Electronic Supplementary Material (ESI) for ChemComm.

Supporting Information

A coupled chlorinase-fluorinase system with high efficiency of trans-halogenation and a shared substrate tolerance

Huihua Sun, Huimin Zhao,* Ee Lui Ang*

Experimental Methods

Chemicals

5'-ClDA, 5'-FDA, 5'-ClDEA and 5'-FDEA were chemically synthesized as described previously[1, 2]. L-Met, L-SeMet, and SAM were purchased from Sigma-Aldrich.

Halogenase expression in *E. coli* and purification

The genes encoding the halogenases (protein sequences are listed on Page 8) were synthesized (GenScript) and cloned into the expression vector pET28a (Novagen) using the *Nde*I and *Xho*I restriction sites. The plasmids harboring the halogenase genes were transformed into *E. coli* BL21 (DE3) strain. The *E. coli* containing the plasmid was grown in LB medium with 50 μg mL^{-1} kanamycin at 37°C until an absorbance of 0.6 at 600 nm was reached. Overexpression was induced by adding 0.2 mM isopropylthiogalactoside (IPTG) and incubation was continued at 18°C for 20 h. Cells were harvested by centrifugation (3830 x *g* for 20 min). The cell pellet was resuspended in 20 mL binding buffer (50 mM sodium phosphate buffer, pH 7.8, 300 mM NaCl, 5 mM imidazole, 10% glycerol) and the cells were lysed using a cell disruptor (Constant Systems) at 38.7 kPSI. The lysate was then centrifuged and the clarified supernatant was added to TALON metal affinity resin (Clontech). The supernatant-resin mixture was incubated at 4°C with rotation for 1 h. The His-tagged protein bound resin was washed with the binding buffer and eluted with the elution buffer (50 mM sodium phosphate buffer, pH 7.8, 300 mM NaCl, 250 mM imidazole, 10% glycerol). The elution was desalted using PD-10 desalting column (GE Healthcare) in storage buffer (50 mM sodium phosphate buffer, pH 7.8, 10% glycerol). The eluted protein was then concentrated using a 30 kDa concentrator (Sartorius) with centrifugation at 6000 x *g*. Its purity was determined to be > 90% by SDS-PAGE (**Figure S8**).The protein concentration was measured using a NanoDrop 2000 spectrophotometer (Thermo Scientific). The extinction coefficient was determined using the ExPAsy ProtParam tool. The protein yields for ClA1 and ClA2 were about 10 mg from 500 mL culture.

Rate-limiting step determination (as shown in scheme 1):

For all the enzymatic reactions, the reaction buffer is 50 mM sodium phosphate buffer with 10% glycerol at pH 7.8.

Conversion of 5'-ClDA to 5'-FDA: 0.2 mM 5'-ClDA, 80 mM NaF, 0.1 mM L-Met, 50 μM FlA4.
Conversion of 5'-ClDA to SAM: 0.2 mM 5'-ClDA, 0.1 mM L-Met, 50 μM FlA4.
Conversion of SAM to 5'-FDA: 0.2 mM SAM, 80 mM NaF, 50 μM FlA4.

All reactions were carried out in triplicate and were incubated at 47°C for 10 min. Reactions were stopped by heating the samples at 95°C for 1 min. The precipitated protein was then removed by centrifugation at 20000 x *g* for 10 min. The supernatant was used for HPLC analysis.

Kinetic assays

Conversion of SAM to 5'-ClDA: To determine the enzyme kinetics for the substrate Cl^-, the enzyme was incubated with various concentrations of NaCl and 200 μM SAM at initial time points. To determine the enzyme kinetics for the substrate SAM, the enzyme was incubated with various concentrations of SAM and 200 mM NaCl at initial time points.

Conversion of 5'-ClDA to SAM: To determine the enzyme kinetics for the substrate 5'-ClDA, the enzyme was incubated at various concentrations of 5'-ClDA and 20 mM L-Met at initial time points.

1 μM enzyme was used for kinetic assays for SalL, ClA1 and ClA2. 5 μM enzyme was used for kinetic assays for FlA4. All the reactions were carried out at 37°C and were stopped by heating the samples at 95°C for 1 min. Kinetic parameters were obtained by the best-fit model of initial velocity against substrate concentrations based on Michaelis-Menten equation using GraphPad Prism 7 (GraphPad Software).

Trans-halogenation by the coupled fluorinase-chlorinase system

Coupled fluorinase-chlorinase system: 50 μM FlA4, and 30 μM chlorinase, 0.2 mM 5'-ClDA/5'-ClDEA, 80 mM NaF, 0.1 mM L-Met/L-seMet.
Reactions without chlorinase were run with 50 μM or 80 μM FlA4 for comparison.

