A Non-conventional Archaeal Fluorinase Identified by *In Silico* Mining Catalyzes the Fastest Natural Nucleophilic Fluorination

Isabel Pardo[1]†, David Bednar[2,3], Patricia Calero[1], Daniel C. Volke[1], Jiří Damborský[2,3] & Pablo I. Nikel[1]*

[1] The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark; [2] Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, Brno, Czech Republic & [3] International Clinical Research Centre, St. Anne's University Hospital, Brno, Czech Republic.

KEYWORDS: fluorinase, genome mining, fluorine, synthetic biology, biocatalysis, metabolic engineering

Corresponding author

* E-mail: pabnik@biosustain.dtu.dk

Present Address

† Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas Margarita Salas, CSIC, Madrid, Spain

ABSTRACT

Fluorinases, the only enzymes known to catalyze the transfer of fluorine to an organic molecule, are essential biocatalysts for the sustainable synthesis of valuable organofluorines. However, the few fluorinases identified so far have low turnover rates that hamper biotechnological applications. Here, we isolated and characterized putative fluorinases retrieved from systematic *in silico* mining and identified a non-conventional archaeal enzyme from *Methanosaeta* sp. that mediates the highest fluorination rate reported to date. Furthermore, we demonstrate enhanced production of fluoronucleotides *in vivo* in a bacterial host engineered with this archaeal fluorinase, paving the way towards synthetic metabolism for efficient biohalogenations.

Fastest biological S_N2 fluorination

INTRODUCTION

Fluorinated organic compounds (organofluorines), containing at least one fluorine (F) atom, are chemicals of enormous industrial interest—as evidenced by their increasing prevalence in newly developed pharmaceuticals and agrochemicals (Harsanyi and Sandford, 2015; Inoue *et al.*, 2020; Ogawa *et al.*, 2020). The unique physicochemical properties of F endow organofluorines with advantageous properties with respect to their non-fluorinated counterparts, e.g. increased chemical stability or improved bioavailability (O'Hagan, 2008). However, the abundance of human-made organofluorines contrasts with their relative scarcity in Nature (Carvalho and Oliveira, 2017). 5'-Fluoro-5'-deoxyadenosine (5'-FDA) synthase, or fluorinase (FlA), is the only one enzyme known to naturally catalyze the formation of the high-energy C–F bond. This enzyme, originally identified in *Streptomyces cattleya* (O'Hagan *et al.*, 2002; Schaffrath *et al.*, 2003; Dong *et al.*, 2004), catalyzes the S_N2 transfer of a fluoride ion (F^-) to the C5' of the essential methyl donor *S*-adenosyl-L-methionine (SAM). As indicated in step I of **Fig. 1**, this reaction generates 5'-FDA and L-methionine (L-Met) as products (Zhu *et al.*, 2007). Since the discovery of FlA in 2003, only six other fluorinases have been reported in the literature, all of them sourced from Actinomycetes (Deng *et al.*, 2014; Ma *et al.*, 2016; Sooklal *et al.*, 2020). Additionally, a chlorinase closely related to FlAs, catalyzing 5'-chloro-5'-deoxyadenosine (5'-ClDA) synthesis (step II in **Fig. 1**), has also been identified in the marine Actinomycete *Salinispora tropica* (Eustáquio et al., 2008). FlA from *S. cattleya* is capable of catalyzing the chlorination reaction as well, albeit much less efficiently than fluorination (Deng *et al.*, 2006). Conversely, the SalL chlorinase cannot catalyze the formation of C–F bonds. This activity difference has been attributed to a 23-residue loop, present in all known FlAs but absent in SalL. It was hypothesized that this loop, located near the catalytic site, could influence halide specificity by modifying the architecture of the binding pocket.

Considering the environmentally-harsh conditions currently required for the chemical synthesis of organofluorines (Cros *et al.*, 2022), FlAs are promising biocatalysts for "green" production of new-to-Nature, bio-derived organofluorines and the implementation of synthetic metabolism with fluorinated intermediates in living cells (Martinelli and Nikel, 2019; Nieto-Domínguez and Nikel, 2020). However, all known FlAs are poor biocatalysts, with turnover rates < 1 min^{-1}. So far, the handful of protein engineering efforts aimed at the improvement of FlA activity have had limited success (Thomsen *et al.*, 2013; Sun *et al.*, 2016). Furthermore,

these studies mostly relied on employing surrogate substrates, e.g. 5'-ClDA, to select for enzyme variants with improved transhalogenation activity (steps III and I in **Fig. 1**). This strategy hampers the wide applicability of FIAs in a consolidated, whole-cell bioprocess where only F^- and an appropriate carbon substrate would be supplied as feed-stock (Calero *et al.*, 2020; Markakis *et al.*, 2020).

Genome-wide databases are a rich source of potentially valuable enzymes, yet their continuous, exponential expansion makes the selection of catalytically-attractive candidates challenging. The EnzymeMiner platform has been recently developed to address this issue (Hon *et al.*, 2020), as an inter-active and user-friendly website (https://loschmidt.chemi.muni.cz/enzymeminer). This bioinformatic tool searches through databases upon submitting a sequence of at least one representative member of the target enzyme family, together with the identification of essential (i.e. catalytic) residues. EnzymeMiner conducts multiple database searches and accompanying calculations, which provide a set of hits and their systematic annotation based on protein solubility, possible extremophilicity, domain structures and other structural information. These collected and calculated annotations provide users with key information needed for the selection of the most promising sequences for gene synthesis, small-scale protein expression, purification, and functional characterization (Vanacek *et al.*, 2018).

With the goal of expanding the FIA toolset for the biological production of organofluorines, here we describe the systematic screening and *in vitro* characterization of hitherto unknown FIAs retrieved from genome databases using EnzymeMiner. This approach led to the isolation of a non-conventional fluorinase enzyme (identified as a chlorinase) that catalyzes the fastest S_N2 fluorination reaction reported thus far. Likewise, the implementation of this enzyme in engineered bacterial cell factories gave rise to the highest fluorometabolite content produced *via de novo* biofluorination.

