Revisiting tradeoffs in Rubisco kinetic parameters

Avi Flamholz[a], Noam Prywes[a,b], Uri Moran[b], Dan Davidi[b], Yinon Bar-On[b], Luke Oltrogge[a], Rui Alves[c,d], David Savage[a] and Ron Milo[b]

[a]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720;
[b]Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot 76100, Israel;
[c]Institute of Biomedical Research of Lleida IRBLleida, 25198, Lleida, Catalunya, Spain;
[d]Departament de Ciències Mèdiques Bàsiques, University of Lleida, 25198, Lleida, Catalunya, Spain;

Abstract

Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase (Rubisco) is not only the dominant enzyme in the biosphere, responsible for the vast majority of carbon fixation, but also one of the best characterized enzymes. Enhanced Rubisco catalysis is expected to increase crop yields, but a substantially improved enzyme has evaded bioengineers for decades. Based on correlations between Rubisco's kinetic parameters, it is widely posited that tradeoffs stemming from the catalytic mechanism strictly constrain Rubisco's maximum catalytic potential. Though compelling, the reasoning that established that view was based on data from only ≈20 organisms. Here we re-examine these tradeoffs with an expanded dataset including data from >200 organisms. We find that most correlations are substantially attenuated, with the inverse relationship between carboxylation k_{cat} and specificity $S_{C/O}$ being a key example. However, the correlation predicted by one tradeoff model is stronger and more significant in our expanded dataset. In this model, increased catalytic efficiency (k_{cat}/K_M) for carboxylation requires a similar increase in catalytic efficiency for the competing oxygenation reaction, evidenced here by a strong power-law correlation between those catalytic efficiencies. In contrast to previous work, our results imply that Rubisco evolution is constrained mostly by the physicochemical limits of O_2/CO_2 recognition, which should reframe efforts to understand and engineer this very central enzyme.

Significance

All plants, algae and cyanobacteria rely on the Calvin-Benson-Bassham (CBB) cycle for growth. Rubisco is the central enzyme of the CBB cycle and the most abundant enzyme on Earth. While it is often claimed that Rubisco is slow, its catalytic rate is just below the average enzyme. Yet it is surprising that Rubisco is not faster given its centrality and abundance. Previous analyses of Rubisco kinetic parameters raised doubts that the enzyme can be improved. Here we examine a new compendium of Rubisco kinetic parameters for evidence of proposed constraints on

Rubisco catalysis. Only one proposal is strongly supported by the data, which argues for re-evaluation of our understanding of one of the most impactful proteins in Earth's history.

Introduction

Rubisco is the primary carboxylase of the Calvin-Benson-Bassham (CBB) cycle - the carbon fixation cycle responsible for growth throughout the green lineage and many other autotrophic taxa - and the ultimate source of nearly all carbon atoms entering the biosphere (1). Typically, 20-30% of total soluble protein in C3 plant leaves is Rubisco (2). As Rubisco is so highly expressed and plants are the dominant constituents of planetary biomass (3), it is often said that Rubisco is the most abundant enzyme on Earth (1). As Rubisco is ancient (>2.5 billion years old), abundant, and remains central to contemporary biology, one might expect it to be exceptionally fast. But Rubisco is not fast (4–8). Typical carboxylation k_{cat} values range from 1-10 s^{-1} and all known Rubiscos are capable of reacting "wastefully" with O_2 in a process called oxygenation (Figure 1A-B). Improved Rubisco catalysis is expected to increase crop yields (9), but a substantially improved enzyme has evaded bioengineers for decades (10). The multiple evolution of CO_2 concentrating mechanisms, which ensure Rubisco operates near its maximum rate, also raises doubts about whether Rubisco catalysis can be strictly improved (11).

Rubisco is a notoriously complex enzyme that is inhibited by its five-carbon substrate, ribulose 1,5,-bisphosphate (RuBP) and requires a covalent modification for activation - carbamylation of an active-site lysine (12, 13). Moreover, Rubisco depends on chaperones for folding, assembly and catalysis (12, 14–16). Once folded and activated, all known Rubiscos catalyze both carboxylation and oxygenation of RuBP through a multistep mechanism (Figure 1A, S1). Both carboxylation and oxygenation of RuBP are energetically favorable, but only carboxylation is considered productive because it incorporates carbon from CO_2 into precursors that can generate biomass. Oxygenation is often portrayed as counterproductive as it occupies Rubisco active sites and yields a product (2-phosphoglycolate, 2PG) that is not part of the CBB cycle and must be recycled through metabolically-expensive photorespiration at a loss of carbon (17, 18). Despite the fact that many autotrophs depend on Rubisco carboxylation for growth, all known Rubiscos are relatively slow carboxylases and fail to exclude oxygenation (Figure 1A-B).

Over the decades since its discovery, various nomenclature has been used to describe the kinetics of Rubisco carboxylation and oxygenation. Here we use $k_{cat,C}$ and $k_{cat,O}$ to denote the turnover numbers (maximum per active site catalytic rates in units of s^{-1}) for carboxylation and oxygenation respectively. K_C and K_O denote the Michaelis constants (half-saturation concentrations in μM) for carboxylation and oxygenation. The specificity factor $S_{C/O} = (k_{cat,C}/K_C) / (k_{cat,O}/K_O)$ is a unitless measure of the relative preference for CO_2 over O_2 (Figure 1A-C).

The long-standing observation of correlations between the specificity factor $S_{C/O}$ and other Rubisco kinetic parameters (19–21) is often cited to motivate the notion that tradeoffs inherent to the catalytic mechanism strictly constrain Rubisco's carboxylation potential (4, 5, 7, 21). Indeed, we expect any physicochemical tradeoffs embedded in the Rubisco catalytic

mechanism to manifest as correlations between kinetic parameters (Figure 1C) precisely because Rubisco is so central to autotrophic life and has, therefore, experienced substantial selection pressure (22).

Two tradeoff models have been proposed to explain the observed correlations (4, 5). Although the proposed models are substantively different, both models imply limitations on the concurrent improvement of the maximum carboxylation rate ($k_{cat,C}$) and specificity (S) of natural Rubiscos (5). While these hypotheses appeal to physical and chemical intuition, they are based on data from only ≈20 organisms. Here we take advantage of the accumulation of new data - more than 200 Rubisco variants have been characterized since 2010 - to examine whether new data evidence the same correlations. We find that most of the previously-reported correlations between Rubisco kinetic parameters are substantially weakened by the addition of new data. However, correlation between the catalytic efficiency for carboxylation ($k_{cat,C}/K_C$) and the catalytic efficiency for oxygenation ($k_{cat,O}/K_O$) remains very strong and statistically significant (5), which may have important implications for efforts to understand and engineer Rubisco.

