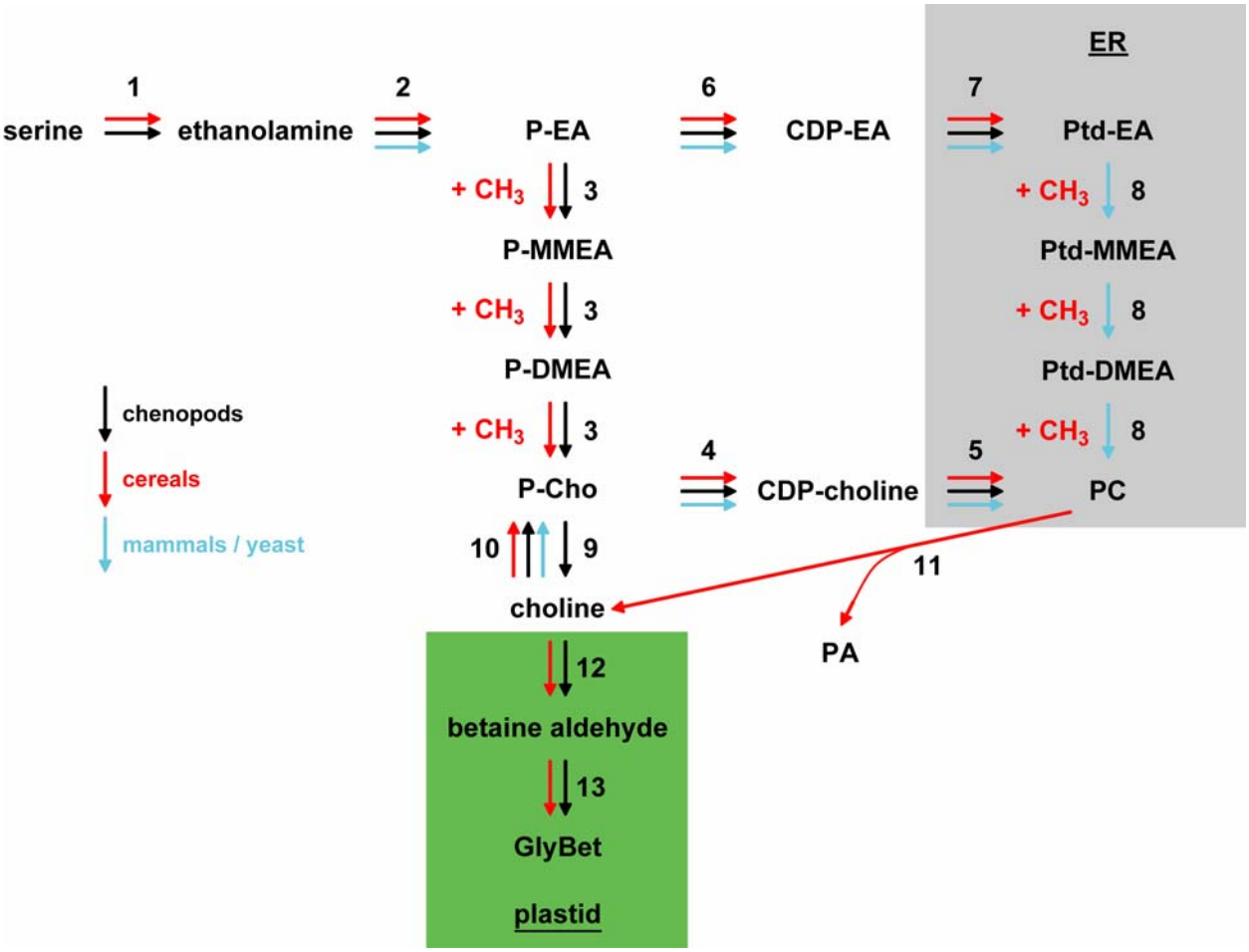
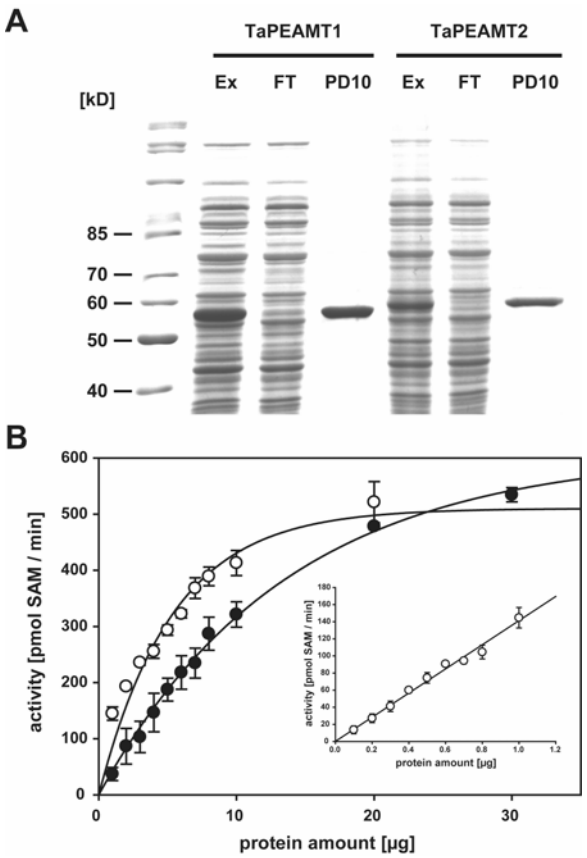
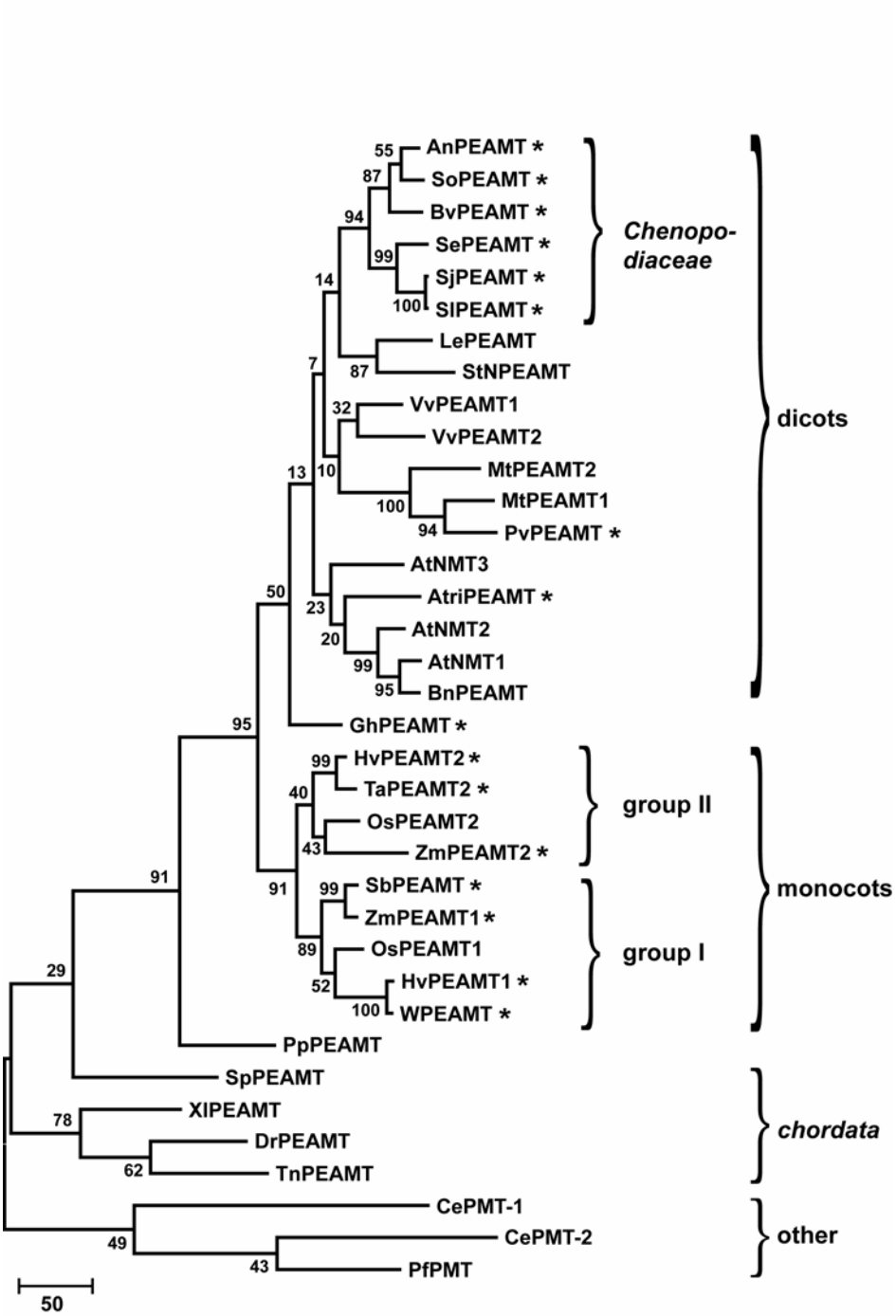
supplemental data