All reactions were carried out in triplicate at 37°C for 1 h. Reactions were stopped by heating the samples at 95°C for 1 min. The precipitated protein was then removed by centrifugation at 20000 x *g* for 10 min. The supernatant was used for HPLC analysis.

HPLC analysis

HPLC analysis was carried out using Shimadzu Prominence system UFLC (Shimadzu). Mobile phase A: 0.1% formic acid in water; B: 0.1% formic acid in methanol. Phenomenex, Kinetex® 2.6 μm Biphenyl 100 Å, LC Column 150 x 4.6 mm, 0.6 mL min^{-1} flow rate, gradient elution, 5 – 95% B for 15 min. Quantification is based on the standard curve of the peak area against various concentrations of the compound.

Supplementary Figures

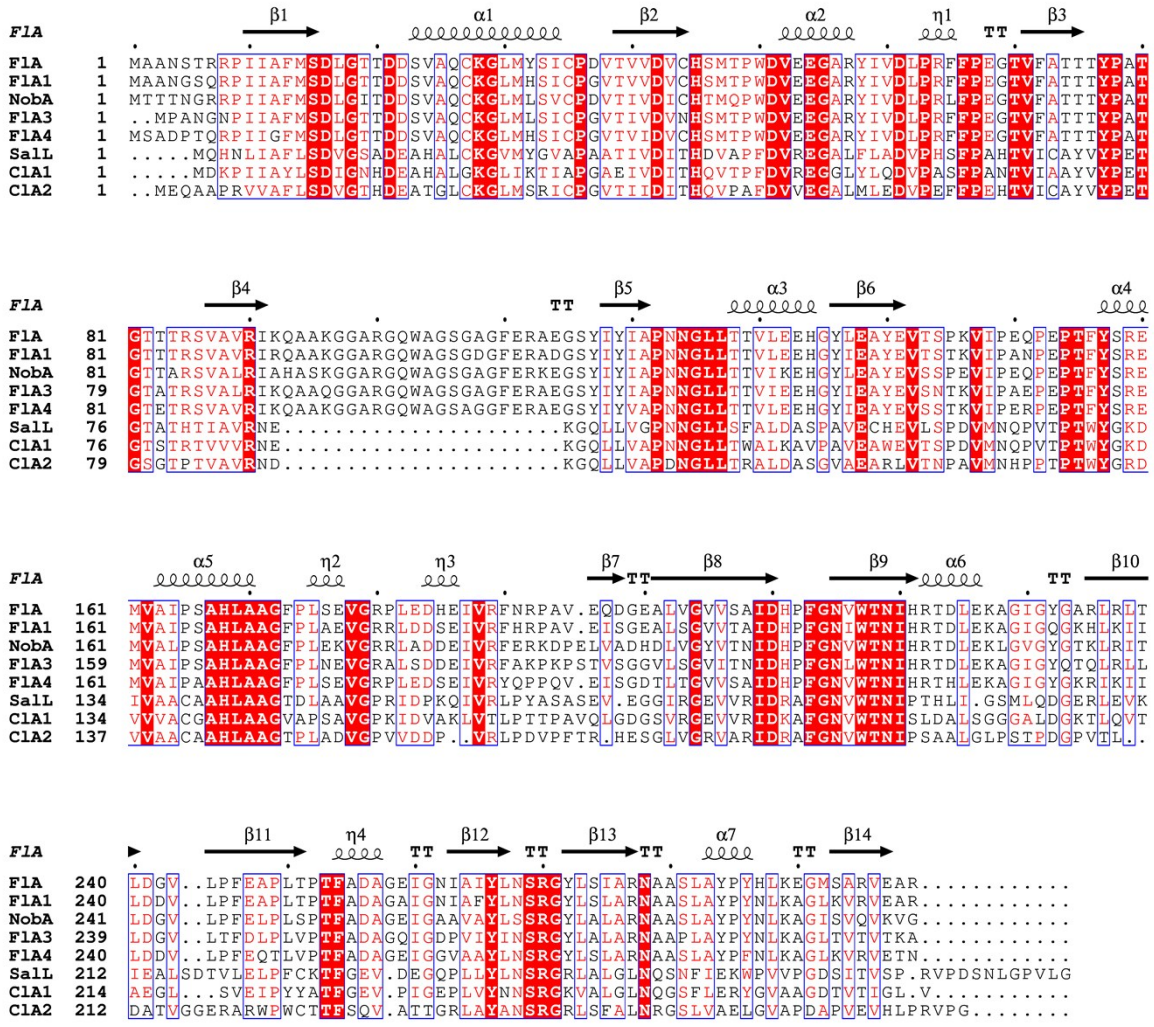

Figure S1. Multiple sequence alignment of 3 chlorinases and 5 fluorinases (FIA[3, 4] from *Streptomyces cattleya*, FIA1[5] from *Streptomyces* sp. MA37, NobA[5, 6] from *Nocardia brasiliensis*, FIA3[5] from *Actinoplanes* sp. N902-109, and FIA4[7, 8] from *Streptomyces xinghaiensis*). The secondary structure of FIA is based on PDB 1RQP. Sequence alignment was produced by ESPript 3.0, with pre-alignment input from Clustal Omega. The protein sequences are in Page 10 of the supplementary information.