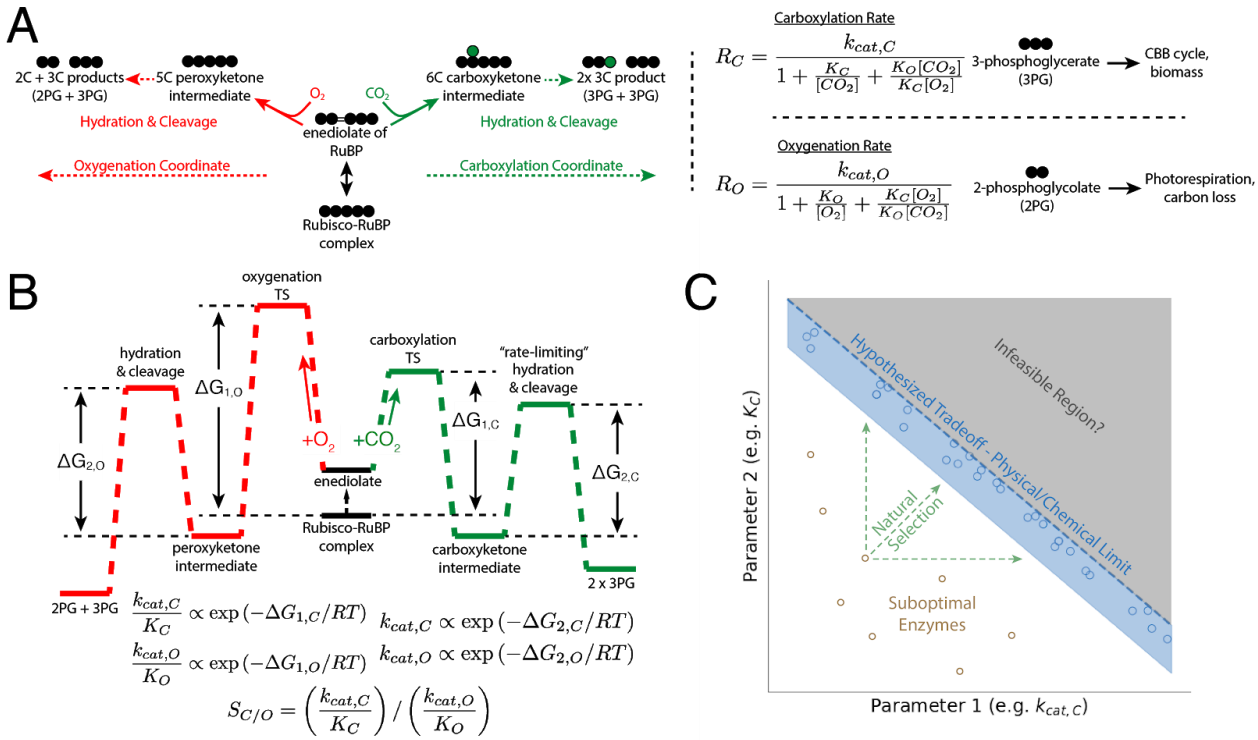

Figure 1: *Description of the catalytic mechanism of Rubisco. The "middle-out" diagram in Panel A shows the ordered mechanisms of carboxylation and oxygenation. Circles represent carbon atoms. RuBP is first isomerized to an enediolate before carboxylation or oxygenation can occur. Addition of CO_2 or O_2 to the enediolate of RuBP are considered irreversible as are the subsequent hydration and cleavage steps of carboxylation and oxygenation arms. Carboxylation displays effective Michaelis-Menten kinetics (maximum catalytic rate $k_{cat,C}$, half-maximum CO_2 concentration $K_M = K_C$) with competitive inhibition by O_2 (half-maximum inhibitory O_2 concentration $K_i = K_O$). Carboxylation results in net addition of carbon to the five-carbon RuBP, which produces two 3PG molecules. 3PG is part of the CBB cycle and can therefore be used to continue the cycle and produce biomass. Oxygenation also displays effective*

Michaelis-Menten kinetics ($k_{cat,O}$, $K_M = K_O$, half-max inhibitory CO_2 concentration $K_I = K_C$). Oxygenation of RuBP produces one 3PG and one 2PG. 2PG is not part of the CBB cycle and must be recycled through photorespiration to avoid total loss of two carbons. Rates of carboxylation (R_C) and oxygenation (R_O) can be calculated from kinetic parameters and the CO_2 and O_2 concentrations. The reaction coordinate diagram in panel (B) mirrors panel A and describes Rubisco carboxylation and oxygenation as a function of two "effective" barriers as in (5). The first effective barrier describes enolization and gas addition while the second describes hydration and bond cleavage. Given standard assumptions (described in SI), the respective catalytic efficiencies (k_{cat}/K_M) are related to the height of the first effective barrier while the k_{cat}s are related to the second. The first barrier to oxygenation is drawn higher than for carboxylation because Rubisco oxygenation is typically much slower than carboxylation. The net reactions of RuBP carboxylation and oxygenation are both quite thermodynamically favorable ($\Delta_r G'^m \approx -28$ kJ/mol and -520 kJ/mol respectively (23)). As kinetic parameters are linearly related to the log of effective energy barriers, energetic tradeoffs should manifest as linear correlations in a log-log plot of kinetic parameters (C). As Rubisco is central to photoautotrophic growth, we expect that natural selection has pushed the enzyme towards the upper limits of its catalytic capacity - i.e. towards the blue shaded region.

An extended dataset of Rubisco kinetic parameters

To augment existing data, we collected literature data on ≈250 Rubiscos including representatives of clades and physiologies that had been poorly represented in earlier datasets e.g. diatoms, ferns, CAM plants and anaerobic bacteria (Figure 2A). We collected kinetic parameters associated with carboxylation and oxygenation - S, K_C, $k_{cat,C}$, K_O and $k_{cat,O}$ - as well as measurements of the RuBP Michaelis constant (half-maximum RuBP concentration, K_{RuBP}) and experimental uncertainty for all values where available. All data considered were measured at 25°C and near pH 8 to ensure that measured values are comparable (Methods).

The resulting dataset contains Rubisco kinetic parameters from a total of 286 distinct species including 319 distinct $S_{C/O}$ values, 275 $k_{cat,C}$ values, 310 K_C values, 198 $k_{cat,O}$ values and 256 K_O values (Figure 2B). In 198 cases there was sufficient data to calculate catalytic efficiencies for carboxylation ($k_{cat,C}/K_C$) and oxygenation ($k_{cat,O}/K_O$, Methods). Though the data include measurements of some Form II, III and II/III Rubiscos, they remain highly focused on the Form I Rubiscos found in cyanobacteria, diatoms, algae and higher plants, which make up > 95% of the dataset (Figure 2B). As such, we focus primarily on the kinetic parameters of Form I Rubiscos here.

Rubisco kinetic parameters display very narrow dynamic range, with multiplicative variation (the standard deviation in log scale, denoted σ^*) being well below one order-of-magnitude for all parameters (Figure 2C). As compared to other enzymes for which multiple k_{cat} measurements are available, Rubisco displays extremely low variation in $k_{cat,C}$ ($\sigma^* = 0.4$, Figure S4). Specificity $S_{C/O}$ displays the least variation ($\sigma^* = 0.2$) of all parameters, though this may be due in part to overrepresentation of C3 plants in the dataset, which occupy a narrow range of $S_{C/O} \approx 80$-120. Nonetheless, measurements of $S_{C/O}$ for Form I and Form II enzymes are clearly distinct, with values ranging from ≈ 40-160 for Form Is and 7-15 for Form IIs (Figure 2C, SI).

