# Supporting Information

# Biosynthetic incorporation of fluorinated amino acids into the nonribosomal peptide gramicidin S

Maximilian Müll[1], Farzaneh Pourmasoumi[1], Leon Wehrhan[2], Olena Nosovska[3], Philipp Stephan[1], Hannah Zeihe[1], Ivan Vilotijevic[3], Bettina G. Keller[2], Hajo Kries[1,4*]

[1]Junior Research Group Biosynthetic Design of Natural Products, Leibniz Institute for Natural Product Research and Infection Biology, Hans Knöll Institute (HKI Jena), 07745 Jena, Germany

[2]Freie Universität Berlin, Department of Biology, Chemistry, and Pharmacy, Institute of Chemistry and Biochemistry, Arnimallee 20, 14195 Berlin, Germany

[3]Institute of Organic Chemistry and Macromolecular Chemistry, Friedrich Schiller University Jena, Humboldtstr. 10, 07743 Jena, Germany

[4]University of Bayreuth, Organic Chemistry I, 95440 Bayreuth, Germany

*hajo.kries@leibniz-hki.de

## Contents

# Supporting Methods

## Cloning

In-Fusion cloning was performed following the supplier's instructions. Primers for PCR amplification of inserts (Table S1) were designed with 15-20 bp long overlaps complementary to the insertion position in the vector DNA and ordered as synthetic oligonucleotides (Eurofins GmbH). Competent cells were transformed by heat shock. The insert was checked by Sanger sequencing (Genewiz). Restriction enzymes for linearization of vectors were purchased from New England Biolabs (NEB). PCR reactions were performed with Phusion or Q5 High-Fidelity polymerase (NEB).

**Construction of pSU18-grsTAB/W239S.** For the construction of pSU18-grsTAB/W239S, the corresponding region of the *grs*A gene was amplified in two fragments with primers grsA/W239S-P1-F-b, grsA/W239S-P1-R-a, grsA/W239S-P2-F-b, and grsA/W239S-P2-R-b. The vector pSU18-grsTAB was linearized with PmlI and EcoO01091.

**Construction of pTrc99a-SrfA-B1.** For the construction of pTrc99a-SrfA-B1 the corresponding region of the *srf*A-B gene was amplified in two PCR steps with the primers SrfA_Nested1_f and SrfA_Nested1_r for the first step and B1_isolation_Nested2_CATCf3 and B1_isolation_Nested2_CATr2 for the second step.

**Construction of pSU18-GrsA/239NNK.** For the construction of the plasmid pSU18-GrsA/NNK a linker fragment of the grsA gene was amplified from pSU18-GrsA using the primer pair GrsA_f and GrsA_r. The NNK fragment was amplified from pSU18-GrsA using the primer pair GrsA_W239NNK_f and GrsA_W239NNK_r. Linearized pSU18 was used as a vector.

**Construction of pOPINF-PheA-ND.** For the construction of plasmid pOPINF-PheA_ND the corresponding region of the *grsA* gene was PCR amplified with the primer pair GrsA-ND_f and GrsA-ND_r. For use as vector, plasmid pOPINF was linearized with KpnI and HindIII.

## Protein expression and purification

Proteins were expressed and purified as described previously.[1] Precultures of *E. coli* HM0079 carrying either the plasmid pSU18-mGrsA, pSU18-GrsBMtoL, pSU18-GrsA/W239S, pTrc99a-SrfA-B1 or *E. coli* BL21 containing pOPINE-deoD, which is coding for purine nucleoside phosphorylase (PNP), were prepared by inoculation of 3 mL LB media containing ampicillin (K029.4, Roth) or chloramphenicol (3886.2, Roth) as resistance marker. Precultures were incubated at 37 °C at 180 rpm overnight in a rotary shaker. Main cultures were prepared by inoculation of 400 mL 2xYT media in 2 L flasks, containing the corresponding resistance marker. The cultures were incubated at 37 °C at 250 rpm on a rotary shaker for approximately 4 h until they reached an $OD_{600}$ of approximately 0.6. The cultures were cooled down to 18 °C and induced with 0.25 mM isopropyl-β-D-thiogalactoside (IPTG; BP1755-10, Fisher Scientific). Proteins were expressed overnight at 18 °C at 250 rpm. Cells were harvested by centrifugation and the supernatant was discarded. After resuspending the cell pellet in 30 mL lysis buffer (50 mM TRIS [pH 7.4], 500 mM NaCl, 20 mM imidazole, 2 mM TCEP), 100 µL protease inhibitor mix (APE-K1010, APExBio) were added, and cells were lysed by sonication while cooling on ice. The lysate was cleared by centrifugation at 19,000 g for 30 min at 4°C and the supernatant was loaded onto a column packed with 2 mL of Ni-IDA suspension (1308.2, Roth) and equilibrated with lysis buffer. After washing the column twice with 20 mL of the lysis buffer, the target protein was eluted with 6 x 0.75 mL elution buffer (50 mM TRIS [pH 7.4], 500 mM NaCl, 300 mM imidazole, 2 mM TCEP). After pooling the protein-containing fractions, they were buffer exchanged with storage buffer (100 mM TRIS [pH 7.4], 500 mM NaCL, 2 mM TCEP), using Vivaspin 6 (Sartorius) filters with a cutoff of 10 kDa for the PNP, 30 kDa for GrsA, GrsA-W239S and SrfA-B1 or 100 kDa for GrsB. Proteins were aliquoted in 10% glycerol and flash frozen in liquid nitrogen for storage at -80 °C. Protein concentrations were determined from

the absorbance at 280 nm measured in Take3 plates on an epoch2 microplate reader (Biotek) using calculated extinction coefficients (www.benchling.com). PNP was stored at a concentration of 500 µM in aliquots of 1.3 mg.

## Adenylation kinetics

Michaelis-Menten parameters of adenylation reactions were determined from kinetic data recorded with the MesG/hydroxylamine assay which was performed as described previously with minor modifications.[2] Reactions contained 50 mM TRIS (pH 7.6), 5 mM $MgCl_2$, 100 µM 7-methylthioguanosine (MesG; PR3790-B100, Biosearch Technologies), 150 mM hydroxylamine (adjusted to pH 7.5-8 with NaOH), 5 mM ATP (A2383, Sigma), 1 mM TCEP, 0.4 U/mL inorganic pyrophosphatase (I1643, Sigma), 50 µM of PNP purified from *E. coli* in-house, the NRPS protein of interest (0.1 µM), and a suitable amino acid substrate. The amino acids *rac*-3,5-$F_2$-Phe, *rac*-2-F-Phe, and *rac*-3-F-Phe were purchased from abcr GmbH, while *rac*-2,4-$F_2$-Phe and *rac*-4-F-Phe were purchased from Fluorochem Ltd. In flat-bottom 384-well plates (781620, Brand), reactions were started in a total volume of 100 µL by addition of substrate. Then, the absorbance was followed at 355 nm on a Synergy H1 (BioTek) microplate reader at 30 °C. Slopes for the background activity were recorded in wells containing buffer but not substrate and were subtracted. Each substrate concentration was measured as biological triplicate. Initial velocities were divided by the slope of a pyrophosphate calibration curve to obtain the pyrophosphate release rate. Michaelis-Menten parameters were extracted from the initial velocity $v_0/[E_0]$ data through nonlinear regression in the R software package version 3.4.2.[3]