Supplemental figure 1: Eukaryotic pathways for choline and phospholipid biosynthesis. In yeast and mammals reactions 2, 6, 7 and 8 constitute the Bremer-Greenberg pathway, while steps 10, 4 and 5 are referred to as the Kennedy pathway. In plants, but also insects, nematodes and *Plasmodium falciparum* the enzyme for the triple methylation of phosphatidylethanolamine (8) has not been identified to date, therefore the putatively sole entry point for *de novo* phosphatidylcholine biosynthesis is via the PEAMT catalysed reactions that produce phosphocholine (3). In the plant genus of Chenopodiaceae step 9 can generate choline directly from phosphocholine, while this enzyme seems to be missing not only in *Gramineae*, but also most certainly in tobacco and *Arabidopsis* ((1), (2)). Therefore in these species choline can only be produced via the hydrolysis of phosphatidylcholine (11). All eukaryotes seem to be capable of recycling phosphocholine via step 10. Chenopods as well as some cereals such as wheat and barley can accumulate glycine betaine in the chloroplast (steps 12 + 13). ER = endoplasmic reticulum, 1 = serine decarboxylase (SDC), 2 = ethanolamine kinase (EK), 3 = phosphoethanolamine N-methyltransferase (PEAMT), 4 = CTP:phosphocholine cytidylyltransferase (CCT), 5 = CDP-choline:diacylglycerol phosphotransferase (CPT), 6 = CTP:phosphoethanolamine cytidylyltransferase (ECT), 7 = CDP-ethanolamine:diacylglycerol phosphotransferase (EPT), 8 = phosphatidylethanolamine N-methyltransferase (PEMT), 9 = phosphocholine phosphatase (PCP), 10 = choline kinase (CK), 11 = phospholipase D (PLD), 12 = choline monooxygenase (CMO), 13 = betaine aldehyde dehydrogenase (BADH).