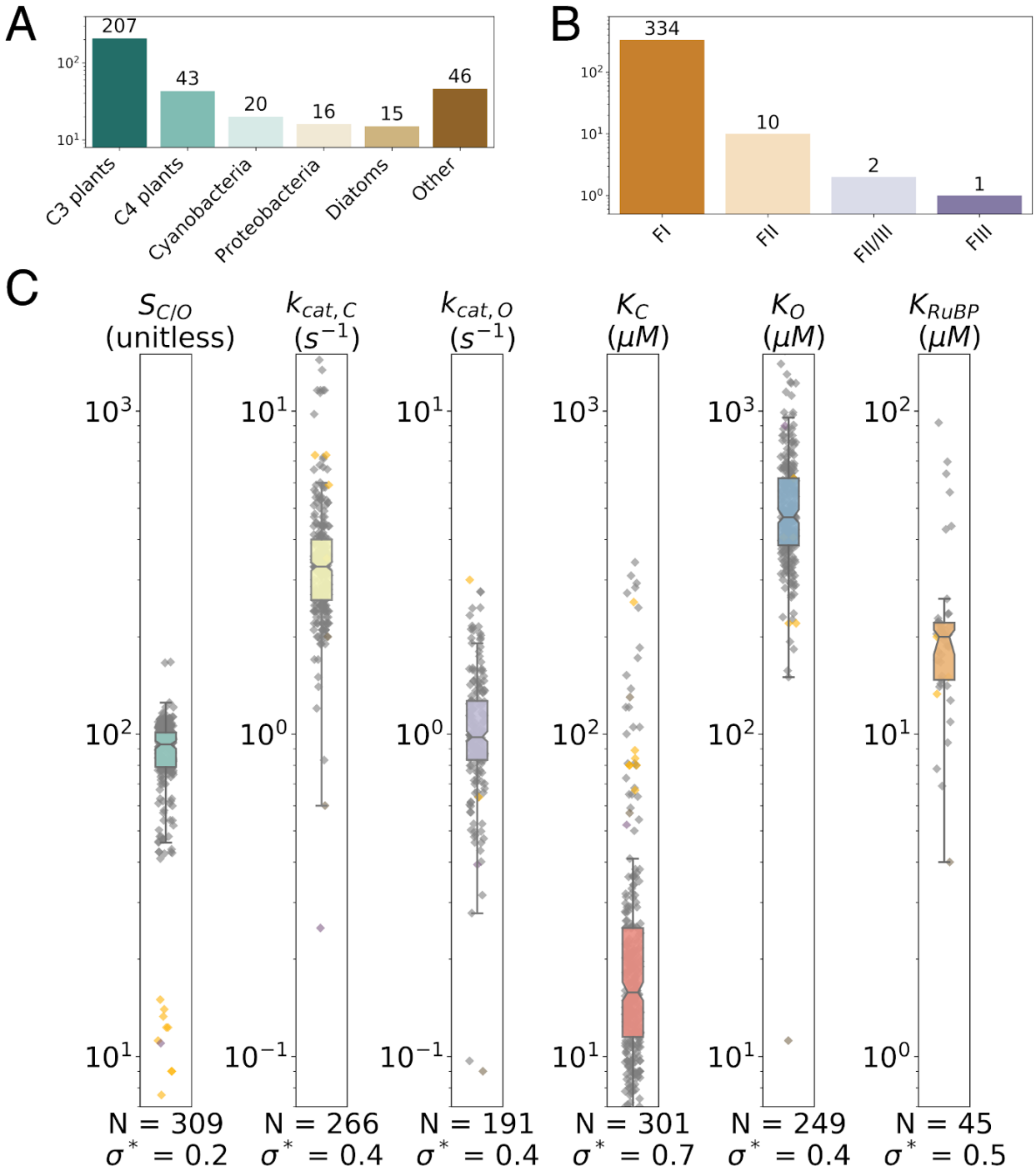

Figure 2: *Summary of the extended dataset. We collected measurements of Rubisco kinetic parameters from a variety of organisms (A) representing four classes of Rubisco isoforms (B). The bulk of data represent Form I enzymes from green lineage organisms (A-B). We collected 200-300 values for each of the various kinetic parameters, with the exception of the RuBP Michaelis constant (K_{RuBP}), which is less often measured. As shown in panel C, the assembled kinetic parameters display relatively narrow dynamic range. The box-plot and grey points describe the distribution of Form I Rubiscos while Form II Rubiscos are in yellow. The log-scale standard deviation of Form I data is reported as σ^*. We find that σ^**

< 1 in all cases, meaning that a single standard deviation of all parameters varies over less than one order-of-magnitude.

Reanalysis of correlations between kinetic parameters of Form I Rubiscos

As in (4, 5) we performed a correlation analysis to investigate relationships between Rubisco kinetic parameters. $S_{C/O}$, $k_{cat,C}$, K_C and other kinetic parameters are mathematically related to the microscopic rate constants of the Rubisco mechanism (SI). Given common assumptions, this multi-step mechanism can be mathematically simplified so that log of measured kinetic parameters are proportional to effective transition state barriers (Figure 1B, SI). Therefore we expect to observe strong power-law (log-linear) correlations between pairs of kinetic parameters when when three conditions are met: (I) the energy barriers associated with each parameter trade-off with each other - e.g. if lowering the effective barrier to CO2 addition ($\Delta G_{1,C}$) requires that the barrier to O2 addition ($\Delta G_{1,O}$) to decrease as well (Figure 1B-C); (II) those tradeoffs affect the rate of net carboxylation by Rubisco by affecting either carboxylation or oxygenation rates appreciably; and (III) the selection pressure imposed on Rubisco evolution was sufficient to reach limits imposed by this tradeoff (as diagrammed in Figure 1D). As Rubisco is the central enzyme of photoautotrophic growth, we assume here that it evolved under strong selection. As such, we interpret weakened correlations as calling into question whether the proposed tradeoffs strongly constrain Rubisco evolution.

We also note that some level of correlation is expected because measured kinetic parameters (e.g. $k_{cat,C}$ and K_C) are mathematically interrelated through the microscopic mechanism of Rubisco as it is commonly understood (SI). For example, when we derive expressions for $k_{cat,C}$ and K_C from the Rubisco mechanism, they share common factors that could drive correlation even in the absence of any tradeoff (SI). Similarly, $S_{C/O}$ is defined as $(k_{cat,C}/K_C) / (k_{cat,O}/K_O)$ and might correlate with $k_{cat,C}$ for this reason. Yet because $S_{C/O}$ has physiological importance as it determines the ratio of carboxylation to oxygenation rates (R_C/R_O) in the limit of low CO_2 and O_2 (SI), high $S_{C/O}$ might be independently selected for. Because we cannot predict a priori which parameters will correlate, we examine log-scale correlations between all pairs of parameters, as shown in Figure 3.

Correlations between $k_{cat,C}$ and $S_{C/O}$ as well as $k_{cat,C}$ and K_C were previously highlighted to support the hypothesis that Rubisco evolution is constrained by the enzyme's catalytic mechanism (4, 5). However, these correlations are substantially attenuated by the addition of new data. In general the set of kinetic parameters do not appear to fall along a one-dimensional curve (Figures 3 and 4). Figure 4A replots the focal correlation of Tcherkez et al. (4) - $k_{cat,C}$ vs. $S_{C/O}$ - and shows that these parameters are much less correlated in the extended dataset ($R < 0.6$ and extremely sensitive to outliers). Similarly, Figure 4B shows that the focal correlation of Savir et al. (5) - $k_{cat,C}$ vs K_C - is weakened, with $R \approx 0.66$ as opposed to $R \approx 0.9$. By far the strongest correlation is between the catalytic efficiencies for carboxylation and oxygenation,

$k_{cat,C}/K_C$ and $k_{cat,O}/K_O$ (R = 0.93, Figure 3). We discuss possible explanations for this very strong correlation in detail below.

Principal components analysis (PCA) of Rubisco kinetic parameters was previously used to interrogate constraints on Rubisco evolution. It was argued that Rubisco adaptation is constrained to a one-dimensional landscape because the first principal component (PC1) explained > 90% of the variance in Rubisco kinetics. In a one-dimensional landscape model all kinetic parameters are tightly interrelated so that changing one (e.g. $k_{cat,C}$) forces all others to assume predetermined values (5). However, the extended dataset is not well-approximated as one-dimensional. While the orientation of PC1 is not substantially altered by the addition of tenfold more measurements, it now explains ≈70% instead of >90% of the variance in Rubisco kinetics (5). More than 2 principal components are required to explain >90% of the variation in the extended dataset (SI). We proceed to ask whether the lower correlations in the extended dataset is consistent with the specific tradeoffs hypothesized to constrain Rubisco evolution.

	$k_{cat,C}$	$k_{cat,O}$	K_C	K_O	$\frac{k_{cat,C}}{K_C}$	$\frac{k_{cat,O}}{K_O}$	K_{RuBP}
$S_{C/O}$	-0.59	-0.17	-0.66	-0.28	0.36	-0.01	0.10
$k_{cat,C}$		0.38	0.62	0.12	0.02	0.22	0.08
$k_{cat,O}$			-0.26	0.25	0.56	0.67	-0.15
K_C				0.64	-0.83	-0.66	0.23
K_O					-0.59	-0.56	0.13
$\frac{k_{cat,C}}{K_C}$						0.93	-0.15
$\frac{k_{cat,O}}{K_O}$							-0.24

Figure 3: *Correlations between kinetic parameters are mostly attenuated by addition of new data. The figure summarizes Pearson correlations (R) between pairs of log-transformed Form I Rubisco kinetic parameters. The S-K_C, S-$k_{cat,C}$, and K_C-$k_{cat,C}$ correlations are of particular interest because they were highlighted in previous works. None of these pairs give R > 0.7. The strongest observed correlation is between the catalytic efficiencies for carboxylation and oxygenation, $k_{cat,C}/K_C$ and $k_{cat,O}/K_O$ (R = 0.93).*

