Electronic Supporting Information

Biological fluorination from the sea: Discovery of a SAM-dependent nucleophilic fluorinating enzyme from the marine-derived bacterium *Streptomyces xinghaiensis* NRRL B24674

Long Ma [*a,b,c], Yufeng Li[b,c], Lingpei Meng[b,c], Hai Deng[d], Yuyin Li[b,c], Qiang Zhang[e], Aipo Diao[b,c]

[a]Biomolecular Sciences Research Complex, EaStCHEM School of Chemistry, University of St Andrews, Fife KY16 9ST, United Kingdom. E-mail: lm228@st.andrews.ac.uk
[b]Key Laboratory of Industrial Fermentation Microbiology, Ministry of Education, School of Biotechnology, Tianjin University of Science & Technology, Tianjin 300457, China. Fax: +86 22 60602298; Tel: +86 22 60602948; E-mail: malong@tust.edu.cn
[c]Tianjin Key Laboratory of Industry Microbiology, School of Biotechnology, Tianjin University of Science & Technology, Tianjin 300457, China. Tel: +0133 446 7254
[d]Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Meston Walk, Aberdeen AB24 3UE, United Kingdom.
[e]Tianjin 3rd center hospital. Tianjin 300170, China.

1. Plasmid construction

Gene sequence of the fluorinase was deduced from the whole genome of *Streptomyces xinghaiensis* NRRL B-24674. The *Streptomyces xinghaiensis* fluorinase (*flA*) and other fluorinase genes were all codon-optimised and synthesised by Shanghai Generay Biotech Co., Ltd. The synthetic gene *flA* was then cloned into a pET-28a(+) vector with two restriction sites EcoRI/HindⅢ located in the upstream and downstream of *flA* gene respectively.

2. Site-directed mutagenesis

The overlap PCR method was used to introduce D16E, Y77A, T80S, F156A, S158A mutants as previously described with slight modifications.[1] The amplified PCR product was then cloned into the pET-28a(+) vector to generate pET-28a(+)-*flA* mutants respectively, which were all sequenced and confirmed by Beijing BGI company. The primers used in site-directed mutagenesis were all listed in Tab S2.

3. Protein expression and purification

The plasmids were transformed into *E.coli* BL21(DE3) using the heat shock method. Simply, a mixture of chemically competent bacteria and plasmid DNA was placed at 42°C for 90 s and then placed back on ice. *E. coli* BL21(DE3) cells, transformed with the synthetic fluorinase plasmids were grown in Luria broth containing 100 mg/mL kanamycin at 37°C until the cell density reached an absorbance at ~0.6 at 600 nm. The culture was then cooled down on ice for 30 min. Over-expression was induced with 0.01 mM isopropylthiogalactoside (IPTG) and continued the incubation at 16°C for around 30 h. Cells were harvested by centrifugation at 6000 × g for 15 min at 4°C. The cell pellets were re-suspended in a lysis buffer (25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 2.5 mM β-mercaptoethanol) and lysed by sonication. Cell debris was removed by centrifugation (8000 × g, 20 min, 4°C). The supernatant was filtered and loaded onto a 5 mL pre-packed His-Trap FF column (GE Healthcare), which was connected to a AKTA Purifier (GE Healthcare). Protein-bound Ni^{2+} beads were washed with washing buffer (25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 50 mM imidazole) and eluted with elution buffer (25 mM Tris-HCl, pH 8.0, 300 mM NaCl, 200 mM imidazole). The eluted protein was dialysed against buffer (50 mM Tris-HCl, pH 7.0) and concentrated using a Millipore concentrator. The flow rates of AKTA Purifier for sample loading, washing and elution were all 1 mL/min. The protein concentration was measured based on Lambert-Beer Law and the extinction coefficient was determined

using the ExPAsy ProtParam tool.

