Desensitizing plant EPSP synthase to glyphosate: Optimized global sequence context accommodates a glycine-to-alanine change in the active site

Yuxia Dong[1], Emily Ng[1], Jian Lu, Tamara Fenwick, Yumin Tao, Sean Bertain, Marian Sandoval, Ericka Bermudez, Zhenglin Hou, Phil Patten[2], Michael Lassner[3], Daniel Siehl*

From Corteva Agriscience™, Agriculture Division of DowDuPont, 4010 Point Eden Way, Hayward, CA 94545

Running title*: Optimized, glyphosate-desensitized plant EPSPS*

[1]Equal contributors
[2]Present address; DiCE Molecules, 220 Penobscot Dr., Redwood City, CA 94063
[3]Present address; Amfora, Inc., San Francisco, CA 94104

*To whom correspondence should be addressed: Daniel L. Siehl, Corteva Agriscience™, Agriculture Division of DowDuPont, 4010 Point Eden Way, Hayward, CA, 510 780-7156, dan.siehl@pioneer.com

Abstract
5-Enolpyruvylshikimate-3-phosphate synthase (EPSPS) catalyzes the transfer of a carboxyvinyl group from phosphoenolpyruvate (PEP) to shikimate-3-phosphate and in plants is the target of the herbicide glyphosate. EPSPSs with high catalytic efficiency and insensitivity to glyphosate are of microbial origin, including the enzyme from *Agrobacterium* strain CP4, in which insensitivity is conferred by an active site alanine. In the sequence context of plant EPSPSs, alanine in place of glycine at the equivalent position interferes with the binding of both glyphosate and PEP. We show here that iterative optimization of maize EPSPS containing the G101A substitution yielded variants on par with CP4 in terms of catalytic activity in the presence of glyphosate. The improvement relative to G101A alone was entirely due to reduction in K_m for PEP from 333 μM to 18 μM, *versus* 9.5 μM for native maize EPSPS. A large portion of the reduction in K_m was conferred by two down-sizing substitutions (L97C and V332A) within 8 Å of glyphosate, which together reduced K_m for PEP to 43 μM. Though the original optimization was conducted with maize EPSPS, contextually homologous substitutions conferred similar properties to the EPSPSs of other crops. We also discovered a variant having the known glyphosate-desensitizing substitution P106L plus three additional ones that reduced the K_m for PEP from 47 μM, observed with P106L alone, to 10.3 μM. The improvements obtained with both Ala-101 and Leu-106 have implications regarding glyphosate-tolerant crops and weeds.

5-Enolpyruvylshikimate-3-phosphate synthase (EPSPS; EC 2.5.1.19) catalyzes the transfer of a carboxyvinyl group from phosphoenolpyruvate (PEP) to shikimate-3-phosphate (S3P, Figure 1). The enzyme is inhibited by glyphosate, competitively with PEP (1,2). The herbicidal efficacy of glyphosate against all plant species, its low cost, low mammalian toxicity and benign environmental

impact favor its use in crops that are endowed with a tolerance mechanism (3-7). Glyphosate tolerant crops have been created by expressing glyphosate-insensitive bacterial EPSPS encoded on a transgene (8,9).

With gene editing facilitated by Cas9/CRISPR technology (10), it is theoretically possible to create glyphosate tolerance without a foreign transgene by introducing a set of mutations into the EPSPS gene that effectively desensitizes the enzyme to glyphosate while maintaining sufficient catalytic capacity. USDA, the American agency regulating genetically modified crops, has given "unregulated" status to many traits created with gene editing, without the usual lengthy and costly (~$35 million) process required for a transgenic trait (https://www.aphis.usda.gov/aphis/ourfocus/biotechnology/am-i-regulated). Thus, gene editing could make glyphosate tolerance economically accessible for low volume crops or for staple crops in developing countries. In the latter case, the hand weeding prevalent in sub-Saharan Africa is typically far from optimal for preserving yield, is enormously time-consuming and often results in spinal deformation (11). However, enabling glyphosate tolerance by editing the native EPSPS gene will require a set of mutations that result in a much fitter enzyme than the plant variants currently known.

The known mutations in plant EPSPS and close homologs (termed Class I EPSPS) all involve modulations of active site Gly[101] (numbering according to mature maize EPSPS [CAA44974.1, Figure 2]; 96 in *E. coli*) that can create interference with the binding of glyphosate through one of its phosphonate oxygens (2). Changing Pro[106] to glycine, serine, threonine, leucine or alanine has the effect of moving the alpha carbon of Gly[101] closer to glyphosate, causing moderately reduced affinity (12). These mutations were discovered by mutagenesis of bacterial genes (13) and are present in seven species of glyphosate-resistant weeds (14-16).

Depending on the substitution, the degree of desensitization (increase in K_i) and whole plant resistance are proportional and are in the range of two- to nine-fold (14,15,17).

A second glyphosate-desensitized EPSPS variant is a double-mutant maize enzyme in which threonine at position 102 is changed to isoleucine in concert with the P106S mutation. The enzyme, termed TIPS, is highly desensitized to glyphosate, but while its K_m for phosphoenolpyruvate is nearly normal, it has only 5% of the k_{cat} of the native enzyme, as shown here and in earlier studies (18,19). Like the Pro[106] mutations, the TIPS mutations exert their effect by shifting Gly[101] closer to the glyphosate binding site (18). While catalytic capacity for TIPS EPSPS is insufficient if the mutations are created by natural mutagenesis and gene editing (see Discussion), sufficiently high transgenic expression can result in competitive performance, as in GA21 maize (20). Very recently, a mutation resulting in the same threonine (position 102) being changed to serine was shown to account for glyphosate resistance in a tropical weed (21). We here provide the first kinetic characterization of the variant, which shows it to be no fitter than P106S and thus unsuitable for gene editing.

The other known de-sensitizing substitution is alanine for glycine at the equivalent of our maize position 101, first reported with the enzyme from a glyphosate-resistant strain of *Klebsiella* (22). Alanine occurs naturally at this position in EPSPS from *Agrobacterium* strain CP4, the enzyme present in RoundupReady™ crops (8,9). The CP4 enzyme is highly insensitive to inhibition by glyphosate and has 11-15% of the catalytic efficiency (k_{cat}/K_m) of the plant enzyme (23,24). Plant EPSPS with the G101A mutation has similar insensitivity to CP4 but only 1.4% of the catalytic efficiency of the native plant enzyme, largely due to the 40-fold increase in K_m for PEP imposed by the additional methyl group (23,25). Though the structures of the CP4 and *E. coli* enzymes show

2

that they share the same structural fold and topology (26), they share only 23% identity. In contrast, the *E. coli* enzyme shares >50% identity with plant EPSPS. This divergence in sequence homology was the basis for designating CP4 and homologous microbial EPSPSs as Class II. The divergent amino acid sequence of CP4 has been thought to provide the structural context for an optimal spatial location of the alanine methyl group that results in interference with the longer glyphosate but not with PEP.

The G101A mutation has not been found in a glyphosate resistant weed, likely due to its poor affinity for PEP. We here show that with 17 or more additional mutations, the enzyme from maize, a Class I EPSPS, can be adapted to accommodate the G101A mutation, resulting in kinetic parameters equal to or better than those of CP4. The variants are no closer in homology to CP4 than is the native enzyme.

RESULTS

Mutagenesis, maize native EPSPS – Because the use of degenerate oligonucleotides for saturation mutagenesis theoretically can access all 19 changes at each position, we began by searching for single desensitizing mutations in native maize EPSPS that may have been missed by earlier methodology. Saturation mutagenesis was performed and positional libraries screened as described in Experimental Procedures. No new desensitizing mutations were found. The only ones found that showed appreciable desensitization were the known mutations at Pro[106] (further evaluated below). However, many neutral or slightly beneficial mutations were identified.

Combinatorial shuffling, initial library – The mutations discovered were used to construct a combinatorial library designed to explore combinations of functional diversity at the known desensitizing positions [G101A, T102(IALGV) and P106(SGLVQWA)] in novel sequence contexts provided by the newly identified neutral

mutations; in all, 43 substitutions at 29 positions. The library was synthesized entirely from oligonucleotides, using the technique of synthetic shuffling (27). The vector DNA of the library was transformed into our BL21(DE3) Tuner-AroA knockout strain and the cells were plated onto minimal medium with glyphosate (see Figure S1 for details on library design, host cells, screening conditions, advancement criteria). The variant forms of EPSPS from the 184 colonies that grew were evaluated by enzyme assay and the best of those were subjected to substrate saturation kinetic analysis.

Three variants each included one of the previously known mutations or pair of mutations that reduce sensitivity of EPSPS to inhibition by glyphosate. Variant P106L+3 has leucine substituted for proline at position 106 and three additional mutations (A100S, E301S and A337S). Alone, the P106L mutation raised K_i for glyphosate 60-fold, but also raised K_m for PEP five-fold (Table 1). The three additional mutations served to lower K_m for PEP from 47 μM, seen with P106L alone, to 10.3 μM with 60% retention of K_i. The overall result was 30-fold improved fitness [$(k_{cat}/K_m)*K_i$] compared to native maize EPSPS and seven-fold improved relative to P106S.

