Supporting Information

Exploiting Enzyme Plasticity and Reactivity in Virtual Screening and Ligand Discovery: the discovery of glutamate racemase inhibitors with high ligand efficiency values.

Katie L. Whalen[§], Katherine L. Pankow[§], Steven R. Blanke[†] and M. Ashley Spies[‡§*]

[§]*Department of Biochemistry,* [‡]*Institute for Genomic Biology, and* [†]*Department of Microbiology, University of Illinois, Urbana, Illinois 61801*

RECEIVED DATE (automatically inserted by publisher); ***E-mail: aspies@life.uiuc.edu**

Computational Methods

High-throughput database virtual screening. The docking of the Chemical Computing Group's Conformational Data Base, version 2007, (lead-like compounds, based on the parameters set forth by Oprea[1], and consisting of ~ 1 million compounds from public catalogs of 45 chemical suppliers) into the active site of the reactive conformation of *B. subtilis* glutamate racemase was performed as described in Spies *et al*.,[2] However, the method is briefly outlined as follows: Docking was performed using the program LigandFit[3] within the Cerius2 suite (Accelrys, Inc.) The receptor was an energy minimized snap shot for MD simulations with the glutamate carbanion, which was later validated to lead to *ab initio* QM/MM transition states, for Cα proton transfer, upon geometry optimization, as described in Spies *et al*.[2] All compounds in the database were protonated, energy minimized, and docked in the binding site, performing *in-situ* ligand conformational search and rigid-body minimization of the docked poses in the receptor's binding site. All docked compounds were scored using the LigScore scoring

function available in LigandFit. The resulting docked poses were analyzed for biological and computational consistency by i) verifying that the known carbanion ligand was ranked among the best compounds, and ii) verifying that the carbanion was docked in a geometry that reproduces the initial pose.

Scoring and Ranking of Docked Poses. In the initial virtual screen, detailed above, LigScore was used in ranking docked poses, and the top 500 compounds were selected for further analysis. A more rigorous and computationally expensive ranking was employed hereafter to re-rank these compounds in order to determine compounds for actual experimental examination for GR inhibition activity. The LigX utility within the Molecular Operating Environment (MOE)[4], Version 2008.10, was used to predict the binding free energy for the docked complex (i.e. the pK_i value). The scoring procedure involved first a local energy minimization, in which only the ligand and all atoms (including protein receptor and water) in a sphere of 8 Å from the ligand are allowed to move in the geometry optimization. The force field used in the energy minimization was MMFF94x, using a nonbonded cutoff of 10 Å. After energy minimization, the complex is scored using the MOE's London dG scoring function[4] (*vide infra*), which is similar to the often cited scoring function used in AutoDock[5]. Briefly, the functional form of the London dG scoring function is as follows:

$$\Delta G = c + E_{flex} + \sum_{h-bonds} c_{HB} * f_{HB} + \sum_{atoms\ i} \Delta D_i$$

where c is the average gain/loss of translational and rotational entropy; E_{flex} equals the energy from loss of flexibility of the ligand; f_{HB} is a measure of the degree of imperfections of hydrogen bonds (ranging from zero to 1); D_i is the desolvation energy of atom i.

$$\Delta D_i = c_i R^3 \{ \iiint_{u \in A \cup B} |u|^{-6} du - \iiint_{u \in B} |u|^{-6} du \}$$

where A and B are the protein and ligand volumes, respectively, with atom *i* belonging to B; R is the solvation radius of atom *i*; c_i is the desolvation coefficient of atom *i*. The values of *c*, c_{HB} and c_i were fitted from ~ 400 x-ray crystal structures of protein-ligand complexes and their corresponding experimentally determined pK_i values. The triple integrals were approximated using the Generalized Born integral. For additional information the reader is referred to the MOE (2008.10) manual (Chemical Computing Group, Inc.)[4].

The London dG scoring function was a far superior ranking system than sampling using MMFF94 protein-ligand interaction energies from globally energy minimized complexes, as had been employed on GR studies in the past, which led to only weak competitive inhibitors with low mM binding dissociation constants, as described in Spies *et al.*,[2] The MMFF94 interaction energy approach, as employed previously, was computationally more expensive, and did not locate any μM lead compounds. In the current approach, using the ranking based on the London dG scoring led to the discovery of two μM competitive inhibitors, one weak inhibitor and one compound which was shown to be a colloidal aggregator (by employing the well-established detergent tests of Shoichet and coworkers[6, 7]).