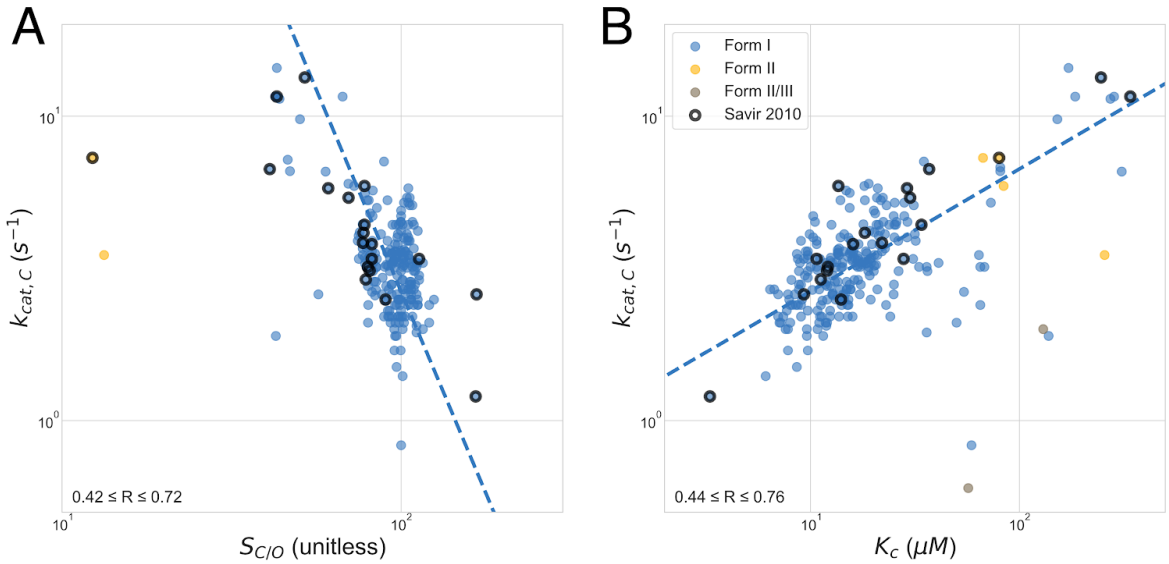

Figure 4: *Focal correlations of previous analyses are not robust to new data. Points with black outlines are those in Savir et al. 2010 and dashed blue lines represent the best fit to all Form I Rubiscos, which are also in blue. Panel A plots the maximum carboxylation rate $k_{cat,C}$ against specificity $S_{C/O}$ as in (4). Considering only Form I Rubiscos, $k_{cat,C}$ and $S_{C/O}$ correlate with $R \approx 0.6$. Bootstrapping gives very wide 95% confidence intervals (CIs) of (-3.3, -1.9) for the fit exponent and ($2x10^4$, $1x10^7$) for the exponential prefactor (the slope and intercept in log-log scale respectively) indicating that the form of $k_{cat,C}$-$S_{C/O}$ correlation is very uncertain. Panel B plots $k_{cat,C}$ against the Michaelis constant for CO_2 (K_C) as in (4, 5). R = 0.62 as compared to the previously reported value of 0.92. This fit is substantially more robust to outliers with bootstrapping giving 95% CIs of (0.3, 0.5) and (0.9, 1.5) for the fit exponent and prefactor respectively. The dashed blue line in (A) and (B) is the best fit to Form I Rubiscos.*

Re-evaluation of Energy Partitioning Models

Two mechanistic tradeoff models were advanced in (4, 5). Savir et al. 2010 (5) cast these proposals in energetic terms by relating the measured catalytic parameters to effective transition state barrier heights (Figure 1B, SI). This energetic interpretation of the first mechanistic model (4, 5) - that increased specificity towards CO_2 necessitates a slower maximum carboxylation rate (Figure S1) - is not supported by the extended dataset. This model was previously supported by an inverse relationship between $k_{cat,C}$ and $k_{cat,C}/K_C$. Since $k_{cat,C}/K_C$ is exponentially related to the first effective carboxylation barrier ($\ln(k_{cat,C}/K_C) \propto -\Delta G_{1,C}$) and $k_{cat,C}$ to the second ($\ln(k_{cat,C}) \propto -\Delta G_{2,C}$), the previously-observed relationship was taken to imply that the effective carboxylation barriers must sum to a constant ($\Delta G_{1,C} + \Delta G_{2,C} = C$, Figure 5A, (5)). The extended dataset does not, however, conform to the reported power law (Figure 5B). Moreover, correlation is not improved by restricting focus to C3 plants for which data is abundant and the measured leaf CO_2 concentration varies by only 20-30% (Figure 5C, (24, 25)).

Absence of correlation does not necessarily imply the absence of a tradeoff. Rather, if the Rubisco mechanism couples $k_{cat,C}$ and $k_{cat,C}/K_C$, much decreased correlation over the extended dataset ($R < 0.3$) could result from several factors (SI) including bias in data collection leading to undersampling of faster Rubiscos (e.g. from cyanobacteria) or, alternatively, insufficient selection pressure (as diagrammed in Figure 1D).

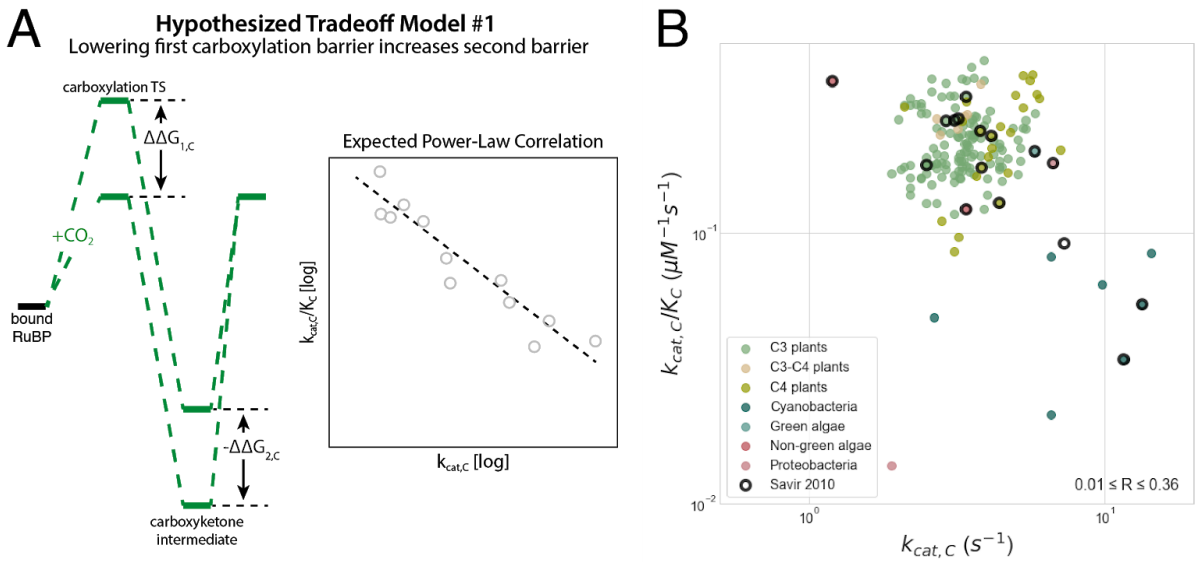

***Figure 5**: Negative correlation between $k_{cat,C}$ and $k_{cat,C}/K_C$ is not supported by the extended dataset. Under the hypothesized tradeoff model in panel (A), CO_2-specific Rubiscos have low barriers to enolization and CO_2 addition (first effective carboxylation barrier $\Delta G_{1,C}$), but lowering the first effective barrier necessarily increases the height of the effective barrier to subsequent carboxylation steps ($\Delta G_{2,C}$). This tradeoff might be due to coupling between the carboxylation transition state and the carboxyketone carboxylation intermediate (4), where stabilizing the TS also stabilizes intermediate diagrammed in panel (A) and described in Figure S2. In this case we would expect $\Delta G_{1,C}$ and $\Delta G_{2,C}$ to be negatively correlated, which would manifest as negative linear correlation on a log-log plot of $k_{cat,C}$ vs. $k_{cat,C}/K_C$. (B) The extended dataset does not evidence the expected power-law correlation ($R = 0.02$, $P = 0.8$ for Form I enzymes). Fitting the entire dataset gives $R = 0.28$ as compared to the previously-reported $R = -0.95$ in (5), where two outliers were omitted. Restricting focus to particular physiologies like C3 plants also does not produce the expected correlation (SI).*

The second mechanistic tradeoff model (5) - wherein faster CO_2 addition entails faster O_2 addition as well - is extremely well-supported by the addition of new data (Figure 6). This model was previously supported by a power-law relationship between $k_{cat,C}/K_C$ and $k_{cat,O}/K_O$ with an exponent of 0.5 ($k_{cat,O}/K_O \propto (k_{cat,C}/K_C)^{0.5}$). As $k_{cat,C}/K_C$ is exponentially related to the first effective carboxylation barrier ($\ln(k_{cat,C}/K_C) \propto -\Delta G_{1,C}$) and $k_{cat,O}/K_O$ to the first effective oxygenation barrier ($\ln(k_{cat,O}/K_O) \propto -\Delta G_{1,O}$), the power-law relationship was taken to imply that decreasing the barrier to CO_2 addition will also decrease the barrier to O_2 addition ($0.5\,\Delta G_{1,C} - \Delta G_{1,O} = C$, Figure 6A).

