

Supporting Information

Vosyka et al. 10.1073/pnas.1215076110

SI Methods

Protein Purification. Transformed bacteria [*Escherichia coli* strain BL21 (DE3) for AarA, GlpG, and deformylase overexpression; *E. coli* Δ glpEGR::kan derived from *E. coli* MG1655 for TatA overexpression] were induced at an $OD_{600} = 0.8$ (TatA and deformylase) and $OD_{600} = 0.6$ (GlpG and AarA). Protein expression was performed for 4 h at 37 °C (TatA and deformylase) and 18 h at 16 °C (GlpG and AarA). The pelleted cells were lysed in buffer (20 mM Hepes, pH 7.4; 10 mM NaCl; 10% glycerol; and Roche Complete inhibitor mix) by French press and sequentially centrifuged 3,000 \times g, 15 min, 4 °C to remove unbroken cells, and at 100,000 \times g, 30 min, 4 °C to pellet membrane fraction. The membrane pellet was resuspended in buffer (20 mM Hepes-NaOH, pH 7.4; 10% (vol/vol) glycerol; 300 mM NaCl; and 10 mM imidazole) and solubilized by adding 1.5% (wt/vol) β -D-dodecyl-maltoside (DDM) and left shaking for 3 h at 4 °C. Centrifugation at 100,000 \times g, 30 min, 4 °C removes the unsolubilized membrane debris, and supernatant is subjected to nickel-nitrilotriacetic acid beads (Qiagen; 0.5 mL beads per 2 l bacterial expression culture) for several hours and eluted by standard imidazole washing steps (25 mM, 50 mM, 100 mM, and 750 mM final elution). Dialysis was performed against buffer (20 mM Hepes-NaOH, pH 7.4; 10% (vol/vol) glycerol; 300 mM NaCl; 10 mM imidazole; and 0.05% DDM for rhomboids AarA and GlpG; 0.0125% DDM for TatA and deformylase; dialysis membrane with 1-kDa MW cutoff).

Deformylation of TatA and Ionization Factor. For deformylation of TatA, peptide deformylase enzyme (PDF) was added in a 1:5 molar ratio and incubated until full deformylation of TatA was achieved (~2 h, tested by MALDI-MS; Fig. S2). The ionization factor was determined by mixing formylated or deformylated TatA and the proteolytic product (100% cleavage by AarA) in a 1:1 molar ratio. Differences in signal intensities of these molecules during MALDI-MS analysis were used to calculate a correction factor for the different ionizability of these species. This factor was used to normalize screening and titration data.

Crystallization. Crystals of GlpG were obtained by mixing a solution of 2.5–3.0 M ammonium chloride with protein (~5–7 mg/mL) at a ratio of 1:1 in hanging drops at 25 °C. Inhibitors were dissolved in 100% DMSO at a concentration of 10 mM and stored at –20 °C. For soaking, inhibitors were diluted to 2 mM in buffer resembling the mother liquor (25 mM Bis-Tris, pH7.0, and 2.5 M ammonium chloride) with 10% DMSO. Single crystals were incubated in a diluted solution with final inhibitor concentrations between 0.35 mM and 1 mM for a wide period ranging from 30 min to 48 h. All crystals were cryoprotected by adding 25% glycerol (by volume) to the mother liquor and flash frozen in liquid nitrogen. There was a general tendency of crystals to lose their diffracting ability upon soaking with more hydrophobic substitutions, and incubation for a short period showed either no reaction or very low occupancy. The structure described here was collected from crystals soaked with 4-chloro-isocoumarin (IC) **16** for 48 h at a final concentration of 0.5 mM.

