(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 November 2023 (23.11.2023)

WIPO | PCT

(10) International Publication Number
WO 2023/224560 A1

(51) International Patent Classification:
C12N 9/02 (2006.01) *C12P 13/00* (2006.01)
C12N 15/70 (2006.01)

(21) International Application Number:
PCT/SG2023/050347

(22) International Filing Date:
19 May 2023 (19.05.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10202205361W 20 May 2022 (20.05.2022) SG

(71) Applicant: **AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH** [SG/SG]; 1 Fusionopolis Way, #20-10, Connexis North Tower, Singapore 138632 (SG).

(72) Inventors: **LIM, Yee Hwee**; c/o Institute of Chemical and Engineering Sciences, 8 Biomedical Grove, Neuros, #07-01, Singapore 138665 (SG). **ANG, Ee Lui**; c/o Singapore Institute of Food and Biotechnology, Innovation, 31 Biopolis Way, Nanos, #01-02, Singapore 138669 (SG). **WONG, Fong Tian**; c/o Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, #07-06, Singapore 138673 (SG). **PEH, Guangrong**; c/o Institute of Chemical and Engineering Sciences, 8 Biomedical Grove, Neuros, #07-01, Singapore 138665 (SG). **TAY, Song Buck**; c/o Singapore Institute of Food and Biotechnology, Innovation, 31 Biopolis Way, Nanos, #01-02, Singapore 138669 (SG).

(74) Agent: **DAVIES COLLISON CAVE ASIA PTE. LTD.**; 10 Collyer Quay #07-01, Ocean Financial Centre, Singapore 049315 (SG).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title:** ENZYMES AND USES IN BIOCATALYTIC HALOGENATION OF N-HETEROARYLS THEREOF

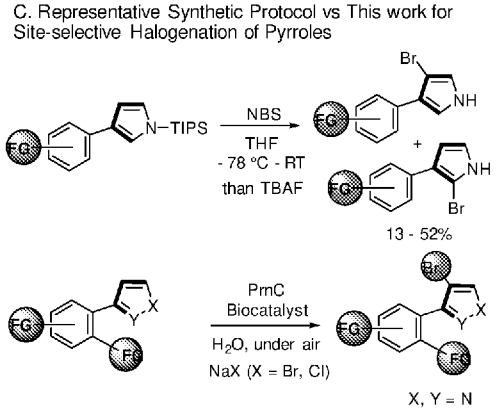


Figure 2C

(57) **Abstract:** The present disclosure concerns enzymes and their uses in biocatalytic halogenation of pyrrolic heterocycles thereof. The method of halogenating a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof comprises contacting the N-heteroaryl or derivative thereof with a halogen and a flavin-dependent halogenase under co-enzyme regeneration conditions in order for the halogen to be substituted at a 4' position on the N-heteroaryl or derivative thereof.

WO 2023/224560 A1

Enzymes and Uses in Biocatalytic Halogenation of N-Heteroaryls Thereof

Technical Field

The present invention relates, in general terms, to enzymes and their uses in biocatalytic halogenation of pyrrolic heterocycles thereof.

Background

Pyrroles are one of the privileged biological motif present in numerous natural products. Pyrrolic compounds have diverse uses in medicine and agrochemistry (antifungal, antibacterial, and anti-cancer) (Figure 1), advanced materials (organic semiconductors, solar cells), and organic catalysts. The pyrrole ring is electron-rich and, when not part of a larger conjugated system, is susceptible to air oxidation. Many pyrrole natural products bear electron-withdrawing substituents, such as carbonyl, nitro and halides. Chemical halogenation of pyrroles is challenging due to their tendency to undergo uncontrolled poly-halogenation. Furthermore, due to the differing stability of the intermediates, pyrroles preferentially undergo C2 halogenation, making C3-selective halogenation difficult. (Figure 2A).

It would be desirable to overcome or ameliorate at least one of the above-described problems.

Summary

The present invention provides a method of halogenating a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof, comprising:

a) contacting the N-heteroaryl or derivative thereof with a halogen and a flavin-dependent halogenase under co-enzyme regeneration conditions in order for the halogen to be substituted at a 4' position on the N-heteroaryl or derivative thereof;

wherein the co-enzyme regeneration conditions comprises flavin adenine dinucleotide (FAD) at about 0.01 equivalents to about 0.2 equivalents relative to a flavin-dependent halogenase concentration;

E. coli flavin reductase (Fre) at about 0.1 equivalents to about 0.5 equivalents relative

to a flavin-dependent halogenase concentration;
nicotinamide adenine dinucleotide hydrogen (NADH) at about 2 equivalents to about 250 equivalents relative to a flavin-dependent halogenase concentration;
glucose 1-dehydrogenase (GdHi) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration;
monosaccharide at about 5 equivalents to about 500 equivalents relative to a flavin-dependent halogenase concentration.