The extended dataset evidences clear power-law correlation between $k_{cat,C}/K_C$ and $k_{cat,O}/K_O$ (Figure 6B). While some Form II enzymes appear to be strictly inferior to the Form I enzymes on these axes, there is a clear “front” in the $k_{cat,C}/K_C$ vs. $k_{cat,O}/K_O$ plot. Most measurements lie along a robust line of positive correlation in a log-log plot (22). Fitting the Form I enzymes gives a remarkably high-confidence ($R = 0.93$, $P < 10^{-10}$) power-law relationship with $k_{cat,C}/K_C \propto (k_{cat,O}/K_O)^{1.06}$ (Figure 6B). The ratio of $k_{cat,C}/K_C$ to $k_{cat,O}/K_O$ is defined as the $S_{C/O}$ and so an exponent of 1 implies constant specificity. Subdividing the Form I enzymes by host physiology (e.g. C3 plants, C4 plants, cyanobacteria, etc.) reveals that all groups with sufficient data display a strong and statistically-significant power-law relationship between $k_{cat,C}/K_C$ and $k_{cat,O}/K_O$ (Figure 6C, SI (26)). The power-law exponent differs consistently from the value of 0.5 given by (5). We now find a roughly 1:1 relationship of $\Delta G_{1,C} - \Delta G_{1,O} = C$, meaning that a decrease in the CO_2 addition barrier is associated with an *equal* decrease in the barrier to O_2 addition. We estimate a 95% confidence interval (CI) of 0.98-1.24 for the exponent of this power law relationship for Form I enzymes, about twice the previously reported value.

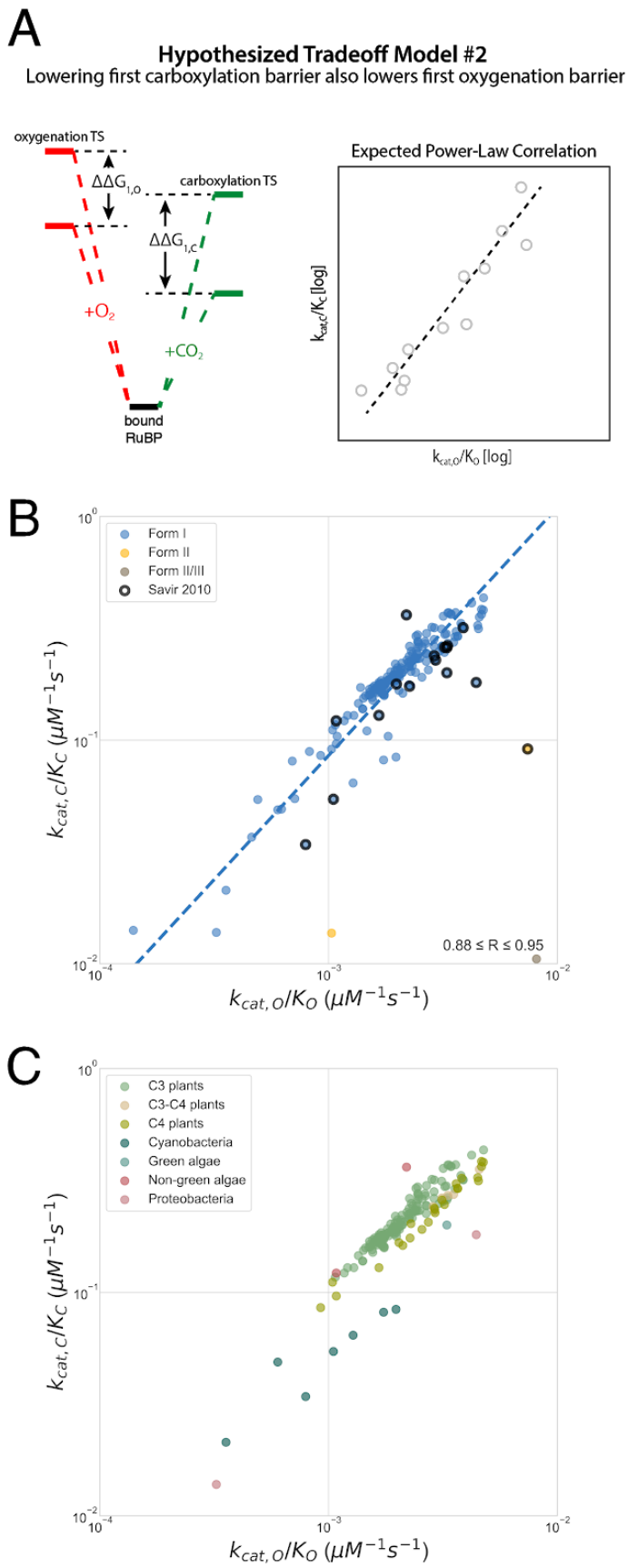

Figure 6: *The second mechanistic proposal of (5) is remarkably well-supported by the extended dataset, but with a different power-law exponent. (A) In this proposal, CO_2 and O_2 addition rates are coupled, with faster CO_2 addition necessitating faster O_2 addition. This can be framed in energetic terms, where lowering the effective barrier to enolization and CO_2 addition ($\Delta G_{1,C}$) lowers the first effective barrier to O_2 addition ($\Delta G_{1,O}$) as well. Given this model, we would expect the barrier heights to be positively correlated, which would manifest as a positive linear correlation on a log-log plot of $k_{cat,C}/K_C$ vs $k_{cat,O}/K_O$. (B) most measurements cluster along a power-law (linear in log-log) front in the $k_{cat,C}/K_C$ vs $k_{cat,O}/K_O$ plot (dashed blue line, R = 0.93). While some Form I Rubiscos appear to lie beneath this front, Form II and Form II/III enzymes deviate most profoundly. A total least squares fit to the Form I enzymes produces a very strong power-law correlation ($P < 10^{-10}$, blue dashed line). 95% CIs for the exponent and prefactor are (1.0, 1.2) and (96, 363), respectively. The best fit power law is $k_{on,C} \sim (k_{on,O})^{1.08}$, but forcing $k_{on,C} \sim (k_{on,O})^{1.0}$ gives a fit of nearly identical quality. (C) Restricting focus to particular physiologies - e.g. C3 and C4 plants, cyanobacteria - reveals that each grouping obeys a distinct power law. These power laws differ primarily in the exponential prefactor, which causes variation in the Y-intercept but not the slope on a log-log plot (SI). 95% CIs on the power-law exponent are (0.87, 0.98) for C3 plants, (0.82, 1.02) for C4 plants and (0.46, 1.09) for cyanobacteria.*

Implications for the mechanism of CO_2/O_2 discrimination by Rubisco

Figure 6 shows that the difference between the first effective barriers to carboxylation and oxygenation is roughly constant, such that they are constrained to vary in proportion ($\Delta G_{1,C}$ - $\Delta G_{1,O}$ = constant). A roughly 1:1 correlation between effective barriers to CO_2 and O_2 addition suggests that a single factor controls both. We offer a model based on the known catalytic mechanism of Rubisco that could produce a 1:1 relationship between barriers. In this model, the RuBP-bound Rubisco active site fluctuates between reactive and unreactive states (Figure 7A). The fraction of enzyme in the reactive state is denoted ϕ. In the unreactive state neither oxygenation or carboxylation is possible. In the reactive state, either gas can react at an intrinsic rate that does not vary across Rubiscos of the same class.

This model can be phrased quantitatively as $\frac{k_{cat,C}}{K_C} \propto \phi \exp(-\Delta G^*_{1,C}/RT)$ and $\frac{k_{cat,O}}{K_O} \propto \phi \exp(-\Delta G^*_{1,O}/RT)$ where $\Delta G^*_{1,C}$ and $\Delta G^*_{1,O}$ are the intrinsic reactivities of the enediolate of RuBP to CO_2 and O_2 respectively (SI, Figure S8). ϕ is likely determined by the degree of enolization of RuBP - i.e. $\phi = f(\Delta G_E)$ where ΔG_E is the equilibrium constant of on-enzyme enolization of RuBP (SI). Given this model, we expect to observe a power-law relationship with exponent 1.0 between $\frac{k_{cat,C}}{K_C}$ and $\frac{k_{cat,O}}{K_O}$ (SI). As specificity $S_{C/O} = (k_{cat,C}/K_C) / (k_{cat,O}/K_O)$, $S_{C/O}$ should be roughly constant under this model. Though $S_{C/O}$ varies the least of all measured Rubisco kinetic parameters (Figure 2C), it is not constant. Rather, $S_{C/O}$ varies over 3-4 fold among Form I Rubiscos and more than tenfold over the entire dataset (Figure 2C). However, Rubiscos isolated from hosts belonging to the same physiological grouping - e.g. C3 or C4 plants - do display a characteristic and roughly constant $S_{C/O}$ value independent of $k_{cat,C}/K_C$ (Figure 7B).