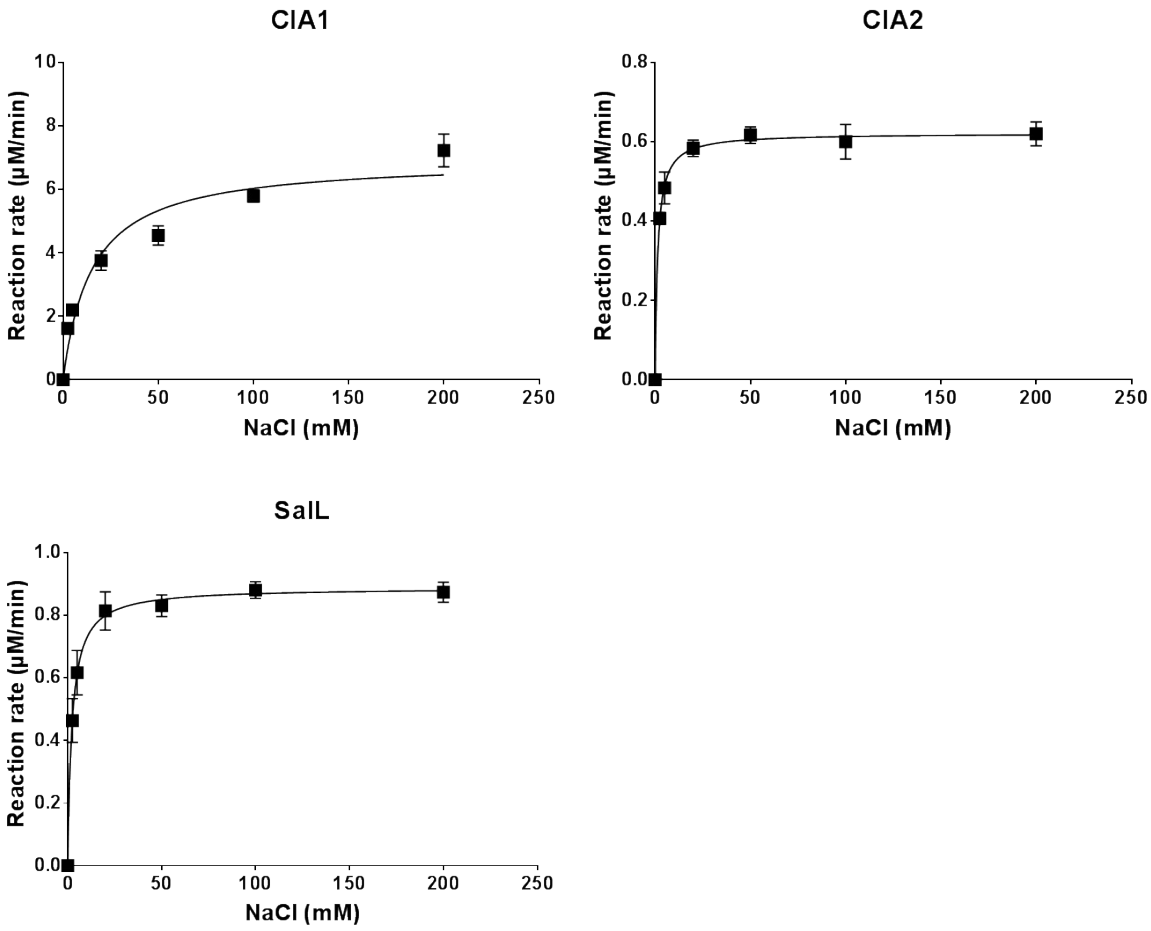

Figure S2. Kinetic assays of ClA1, ClA2 and SalL for the conversion of SAM to 5'-ClDA. Assays contain 100 μM SAM and various concentrations of NaCl.