## Isothermal titration calorimetry

For isothermal titration calorimetry (ITC), the preparation of protein was slightly modified. Protein expression and purification was done as described above from precultures of *E. coli* BL21 carrying plasmid pOPINF-PheA_ND for expression of the core-domain of GrsA-A. Then, a Vivaspin filter with a 10 kDa cutoff was used for buffer exchange into low salt buffer (100 mM TRIS [pH 7.4], 20 mM NaCl). To achieve the required purity for ITC, the protein sample was further purified by anion exchange chromatography on an NGC chromatography system (Bio-Rad Laboratories) using a MonoQ 5/50 GL column (GE Healthcare). Protein was eluted with a gradient of 20 to 600 mM NaCl in 20 mM TRIS (pH 8). After anion exchange, buffer was exchanged into storage buffer (100 mM HEPES [pH 8], 10% glycerol) using a Vivaspin filter. As before, aliquots were flash frozen in liquid nitrogen and stored at -80 °C.

For ITC measurements, 800 µL of a 60 µM protein solution were prepared in HEPES buffer (100 mM HEPES [pH 8], 10% glycerol). Substrate stocks (*rac*-Phe, *rac*-4-F-Phe and *rac*-2,4-$F_2$-Phe) were prepared at 6.25 mM in HEPES buffer. Protein samples were loaded into the cell of a MicroCal PEAQ-ITC (Malvern Panalytical) and titrated against the substrate solution at 25 °C at 750 rpm stirrer speed with 15 injections of 2 µL substrate. The initial delay was 60 s, the reference power was set to 10.0 µcal/s, the feedback to high, the injection spacing to 150 s and the injection duration to 4 s. Data analysis was performed with the MicroCal PEAQ-ITC Analysis software from Malvern Panalytical.

## Hydroxamate assay (HAMA)

Hydroxamate assays for adenylation specificity were performed as previously reported.[1] In brief, HAMA was conducted at room temperature in 100 µL volume containing 50 mM TRIS (pH 7.6), 5 mM $MgCl_2$, 150 mM hydroxylamine (pH 7.5-8, adjusted with NaOH), 5 mM ATP (A2383, Sigma), 1 mM TCEP, and 1 mM proteinogenic amino acids with D-Val, D-Phe, deuterated L-Leu-d7, deuterated L-Phe-d5, deuterated L-Val-d8, *rac*-4-F-Phe, and *rac*-2,4-$F_2$-Phe. Reactions were started by adding the NRPS protein to a final concentration of 0.1 µM. Samples were incubated for 30 min and 10 µL of the reaction was quenched in 95% ACN containing 0.1% formic acid (A117-50, Fisher Scientific), cooled down and

centrifuged to remove the precipitated hydroxylamine. As control, heat denatured enzyme was used. Samples were measured as biological duplicates.

Samples were analyzed on a UPLC-MS/MS (Xevo TQ-S micro, Waters) as described previously[1] with the same mass transitions and MS conditions. The mass transitions for the detection of 4-F-PheHA and 2,4-$F_2$-PheHA are 199.18 → 138.08 and 217.17 → 156.13, respectively.

## Screening a GrsA-W239X library using HAMA

The screening method has been adapted from a published procedure.[4]

**Protein expression.** *E. coli* HM0079 hosting the plasmid library pSU18-GrsA-W239/NNK was used to overexpress GrsA-W239 variants in a 96-well plate format. Precultures were prepared by inoculating the transformants picked from an agar plate into a round bottom 96-well plate (310 µl, Sarstedt) filled with 150 µl of 2xYT medium supplemented with 100 µg/ml of chloramphenicol (3886.2, Roth). The 96-well plate contained four wells with *E. coli* HM0079::pSU18-GrsA and four wells with *E. coli* HM0079::pSU18-GrsA/ W239S as controls. Plates were covered with a breathable polyurethane film (Breathe-Easy, Sigma-Aldrich) and incubated for 18 h at 30 °C and 400 rpm in an orbital shaker. The following liquid handling steps were typically performed using a Gilson Platemaster 220 µL as 96-well pipette. For protein expression, 20 µl of the preculture was inoculated into a 96-deep-well plate (2 mL, Sarstedt) containing 1 ml 2xYT medium supplemented with 100 µg/ml chloramphenicol and incubated for 5 h at 30 °C and 400 rpm. A 20 µL aliquot was taken from the culture and stored with 25% glycerol at -70 °C for sequencing. To start the induction process, the cultures were cooled down to 18 °C for 30 min, followed by addition of 0.25 mM IPTG (BP1755-10, Fisher Scientific). Cells were incubated overnight for 18 h at 18 °C at 400 rpm. Cells were harvested by centrifugation at 3,220 g and 10 °C for 5 min and the supernatant was discarded. Lysis buffer (50 mM TRIS [pH 8.0], 100 mM NaCl, 10 mM imidazole, 1.5 mg/mL lysozyme) was prepared freshly by adding 1 µL/mL of protease inhibitor mix (APE-K1010, APExBio). Per well, 400 µL lysis buffer was used for resuspension of the cells. Cells were incubated at room temperature for 30 min, followed by freezing at -20 °C. Lysis was achieved by thawing the cells for 2 h at room temperature.

**Protein purification.** After thawing, 100 µL of DNA removal mix (50 mM TRIS [pH 8.0], 100 mM NaCl, 10 mM imidazole, 10 mM $MgCl_2$, 10 mM TCEP, 15 U/mL Turbonuclease [Jena Bioscience]) was added to reduce the viscosity of the lysate. Cell debris was removed by centrifugation at 3,220 g and 6 °C for 30 min. In a separate, 96-well plate (1.8 mL, Sarstedt) compatible with the magnetic separation rack (S1511S, New England Biolabs), 20 µl of a 25% Ni-IDA MagBeads (PureCube) suspension was added. For equilibration of the beads, 700 µL of lysis buffer was used and supernatant was discarded. Next, 400 µL of lysate was added to the equilibrated beads. The plate was covered with a silicon lid and kept at 6 °C for 20 min. Every 5 min the plate was shaken vigorously to resuspend the beads. Beads were pelleted on a magnetic rack and supernatant was discarded. Beads were washed twice with 700 µl of wash buffer (50 mM TRIS [pH 8.0], 100 mM NaCl) using the magnetic rack for separation of beads and wash fraction.

**HAMA in 96-well plate format.** After the second washing step, 100 µl of freshly prepared HAMA master mix (50 mM TRIS [pH 8.0], 5 mM ATP, 5 mM $MgCl_2$, 100 mM hydroxylamine adjusted to pH 7.5-8 with NaOH, 1 mM TCEP, 1 mM amino acid mix) was added directly to the beads containing the adsorbed protein and incubated at room temperature for 1.5 h. The amino acid mix contained the proteinogenic amino acids with D-Val, D-Phe, deuterated L-Leu-d7, deuterated L-Phe-d5, deuterated L-Val-d8, *rac*-4-F-Phe and *rac*-2,4-$F_2$-Phe. After incubation, 6 µl of the reaction mixture was diluted with 54 µl of analysis solution (acetonitrile with 0.1% formic acid) in a 384-well plate (100 µL, Brandt). After the dilution step, the 384-well plate was immediately placed on ice and covered with aluminum foil to

minimize evaporation of the solvent. The plate was analysed immediately by UPLC-MS/MS according to the general HAMA procedure.