Detailed Docking of **3** *into the Reactive Conformation of GR versus Docking into GR Crystal Structure Using Docking-Energy Minimization*. The goal was to look at all reasonable poses of **3** docked into the reactive form of GR and the crystal structure of GR (PDB 1ZUW), retaining all docked poses and allowing each pose to undergo minimization and rescoring with London dG, as above. The Dock utility of MOE (Version 2008.10) was employed. Briefly, the Alpha Triangle method[4] was employed for the placement stage, followed by scoring using the London dG scoring function. The next step was a refinement stage, consisting of a full force field (MMFF94) energy minimization of each pose into the rigid receptor, followed by a rescoring using the London dG scoring function. This procedure was followed for both the reactive form of GR as receptor and the crystal structure of GR as receptor, respectively. All explicit solvent was included in the receptor, and the energy minimization convergence is defined

as an RMSD gradient of 0.01 kcal/mol/Å^2, with an upper limit of 500 iterations. The MMFF94 force field is used in the energy minimization; receptor residues over 6 Å away from ligand poses are not included in the energy calculation. The final energy was calculated using the Generalized Born solvation model[4].

Calculation of Tanimoto Volumes for **3** *Docked into Reactive Conformation of GR and the Crystal Structure of GR*. Tanimoto volumes (T_{vol}) is a measure of the goodness of overlap between docked poses (or between a docked pose and an experimentally determined receptor-ligand complex), which has the following form[8-10]:

$$T_{vol} = \frac{O_{xy}}{I_x + I_y - O_{xy}}$$

where, O_{xy} is the overlap in molecular volume between two docked poses; I_x and I_y are the molecular volumes of docked poses x and y, respectively. For the case of perfect overlap between pose x and y, O_{xy} is equal to the molecular volume overlap between compounds x and y, which reduces T_{vol} to a value of unity. When there is no overlap at all between poses, then O_{xy} approaches zero, and T_{vol} approaches zero. T_{vol} is considered a superior method for the similarity of placement between complexes, because it accounts for cases where a ligand may be rotated ~ 180˚ (i.e. flipped), but which accurately describes the available and occupied active site space[8-10].

Calculations of Molecular Volumes Using Yasara. The solvent excluded (ie Connolly) surface of **3**, from docked poses, was calculated using the program YASARA[11] (Version 9.5.10), using the 'Numeric' algorithm option, with the default grid resolution.

Experimental Methods

Materials. Materials for Luria-Bertani (LB) medium were purchased from BD Diagnostics (Franklin Lakes, NJ). Isopropyl-β-D-thiogalactoside (IPTG) was purchased

from Fisher Scientific (Fair Lawn, NJ). Ampicillin, dithiothreitol (DTT), and β-mercaptoethanol were obtained from Sigma Aldrich (St. Louis, MO). Amicon centrifugal filter devices with a molecular weigh cutoff of 10,000 were purchased from Millipore (Billerica, MA). Ni-NTA His-Bind Resin was obtained from Invitrogen. Materials for preparing the Ni-NTA bind, wash, and elute buffers were obtained from Sigma Aldrich. Iodonitrotetrazolium chloride, Nicotinamide adenine dinucleotide hydrate (NAD^+), Diaphorase from *Clostridium kluyveri* and L-Glutamate Dehydrogenase from bovine liver (obtained as an ammonium sulfate suspension) was purchased from Sigma Aldrich. Adenosine-5'-Diphosphate, Disodium Salt, Dihydrate was obtained from USB Corporation (Cleveland, OH). Compound **4** (ID# S364169) was obtained from Sigma-Aldrich. Compound **5** (ID# BAS 00124393) was obtained from Asinex (Salem, NC). Compound **1** (ID# MMSA-1076) was obtained from Scientific Exchange (Center Ossipee, NH). Compound **3** (ID# LT00453399) was obtained from Labotest OHG (Niederschöna, Germany). Compound **7** was obtained from InterBioschreen (ID# BB_NC_0417). Compound **8** was obtained from Sigma (ID# 297763). Compound **2** was obtained from Sigma's Rare Chemical Library (ID# S819670).

