

Universal Design Principle to Enhance Enzymatic Activity using the Substrate Affinity

Hideshi Ooka^{1*}, Yoko Chiba¹, Ryuhei Nakamura^{1,2}

¹ Biofunctional Catalyst Research Team, Center for Sustainable Resource Science, 2-1 Hirosawa, Wako, Saitama 351-0198 Japan

² Earth-Life Science Institute (ELSI), Tokyo Institute of Technology, 2-12-IE-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan

Corresponding Author: Hideshi Ooka

Email: hideshi.ooka@riken.jp

Author Contributions: H.O. performed the mathematical calculations, numerical simulations, and bioinformatic analysis. The research was conceived by all coauthors. All coauthors wrote the manuscript and have approved of the final version.

Competing Interest Statement: The authors declare no competing interests.

ABSTRACT

Design principles to improve enzymatic activity are essential to promote energy-material conversion using biological systems. For more than a century, the Michaelis-Menten equation has provided a fundamental framework of enzymatic activity. However, there is still no concrete guideline on how the parameters should be optimized to enhance enzymatic activity. Here, we demonstrate that tuning the Michaelis-Menten constant (K_m) to the substrate concentration (S) maximizes enzymatic activity. This guideline ($K_m = S$) was obtained by applying the Brønsted (Bell)-Evans-Polanyi (BEP) principle of heterogeneous catalysis to the Michaelis-Menten equation, and is robust even with mechanistic deviations such as reverse reactions and inhibition. Furthermore, K_m and S are consistent to within an order of magnitude over an experimental dataset of approximately 1000 wild-type enzymes, suggesting that even natural selection follows this principle. The concept of an optimum K_m offers the first quantitative guideline towards improving enzymatic activity which can be used for highthroughput enzyme screening.

MAIN TEXT

Introduction

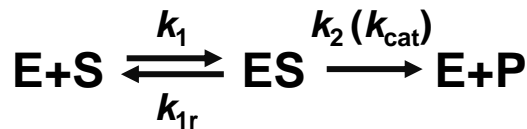
Enzymes are responsible for catalysis in virtually all biological systems,^[1,2] and a rational framework to improve their activity is critical to promote biotechnological applications. Since the early 20th century, a reaction mechanism where the enzyme first binds to the substrate ($E+S \rightarrow ES$) before releasing the product ($ES \rightarrow P$) has been used as the conceptual basis to understand enzyme catalysis (Scheme 1).^[3-6] The reaction rate of this mechanism is given by the Michaelis-Menten equation:

$$v = \frac{k_2 S}{K_m + S} E_T \quad (1)$$

Here, the reaction rate (v) is expressed as a function of a rate constant (k_2), the Michaelis-Menten constant (K_m), and the substrate (S) and enzyme (E_T) concentrations. K_m can be interpreted as a quasi-equilibrium constant for the formation of the enzyme-substrate complex, defined as:

$$K_m \equiv \frac{k_{1r} + k_2}{k_1} \quad (2)$$

with rate constants defined based on the mechanism shown in Scheme 1. k_2 is the rate constant for releasing the product from the enzyme-substrate complex ($ES \rightarrow P$), routinely expressed as k_{cat} in the enzymology literature. These parameters are experimentally accessible by fitting the theoretical rate law (Eq. (1)) with experimental data^[7-10] and are subsequently registered in databases such as BRENDA^[11] and Sabio-RK.^[12] In principle, the accumulated data may help rationalize and improve the activity of existing enzymes.



Scheme 1. The standard reaction mechanism of enzyme catalysis.

However, there is no concrete understanding on how these parameters influence enzymatic activity. For example, increasing k_2 may enhance activity according to Eq. (1), or diminish it due to a larger K_m (Eq. (2)).^[13] Thus, the mutual dependence between k_2 and K_m complicates their influence on the enzymatic activity (v), hindering the rational design of enzymes towards biotechnological applications such as the synthesis of commodity chemicals,^[14] antibiotics,^[15] or pharmaceuticals,^[16] increasing the nutritional content of crops,^[17] and restoring the environment.^[18]

In this study, we analyzed the Michaelis-Menten equation to clarify the relationship between the enzyme-substrate affinity (K_m) and the activity (v). The key ingredient of our mathematical analysis is the Brønsted (Bell)-Evans-Polanyi (BEP) relationship,^[19-23] which models the activation barrier as a function of the driving force. This is a well-known concept in heterogeneous catalysis, and in conjunction with the Arrhenius equation,^[24] can be used to evaluate the mutual dependence between k_2 and K_m to quantitatively. This allowed us to calculate the optimum value of K_m required to maximize enzymatic activity (v), a finding which is supported by our bioinformatic analysis of approximately 1000 wild-type enzymes.

Results

Construction of the Thermodynamic Model

In principle, an ideal enzyme with low K_m and large k_2 can be realized if both k_1 and k_2 are increased simultaneously. However, this is physically unrealistic, because the driving force which can be allocated to k_1 and k_2 is limited by the free energy change of the entire reaction. Within this thermodynamic context, maximum activity is realized by optimizing the distribution of the total driving force between the first ($E + S \rightarrow ES$) and second ($ES \rightarrow P$) steps shown in Scheme 1. To quantitatively evaluate the relationship between the driving force and the activity, we have used the BEP relationship^[19-23] to convert driving forces (ΔG) into activation barriers (E_a), and the Arrhenius^[24] equation to convert activation barriers to rate constants.