RESULTS & DISCUSSION

In an effort to identify "Nature's best" biocatalyst, the fluorinase from *Streptomyces* sp. MA37 (FlAMA37) was used as the query sequence (Uni-Prot W0W999), and the residues D16, Y77, S158, D210, and N215 were specified as essential based on their implication in catalysis and substrate binding (**Fig. 2a**). After curing redundant sequences, sixteen unique candidates were obtained (**Table 1** and **Fig. 2b**). Some of the retrieved amino acid sequences were found to be missing several *N*-terminal residues, which were added after manually curating the deposited genome sequences from which the fluorinase genes had been predicted (**Table S1**). Out of the sixteen sequences retrieved, five corresponded to fluorinases previously reported in the literature (thus serving as an internal quality control of the prediction routine), while nine corresponded to new putative fluorinases. Another two sequences corresponded to the chlorinase from *Salinispora tropica* CNB-440 (SalL) and a putative chlorinase from the archaea *Methanosaeta* sp. PtaU1.Bin055 (FlAPtaU1). Both of these sequences lack the 23-residue loop previously hypothesized to differentiate fluorinases from chlorinases (**Fig. 2c-e**). Notably, only four of all the retrieved sequences were not sourced from Actinobacteria. These included the putative enzymes from a *Chloroflexi* bacterium (*Chloroflexi*), *Peptococcaceae* bacterium CEB3 (Clostridia), *Thermosulforhabdus norvegica* (Deltaproteobacteria) and *Methanosaeta* sp. PtaU1.Bin055 (Methanomicrobia). Phylogenetic analysis of the 16S rRNA sequences of the fluorinase-encoding organisms gave a similar result to that obtained when using the fluorinase amino acid sequences, except that, as expected, *Salinispora tropica* groups together with the other Actinomycetes, in a clade separate from the one formed by *Streptomyces* sp. (**Table S2** and **Fig. S1**).

The genomic context of the different *flA* genes was likewise examined (**Table S3**). As reported for the fluorination gene clusters of *Streptomyces* sp. MA37, *N. brasiliensis*, *Actinoplanes* sp. N902-109, and *S. xinghaiensis*, all Actinomycetes presented gene clusters resembling that of *S. cattleya* (**Fig. 3**), the most studied source of *fl* genes to date (Huang *et al.*, 2006; O'Hagan and Deng, 2015). The genes *flB* (encoding a 5'-FDA phosphorylase), *flG* (encoding a response regulator), *flH* (encoding a putative cation:H^+ antiporter) and *flI* (encoding a *S*-adenosyl-L-homocysteinase) were conserved in all Actinomycetes. Most of them also presented the genes *flIso* (5'-fluoro-5'-deoxyadenosine-1-phosphate isomerase) and *flFT* (4-fluoro-L-threonine transaldolase), involved in the synthesis of fluoroacetate and 4-fluoro-L-threonine, respectively. In

these cases, genes encoding a prolyl-tRNA synthetase-associated protein and an EamA family transporter were usually found in proximity to *flFT*. In *S. cattleya*, the products of these genes (termed *fthB* and *fthC*, respectively) play a role in detoxification by deacylation of 4-fluoro-L-threoninyl·tRNA and export of 4-fluoro-L-threonine (McMurry and Chang, 2017). Interestingly, *Amycolatopsis bartoniae* and *Goodfellowiella* sp. AN110305 lacked either *flIso* and *flFT* orthologues within their *fl* clusters, presenting, instead, orthologues to the *fdr* genes of *Streptomyces* sp. MA37 that are probably involved in the biosynthesis of 5-fluoro-2,3,4-trihydroxypentanoic acid *via* the fluorosugar intermediate 5-fluoro-5-deoxy-D-ribose (Ma *et al.*, 2015). Activities encoded in these clusters include phosphoesterases, short chain dehydrogenases, dihydroxyacid dehydratases and cyclases, suggesting that the main fluorinated compounds produced by these organ-isms could be different from the canonical fluoroacetate and 4-fluoro-L-threonine. Similar activities seem to be also encoded by genes in the vicinity of *flA* in *Chloroflexi* bacterium and *salL* in *S. tropica* (Eustáquio *et al.*, 2009). Other genes widely distributed amongst the different actinomycotal clusters encoded activities related to SAM synthesis (i.e. SAM synthetase) and *S*-adenosyl-L-homocysteine degradation (i.e. *S*-adenosyl-L-homocysteinase), a competitive inhibitor of fluorinase activity (Schaffrath *et al.*, 2003). As indicated above, the latter gene (*flI*) was present in all actinomycotal clusters. Since SAM and *S*-adenosyl-L-homocysteine are involved in essential cellular reactions, it is likely that these enzymes modulate the levels of these compounds during secondary metabolism, when organofluorines are produced. Further analysis of the genes found in these *fl* clusters will provide clues as to what activities are needed to establish robust and efficient biofluorination pathways in heterologous hosts.

Next, coding sequences of all FIA candidates were codon-optimized for production in *Escherichia coli* as *N*-terminal His-tag fusions (**Tables S4** and **S5**)—*flA*^{MA37}, *flA*^{Scat} and *flA*^{Sxin} had been previously codon-optimized for expression in Gram-negative hosts (Calero *et al.*, 2020). SalL^{Stro} was not included in this experimental set since it is reportedly inactive on F^- (Eustáquio *et al.*, 2008). The expression of the 16 candidate genes was initially evaluated in 96-well microtiter plate cultures. FIA^{Tnor}, FIA^{Amza} and FIA^{Pbac} could not be obtained as soluble enzymes and were not included in further analyses. Moreover, very faint bands of the expected size were observed in SDS-PAGE of *E. coli* extracts producing FIA^{Tnor} and FIA^{Amza}, suggesting limited expression levels (**Fig. S2**). Therefore, we proceeded to obtain the remaining 13 candidates in medium-scale shaken-flask cultures for His-tag purification and activity assays. The purified enzymes were incubated in the

presence of increasing SAM concentrations for 1 h, after which the 5'-FDA produced was measured by HPLC. 5'-FDA synthase activity could be detected for 12 out of the 13 candidates (**Fig. S3**). The protein concentration was normalized for these assays, although the enzymes were recovered with varying degrees of purity due to differences in solubility—typical of proteins isolated from high-G+C-content species when produced in a Gram-negative host. Notably, the enzyme from *Methanosaeta* sp. (FIA^{PtaU1}, predicted to be a chlorinase), was one of the top performers, together with FIA^{SAJ15}. These enzymes had specific activities similar to those of FIA^{MA37} and FIA^{Sxin}, which present the highest catalytic efficiencies on SAM fluorination reported to date.