Along with two other mutations (T277G, A337S), variant TIPS+2 has the T102I and P106S (TIPS) mutations present in the GA21 maize transformation event (20). Kinetic analysis indicated a high level of insensitivity to glyphosate (K_i), near native affinity for PEP (K_m) but only ~ 5% of the native k_{cat} (Table 1), reflecting results obtained previously for TIPS (18,19).

Variant G101A+3 has alanine substituted for glycine at position 101, plus three other substitutions (E301S, E390G, V437R) and was the starting point for the optimization scheme (Figure S1) that resulted in the fittest glyphosate-desensitized variants (see below).

P106L and T102I-P106S variants could not be improved – P106L+3 and TIPS+2 were each subjected to a cycle of saturation mutagenesis and combinatorial shuffling, attempting to improve the K_i of P106L+3 or to improve the k_{cat} of TIPS+2. Neither attempt was successful (not shown). We found no mutation(s) that could work in concert with P106L (and presumably P106S or A) to enable any further desensitization. Nor did we find any mutation(s) that could compensate for the low k_{cat} imposed by the TIPS mutations.

Very recently, glyphosate resistance in a population of *Tridax procumbens* (Tribe *Heliantheae*) was attributed to a change from threonine to serine at position 102 (21). We created the T102S mutation in the sequence contexts of maize and *Helianthus annuus* (XP_022017499.1) EPSPS. The mutation had the effect of increasing K_i by 10-fold, but also elevating K_m for PEP by 3-fold (Table 1). We made no attempt to optimize this variant because it was unknown at the outset of our work.

Further optimization of maize EPSPS-G101A – In the context of the maize enzyme, G101A is highly insensitive to glyphosate, but has 35-fold elevated K_m for PEP relative to native EPSPS (Table 2), confirming earlier results with Class I EPSPS (22,23,28). However, unlike the situation with variants containing P106L or the TIPS mutations, G101A was amenable to improvement through iterative cycles of diversity generation and combinatorial shuffling. The process is shown schematically in Figure S1. For every application of saturation mutagenesis, at no position was the amino acid fixed. Rather, for each position, all 20 options were created in the sequence context of the current fittest variant and subjected to selection at the prevailing level of stringency. By not fixing the backbone, detrimental mutations can change to a different amino acid or revert to the native one.

Although G101A+3 was our starting point, the process was effectively the progressive

addition of mutations to G101A. The kinetic parameters for G101A+3 are poorer than those of G101A (Table 2), suggesting that one or more of its three additional mutations (E301S, E390G, V437R) was detrimental. Interestingly, all three were eliminated in the next phase of optimization. H6 and C2 were variants captured from a combinatorial library of diversity generated in the context of G101A+3, having 15 and 16 mutations, respectively. To fully explore the diversity present in H6 and C2, a small library was constructed in which the amino acids at their variable positions were toggled with the native amino acid. This allowed all positions encompassed by H6 and C2 to acquire a mutation or revert to the native amino acid, thus creating all possible combinations of the mutations present in the two variants and eliminating non-essential or deleterious ones. While not resulting in significant improvement relative to H6 (the better of the two parents), the procedure yielded variant C1, which had kinetic parameters similar to those of H6 (Table 2) but only 9 mutations compared with the 15 present in H6. Five of the six substitutions eliminated from H6 by the backcross procedure (including the three from G101A+3) did not reappear in subsequent cycles of saturation mutagenesis and combinatorial shuffling (Figure 3). From variant C1 on, with one exception (R1S in D2c-A5), the process was the progressive addition of new mutations to the previous best backbone sequence.

About the stage of variant C1, we created a parameter for predicting performance in the treated plant that would be a more rapid and informative representation of fitness than $(k_{cat}/K_m)*K_i$. Though useful for assessing the intrinsic potential for activity in the presence of a competitive inhibitor (29), $(k_{cat}/K_m)*K_i$ is inadequate for predicting the reaction velocity under the conditions of the application (plants sprayed with glyphosate) because it omits concentrations of substrates and inhibitor, factors that are not intrinsic to the enzyme, but on which

4

the reaction rate depends. Therefore, libraries subsequent to variant C1 in Figure S1 were evaluated with a single rate measurement designed to take all factors in the rate equation for competitive inhibition into account. The concentrations of PEP and S3P were set as nearly as possible (30 μM, limited by the sensitivity of our assay) to the presumed intracellular concentrations of 10-15 μM, the approximate values of K_m for both PEP and S3P for the native enzyme (Table 1). Glyphosate was included at 1 mM, a concentration attainable in tissues, especially meristems, receiving metabolite flow from treated leaves (30). The pH (7.0), ionic strength (100 mM KCl) and co-solvent concentration (5% ethylene glycol) were also intended to mimic *in vivo* conditions (31,32). The unit for the parameter is reaction rate (μM·min⁻¹) per enzyme concentration (μM), or min⁻¹, describing the enzyme turnover under application conditions, which we abbreviate as "k_{gly}". Though we continued to determine individual kinetic parameters for key variants, k_{gly} was adopted as the one parameter needed both for medium throughput screening and for ultimate evaluation of fitness.

The progressive increase in k_{gly} is shown graphically in Figure 4. The largest step in the progression was the 4-fold increase found in the first combinatorial library involving G101A. Variants from H6 on show dramatically reduced K_m for PEP, with some variants falling within 1.5-fold of the native value (Table 2). Most of the insensitivity to glyphosate conferred by the G101A mutation was retained, with K_i values clustering around 900 μM for all variants but C1. Optimization culminated with variants D2-124 and D2c-A5, with 18 and 21 mutations, respectively.

Large contribution from two near-active site substitutions – Examination of a model in which the maize amino acid sequence was threaded onto PDB:1G6S (*E. coli* EPSPS) provided clues as to how substitutions outside the active site could result in the dramatic reduction in K_m for PEP from 333 μM in G101A to 18 μM in variants D2-124 and D2c-A5. Three down-sizing substitutions occurred that were within 9 Å of glyphosate. L97C and D401G became incorporated in variant H6, and V332A, in D2 (Figures S1 and 3). In combination with G101A, D401G had little if any effect on K_m for either PEP or S3P relative to G101A alone and was detrimental to k_{cat} (Table 2). In contrast, when L97C and V332A were each tested in combination with G101A, each reduced K_m for PEP relative to G101A alone by half or more with no impact on k_{cat} (Table 2). The two together with G101A resulted in a K_m for PEP of 47 μM (variant ACA, Table 2), accounting for half of the improvement in k_{cat}/K_m of the best variant, D2c-A5, relative to G101A. As judged by k_{gly}, the triple mutant ACA was 63% as fit as the best variant.

Mutations in maize are transferable to EPSPS from other plant species – The high degree of conservation among plant EPSPS (identity at 66.4% of the positions with consensus at 95% of positions among 11 species) suggests that the mutations defined in the maize background would have a similar effect in EPSPS from other crop species. To map the D2-124 substitutions to EPSPS from other crop species, we used ClustalW to align native maize EPSPS, maize variant D2-124 and the native sequences from the target species. We used the ChloroP Prediction Server (http://www.cbs.dtu.dk/services/ChloroP/) to approximate the amino terminus of the mature proteins. Nucleotide sequences were optimized for expression in *E. coli* and synthesized commercially. The synthetic genes were cloned into pHD2114 and expressed, purified and analyzed. In most but not all cases, the substitutions defined in maize endowed a high level of fitness (k_{gly}) in the alternative plant EPSPS (Table 3). In the case of wheat, the K_m for PEP was close to that of the maize enzyme (18.9

5

μM *versus* 18.1 μM for maize), but the substitutions conferred a deficiency in k_{cat} (223 min^{-1} *versus* 631 min^{-1} for maize). In the other case of inefficient translation, the result was the opposite. The k_{cat} of soybean D2-124 was the same as maize D2-124, but the K_m for PEP was 5-fold higher, indicating that the D2-124 substitutions in the sequence context of the soy enzyme were only partially effective in accommodating the G101A substitution.

Discussion

No novel mutations – Extensive mutagenesis with a method (oligonucleotides with the "NNK" degenerate codon usage) theoretically capable of capturing all changes at all positions revealed no new mechanisms for desensitizing plant EPSPS to glyphosate. All known mutations function directly or indirectly at the contact between G101 (position 96 in *E. coli*) and a phosphonate oxygen of glyphosate. We tentatively conclude that the other amino acids with contacts to glyphosate in the *E coli* structure [PDB 1G6S, (2)], namely, K^{22}, N^{94}, R^{124}, Q^{171}, E^{341}, R^{344}, R^{386} and K^{411}, corresponding to K^{24}, N^{99}, R^{131}, Q^{180}, E^{359}, R^{362}, R^{404} and K^{429} in our maize model, are too important for binding or catalysis to be mutated or have no means through which their spatial positions can be adjusted to improve discrimination between PEP and glyphosate.