In some embodiments, the co-enzyme regeneration conditions comprises NADH at about 0.05 equivalents to about 5 equivalents relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration.

In some embodiments, the co-enzyme regeneration conditions comprises monosaccharide at about 5 equivalents to about 20 equivalents relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration.

In some embodiments, a ratio of FAD: Fre : GDH is about 1: 2.5 : 2.5.

In some embodiments, a ratio of FAD: Fre : NADH : GDH is about 1: 2.5 : 2500: 2.5.

In some embodiments, the co-enzyme regeneration conditions further comprises a halogen at about 10 equivalents to about 30 equivalents relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration.

In some embodiments, the halogen is derived from a halide, the halide selected from Cl^- or Br^-.

In some embodiments, the co-enzyme regeneration conditions further comprises a phosphate buffer or a Tris HCl buffer.

In some embodiments, the flavin-dependent halogenases is monodechloroaminopyrrolnitrin halogenase (PrnC).

In some embodiments, the flavin-dependent halogenases comprises a N-terminal 11 amino acid solubility tag.

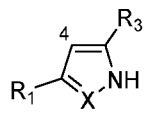
In some embodiments, the N-terminal 11 amino acid solubility tag is derived from a first 11 amino acid residues within a N-terminal N-half domain of a duplicated carbonic anhydrase (dCA) from Dunaliella species.

In some embodiments, the 11 amino acid solubility tag has an amino acid sequence of VSEPHDYNYEK.

In some embodiments, the N-heteroaryl or derivative thereof is a 5 membered N-heteroaryl or derivative thereof.

In some embodiments, the substituent on the 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof is each selected from optionally substituted aryl or optionally substituted heteroaryl.

In some embodiments, the N-heteroaryl or derivative thereof is a compound of formula (I)



(I)

wherein X is selected from CR_2 or N;

R_1 is H, optionally substituted aryl, or optionally substituted heteroaryl;

R_2 is H, optionally substituted aryl, or optionally substituted heteroaryl;

R_3 is H, optionally substituted aryl, or optionally substituted heteroaryl;

wherein at least one of R_1, R_2 and R_3 is optionally substituted aryl or optionally substituted heteroaryl; or

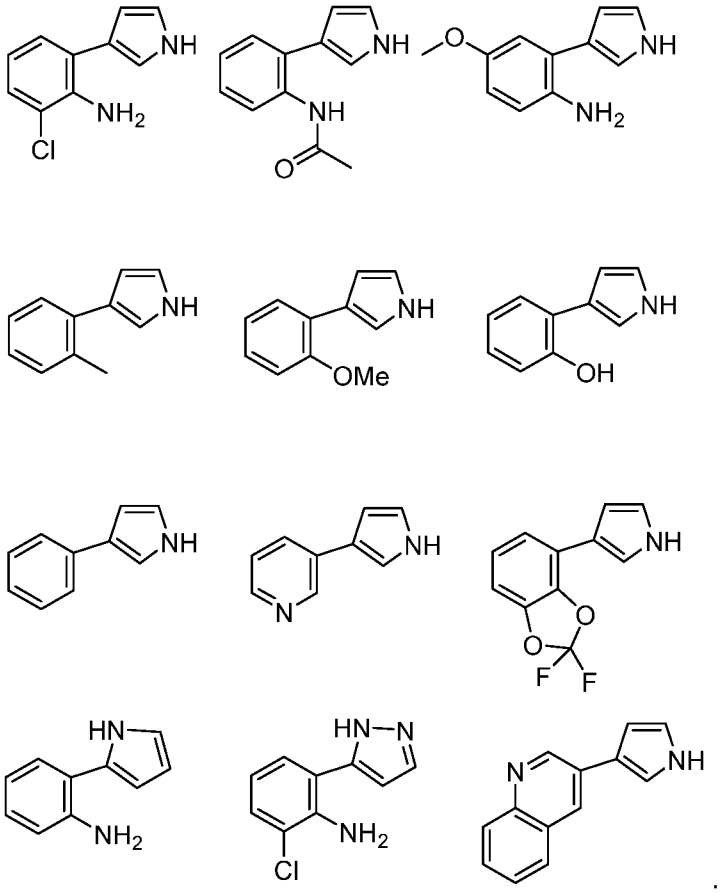
R_1 and R_2 are linked to form optionally substituted aryl, or optionally substituted heteroaryl.