<u>Supplemental figure 2:</u> Enzyme purification and protein dependence of PEAMT activity. A. The two wheat PEAMT isoforms were expressed as N-terminal translational 6xHis fusion proteins and purified from 600 ml bacterial culture induced with 0.5 mM IPTG at an OD_{600} of 0.8, then grown overnight at 25ºC. Bacterial pellets were shock-frozen in liquid N_2 and resuspended in buffer (100mM Tris HCl 7.5, 300 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol, 0.3 mM phenyl-methylsulfonyl fluoride) adding 500 units of Benzonase Nuclease HC (Novagen) followed by sonication with a probe tip sonicator. Total cell lysates were centrifuged for 30 min at 10,000 x g. The supernatant was loaded onto a 1-ml HiTrap Ni NTA affinity column (GE Healthcare) and washed with a 5 ml of buffer containing 30mM imidazole, followed by a more stringent wash using buffer containing 50 mM (TaPEAMT1) or 35mM (TaPEAMT2) imidazole. Protein was eluted from the column in buffer containing 250 mM imidazole. The buffer of the eluates was exchanged to 50 mM HEPES pH 7.8, 1 mM EDTA, 15% (v/v) glycerol and 5 mM DTT using a PD-10 column (GE Healthcare). Ex = bacterial protein extract, FT = flow-through, PD10 = desalted purified protein. The size of the fusion proteins matched the expected sizes of 59 and 61 kDa for TaPEAMT1 and 2, respectively. B. Protein dependence of PEAMT activity using 200 µM P-EA, 200 µM SAM spiked with 6.5 µM ^{14}C-SAM and between 1 and 20 µg of recombinant TaPEAMT1 (●) or TaPEAMT2 (○) protein, respectively. In this physiologically relevant substrate concentration range TaPEAMT2 is 4-times more active than TaPEAMT1 showing an increase of initial velocity proportional to the amount of protein used between 0.1 and 1 µg protein (see insert), while the linear section of the TaPEAMT1 plot is found between 1 and 8 µg of protein (sp.A.(TaPEAMT1) = 570 ± 62 pkat / mg protein, sp.A. (TaPEAMT2) = 2353 ± 122 pkat / mg protein, means ± SE , n = 3).