The thermodynamic model which served as the basis of our calculations is shown in Fig. 1. In a classical Michaelis-Menten reaction, the enzyme and substrate first form an enzyme-substrate complex ($E + S \rightarrow ES$) before producing the product in the second step ($ES \rightarrow P$). This mechanism is conceptually similar to reactions that occur on a heterogeneous catalyst surface, where the substrate molecule first binds to the catalyst surface before being converted into the product.^[19-23] The Gibbs free energies for the formation of the enzyme-substrate complex and the product are denoted as ΔG_1 and ΔG_2 , respectively. By definition, their sum must equal the total free energy change of the reaction ΔG_T :

$$\Delta G_T = \Delta G_1 + \Delta G_2 \quad (3)$$

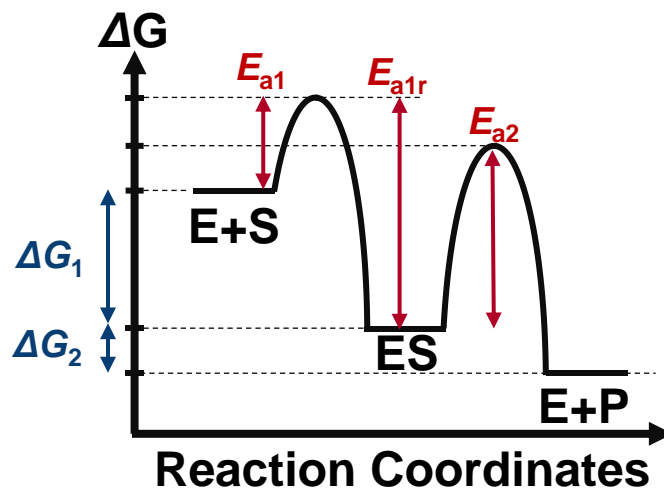


Figure 1. The free energy landscape corresponding to the mechanism shown in Scheme 1. Each reaction in the mechanism is labeled by its corresponding rate constant. The free energy landscape below indicates the free energy changes ($\Delta G_1, \Delta G_2$) and activation barriers (E_{a1}, E_{a1r}, E_{a2}).

From these thermodynamic constraints, we will use the BEP relationship^[19-23] to obtain activation barriers (E_a), and then the Arrhenius^[24] equation to obtain rate constants, which ultimately yields quantitative insight on the relationship between k_1 , k_2 , and K_m . Based on the BEP relationship, the activation barrier corresponding to k_1 can be written as:

$$E_{a1} = E_{a1}^0 + \alpha_1 \Delta G_1 \quad (4)$$

where E_{a1}^0 represent the activation barriers when the elementary reaction is in equilibrium ($\Delta G_1 = 0$). They are positive constants which reflect the inherent favorability of this elementary step. α_1 is a positive constant coefficient which indicates the sensitivity of the activation barrier with respect to the driving force. Recently, Kari et al have shown that fungal cellulases indeed satisfy such linear free energy relationships between the activation barrier and the driving force.^[25] Next, activation barriers can be converted to rate constants based on the Arrhenius equation^[24] as follows:

$$k_1 = A_1 \exp \frac{-E_{a1}}{RT} \quad (5)$$

Here, A_1 is a pre-exponential factor, and R and T are the gas constant and absolute temperature, respectively. Using Eqs. (4) and (5), k_1 can be expressed as:

$$\begin{aligned} k_1 &= k_1^0 \exp \frac{-\alpha_1 \Delta G_1}{RT} \\ &= k_1^0 g_1^{-\alpha_1} \end{aligned} \quad (6)$$

where $k_1^0 \equiv A_1 \exp \frac{-E_{a1}^0}{RT}$ and $g_1 \equiv \exp \frac{\Delta G_1}{RT}$ were used to aggregate factors independent and dependent on the driving force, respectively (see Supporting Information, Appendix 1 for details). k_{1r} and k_2 can also be written similarly as:

$$k_{1r} = k_1^0 g_1^{\alpha_{1r}} = k_1^0 g_1^{1-\alpha_1} \quad (7)$$

$$k_2 = k_2^0 g_2^{-\alpha_2} = k_2^0 \left(\frac{g_1}{g_T} \right)^{\alpha_2} \quad (8)$$

using notations similar to those defined for k_1 (See Appendices 2 and 3 for details). Substituting these rate constants into Eq. (2) yields the following expression for K_m :

$$\begin{aligned} K_m &\equiv \frac{k_{1r} + k_2}{k_1} \\ &= g_1(1 + K) \end{aligned} \quad (9)$$

where K was defined as $K \equiv \frac{k_2^0 g_1^{\alpha_1 + \alpha_2 - 1}}{k_1^0 g_T^{\alpha_2}}$. Finally, based on Eqs. (8) and (9), the enzymatic activity (v) can be expressed as:

$$\begin{aligned} v &= \frac{k_2 S}{K_m + S} E_T \\ &= \frac{k_2^0 g_1^{\alpha_2} g_T^{-\alpha_2} S}{g_1(1 + K) + S} E_T \end{aligned} \quad (10)$$

To illustrate how Eq. (10) captures the tradeoff relationship between k_2 and K_m , numerical simulations were performed (Fig. 2A). Hereafter, all simulations will be performed at $\alpha_1 = \alpha_{1r} = \alpha_2 = 0.5$, which is a common assumption used to make baseline models in heterogeneous catalysis.^[22,26-28] Physically, this means that when the driving force of an elementary reaction is increased by 1 kJ/mol, its activation barrier decreases by 0.5 kJ/mol. In reality, typical experimental values of α range between 0.3 and 0.7 for artificial catalysts,^[29-31] and the experimental value reported for cellulases is 0.74.^[25] Therefore, the influence of α deviating from 0.5 will be discussed in detail in Fig. 5D.