In some embodiments, when X is CR_2, R_3 is H.

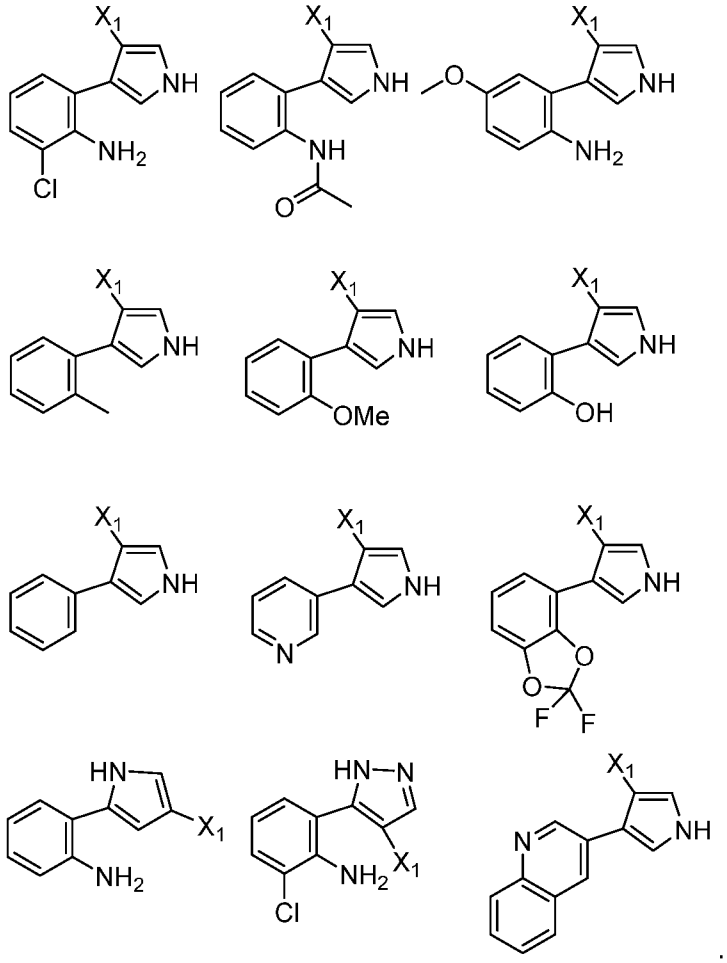
In some embodiments, when X is CR_2, R_2 is H and R_3 is H.

In some embodiments, when X is N, R_1 is H.

In some embodiments, the N-heteroaryl or derivative thereof is selected from:



In some embodiments, the halogenated N-heteroaryl or derivative thereof is (where X_1 represents halo):



In some embodiments, the method is characterised by a regioisomeric ratio of 4' substitution to 2', 3' or 5' substitution is about 3:1 to about 1.1:1.

In some embodiments, the method is characterised by a ratio of 4' monohalogenation to 2', 4' dihalogenation, 3', 4' dihalogenation, and/or 4', 5' dihalogenation of about 20:1 to about 4:1.

In some embodiments, the method is characterised by a K_{cat} of about 7 x 10^{-3} s^{-1} to about 8 x 10^{-3} s^{-1}.

In some embodiments, the method is characterised by a K_m of about 1 x 10^{-5} M to about 2 x 10^{-5} M.

In some embodiments, the method is characterised by a K_{cat}/K_m of about 4 x 10^2 s^{-1} M^{-1} to about 5 x 10^2 s^{-1} M^{-1}.

The present invention also provides a method of producing a flavin-dependent halogenase, comprising:

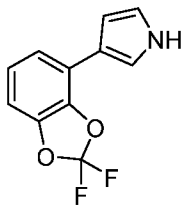
a) expressing the flavin-dependent halogenase in a cell;

wherein the flavin-dependent halogenase comprises an 11 amino acid solubility tag at a N-terminus thereof and a His6 tag at a C-terminus thereof.

In some embodiments, the method further comprises a step of purifying the flavin-dependent halogenase by metal affinity chromatography.

The present invention also provides a method of synthesising Fludioxonil, comprising:

a) contacting a compound of formula (II) or derivative thereof with a halogen and a flavin-dependent halogenase under co-enzyme regeneration conditions in order for the halogen to be substituted at a 4' position on the compound or derivative thereof;



(II)

wherein the co-enzyme regeneration conditions comprises flavin adenine dinucleotide (FAD) at about 0.01 equivalents to about 0.2 equivalents relative to a flavin-dependent halogenase concentration;

E. coli flavin reductase (Fre) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration;

nicotinamide adenine dinucleotide hydrogen (NADH) at about 2 equivalents to about 250 equivalents relative to a flavin-dependent halogenase concentration;

glucose 1-dehydrogenase (GdHi) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration;

monosaccharide at about 5 equivalents to about 500 equivalents relative to a flavin-dependent halogenase concentration;

b) contacting the 4' halogenated compound of step a) with a palladium precatalyst and a cyanide precursor.