Supplemental figure 3: Phylogenetic analysis of PEAMT proteins. Plant PEAMT proteins fall into two well supported clades, one for monocotyledonous and one for dicotyledonous species. Within the monocot branch, two isoforms can be found for most species which fall into two distinct subgroups. Enzymes from GlyBet accumulator (*****, (3)) and non-accumulator plants appear to be randomly distributed throughout the two clades with some clustering observed for the *Chenopodiaceae*. It is interesting to note that PEAMT proteins can be found in some vertebrates, while missing in mammalian genomes. For the analysis only full-length sequences were aligned using ClustalX1.8b. The midpoint rooted tree was generated with MEGA4 software using the Maximum Parsimony method with 1000 bootstrap repetitions (4). The percentage of replicate trees in which the associated PEAMT proteins clustered together is shown next to the branches.

supplemental table 1: Protein accessions used to construct the phylogenetic tree in supplemental figure 3.

identifier	Species	Acc. No.
AnPEAMT	*Atriplex nummularia*	GenBank BAD80838
AtNMT1	*Arabidopsis thaliana*	TAIR At3g18000
AtNMT2	*Arabidopsis thaliana*	TAIR At1g48600
AtNMT3	*Arabidopsis thaliana*	TAIR At1g73600
AtriPEAMT	*Aster tripolium*	GenBank BAC57960
BnPEAMT	*Brassica napus*	GenBank AY319479
BvPEAMT1	*Beta vulgaris*	GenBank BAE07178
GhPEAMT	*Gossypium hirsutum*	GenBank ABX10444
HvPEAMT2	*Hordeum vulgare*	TIGR TC131933
HvPEAMT1	*Hordeum vulgare*	TIGR TC146761
LePEAMT	*Lycopersicon esculentum*	GenBank AAG59894
OsPEAMT1	*Oryza sativa*	MIPS Os05g47540
OsPEAMT2	*Oryza sativa*	MIPS Os01g50030
MtPEAMT1	*Medicago truncatula*	TIGR AC137078_26.4
MtPEAMT2	*Medicago truncatula*	TIGR AC138199_16.4
PpPEAMT	*Physcomitrella patens*	GenBank XP_001763293
PvPEAMT	*Phaseolus vulgaris*	TIGR TC317
SbPEAMT	*Sorghum bicolour*	TIGR TC102830
SePEAMT	*Salicornia europaea*	GenBank ABG57185
SjPEAMT	*Suaeda japonica*	GenBank BAC57432
SlPEAMT	*Suaeda liaotungensis*	GenBank ABK42071
SoPEAMT	*Spinacia oleracea*	Gen Bank Q9M571
StPEAMT	*Solanum tuberosum*	TIGR TC165642
TaPEAMT1/WPEAMT	*Triticum aestivum*	GenBank AAL40895
TaPEAMT2	*Triticum aestivum*	GenBank FJ803924
VvPEAMT1	*Vitis vinifera*	GenBank CAO68187
VvPEAMT2	*Vitis vinifera*	GenBank CAO42308
ZmPEAMT1	*Zea mays*	GenBank ABF83429
ZmPEAMT2	*Zea mays*	GenBank NP_001105267
Non-plant sequences		
CePMT-1	*Caenorhabditis elegans*	GenBank NP_494990
CePMT-2	*Caenorhabditis elegans*	GenBank NP_504248
DrPEAMT	*Danio rerio*	GenBank NP_001070105
PfPMT	*Plasmodium falciparum*	GenBank AAR08195
SpPEAMT	*Strongylocentrotus purpuratus*	GenBank XP_794381
TnPEAMT	*Tetraodon nigroviridis*	GenBank CAG09731
XlPEAMT	*Xenopus laevis*	GenBank NP_001087172

<u>supplemental table 2:</u> Published enzyme kinetic data from recombinant *Plasmodium falciparum* and *C. elegans* PEAMT proteins (PfPMT and CePMT, respectively).