## *In vitro* gramicidin S formation

For *in vitro* biosynthesis of GS, reactions were performed for 2 h at 37 °C in 100 µL PCR tubes in reaction buffer (50 mM HEPES [pH 8], 1 mM $MgCl_2$, and 100 mM NaCl). Amino acids L-Leu, L-Pro, L-Orn, and L-Val were added to a final concentration of 20 mM each. The amino acids *rac*-4-F-Phe or *rac*-2,4-$F_2$-Phe were added to a concentration of 5 mM. As positive control, 5 mM *rac*-Phe was used. GrsA or GrsA/W239S were added to a concentration of 1 µM. GrsB was added to a concentration of 3 µM, pretending that the protein was homogenous (for SDS-PAGE, see Fig. S1). Denatured GrsB was used as negative control. Glycerol was added to a final concentration of 2%. To start the reaction, ATP was added to a concentration of 5 mM. Reactions were quenched by the addition of 100 µL MeOH. Samples were cooled down for 20 min at -20 °C and centrifuged at 12,000 g for 5 min. From the supernatant, 160 µL were taken and mixed with 240 µL of 50% EtOH in water containing 0.1% formic acid. Samples were run on a UPLC-MS/MS (Xevo TQ-S micro, Waters) with an H-class UPLC. For chromatography, a CSH C18 column (186005296, Waters) was used with a linear gradient of 40% ACN and 60% water containing 0.1% formic acid to 98% ACN and 2% water containing 0.1% formic acid over 1 min, followed by 1.2 min reequilibration. Analytes were detected in MRM mode based on the mass transitions for GS (571.696→70.099) and the fluorinated analogs (589.85→70.0991, 607.6757→70.0991). These MRMs use the pyrrolidinium fragment (70.0991) of Pro for quantification. Authentic GS purified from the natural producer *Aneurinibacillus migulanus*[5] was used as a standard for quantification assuming an identical response for GS and the fluorinated analogs.

## *In vivo* production of GS, 4-F-Phe-GS and 2,4-$F_2$-Phe-GS

Samples containing complete TB (3 mL) with chloramphenicol (25 µg/mL) were inoculated with a starter culture (1/500 v/v) of *E. coli* HM0079 cells transformed with plasmids pSU18-grsTAB/W239S and incubated at 30 °C, 230 RPM until $OD_{600}$ = 2 was reached. Samples were divided into three groups with two biological replicates each. To group one, *rac*-2,4-$F_2$-Phe (2.5 mM) was added. To group two, *rac*-4-F-Phe (4 mM), and to group three (control group), only deionized water was added. All cultures were sampled after 96 h (1 mL), and clarified by centrifugation (11,000 g, 4 min, RT). The supernatants were removed and the cell pellets, which contain most of the GS, were resuspended in 70% ethanol, followed by sonication for 10 min and further incubation at 60°C for 30 min. The cell debris was removed by centrifugation (19,000 g, 4 min, RT). The supernatants were further diluted 50-fold in 50% ethanol containing 0.1% formic acid and the concentrations of GS, 4-F-Phe-GS and 2,4-$F_2$-Phe-GS were quantified using UPLC-MS/MS in comparison with a GS standard purified from *A. migulanus.*[5] Concentrations were calculated assuming that GS, 4-F-Phe-GS, and 2,4-$F_2$-Phe-GS were homogenously distributed in the culture volume. A control experiment was conducted in parallel with *E. coli* HM0079 cells transformed with plasmid pSU18-grsTAB instead of the mutated variant.

UPLC-MS/MS analysis was performed on a Waters ACQUITY H-class UPLC system coupled to a Xevo TQ-S micro (Waters) tandem quadrupole instrument under optimized conditions. The injection volume was 2 µL and the flow rate was 0.5 mL $min^{-1}$. Acetonitrile (B) and water with 0.1% formic acid (A) were used as strong and weak eluent, respectively. Acetonitrile was used as the needle wash between the samples. Data acquisition and quantification were done using the MassLynx software (version 4.1). MS/MS analyses were performed using an ESI source in positive ion mode. Nitrogen was used as desolvation gas and argon as collision gas. The following source parameters were used: capillary voltage 0.5 kV, desolvation temperature 600 °C, desolvation gas flow 1000 L $h^{-1}$.

Column: ACQUITY UPLC CSH C18, 1.7 µm particle size, 2.1 × 50 mm

Elution profile: linear gradient of 40 to 98% B over 1 min followed by 1.2 min re-equilibration.

GS MRM transition: 571.696 > 70.099

4-F-Phe-GS MRM transition: 589.85 > 70.099

2,4-$F_2$-Phe-GS MRM transition: 607.676 > 70.099

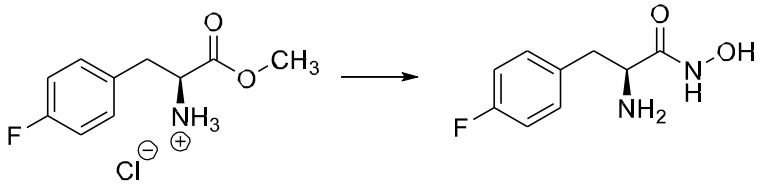
GS calibration curve range: 0.01 µM to 10 µM

## Synthesis of amino acid hydroxamates

All chemicals for synthesis were purchased from Merck, Alfa Aesar, Acros Organics, Fluorochem or TCI and used without further purification. The solvents were dried according to standard conditions if needed. The TLC-glass-plates DURASIL consisted of a 0.25 mm layer of silica 60 with fluorescence indicator UV254. TLCs were checked under UV-light (254 nm or 365 nm) and stained with an aq. $KMnO_4$ solution. $^1H$, $^{13}C$ and $^{19}F$ NMR spectra were measured on BRUKER Fourier 500 or a BRUKER Avance 400 spectrometers. The chemical shift of each signal was reported in ppm. For $^1H$ and $^{13}C$ measurements, the chemical shift refers to TMS, showing a signal at 0 ppm. As an internal standard, the residual $^1H$ or $^{13}C$ nuclei of the corresponding deuterated solvents were used (DMSO-$d_6$, 2.50 ppm [$^1H$-NMR], 39.51 ppm [$^{13}C$-NMR]). The chemical shift of the fluorine NMR was determined indirectly. For carbon spectra, a broadband decoupling was performed. High-resolution mass spectra (HRMS) were measured using a Thermo Q-Exactive plus device with an ESI source coupled to a binary UHPLC system. IR spectra were measured using the Shimadzu IR-Affinity-1 (FTIR) device.

**Synthesis of 4-F-Phe hydroxamate**

Hydroxamate **2** was prepared from commercially available methyl ester **1**.