Brief description of the drawings

Embodiments of the present invention will now be described, by way of non-limiting example, with reference to the drawings in which:

Figure 1 shows examples of pyrrolic motifs in agrochemicals and natural products.

Figure 2A-2C shows application of PrnC biocatalyst for the regioselective halogenation on the pyrrolic backbone.

Figure 3A and 3B shows preliminary optimisation results.

Figure 4 shows conversion yields of in-vitro regioselective chlorination.

Figure 5 shows a PrnC/MDA (1) complex model from the molecular modelling and molecular docking. K97, E129, and E60 are the proposed key residues.

Figure 6 shows an application of the biocatalyst in the synthesis of agricultural and pharmaceutical drug molecules.

Figure 7 shows an application of the biocatalyst in the synthesis of agricultural and pharmaceutical drug molecules. Fludioxonil is shown as an example. Without PrnC, no bromination occurred.

Figure 8 shows Michaelis-Menten plot and table of kinetic parameters for PrnC chlorination of monodechloroaminopyrrolnitrin (1).

Figure 9 shows a chart and table of MDA-Cl (2) product formation by PrnC enzyme variants; N.D.: not detectable.

Figure 10 shows the sequence alignment and the template crystal structure of halogenase PltM (PDBCODE: 6BZA) were provided by BLASTP on the NCBI server.

Figure 11 shows the native substrate MDA (1) in the binding pocket. Brown color refers to the hydrophobic residues; green color standing for the polar residues.

Figure 12 shows calibration curves of MDA (1) and MDACl (2) for analytical HPLC.

Figure 13 shows additional co-factors optimization parameters on 1 using NT11-PrnC.

Figure 14 shows determination of optimum lysate amount for PrnC lysate runs.

Detailed description

The present disclosure is predicated on the understanding that in nature, aryl halides are usually installed by flavin-dependent halogenases. Many phenol and indole halogenases are known, and several have been applied in organic synthesis. Conversely,

much fewer pyrrole halogenases are known, and their synthetic applications have not been investigated to date. Enzymatic pyrrole halogenation may offer mild reaction conditions, site-selective mono-halogenation without the need for any protective and/or directing groups, and avoids toxic halogenating reagents such as N-bromosuccinamide (NBS), iodine, and mercuric salts. In addition, the mono-halogenated site may provide a useful handle for late-stage functionalization through a plethora of metal-catalyzed coupling reactions.

Of the known pyrrole halogenases, nearly all halogenate pyrrole forms a part of a biosynthetic intermediate that is tethered to a carrier protein. Monodechloroaminopyrrolnitrin 3-halogenase (PrnC) is the only flavin-dependent halogenase reported to act on a free-standing pyrrole substrate, making it a prime candidate for further investigations for synthetic applications. PrnC is part of the prnABCD cluster responsible for the biosynthesis of pyrrolnitrin (Figure 2B). In the mechanism, the two chlorine atoms in pyrrolnitrin are introduced sequentially by the tryptophan halogenase PrnA, and the pyrrole halogenase PrnC, which is hypothesized to regioselectively chlorinate the C3 position of the pyrrolic backbone. While PrnA is well studied, there are no reports of heterologous expression of PrnC for biochemical study or *in-vitro* screening.

Without wanting to be bound by theory, the inventors believed that PrnC may be used to halogenate structurally diverse aryl or biaryl pyrroles. Towards this end, PrnC was characterized and a library of biaryl pyrroles were synthesized. The utility of the biocatalyst was also demonstrated in the synthesis of an agrochemical compound.

The present disclosure concerns a method of halogenating a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof, comprising:

a) contacting the N-heteroaryl or derivative thereof with a halogen and a flavin-dependent halogenase under co-enzyme regeneration conditions in order for the halogen to be substituted at a 4' position on the N-heteroaryl or derivative thereof;

wherein the co-enzyme regeneration conditions comprises:

i) flavin adenine dinucleotide (FAD) at about 0.01 equivalents to about 0.2 equivalents relative to a flavin-dependent halogenase concentration;

ii) *E. coli* flavin reductase (Fre) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration;

iii) nicotinamide adenine dinucleotide hydrogen (NADH) at about 2 equivalents to about 250 equivalents relative to a flavin-dependent halogenase concentration;

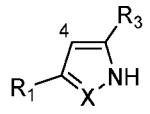
iv) glucose 1-dehydrogenase (GdHi) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration; and

v) monosaccharide at about 5 equivalents to about 500 equivalents relative to a flavin-dependent halogenase concentration.