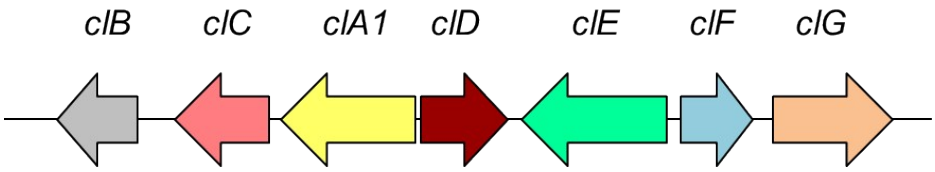

Figure S3. *clA1* and neighboring genes shared among multispecies of *Streptomyces*. Genes and the putative proteins encoded by the genes: *clB*, hypothetical protein (WP_016573067); *clC*, methyltransferase type 11 (WP_016573068.1); *clD*, NAD(P)-dependent oxidoreductase (WP_038524328, WP_037634555, WP_044380525); *clE*, γ-butyrolactone biosynthesis enzyme (WP_016573071); *clF*, γ-butyrolactone receptor protein (WP_016573072); *clG*, short chain dehydrogenase (WP_078486931),

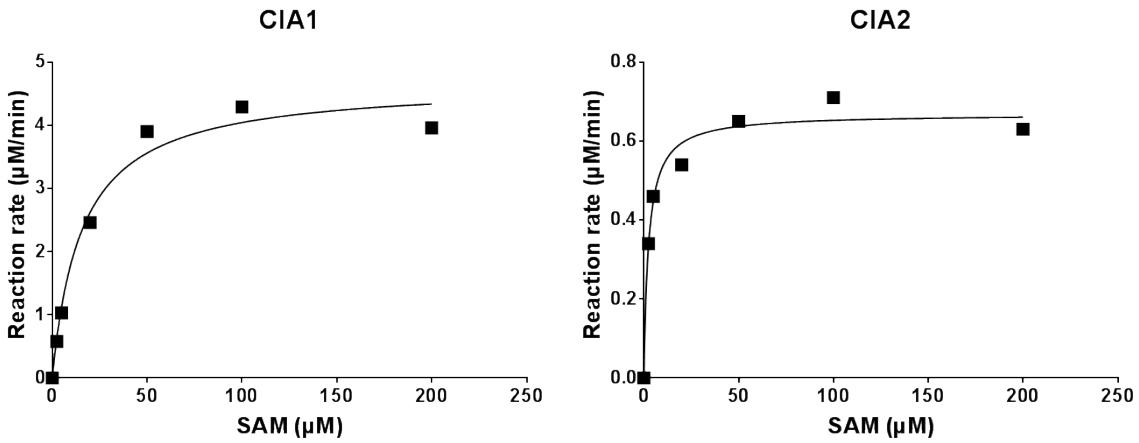

Figure S4. Kinetic assays of ClA1 and ClA2 for the conversion of SAM to 5'-ClDA. Assays contain 200 mM NaCl and various concentrations of SAM.

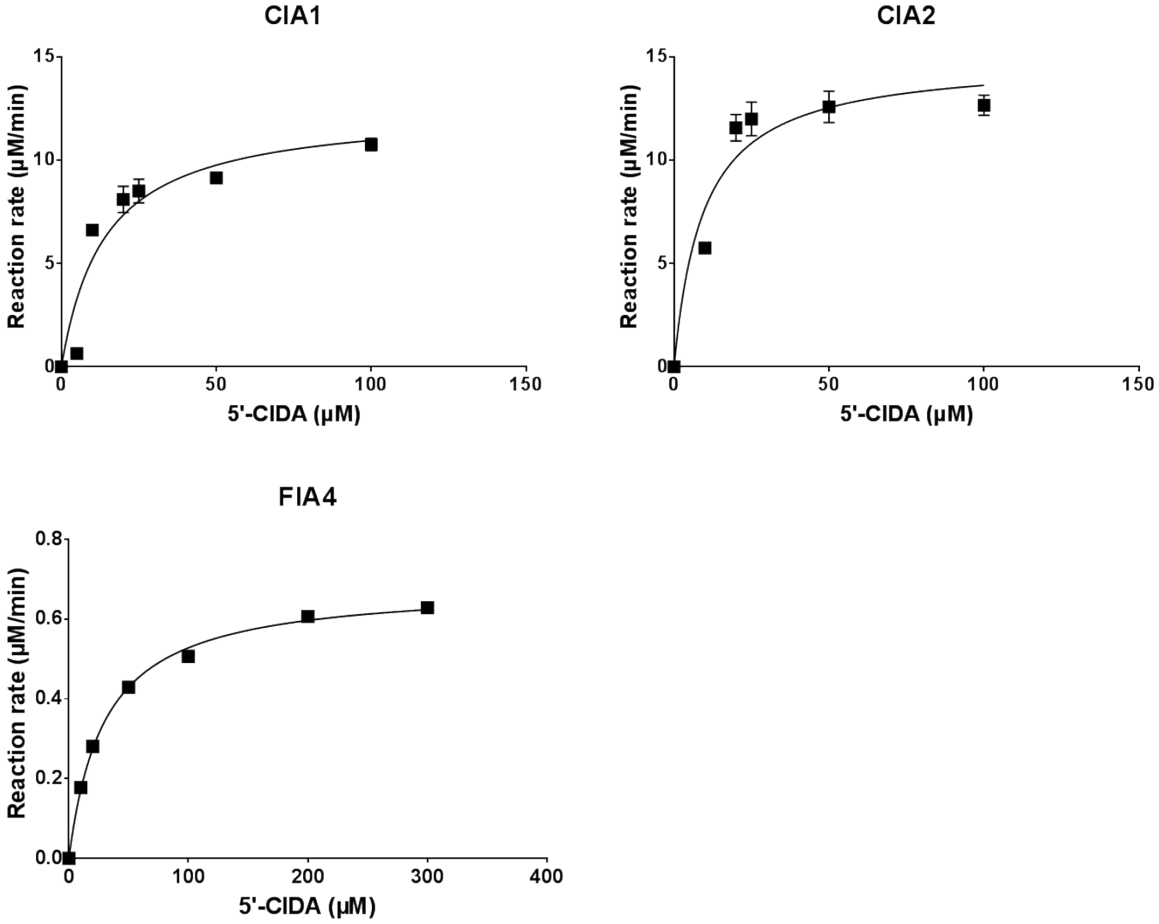

Figure S5. Kinetic assays of ClA1, ClA2 and FlA4 for the conversion of 5'-ClDA to SAM. Assays contain 20 mM L-Met and various concentrations of 5'-ClDA.

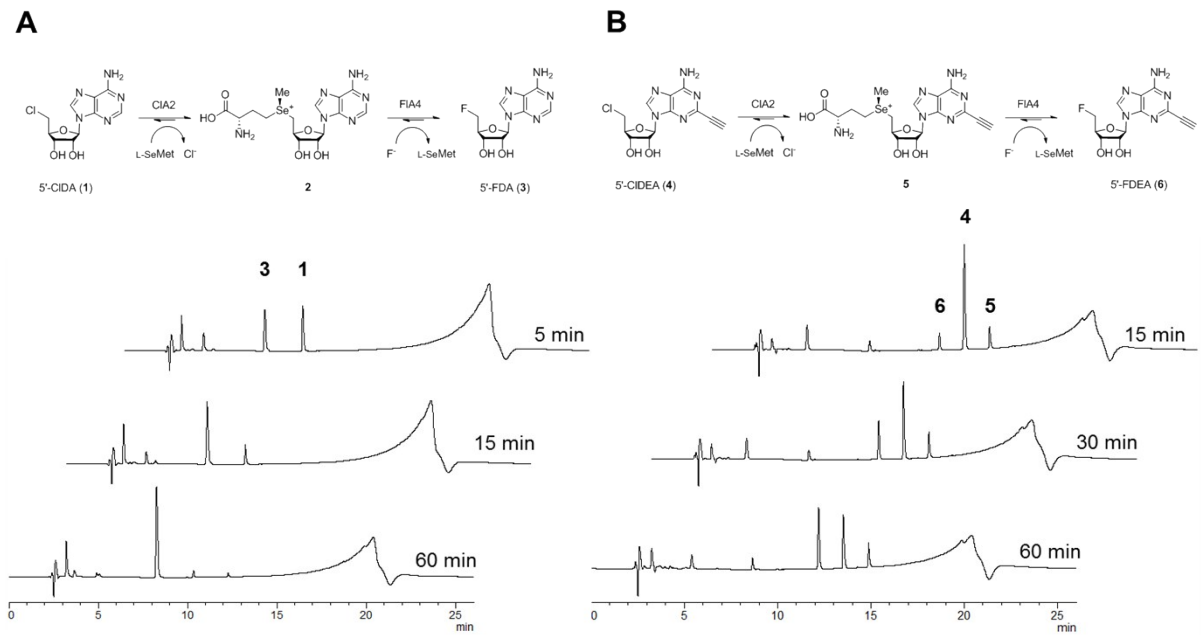

Figure S6. (A) HPLC traces (258 nm) showing the trans-halogenation of 5'-ClDA (**1**) to 5'-FDA (**3**) by FIA4 coupled to ClA2. (B) HPLC traces (268 nm) showing the trans-halogenation of 5'-ClDEA (**4**) to 5'-FDEA (**6**) by FIA4 coupled to ClA2.

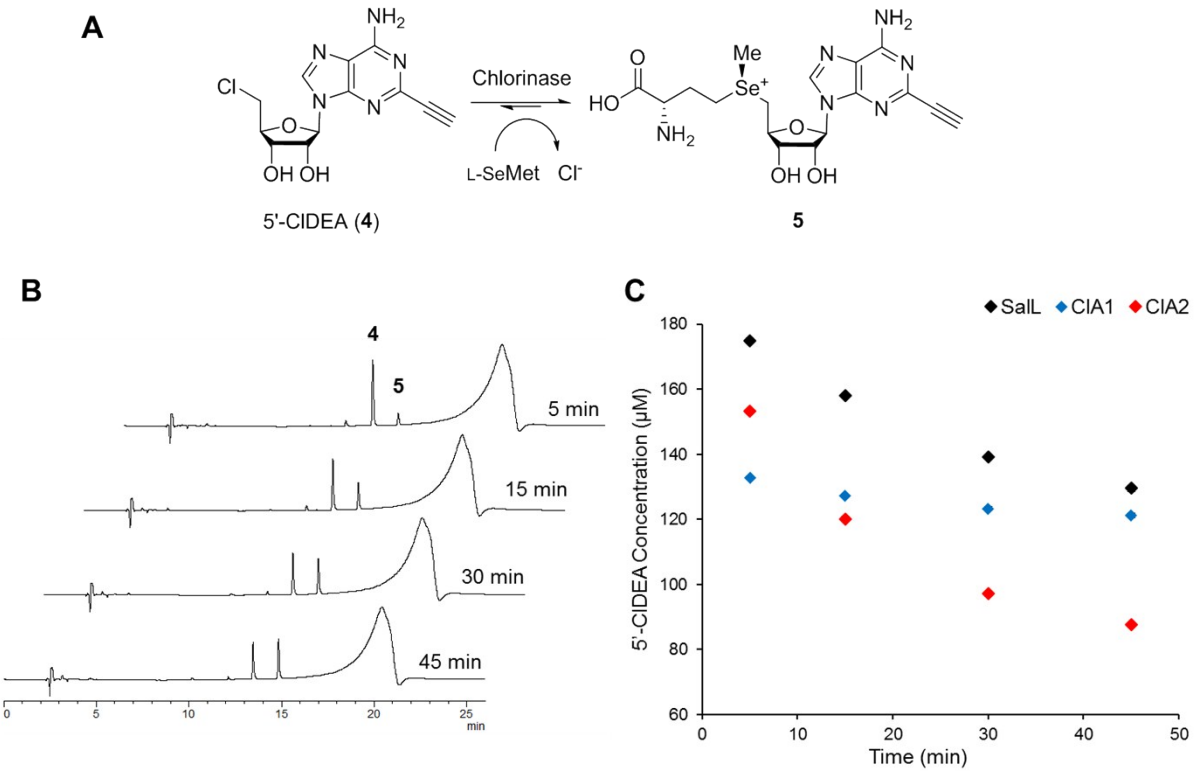

Figure S7. (A) Scheme for the conversion of 5'-ClDEA (**4**) to its SAM derivative **5** by the chlorinase. (B) HPLC traces (268 nm) showing the conversions of **4** to **5** by ClA2 at 5, 15, 30 and 45 min. (C) Comparison of the consumption rates of 5'-ClDEA by the three chlorinases (SalL, ClA1 and ClA2). Each reaction contains 30 μM chlorinase, 0.2 mM 5'-ClDEA and L-SeMet. Reactions were incubated at 37°C

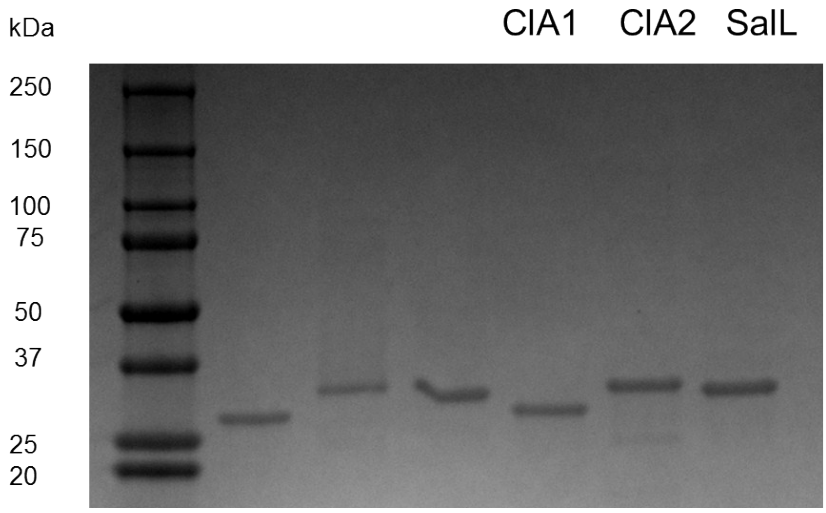

Figure S8. SDS-PAGE gel of purified ClA1, ClA2 and SalL.

Supplementary Table

Table S1. Reactions catalyzed by FlA4 as shown in Scheme 1.

Reaction	Product yield (%)
Conversion of 5'-ClDA to 5'-FDA[a]	1.04 ± 0.01 (5'-FDA)
Conversion of 5'-ClDA to SAM[b]	2.60 ± 0.13 (SAM)
Conversion of SAM to 5'-FDA[c]	55.31 ± 0.33 (5'-FDA)

All reactions were carried out in triplicate and were incubated at 47°C for 10 min. Reaction conditions: [a] 0.2 mM 5'-ClDA, 80 mM NaF, 0.1 mM L-Met, 50 μM FlA4. [b] 0.2 mM 5'-ClDA, 0.1 mM L-Met, 50 μM FlA4. [c] 0.2 mM SAM, 80 mM NaF, 50 μM FlA4.

Table S2. The newly identified SAM-dependent chlorinases.

Protein Name	Identity / Similarity to SalL[a]	Secondary Metabolite Cluster[b]	Organism
ClA1	59.6% / 73.3%	γ-Butyrolactone	*Streptomyces albulus* CCRC 11814 *S. albulus* PD-1 *S. albulus* ZPM *S. albulus* NK660 *S. ahygroscopicus* subsp. *wuyiensis* CK-15
ClA2	52.7% / 62.7%	Unknown[c]	*Umezawaea tangerina* NRRL B-24463

[a] The identity and similarity were determined using Pairwise Sequence Alignment via EMBOSS Needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/).
[b] Secondary metabolite clusters were predicted by antiSMASH 4.0.
[c] Assembly gaps in the genome sequence around *clA2* made it difficult for prediction.

Table S3. Comparative kinetic data for SalL, ClA1, and ClA2 for the chlorination reactions with 200 μM SAM and various concentrations of NaCl.

Enzyme	Cl^- K_M (mM)	k_{cat} (min^{-1})	[k_{cat}/K_M] (mM^{-1} min^{-1})
SalL	2.24	0.89 ± 0.01	0.40
ClA1	15.13	6.94 ± 0.34	0.46
ClA2	1.35	0.62 ±0.01	0.46

Protein Sequences

>SalL
MQHNLIAFLSDVGSADEAHALCKGVMYGVAPAATIVDITHDVAPFDVREGALFLADVPHSFPAHTVIC
AYVYPETGTATHTIAVRNEKGQLLVGPNNGLLSFALDASPAVECHEVLSPDVMNQPVTPTWYGKDIV
AACAAHLAAGTDLAAVGPRIDPKQIVRLPYASASEVEGGIRGEVVRIDRAFGNVWTNIPTHLIGSMLQD
GERLEVKIEALSDTVLELPFCKTFGEVDEGQPLLYLNSRGRLALGLNQSNFIEKWPVVPGDSITVSPR
VPDSNLGPVLG

>ClA1
MDKPIIAYLSDIGNHDEAHALGKGLIKTIAPGAEIVDITHQVTPFDVREGGLYLQDVPASFPANTVIAAYV
YPETGTSTRTVVVRNEKGQLLVAPNNGLLTWALKAVPAVEAWEVTSPDVMNQPVTPTWYGKDVVVA
CGAHLAAGVAPSAVGPKIDVAKLVTLPTTPAVQLGDGSVRGEVVRIDKAFGNVWTNISLDALSGGGA
LDGKTLQVTAEGLSVEIPYYATFGEVPIGEPLVYNNSRGKVALGLNQGSFLERYGVAAGDTVTIGLV

>ClA2
MEQAAPRVVAFLSDVGTHDEATGLCKGLMSRICPGVTIIDITHQVPAFDVVEGALMLEDVPEFFPEHT
VICAYVYPETGSGTPTVAVRNDKGQLLVAPDNGLLTRALDASGVAEARLVTNPAVMNHPPTPTWYG
RDVVAACAAHLAAGTPLADVGPVVDDPVRLPDVPFTRHESGLVGRVARIDRAFGNVWTNIPSAALGL
PSTPDGPVTLDATVGGERARWPWCTTFSQVATTGRLAYANSRGRLSFALNRGSLVAELGVAPDAPV
EVHLPRVPG

>FlA
MAANSTRRPIIAFMSDLGTTDDSVAQCKGLMYSICPDVTVVDVCHSMTPWDVEEGARYIVDLPRFFP
EGTVFATTTYPATGTTTRSVAVRIKQAAKGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTTVLEEH
GYLEAYEVTSPKVIPEQPEPTFYSREMVAIPSAHLAAGFPLSEVGRPLEDHEIVRFNRPAVEQDGEAL
VGVVSAIDHPFGNVWTNIHRTDLEKAGIGYGARLRLTLDGVLPFEAPLTPTFADAGEIGNIAIYLNSRG
YLSIARNAASLAYPYHLKEGMSARVEAR

>FlA1
MAANGSQRPIIAFMSDLGTTDDSVAQCKGLMHSICPGVTVVDVCHSMTPWDVEEGARYIVDLPRFFP
EGTVFATTTYPATGTTTRSVAVRIRQAAKGGARGQWAGSGDGFERADGSYIYIAPNNGLLTTVLEEH
GYIEAYEVTSTKVIPANPEPTFYSREMVAIPSAHLAAGFPLAEVGRRLDDSEIVRFHRPAVEISGEALS
GVVTAIDHPFGNIWTNIHRTDLEKAGIGQGKHLKIILDDVLPFEAPLTPTFADAGAIGNIAFYLNSRGYLS
LARNAASLAYPYNLKAGLKVRVEAR

>NobA
MTTTNGRRPIIAFMSDLGITDDSVAQCKGLMLSVCPDVTIVDICHTMQPWDVEEGARYIVDLPRLFPE
GTVFATTTYPATGTTARSVALRIAHASKGGARGQWAGSGAGFERKEGSYIYIAPNNGLLTTVIKEHGY
LEAYEVSSPEVIPEQPEPTFYSREMVALPSAHLAAGFPLEKVGRRLADDEIVRFERKDPELVADHDLV
GYVTNIDHPFGNVWTNIHRTDLEKLGVGYGTKLRITLDGVLPFELPLSPTFADAGEIGAAVAYLSSRGY
LALARNAASLAYPYNLKAGISVQVKVG

>FlA3
MPANGNPIIAFMSDLGTTDDSVAQCKGLMLSICPGVTIVDVNHSMTPWDVEEGARYIVDLPRFFPEGT
VFATTTYPATGTATRSVALRIKQAAQGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTTVIEEHGYIE
AYEVSNTKVIPAEPEPTFYSREMVAIPSAHLAAGFPLNEVGRALSDDEIVRFAKPKPSTVSGGVLSGVI
TNIDHPFGNLWTNIHRTDLEKAGIGYQTQLRLLLDGVLTFDLPLVPTFADAGQIGDPVIYINSRGYLALA
RNAAPLAYPYNLKAGLTVTVTKA

>FlA4
MSADPTQRPIIGFMSDLGTTDDSVAQCKGLMHSICPGVTVIDVCHSMTPWDVEEGARYIVDLPRFFP
EGTVFATTTYPATGTETRSVAVRIKQAAKGGARGQWAGSAGGFERAEGSYIYVAPNNGLLTTVLEEH
GYIEAYEVSSTKVIPERPEPTFYSREMVAIPAAHLAAGFPLSEVGRPLEDSEIVRYQPPQVEISGDTLT
GVVSAIDHPFGNVWTNIHRTHLEKAGIGYGKRIKIILDDVLPFEQTLVPTFADAGEIGGVAAYLNSRGYL
SLARNAASLAYPFNLKAGLKVRVETN

References

1. H. Sun, W. L. Yeo, Y. H. Lim, X. Chew, D. J. Smith, B. Xue, K. P. Chan, R. C. Robinson, E. G. Robins, H. Zhao and E. L. Ang, *Angew. Chem. Int. Ed.*, 2016, **55**, 14277-14280.
2. W. L. Yeo, X. Chew, D. J. Smith, K. P. Chan, H. Sun, H. Zhao, Y. H. Lim and E. L. Ang, *Chem. Commun.*, 2017, **53**, 2559-2562.
3. D. O'Hagan, C. Schaffrath, S. L. Cobb, J. T. G. Hamilton and C. D. Murphy, *Nature.*, 2002, **416**, 279-279.
4. C. Schaffrath, H. Deng and D. O'Hagan, *FEBS Lett.*, 2003, **547**, 111-114.
5. 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.
6. Y. Wang, Z. Deng and X. Qu, *F1000Res.*, 2014, **3**, 61.
7. S. Huang, L. Ma, M. H. Tong, Y. Yu, D. O'Hagan and H. Deng, *Org. Biomol. Chem.*, 2014, **12**, 4828-4831.
8. L. Ma, Y. Li, L. Meng, H. Deng, Y. Li, Q. Zhang and A. Diao, *RSC Adv.*, 2016, **6**, 27047-27051.