Improved context for P106L, implications for weed resistance – In screening our first combinatorial library, we found variants representing all three of the previously known mechanisms by which Class I EPSPS (e.g., plant, *E. coli*) can be desensitized to inhibition by glyphosate. Variant P106L+3 has leucine substituted for proline at position 106 plus three other mutations. Substitutions at position 106 that reduce glyphosate affinity are present in seven species of glyphosate-resistant weeds (14,15). Though Pro^{106} is not directly involved in molecular interaction with PEP or glyphosate,

mutations at 106 have the effect of moving other amino acids, particularly Gly^{101}, sufficiently closer toward the phosphonate end of glyphosate to moderately impede its binding (12). While the P106S mutation did not perturb kinetic parameters, P106L caused a 5-fold increase in K_m for PEP and doubled K_m for S3P [Table 1; also, (12,33)]. P106L conferred 60-fold reduced sensitivity (elevated K_i) to glyphosate compared to native EPSPS and 12-fold *versus* P106S (Table 1). Until recently, only P106S and P106T were known to confer resistance without involvement of another mechanism (14), leading to conjecture that the high K_m for PEP imposed by P106L precluded its emergence as a resistance mechanism. However, two weed species where resistance was due to P106L recently have been reported (15,16). Interestingly, in our variant P106L+3, K_m for PEP is restored nearly to normal, yielding intrinsic fitness [$(k_{cat}/K_m)*K_i$] that is 30-fold improved over native maize EPSPS and a k_{gly} value four-fold improved over P106S (Table 1). Were the P106L mutation to occur in a plant having the right sequence context, a plant with minimal fitness penalty and proportionally greater resistance factor (effective dose of glyphosate relative to wild type) than that conferred by P106S would be expected.

T102S, the recently reported third mutation in glyphosate resistant weeds – Our established methods presented the opportunity to evaluate this variant, recently reported as the basis for glyphosate resistance in *Tridax procumbens*, of the Tribe *Heliantheae* (21). In the sequence context of either maize or *Helianthus annuus* (sunflower), the mutation had the effect of elevating K_i 10-fold while elevating K_m for PEP 3-fold (Table 1). These results are supportive of the prediction, based on molecular dynamic simulations in a homology model, that the mutation increases the affinity for PEP relative to that of glyphosate (21). Compared with P106S, the combined effects of the T102S mutation (greater desensitization to glyphosate and

6

reduced affinity for PEP) leaves the two with about equal fitness as measured by k_{gly} (Table 1).

Combinations involving Thr[102] and P[106] could not be further improved – Variant TIPS+2 contains the T102I and P106S (TIPS) mutations present in the GA21 maize transformation event (20). The variant had a high level of insensitivity to glyphosate while retaining near native affinity for PEP but with only ~ 5% of the native k_{cat} (Table 1), confirming earlier reports with TIPS variants (18,19). The low catalytic efficiency of TIPS EPSPS evidently makes it unsuitable as a resistance mechanism except at high-level transgenic expression. Step-wise acquisition of both T102I and P106S mutations was documented in a population of *Eleusine indica* (19). However, out of a population of 193 individuals, only 1.6% were homozygous for TIPS. The highest frequency allelic combination was TIPS/P106S, suggesting that the normal catalytic capacity contributed from the P106S allele was more important for fitness than having the second allele encode a highly insensitive, but catalytically deficient enzyme. In controlled experiments, the same group later confirmed the extreme fitness penalty imposed by the TIPS mutations (34). Similarly, unacceptable agronomics were observed in cassava edited to create the TIPS or TIPA mutations, unless highly expressed by a constitutive viral promoter (35,36). Further, homozygosity for TIPS where the T0 rice plants were produced by CRISPR/Cas9 gene editing could not be found among 93 T2 plants analyzed, and was suggested to be lethal (37). Our attempt to improve k_{cat} in variants containing the T102I and P106S mutations was unsuccessful. The narrow range of parameters seen among the 32 variants analyzed (i.e., those that supported colony growth under selection conditions) suggests that the properties conferred by the T102I and P106S mutations are not significantly influenced by the sequence context but rather, are strongly determinant for a

narrow range of kinetic parameters (data not shown).

Class II properties from an optimized Class I EPSPS – Only variant G101A+3, having alanine substituted for glycine at position 101, was amenable to further optimization. G101A was associated with a 30,000-fold increase in K_i, but also with a 35-fold increase in K_m for PEP (Table 2), confirming earlier results (23,25,28,38,39). Alanine is present naturally at the homologous position in the Class II EPSPS from *Agrobacterium* sp. strain CP4. CP4 EPSPS exhibited a high degree of insensitivity to glyphosate but with a K_m for PEP of just 15.5 μM (Table 2), in good agreement with the 12 μM obtained earlier (23). Comparison of the crystal structures of CP4 ligated with S3P and glyphosate [PDB 2GGA; (24)] and *E. coli* EPSPS with the contextually equivalent glycine changed to alanine, ligated with S3P and glyphosate [modeled, (28)], indicate that the alanine methyl group in CP4 is positioned 0.3 Å further away from the phosphonate group of glyphosate than in the *E. coli* structure. Because PEP is shorter than glyphosate, it is hypothesized that the alanine methyl group in CP4 EPSPS is ideally positioned to interfere with binding of glyphosate but not PEP. Though there is only 24-26% homology between the CP4 enzyme and *E. coli* or maize EPSPS, structures of the CP4 and *E. coli* enzymes show that they share the same structural fold and topology (2,24). Presumably, the amino acid sequence of CP4 creates an overall structural context that places the alanine methyl group in its favorable position. The novel global sequence context that we arrived at resulted in Class 1 EPSPSs (D2-124 and D2c-A5) that match CP4 in catalytic proficiency in the presence of glyphosate, yet are no closer in sequence homology to CP4 than is the native maize EPSPS.

Structural rationale for the effect of substitutions outside the active site –We surmise that the new sequence context we arrived at through our process had a similar effect of

increasing the distance between the Ala[101] methyl and the phosphonate group of glyphosate. Two downsizing substitutions (L97C and V332A) when present together with G101A resulted in a K_m for PEP of 47 µM (variant ACA, Table 2), accounting for 50% the improvement in k_{cat}/K_m of the best variant relative to G101A. Both are within 8 Å of glyphosate in our model of maize EPSPS threaded onto the *E. coli* enzyme (PDB: 1G6S). Leu[97] is in a loop that links a β-strand to helix-3, the key catalytic element harboring G101A. The bulky Leu[97] side-chain packs tightly against the helix, ~6 Å from Ala[101]. Downsizing the side chain to cysteine would create a void that could create space for Ala[101] to move outward from the PEP-glyphosate binding site. The 18 remaining substitutions in D2c-A5 are remote from the active site, yet somehow transmit a collective effect through the structure, further optimizing the position of the methyl group so as to provide a further two-fold improvement in k_{cat}/K_m while retaining robust insensitivity to glyphosate.

Further evidence that the active site of the best variant (D2c-A5) is enlarged is that when its Ala[101] was changed back to glycine, the kinetic parameters obtained were very different from those of the native enzyme. As expected, sensitivity to glyphosate was almost completely restored. Surprisingly, however, K_m for PEP fell from 18 to 3.9 µM (*versus* 9.5 µM for the native enzyme), while k_{cat} fell to 20% of the optimized variant (Table 2). The effects on both k_{cat} and K_m for PEP can be rationalized by presuming that the re-modeled active site has become enlarged and that in the absence of the extra methyl group, is too large. We speculate that K_m is lower in the methyl group-free optimized active site than in the native because in the native active site, PEP must expend some of its binding energy in binding to a relatively constricted pocket compared to that of the methyl group-free evolved variant, manifesting in a relatively higher K_m. The negative effect on k_{cat} of the expanded active site may be a consequence of the substrates being held in positions that are out of range for optimal reaction rate. A comparison of crystal structures of maize native and variant EPSPS may confirm this conjecture and provide further details of how global sequence context creates an active site appropriately re-modeled to accommodate the extra methyl group.

In addition to its 35-fold effect on affinity for PEP, the G101A mutation also reduced affinity for S3P, but only by 6-fold ($K_m = 84$ µM *versus* 13.2 µM for native EPSPS, Table 2). Reduced affinity for S3P was also observed with the same mutation in the *Klebsiella* enzyme (38). Optimization resulted in the K_m for S3P returning to near normal in parallel with the reduction in K_m for PEP. We hypothesize that in the native sequence context, the G101A mutation does not directly interfere with S3P binding, but that the extra methyl group simply reduces the space available for S3P when PEP is bound.