In some embodiments, the N-heteroaryl or derivative thereof is a 5 membered N-heteroaryl or derivative thereof. Accordingly, the method concerns halogenation of a 2', 3'and/or 5' substituted 5 membered N-heteroaryl or a derivative thereof.

In some embodiments, the substituent on the 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof is each selected from optionally substituted aryl or optionally substituted heteroaryl.

In some embodiments, the N-heteroaryl or derivative thereof is a compound of formula (I)


(I)

wherein X is selected from CR_2 or N;

R_1 is H, optionally substituted aryl, or optionally substituted heteroaryl;

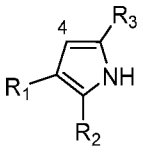
R_2 is H, optionally substituted aryl, or optionally substituted heteroaryl;

R_3 is H, optionally substituted aryl, or optionally substituted heteroaryl;

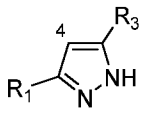
wherein at least one of R_1, R_2 and R_3 is optionally substituted aryl or optionally substituted heteroaryl; or

R_1 and R_2 are linked to form optionally substituted aryl, or optionally substituted heteroaryl.

In some embodiments, X is CR_2. Accordingly, the compound of formula (I) may be


(Ia).

In some embodiments, X is N. Accordingly, the compound of formula (I) may be



(Ib).

In some embodiments, at least one of R_1, R_2 and R_3 is optionally substituted aryl or optionally substituted heteroaryl. In some embodiments, the aryl or heteroaryl is independently 5 membered or 6 membered. In some embodiments, the aryl is phenyl. In some embodiments, the heteroaryl is selected from pyridinyl, benzodioxolyl or quinolinyl. The optional substituent may be an electron withdrawing group. The optional substituent may be substituted one or two times on the aryl or the heteroaryl. The optional substituent may be selected from halo, amino, alkyl, alkoxy, oxo, alkylacylamino, acyloxy, cycloalkyl, cycloalkenyl, or heterocyclyl.

In some embodiments, R_1 is optionally substituted aryl, or optionally substituted heteroaryl, and R_2 is H, optionally substituted aryl, or optionally substituted heteroaryl. In some embodiments, R_1 is optionally substituted aryl, or optionally substituted heteroaryl, and R_2 is H.

In some embodiments, R_2 is optionally substituted aryl, or optionally substituted heteroaryl, and R_1 is H, optionally substituted aryl, or optionally substituted heteroaryl. In some embodiments, R_2 is optionally substituted aryl, or optionally substituted heteroaryl, and R_1 is H.

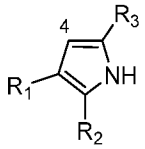
In some embodiments, when X is CR_2, R_3 is H. In some embodiments, when X is CR_2, R_2 is H and R_3 is H.

In some embodiments, R_1 and R_2 are linked to form optionally substituted aryl, optionally substituted heteroaryl, optionally substituted cycloalkyl, or optionally substituted heterocyclyl. For example, the aryl, heteroaryl, cycloalkyl or heterocyclyl may be quinolinyl, aza-2-cycloheptanonyl, or aza-2-cycloheptanone-dienyl.

In some embodiments, when X is N, R_3 is optionally substituted aryl, or optionally substituted heteroaryl, R_1 is H, optionally substituted aryl, or optionally substituted

heteroaryls. In some embodiments, when X is N, R_3 is optionally substituted aryl, or optionally substituted heteroaryl, R_1 is H. In some embodiments, when X is N, R_1 is H.

In some embodiments, the N-heteroaryl or derivative thereof is a compound of formula (Ia)



(Ia)

wherein R_1 is H, optionally substituted aryl, or optionally substituted heteroaryl;

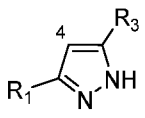
R_2 is H, optionally substituted aryl, or optionally substituted heteroaryl;

R_3 is H;

wherein at least one of R_1, and R_2 is optionally substituted aryl or optionally substituted heteroaryl; or

R_1 and R_2 are linked to form optionally substituted aryl, or optionally substituted heteroaryl.

In some embodiments, the N-heteroaryl or derivative thereof is a compound of formula (Ib)



(Ib);

wherein R_1 is H; and

R_3 is optionally substituted aryl, or optionally substituted heteroaryl.