Datasets were collected on the I02 beam line at the Diamond Light Source (Harwell). Diffraction data were indexed and integrated with XDS and reduced with SCALA (1, 2). The structure of WT GlpG in trigonal crystal form [Protein Data Bank (PDB) ID code 2XOV], with residues 245–249 corresponding to L5, active-site serine 201 and H150 omitted, was used as an initial input model for Phaser (3). The ligand dictionary file was generated with

GRADE server, and the link file was generated with JLigand (4). Manual model building with Coot (5) and restrained refinement of the structure was performed with Phenix (6) and Refmac with riding hydrogen atoms included in the refinement (7). Structural figures were generated with PyMol (8).

Docking of Substrate. The tetrapeptide substrate formyl-Thr-Ala-Ala-Phe-NH₂ was built as a tetrahedral intermediate resulting from si-face attack onto the scissile peptide bond between the two alanine residues. The substrate was geometry optimized with an MMFF94 force field. The substrate was then defined as an extension of the side chain of the GlpG S201 (PDB ID code 2XOV, from which the inhibitor structure was deleted). Docking of the substrate as a flexible side chain of S201 was performed with AutoDock Vina (9).

In Vivo Labeling of GlpG. *E. coli* cells carrying a plasmid for GlpG expression were grown to $OD = 0.6$, and expression was induced by adding 1 mM isopropyl β -D-1-thiogalactopyranoside for 2 h. Cells were harvested by centrifugation (45 min, 4 °C, 30,000 \times g), washed with 1 \times PBS buffer, resuspended in 4 mL 1 \times PBS containing 1 mM EDTA, and incubated at 37 °C for 30 min. First, 100 μ L of this suspension was treated with 200 μ M IC **16** or DMSO for 30 min, then it was incubated with 10 μ M activity-based probe (ABP) **36** for 30 min and subsequently lysed by adding SDS-sample buffer.

Labeling of Endogenous GlpG. *E. coli* cell strains BL21(DE3)gold and Δ glpG were grown at 37 °C to $OD_{600} = 0.8$ and subsequently lysed with a French press. Cell debris was removed by centrifugation, and membranes were isolated by ultracentrifugation (100,000 \times g, 4 °C, 30 min), resuspended in 50 mM Hepes buffer, and stored at –80 °C until use. Next, 1 mg/mL of total protein was incubated with 2 μ M IC **36** at 37 °C for 30 min and resolved by SDS/PAGE.

In Vivo Inhibition of AarA in *Providencia stuartii*. *P. stuartii* cells (strain DSM4539) were grown in LB medium at 37 °C to $OD = 0.3$, diluted (1:1) with LB containing 1 mM EDTA, and incubated for 30 min. Thereafter, 1 mL each was transferred to 24-well culture plates and incubated with 100 μ M AarA inhibitors or DMSO for 1 h. Cells then were analyzed directly by microscopy (100 \times magnification).

ABP Labeling of Rhomboids in Different Detergents. While shaking, 10 ng GlpG in 50 mM Hepes buffer (pH 7.5) containing 5 \times critical micelle concentration of octyl maltoside, decyl maltoside, DDM, or tetradecyl maltoside was incubated with ABP **36** at 37 °C in the dark, and samples were taken after 10, 30, and 60 min. The labeling reaction was quenched by the addition of 1 \times sample buffer. The samples were separated by a 15% Tris-glycine SDS-polyacrylamide gel and visualized on a fluorescent scanner at 546 nm excitation and 574 nm emission. Fluorescent band intensity was determined densitometrically using ImageJ (10).

SI Synthesis of Compounds

Compound 35 (Scheme S1). To a solution of **32** (55 mg, 0.20 mmol) in dichloromethane (2.0 mL) was added a solution of HBr in acetic acid (33%, 2.0 mL). After stirring at room temperature for 30 min, the solvents were removed in vacuo. The residue was coevaporated twice with chloroform to afford the crude HBr-salt. To this salt was directly added azido-phenylalanine (**11**) (23 mg, 0.12 mmol), benzotriazol-1-yloxy-tris(dimethylamino)-phos-