Fitness as quantified by k_{gly} versus (k_{cat}/K_m)*K_i – The fitness parameter k_{gly}, which takes into account anticipated concentrations of substrates and inhibitor as well as kinetic parameters, is intended to better capture enzyme fitness under the conditions of the application than (k_{cat}/K_m)*K_i. In fact, the two measures of fitness correlated rather well with a few exceptions (Figure 5). CP4, with its very high K_i, displayed a disproportionately high (k_{cat}/K_m)*K_i. This is due to the greater impact of K_i on that parameter compared with its impact on the velocity equation for competitive inhibition, $v = k_{cat} [E] [S] / K_m (1 + [I] / K_i + [S]$, which our k_{gly} parameter seeks to represent. The relatively low k_{gly} for CP4 given its outstanding selectivity (K_i/K_m) is attributable to its low value for k_{cat}. The converse, high k_{cat} but poor selectivity (G101A), is likewise unsatisfactory. D2c-A5 incorporates the best combination of the parameters under optimization. Initially (H6), substitutions that rectified the high K_m for PEP imposed by G101A were also detrimental to k_{cat}. Improvement from

8

D2 on occurred mainly by improving k_{cat} while retaining the selectivity established earlier.

Slow release of glyphosate from maize native EPSPS – Our value of 66 nM for K_i for the native enzyme is in accord with the 80 nM reported for EPSPS from *Pisum sativum* (40) and the 48 nM obtained with the *Eleusine indica* enzyme (17). Though glyphosate has not been characterized as a "slow, tight-binding" inhibitor (41), its dissociation from an E:S3P:glyphosate complex was slow enough to be observed over a 40-sec span using our continuous assay (Figure S3, Supporting Information). The slow release effectively removed a fraction of the enzyme from the reaction, causing the substrate saturation curves with and without glyphosate to resemble non-competitive inhibition (Figure S2). When the protocol was revised to avoid a pre-incubation with enzyme, glyphosate and S3P, classic competitive inhibition was observed (Figures S4-S5). In view of nM values for K_i and our observation of time-dependent dissociation, the rather high application rates of glyphosate (relative to, e.g., AHAS inhibiting herbicides) may have led to an under appreciation of glyphosate's potency as an inhibitor of plant EPSPS.

Previous attempts to use genome engineering to introduce known EPSPS mutations into crops to develop glyphosate tolerant plants met with limited success unless a strong constitutive promoter was used to drive EPSPS expression (35,36). It will be interesting to observe whether converting native maize EPSPS to an optimized variant by genome editing will result in commercial level trait efficacy and normal phenotype. If so, the facile transfer of properties obtained by mapping the mutations identified in this work onto EPSPS from other species should enable non-transgenic glyphosate tolerance in many diverse crops.

EXPERIMENTAL PROCEDURES

Reagents – S3P was prepared from cultures of *Klebsiella pneumonia* aroA- (ATCC 25597). Cells from a 500 ml culture grown in 2xYT were used to inoculate 6 L of minimal medium augmented with 55 μM tyrosine, 60 μM phenylalanine, 25 μM tryptophan, 0.1 μM 4-aminobenzoate and 0.1 μM 4-hydroxybenzoate (42). Accumulation of S3P was monitored by anion exchange HPLC. After 4 days shaking at 37 C, the concentration reached ~1 mM. S3P was purified from the culture supernatant by anion exchange chromatography in ammonium bicarbonate at pH 7.3, with gradient elution up to 0.7 M. S3P was cleanly separated from phosphate, which eluted earlier. 2-Amino-6-mercapto-7-methylpurine ribonucleoside (MESG) was from Setareh Biotech, Eugene Oregon. Glyphosate was from ORICO (oricoglobal.com). All other reagents were from Sigma.

Vector for expressing plant EPSPS in E. coli – The amino acid sequence of mature *Zea mays* (maize) EPSPS was obtained from GenBank entry CAA44974.1. A nucleotide sequence was created to add an N-terminal methionine and optimize codon usage for expression in *E. coli*. The synthetic gene was supplied by a commercial vendor and was cloned into pHD2114, a derivative of pET16 *E coli* expression vector that provides a T7 promoter driving expression of the protein (Figure S6, Supporting Information). The vector was modified to change the 6x N-terminal histidine tag to a 10x tag. Later, to reduce leaky expression and increase stringency, the origin of replication was altered to produce a lower copy number (see below, *Screening procedures*). The resulting coding region of the vector yields an expressed protein with the amino acid sequence shown in Figure 2. CP4, the EPSPS derived from *Agrobacterium sp.* strain CP4 (9), was accessed by gene synthesis.

Saturation mutagenesis – Native or improved variants of EPSPS were subjected to

saturation mutagenesis to discover novel mutations that reduce sensitivity to glyphosate. Libraries of substitutions for each position in the EPSPS polypeptide chain were created using NNK (where N represents a 25% mix each of adenine, thymine, guanine, and cytosine nucleotides; and K represents a 50% mix each of thymine and guanine nucleotides) as the degenerate codon for the position to be mutagenized. PCR reaction mixtures contained a mutagenic forward primer (NNK codon flanked by 28 nucleotides matching with template at each side of the NNK) and a reverse primer that was the complement of the sequence preceding the forward primer, 28 nucleotides in length. To make circular double stranded DNA plasmids from the blunt ended PCR products, the products were digested with T4 polynucleotide kinase, T4 DNA ligase, and DpnI (to disrupt the parental DNA template). After desalting by ultrafiltration, the ligation products were ready for transformation and downstream applications.

Combinatorial library construction – The initial combinatorial library (neutral diversity identified in the native EPSPS plus known mutations at positions 101, 102 and 106) was constructed entirely from oligonucleotides, using the technique of synthetic shuffling (27). All other libraries were made by random toggling of desired substitutions into a specific sequence, termed the backbone, by the technique of semi-synthetic shuffling (27).

Screening procedures – We used a tiered screening cascade to identify the fittest variants in each library. At the outset of optimization beginning with native maize EPSPS, plasmid libraries of single or multiple mutations were transformed into BL21(DE3)-Tuner cells in which the native *AroA* gene was functionally deleted by phage transduction. The donor strain was JW0891 (CGSC), in which the *AroA* gene is disrupted with a Kanamycin resistant gene. Variant H6 (Figure3, Table 2) had a level of fitness that allowed growth on >10 mM

glyphosate (at which the native EPSPS is completely inhibited), allowing its descendants to be expressed in standard BL21(DE3) cells. Cells were plated on minimal agar medium containing M9 salts, 2% glucose, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, and 10 mg/L thiamine, creating the requirement for a functional, plasmid-encoded EPSPS to enable colony growth. As fitness improved, amendments were added to increase the stringency of selection, as indicated in Figure S1. Glyphosate was added up to 300 mM, its limit of solubility in agar. Polymyxin B nonapeptide was included at 1 mg/L to increase membrane permeability to glyphosate (43). Betaine was added to provide an intracellular osmolyte to support growth in the presence of high concentrations of glyphosate (44). Later, the pH of the agar medium was reduced from 7.0 to 5.5, under the hypothesis that a larger fraction of glyphosate would be protonated and thus be more likely to be transported by proton cotransporters.

The plates were incubated at 37 C until colonies grew to a size conducive for automated picking. Saturation mutagenesis generates a small library of 19 variants at each position. For screening, 96 positional libraries were pooled, combining for roughly 2000 variants per pool, easily over-sampled by plating on selective medium. For combinatorial libraries, sufficient cell transformation reactions were performed to generate as many colony-forming units as theoretical unique members in the library, typically ~10^6, all of which were plated and screened.

At the fitness level attained by variant D2, two more adaptations were made to increase stringency of the plate selection. The origin of replication in the expression vector was replaced with that from pSC101, which typically generates ~5 copies of the plasmid instead of ~20 (45). In the final stages, vector DNA encoding the libraries was transformed into BL21(DE3)-C41 cells. In that strain, mutations in the promoter of T7 RNA polymerase results in lower levels of

10

induced, and presumably un-induced, expression (46).

EPSPS production and purification – Proteins were produced in *E. coli* strain BL21(DE3). Transformed cells were grown in Magic Medium (Thermo Fisher Scientific). After 4 hours of growth at 37 C, cells were transferred to 16 C and grown another 48 hrs. Pelleted cells were lysed with BPER (Pierce) protein extraction reagent containing 25 mM Tris (to elevate pH to 7.8), 0.2 mg/ml lysozyme, 1 mM dithiothreitol, protease inhibitor cocktail (Sigma, bacterial cocktail) and endonuclease. Insoluble cellular debris was removed by centrifugation. EPSPS protein was purified from the soluble protein solution by affinity chromatography on the nickel form of nitrilotriacetic acid (Ni-NTA) resin (Qiagen). Protein concentration was measured by absorbance at 280 nm using the molar extinction coefficient calculated by Vector NTI for the amino acid sequence of each variant, including the additional N-terminal amino acids coded on the vector.