In some embodiments, when R_1, R_2 and R_3 are independently N-heteroaryl, the N-heteroaryl is substituted at its 4' position. In some embodiments, R_1, R_2 and R_3 is not N-heteroaryl. In some embodiments, R_1, R_2 and R_3 is not 5 membered N-heteroaryl. In some embodiments, R_1, R_2 and R_3 is not pyrrolyl or pyrazolyl.

"Aryl" refers to an unsaturated aromatic carbocyclic group having a single ring (eg. phenyl) or multiple condensed rings (eg. naphthyl or anthryl), preferably having from 6 to 14 carbon atoms. Examples of aryl groups include phenyl, naphthyl and the like.

"Heteroaryl" refers to a monovalent aromatic heterocyclic group which fulfils the Hückel criteria for aromaticity (ie. contains 4n + 2 π electrons) and preferably has from 2 to 10 carbon atoms and 1 to 4 heteroatoms selected from oxygen, nitrogen, selenium, and sulfur within the ring (and includes oxides of sulfur, selenium and nitrogen). Such heteroaryl groups can have a single ring (eg. pyridyl, pyrrolyl or N-oxides thereof or furyl) or multiple condensed rings (eg. indolizinyl, benzoimidazolyl, coumarinyl, quinolinyl, isoquinolinyl or benzothienyl).

Examples of heteroaryl groups include, but are not limited to, oxazole, pyrrole, imidazole, pyrazole, pyridine, pyrazine, pyrimidine, pyridazine, indolizine, isoindole, indole, indazole, purine, isoquinoline, quinoline, phthalazine, naphthylpyridine, quinoxaline, quinazoline, cinnoline, pteridine, carbazole, carboline, phenanthridine, acridine, phenanthroline, isothiazole, phenazine, isoxazole, isothiazole, phenoxazine, phenothiazine, thiazole, thiadiazoles, oxadiazole, oxatriazole, tetrazole, thiophene, benzo[b]thiophene, triazole, imidazopyridine and the like.

"Halo" or "halogen" refers to fluoro, chloro, bromo and iodo.

"Oxo/hydroxy" refers to groups =O, HO-.

"Alkyl" refers to monovalent alkyl groups which may be straight chained or branched and preferably have from 1 to 10 carbon atoms or more preferably 1 to 6 carbon atoms. Examples of such alkyl groups include methyl, ethyl, *n*-propyl, *iso*-propyl, *n*-butyl, *iso*-butyl, *n*-hexyl, and the like.

"Alkoxy" refers to the group alkyl-O- where the alkyl group is as described above. Examples include, methoxy, ethoxy, *n*-propoxy, *iso*-propoxy, *n*-butoxy, *tert*-butoxy, *sec*-butoxy, *n*-pentoxy, *n*-hexoxy, 1,2-dimethylbutoxy, and the like.

"Amino" refers to the group –NR"R" where each R" is independently hydrogen, alkyl, cycloalkyl, aryl, heteroaryl, and heterocyclyl and where each of alkyl, cycloalkyl, aryl, heteroaryl and heterocyclyl is as described herein.

"Acylamino" refers to the group –NR"C(O)R" where each R" is independently hydrogen, alkyl, cycloalkyl, aryl, heteroaryl and heterocyclyl and where each of alkyl, cycloalkyl,

aryl, heteroaryl, and heterocyclyl are as described herein.

"Acyloxy" refers to the groups -OC(O)-alkyl, -OC(O)-aryl, -C(O)O-heteroaryl, and -C(O)O-heterocyclyl where alkyl, aryl, heteroaryl and heterocyclyl are as described herein.

"Cycloalkyl" refers to cyclic alkyl groups having a single cyclic ring or multiple condensed rings, preferably incorporating 3 to 11 carbon atoms. Such cycloalkyl groups include, by way of example, single ring structures such as cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cyclooctyl, and the like, or multiple ring structures such as adamantanyl, indanyl, 1,2,3,4-tetrahydronapthalenyl and the like.

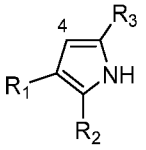
"Cycloalkenyl" refers to cyclic alkenyl groups having a single cyclic ring or multiple condensed rings, and at least one point of internal unsaturation, preferably incorporating 4 to 11 carbon atoms. Examples of suitable cycloalkenyl groups include, for instance, cyclobut-2-enyl, cyclopent-3-enyl, cyclohex-4-enyl, cyclooct-3-enyl, indenyl and the like.