phonium hexafluorophosphate (57 mg, 0.13 mmol), dichloromethane (2.0 mL), and diisopropylethylamine (49 μ L, 0.28 mmol). After stirring for 1 h at room temperature, the solvent was evaporated. The residue was dissolved in EtOAc (20 mL) and washed with KHSO_4 (1.0 M, 2×10 mL), NaHCO_3 (5% m/m, 10 mL), and brine. Drying (Na_2SO_4) followed by column chromatography (gradient from 2% acetone/ CH_2Cl_2 to 2.5%, and finally to 3%) afforded **35** as a white solid (19.7 mg, 33%). ^1H NMR (300 MHz, acetone- d_6): δ = 2.84, [dd, J_{gem} = 14.1 Hz, J_{vic} = 9.4 Hz, 1H, CH^aPh (N_3Phe)], 2.96 (dd, J_{gem} = 13.8 Hz, J_{vic} = 8.4 Hz, 1H, CH^aPh ($\text{Phe-}\psi[\text{CH}_2\text{SO}_2]$)], 3.06 [dd, J_{gem} = 13.8 Hz, J_{vic} = 6.0 Hz, 1H, CH^bPh ($\text{Phe-}\psi[\text{CH}_2\text{SO}_2]$)], 3.16, [dd, J_{gem} = 14.1 Hz, J_{vic} = 4.6 Hz, 1H, CH^bPh (N_3Phe)], 3.34 (dd, J_{gem} = 14.6 Hz, J_{vic} = 5.0 Hz, 1H, $\text{NCHCH}^a\text{SO}_2$), 3.42 (dd, J_{gem} = 14.6 Hz, J_{vic} = 7.6 Hz, 1H, $\text{NCHCH}^b\text{SO}_2$), 4.01 (m, 2H, $\text{NCHCH}_2\text{SO}_2$), 4.07, 4.10 (2d, 2H, N_3CH), 4.61 [m, 1H, NCH ($\text{Phe-}\psi[\text{CH}_2\text{SO}_2]$)], 6.70 (t, J = 6.2 Hz, 1H, NHCH_2CH_2), 7.26 (m,

10H, $2 \times \text{C}_6\text{H}_5$), 6.70 (d, J = 8.4 Hz, 1H, NHCHCH_2); ^{13}C NMR (75 MHz, acetone- d_6): δ = 38.3, 38.5 (NCH_2), 40.4 (NCHCH_2Ph), 38.3 (NHCH), 51.9, 52.1 ($\text{CH}_2\text{SO}_2\text{F}$), 55.2 (CHCH_2SO_2), 65.2 (N_3CH), 127.5, 127.7, 129.3, 130.2, 130.3, 138.0, 138.4 ($2 \times \text{C}_6\text{H}_5$), 169.9 (C = O); ^{19}F NMR (282 MHz, acetone- d_6): δ = -144.6 (s); IR (KBr): 2,122 cm^{-1} (N_3). High resolution mass spectrometry (HRMS) calcd for $\text{C}_{20}\text{H}_{24}\text{FN}_5\text{NaO}_5\text{S}_2$ [$\text{M} + \text{Na}$] 520.1098; found 520.1038.

Compound 36 (Scheme S2). To a solution of **6** (0.82 mg; 1.5 eq) and **5** (6)-carboxamido-(3-azidopropyl)-tetramethylrhodamine (1.06 mg; 1 eq) in THF (0.5 mL) was added 0.1 mg CuBr and 5 μ L of a 100-mM tris-(benzyltriazolylmethyl)amine solution. The mixture was stirred at 50 $^\circ\text{C}$ overnight. The solvent was evaporated, and the residue was dissolved in DMSO and purified by HPLC, giving 0.80 mg of a red solid (50% yield). HRMS calcd for $\text{C}_{41}\text{H}_{39}\text{ClN}_7\text{O}_7$ [$\text{M} + \text{H}$] 776.2599; found 776.2618.

- Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1):72–82.
- Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):133–144.
- McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* 63(Pt 1):32–41.
- Lebedev AA, et al. (2012) Jligand: A graphical tool for the CCP4 template-restraint library. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):431–440.
- Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
- Adams PD, et al. (2011) The Phenix software for automated determination of macromolecular structures. *Methods* 55(1):94–106.
- Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53 (Pt 3):240–255.
- Delano WL (2002) The Pymol Molecular Graphics Systems. Available at www.pymol.org.
- Trott O, Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31(2):455–461.
- Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675.
- Goddard-Borger ED, Stick RV (2007) An efficient, inexpensive, and shelf-stable diazotransfer reagent: Imidazole-1-sulfonyl azide hydrochloride. *Org Lett* 9(19): 3797–3800.

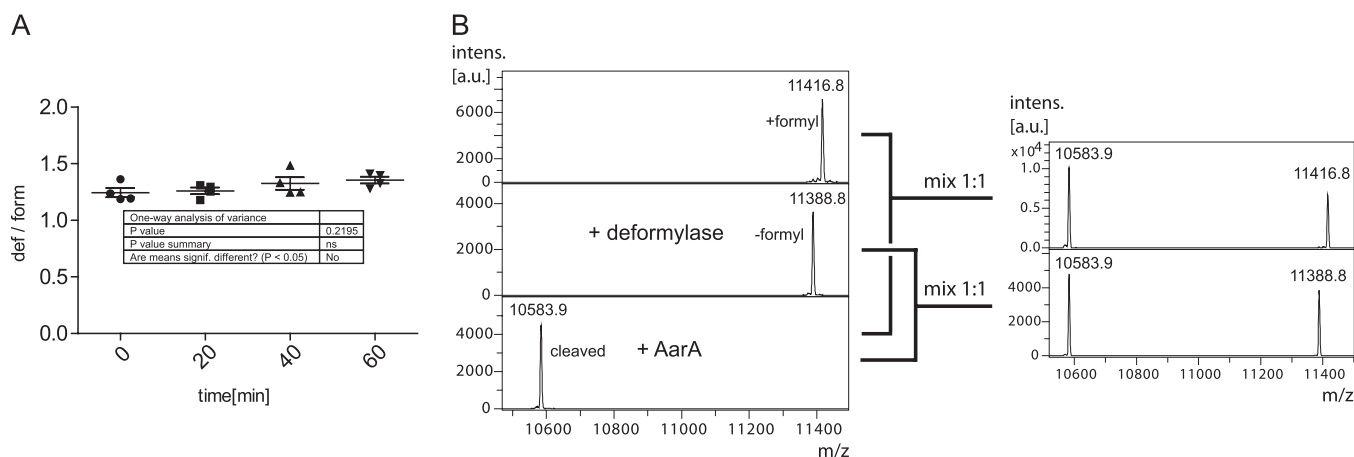


Fig. S1. Comparison of rhomboid cleavage kinetics and determination of the ratio of ionizability of the uncleaved and cleaved TatA substrate species. (A) The ratio (def/form) between the signal intensities of deformylated and formylated TatA over time was calculated from four independent cleavage reactions by rhomboid AarA. One-way analysis of variance shows there is no significant difference in the ratio at different time points. (B) Fully formylated TatA (87 μM) was in vitro deformylated using recombinant *E. coli* peptide deformylase (PDF; 10 μM) until full deformylation was achieved (2–3 h; tested by MALDI-MS analysis). TatA was processed further by adding AarA (5 μM) until quantitative cleavage was achieved (0.5–1 h; tested by MALDI-MS analysis). (Right) All TatA species were mixed with 2% of TFA to quench enzymatic activity of AarA and PDF. In four independent experiments, each formylated and deformylated TatA was mixed with fully cleaved TatA in a 1:1 molar ratio and analyzed directly by MALDI-MS. The ratio between signal intensities of the cleaved and the formylated or the deformylated TatA species was further used as ionization factors ($f_{\text{formylated}} = 1.56$ and $f_{\text{deformylated}} = 1.32$) to normalize all MALDI-MS measurements of TatA cleavage.