EPSPS assay – Activity was determined by quantifying release of inorganic phosphate, coupled to reaction with 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG), catalyzed by purine-nucleoside phosphorylase (44). The absorbance change that occurs was monitored with a Spectramax plate reader (Molecular Devices) at 360 nm, where the extinction is 11,200 M-1cm-1. Reaction mixtures contained 25 mM Hepes, pH 7, 100 mM KCl, 5% (v/v) ethylene glycol, 0.15 mM MESG, 1.5 unit/ml purine nucleoside phosphorylase (Sigma N8264) and concentrations of EPSPS customized to result in linear initial reaction rates. In addition to these common components, reaction mixtures contained substrates and glyphosate as required by the fitness of the variant. To determine kinetic parameters with substrate saturation analysis, the varied substrate was present at eight concentrations typically ranging from zero (blank) to 800 μM, but as high as 3.2 mM for G101A, and the unvaried substrate was supplied at the saturating concentration of 10 times its K_m. Five microliters of 60-fold concentrated stock solutions of the varied substrate were placed in the wells of the 96-well assay plate and reactions were started with the addition of 295 μl of a mixture of all other components. Reactions were observed at 360 nm for one minute, taking readings at 4 sec intervals.

Data collection and analysis – All the data in Tables 1 and 2 were obtained from a single panel of purified enzymes prepared together and meeting the purity criterion of no additional bands on a Coomassie Blue-stained SDS gel loaded with 3 ug of each enzyme. The initial, linear reaction rate was captured using the first 6 or 7 readings. Data from runs in which a linear initial rate was not captured due to excessive enzyme concentration were excluded. Kinetic parameters for S3P were obtained by non-linear regression analysis (GraphPad Prism, graphpad.com) of at least four sets of rate measurements at 7 substrate concentrations, fitted to the Michaelis-Menten equation. For PEP, the procedure was repeated with one or two concentrations of glyphosate and the data were processed by non-linear regression analysis fitted to the Michaelis-Menten equation for competitive inhibition.

REFERENCES

1. Boocock, M. R., and Coggins, J. R. (1983) Kinetics of 5-enolpyruvylshikimate-3-phosphate synthase inhibition by glyphosate. *FEBS Lett* **154**, 127-133

2. Schonbrunn, E., Eschenburg, S., Shuttleworth, W. A., Schloss, J. V., Amrhein, N., Evans, J. N., and Kabsch, W. (2001) Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. *Proc Natl Acad Sci U S A* **98**, 1376-1380

3. Franz, J. E., Mao, M. K., and Sikorski, J. A. (1997) *Glyphosate: A unique global herbicide*, American Chemical Society, Washington, DC

4. Solomon, K. (2017) What is the problem with glyphosate? *Outlooks on Pest Management* **28**, 173-174

5. Duke, S. O., and Powles, S. B. (2008) Glyphosate: a once-in-a-century herbicide. *Pest Manag Sci* **64**, 319-325

6. Nielsen, L. N., Roager, H. M., Casas, M. E., Frandsen, H. L., Gosewinkel, U., Bester, K., Licht, T. R., Hendriksen, N. B., and Bahl, M. I. (2018) Glyphosate has limited short-term effects on commensal bacterial community composition in the gut environment due to sufficient aromatic amino acid levels. *Environ Pollut* **233**, 364-376

7. Tarone, R. E. (2018) Conflicts of interest, bias, and the IARC Monographs Program. *Regul Toxicol Pharmacol*

8. Heck, G. R., Armstrong, C. L., Astwood, J. D., Behr, C. F., Bookout, J. T., Brown, S. M., Cavato, T. A., DeBoer, D. L., Deng, M. Y., George, C., Hillyard, J. R., Hironaka, C. M., Howe, A. R., Jakse, E. H., Ledesma, B. E., Lee, T. C., Lirette, R. P., Mangano, M. L., Mutz, J. N., Qi, Y., Rodriguez, R. E., Sidhu, S. R., Silvanovich, A., Stoecker, M. A., Yingling, R. A., and You, J. (2005) Development and Characterization of a CP4 EPSPS-Based, Glyphosate-Tolerant Corn Event. *Crop Science* **45**, 329-339

9. Padgette, S. R., Kolacz, K. H., Delanny, X., Re, D. B., AlVallee, B. J., Tinius, C. N., Rhodes, W. K., Otero, Y. I., Barry, G. F., Eichholtz, D. A., Peschke, V. M., Nida, D. L., Taylor, N. B., and Kishore, G. M. (1995) Development, identification, and characterization of a glyphosate-tolerant soybean line. *Crop Science* **35**, 1451-1461

10. Sander, J. D., and Joung, J. K. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* **32**, 347-355

11. Gianessi, L. P. (2013) The increasing importance of herbicides in worldwide crop production. *Pest Manag Sci* **69**, 1099-1105

12. Healy-Fried, M. L., Funke, T., Priestman, M. A., Han, H., and Schonbrunn, E. (2007) Structural basis of glyphosate tolerance resulting from mutations of Pro101 in Escherichia coli 5-enolpyruvylshikimate-3-phosphate synthase. *J Biol Chem* **282**, 32949-32955

13. Stalker, D. M., Hiatt, W. R., and Comai, L. (1985) A single amino acid substitution in the enzyme 5-enolpyruvylshikimate-3-phosphate synthase confers resistance to the herbicide glyphosate. *J Biol Chem* **260**, 4724-4728

14. Sammons, R. D., and Gaines, T. A. (2014) Glyphosate resistance: state of knowledge. *Pest Manag Sci* **70**, 1367-1377

15. Ngo, T. D., Krishnan, M., Boutsalis, P., Gill, G., and Preston, C. (2018) Target-site mutations conferring resistance to glyphosate in feathertop Rhodes grass (Chloris virgata) populations in Australia. *Pest Manag Sci* **74**, 1094-1100

16. Chen, J., Huang, H., Zhang, C., Wei, S., Huang, Z., Chen, J., and Wang, X. (2015) Mutations and amplification of EPSPS gene confer resistance to glyphosate in goosegrass (Eleusine indica). *Planta* **242**, 859-868

17. Baerson, S. R., Rodriguez, D. J., Tran, M., Feng, Y., Biest, N. A., and Dill, G. M. (2002) Glyphosate-resistant goosegrass. Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. *Plant Physiol* **129**, 1265-1275

18. Funke, T., Yang, Y., Han, H., Healy-Fried, M., Olesen, S., Becker, A., and Schonbrunn, E. (2009) Structural basis of glyphosate resistance resulting from the double mutation Thr97 -> Ile and Pro101 -> Ser in 5-enolpyruvylshikimate-3-phosphate synthase from Escherichia coli. *J Biol Chem* **284**, 9854-9860

19. Yu, Q., Jalaludin, A., Han, H., Chen, M., Sammons, R. D., and Powles, S. B. (2015) Evolution of a double amino acid substitution in the 5-enolpyruvylshikimate-3-phosphate synthase in Eleusine indica conferring high-level glyphosate resistance. *Plant Physiol* **167**, 1440-1447

20. Anonymous. (2002) Safety assessment of Roundup Ready corn event GA21. http://cera-gmc.org/files/cera/GmCropDatabase/docs/decdocs/02-269-006.pdf

21. Li, J., Peng, Q., Han, H., Nyporko, A., Kulynych, T., Yu, Q., and Powles, S. (2018) Glyphosate Resistance in Tridax procumbens via a Novel EPSPS Thr-102-Ser Substitution. *J Agric Food Chem* **66**, 7880-7888

22. Sost, D., and Amrhein, N. (1990) Substitution of Gly-96 to Ala in the 5-enolpyruvylshikimate-3-phosphate synthase of Klebsiella pneumoniae results in a greatly reduced affinity for the herbicide glyphosate. *Arch Biochem Biophys* **282**, 433-436

23. Sikorski, J. A., and Gruys, K. J. (1997) Understanding Glyphosate's Molecular Mode of Action with EPSP Synthase: Evidence Favoring an Allosteric Inhibitor Model. *Accounts of Chemical Research* **30**, 2-8

24. Funke, T., Han, H., Healy-Fried, M. L., Fischer, M., and Schonbrunn, E. (2006) Molecular basis for the herbicide resistance of Roundup Ready crops. *Proc Natl Acad Sci U S A* **103**, 13010-13015

25. Padgette, S. R., Re, D. B., Gasser, C. S., Eichholtz, D. A., Frazier, R. B., Hironaka, C. M., Levine, E. B., Shah, D. M., Fraley, R. T., and Kishore, G. M. (1991) Site-directed mutagenesis of a conserved region of the 5-enolpyruvylshikimate-3-phosphate synthase active site. *J Biol Chem* **266**, 22364-22369

26. Pollegioni, L., Schonbrunn, E., and Siehl, D. (2011) Molecular basis of glyphosate resistance-different approaches through protein engineering. *FEBS J* **278**, 2753-2766

27. Ness, J. E., Kim, S., Gottman, A., Pak, R., Krebber, A., Borchert, T. V., Govindarajan, S., Mundorff, E. C., and Minshull, J. (2002) Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. *Nat Biotechnol* **20**, 1251-1255