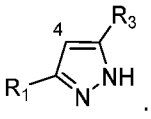
"Heterocyclyl" refers to a monovalent saturated or unsaturated group having a single ring or multiple condensed rings, preferably from 1 to 8 carbon atoms and from 1 to 4 hetero atoms selected from nitrogen, sulfur, oxygen, selenium or phosphorous within the ring. The most preferred heteroatom is nitrogen. It will be understood that where, for instance, R_2 or R' is an optionally substituted heterocyclyl which has one or more ring heteroatoms, the heterocyclyl group can be connected to the core molecule of the compounds of the present invention, through a C-C or C-heteroatom bond, in particular a C-N bond.

Examples of heterocyclyl and heteroaryl groups include, but are not limited to, oxazole, pyrrole, imidazole, pyrazole, pyridine, pyrazine, pyrimidine, pyridazine, indolizine, isoindole, indole, indazole, purine, quinolizine, isoquinoline, quinoline, phthalazine, naphthylpyridine, quinoxaline, quinazoline, cinnoline, pteridine, carbazole, carboline, phenanthridine, acridine, phenanthroline, isothiazole, phenazine, isoxazole, isothiazole, phenoxazine, phenothiazine, imidazolidine, imidazoline, piperidine, piperazine, indoline, phthalimide, 1,2,3,4-tetrahydroisoquinoline, 4,5,6,7-tetrahydrobenzo[b]thiophene,

thiazole, thiadiazoles, oxadiazole, oxatriazole, tetrazole, thiazolidine, thiophene, benzo[b]thiophene, morpholino, piperidinyl, pyrrolidine, tetrahydrofuranyl, triazole, and the like.

The halogenation occurs primarily at a 4' position of the N-heteroaryl or a derivative thereof. For clarity, the 4' position of a pyrrole and pyrazole is



and


.

It was found that other isomers may form, but in lower quantities. In some embodiments, a regioisomeric ratio of 4' substitution to 2', 3' or 5' substitution is about 3:1 to about 1.1:1. In other embodiments, the ratio is about 3:1 to about 1.2:1, about 3:1 to about 1.3:1, about 3:1 to about 1.4:1, about 3:1 to about 1.5:1, about 2.5:1 to about 1.5:1, or about 2:1 to about 1.5:1.

In some embodiments, the method is characterised by a ratio of 4' monohalogenation to dihalogenation of about 20:1 to about 4:1. The dehalogenation may occur at 2' and 4' position, 3' and 4' position, or 4' and 5' position. In other embodiments, the ratio is about 20:1 to about 5:1, about 20:1 to about 6:1, about 20:1 to about 7:1, or about 20:1 to about 10:1.

In some embodiments, the co-enzyme regeneration conditions comprises nicotinamide adenine dinucleotide hydrogen (NADH) at about 2 equivalents to about 250 equivalents relative to a flavin-dependent halogenase concentration. In other embodiments, the concentration is about 2 equivalents to about 200 equivalents, about 2 equivalents to about 150 equivalents, about 2 equivalents to about 100 equivalents, about 2 equivalents to about 80 equivalents, about 2 equivalents to about 60 equivalents, about 2 equivalents to about 50 equivalents, about 2 equivalents to about 40 equivalents, about 2 equivalents to about 30 equivalents, about 2 equivalents to about 20 equivalents, about 2 equivalents to about 10 equivalents, about 2 equivalents to about 8 equivalents, about 2 equivalents to about 6 equivalents, or about 4 equivalents to about 6 equivalents.

In some embodiments, the co-enzyme regeneration conditions comprises NADH at about 0.05 equivalents to about 5 equivalents relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration. In other embodiments, the NADH concentration is about 0.1 equivalents to about 5 equivalents, about 0.5 equivalents to about 5 equivalents, about 1 equivalents to about 5 equivalents, about 1.5 equivalents to about 5 equivalents, about 2 equivalents to about 5 equivalents, about 2.5 equivalents to about 5 equivalents, about 3 equivalents to about 5 equivalents, or about 4 equivalents to about 5 equivalents.

In some embodiments, the co-enzyme regeneration conditions comprises monosaccharide at about 5 equivalents to about 500 equivalents relative to a flavin-dependent halogenase concentration. In other embodiments, the concentration is about 5 equivalents to about 450 equivalents, about 5 equivalents to about 400 equivalents, about 5 equivalents to about 350 equivalents, about 5 equivalents to about 300 equivalents, about 5 equivalents to about 250 equivalents, about 5 equivalents to about 200 equivalents, about 5 equivalents to about 150 equivalents, about 5 equivalents to about 100 equivalents, about 5 equivalents to about 80 equivalents, about 5 equivalents to about 60 equivalents, about 5 equivalents to about 50 equivalents, about 5 equivalents to about 40 equivalents, about 5 equivalents to about 30 equivalents, about 5 equivalents to about 20 equivalents, or about 5 equivalents to about 10 equivalents.