4. Kinetics assay of fluorinase

For determining the kinetics of fluorinase for SAM, the fluorinase (14.0 mM) was incubated at various concentrations of SAM (from 0 to 800 μM) and a saturating concentration of KF (200 mM) in 20 mM phosphate buffer (pH 7.8) at various time points. The reaction temperature was 37ºC. Similarly, for determining the kinetics of fluorinase for KF, the fluorinase (15.5 μM) was incubated at various concentrations of KF (from 0 to 30 mM) and a saturating concentration of SAM (500 μM) in 20 mM sodium phosphate buffer (pH 7.8) at various time points. All reactions were quenched by addition of 200 μL 10% trichloroacetic acid (TCA). All reactions were centrifuged (14000 × g, 10 min at 4°C) to remove the precipitated protein. The samples were further filtered using a 0.22 μM Waterman filter membrane and subjected to HPLC analysis. The level of 5'-FDA production was determined against a standard curve of 5'-FDA. All experiments were performed in triplicate. Kinetic parameters were obtained by least-square fitting of the initial velocity against substrate concentrations based on Michaelis-Menten equation using GraphPad Prism 5.

5. Verification of fluorinase

Streptomyces xinghaiensis fluorinase activity was tested at 37°C by monitoring 5'-FDA production using analytical HPLC. Assays were performed at 37°C for 10 min in 100 μL of 20 mM sodium phosphate buffer (pH 7.8), 400 μM SAM, 0.5 mg/mL purified protein, 200 mM KF. Reactions were quenched by 200 μL 10% trichloroacetic acid (TCA) and then centrifuged (14000 × g, 10 min) to obtain the supernatant. The supernatant was further filtered using a 0.22 μM Waterman filter and subjected to analytical HPLC or LC-MS analysis.

6. Verification of chlorinase activity

Assays were performed as same as previously described for verification of fluorinase, but additional L-amino acid oxidase (Sigma-Aldrich) was added in the reaction mixture with concentration at 0.5 mg/mL for the purpose of pushing the reaction forward. Reactions were incubated at 37°C for 60 to 180 min, quenched by addition of 200 μL 10% trichloroacetic acid (TCA). The supernatant was obtained by centrifugation, which was filtered and subjected to HPLC or LC-MS analysis.

7. HPLC and LC-MS identification for 5'-FDA

The supernatant was subjected to a 1260 infinity HPLC (Agilent Technologies) or a 6120 Quadrupole LC-MS (Agilent Technologies) analysis on an analytic ZORBAX SB-C18 column (4.6×150 mm, 5 μm). The mobile phases for HPLC were composed of water containing 0.05% trifluoroacetic acid (TFA) and CH_3CN containing 0.05% TFA ; while the buffer for LC-MS was exactly the same apart from the fact that the both water and CH_3CN buffer were void of TFA. The gradient applied for both HPLC and LC-MS was shown as follows: 5% CH_3CN (with 0.05% trifluoroacetic acid) for 2 min, 5% to 25% CH_3CN over 10 min and 25% CH_3CN for 2 min, 25% to 50% CH_3CN over 4 min and 50% CH_3CN for 2 min, 50% to 95% CH_3CN over 3 min and 95% CH_3CN for 2 min, 95% to 5% CH_3CN over 3 min and 5% CH_3CN for 2 min at a flow rate of 1 mL/min under 260 nm UV spectral monitoring.

8. Circular dichroism

Circular dichroism was executed with protein concentration at 1.5 mg/mL in 20 mM sodium phosphate

buffer (pH 7.8) using MOS-450 circular dichroism instrument (BioLogic) at ambient temperature.

9. Analysis of protein stability under different conditions by thermal shift assay

Thermal shift analysis of protein stability was carried out in the presence of SYPRO Orange dye using an Applied Biosystems StepOneTM Real-Time PCR (Life Technologies). Assays were performed in a 25 μL reaction containing 3 μM fluorinase, 3 μM SYPRO Orange dye, in the presence and absence of 3 μM relevant small molecule compounds. The ROX channel was chosen for fluorescence recording. When the fluorinase melted (experienced thermal denaturation), it was able to give rise enhanced fluorescence; thus, a graph of fluorescence intensity *vs*. temperature was plotted to calculate T_m (melting temperature) values under different conditions. All experiments were performed in triplicate.