28. Eschenburg, S., Healy, M. L., Priestman, M. A., Lushington, G. H., and Schonbrunn, E. (2002) How the mutation glycine96 to alanine confers glyphosate insensitivity to 5-enolpyruvyl shikimate-3-phosphate synthase from Escherichia coli. *Planta* **216**, 129-135

29. Lu, J., Dong, Y., Ng, E. C., and Siehl, D. L. (2017) Novel form of the Michaelis-Menten equation that enables accurate estimation of (kcat/KM)*KI with just two rate measurements; utility in directed evolution. *Protein Eng Des Sel* **30**, 395-399

30. Kirkwood, R., Hetherington, R., Reynolds, T. L., and Marshall, G. (2000) Absorption, localisation, translocation and activity of glyphosate in barnyardgrass (Echinochloa crus-galli (L) Beauv): influence of herbicide and surfactant concentration. *Pest. Manag. Sci.* **56**, 359-367

31. Kronzucker, H. J., Szczerba, M. W., and Britto, D. T. (2003) Cytosolic potassium homeostasis revisited: 42K-tracer analysis in Hordeum vulgare L. reveals set-point variations in [K+]. *Planta* **217**, 540-546

32. Cuin, T. A., Miller, A. J., Laurie, S. A., and Leigh, R. A. (2003) Potassium activities in cell compartments of salt-grown barley leaves. *J Exp Bot* **54**, 657-661

33. Zhou, M., Xu, H., Wei, X., Ye, Z., Wei, L., Gong, W., Wang, Y., and Zhu, Z. (2006) Identification of a glyphosate-resistant mutant of rice 5-enolpyruvylshikimate 3-phosphate synthase using a directed evolution strategy. *Plant Physiol* **140**, 184-195

34. Han, H., Vila-Aiub, M. M., Jalaludin, A., Yu, Q., and Powles, S. B. (2017) A double EPSPS gene mutation endowing glyphosate resistance shows a remarkably high resistance cost. *Plant Cell Environ*

35. Chauhan, R. D., Hummel, A., Cermak, T., Starker, C., Bart, R., Voytas, D., and Taylor, N. (2017) Generation of glyphosate tolerant cassava plants through CRISPR/Cas9-mediated gene editing. *In Vitro Cell. Dev. Biol.*

36. Hummel, A. W., Chauhan, R. D., Cermak, T., Mutka, A. M., Vijayaraghavan, A., Boyher, A., Starker, C. G., Bart, R., Voytas, D. F., and Taylor, N. J. (2018) Allele exchange at the EPSPS locus confers glyphosate tolerance in cassava. *Plant Biotechnol J* **16**, 1275-1282

37. Li, J., Meng, X., Zong, Y., Chen, K., Zhang, H., Liu, J., Li, J., and Gao, C. (2016) Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. *Nat Plants* **2**, 16139

38. Sost, D., Schultz, A., and Amrhein, N. (1984) Characterization of a glyphosate-insensitive 5-enolpyruvylshikimic acid-3-phosphate synthase. *FEBS Lett* **173**, 238-242

39. Barry, G. F., Kishore, G. M., Padgette, S. R., and Stallings, W. C. (1997) Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.

40. Mousdale, D. M., and Coggins, J. R. (1984) Purification and properties of 5-enolpyruvylshikimate 3-phosphate synthase from seedlings of Pisum sativum L. *Planta* **160**, 78-83

41. Morrison, J. F. (1982) The slow-binding and slow, tight-binding inhibition of enzyme-catalysed reactions. *Trends in Biochemical Sciences* **7**, 102-105

42. Weiss, U., Davis, B. D., and Mingioli, E. S. (1953) Aromatic biosynthesis X. Identification of an early precursor as 5-dehydroquinic acid. *J. Amer.Chem Soc.* **75**, 5572-5576

43. Vaara, M. (1992) Agents that increase the permeability of the outer membrane. *Microbiol Rev* **56**, 395-411

44. Record, M. T., Jr., Courtenay, E. S., Cayley, D. S., and Guttman, H. J. (1998) Responses of E. coli to osmotic stress: large changes in amounts of cytoplasmic solutes and water. *Trends Biochem Sci* **23**, 143-148

45. Manen, D., and Caro, L. (1991) The replication of plasmid pSC101. *Mol Microbiol* **5**, 233-237

46. Schlegel, S., Genevaux, P., and de Gier, J. W. (2015) De-convoluting the Genetic Adaptations of E. coli C41(DE3) in Real Time Reveals How Alleviating Protein Production Stress Improves Yields. *Cell Reports* **10**, 1758-1766

FOOTNOTE

Amino acids are written as either the one- or three-letter abbreviations approved by the International Union of Pure and Applied Chemists

(https://www.ncbi.nlm.nih.gov/Class/MLACourse/Modules/MolBioReview/iupac_aa_abbreviations.html)

Table 1. Kinetic parameters of maize EPSPS with mutations found in glyphosate resistant weeds

	k_{cat}, min^{-1}	K_m PEP, µM	k_{cat}/K_m	K_i, µM	K_m S3P, µM	k_{gly}[1], min^{-1}	$(k_{cat}/K_m)*K_i$[2]
Zm native[4]	1630 ± 14	9.5 ± 0.3	172	0.066 ± 0.003	13.2 ± 0.6	< lod[3]	11
Zm-P106S[5]	1540 ± 12	11.5 ± 0.5	134	0.33 ± 0.02	15.4 ± 0.4	2.3 ± 0.07	44
Zm-P106L[6]	1760 ± 10	47.0 ± 1.1	37.5	3.94 ± 0.17	27.6 ± 0.7	5.7 ± 0.17	148
Zm-P106L+3[7]	1450 ± 22	10.3 ± 0.7	140	2.34 ± 0.18	17.5 ± 0.6	10.1 ± 0.1	329
Zm-TIPS+2[8]	105 ± 1.0	16.2 ± 0.7	6.5	731 ± 36	27.5 ± 0.7	25.5 ± 0.4	4740
Zm-T102S[9]	1600 ± 13	30.9 ± 1.0	51.8	0.69 ± 0.02	28.8 ± 0.6	2.0 ± 0.09	35.7
Ha-T102S[10]	1800 ± 13	30.4 ± 0.8	59.1	0.75 ± 0.03	21.5 ± 0.5	2.4 ± 0.08	44.3

Values represent mean and standard error for at least 4 sets of rate measurements. For assay procedure and analysis, see Experimental Procedures. Values for K_m were obtained by varying one substrate, with the other present at saturation (10 times its K_m) and therefore are apparent K_m. Values shown for k_{cat} are those obtained from substrate saturation with PEP, which in no case differed significantly from those obtained with S3P.

[1]Enzyme turnover (min^{-1}) at 30 µM PEP and S3P, 1 mM glyphosate (see Results for rational as a fitness parameter)

[2]Calculated with K_m for PEP

[3]lod; limit of detection calculated as Blank + 3xSD of blank. The value is 1.5 min^{-1}.

[4]Native *Zea mays* EPSPS having the amino acid sequence described in Figure 2.

[5]A mutation found in multiple species of glyphosate resistant weeds, here created in maize EPSPS

[6]A mutation found in several species of glyphosate resistant weeds, here created in maize EPSPS

[7]Variant captured from the initial combinatorial library (see Results), having P106L and three additional mutations (A100S, E301S, A337S).

[8]Variant captured from the initial combinatorial library, having T102I, P106S and two other mutations (T277G, A337S)

[9]A mutation found in a glyphosate resistant population of *Tridax procumbens*, here created in maize EPSPS

[10]A mutation found in a glyphosate resistant population of *Tridax procumbens*, here created in *Helianthus annuus* (sunflower) EPSPS.