GR Expression and Purification. The expression and purification of RacE (*B. subtilis* GR)[2], RacE1 and RacE2 (isozymes of *B. anthracis* GR)[12] were as described previously. A 10 mL starter culture of LB medium with 100 μg/mL ampicillin was prepared from the stock *E. coli* BL21 (DE3) cells containing a pET-15b-*racE* plasmid and grown overnight at 37 degrees Celsius with rotation. The 10 mL starter culture was back-diluted into 1 L fresh LB medium with 100 μg/mL ampicillin. Cells were grown at 37 degrees Celsius with shaking until the optical density at 600 nm reached ~ 0.5. *RacE* expression was induced upon addition of a final concentration of 0.1 mM IPTG. Following induction, cells were grown for an additional 16 h at 37 degrees Celsius with shaking. Cells were harvested by centrifugation at 5,000 x *g* for 15 min. Supernatant was discarded and cell pellets were resuspended in Ni-NTA bind buffer (50 mM phosphate, 300 mM NaCl, 20 mM imidazole, 0.5 mM Tris carboxyethyl phosphine, pH 8.0). Cell lysis was achieved through sonication (3x 20 sec cycles, 23 kHz and 20 W), using a 100 Sonic Dimembrator from Fisher Scientific. Insoluble materials were pelleted by centrifugation at 30,000 x *g*

for 30 min and clarified lysate was applied to 4 mL bed volume of Ni-NTA His-Bind resin. Bound protein was washed with 2X 16 mL of wash buffer (50 mM phosphate, 300 mM NaCl, 40mM imidazole, 0.5 mM Tris carboxyethyl phosphine, pH 8.0). Bound protein was eluted twice with 4 mL elution buffer (50 mM phosphate, 300 mM NaCl, 250 mM imidazole, 0.5 mM Tris carboxyethyl phosphine, pH 8.0) and the collected eluant was exchanged into protein storage buffer (50 mM Tris, 100 mM NaCl, 0.2 mM DTT, pH 8.0) utilizing a 10,000 MWCO Amicon centrifugal filter device from Millipore. Concentration by centrifugation and resuspension in protein storage buffer was repeated twice more for a total of 3 buffer exchanges.

Protein concentration was quantified by absorbance spectroscopy based on previously employed methods[13]. Based on the primary amino acid sequence, the extinction coefficient for RacE was calculated to be 24,401 M^{-1} cm^{-1} at 280nm. Absorbance readings were acquired using a Cary 300 Bio UV-Visible Spectrophotometer from Varian Incorporated (Palo Alto, CA). Finally, protein stocks (final concentration = 0.420 mM) were stored in protein storage buffer with 20% glycerol at -20 degrees Celsius.

Enzyme Kinetics – Circular Dichroism. Stereoisomerization of D-glutamate by glutamate racemase was assayed using a J-720 CD spectropolarimeter from JASCO (Easton, MD). A jacketed cylindrical cuvette with a volume of 750 μL and a path length of 10 mm was used for each assay. Readings were measured at 220nm or 225nm depending on contributions to the signal by the inhibitor. All measurements were conducted at 25 ˚C. Concentrations of D-glutamate were varied from 0.25–5 mM in an optically clear borate buffer (50 mM boric acid, 100 mM KCl, 0.7 mM DTT; pH 8.0). Reactions were initiated upon addition of RacE (0.84 μM), RacE2 (0.47 μM) or RacE1 (0.49 μM) . Data acquisition was performed using a JASCO Spectra Manager v1.54A software and Excel, and fitting was performed using GraphPad Prism v5.0 software from GraphPad Software (San Diego, CA). For weaker inhibitors with high absorbance, where higher concentrations were employed, a 1 mm path length cuvette was used in order to

avoid saturating the instrument. Data was also presented as Lineweaver-Burke plots, using KaleidaGraph v4.03 by Synergy Software (Reading, PA).

No Observed Time-Based Inhibition of GR by 5. Inhibitors were analyzed for time-dependent inhibition by mixing RacE (0.56 μM) and inhibitor (50 μM) in optically clear borate buffer (50 mM boric acid, 100 mM KCl, 0.7 mM DTT; pH 8.0) and incubating on ice for various amounts of time (0 min, 2.5 min, 5 min, and 10 min). Reactions were initiated upon addition of D-glutamate (1 mM) and initial velocity was measured using circular dichroism as described above. Differences in initial velocity are negligible (less than 10%) for both compounds (Figure S1).

Testing for Colloidal Aggregation. It has been well established that the identification of non-competitive inhibition, which has often arisen in high throughput screenings (both virtual and non-virtual), may be the result of colloidal aggregation of the ligand of interest. This phenomenon occurs exclusively with (apparent) non-competitive, not competitive inhibitors, and has been extensively characterized and reviewed by Shoichet and co-workers[6, 7]. In the current study, a single non-competitive inhibitor was identified (**1**; Table 1). Thus, **1** was analyzed for the possibility of colloidal aggregation using a previously established detergent-based assay[6]. First, we describe the coupled enzyme assay used in this study[14], followed by the details of the detergent assay for colloidal aggregation.

Enzyme Kinetics – Coupled-Enzyme Assay. The D- to L-glutamate racemization activities of RacE, RacE1 and RacE2 were assayed through a previously established coupled method[14], utilizing L-glutamate dehydrogenase and diaphorase. Absorption data was collected on a Cary 300 UV-VIS Spectrophotometer from Varian Incorporated (Palo Alto, CA). In this coupled-enzyme assay, the stereoinversion of D-glutamate by glutamate racemase provided the L-glutamate required by L-glutamate dehydrogenase. The NADH by-product of dehydrogenation is then oxidized by diaphorase to produce the

reduced form of iodonitrotetrazolium (INT). Reduced iodonitrotetrazolium can be quantified by measuring the absorption at 500 nm.

Coupled-Enzyme Assay Calibration. The stoichiometric relationship between absorption at 500 nm and the concentration of L-glutamate produced was calculated to be approximately 0.00103 AU/nmol, with some variance depending on the absorptivity of each new batch of iodonitrotetrazolium (INT). Absorptivity of INT was determined by titrating 10 nmol-aliquots of L-glutamate into a cuvette containing diaphorase, INT, L-glutamate dehydrogenase, ADP and NAD^+. Absorption at 500 nm is plotted as a function of L-glutamate added. The slope of this plot is the conversion factor used in the coupled-enzyme assay to convert AU/min to nmol/min.

Coupled-Enzyme Assay: test for inhibition of glutamate dehydrogenase coupling enzyme. The coupled-enzyme assay was utilized in order to determine any possible incidence of L-glutamate dehydrogenase inhibition by the added inhibitor. Michaelis-Menten curves were composed under three sets of conditions: without inhibitor, with a concentration of inhibitor that causes significant inhibition (of the full coupled system), and with the same concentration of inhibitor and twice the concentration of glutamate racemase. All other reagent concentrations remain the same. If L-glutamate dehydrogenase is not being inhibited, one would expect to see an exactly two-fold increase in activity in the presence of two-fold more glutamate racemase[15]. In other words, the ratio of initial velocity in the presence of 1X enzyme to the initial velocity in the presence of 2X enzyme should be 1:2. For **1**, the ratio is 1: 1.91 ± 0.07, which is within error due to factors of the coupled-enzyme system. Thus, **1** would not appear to be inhibiting the coupled enzyme and any inhibition witnessed is that of glutamate racemase alone. For **4**, the ratio is 1: 1.75 ± 0.01 indicating that some, but not the majority, of the total inhibition is contributed by inhibition of L-glutamate dehydrogenase. It should be noted that neither compounds **1** nor **4** could be studied using circular dichroism due to their high absorbance in the UV range.