10. pH, temperature and metal ion effect on the activity of fluorinase

In order to find optimal pH value, a reaction mixture (100 μL) containing 200 mM KF, 0.4 mM SAM in 20 mM sodium acetate (pH 4.0-5.5), sodium phosphate (pH 6.0-7.0), Tris-HCl (pH 7.5-9.0), or Gly-sodium hydroxide (pH 9.5-10.0) buffer was prepared. The reaction was at 37°C for 10 min with fluorinase at 0.5 mg/mL. The supernatant of each reaction was then subjected to HPLC analysis. Effect of temperature on fluorinase activity was performed in 20 mM Tris-HCl buffer (pH 7.8) and at varied temperature. To measure the metal ion effects on the activity, a group of 100 μL reaction mixtures containing 200 mM KF, 0.4 mM SAM in 20 mM Tris-HCl (pH 7.8) buffer were mixed with 1 mM different divalent metal chloride salts (Mg^{2+}, Mn^{2+}, Fe^{2+}, Cu^{2+}, Zn^{2+}, Ca^{2+}, Ni^{2+} and Sn^{2+}) respectively or 1 mM EDTA (to remove the potential associated metal ions of the reaction). No metal ions nor EDTA was added in the 'control' reaction. The relative fluorinase activity was reflected by the production of 5'-FDA, which can be monitored by HPLC and quantified based on a standard curve of 5'-FDA. All experiments were performed in triplicate.

11. **Protein stability at room temperature**

The purified fluorinase was maintained at 25°C for 0 day, 1 day, 2 days, 3 days and 4 days respectively. Then assays were executed at 37°C for 10 min in 100 μL reaction containing 50 mM Tris-HCl buffer (pH 7.0), 400 μM SAM, 200 mM KF and 0.5 mg/mL fluorinase. Reactions were quenched by addition of 200 μL 10% trichloroacetic acid. The samples were further processed for HPLC analysis. The enzyme activity can be reflected by the production of 5'-FDA. All experiments were performed in triplicate.

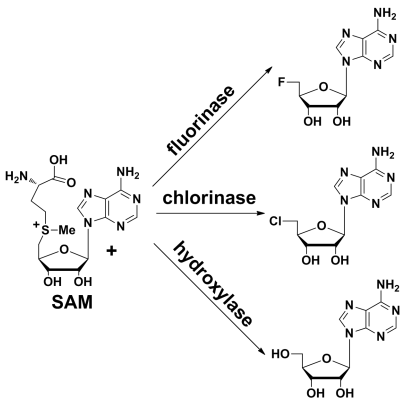

Fig. S1. A comparison of fluorinase, chlorinase and duf-62 enzymes.