Table 2. Kinetic parameters of variants at each turn of the optimization cycle of maize EPSPS with G101A as the key desensitizing mutation

	k_{cat}, min^{-1}	K_m PEP, µM	k_{cat}/K_m	K_i, µM	K_m S3P, µM	k_{gly}[1], min^{-1}	$(k_{cat}/K_m)*K_i$[2]
Zm native	1630 ± 14	9.5 ± 0.3	172	0.066 ± 0.003	13.2 ± 0.6	$<$ lod[3]	11
G101A+3[4]	824 ± 15	347 ± 6.8	2.4	1430 ± 137	153 ± 3.9	17.7 ± 0.6	3400
G101A	1000 ± 35	333 ± 26	3.0	1930 ± 40	84.0 ± 2.5	25.7 ± 0.2	5780
H6	397 ± 4	25.1 ± 0.9	15.8	989 ± 31	14.0 ± 0.5	95.5 ± 2.0	15700
C1	517 ± 7	18.9 ± 0.5	27.4	449 ± 48	11.4 ± 0.5	104 ± 1.6	12300
D2	414 ± 38	14.5 ± 2.2	28.6	935 ± 66	11.3 ± 0.3	119 ± 2.0	26700
D2-67	530 ± 17	14.4 ± 0.6	36.8	945 ± 98	10.9 ± 0.4	146 ± 2.0	34800
D2-124	631 ± 5	18.1 ± 0.7	34.9	893 ± 53	11.1 ± 0.7	175 ± 2.5	31130
D2c-A5	741 ± 6	18.1 ± 0.6	40.9	839 ± 40	12.6 ± 1.4	186 ± 4.7	34350
CP4[5]	411 ± 5	15.5 ± 0.4	26.5	1970 ± 276	5.2 ± 0.5	176 ± 1.6	52240
101A-97C[6]	477 ± 3.5	93.0 ± 2.7	5.1	646 ± 30	36.7 ± 1.0	79.3 ± 2.2	3300
101A-332A[6]	764 ± 10	123 ± 6.6	6.2	2390 ± 224	52.5 ± 3.1	47.8 ± 0.5	14800
101A-401G[6]	222 ± 2.3	291 ± 9.0	0.8	2970 ± 183	91.8 ± 4.0	23.4 ± 1.9	2290
ACA[7]	408 ± 3.7	42.5 ± 1.8	9.6	1470 ± 96	31.4 ± 1.1	117 ± 1.8	14100
A5-A101G[8]	135 ± 0.8	3.9 ± 0.2	34.7	0.17 ± 0.01	11.0 ± 0.6	0.2 ± 0.08	5.9

For details on data collection and presentation, see the legend to Table 1.

[1]Enzyme turnover (min^{-1}) at 30 µM PEP and S3P, 1 mM glyphosate (see Results for rational as a fitness parameter)

[2]Calculated with K_m for PEP

[3]lod; limit of detection calculated as Blank + 3xSD of blank. The value is 1.5 min^{-1}.

[4]Variant captured from the initial combinatorial library, having G101A plus three other mutations (E301S, E390G, V437R)

[5]EPSPS from Agrobacterium sp. strain CP4, present in RoundUp Ready crops

[6]Variants of maize EPSPS having only the two mutations shown

[7]Variant having G101A, L97C and V332A. ACA, G101A-L98C and G101A-V332A were accessed commercially by gene synthesis.

[8]D2c-A5 with Ala^{101} changed back to glycine

16

Table 3. Efficiency of translation of mutations from maize variant D2-124 to EPSPS from other crop species

Species[1]	k_{gly}[2], min^{-1}	k_{cat}, min^{-1}	K_m PEP, μM	K_i, μM
Maize	175 ± 2.5	631 ± 5	18.1 ± 0.7	893 ± 53
Rice	174 ± 3.1	707 ± 17	22.3 ± 2.7	893 ± 89
Sorghum	172 ± 8.5	846 ± 2.9	20.7 ± 1.7	404 ± 44
Sunflower	155 ± 7.7	381 ± 7.0	18.0 ± 1.8	918 ± 37
Grapevine	144 ± 11	652 ± 8.1	30.4 ± 1.7	1150 ± 98
Cotton	143 ± 8.4	413 ± 2.7	15.3 ± 0.6	940 ± 85
Cassava	133 ± 6.3	688 ± 8.6	24.7 ± 1.5	1400 ± 228
Soybean	114 ± 4.1	618 ± 4.8	90.3 ± 2.3	3550 ± 374
Wheat	102 ± 6.1	223 ± 4.3	18.9 ± 1.7	820 ± 119

The 18 mutations present in maize variant D2-124 were mapped onto the amino acid sequence of the predicted mature form of EPSPS from the species shown (see Results for details). Proteins were expressed, purified and evaluated as described in Experimental Procedures and the legend to Table 1.
[1]Species, accession: *Zea mays*, CAA44974.1; *Oryza sativa*, AF413082.1; *Sorghum bicolor*, XM_002436379.2; *Helianthus annuus*, XM_022161807.1; *Vitis vinifera*, NC_012021.3; *Gossypium hirsutum*, UniProt A7Y7Y2; *Manihot esculenta*, XM_021758443.1; *Glycine max*, XM_003516991.3; *Triticum aestivum*, ACH72672.1
[2]k_{gly} is the enzyme turnover (min^{-1}) at 30 μM PEP and S3P, 1 mM glyphosate (see Results for rational as a fitness parameter).
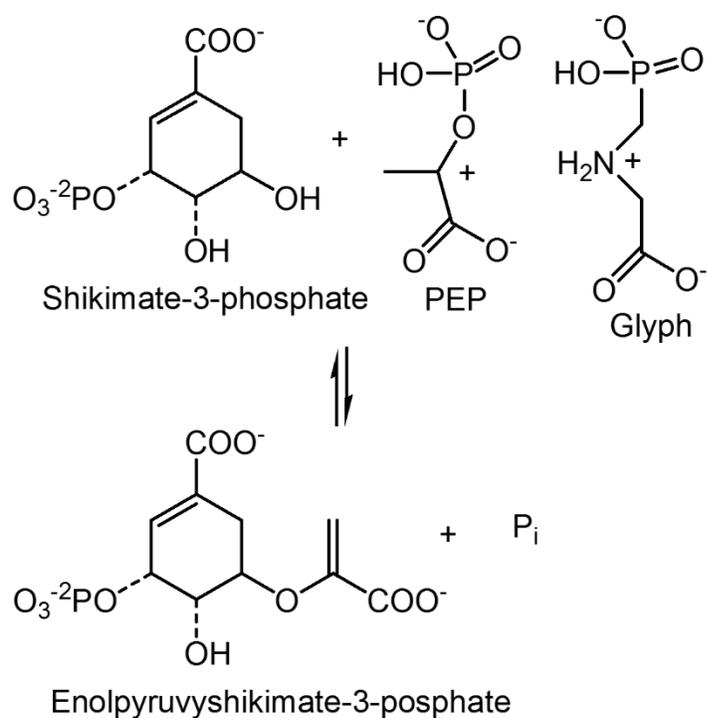
Figure 1. EPSPS reaction

The reaction is an addition/elimination in which an enzymic base deprotonates the 5-hydroxyl of S3P, allowing the electron pair to attack the oxocarbenium ion of PEP (shown to suggest the species mimicked by glyphosate), generated by the enzyme.

18

```
1            11          21          31          41
AGAEEIVLQP  IKEISGTVKL  PGSKSLSNRI  LLLAALSEGT  TVVDNLLNSE
51           61          71          81          91
DVHYMLGALR  TLGLSVEADK  AAKRAVVVGC  GGKFPVEDAK  EEVQLFLGNA
101          111         121         131         141
GTAMRPLTAA  VTAAGGNATY  VLDGVPRMRE  RPIGDLVVGL  KQLGADVDCF
151          161         171         181         191
LGTDCPPVRV  NGIGGLPGGK  VKLSGSISSQ  YLSALLMAAP  LALGDVEIEI
201          211         221         231         241
IDKLISIPYV  EMTLRLMERF  GVKAEHSDSW  DRFYIKGGQK  YKSPKNAYVE
251          261         271         281         291
GDASSASYFL  AGAAITGGTV  TVEGCGTTSL  QGDVKFAEVL  EMMGAKVTWT
301          311         321         331         341
ETSVTVTGPP  REPFGRKHLK  AIDVNMNKMP  DVAMTLAVVA  LFADGPTAIR
351          361         371         381         391
DVASWRVKET  ERMVAIRTEL  TKLGASVEEG  PDYCIITPPE  KLNVTAIDTY
401          411         421         431         441
DDHRMAMAFS  LAACAEVPVT  IRDPGCTRKT  FPDYFDVLST  FVKN
```

Figure 2. Amino acid sequence of the variant termed "native maize EPSPS" as expressed in vector pHD2114. The sequence shown is the reference for all position numbers mentioned in the text. The actual expressed protein has an N-terminal extension consisting of MGHHHHHHHHHHHHSSGHIEGRHM.