PfPMT kinetics	according to Pessi *et al.* (5)
P-EA [μM]	25 – 1000 (at 2 mM SAM)
SAM [μM]	25 – 1000 (at 3 mM P-EA)
V_{max} [nmol / (mg * min)]	1.2
K_m(app., P-EA) [μM]	79
K_m(app., SAM) [μM]	153
IC_{50} (P-Cho) [μM]	50
IC_{50} (SAH) [μM]	50
CePMT-1 kinetics	according Brendza et al. (6)
P-EA [μM]	5 – 200
SAM [μM]	15 – 1500
V_{max} [nmol / (mg * min)]	30.4 ± 1.6
K_m(app., P-EA) [μM]	24.1 ± 3.6
K_m(app., SAM) [μM]	223 ± 33
mechanism	random Bi Bi, sequential
K_i(P-Cho) [μM]	6940 ± 615 (0 – 30 mM P-Cho, 500μM SAM)
K_i(SAH) [μM]	9.1 ± 1.2 (0 – 20 μM SAH, 50 μM P-EA)
CePMT-2 kinetics	according to Palavalli et al. (7)
P-DME [μM]	250 – 2000
SAM [μM]	15 – 2500
V_{max} [nmol / (mg * min)]	555 ± 19
K_m(app., P-DME) [μM]	4560 ± 380
K_m(app., SAM) [μM]	660 ± 50
α	0.3 ± 0.09
mechanism	random Bi Bi, sequential
K_i(P-Cho) [μM]	1540 ± 268 (0.5 – 6 mM P-Cho, 2.5 mM SAM)
K_i(SAH) [μM]	9.8 ± 1.2(10 – 500 μM SAH, 10 mM P-DME)

A

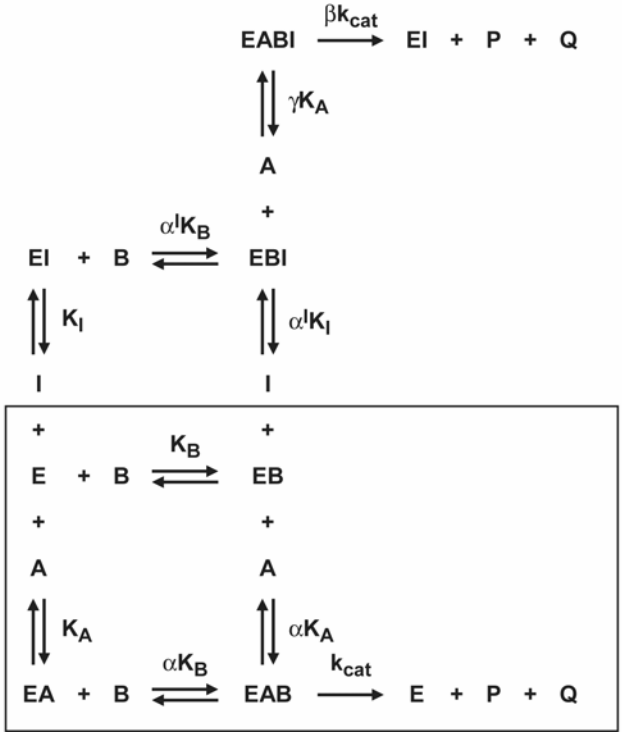
$$(1)\quad v = \frac{V_{max}\,[A][B]}{\alpha\, K_{mA} K_{mB} + K_{mB}[A] + K_{mA}[B] + [A][B]}$$

$$(2)\quad v = V_{max}\left(\frac{1+(\beta\,[I])/(\alpha\, K_i)}{1+[I]/(\alpha\, K_i)}\right) \Big/ \left(1+\frac{(K_m/[S])(1+[I]/K_i)}{1+[I]/(\alpha\, K_i)}\right)$$