The implication of these data and model is that ϕ varies within the various Rubisco groups, perhaps by varying ΔG_E, which does not appear in the expression for $S_{C/O}$ (SI). In contrast, the difference between intrinsic reactivities $\Delta G^*_{1,O} - \Delta G^*_{1,C}$ appears to vary between the groupings. This would produce roughly constant $S_{C/O}$ among C3 plants while allowing large variation in $S_{C/O}$ between C3 plants, cyanobacteria and proteobacterial Form I Rubiscos. Characteristic variation in $S_{C/O}$ between groups of Form I Rubiscos might be understood via the conformational proofreading model (27). In this model, intentionally reducing complementarity between enzyme and substrate can lead to increased specificity if the change affects off-target substrates (e.g. O_2) even more than it affects on-target ones (e.g. CO_2). A full derivation of this model and discussion of its potential implications is given in the SI.

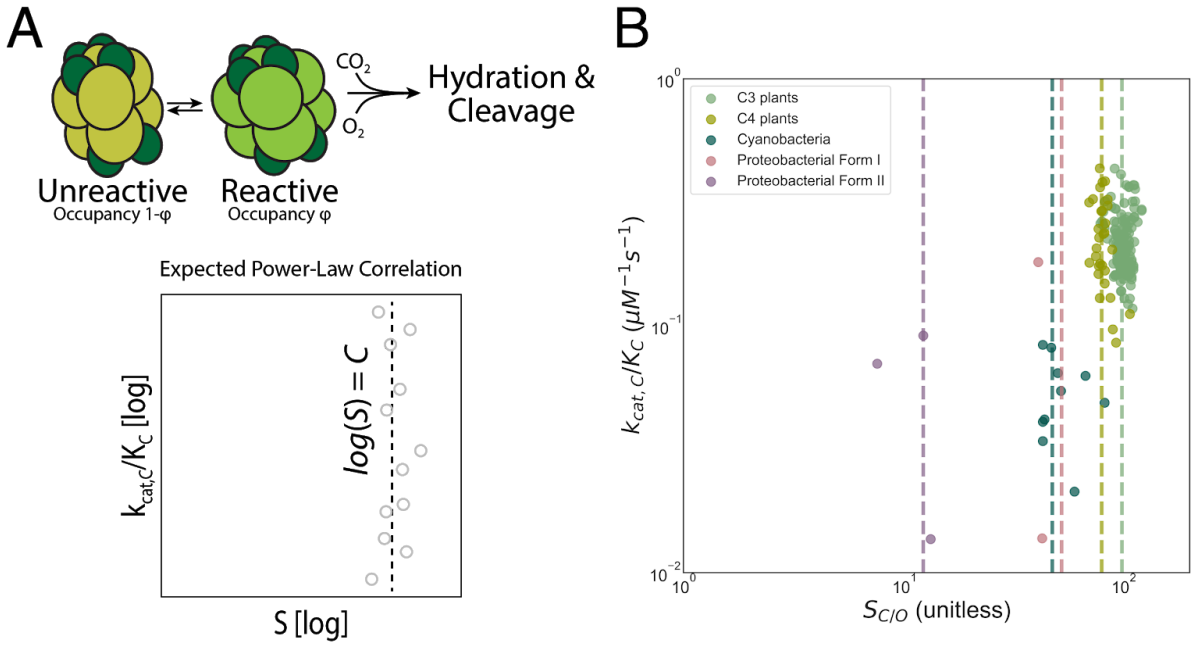

Figure 7: *A power-law relationship between $k_{cat,C}/K_C$ vs $k_{cat,O}/K_O$ with an exponent of roughly 1.0 can be explained by an active site that fluctuates between "reactive" and "unreactive" states. (A) In this model CO_2 and O_2 can react with the bound RuBP only if the enzyme is in the reactive state. If the difference in intrinsic reactivities of the active site complex ($\Delta G^*_{1,O} - \Delta G^*_{1,C}$) is organism-independent, we derive a power-law relationship between $k_{cat,C}/K_C$ vs $k_{cat,O}/K_O$ that has an exponent of 1.0 (SI, Figure S8). This model predicts that $S_{C/O}$ is constant. However, $S_{C/O}$ varies 3-4 fold across Form I Rubiscos. (B) Rubiscos within the same physiological grouping - e.g. C3 or C4 plants - have roughly constant $S_{C/O}$ independent of $k_{cat,C}/K_C$. The dashed vertical line is drawn at the median $S_{C/O}$ value for each group.*

Discussion

We collected and analyzed >200 literature measurements of Rubisco kinetic parameters (Figure 2A). The collection is quite biased, with the readily-purified Rubiscos of land plants making up ≈80% of the data (Figure 2B). Better sampling of Rubisco diversity including more algal, bacterial and archaeal Rubiscos would greatly improve our understanding of the evolution and capacity of this enzyme (28). Despite incomplete coverage, some trends are clear. All Rubisco kinetic parameters display limited dynamic range, with standard deviations in log-scale being less than one order-of-magnitude in all cases (Figure 2C). Rubisco is also noteworthy in displaying much less variability in k_{cat} than any other enzyme for which multiple measurements are available (Figure S4). The highest measured $k_{cat,C}$ at 25°C is 14 s^{-1} (*S. elongatus* PCC 7942) and the enzyme with the greatest affinity for CO_2 has K_C ≈ 3.3 μM at 25°C (*G. sulphuraria*). Many Rubiscos are quite slow oxygenators with half of measurements having $k_{cat,O}$ < 1 s^{-1} (Figure 2A). Similarly, many Rubiscos have relatively low affinity for O_2 - the median K_O is 465 μM. The Rubisco of the diatom *Thalassiosira weissflogii*, for example, has a K_O ≈ 2 mM (29), corresponding to roughly ten times the ambient O_2 concentration (30).

Specificity $S_{C/O}$ varies the least of all Rubisco parameters (Figure 2C). Form I Rubiscos are typically much more CO_2-specific than their Form II, III and II/III counterparts (Figure 7B, SI). This might be explained by the prevalence of Form II, III and II/III enzymes in bacteria and archaea that fix CO_2 in anaerobic conditions, where oxygenation should not appreciably compete with carboxylation. We note, however, that there is substantial variation among measurements of the model Form II Rubisco from *R. rubrum*. This and the general paucity of data on non-Form I Rubiscos indicates that more measurements are required to evaluate stereotyped differences within and between Form II, III and II/III Rubiscos.

In order to understand the limits of the Rubisco catalytic mechanism, we examined correlations between Rubisco kinetic parameters (Figure 3). Given the wide range of organisms studied, various methods of measurement applied in inconsistent buffer conditions, etc. we focused here on log-scale correlations among Form I Rubiscos for which abundant data is available (Figure 2B). This approach enables us to examine constraints within that group, but does not directly address trends between the Rubisco isoforms (e.g. stereotyped differences between Form I, II, II/III and III) as data is scant for these comparisons. Overall, the addition of new measurements weakens correlations between Rubisco kinetic parameters (Figure 3). The first principal axis of variation in the data (PC1) explains substantially less of the variance in Rubisco kinetics than in previous analyses (≈70% as compared to > 90%, SI). Increased variation within the extended dataset manifests as weaker correlations between some pairs of kinetic parameters. A plot of $k_{cat,C}$ against $S_{C/O}$ (Figure 4A) shows that the focal correlation of (4) is not strongly supported by the data. Similarly, a plot of $k_{cat,C}$ vs K_C shows that the focal correlation of (5) is substantially weakened (Figure 4B). Our understanding of weakened correlations is that the tradeoff model of (4) does not capture as much of the variation in the Rubisco mechanism as could be previously surmised. Another interpretation, however, is that natural Rubiscos are not “perfectly optimized” - that factors other than Rubisco limit autotrophic growth and natural selection has not pushed Rubiscos to the limits of their catalytic capacity (Figure 1C).

Examining the two mechanistic tradeoff models in (4, 5) we showed that only one is supported by our larger dataset (Figures 5, 6). The chemical intuition underlying both proposals is that the intrinsic difficulty of binding and discriminating between CO_2 and O_2 requires the enzyme to differentiate between carboxylation and oxygenation transition states. The requirement of TS discrimination is a direct consequence of two common assumptions that are supported by experimental evidence (31). Briefly, it is assumed that addition of either gas is irreversible and that there is no binding site for CO_2 or O_2 and, thus, no so-called Michaelis complex (4, 5, 31–33). If CO_2 bound a specific site on Rubisco before reacting, K_C could be modulated by mutation without substantially affecting the kinetics of subsequent reaction steps. In the less likely case that gas addition is substantially reversible, (33, 34) we would expect to find Rubiscos that evolved enhanced selectivity by energy-coupled kinetic proofreading. Energy coupling would enable amplification of selectivity determined by differential CO_2 and O_2 off-rates (35). The fact that no such Rubiscos have been found suggests that gas addition is irreversible or that the off-rates of CO_2 and O_2 are incompatible with kinetic proofreading in some other way.