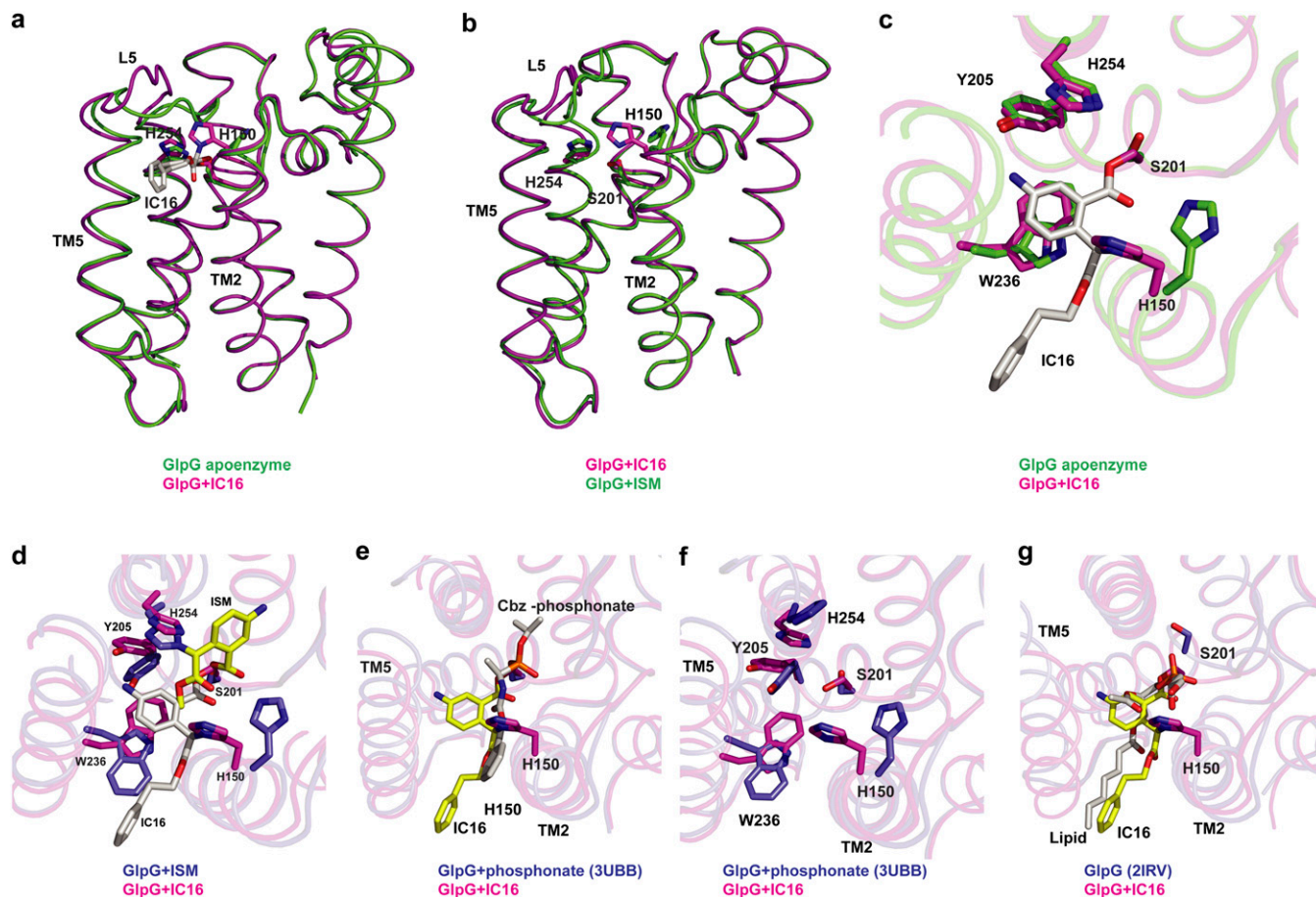
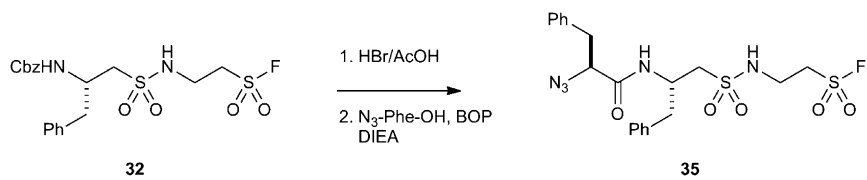
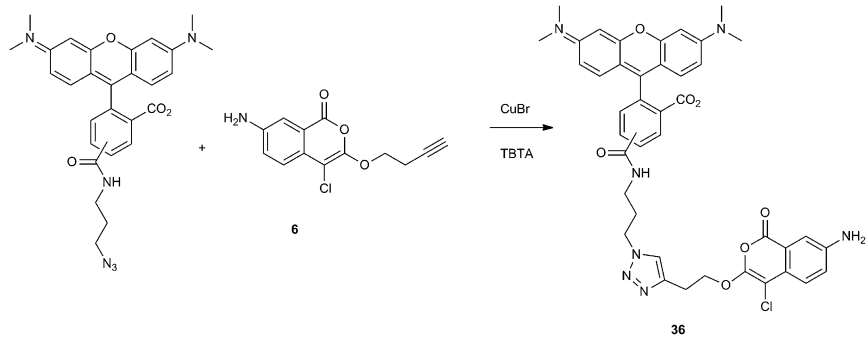