**1** **2**

(*S*)-2-amino-3-(4-fluorophenyl)-N-hydroxypropanamide **2**

A suspension of 4-fluoro-L-phenylalanine methyl ester hydrochloride **1** (91.60 mg, 0.39 mmol) in dichloromethane (10 mL) was washed with saturated aqueous $K_2CO_3$ solution (2 x 10 mL). The combined aqueous layers were washed with dichloromethane (3 x 5 mL). The combined organic layers were dried over anhydrous $K_2CO_3$ and concentrated under reduced pressure to provide 4-fluoro-L-phenylalanine methyl ester free base as a colourless oil (72 mg, 0.37 mmol, 93%). It was subjected to the next transformation without further purification.

A 5.0 M solution of KOH in dry methanol (0.2 mL, 1.09 mmol, 3 equiv.) was added to a 1.0 M solution of hydroxylamine hydrochloride in dry methanol (1 mL, 1.09 mmol, 3 equiv.) at 0 °C. The resulting mixture was kept at 0 °C for 15 min and was filtered through a Teflon 2.5 µm filter to remove the precipitate. The filtrate was then added to a solution of 4-fluoro-L-phenylalanine methyl ester free base (72 mg, 0.37 mmol, 1 equiv.) in dry methanol (1 mL) and kept at -20 °C without stirring to facilitate crystallization. After 5 days, precipitate was filtered, washed with dry methanol, and dried under vacuum to afford the desired hydroxamate **2** as a colourless solid (10 mg, 0.05 mmol, 10%). The product was stored at -20 °C.

**$^1$H NMR** (500 MHz, DMSO-$d_6$) δ ppm 7.16 - 7.26 (m, 2 H), 7.08 (br t, *J* = 8.79 Hz, 2 H), 3.14 - 3.27 (m, 1 H), 2.80 (br dd, *J* = 13.27, 5.97 Hz, 1 H), 2.61 (br dd, *J* = 13.35, 7.71 Hz, 2 H).

**$^{13}$C NMR** (126 MHz, DMSO-$d_6$) δ ppm 170.98, 160.31 (br d, *J* = 90.76 Hz), 134.85 (d, *J* = 2.99 Hz), 131.00 (d, *J* = 7.98 Hz), 114.67 (d, *J* = 20.94 Hz), 54.39.

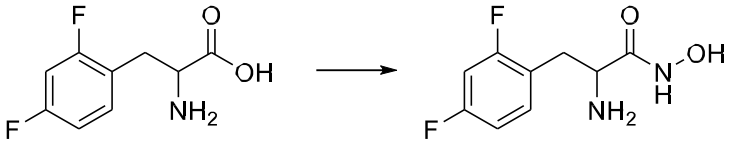
**$^{19}$F NMR** (471 MHz, DMSO-$d_6$) δ ppm -117.35.

**HRMS** [ESI]: m/z calculated for $C_9H_{11}FN_2O_2$ $[M+H]^+$ 199.0877, found 199.0875.

**IR** (ATR): $\nu$ = 2825 (w), 1608 (s), 1508 (s), 1489 (m), 1474 (m), 1383 (m), 1289 (w), 1161 (w), 891 (m).

**Synthesis of 2,4-$F_2$-Phe hydroxamate**

Hydroxamate **4** was prepared from commercially available amino acid **3.**




**3** **4**

(*R,S*)-2-amino-3-(2,4-difluorophenyl)-N-hydroxypropanamide **4**

To a suspension of amino acid **3** (201 mg, 1 mmol, 1 equiv.) in dry methanol (4 mL), thionyl chloride (177 µL, 2.5 mmol, 2.5 equiv.) was added. The resulting mixture was stirred at 80 °C. After the reaction was completed, the reaction mixture was concentrated under reduced pressure. The residue was dissolved in saturated aqueous $K_2CO_3$ solution and extracted with dichloromethane (3 x 15 mL). The combined organic layers were dried over anhydrous $K_2CO_3$ and concentrated under reduced pressure to give methyl 2-amino-3-(2,4-difluorophenyl)propanoate as a colourless oil (186 mg, 0.87 mmol, 87%). The product was subjected to the next transformation without further purification.

A 5.0 M solution of KOH in dry methanol (0.5 mL, 2.5 mmol, 3 equiv.) was added to a 1.0 M solution of hydroxylamine hydrochloride in dry methanol (2.5 mL, 2.5 mmol, 3 equiv.) at 0 °C. The resulting mixture was kept at 0 °C for 15 min and was filtered through a Teflon 2.5 µm filter to remove the precipitate. The filtrate was then added to a solution of methyl 2-amino-3-(2,4-difluorophenyl)propanoate (176 mg, 0.81 mmol, 1 equiv.) in dry methanol (0.8 mL) and kept at -20 °C without stirring to facilitate crystallization. After 2 days, precipitate was filtered, washed with dry methanol, and dried under vacuum to give the desired hydroxamate **4** as a colourless solid (55.3 mg, 0.25 mmol, 31%). The product was stored at -20 °C.

**$^1$H NMR** (400 MHz, DMSO-$d_6$) δ ppm 8.04 - 10.65 (m, 1 H), 7.29 (br d, *J* = 7.31 Hz, 1 H), 7.14 (br t, *J* = 8.77 Hz, 1 H), 6.99 (br s, 1 H), 3.10 - 3.27 (m, 1 H), 2.80 (br dd, *J* = 12.42, 5.70 Hz, 1 H), 2.57 - 2.71 (m, 1 H).

**$^{13}$C NMR** (101 MHz, DMSO-$d_6$) δ ppm 170.87, 160.93 (dd, *J* = 244.49, 13.00 Hz), 160.63 (dd, *J* = 247.52, 13.44 Hz), 132.72 (dd, *J* = 9.10, 6.50 Hz), 121.56, 110.99 (dd, *J* = 20.37, 3.03 Hz), 103.39 (t, *J* = 26.01 Hz), 53.19, 33.81.

**$^{19}$F NMR** (376 MHz, DMSO-$d_6$) δ ppm -113.30, -113.52.

**HRMS** [ESI]: m/z calculated for $C_9H_{10}F_2N_2O_2$ $[M+H]^+$ 217.0783, found 217.0778.

**IR** (ATR): $\nu$ = 3186 (w), 2897 (w), 1624 (s), 1605 (m), 1541 (m), 1508 (s), 1379 (s), 1292 (m), 1265 (m), 1128 (s), 978 (m), 891 (s), 743 (w).

# Computational Methods

## Protein structure preparation

The X-ray protein structure of the GrsA A-domain in complex with AMP and Phe (PDB ID 1amu)[6] was loaded into Schrödinger's Maestro and chain B of the dimer was deleted. The remaining structure was prepared with the Protein Preparation Wizard[7] in default settings, which includes adding missing hydrogens and sidechains, deleting waters far from the natural ligand and running a short MM energy minimization. The W239S mutant was modeled by mutating the residue and briefly minimizing the obtained structure inside Maestro. The fluorinated ligands 4-F-Phe, 2,4-$F_2$-Phe, 2-F-Phe, 3-F-Phe, and 3,5-$F_2$-Phe were constructed by replacing the respective hydrogens of the natural ligand Phe with the Maestro Build tool.