For this reason we suggest that the higher specificity of C3 plant and red algal Rubiscos is achieved through a mechanism like conformational proofreading, which does not require energy coupling or reversible gas addition (5, 27).

Tcherkez et al. 2006 suggest that high specificity (i.e. large $S_{C/O}$) is realized via a late carboxylation TS which is maximally discriminable from the oxygenation TS (4). As a late TS resembles the carboxyketone carboxylation intermediate, specific Rubiscos must tightly bind the carboxyketone, which throttles the subsequent hydration and cleavage steps (Figure S2). The extraordinarily tight binding of the carboxyketone analog CABP to plant Rubisco provides strong support for this model. Savir et al. 2010 articulates a related model, noting that $k_{cat,C}$ and $k_{cat,C}/K_C$ are inversely correlated in their dataset (5). Since $k_{cat,C}/K_C$ is related to effective barrier to enolization and CO_2 addition and $k_{cat,C}$ is related to the effective barrier to hydration and cleavage (Figure 1B), Savir et al. argued that a lower effective barrier to CO_2 addition entails a higher barrier for the subsequent steps (i.e. a lower $k_{cat,C}$, Figure 5A). In both of these descriptions, the initial steps of carboxylation are negatively coupled to the subsequent steps in a manner that produces the correlations. However, those correlations - between $S_{C/O}$ and $k_{cat,C}$, K_C and $k_{cat,C}$ and $k_{cat,C}/K_C$ and $k_{cat,C}$ - are attenuated by the addition of new measurements (Figures 3, 4) which calls these proposals into question. Importantly, we do not argue that the chemical logic advanced by (4) is incorrect, but rather that the assembled data do not support such a tradeoff being optimized by the evolution of Form I Rubiscos.

The second tradeoff model posited by (5) is that faster CO_2 addition to the Rubisco-RuBP complex necessarily allows faster O_2 addition. This was evidenced by a positive power-law correlation between the catalytic efficiencies for carboxylation and oxygenation ($k_{cat,C}/K_C$ and $k_{cat,O}/K_O$ respectively), which can be understood as a positive coupling of the effective barrier to enolization and gas addition (Figure 6A, SI). We have shown that the extended dataset strongly supports this power-law relation and suggests that there exists a front along which lowering the effective CO_2 addition barrier (enabling faster carboxylation) requires a *roughly equal* reduction in the effective O_2 addition barrier (i.e. causing faster oxygenation as well). A power law relation with an exponent of 1.0 can be seen as resulting from an active site that fluctuates between a reactive and unreactive state (Figure 7A). In this model, the average occupancy of the reactive state dictates the rate of CO_2 and O_2 addition and throttles the subsequent steps of carboxylation and oxygenation equally (Figure 7). This model can be mapped onto the Rubisco mechanism by noting that RuBP must be enolized before CO_2 or O_2 can react, suggesting that the occupancy of the reactive state (φ) is related to the degree of enolization of RuBP (SI, Figure S8). One implication of this model is that $S_{C/O}$ is roughly constant. While $S_{C/O}$ does vary over roughly tenfold across the entire dataset and 3-4 fold across Form I enzymes (Figure 1B), Rubiscos from the same physiological groupings do display roughly constant $S_{C/O}$ values independent of $k_{cat,C}/K_C$ (Figure 7B). More measurements of bacterial Form I, II and III enzymes will be crucial to evaluate the generality of this observation.

We note, however, that the power-law relationship between $k_{cat,C}/K_C$ and $k_{cat,O}/K_O$ may not have an exponent of exactly 1.0. The Form I data are consistent with exponents ranging from

0.98-1.2 and fits to data from C3 and C4 plants gives exponents ranging from 0.8-1.0 (Figure 6C). Therefore, in order to better resolve the evolutionary constraints imposed on Rubisco kinetics, we suggest several avenues for future research. First, the kinetics of non-plant Rubiscos should be characterized much more thoroughly. These should include the Form II, III and II/III enzymes of bacteria and archaea as well as Form I enzymes of cyanobacteria and diverse Eukaryotic autotrophs (28). Ideally these enzymes would be sampled from accumulated genomic data in a manner that maximizes sequence and phylogenetic diversity (36) and characterized for their binding (e.g. of RuBP and CABP) and catalytic activity (i.e. measuring $k_{cat,C}$, K_C, $k_{cat,O}$, K_O and $S_{C/O}$) as a function of temperature and pH as in (37, 38). These data would likely resolve whether different Rubisco isoforms have characteristic differences in their catalytic potential.

Furthermore, it is important to revisit the classic experiments undergirding our understanding of the Rubisco catalytic mechanism, especially those supporting the central assumptions that (a) there is no Michaelis complex for CO_2 or O_2 and (b) that gas addition is irreversible (31, 33, 34). Other carboxylases, for example crotonyl-coa carboxylase/reductase, accept CO_2 as a substrate without any apparent affinity for O_2 (39), which leads us to wonder what makes Rubisco unique in this regard (8). One avenue for deeper study of the Rubisco mechanism would be measurement of carbon and oxygen kinetic isotope effects (KIEs) for a wide variety of Rubiscos. Kinetic isotope effects report indirectly on transition state barrier heights (40, 41) and so investigating the relationship between kinetic isotope effects and kinetic parameters will refine our current understanding of the catalytic mechanism (4).

There remains some disagreement about the precise ordering of the carboxylation mechanism (4, 12, 13) and the mechanism of oxygenation is not well understood (26). Chemical reasoning about the mechanisms of Rubisco carboxylation and oxygenation would benefit from progress in structural biology - intermediates and transition state analogues should be used to capture the active site at various points along the reaction trajectory (12, 26, 42–44). If experiments and structural analyses confirm that the above assumptions hold for all Rubiscos, it would greatly limit our capacity to engineer Rubisco and strongly suggest that alternative strategies for improving carbon fixation should be pursued (45, 46). If, however, these assumptions are invalidated, many enzyme engineering strategies would be viable. Such data and analyses will be instrumental in guiding the engineering of carbon fixation for the next decade.

Methods

Data collection and curation. We reviewed the literature to find Rubisco kinetic data measured at 25ºC and near pH 8. Ultimately 61 primary literature studies were included, yielding 319 $S_{C/O}$, 275 $k_{cat,C}$, 310 K_C, and 256 K_o values for Rubiscos from 286 distinct organisms (Datasets S1 and S2). We also recorded 51 measurements of the Michaelis constant for RuBP (K_{RuBP}). Experimental error was recorded for all of these values along with the pH, temperature and assumed pKa. Cases where the soluble CO_2 concentration was derived in a different manner were specifically noted. Data was filtered as described in SI. $k_{cat,O}$ is usually not measured

directly, but is rather inferred as $k_{cat,O} = (k_{cat,C}/K_C) / (S_{C/O}/K_O)$. When an uncertainty is reported, we assumed that the underlying experimental noise is normally distributed and used 10^4-fold bootstrapping to estimate 198 $k_{cat,O}$ values and 95% confidence intervals thereof. We used an identical procedure to estimate $k_{cat,C}/K_C$ and $k_{cat,O}/K_O$ (SI). Datasets S1 and S2 provide all source and inferred data.

Fitting power laws. In contrast to textbook examples with one independent and one dependent variable, there is experimental error associated with both variables in all correlation calculations here. As such we used total least squares linear regression in log scale to fit power-law relationships between Rubisco parameters. Because R^2 values of total least squares fits do not convey the explained fraction of Y axis variance, they are challenging to interpret. As such, we report the quality of correlations as Pearson R values. Bootstrapping was used to determine 95% confidence intervals for power-law exponents and prefactors (slopes and intercepts of linear fits in log scale). In each iteration of the bootstrap, data were subsampled to 90% with replacement. total least squares regression was applied to each subsample to determine a point estimate of R, power-law exponent and prefactor. This procedure was repeated 10^4 times to determine a 95% confidence interval on the above parameters. Python source code is online at github.com/flamholz/rubisco.