Detergent Assay for Colloidal Aggregation (false positive for non-competitive inhibition)-Activity of RacE was measured in the presence and absence of inhibitor in buffer containing 0.01% Triton X-100 (vol/vol). The percent inhibition was compared to

that acquired when conducting the same measurements in buffer without Triton-X 100. If the inhibitor is indeed aggregating, one would expect to see a decrease in the percent inhibition in the presence of detergent. A greater than two-fold decrease in inhibition confirms colloidal aggregation[6]. These tests were all performed using the coupled enzyme assay[14] (*vide supra*), together with controls and calibration. Figure 3 shows that **1** is, indeed, a colloidal aggregator (**1** was the only non-competitive inhibitor found in the current study). As a further control for this study, we also performed said detergent test on the competitive inhibitors **3** and **5**, which did not have any effect on rate as a function of increasing concentrations of detergent, as expected.

Figures

Figure 1

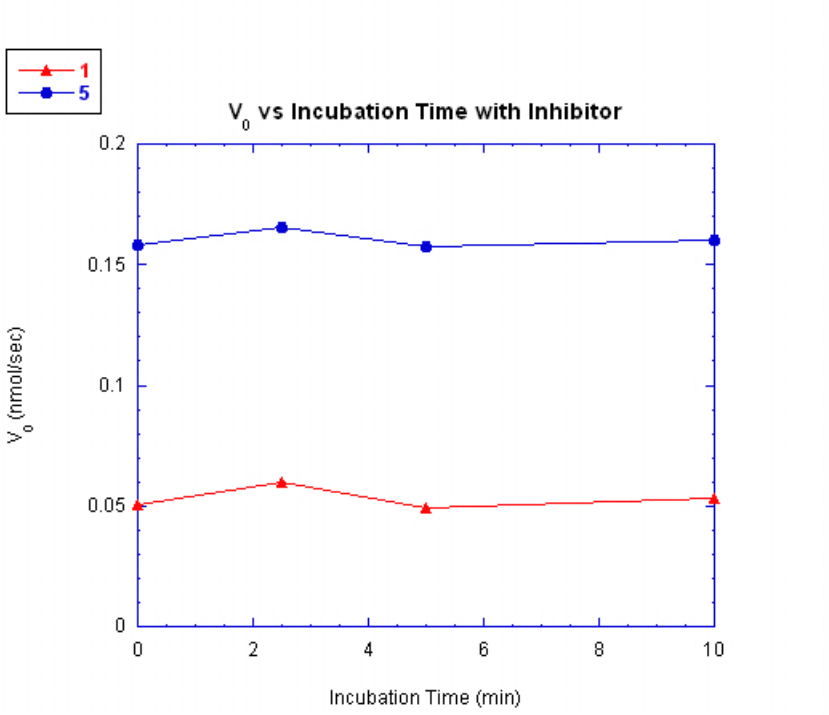

Figure S1. Initial velocity of stereoinversion of D-glutamate catalyzed by RacE in the presence of inhibitor (**3** or **5**) as a function of increasing incubation time. Differences in initial velocity are negligible (less than 10%) for both compounds.

Figure S2 A.

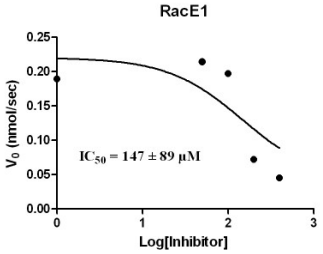

Figure S2 B.

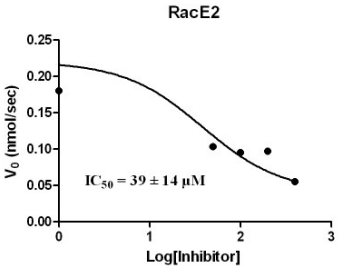

Figure S2. IC_{50} curves for compound **3** against *Bacillus anthracis* glutamate racemase isozymes, RacE1 **(a)** and RacE2 **(b)**. The Log[Inhibitor] vs. Response equation (constraints: TOP = 0.22, BOTTOM = 0.04, and SLOPE = -1.0) was fit using GraphPad Prism v5.0 software.

Figure S4

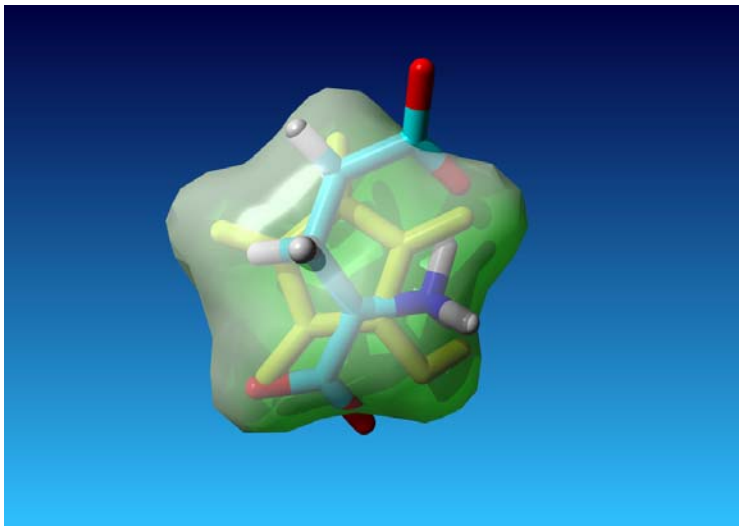

Figure S3. *Superpose of GR complex with* **3** *onto GR complex with Cyclic Glutamate Carbanion*. The top scoring complex from docked compound **3** into the reactive form of GR (as outlined in Figure 1) is shown with its molecular volume, while the cyclic glutamate carbanion is rendered in stick form. The overlap image derives from a full superpose of the two complexes, and only the ligands are shown for clarity. The T_{vol} value (see Computational Methods section, above, for a full description of this metric) for this overlap was calculated to be 0.55. This indicates a reasonable volume overlap between the two complexes; however, there is clearly a far from ideal electrostatic mapping between the two complexes, especially around the C_β and C_γ carbons of the carbanion. The docking procedure is described above in the Computational Methods section.

References Cited

1. Oprea, T. I. Property distribution of drug-related chemical databases. *J Comput Aided Mol Des* **2000,** 14, 251-64.
2. Spies, M. A.; Reese, J. G.; Dodd, D.; Pankow, K. L.; Blanke, S. R.; Baudry, J. Determinants of Catalytic Power and Ligand Binding in Glutamate Racemase. *J Am Chem Soc* **2009**.
3. Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J Mol Graph Model* **2003,** 21, 289-307.
4. *Molecular Operating Environment*, Version 2008.10; Chemical Computing Group: Montreal, Quebec, Canada.
5. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; K., B. R.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998,** 19, 1639-1662.
6. Feng, B. Y.; Shoichet, B. K. A detergent-based assay for the detection of promiscuous inhibitors. *Nat Protoc* **2006,** 1, 550-3.
7. Shoichet, B. K. Screening in a spirit haunted world. *Drug Discov Today* **2006,** 11, 607-15.
8. Willett, P.; Javidi, B.; Lops, M. Analysis of image detection based on fourier plane nonlinear filtering in a joint transform correlator. *Appl Opt* **1998,** 37, 1329-41.
9. Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *J Med Chem* **2006,** 49, 5912-31.
10. Muchmore, S. W.; Souers, A. J.; Akritopoulou-Zanze, I. The use of three-dimensional shape and electrostatic similarity searching in the identification of a melanin-concentrating hormone receptor 1 antagonist. *Chem Biol Drug Des* **2006,** 67, 174-6.
11. Krieger, E.; Koraimann, G.; Vriend, G. Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field. *Proteins* **2002,** 47, 393-402.
12. Dodd, D.; Reese, J. G.; Louer, C. R.; Ballard, J. D.; Spies, M. A.; Blanke, S. R. Functional comparison of the two Bacillus anthracis glutamate racemases. *J Bacteriol* **2007,** 189, 5265-75.
13. Gill, S. C.; von Hippel, P. H. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* **1989,** 182, 319-26.
14. Rej, R. A convenient continuous-rate spectrophotometric method for determination of amino acid substrate specificity of aminotransferases: application to isoenzymes of aspartate aminotransferase. *Anal Biochem* **1982,** 119, 205-10.
15. Cook, P. F.; Cleland, W. W. *Enzyme kinetics and mechanism*. Garland Science: London ; New York, 2007; p xxii, 404 p.