Fig. S3. Structural comparison of GlpG–IC 16 complex with apoenzyme and other structures. (A) Comparison of GlpG apoenzyme (2XOV) and GlpG in complex with IC 16 (3ZEB). The inhibitor molecule and key residues that interact with the inhibitor are shown in stick representation. The carbon atoms of IC 16 are colored in white and the amino acids in magenta. Major differences are observed in transmembrane (TM) helix 5, L5 and L1 (residues 128–135). (B) Comparison of GlpG in complex with IC 16 and methoxy isocoumarin (ISM; ring-opened form of JLK-6) (2XOW). In both these structures, the inhibitor is doubly bonded to the enzyme. The structural change in TM5 is very similar in both these structures. However, in the IC 16 structure, L5 adapts a slightly different conformation. As the inhibitor doubly bonds to S201 and H150, it leaves the active site H254 and, consequently, Y205 unperturbed. The double-bonded ISM structure (with S201 and H254) also results in a change in TM6, which is minimal in the IC 16 structure. (C) An overlay of the active-site residues of the apoenzyme and GlpG in complex with IC 16. Movement of H150 and the formation of a covalent bond with the inhibitor are seen clearly. (D) Two distinct modes of isocoumarin binding in GlpG. The active-site residues of GlpG in complex with either ISM (2XOW) or IC 16 are highlighted. The carbon atoms of ISM molecule are colored yellow, and those of IC 16 are in white. In the complex with ISM, the second bond is formed with H254, whereas in IC 16, the second bond is formed with H150. This leaves the side chains of Y205, W236, and H254 closer to the conformations of the apoenzyme. Perhaps because of this arrangement, a hydrophobic cavity observed in ISM structure is not formed in complex with IC 16. (E) The formation of the alkylated acyl enzyme of GlpG with IC 16 and the position of the inhibitor is comparable with the position of Cbz-phosphonate that forms a single covalent bond to S201 (PDB ID code 3UBB). The carbon atoms of Cbz-phosphonate are shown in white and those of IC 16 in yellow. (F) The active site of GlpG in complex with IC 16 or Cbz-phosphonate. The reaction of the phosphonate with S201 results in displacement of H254, which in turn leads to a different conformation of Y205, whereas these changes are not observed in IC 16 structure. (G) The position of the IC 16 inhibitor also is comparable with the position of the lipid molecule observed previously in a structure of GlpG by Ben-Shem et al. (1) (PDB ID code 2IRV). However, in the IC 16 structure, the extent of change in TM5 and L5 is minimal compared with the structure with the lipid at the active site.