FIA^{PtaU1} and FIA^{SAJ15} were selected for large-scale shaken-flask production and a more detailed biochemical characterization. Steady-state kinetics assays with 1 μM of the purified protein, varying concentrations of SAM (1.5 to 800 μM) and 75 mM KF revealed that both of these enzymes presented higher turnover rates (k_{cat}) than FIA^{MA37} and FIA^{Sxin} (**Fig. 4a** and **Table 2**). Surprisingly, K_M^{SAM} values were < 10 μM, much lower than what had been previously reported in the literature for fluorinases (Schaffrath *et al.*, 2003; Zhu *et al.*, 2007; Sooklal *et al.*, 2020). Previous studies used high enzyme concentrations (>10 μM), which impedes reaching a steady state of the reaction for substrate concentrations below 10 μM. Additionally, we have used a KF concentration that ensures F^- saturation without causing any inhibitory effect (previous studies have used KF concentrations > 200 mM).

To gain insight on the structural factors that could determine these differences in activity, we inspected the predicted crystal structures of FIA^{MA37}, FIA^{Sxin}, FIA^{SAJ15}, FIA^{PtaU1} and $SalL^{Stro}$. Examination of the residues potentially interacting with SAM (distances < 5 Å) revealed important variations between the substrate binding pocket of FIAPtaU1 and that of the other fluorinases (**Fig. 4c-f**). These differences were mostly located near the adenyl moiety of SAM, and involve the substitution of a conserved proline for an arginine residue and an RNAA motif for YYGG. This motif is found in the *C*-terminal domain of fluorinases, which is more variable than the *N*-terminal domain (**Fig. S4**) and is presumably also involved in hexamer formation (Kittilä *et al.*, 2022). Interestingly, the variations found in FIA^{PtaU1} do not resemble those seen in the $SalL^{Stro}$ chlorinase. The evaluation of the effect of these amino acid differences in fluorinase activity will be of interest for future enzyme engineering efforts.

Since FIA^{PtaU1} was predicted to be a chlorinase, we evaluated whether it was also active towards S_N2-dependent addition of Cl^- onto SAM. Unexpectedly, no 5'-ClDA accumulation could be detected in enzymatic reactions in which KF was replaced by KCl—in contrast to SalL^{Stro} (Eustáquio *et al.*, 2008). Previous studies have shown that FIA^{Scat} can also catalyze the chlorination reaction (Deng *et al.*, 2006). However, this feature requires the simultaneous removal of L-Met or 5'-ClDA, the reaction products, since the reverse dehalogenation reaction is favored. We could observe transhalogenation on 5'-ClDA (i.e. 5'-FDA production in the presence of L-Met and F^-, steps III and I in **Fig. 1**, respectively; see **Fig. 3b**). Again, FIA^{PtaU1} catalytically outperformed all other fluorinases, with a 3-fold higher V_{max} value. Although we cannot rule out that FIAPtaU1 could execute *de novo* chlorination, it is clear that the 23-residue loop reportedly found in «conventional» fluorinases is not essential for the activity towards F^-.

On this background, we tested the biosynthesis of fluorometabolites both *in vitro* and *in vivo* by engineering selected fluorinases in *Pseudomonas putida*. We have previously designed a fluoride-responsive genetic circuit that enabled biofluorination in this heterologous host (Calero *et al.*, 2020). Here, this system was adapted to express either *flA*^{PtaU1} or *flA*^{SAJ15}, codon-optimized to facilitate expression in *P. putida*. Upon induction of the system with NaF and expression of the fluorinase genes for 20 h at 30°C, production of 5'-FDA was determined by LC-MS to evaluate *de novo* fluorination activity *in vivo* (**Fig. 4**). Production of 5'-FDA could be detected for cells expressing either *flA*^{PtaU1}, *flA*^{SAJ15}, *flA*^{MA37} and *flA*^{Sxin}. Notably, intracellular 5'-FDA, indicative of *in vivo* biofluorination, was 6- to 12-fold higher in cells expressing *flA*^{PtaU1} with respect to the other three fluorinase genes. On the other hand, fluorinase activity from cell-free extracts incubated for 20 h at 30°C in the presence of exogenously-added 200 μM SAM and 5 mM NaF was similar in all cases (data not shown).

CONCLUSION

In conclusion, out of the 10 newly identified enzymes, the non-conventional FIA from the archaea *Methanosaeta* sp. PtaU1.Bin055 was found to present turnover rates far superior than those of all FIAs reported to date. Surprisingly, this enzyme lacks the loop that was hypothesized to be a differentiating feature between fluorinases and chlorinases, challenging the prevailing view that this loop is required for activity

towards F^-. This work highlights the importance of systematic and efficient biocatalyst selection across ever-expanding genomic databases followed by careful characterization *in vitro* and cell factory engineering *in vivo*. These results open avenues for the implementation of neo-metabolic pathways incorporating F in the host cell by synthetic biology approaches (Cros *et al.*, 2022).

ABBREVIATIONS

5'-ClDA, 5'-chloro-5'-deoxyadenosine; 5'-FDA, 5'-fluoro-5'-deoxyadenosine; FIA, fluorinase; L-Met, L-methionine; SAM, *S*-adenosyl-L-methionine; SalL, chlorinase.

REFERENCES

Calero, P., Volke, D.C., Lowe, P.T., Gotfredsen, C.H., O'Hagan, D., and Nikel, P.I. (2020) A fluoride-responsive genetic circuit enables *in vivo* biofluorination in engineered *Pseudomonas putida*. *Nat. Commun.* **11**: 5045.