## Classical force field distance scan

The potential energy of a system of capped fluorinated and unfluorinated Phe, which is placed perpendicularly above the aromatic system of capped Trp, was calculated at various distances in vacuum. The capped Trp molecule was constructed by extracting Trp239 from the prepared protein structure of WT GrsA and then capping the N-terminal end with an acetyl cap (ACE) and the C-terminal end with an N-methyl cap (NME). The capped Phe and fluorinated Phe analogues were constructed in a similar way, by extracting the ligands from the prepared PDB structure and then applying ACE and NME caps. The Phe molecule was placed directly above the center between CD2 and CE2 of the Trp molecule. The distance between this center and CZ of the Phe molecule was adjusted to cover distances between 0.3 nm and 1.1 nm in steps of 0.05 nm. For each of the systems one structure was energy minimized with the Gromacs[8–10] 2021 software using a steepest descent algorithm and positional restraints on every heavy atom of the structure. The force field used was Amber14SB[11] with GAFF2[12] parameters for the fluorinated amino acids. All parameters were generated using Acpype.[13–15] The total potential energy of the system after energy minimization was then calculated using Gromacs energy. For each system the total potential energy for the distance of 1.1 nm was set to zero and the difference to this energy was calculated for the other distances.

## Molecular dynamics simulation

Molecular dynamics (MD) simulations were run using the Gromacs[8–10] 2021 software and the Amber14SB[11] force field, combined with GAFF2[12] parameters, generated with Acpype,[13–15] for the fluorinated amino acid. The simulations were initiated with the prepared structures of GrsA (PDB ID 1amu) in complex with Phe and —F-Phe as starting structures. The three systems that were simulated were WT GrsA in complex with Phe, GrsA-W239S in complex with Phe and GrsA-W239S in complex with 4-F-Phe. The structures were solvated in a cubic box with periodic boundary conditions and with 1.1 nm between the solute and box edges in TIP3P[16] water. The systems were energy minimized using a steepest descent algorithm, followed by an equilibration in the NVT ensemble for 100 ps at 300 K with restraints on all solute heavy atoms and a subsequent unrestrained equilibration in the NPT ensemble at 300 K and 1.0 bar for 1 ns. The production simulations were run in the NPT ensemble for 10 ns. The temperature was kept constant with a velocity rescaling scheme with a stochastic term[17] and the employed barostat was the Parinello-Rahman barostat.[18]

The special region of the cavity next to Trp/Ser239 was modeled by a sphere of 0.4 nm radius. The center of the sphere was placed at the center of the straight line between the C-alpha atoms of Trp/Ser239 and Thr334 in the prepared starting structure. The position of the sphere was not moved throughout the simulation. The number of water molecules inside the sphere was counted throughout the simulation and divided by the number of simulation snapshots to obtain the percentages (Table S3).

## Molecular docking of Phe analogs to GrsA-W239S

The ligands Phe, 4-F-Phe, 2,4-$F_2$-Phe and O-propargyl-Tyr were docked against the W239S mutant of GrsA using the GLIDE[19–21] docking software. A receptor grid was generated based on the prepared structure of GrsA-W239S. The docking was run in extra precision mode with flexible ligand sampling and penalized non-planar amide conformations. Core constraints were applied by restricting the ligands to the reference position of the natural Phe ligand in the PDB structure 1amu with a tolerance of 0.1 Å. The core atoms were defined as the maximum common substructure. The preferred docking pose is shown in Fig. S4 and the docking scores in Table S4.