19

Variant	1	2	3	53	68	70	71	83	91	97	101	161	207	223	242	278	292	301	332	353	360	390	401	415	437	440
Native	A	G	A	H	A	K	A	K	E	L	G	N	I	K	K	T	M	E	V	A	T	E	D	A	V	T
G101A+3	A	G	A	H	A	K	A	K	E	L	A	N	I	K	K	T	M	S	V	A	T	G	D	A	R	T
H6	R	G	W	H	A	K	Q	R	E	C	A	N	L	K	K	A	M	S	V	A	S	G	G	G	R	Q
C1	R	G	W	H	A	K	A	R	E	C	A	N	L	K	E	T	M	E	V	A	T	P	G	A	V	T
D2	R	G	W	H	H	K	A	R	G	C	A	N	L	K	E	T	M	E	A	A	T	P	G	A	V	T
D2-67	R	K	W	H	H	K	A	R	G	C	A	N	L	K	E	T	M	E	A	G	T	P	G	A	V	T
D2-124	R	K	W	G	H	E	A	R	G	C	A	N	L	R	E	T	M	E	A	G	T	P	G	G	V	T
D2cA5	S	K	W	G	H	E	A	R	G	C	A	R	L	R	E	T	L	P	A	G	T	P	G	G	V	T

Figure 3. Amino acid substitutions present in variants in the progressive optimization of maize EPSPS for activity in the presence of glyphosate. Numbers are the amino acid positions in the reference sequence (Figure 2). Amino acids that differ from those in the native enzyme are highlighted.
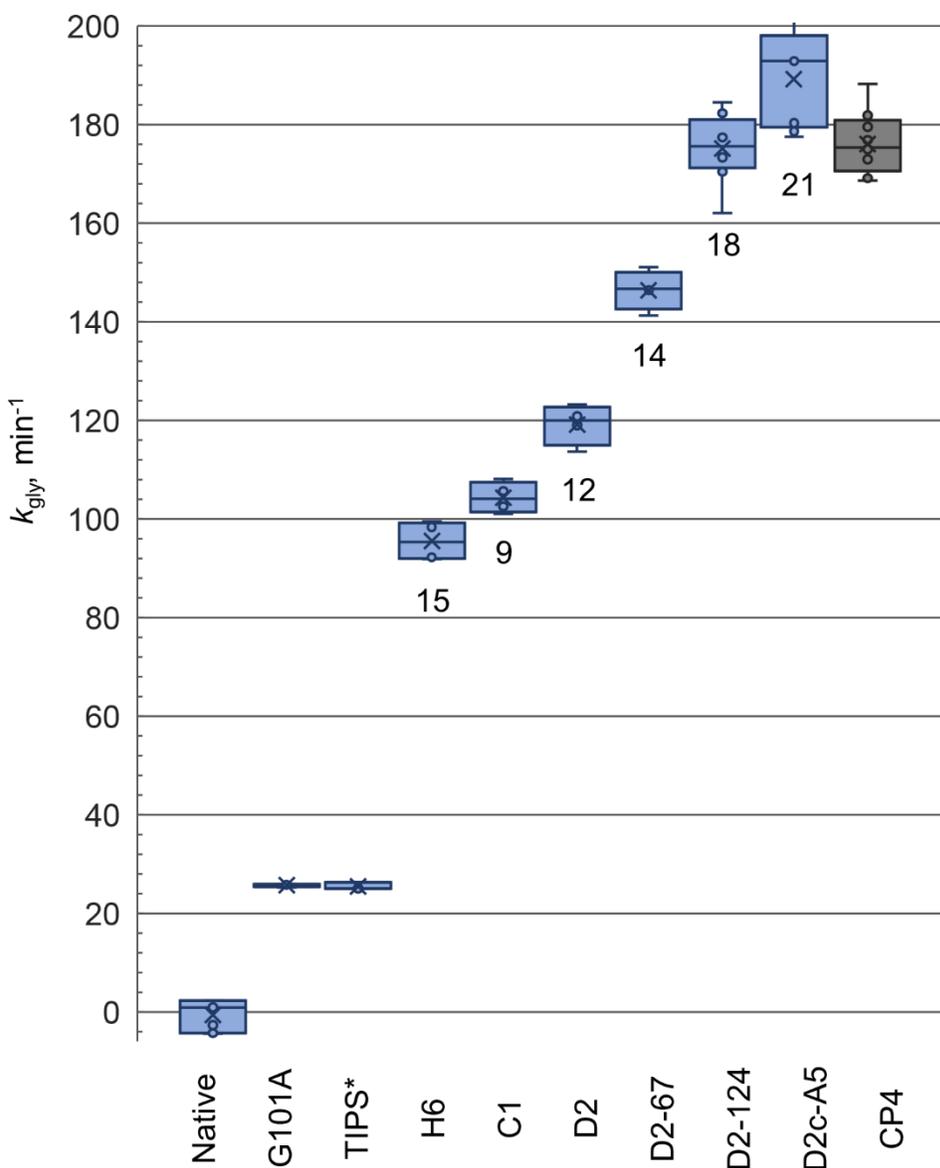
20

Figure 4. Progressive fitness resulting from optimization of maize EPSPS with the G101A mutation.
Values are mean and standard error of 3-6 rate measurements, taken from Table 2.
k_{gly} is enzyme turnover, min^{-1}, under simulated *in vivo* application conditions (30 μM PEP, 30 μM S3P and 1 mM glyphosate; for rationale, see Results).
Number below the box indicates the number of amino acid changes relative to native maize EPSPS. The reduction from 15 to 9 occurred in the H6-C2-native backcross library (see text).
TIPS: Maize EPSPS with the T102I and P106S mutations plus an additional N444G mutation at the C-terminus. TIPS and CP4 are included for reference.
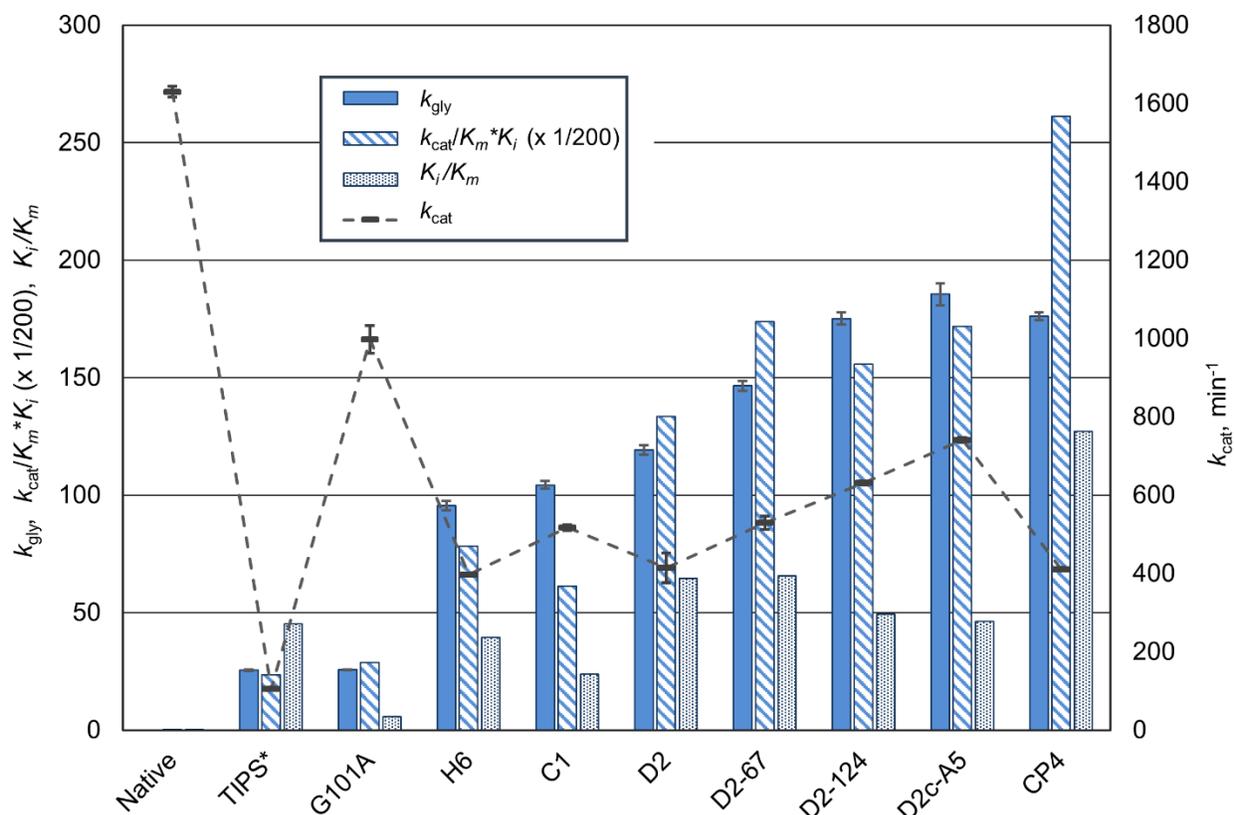
Figure 5. Impact of improvements in k_{cat} and selectivity (K_i/K_m) on fitness of optimized maize EPSPS k_{gly} is enzyme turnover, min^{-1}, under simulated *in vivo* application conditions (30 μM PEP, 30 μM S3P and 1 mM glyphosate; for rationale, see Results). Units for k_{cat} and k_{cat}/K_m*K_i are also min^{-1}. *TIPS: Maize EPSPS with the T102I and P106S mutations plus an additional N444G mutation at the C-terminus.

Desensitizing plant EPSP synthase to glyphosate: Optimized global sequence context accommodates a glycine-to-alanine change in the active site

Yuxia Dong, Emily C. Ng, Jian Lu, Tamara K. Fenwick, Yumin Tao, Sean Bertain, Marian Sandoval, Ericka Bermudez, Zhenglin Hou, Phil Patten, Michael W. Lassner and Daniel L. Siehl

Access the most updated version of this article at doi: 10.1074/jbc.RA118.006134

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts