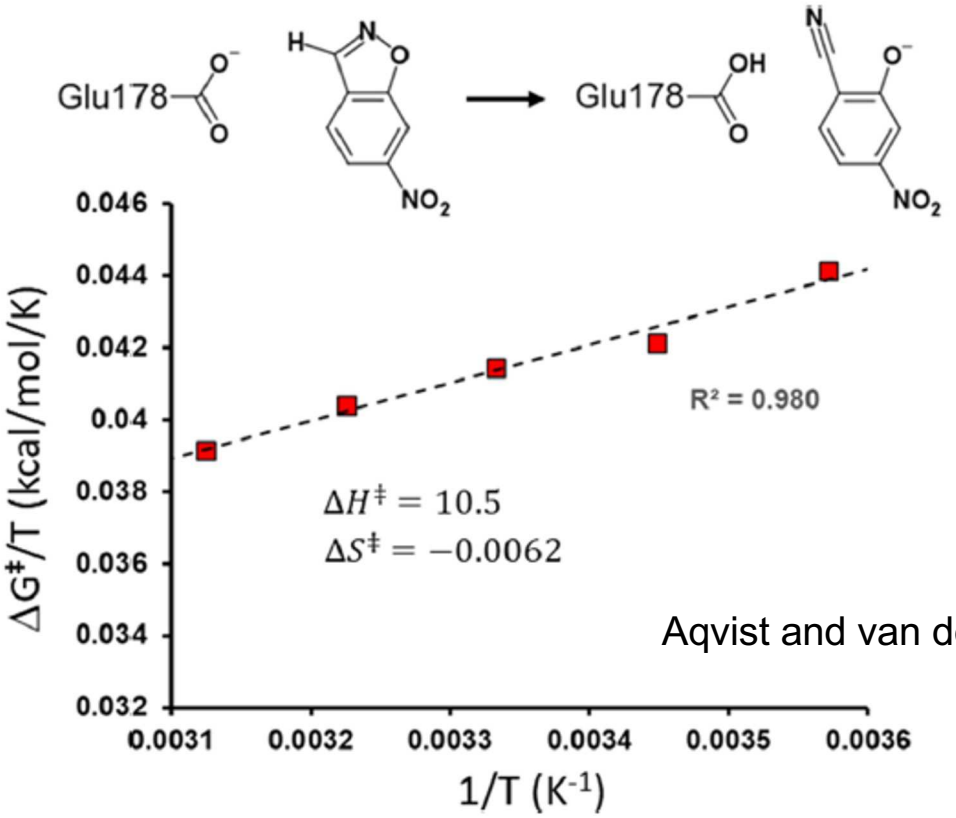
The chemical kinetics theory performance for empirical data



Optimal temperature is higher at more substrate



Molecular dynamics simulations: The Gibbs free energy of activation is a linear function of temperature

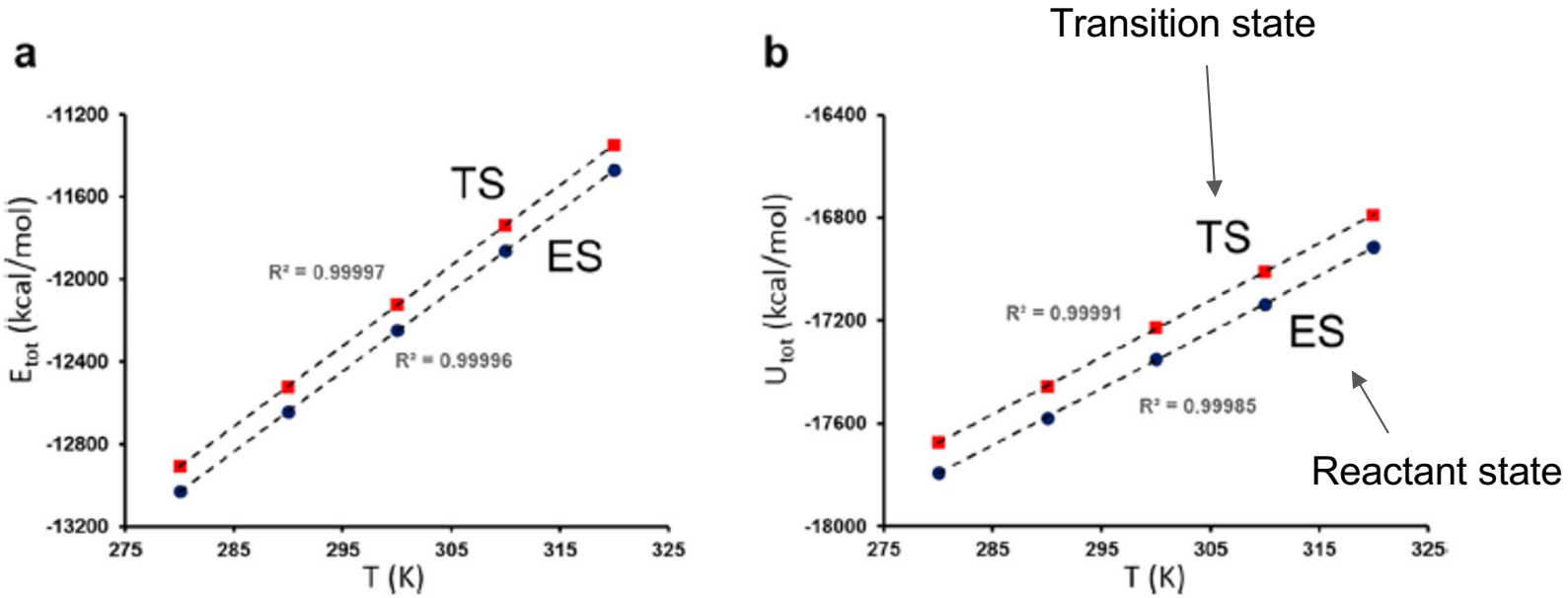


$$\ln(k) = \ln\left(\frac{k_B T}{h}\right) - \frac{[\Delta H^{\ddagger}_{T_0} + \Delta C^{\ddagger}_p (T - T_0)]}{RT} + \frac{[\Delta S^{\ddagger}_{T_0} + \Delta C^{\ddagger}_p \ln(T/T_0)]}{R}$$

Should be zero

Aqvist and van der Ent (2022)

Molecular dynamics simulations: The component energy are a linear function of temperature



Aqvist and van der Ent (2022)

Summary: issues with MMRT

$$\ln(k) = \ln\left(\frac{k_B T}{h}\right) - \frac{[\Delta H^{\ddagger}_{T_0} + \Delta C^{\ddagger}_p (T - T_0)]}{RT} + \frac{[\Delta S^{\ddagger}_{T_0} + \Delta C^{\ddagger}_p \ln(T/T_0)]}{R}$$

$$F_\infty = v_{max,0} E_T f_v(T) f_E(T) = r_0 f_v(T) f_E(T)$$

is equally good (Ratkowsky et al., 2005)

- MMRT ignores the existence of reversible denaturation (well-established fact from protein physics).
- It introduces (negative) heat capacity into reaction rate.
- The original parametric fitting ignores the substrate dependence of the temperature dependence of K.
- It predicts fixed optimal temperature.

Overall summary of MMRT

- It is just another empirical approximation that slightly better than Q_{10}.
- Mechanistic modeling should follow the chemical kinetics theory.

Implication for interpreting temperature dependence

- Optimal temperature is not a biochemical/phenological trait.
- Acclimation/adaptation can occur through changes in substrate availability/type.

  E.g. for plant respiration, the model should represent internal carbon and nutrient reserve, and use them to drive the growth. The NPP or GPP driven algorithm for growth could be problematic for simulating adaptation to warming.

- Implications for CO2 fertilization effect?

- The optimal temperature of leaf photosynthesis is a combination of dark respiration and CO2-dependent light reaction.

Strategies to better characterize the temperature dependence

- For enzyme assays: measure the temperature response at multiple substrate levels (at least 2).
- For general reaction rate (e.g. respiration): measure substrate availability together with the temperature response.
- For organisms, calibrating/benchmarking with as much data as possible.
- Proteomic data may help (Ghosh and Dill, 2010).

Philosophical lessons learned

- When empirical data are insufficient to reveal a model's problem, it is helpful to reanalyze the theoretical foundations.
- Known first principles are by and large consistent with each other, it is important to theoretically benchmark many of the ideas we put into the models.

Questions?

Scaling problem with MMRT

$$E_a + S \underset{k_1^-}{\overset{k_1^+}{\leftrightarrow}} E_aS \xrightarrow{v_{max}} E_a + P,$$

Quasi-steady-state-approximation:

$k_1^+ E_a{*}S = v_{max} E_a S$ (*)

- With MMRT, v_{max} approaches zero at high temperatures, so does the right hand side of equation (*).k_1^+ increases with temperature, unless E_a decreases with temperature accordingly with v_{ma}, equation (*) will breakdown. If E_a decreases with temperature, and MMRT describes temperature response of the reaction rate, then enzyme E_a should undergo reversible denaturation.
- The other way is that k_1^- is comparable to v_{max}, but then it predicts ecological advantage of having non-effective substrate binding/capture rate, contrary to the existences of fast growers.

Scaling problem with MMRT

$$E_a + S \underset{k_1^-}{\overset{k_1^+}{\leftrightarrow}} E_aS \xrightarrow{v_{max}} E_a + P,$$

Quasi-steady-state-approximation:

$k_1^+ E_a{*}S = v_{max} E_a S$ (*)