Fig. 2A shows three possible thermodynamic landscapes for a reaction with a total driving force of $\Delta G_T = -40$ kJ/mol. This parameter was chosen as a representative value based on the fact that the ΔG_T of typical biochemical reactions is between $-80 \sim +40$ kJ/mol.^[32,33] Similar calculations with different values of ΔG_T can be found in Figs. S1-S3. When the first reaction is thermodynamically favorable compared to the second ($\Delta G_1 < \Delta G_2$; Fig. 2A, black lines), the activity increases rapidly from low substrate concentrations (Fig. 2B, solid black line), consistent with the small K_m value. However, an enzyme with a small K_m suffers from a small k_2 value, which is evident from the saturating behavior at $S > 1 \mu\text{M}$. Increasing the driving force of the second step (blue and red lines) leads to a larger k_2 and thus higher activity at large S values ($S > 1 \mu\text{M}$) compared to the enzyme shown in black. At the same time, however, K_m increases, which decreases the enzymatic activity at low S ($S < 1 \mu\text{M}$). The difference in activity at low and high substrate concentrations occurs because the substrate participates in only the first elementary step. For example, even if $k_1 < k_2$ ($\Delta G_1 > \Delta G_2$), the rates of the two forward reactions ($k_1 E \cdot S$ and $k_2(ES)$) can be matched if the substrate concentration (S) is sufficiently large. However, at low substrate concentrations, a small k_1 can no longer be compensated, resulting in the first step being rate-limiting. For this reason, a large k_1 is necessary to increase the enzymatic activity at low substrate concentrations, whereas a large k_2 is more desirable when the substrate concentration is sufficient. The balance in tradeoff changes when the rates of the two forward reactions are equal ($k_1 E \cdot S = k_2(ES) \leftrightarrow S = \frac{k_2(ES)}{k_1 E}$). As the optimum values of k_1 and k_2 are dependent on the substrate concentration (S), the K_m value necessary to maximize the activity must also be dependent on (S).

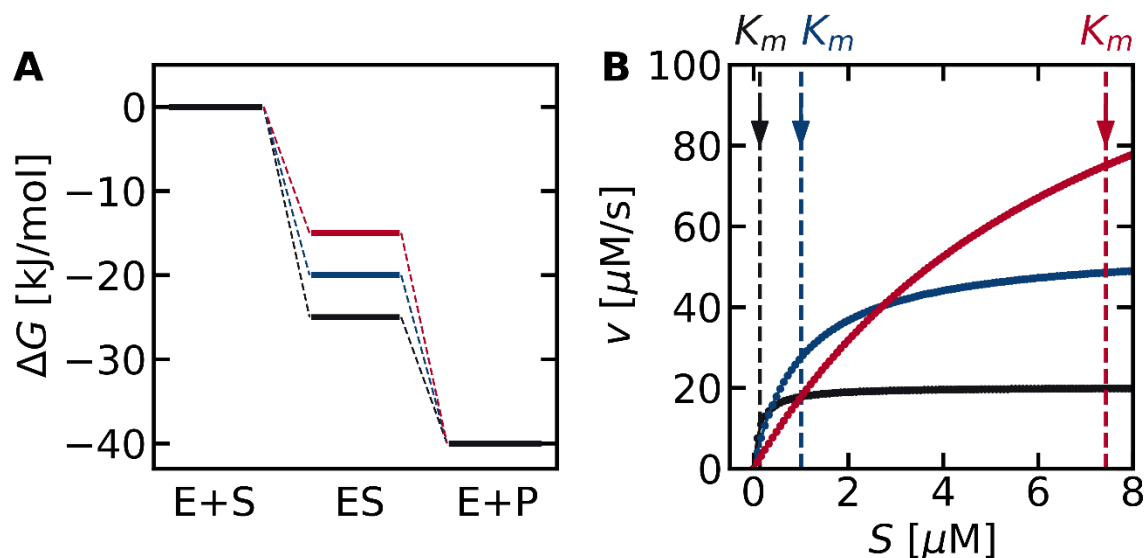


Figure 2. Thermodynamic landscapes (A) and their corresponding activity shown in the form of Michaelis-Menten plots (B). The K_m values are indicated as vertical dashed lines in (B). Increasing the driving force of the first step increases activity at low substrate concentrations but lowers the activity at high substrate concentrations. Therefore, the thermodynamic landscape of an optimum enzyme depends on the substrate concentration (S). The free energies of the enzyme-substrate complex (ΔG_1) were -25 , -20 , and -15 kJ/mol for the black, blue, and red lines, respectively, and that of the total reaction (ΔG_T) was -40 kJ/mol. All numerical simulations in this study were performed at $E_T = 1 \mu\text{M}$, $k_1^0 = k_2^0 = 1$ ($1/\mu\text{M/s}$ and $1/\text{s}$ units, respectively).

Analysis of the Activity – Driving Force Relationship

To directly illustrate the influence of driving force (ΔG_1 and ΔG_T) on enzymatic activity, we performed numerical simulations using Eq. (10) at various substrate concentrations (Fig. 3). At a substrate concentration of $0.1 \mu\text{M}$ (Fig. 3A), the region of highest enzymatic activity (orange) was observed in the bottom left region. It is reasonable for activity to be higher in the lower half of the panel, due to the more negative ΔG_T . A negative ΔG_1 is also beneficial for activity at a low substrate concentration ($S=0.1 \mu\text{M}$), leading to enzymatic activity being higher in the left half of the panel. At higher substrate concentrations, the overall color within each panel changed from blue to red, because a higher substrate concentration always increases activity (Figs. 3B-3D). At the same time, the ΔG_1 corresponding to maximum activity gradually shifted positively (black dashed lines). This finding is consistent with Fig. 2 which shows that a more positive ΔG_1 is desirable when the substrate concentration is increased. In each panel, the location with the highest activity at a given ΔG_T value is shown as a dashed black line. Notably, when the K_m value was calculated at the $(\Delta G_1, \Delta G_T)$ values under the dashed line using Eq. (9), the obtained value was always equal to the substrate concentration S in each panel. In other words, the dashed line is not only the ridge of the volcano plot, but also the contour line showing $K_m = S$. This suggests that the condition for maximizing enzymatic activity can be represented by $K_m = S$.

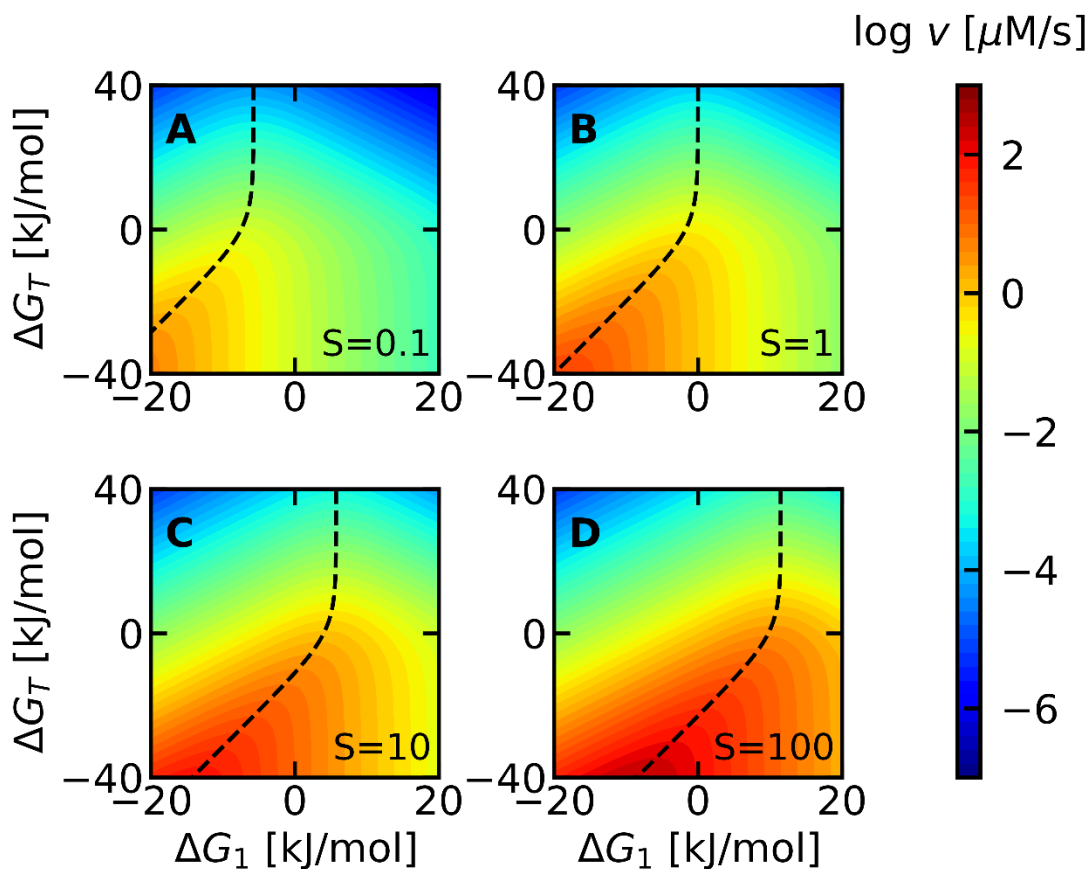


Figure 3. Enzymatic activity (v) plotted against ΔG_1 and ΔG_T based on Eq. (10). The substrate concentration (S) in each panel was (A) 10^{-1} , (B) 1, (C) 10, and (D) $10^2 \mu\text{M}$, as indicated in the bottom right of each panel. In all panels, the black dashed line corresponding to $K_m = S$ overlaps with the region with the highest enzyme activity.

To examine why $K_m = S$ leads to maximum activity, Eq. (10) was rearranged to give the following expression for the activity (v):

$$v = \frac{k_2^0 g_T^{-\alpha_2} S}{S g_1^{-\alpha_2} + g_1^{1-\alpha_2} + \frac{k_2^0 g_1^{\alpha_1}}{k_1^0 g_T^{\alpha_2}}} E_T \quad (11)$$

in which g_1 is only in the denominator. The derivative of the denominator, denoted as f is:

$$\frac{df}{dg_1} = -\alpha_2 g_1^{-(\alpha_2+1)} S + (1 - \alpha_2) g_1^{-\alpha_2} + \frac{k_2^0 \alpha_1}{k_1^0 g_T^{\alpha_2}} g_1^{\alpha_1-1} \quad (12)$$

To maximize the activity (v), f must be minimized which is realized at:

$$\frac{df}{dg_1} = 0 \leftrightarrow S = g_1 \left(\frac{1 - \alpha_2}{\alpha_2} + \frac{\alpha_1}{\alpha_2} K \right) \quad (13)$$

Considering that K_m is defined as $K_m \equiv g_1(1 + K)$ (Eq. (10), Eq. (13) yields a surprisingly simple formula for the condition of maximum activity when $\alpha_1 = \alpha_{1r} = \alpha_2 = 0.5$:

$$K_m = S \quad (14)$$

Eq. (14) provides the theoretical basis for why maximum activity was consistently observed along the contour line $K_m = S$ in Fig. 3: The combination of $(\Delta G_1, \Delta G_T)$ necessary to maximize activity guarantees $K_m = S$. This finding is further illustrated in Fig. 4, where the activity (v) is plotted as a function of K_m at different substrate concentrations. In all cases, maximum activity (v) is observed when the binding affinity (K_m) is equal to the substrate concentration (S). Thus, the derivations and simulations so far provide mathematical evidence that having a K_m value equal to the substrate concentration S guarantees maximal enzymatic activity as long as the enzyme follows the Michaelis-Menten mechanism (Scheme 1), and the rate constants follow the BEP relationship with $\alpha_1 = \alpha_{1r} = \alpha_2 = 0.5$.

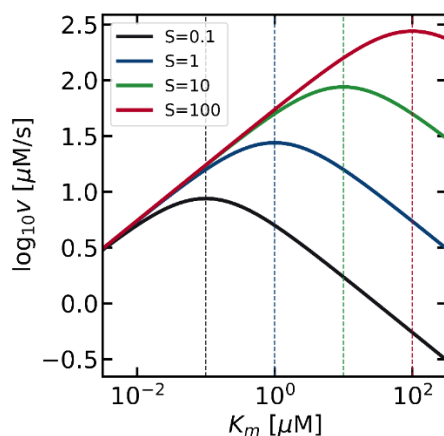


Figure 4. Volcano plots showing how the activity is expected to change with respect to the Michaelis-Menten constant (K_m). As the substrate concentration was increased from 10^{-1} μM (black) to 10^2 μM (red), the volcano plot shifted to the upper right. The apex is located at $K_m = S$, as indicated by the vertical dashed lines of the corresponding color. $\Delta G_T = -40$ kJ/mol and $k_1^0 = k_2^0 = 1$ [$1/\mu\text{M/s}$ and $1/\text{s}$ units, respectively] were used for the numerical simulations. Changing these values did not influence the conclusion that the activity is maximized when $K_m = S$, as shown in Fig. S4.

Robustness of the Theoretical Model

To confirm the robustness of our finding, we have performed numerical simulations by loosening each of the theoretical requirements. Deviation from the Michaelis-Menten mechanism (Scheme 1) are shown in Fig. 5A-C, and deviation of α values from 0.5 are shown in Fig. 5D. The possibility of reverse reactions ($P \rightarrow S$) or inhibition ($E + I \rightarrow EI$ or $ES + I \rightarrow ESI$) are common deviations from Michaelis-Menten kinetics.^[34] The net rate in the presence of a reverse reaction when the substrate and product are in equal concentrations ($S = P = 10 \mu\text{M}$) is shown in Fig. 5A. In terms of maximizing the activity in the forward direction ($S \rightarrow P$), the physically meaningful region is ($\Delta G_T < 0$), where the net reaction proceeds in the forward direction. Under this condition, the dashed line corresponding to $K_m = S$ and the solid line corresponding to the true maximum activity (forward minus reverse reaction rates) overlap almost completely, indicating that $K_m = S$ is a good guideline to enhance activity even in the presence of reverse reactions ($P \rightarrow S$).

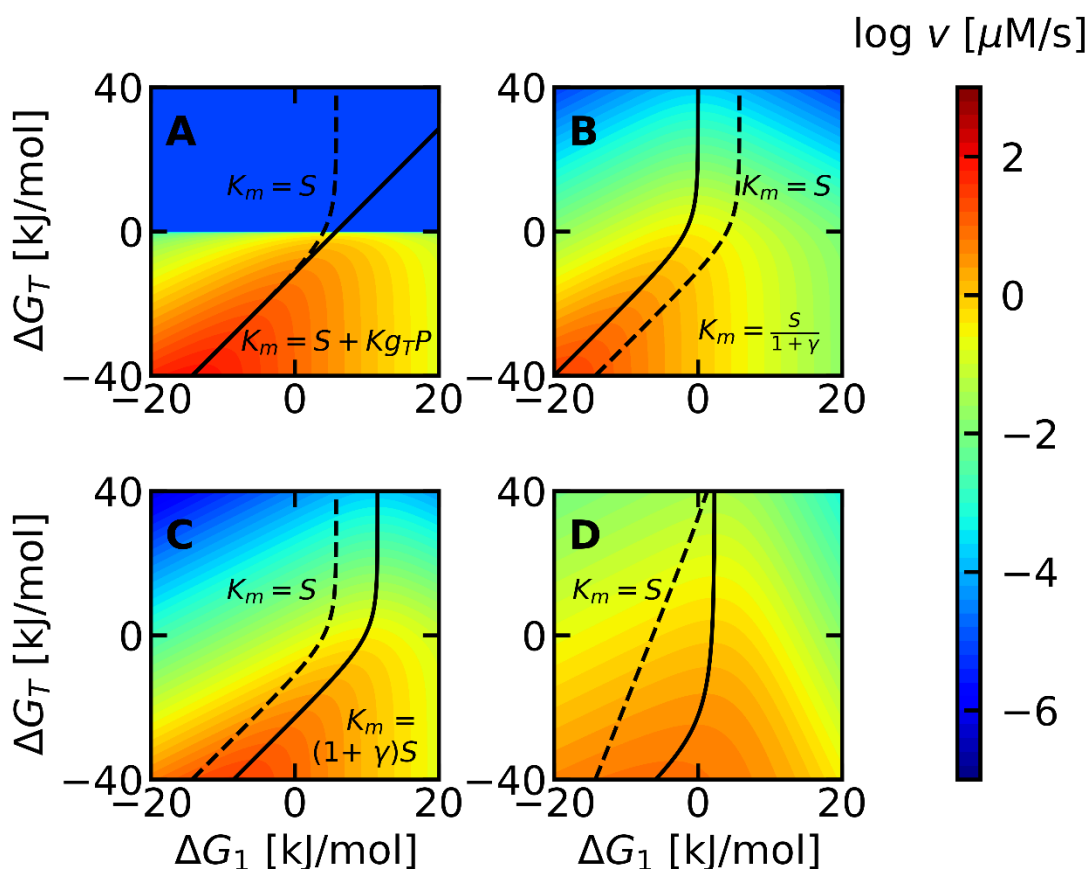


Figure 5. Influence of (A) Reverse reaction, (B) Competitive inhibition, (C) Uncompetitive inhibition, and (D) BEP coefficient, α on the optimal K_m . The dashed line corresponds to $K_m = S$, with $S = 10 \mu\text{M}$. The true optimum K_m for each mechanism is shown as a solid line along with its analytical equation (refer to the SI for the derivations). In panel A, the product concentration (P) was set to $10 \mu\text{M}$. The top half of (A) was colored at an arbitrarily low activity because the reverse reaction is more favorable in this region. The large discrepancy between the dashed and solid lines at $\Delta G_T > 0$ is physically irrelevant, because the activity of the forward reaction cannot be discussed when the net reaction proceeds in the reverse direction. In panels B and C, the degree of inhibition ($\gamma \equiv I/K_i$) was set to 10. In panel D, the BEP coefficients were set to $\alpha_1 = \alpha_2 = 0.2$. No analytical solution was obtained for D.