# Supporting Figures

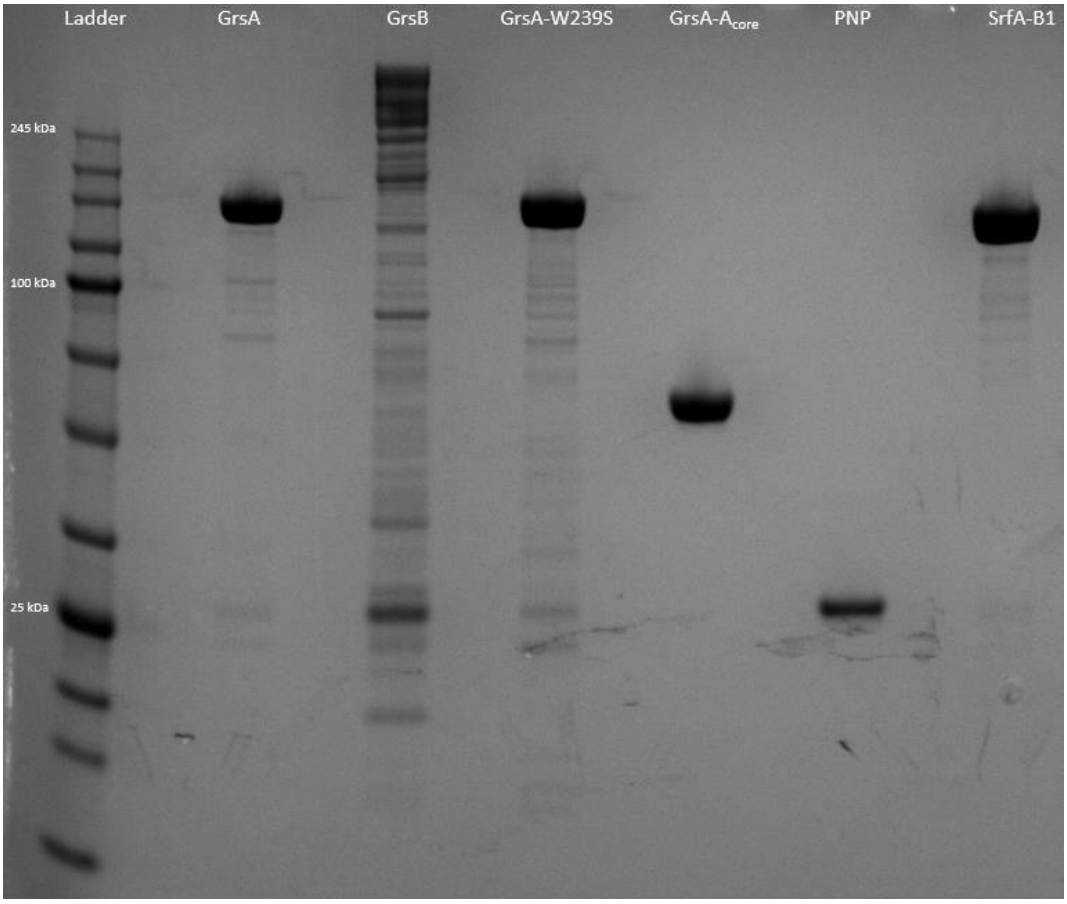



**Figure S1**. Analysis of protein purity by SDS-PAGE.

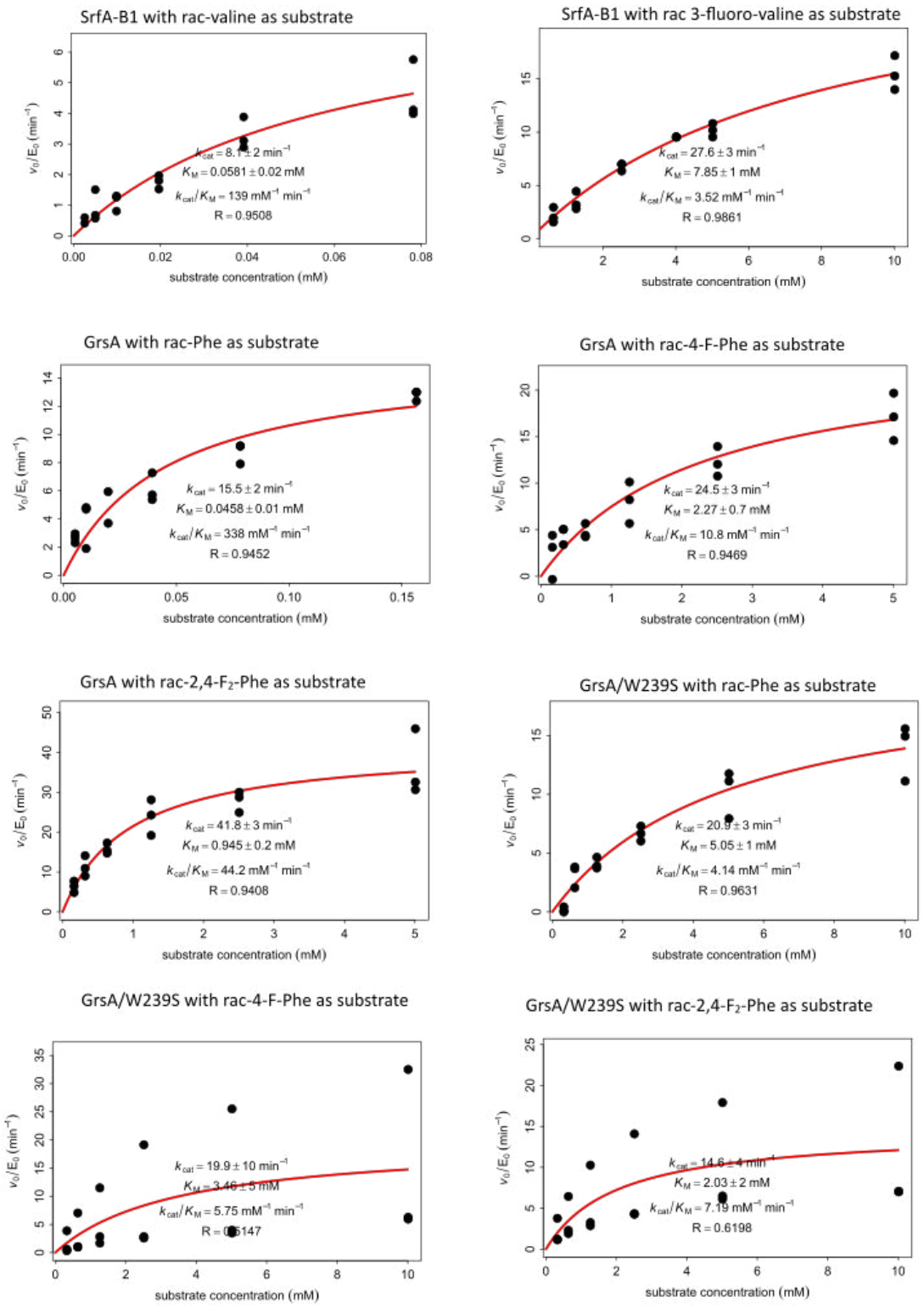



**Figure S2**. Adenylation kinetics (continued on next page).

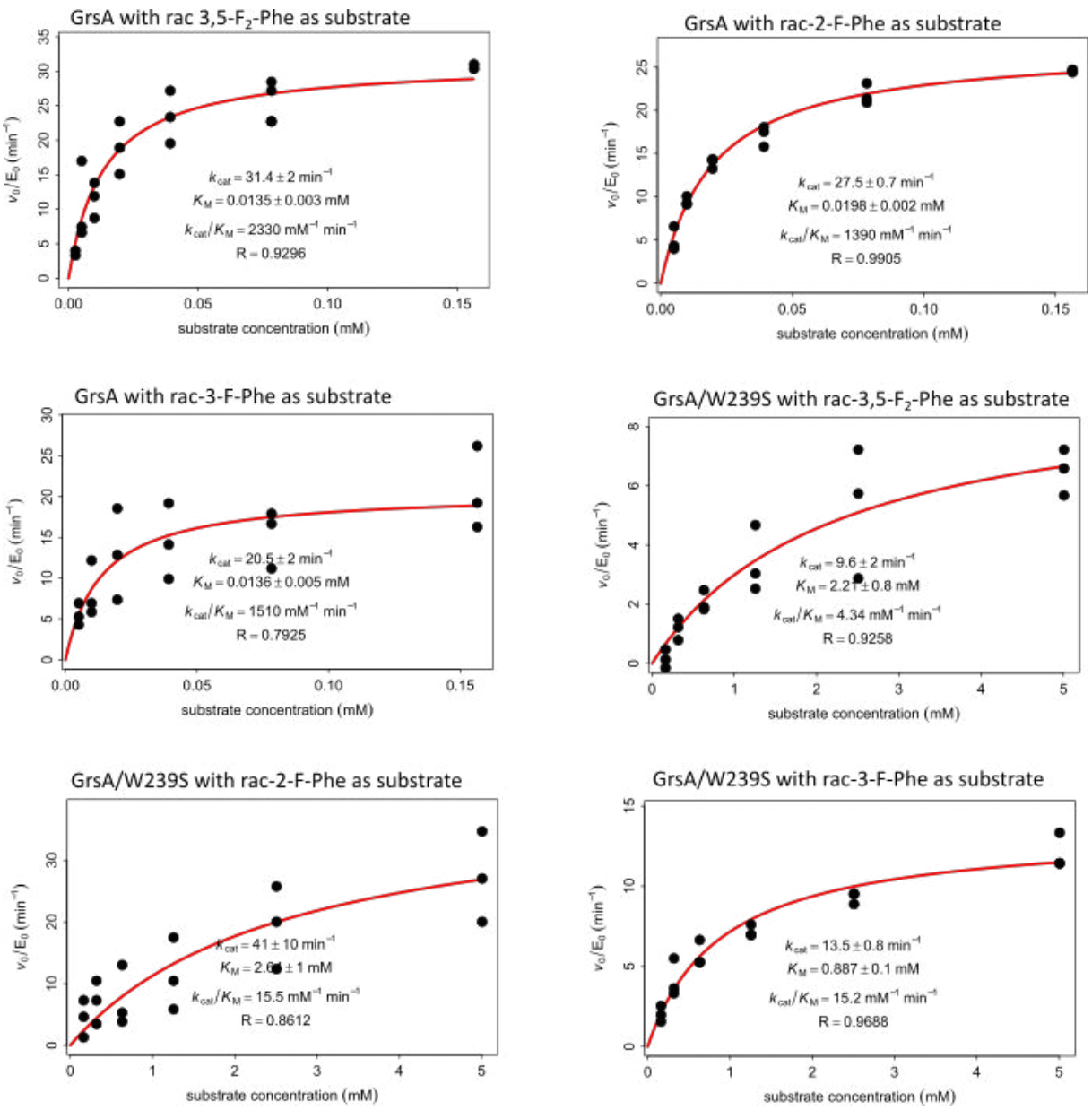



**Figure S2**. Continued.

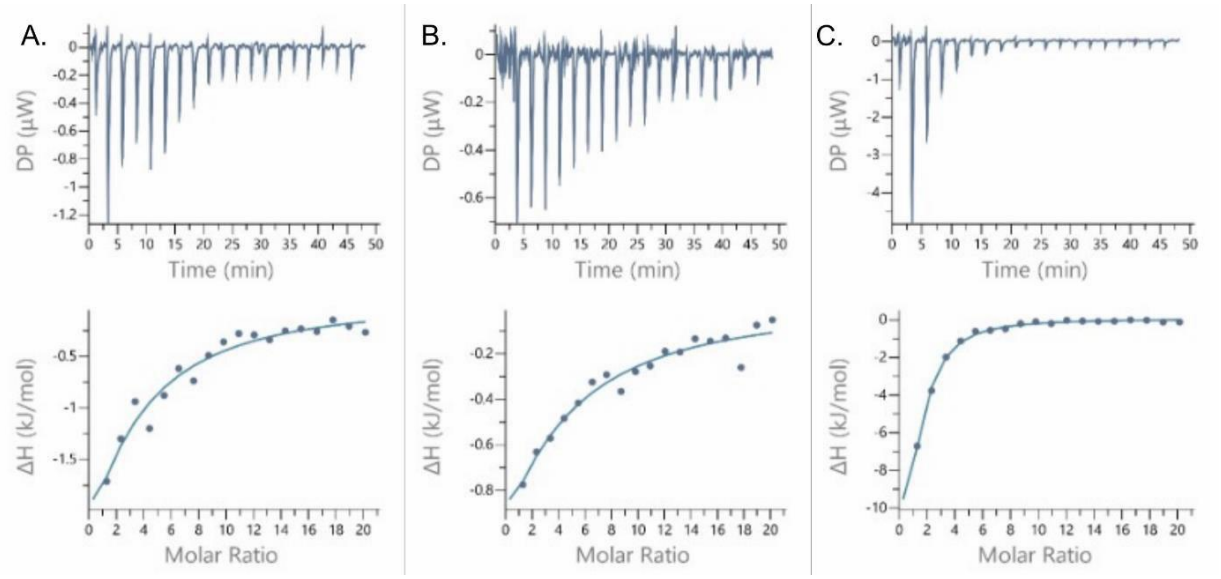



**Figure S3**. Representative thermograms of amino acid binding to GrsA-$A_{core}$ (60 µM) measured by ITC. A) Titration with 6.25 mM *rac*-2,4-$F_2$-Phe. Four replicates were recorded. B) Titration