1. Ben-Shem A, Fass D, Bibi E (2007) Structural basis for intramembrane proteolysis by rhomboid serine proteases. *Proc Natl Acad Sci USA* 104(2):462–466.



Scheme S1. Synthesis of probe 35.



Scheme S2. Synthesis of probe 36.

Table S1. Compounds used in MALDI inhibition assay

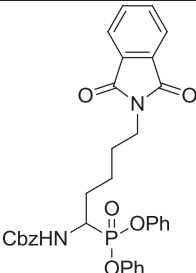
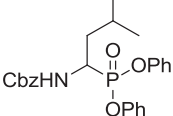
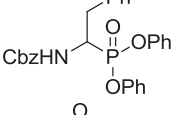
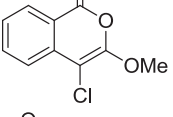
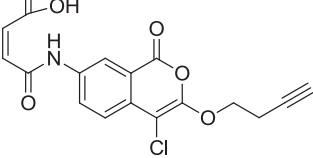
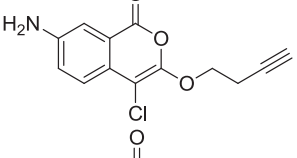
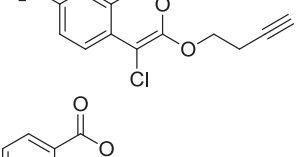
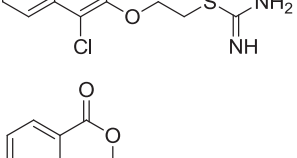
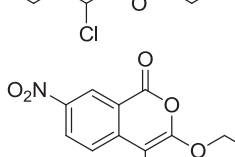
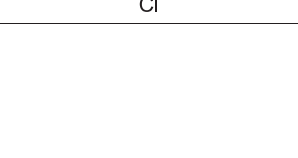
Compound	Structure
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Table S1. Cont.

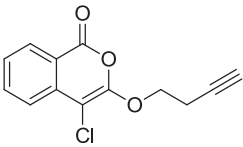
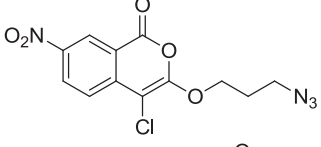
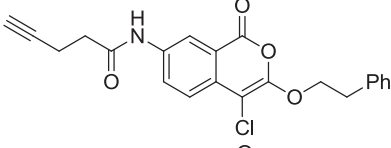
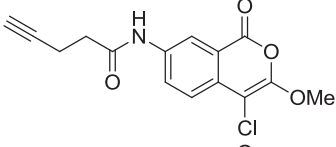
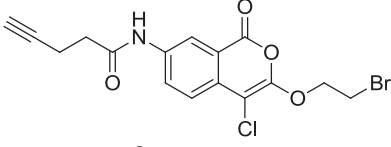
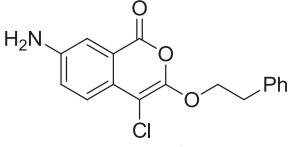
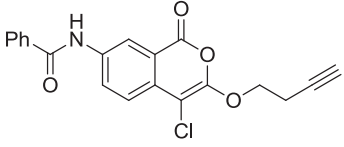
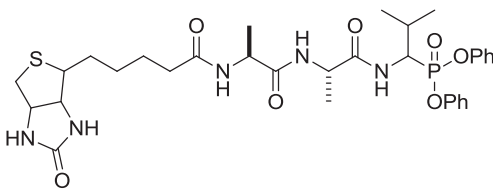
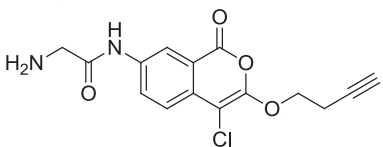
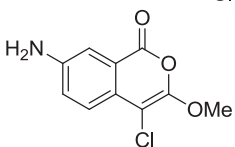
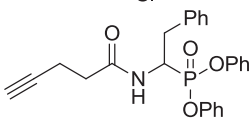
Compound	Structure
11	
12	
13	
14	
15	
16	
17	
18	
19	
20 (JLK-6)	
21	