Carvalho, M.F., and Oliveira, R.S. (2017) Natural production of fluorinated compounds and biotechnological prospects of the fluorinase enzyme. *Crit. Rev. Biotechnol.* **37**: 880-897.

Cros, A., Alfaro-Espinoza, G., de Maria, A., Wirth, N.T., and Nikel, P.I. (2022) Synthetic metabolism for biohalogenation. *Curr. Opin. Biotechnol.* **74**: 180-193.

Deng, H., Cobb, S.L., McEwan, A.R., McGlinchey, R.P., Naismith, J.H., O'Hagan, D.*, et al.* (2006) The fluorinase from *Streptomyces cattleya* is also a chlorinase. *Angew. Chem. Int. Ed. Engl.* **45**: 759-762.

Deng, H., Ma, L., Bandaranayaka, N., Qin, Z., Mann, G., Kyeremeh, K.*, et al.* (2014) Identification of fluorinases from *Streptomyces* sp MA37, *Nocardia brasiliensis*, and *Actinoplanes* sp N902-109 by genome mining. *ChemBioChem* **15**: 364-368.

Dong, C., Huang, F., Deng, H., Schaffrath, C., Spencer, J.B., O'Hagan, D., and Naismith, J.H. (2004) Crystal structure and mechanism of a bacterial fluorinating enzyme. *Nature* **427**: 561-565.

Eustáquio, A.S., Pojer, F., Noel, J.P., and Moore, B.S. (2008) Discovery and characterization of a marine bacterial SAM-dependent chlorinase. *Nat. Chem. Biol.* **4**: 69-74.

Eustáquio, A.S., McGlinchey, R.P., Liu, Y., Hazzard, C., Beer, L.L., Florova, G., *et al.* (2009) Biosynthesis of the salinosporamide A polyketide synthase substrate chloroethylmalonyl-coenzyme A from *S*-adenosyl-L-methionine. *Proc. Natl. Acad. Sci. USA* **106**: 12295-12300.

Harsanyi, A., and Sandford, G. (2015) Organofluorine chemistry: applications, sources and sustainability. *Green Chem.* **17**: 2081-2086.

Hon, J., Borko, S., Stourac, J., Prokop, Z., Zendulka, J., Bednar, D., *et al.* (2020) *EnzymeMiner*: Automated mining of soluble enzymes with diverse structures, catalytic properties and stabilities. *Nucleic Acids Res.* **48**: W104-W109.

Huang, F., Haydock, S.F., Spiteller, D., Mironenko, T., Li, T.L., O'Hagan, D., *et al.* (2006) The gene cluster for fluorometabolite biosynthesis in *Streptomyces cattleya*: a thioesterase confers resistance to fluoroacetyl-coenzyme A. *Chem. Biol.* **13**: 475-484.

Inoue, M., Sumii, Y., and Shibata, N. (2020) Contribution of organofluorine compounds to pharmaceuticals. *ACS Omega* **5**: 10633-10640.

Kittilä, T., Calero, P., Fredslund, F., Lowe, P.T., Tezé, D., Nieto-Domínguez, M., *et al.* (2022) Oligomerization engineering of the fluorinase enzyme leads to an active trimer that supports synthesis of fluorometabolites in vitro. *Microb. Biotechnol. In press,* DOI: 10.1111/1751-7915.14009.

Ma, L., Bartholomé, A., Tong, M.H., Qin, Z., Yu, Y., Shepherd, T., *et al.* (2015) Identification of a fluorometabolite from *Streptomyces* sp. MA37: (2*R*3*S*4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid. *Chem. Sci.* **6**: 1414.

Ma, L., Li, Y., Meng, L., Deng, H., Li, Y., and Diao, A. (2016) Biological fluorination from the sea: discovery of a SAM-dependent nucleophilic fluorinating enzyme from the marine-derived bacterium *Streptomyces xinghaiensis* NRRL B24674. *RSC Adv.* **6**: 27047-27051.

Markakis, K., Lowe, P.T., Davison-Gates, L., O'Hagan, D., Rosser, S.J., and Elfick, A. (2020) An engineered *E. coli* strain for direct *in vivo* fluorination. *ChemBioChem* **21**: 1856-1860.

Martinelli, L., and Nikel, P.I. (2019) Breaking the state-of-the-art in the chemical industry with new-to-Nature products *via* synthetic microbiology. *Microb. Biotechnol.* **12**: 187-190.

McMurry, J.L., and Chang, M.C.Y. (2017) Fluorothreonyl-tRNA deacylase prevents mistranslation in the organofluorine producer *Streptomyces cattleya*. *Proc. Natl. Acad. Sci. USA* **114**: 11920-11925.

Nieto-Domínguez, M., and Nikel, P.I. (2020) Intersecting xenobiology and *neo*-metabolism to bring novel chemistries to life. *ChemBioChem* **21**: 2551-2571.

O'Hagan, D., Schaffrath, C., Cobb, S.L., Hamilton, J.T., and Murphy, C.D. (2002) Biochemistry: biosynthesis of an organofluorine molecule. *Nature* **416**: 279.

O'Hagan, D. (2008) Understanding organofluorine chemistry. An introduction to the C–F bond. *Chem. Soc. Rev.* **37**: 308-319.

O'Hagan, D., and Deng, H. (2015) Enzymatic fluorination and biotechnological developments of the fluorinase. *Chem. Rev.* **115**: 634-649.

Ogawa, Y., Tokunaga, E., Kobayashi, O., Hirai, K., and Shibata, N. (2020) Current contributions of organofluorine compounds to the agrochemical industry. *iScience* **23**: 101467.

Schaffrath, C., Deng, H., and O'Hagan, D. (2003) Isolation and characterisation of 5'-fluorodeoxyadenosine synthase, a fluorination enzyme from *Streptomyces cattleya*. *FEBS Lett.* **547**: 111-114.

Sooklal, S.A., de Koning, C., Brady, D., and Rumbold, K. (2020) Identification and characterisation of a fluorinase from *Actinopolyspora mzabensis*. *Protein Expr. Purif.* **166**: 105508.