with 6.25 mM *rac*-4-F-Phe. To compensate for the weak signal, 13 replicates were recorded. C) Titration with 6.25 mM *rac*-Phe. Three replicates were recorded.

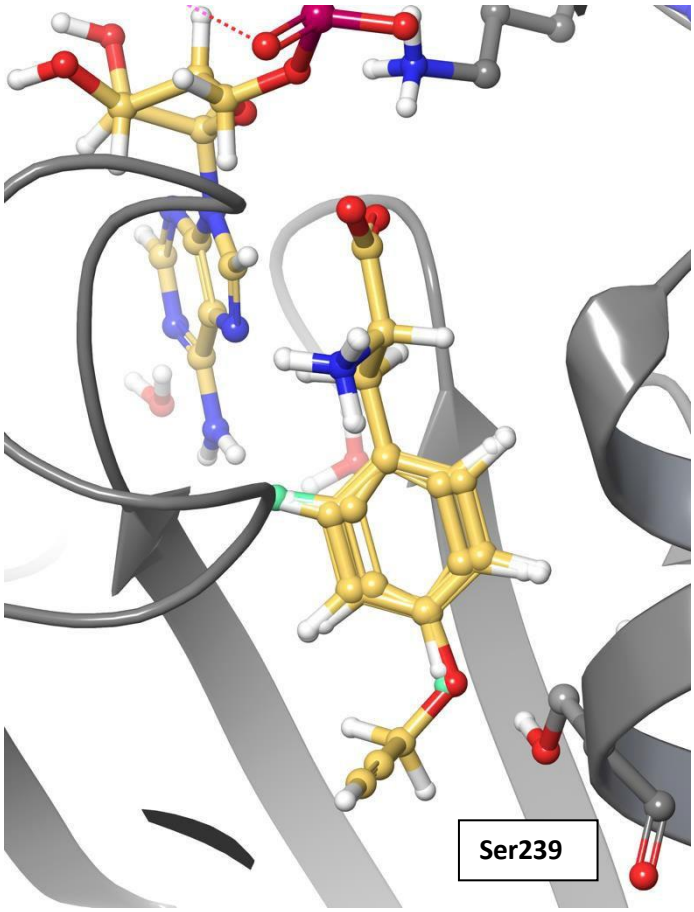



**Figure S4.** Preferred docking poses of the ligands 4-F-Phe, 2,4-$F_2$-Phe and *O*-propargyl-Tyr in the modelled structure of GrsA-W239S are superposed. Core constraints are applied on the docking poses based on the natural substrate in pdb structure 1amu, so the core region of the ligands has little degrees of freedom.