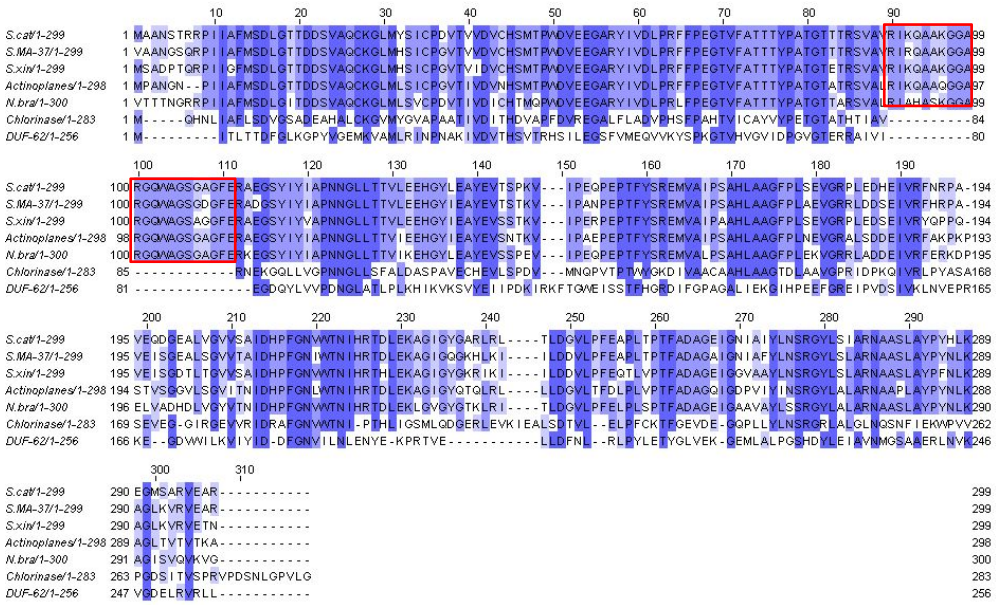

Fig. S2. Amino acid alignment of the known fluorinases and other related proteins generated by Jalview.[2] *S.cat*, *S.MA-37*, *S.xin*, *Actinoplanes* and *N.bra* denoted the fluorinases from *Streptomyces cattleya*, *Streptomyces* sp. MA37, *Streptomyces xinghaiensis* NRRL B-24674, *Actinoplanes* sp. N902-109 and *Nocardia brasiliensis* respectively.[3] Chlorinase is from the marine actinomycete *Salinospora tropica*.[4] SAM hydroxide adenosyltransferase (duf-62) from *Pyrococcus horikoshii*.[5] The diagnostic characteristic of fluorinases to chlorinase and duf-62 enzymes is the boxed 22 amino acid loop.

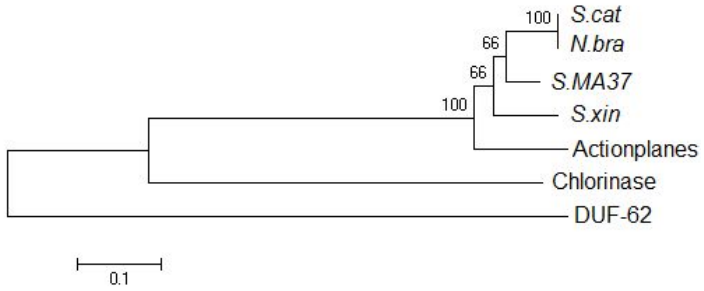

Fig. S3. Phylogenetic tree of some functionally related enzymes in terms of protein sequences generated by MEGA software.[6]

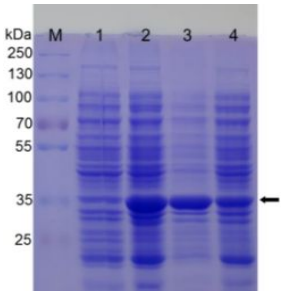

Fig. S4. SDS-PAGE showing the over-expression of fluorinases by BL21(DE3). 1: no IPTG; 2: IPTG induction; 3: insoluble fraction of over-expression; 4: souble fraction of over-expression. The over-expressed fluorinase was indicated by an arrow.

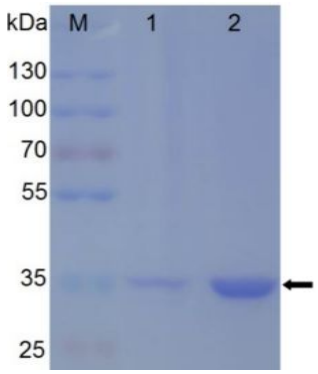

Fig. S5. SDS-PAGE showing purified fluorinase. 1: loaded purified sample; 2: over-loaded sample. The purified fluorinase was indicated by an arrow.

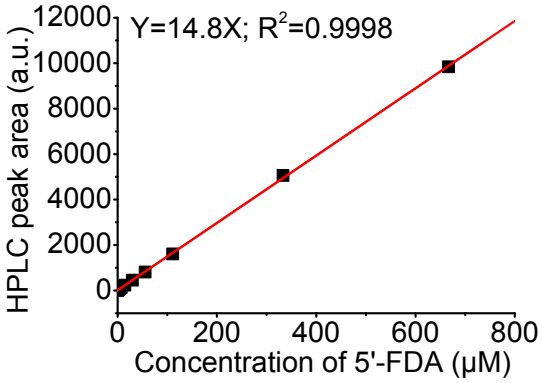

Fig. S6. Standard curve of 5'-FDA HPLC peak area against 5'-FDA concentration (μM).

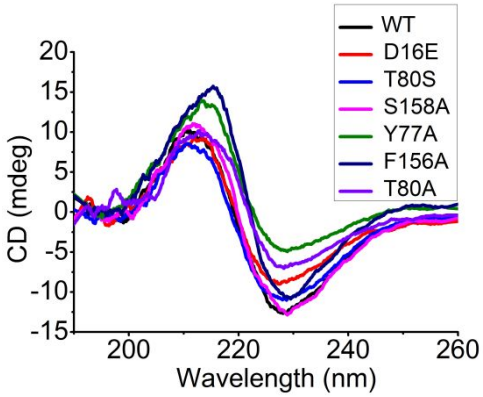

Fig. S7. CD spectra of wild type and mutated fluorinases to monitor secondary structures.

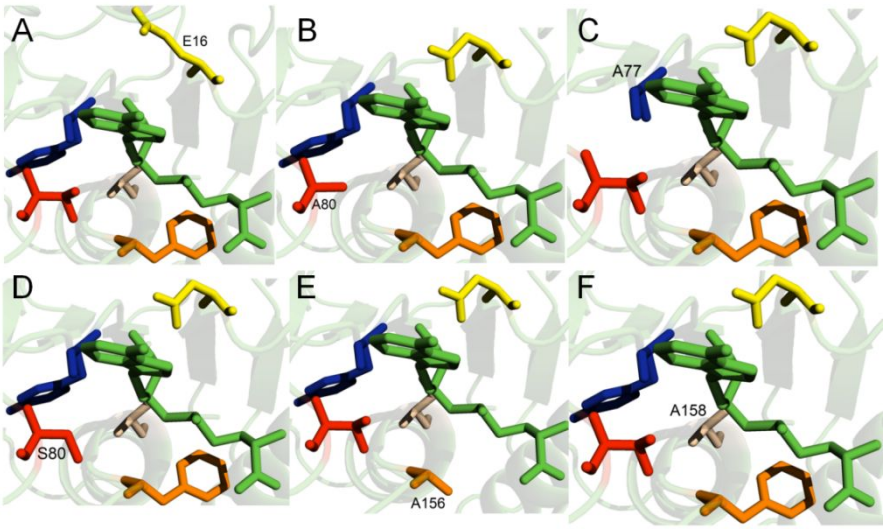

Fig. S8. Wild type and mutated *S. Xinghaiensis* fluorinase. A: D16E; B: T80A ; C: Y77A; D:T80S; E: F156A; F: S158A.

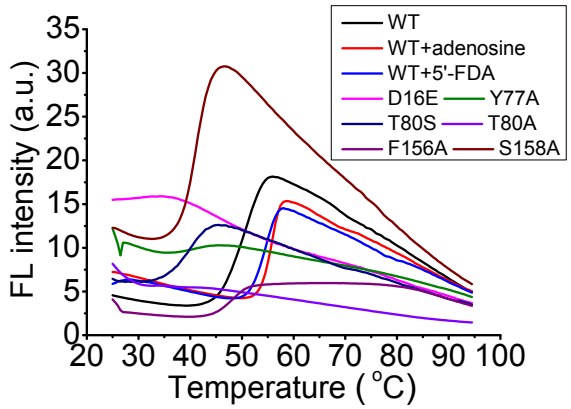

Fig. S9. Melting profile of wild-type, mutant and small molecule complexed fluorinases in 50 mM Tris-HCl buffer (pH 7.0).

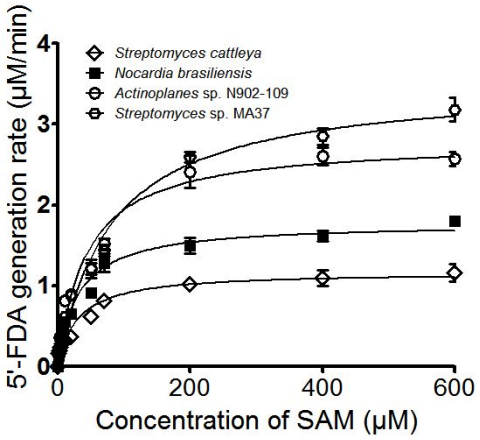

Fig. S10. Initial generation rate of 5'-FDA *vs* [SAM] for fluorinases from different sources.

Samples	Primers
WT	F:5′-GGAATTCATGTCTGCGGACCCGACCCAG-3′ R:5′-CCCAAGCTTAGTTGGTTTCAACACGAACTTTC -3′
D16E	F_m:5′-GGAATTCATGTCTGCGGACCCGACCCAGCGCCCGATC ATTGGCTTCATGTCTGAACTGGGCACTACCGACGACTCCG-3′
Y77A	F_m:5′-GCGACCACCACCGCCCCGGCGACCGG-3′ R_m:5′-TCAGTACCGGTCGCCGGGGCGGTGGTGG-3′
T80S	F_m:5′-CTACCCGGCGAGCGGTACTGAAACC-3′ R_m:5′-TACGGGTTTCAGTACCGCTCGCCGG-3′
F156A	F_m:5′-CGGAACCGACTGCATATTCTCGTGAAATGG-3′ R_m:5′-CGCAACCATTTCACGAGAATATGCAGTCGG-3′
S158A	F_m:5′-CGACTTTCTATTCTCGTGAAATGGTTGCGATCC-3′ R_m:5′-CCATTTCACGAGCATAGAAAGTCGGTTCCGGACG-3′
T80A	F_m:5′-CTACCCGGCGGCCGGTACTGAAACC-3′ R_m:5′-TACGGGTTTCAGTACCGGCCGCCGG-3′

Tab. S1. All primers used for generation of mutant fluorinases.

Samples	T_m (°C)
WT (50 mM Tris-HCl, pH 7.0)	50.9 ± 1.35
WT + Adenosine (50 mM Tris-HCl, pH 7.0)	55.5 ± 0.80
WT + 5′-FDA (50 mM Tris-HCl, pH 7.0)	54.6 ± 1.20
D16E (50 mM Tris-HCl, pH 7.0)	Not able to determine
Y77A (50 mM Tris-HCl, pH 7.0)	41.1 ± 0.39
T80S (50 mM Tris-HCl, pH 7.0)	40.5 ± 2.31
T80A (50 mM Tris-HCl, pH 7.0)	Not able to be determine
F156A (50 mM Tris-HCl, pH 7.0)	49.0 ± 1.03
S158A (50 mM Tris-HCl, pH 7.0)	41.3 ± 0.41

Tab. S2. Calculated T_m values of samples pictured on Fig. S9.

Samples	T_m (°C)
WT (50 mM Tris-HCl, pH 7.0)	50.5 ± 0.65
WT (50 mM Tris-HCl, pH 7.5)	44.3 ± 0.42
WT (50 mM Tris-HCl, pH 8.0)	43.3 ± 1.19
WT (50 mM Tris-HCl, pH 8.5)	46.5 ± 1.13
WT (100 mM phosphate buffer, pH 6.0)	43.3 ± 0.30
WT (100 mM phosphate buffer, pH 6.5)	43.5 ± 0.24
WT (100 mM phosphate buffer, pH 7.0)	44.2 ± 0.55

Tab. S3. Calculated T_m values of *Streptomyces xinghaiensis* fluorinase in different buffers.

References:

1. K. L. Heckman and L. R. Pease, *Nat. Protoc.*, 2007, **2**, 924-932.

2. A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp and G. J. Barton, *Bioinformatics*, 2009, **25**, 1189-1191.

3. H. Deng, L. Ma, N. Bandaranayaka, Z. Qin, G. Mann, K. Kyeremeh, Y. Yu, T. Shepherd, J. H. Naismith and D. O'Hagan, *ChemBioChem*, 2014, **15**, 364-368.

4. A. S. Eustaquio, F. Pojer, J. P. Noel and B. S. Moore, *Nat. Chem. Biol.*, 2008, **4**, 69-74.

5. H. Deng, S. A. McMahon, A. S. Eustáquio, B. S. Moore, J. H. Naismith and D. O'Hagan, *ChemBioChem*, 2009, **10**, 2455-2459.

6. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar, *Mol. Biol. Evol.*, 2011, **28**, 2731-2739.