Similar calculations for competitive and uncompetitive inhibition, where the inhibitor binds to either the free enzyme or the enzyme-substrate complex, are shown in Fig. 5B,C. The degree of inhibition ($\gamma \equiv \frac{I}{K_i}$), is determined by the inhibitor concentration (I) and the equilibrium constant of inhibition (K_i).^[34] Based on the experimental data of Park et al.,^[35] γ can range from 10^{-4} to 10^4 . As γ was less than 10 in approximately 80% of their data, $\gamma = 10$ was used here for the numerical simulations. Again, the optimal K_m (solid line) deviates only slightly from the dashed line ($K_m = S$), and both lines pass through the region of high activity (orange). The K_m values are approximately 1 order of magnitude apart between dashed and solid lines, yet there is only a 57 % difference in activity at a specific ΔG_T . This is much smaller than the scale of the entire diagram (10 orders of magnitude), suggesting that adjusting K_m to the substrate concentration S is a robust strategy to enhance the activity, even in the presence of inhibition. A detailed discussion on the parameter dependence (γ, S), as well as for other mechanisms such as substrate inhibition or allostericity can be found in Section 3 of the supporting information. The derivations for the equations of the true optimal K_m can also be found in the same section.

The influence of the second assumption ($\alpha_1 = \alpha_{1r} = \alpha_2 = 0.5$) is shown in Fig. 5D. As physical constraints require $\alpha_{1r} = 1 - \alpha_1$ (Appendix 2), only α_1 and α_2 are independent. In an extreme case where $\alpha_1 = \alpha_2 = 0.2$, the activity is markedly diminished because rate constants hardly change even if their driving force is increased. However, the dashed line still passes through the region of high activity, and the activity is still less than an order of magnitude away from the true optimum (solid lines). Taken together, these simulations confirm that $K_m = S$ is a robust theoretical guideline to enhance enzymatic activity.

Validation based on Experimental Data

Finally, to evaluate whether $K_m = S$ can rationalize enzymatic properties in nature, we have analyzed their relationship based on the experimental data from Park et al.^[35] The original data consisted of K_m values of wild-type enzymes obtained from BRENDA, and intracellular S values obtained from *Escherichia coli*, *Mus musculus*, and *Saccharomyces cerevisiae* cells, yielding a total of 1703 K_m - S combinations. This dataset was then classified based on the number of entries for each substrate, based on the expectation that a substrate which participates in many reactions is more likely to deviate from Michaelis-Menten kinetics. ATP is the most frequent substrate with 313 entries and is shown in black. Both the raw K_m and S values (Fig. 6A) and their relative distribution (Fig. 6B) shows that $S > K_m$ for ATP. The deviation from $K_m = S$ may be because the Michaelis-Menten mechanism, which is the basis of our mathematical analysis, does not consider scenarios where multiple reactions compete for the same substrate. The next subset shown in blue covers 410 entries and consists of 5 substrates which each appear more than 50 times: NAD⁺, NADH, NADP⁺, NADPH, and acetyl-CoA. These cofactors are less universal than ATP, and S is only slightly larger than K_m . The remaining 980 entries are shown in red. This subset contains 115 substrates such as carbon metabolites and amino acids and appear within the dataset 8 times on average. As the substrate becomes less universal, their K_m and S values become roughly consistent. In particular, the Gaussian distribution fitted to the red histogram (Fig. 6B) has a center at $\log_{10} S/K_m = 0.18$ and a standard deviation of 1.3, which is reasonable considering that influences from inhibitors or the BEP coefficient can change the optimum K_m by roughly an order of magnitude (Fig. 5). Thus, the dataset from wild-type enzymes supports the theoretical prediction that a Michaelis-Menten constant equivalent to the substrate concentration is favorable for the activity, especially when the substrate participates in fewer reactions and Michaelis-Menten kinetics becomes more accurate.

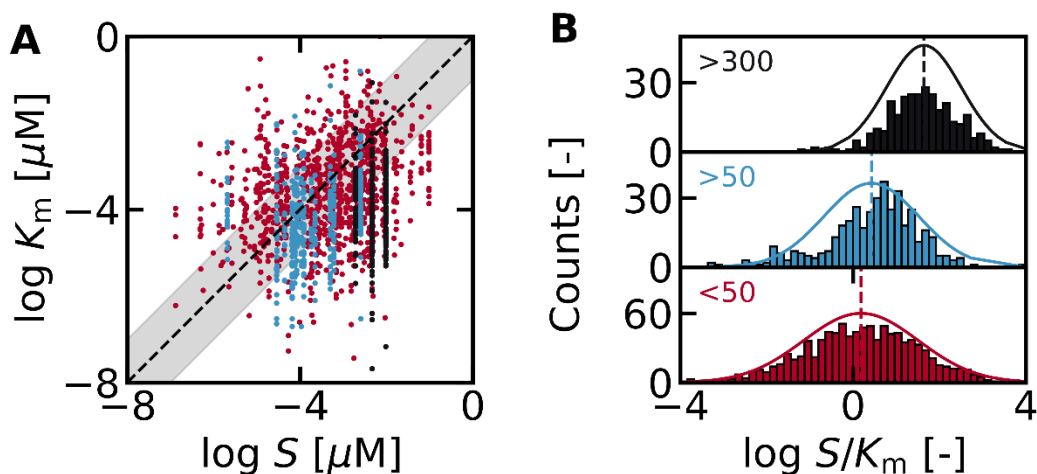


Figure 6. Relationship between K_m and S from the dataset reported by Park et al.^[35] The raw values of K_m and S are shown in (A), and their relative values are plotted in (B). Each entry of K_m and S was categorized based on the number of times the substrate appeared in the entire dataset. Black: > 300 (ATP), blue: > 50 (NAD⁺, NADH, NADP⁺, NADPH, and acetyl-CoA), red: < 50 (others). The number of entries was used as a proxy for the validity of the Michaelis-Menten mechanism of the specific substrate. The dashed line in (A) corresponds to $K_m = S$, and the shaded area shows a deviation of 1 order of magnitude.

Discussion

So far, various criteria^[13,34,36] such as large k_2 (k_{cat}), small K_m , or large k_2/K_m have been proposed to characterize enzymes with high activity, making it difficult to rationally evaluate or engineer the activity of an enzyme. The lack of a universal consensus is largely due to the mutual dependence between k_2 and K_m . As our theoretical model addresses this challenge directly and maximizes the activity within the thermodynamic constraints imposed by k_2 and K_m , we believe that $K_m = S$ is a criterion for high activity which is viable in a wider range of scenarios.

The idea that the Michaelis-Menten constant should be increased at higher substrate concentrations to maximize activity is consistent with the experimental work by Kari et al.^[37] who measured the activity of cellulases with different K_m . When the substrate concentration was increased 6 times, the K_m value of the most active enzyme increased approximately 2.4 times. Considering that K_m can change by roughly 6 orders of magnitude, the experimental trend supports our hypothesis $K_m = S$, especially when their experimental BEP coefficient of 0.74 is considered. The idea of the optimum binding affinity being dependent on the reaction condition and driving force is also consistent with recent theoretical models of heterogeneous catalysis.^[22,38-40]

As a corollary, our model which quantifies the relationship between K_m and k_2 immediately rationalizes the recently reported free energy relationship between them in cellulases.^[25] Namely, the relationship between K_m and k_2 can be written as:

$$K_m = (1 + K)g_T \left(\frac{k_2}{k_2^0} \right)^{1/\alpha_2}$$

$$\therefore \log k_2 = \alpha_2 \log K_m - \log(1 + K)g_T + \log k_2^0 \quad (15)$$

This equation shows that $\log k_2$ and $\log K_m$ are linearly correlated by a factor of α_2 , and provides a physical basis to the high linearity ($R^2 = 0.95$) observed for cellulases.^[25] The consistency between our theoretical model and previously accumulated experimental insight suggests that it may be possible to quantitatively rationalize enzymatic properties based on fundamental principles of physical chemistry.

Online Methods

The mathematical formulas were derived by hand, and the step-by-step derivations for the standard Michaelis-Menten mechanism are explained in the main text. The derivations in the presence of inhibition and allostericity are provided in the supporting information. Numerical simulations and bioinformatic analysis were performed using Python 3.9.12. The code used for the analysis can be found in the extended data or accessed directly at github:

https://github.com/HideshiOoka/SI_for_Publications.

Acknowledgments

H.O. gratefully acknowledges the support from the JST FOREST program (Grant Number JPMJFR213E, Japan). Y. C. is grateful for the support from the JST ACT-X program (Grant Number JPMJAX20BB, Japan).

References

- [1] Inigo M, Deja S, Burgess SC Ins and Outs of the TCA Cycle: The Central Role of Anaplerosis. *Annu. Rev. Nutr.* 41:19–47 (2021).
- [2] Sweetlove LJ, Beard KF, Nunes-Nesi A, Fernie AR, Ratcliffe RG Not Just a Circle: Flux Modes in the Plant TCA Cycle. *Trends Plant Sci.* 15:462–470 (2010).
- [3] Henri V Théorie Générale de l'Action de Quelques Diastases. *C. R. Hebd. Séances Acad. Sci.* 135:916–919 (1902).
- [4] Michaelis L, Menten ML Die Kinetik der Invertinwirkung. *Biochem. Z* 49:352 (1913).
- [5] Cornish-Bowden A The origins of enzyme kinetics. *FEBS Lett.* 587:2725–2730 (2013).
- [6] Johnson KA, Goody RS The Original Michaelis Constant: Translation of the 1913 Michaelis–Menten Paper. *Biochem.* 50:8264–8269 (2011).
- [7] Atkins G, Nimmo I A Comparison of Seven Methods for Fitting the Michaelis-Menten Equation. *Biochem. J.* 149:775–777 (1975).
- [8] Atkins G, Nimmo I Current Trends in the Estimation of Michaelis-Menten Parameters. *Anal. Biochem.* 104:1–9 (1980).
- [9] Sakoda M, Hiromi K Determination of the Best-Fit Values of Kinetic Parameters of the Michaelis-Menten Equation by the Method of Least Squares with the Taylor Expansion. *J. Biochem.* 80:547–555 (1976).
- [10] Chiba Y, et al. Mechanism for Folate-Independent Aldolase Reaction Catalyzed by Serine Hydroxymethyltransferase. *FEBS J.* 279:504–514 (2012).
- [11] Schomburg I, Hofmann O, Baensch C, Chang A, Schomburg D Enzyme Data and Metabolic Information: BRENDA, A Resource for Research in Biology, Biochemistry, and Medicine. *Gene Funct. Dis.* 1:109–118 (2000).
- [12] Rojas I, et al. SABIO-RK: A Database for Biochemical Reactions and their Kinetics. *BMC Syst. Biol.* 1:1–2 (2007).
- [13] Eisenthal R, Danson MJ, Hough DW Catalytic Efficiency and k_{cat}/K_M : A Useful Comparator? *Trends Biotechnol.* 25:247–249 (2007).
- [14] Kato Y, Inabe K, Hidese R, Kondo A, Hasunuma T Metabolomics-Based Engineering for Biofuel and Bio-based Chemical Production in Microalgae and Cyanobacteria: A Review. *Biores. Technol.* 344:126196 (2022).
- [15] Walsh CT Combinatorial Biosynthesis of Antibiotics: Challenges and Opportunities. *ChemBioChem* 3:124–134 (2002).
- [16] Sanchez S, Demain AL Metabolic Regulation of Fermentation Processes. *Enzyme Microb. Technol.* 31:895–906 (2002).
- [17] Britto DT, Kronzucker HJ Bioengineering Nitrogen Acquisition in Rice: Can Novel Initiatives in Rice Genomics and Physiology Contribute to Global Food Security? *BioEssays* 26:683–692 (2004).
- [18] Tripathi V, et al. Biotechnological Advances for Restoring Degraded Land for Sustainable Development. *Trends Biotechnol.* 35:847–859 (2017).
- [19] Ooka H, Huang J, Exner KS The Sabatier Principle in Electrocatalysis: Basics, Limitations, and Extensions. *Front. Energ. Res.* 9:155 (2021).
- [20] Bligaard T, et al. The Brønsted–Evans–Polanyi Relation and the Volcano Curve in Heterogeneous Catalysis. *J. Catal.* 224:206–217 (2004).
- [21] Logadottir A, et al. The Brønsted–Evans–Polanyi relation and the Volcano Plot for Ammonia Synthesis Over Transition Metal Catalysts. *J. Catal.* 197:229–231 (2001).
- [22] Ooka H, Nakamura R Shift of the Optimum Binding Energy at Higher Rates of Catalysis. *J. Phys. Chem. Lett.* 10:6706–6713 (2019).
- [23] Vojvodic A, et al. On the Behavior of Brønsted-Evans-Polanyi Relations for Transition Metal Oxides. *J. Chem. Phys.* 134:244509 (2011).
- [24] Laidler KJ The Development of the Arrhenius Equation. *J. Chem. Educ.* 61:494 (1984).
- [25] Kari J, et al. Physical Constraints and Functional Plasticity of Cellulases. *Nat. Commun.* 12:1–10 (2021).
- [26] Koper M Analysis of Electrocatalytic Reaction Schemes: Distinction Between Rate-Determining and Potential-Determining Steps. *J. Solid State Electrochem.* 17:339–344 (2013).

- [27] Koper M Volcano Activity Relationships for Proton-Coupled Electron Transfer Reactions in Electrocatalysis. *Top. Catal.* 58:1153–1158 (2015).
- [28] Ooka H, Wintzer ME, Nakamura R Non-Zero Binding Enhances Kinetics of Catalysis: Machine Learning Analysis on the Experimental Hydrogen Binding Energy of Platinum. *ACS Catal.* 11:6298–6303 (2021).
- [29] Ooka H, Yamaguchi A, Takashima T, Hashimoto K, Nakamura R Efficiency of Oxygen Evolution on Iridium Oxide Determined from the pH Dependence of Charge Accumulation. *J. Phys. Chem. C* 121:17873–17881 (2017).
- [30] Madhiri N, Finklea HO Potential-, pH-, and Isotope-Dependence of Proton-Coupled Electron Transfer of an Osmium Aquo Complex Attached to an Electrode. *Langmuir* 22:10643–10651 (2006).
- [31] Finklea HO Theory of Coupled Electron- Proton Transfer with Potential-Dependent Transfer Coefficients for Redox Couples Attached to Electrodes. *J. Phys. Chem. B* 105:8685–8693 (2001).
- [32] Weber AL Chemical Constraints Governing the Origin of Metabolism: The Thermodynamic Landscape of Carbon Group Transformations Under Mild Aqueous Conditions. *Orig. Life. Evol. Biosph.* 32:333–357 (2002).
- [33] Li X, Wu F, Qi F, Beard DA A Database of Thermodynamic Properties of the Reactions of Glycolysis, the Tricarboxylic Acid Cycle, and the Pentose Phosphate Pathway. *Database* (2011).
- [34] Atkins P, Paula JD (2011) *Physical Chemistry for the Life Sciences*. (Oxford University Press, USA).
- [35] Park JO, et al. Metabolite Concentrations, Fluxes and Free Energies Imply Efficient Enzyme Usage. *Nat. Chem. Biol.* 12:482–489 (2016).
- [36] Bauer C, Osman AM, Cercignani G, Gparkialluca N, Paolini M A Unified Theory of Enzyme Kinetics Based Upon the Systematic Analysis of the Variations of k_{cat} , K_M , and k_{cat}/K_M and the Relevant ΔG^0 Values-Possible Implications in Chemotherapy and Biotechnology. *Biochem. Pharmacol.* 61:1049–1055 (2001).
- [37] Kari J, et al. Sabatier Principle for Interfacial (Heterogeneous) Enzyme Catalysis. *ACS Catal.* 8:11966–11972 (2018).
- [38] Zeradjanin AR, Grote JP, Polymeros G, Mayrhofer KJ A Critical Review on Hydrogen Evolution Electrocatalysis: Re-exploring the Volcano-Relationship. *Electroanal.* 28:2256–2269 (2016).
- [39] Exner KS Beyond the Traditional Volcano Concept: Overpotential-Dependent Volcano Plots Exemplified by the Chlorine Evolution Reaction over Transition-Metal Oxides. *J. Phys. Chem. C* 123:16921–16928 (2019).
- [40] Exner KS Paradigm Change in Hydrogen Electrocatalysis: The Volcano's Apex is Located at Weak Bonding of the Reaction Intermediate. *Int. J. Hyd. Energ.* 45:27221–27229 (2020).