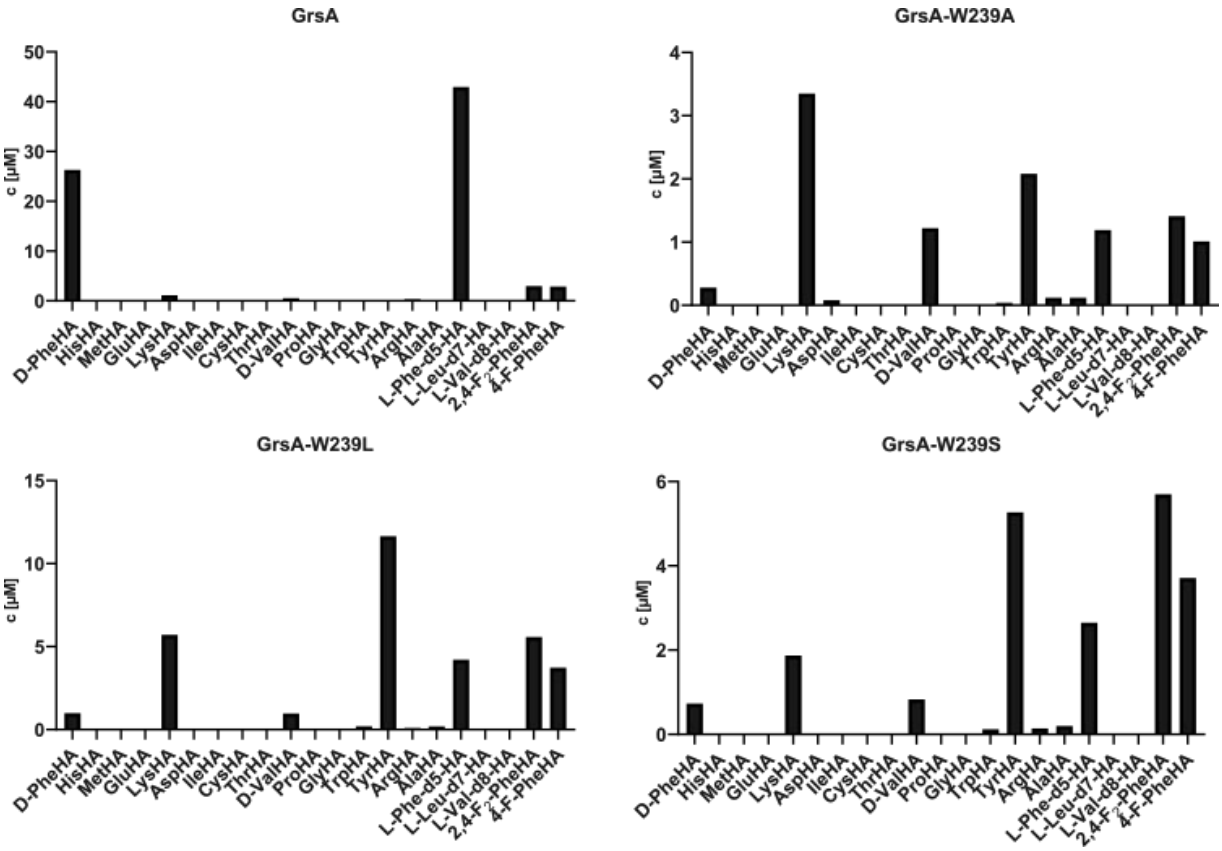



**Figure S5.** Selected mutants in position GrsA-W239 showing improved selectivity for 4-F-Phe in the HAMA 96 well screening experiment. The mutant W239S shows the highest selectivity towards the fluorinated substrates.

# Supporting Tables

**Table S1**. List of primers.

| Primer names | Sequence (5' - 3') |
|---|---|
| GrsA-ND_f | AAG TTC TGT TTC AGG GCC CGA TGT TAA ACA GTT CTA AAA G |
| GrsA-ND_r | ATG GTC TAG AAA GCT TTA TAT TCT TCC GAG ATA TTC AAT ATT TCC |
| grsA/W239S-P1-F-b | AAG CGG AGT ATC CAC GTG ATA AGA CGA TCC ATC AGT TAT TTG AAG AGC AG |
| grsA/W239S-P1-R-a | TAC AGA TGC ATC AAA AGA GAT GCT GGC AAA |
| grsA/W239S-P2-F-b | TCT CTT TTG ATG CAT CTG TAA GCG AGA TGT TTA TGG C |
| grsA/W239S-P2-R-b | TCT TTA CTA GAG GGC CTA CTT CCA AGT TTA TAC TAT TTT GTA ATC GAG CA |
| SrfA_Nested1_f | CTA TTT AGG TCA GTT TGA CGA AAT G |
| SrfA_Nested1_r | CTT GGG CAC GAA GAT GAT G |
| B1_isolation_Nested2_CATCf3 | CAC AGG AAA CAG ACC ATG AGC AAA AAA TCG ATT CAA |
| B1_isolation_Nested2_CATr2 | ATG GTG ATG AGA TCT CAA ATA CAG TGC CAG TTC TTG AAT A |
| GrsA_f | AAG AGG AGA AAT TAA CCA TGT TAA |
| GrsA_r | TAC AGA TGC ATC AAA AGA GAT G |
| GrsA_W239NNK_f | CAT CTC TTT TGA TGC ATC TGT ANN KGA GAT GTT TAT GGC T |
| GrsA_W239NNK_r | GAT GGT GAT GAG ATC TGG A |

**Table S2**. Thermodynamic parameters of amino acid binding to GrsA-$A_{core}$ .

| Enzyme | Substrate | $K_D$ [µM] | $\Delta G$ [kJ/mol] | $\Delta H$ [kJ/mol] | $-T\Delta S$ [kJ/mol] |
|---|---|---|---|---|---|
| **GrsA-$A_{core}$** | Phe | 60±10 | -24.1±0.5 | -20±2 | -4±3 |
| | 4-F-Phe | 600±300 | -19±2 | -11±6 | -7±7 |
| | 2,4-$F_2$-Phe | 420±40 | -19.3±0.2 | -15±2 | -5±2 |

**Table S3.** Percentage of simulation snapshots that show the according number of water molecules inside the cavity throughout a 10 ns MD simulation of GrsA (wild type and W239S mutant) in complex with the substrates Phe and 4-F-Phe.

| GrsA variant | Substrate | Number of water molecules | | | | |
|---|---|---|---|---|---|---|
| | | **0** | **1** | **2** | **3** | **4** |
| wild type | Phe | 100% | 0% | 0% | 0% | 0% |
| W239S | Phe | 0.1% | 4.1% | 41.0% | 44.7% | 9.6% |
| W239S | 4-F-Phe | 18.9% | 39.7% | 33.2% | 8.1% | 0.1% |

**Table S4.** Docking scores of Phe analogs docked to GrsA-W239S.

| Ligand | Docking Score |
|---|---|
| O-propargyl-Tyr | -8.425 |
| 2,4-$F_2$-Phe | -7.560 |
| 4-F-Phe | -7.205 |
| Phe | -7.049 |

# References


1 A. Stanišić, A. Hüsken and H. Kries, *Chem. Sci.*, 2019, **10**, 10395–10399.

2 D. J. Wilson and C. C. Aldrich, *Analytical biochemistry*, 2010, **404**, 56–63.

3 R Core Team, R: A Language and Environment for Statistical Computing http://www.r-project.org 2017.

4 A. Stanišić, C.-M. Svensson, U. Ettelt and H. Kries, *bioRxiv*, 2022, 2022.08.30.505883.

5 F. Pourmasoumi, S. De, H. Peng, F. Trottmann, C. Hertweck and H. Kries, *ACS Chem. Biol.*, 2022, **17**, 2382–2388.

6 E. Conti, T. Stachelhaus, M. A. Marahiel and P. Brick, *The EMBO journal*, 1997, **16**, 4174–83.

7 G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju and W. Sherman, *J Comput Aided Mol Des*, 2013, **27**, 221–234.

8 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1–2**, 19–25.

9 S. Páll, M. J. Abraham, C. Kutzner, B. Hess and E. Lindahl, in *Solving Software Challenges for Exascale*, eds. S. Markidis and E. Laure, Springer International Publishing, Cham, 2015, pp. 3–27.

10 S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845–854.

11 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

12 X. He, V. H. Man, W. Yang, T.-S. Lee and J. Wang, *The Journal of Chemical Physics*, 2020, **153**, 114502.

13 A. W. Sousa da Silva and W. F. Vranken, *BMC Research Notes*, 2012, **5**, 367.

14 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *Journal of Molecular Graphics and Modelling*, 2006, **25**, 247–260.

15 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *Journal of Computational Chemistry*, 2004, **25**, 1157–1174.

16 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *The Journal of Chemical Physics*, 1983, **79**, 926–935.

17 G. Bussi, D. Donadio and M. Parrinello, *The Journal of Chemical Physics*, 2007, **126**, 014101.

18 M. Parrinello and A. Rahman, *Journal of Applied Physics*, 1981, **52**, 7182–7190.

19 R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, *J. Med. Chem.*, 2006, **49**, 6177–6196.

20 R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739–1749.

21 T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, *J. Med. Chem.*, 2004, **47**, 1750–1759.