$$(3)\quad v = \frac{V_{max}}{(K_m/[S])(1+[I]/K_i)+(1+[I]/(\alpha\, K_i))}$$

B

C



Supplemental figure 4: Equilibria in a random sequential Bi Bi mechanism. A. Best least square fits of the initial velocity data were obtained for the sequential random Bi Bi model (equation 1), supported by R^2 and p values for partial (equation 2) or full (equation 3) mixed inhibition models (8). Here V_{max} is the maximal velocity in the absence of inhibitor ([I]). [A], [B] and [S] are substrate concentrations, K_m, K_{mA} and K_{mB} are the Michaelis-Menten constants for each substrate and *v* is the initial velocity. K_i is the inhibition constant for I binding to free E, while α is the K_m factor change for binding of I to ES complex and β is the factor by which the rate constant is reduced when the product is released from the ESI complex. B. Model for phosphoethanolamine N-methyltransferase catalysis adapted from Cleland (9). This can only be considered as a very coarse model for the overall reaction mechanism since plant PEAMTs have a modular structure with two separate catalytic domains that perform the three sequential methylation steps. It has also been demonstrated that the reaction intermediates mono- and dimethyl phosphoethanolamine are released from the enzyme and need to be re-captured by the enzyme to form phosphocholine (10). C. General representation of various rate constants in the absence (box) or presence of inhibitor according to Segel (8). For both PEAMTs α (without inhibitor) is close to 1, i.e. binding of one substrate does not influence the binding of the other, fitting to a random mechanism for substrate addition. For both P-Cho and SAH inhibition α^I values are always substantially bigger than 1, i.e. both inhibitors decrease the affinity of the EI form for the corresponding substrate, as well as the affinity for the inhibitor, once the corresponding substrate is bound. In suppl. table 2 β is always bigger than 0 in the case of TaPEAMT1, i.e. both inhibitors slow down enzyme catalysis, without bringing it to a stand-still. TaPEAMT1 product inhibition therefore is best described by the hyperbolic mixed-type inhibition system C2 with $1<\alpha<\infty$, $0<\beta<1$ (8): E and EI bind S, but with different affinities, both ES and ESI form product, but at different rates, i.e. it can be seen as a mixture of partial competitive and partial non-competitive inhibition. V_{max} will decrease, K_S will increase and because ESI can form product, the velocity cannot be driven to zero by increasing [I]. For both product inhibitors of TaPEAMT2 β is very close to 0, i.e. in this case linear mixed-type inhibition system C1 with $\alpha>1$, $\beta=0$ best describes the mechanism (8): EI has a lower affinity than E for S and the ESI complex is non-productive. This system is a mixture of partial competitive inhibition and pure non-competitive inhibition. V_{max} decreases and K_S will increase. Because ESI is non-productive, velocity can be driven to zero by increasing [I].

REFERENCES

1. Summers, P. S., and Weretilnyk, E. A. (1993) *Plant Physiol* **103**(4), 1269-1276
2. Tanaka, K., Tolbert, N. E., and Gohlke, A. F. (1966) *Plant Physiol* **41**(2), 307-312
3. Lorenzin, D., Webb, C., Summers, P. S., and Weretilnyk, E. A. (2001) *Can. J. Bot.* **79**, 897 - 904
4. Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007) *Mol Biol Evol* **24**(8), 1596-1599
5. Pessi, G., Kociubinski, G., and Ben Mamoun, C. (2004) *P Natl Acad Sci USA* **101**(16), 6206-6211
6. Brendza, K. M., Haakenson, W., Cahoon, R. E., Hicks, L. M., Palavalli, L. H., Chiapelli, B. J., McLaird, M., McCarter, J. P., Williams, D. J., Hresko, M. C., and Jez, J. M. (2007) *Biochem J* **404**, 439-448
7. Palavalli, L. H., Brendza, K. M., Haakenson, W., Cahoon, R. E., McLaird, M., Hicks, L. M., McCarter, J. P., Williams, D. J., Hresko, M. C., and Jez, J. M. (2006) *Biochemistry* **45**(19), 6056-6065
8. Segel, I. H. (1975) *Enzyme kinetics - behavior and analysis of rapid equilibrium and steady-state enzyme systems*, John Wiley & Sons, New York, London, Sydney, Toronto
9. Cleland, W. W. (1963) *Biochim Biophys Acta* **67**, 104-137
10. Nuccio, M. L., Ziemak, M. J., Henry, S. A., Weretilnyk, E. A., and Hanson, A. D. (2000) *J Biol Chem* **275**(19), 14095-14101