Table S1. Cont.

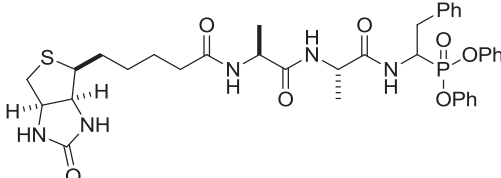
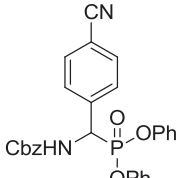
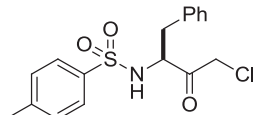
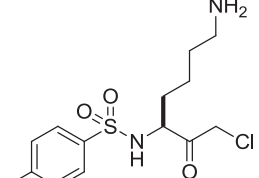
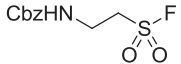
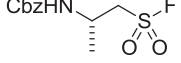
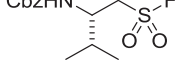
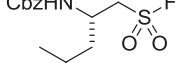
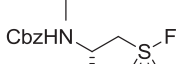
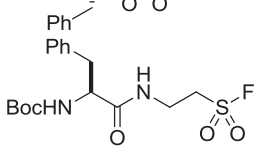
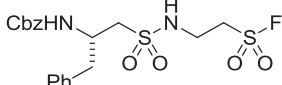
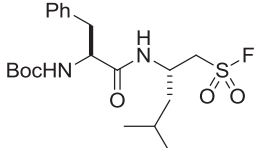
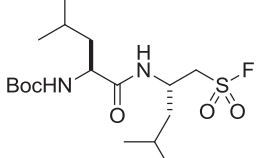
Compound	Structure
22	
23	
24 (TPCK)	
25 (TLCK)	
26	
27	
28	
29	
30	
31	
32	
33	
34	

Table S2. Data collection and refinement statistics for the crystal structure of GlpG with IC 16

Data collection	
Beam line	Diamond/IO2
Space group	R32
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	111.1, 111.1, 126.07
γ (°)	120
Resolution (Å)	44.9–2.2 (2.32–2.2)*
R_{merge}	0.055 (0.48)
I/σ	13.5 (3.1)
Completeness (%)	99.6 (98.1)
Redundancy	4.9 (4.9)
Refinement	
Resolution (Å)	44.9–2.2
No. of reflections	15317
$R_{\text{work}}/R_{\text{free}}^{\dagger}$	0.214/0.247
No. of atoms	
Total	1523
Protein	1436
Ligand [‡]	21
Heteroatoms	48
Water	18
<i>B</i> -factors (Å ²)	
Total	60
Protein	59.3
Ligand	60.2
Heteroatoms	83.6
Water	56.4
rms deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.09

*Values in parentheses are for highest-resolution shell.

[†] R_{free} was calculated using a randomly selected subset of reflections (5%), remaining (95%) reflections was used for calculation of R_{work} .

[‡]Ligand denotes IC16 covalently bound to the enzyme.

Other Supporting Information Files

[Dataset S1 \(PDF\)](#)