Acknowledgements

We would like to thank Uri Alon, Kapil Amarnath, Doug Banda, Arren Bar-Even, Cissi Blikstad, Jack Desmarais, Woodward Fischer, Vahe Galstyan, Laura Helen Gunn, Itai Halevy, Robert Nichols, Elad Noor, Yonatan Savir, Patrick Shih, Daniel Stolper, Dan Tawfik, Guillaume Tcherkez, Tsvi Tlusty and Renee Wang for helpful conversations and comments on the manuscript.

References

1. Raven JA (2013) Rubisco: still the most abundant protein of Earth? *New Phytol* 198(1):1–3.

2. Galmés J, et al. (2014) Expanding knowledge of the Rubisco kinetics variability in plant species: environmental and evolutionary trends. *Plant Cell Environ* 37(9):1989–2001.

3. Bar-On YM, Phillips R, Milo R (2018) The biomass distribution on Earth. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1711842115.

4. Tcherkez GGB, Farquhar GD, Andrews TJ (2006) Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.0600605103.

5. Savir Y, Noor E, Milo R, Tlusty T (2010) Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. *Proc Natl Acad Sci U S A* 107(8):3475–3480.

6. Bar-Even A, et al. (2011) The Moderately Efficient Enzyme: Evolutionary and

Physicochemical Trends Shaping Enzyme Parameters. *Biochemistry*. doi:10.1021/bi2002289.

7. Shih PM, et al. (2016) Biochemical characterization of predicted Precambrian RuBisCO. *Nat Commun* 7:10382.

8. Bathellier C, Tcherkez G, Lorimer GH, Farquhar GD (2018) Rubisco is not really so bad. *Plant Cell Environ* 41(4):705–716.

9. Zhu X-G, Long SP, Ort DR (2010) Improving photosynthetic efficiency for greater yield. *Annu Rev Plant Biol* 61:235–261.

10. Spreitzer RJ, Salvucci ME (2002) Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. *Annu Rev Plant Biol* 53:449–475.

11. Raven JA, Beardall J, Sánchez-Baracaldo P (2017) The possible evolution, and future, of CO2-concentrating mechanisms. *J Exp Bot*. doi:10.1093/jxb/erx110.

12. Cleland WW, Andrews TJ, Gutteridge S, Hartman FC, Lorimer GH (1998) Mechanism of Rubisco: The Carbamate as General Base. *Chem Rev* 98(2):549–562.

13. Andersson I (2008) Catalysis and regulation in Rubisco. *J Exp Bot* 59(7):1555–1568.

14. Durão P, et al. (2015) Opposing effects of folding and assembly chaperones on evolvability of Rubisco. *Nat Chem Biol* (January). doi:10.1038/nchembio.1715.

15. Bhat JY, et al. (2017) Mechanism of Enzyme Repair by the AAA+ Chaperone Rubisco Activase. *Mol Cell*. doi:10.1016/j.molcel.2017.07.004.

16. Mueller-Cajar O (2017) The Diverse AAA+ Machines that Repair Inhibited Rubisco Active Sites. *Front Mol Biosci* 4:31.

17. Busch FA, Sage RF, Farquhar GD (2017) Plants increase CO2 uptake by assimilating nitrogen via the photorespiratory pathway. *Nat Plants*. doi:10.1038/s41477-017-0065-x.

18. Bauwe H, Hagemann M, Fernie AR (2010) Photorespiration: players, partners and origin. *Trends Plant Sci* 15(6):330–336.

19. Jordan DB, Ogren WL (1981) Species variation in the specificity of ribulose biphosphate carboxylase/oxygenase. *Nature* 291:513.

20. Parry MAJ, Keys AJ, Gutteridge S (1989) Variation in the Specificity Factor of C3 Higher Plant Rubiscos Determined by the Total Consumption of Ribulose-P2. *J Exp Bot* 40(3):317–320.

21. Bainbridge G, et al. (1995) Engineering Rubisco to change its catalytic properties. *J Exp Bot* 46(special_issue):1269–1276.

22. Shoval O, et al. (2012) Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. *Science* 336(6085):1157–1160.

23. Flamholz A, Noor E, Bar-Even A, Milo R (2012) eQuilibrator--the biochemical thermodynamics calculator. *Nucleic Acids Res* 40(Database issue):D770–5.

24. Evans JR, Sharkey TD, Berry JA, Farquhar GD (1986) Carbon Isotope Discrimination measured Concurrently with Gas Exchange to Investigate CO2 Diffusion in Leaves of Higher Plants. *Funct Plant Biol* 13(2):281–292.

25. Caemmerer SV, Evans JR (1991) Determination of the Average Partial Pressure of CO2 in Chloroplasts From Leaves of Several C3 Plants. *Funct Plant Biol* 18(3):287–305.

26. Tcherkez G (2016) The mechanism of Rubisco-catalysed oxygenation. *Plant Cell Environ* 39(5):983–997.

27. Savir Y, Tlusty T (2007) Conformational proofreading: the impact of conformational changes on the specificity of molecular recognition. *PLoS One* 2(5):e468.

28. Liu D, Ramya RCS, Mueller-Cajar O (2017) Surveying the expanding prokaryotic Rubisco multiverse. *FEMS Microbiol Lett* 364(16). doi:10.1093/femsle/fnx156.

29. Young JN, et al. (2016) Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *J Exp Bot* 67(11):3445–3456.

30. Sander R (2015) Compilation of Henry's law constants (version 4.0) for water as solvent. *Atmos Chem Phys* 15(8):4399–4981.

31. Pierce J, Lorimer GH, Reddy GS (1986) Kinetic mechanism of ribulosebisphosphate carboxylase: evidence for an ordered, sequential reaction. *Biochemistry* 25(7):1636–1644.

32. Andrews TJ, Lorimer GH (1987) 3 - Rubisco: Structure, Mechanisms, and Prospects for Improvement. *Photosynthesis*, eds Hatch, Boardman NK (Academic Press), pp 131–218.

33. Tcherkez GG, Bathellier C, Farquhar GD, Lorimer GH (2018) Commentary: Directions for Optimization of Photosynthetic Carbon Fixation: RuBisCO's Efficiency May Not Be So Constrained After All. *Front Plant Sci* 9:929.

34. Cummins PL, Kannappan B, Gready JE (2018) Directions for Optimization of Photosynthetic Carbon Fixation: RuBisCO's Efficiency May Not Be So Constrained After All. *Front Plant Sci* 9:183.

35. Hopfield JJ (1974) Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc Natl Acad Sci U S A* 71(10):4135–4139.

36. Akiva E, Copp JN, Tokuriki N, Babbitt PC (2017) Evolutionary and molecular foundations of multiple contemporary functions of the nitroreductase superfamily. *Proc Natl Acad Sci U S A* 114(45):E9549–E9558.

37. Orr DJ, et al. (2016) Surveying Rubisco Diversity and Temperature Response to Improve Crop Photosynthetic Efficiency. *Plant Physiol* 172(2):707–717.

38. Sharwood RE, Ghannoum O, Kapralov MV, Gunn LH, Whitney SM (2016) Temperature

responses of Rubisco from Paniceae grasses provide opportunities for improving C3 photosynthesis. *Nat Plants* 2:16186.

39. Erb TJ, Brecht V, Fuchs G, Müller M, Alber BE (2009) Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase. *Proc Natl Acad Sci U S A* 106(22):8871–8876.

40. Hayes JM (2001) Fractionation of Carbon and Hydrogen Isotopes in Biosynthetic Processes. *Rev Mineral Geochem* 43(1):225–277.

41. McNevin DB, et al. (2007) Differences in Carbon Isotope Discrimination of Three Variants of D-Ribulose-1,5-bisphosphate Carboxylase/Oxygenase Reflect Differences in Their Catalytic Mechanisms. *J Biol Chem* 282(49):36068–36076.

42. Schneider G, Lindqvist Y, Brändén CI (1992) RUBISCO: structure and mechanism. *Annu Rev Biophys Biomol Struct* 21:119–143.

43. Andersson I, Backlund A (2008) Structure and function of Rubisco. *Plant Physiol Biochem* 46(3):275–291.

44. Stec B (2012) Structural mechanism of RuBisCO activation by carbamylation of the active site lysine. *Proceedings of the National Academy of Sciences* 109(46). doi:10.1073/pnas.1210754109.

45. McGrath JM, Long SP (2014) Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. *Plant Physiol* 164(4):2247–2261.

46. Bar-Even A (2017) Daring metabolic designs for enhanced plant carbon fixation. *Plant Sci*. doi:10.1016/j.plantsci.2017.12.007.