Sun, H., Yeo, W.L., Lim, Y.H., Chew, X., Smith, D.J., Xue, B.*, et al.* (2016) Directed evolution of a fluorinase for improved fluorination efficiency with a non-native substrate. *Angew. Chem. Int. Ed.* **55**: 14277-14280.

Thomsen, M., Vogensen, S.B., Buchardt, J., Burkart, M.D., and Clausen, R.P. (2013) Chemoenzymatic synthesis and in situ application of *S*-adenosyl-L-methionine analogs. *Org. Biomol. Chem.* **11**: 7606-7610.

Vanacek, P., Sebestova, E., Babkova, P., Bidmanova, S., Daniel, L., Dvořák, P.*, et al.* (2018) Exploration of enzyme diversity by integrating bioinformatics with expression analysis and biochemical characterization. *ACS Catal.* **8**: 2402-2412.

Zhu, X., Robinson, D.A., McEwan, A.R., O'Hagan, D., and Naismith, J.H. (2007) Mechanism of enzymatic fluorination in *Streptomyces cattleya*. *J. Am. Chem. Soc.* **129**: 14597-14604.

TABLES

Table 1. Putative fluorinases retrieved from EnzymeMiner search using FIA^{MA37} as query.

Name	Organism	ID (%)[a]
FIA^{MA37}	*Streptomyces* sp. MA37	*Query*
FIA^{Scat}	*Streptomyces cattleya*	87.6%
FIA^{Sxin}	*Streptomyces xinghaiensis*	86.0%
FIA^{SAJ15}	*Streptomyces* sp. SAJ15	85.0%
FIA^{N902}	*Actinoplanes* sp. N902-109	80.7%
FIA^{Amza}	*Actinopolyspora mzabensis*	78.9%
FIA^{Abar}	*Amycolatopsis bartoniae*	79.1%
FIA^{CA12}	*Amycolatopsis* sp. CA-128772	78.6%
FIA^{AN11}	*Goodfellowiella* sp. AN110305	77.7%
FIA^{Nbra2}	*Nocardia brasiliensis* IFM 10847	75.7%
FIA^{Nbra3}	*Nocardia brasiliensis* NCTC 11294	75.3%
FIA^{Nbra1}	*Nocardia brasiliensis* ATCC 700358	75.3%
FIA^{Cbac}	*Chloroflexi* bacterium	69.3%
FIA^{Pbac}	*Peptococcaceae* bacterium CEB3	64.8%
FIA^{Tnor}	*Thermodesulforhabdus norvegica*	54.5%
FIA^{PtaU1}	*Methanosaeta* sp. PtaU1.Bin055	49.5%
$SalL^{Stro}$	*Salinispora tropica* CNB-440	35.6%

[a] Sequence identity. References to known FIAs are indicated.

Table 2. Michaelis-Menten kinetic constants of selected fluorinases.[a]

Fluorinase	K_M^{SAM} (μM)	k_{cat} (min^{-1})	k_{cat}/K_M^{SAM} (mM^{-1} min^{-1})
FIA^{MA37}	4.42 ± 0.58	0.16 ± 0.01	36.36 ± 4.82
FIA^{Sxin}	3.76 ± 0.15	0.22 ± 0.01	58.63 ± 2.63
FIA^{SAJ15}	9.62 ± 1.43	0.34 ± 0.01	35.81 ± 5.43
FIA^{PtaU1}	6.99 ± 1.06	0.41 ± 0.01	57.54 ± 8.85

[a] Assays conducted in 50 mM HEPES, pH 7.8, with 75 mM KF and varying SAM concentrations at 37°C. Average and standard deviation are given for triplicate independent measurements.

FIGURES

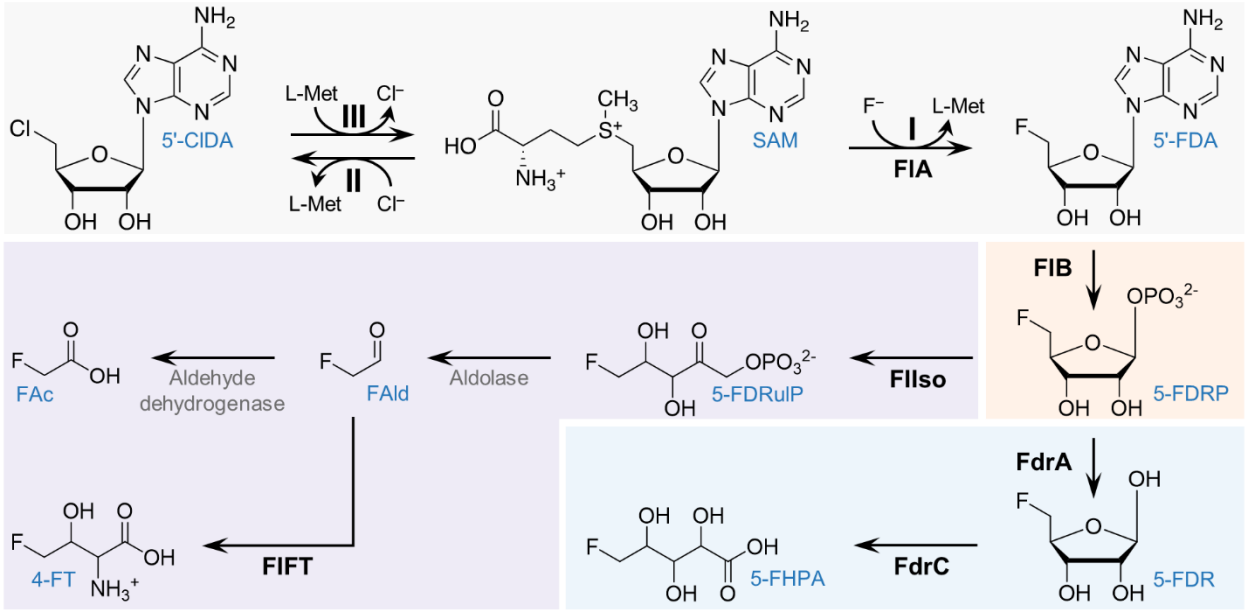

Fig. 1 · In gray, reactions catalyzed by the fluorinase/chlorinase enzyme: (I) forward fluorination reaction; (II) forward chlorination reaction; and (III) reverse chlorination reaction. In orange, common step in fluorometabolite biosynthetic pathways. In purple, canonical fluoroacetate and 4-fluoro-L-threonine biosynthetic pathway. In blue, 5'-fluoro-5'-deoxy-D-ribose biosynthetic pathway. Compound abbreviations (blue font): 5'-ClDA, 5'-chloro-5'-deoxyadenosine; SAM, *S*-adenosyl-L-methionine; 5'-FDA, 5'-fluoro-5'-deoxyadenosine; 5-FDRP, 5-fluoro-5-deoxy-D-ribose-1-phosphate; 5-FDRuIP, 5-fluoro-5-deoxy-D-ribulose-1-phosphate; FAld, fluoroacetaldehyde; FAc, fluoroacetate; 4-FT, 4-fluoro-L-threonine; 5-FDR, 5-fluoro-5-deoxy-D-ribose; 5-FHPA, 5-fluoro-2,3,4-trihidroxypentanoic acid. Enzyme abbreviations (in bold): FIA, fluorinase; FIB, 5'-fluoro-5'-deoxyadenosine phosphorylase; FIIso, 5-fluoro-5-deoxy-D-ribose-1-phosphate isomerase; FIFT, 4-fluoro-L-threonine transaldolase; FdrA, 5-fluoro-5-deoxy-D-ribose-1-phosphate phosphoesterase; FdrC, 5-fluoro-5-deoxy-D-ribose dehydrogenase.

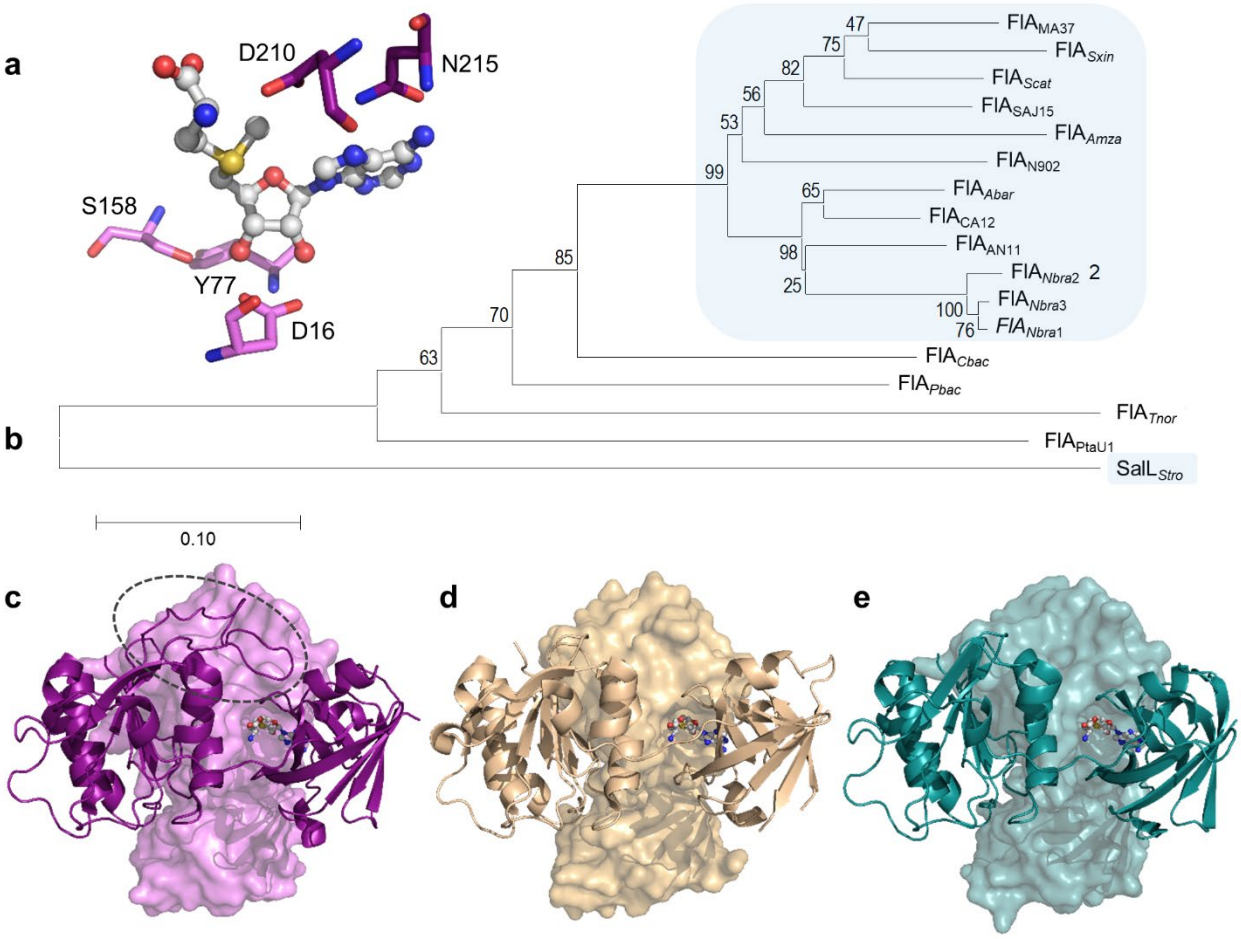

Fig. 2 · Putative fluorinases identified by genome mining. (**a**) Residues specified as essential for the EnzymeMiner search, based on the crystal structure of FIA^{MA37} (PDB ID 5B6I). The SAM substrate is shown in ball-and-stick representation. (**b**) Phylogenetic tree of retrieved fluorinase sequences obtained using MEGA-X software, inferred using the Neighbor-Joining method with a bootstrap of 10,000 iterations. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Sequences sourced from Actinomycetes are highlighted as blue squares. (**c-e**) 3D structures for FIA^{MA37} (**c**), SalL^{Stro} (**d**, PDB ID 6RYZ), and FIA^{PtaU1} (e, modelled from the FIA^{Scat} crystal structure PDB ID 2V7V). The loop hypothesized to differentiate fluorinases from chlorinases is circled in a dashed gray line. Two chains from the homo-trimer for each structure are shown as cartoon and surface representations, respectively.

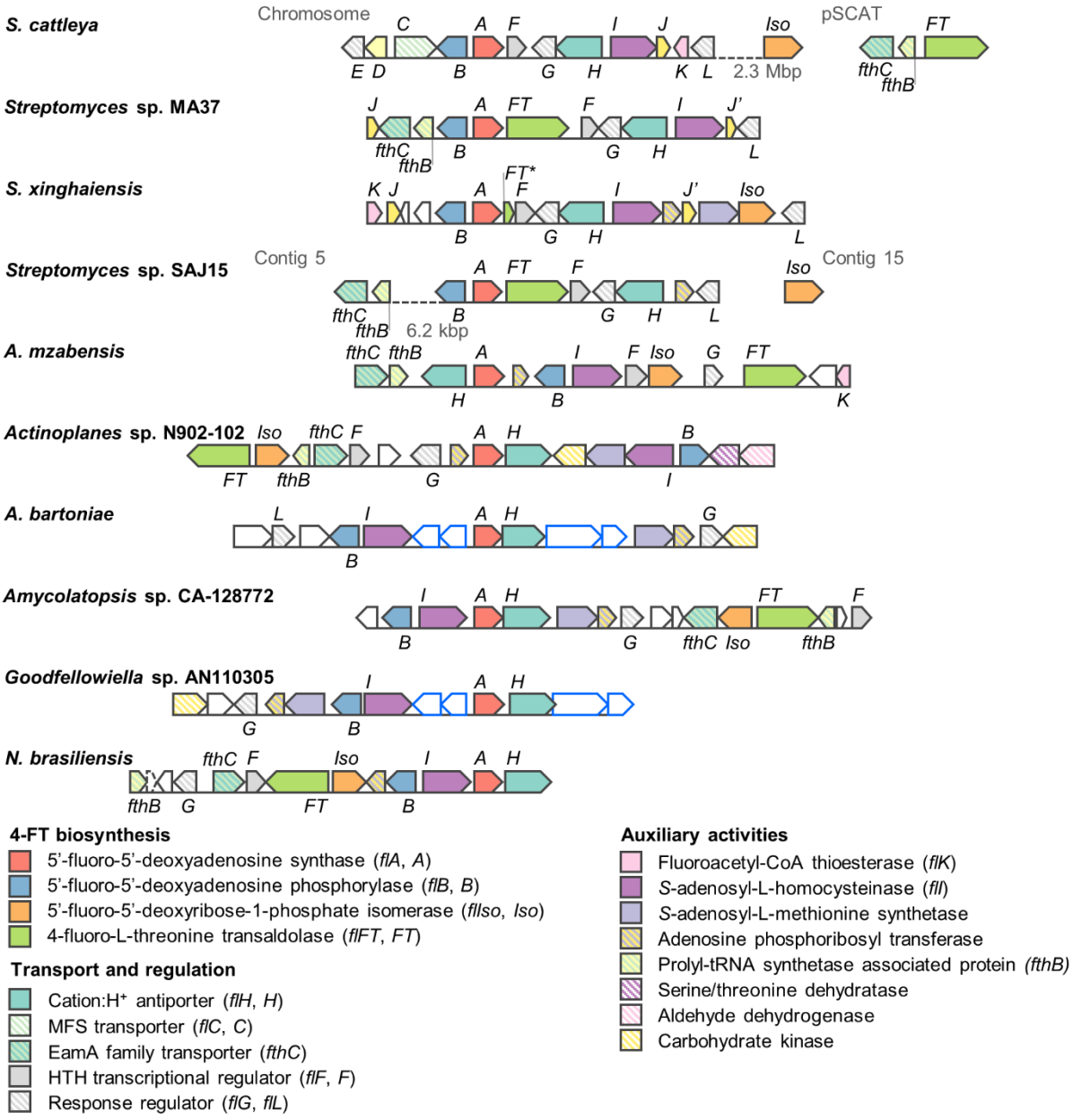

Fig. 3 · Fluorination gene clusters from Actinomycetes. For clarity, the clusters are drawn centered on *flA* (*A*) in the sense orientation. Numbers under dashed lines indicate the distance between open reading frames (ORFs) found in the same sequence entry; ORFs in separate entries are not connected by a line. Italicized letters indicate orthologues to the corresponding *fl* genes from *S. cattleya. J'* indicates duplicate *flJ* copies (encoding DUF190 domain-containing protein). *FT** is a truncated pseudogene homologous to *flFT*. Orthologues to *fdr* genes from *Streptomyces* sp. MA37 are indicated as white blocks with blue outlines. ORFs

outlined in black represent genes with other/unknown functions. MFS, major facilitator superfamily; HTH, helix-turn-helix.

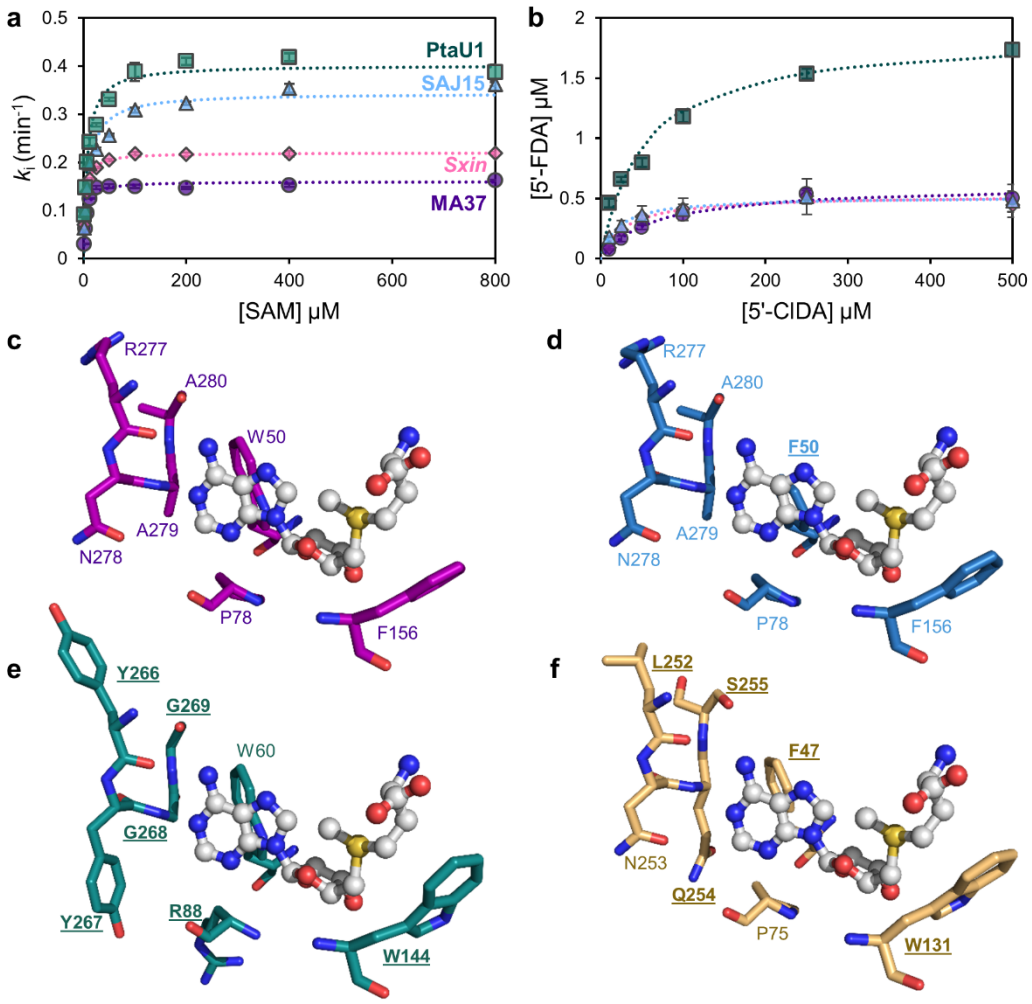

Fig. 4 · Biochemical characterization and residue conservation of selected fluorinases. (**a**) Steady-state fluorination assays using increasing SAM concentrations. Reactions were carried out at 37°C in 50 mM HEPES buffer, pH 7.8, with 75 mM KF. Dotted lines show fits to the Michaelis-Menten equation ($R^2 > 0.95$ in all cases). (**b**) End-point (1 h) transhalogenation assays with increasing 5'-CIDA concentrations. Reactions were carried out at 37°C in 50 mM HEPES buffer, pH = 7.8, with 75 mM KF and 1 mM L-Met. Error bars represent standard deviations from triplicate independent assays. Symbols and color codes are kept in both panels. (**c-f**) Variable residues in the substrate binding pocket of FIA^{MA37} (**c**), FIA^{SAJ15} (**d**), FIA^{PtaU1} (**e**) and SalL^{Stro}. Residues that differ from those of FIA^{MA37} are labelled in underlined bold font. FIA^{Sxin} residues are identical to those of FIA^{MA37}. The SAM substrate is shown in ball-and-stick representation.

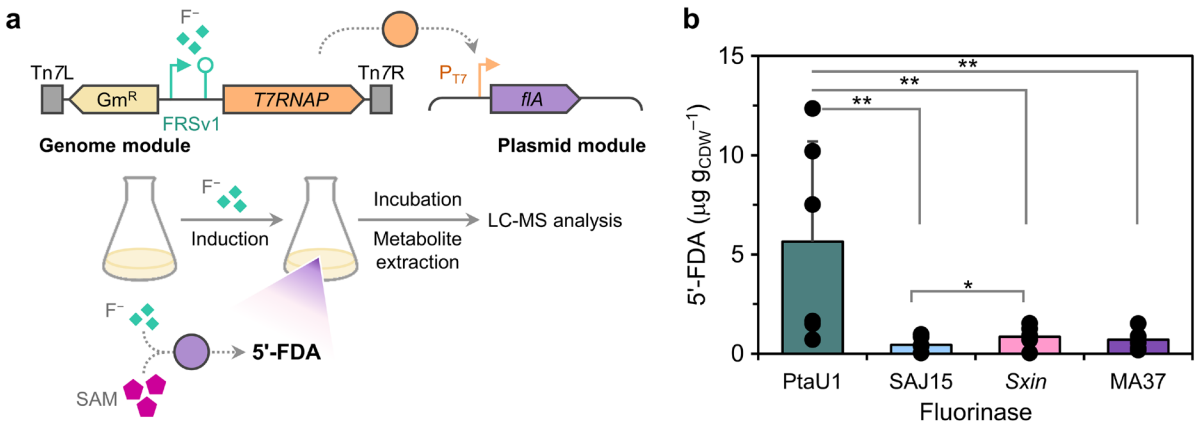

Fig. 5 · Implementing *in vivo* biofluorination in engineered *P. putida* KT2440. **(a)** Schematic representation of the fluoride-responsive genetic circuit based on the T7 phage RNA polymerase (T7RNAP) described by Calero *et al.* (2020) and workflow for the *in vivo* biofluorination assays in *P. putida*. The fluoride-responsive genetic circuit was induced by the addition of 15 mM NaF to the cultures at an OD_{600} = 0.4–0.6. After incubating the cultures at 30°C for 20 h, an aliquot was taken for metabolite extraction, biomass quantification and LC-MS analysis of fluorometabolites. Further details are provided in the Supporting Information. **(b)** Intracellular 5'-FDA content in engineered *P. putida* expressing different fluorinase genes. The fluorometabolite concentration was normalized by the cell dry weight (CDW). Black dots show individual values of independent experiments and error bars represent standard deviations. Asterisks indicate significant differences with *p*-values < 0.1 (*) or <0.05 (**) for a two-sample, one-sided Welch's *t*-test.