In some embodiments, the co-enzyme regeneration conditions comprises monosaccharide at about 5 equivalents to about 20 equivalents relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration. The monosaccharide may be glucose (dextrose), fructose (levulose), or galactose. In some embodiments, the monosaccharide concentration is about 8 equivalents to about 20 equivalents, about 8 equivalents to about 18 equivalents, about 8 equivalents to about 16 equivalents, about 8 equivalents to about 14 equivalents, about 8 equivalents to about 12 equivalents, or about 10 equivalents to about 12 equivalents.

In some embodiments, the co-enzyme regeneration conditions comprises flavin adenine dinucleotide (FAD) at about 0.001 equivalents to about 0.2 equivalents relative to a flavin-dependent halogenase concentration. In other embodiments, the concentration is about 0.02 equivalents to about 0.2 equivalents, about 0.04 equivalents to about 0.2 equivalents, about 0.06 equivalents to about 0.2 equivalents, about 0.08 equivalents to

about 0.2 equivalents, about 0.1 equivalents to about 0.2 equivalents, about 0.12 equivalents to about 0.2 equivalents, about 0.14 equivalents to about 0.2 equivalents, about 0.16 equivalents to about 0.2 equivalents, or about 0.18 equivalents to about 0.2 equivalents. In other embodiments, the concentration is about 0.001 equivalents to about 0.18 equivalents, about 0.001 equivalents to about 0.16 equivalents, about 0.001 equivalents to about 0.14 equivalents, about 0.001 equivalents to about 0.12 equivalents, about 0.001 equivalents to about 0.1 equivalents, about 0.001 equivalents to about 0.05 equivalents, about 0.001 equivalents to about 0.01 equivalents, about 0.001 equivalents to about 0.008 equivalents, about 0.001 equivalents to about 0.006 equivalents, about 0.001 equivalents to about 0.005 equivalents, about 0.001 equivalents to about 0.004 equivalents, or about 0.001 equivalents to about 0.003 equivalents.

In some embodiments, the co-enzyme regeneration conditions comprises flavin adenine dinucleotide (FAD) at about 0.1 mol% to about 1 mol% relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration. In other embodiments, the concentration is about 0.1 mol% to about 0.8 mol%, about 0.1 mol% to about 0.6 mol%, about 0.1 mol% to about 0.5 mol%, about 0.1 mol% to about 0.4 mol%, or about 0.1 mol% to about 0.3 mol%. In some embodiments, the concentration is about 0.2 mol%.

In some embodiments, the co-enzyme regeneration conditions comprises *E. coli* flavin reductase (Fre) at about 0.001 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration. In other embodiments, the concentration is about 0.2 equivalents to about 0.5 equivalents, about 0.3 equivalents to about 0.5 equivalents, or about 0.4 equivalents to about 0.5 equivalents. In other embodiments, the concentration is about 0.001 equivalents to about 0.1 equivalents, about 0.001 equivalents to about 0.05 equivalents, about 0.001 equivalents to about 0.01 equivalents, about 0.001 equivalents to about 0.008 equivalents, about 0.004 equivalents to about 0.008 equivalents, about 0.004 equivalents to about 0.006 equivalents. In other embodiments, the concentration is about 0.005 equivalents.

In some embodiments, the co-enzyme regeneration conditions comprises *E. coli* flavin reductase (Fre) at about 0.1 mol% to about 2 mol% relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration. In other embodiments,

the concentration is about 0.2 mol% to about 2 mol%, about 0.4 mol% to about 2 mol%, about 0.5 mol% to about 2 mol%, about 0.5 mol% to about 1.8 mol%, about 0.5 mol% to about 1.6 mol%, about 0.5 mol% to about 1.4 mol%, about 0.5 mol% to about 1.2 mol%, or about 0.5 mol% to about 1 mol%. In other embodiments, the concentration is about 0.5 mol%.

The method may comprise a co-factor regeneration system using glucose dehydrogenase (GdHi) to regulate a more consistent supply of NADH over time, enabling high conversion as compared to system without the use of GDH.

In some embodiments, the co-enzyme regeneration conditions comprises glucose 1-dehydrogenase (GdHi) at about 0.001 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration. In other embodiments, the concentration is about 0.2 equivalents to about 0.5 equivalents, about 0.3 equivalents to about 0.5 equivalents, or about 0.4 equivalents to about 0.5 equivalents. In other embodiments, the concentration is about 0.001 equivalents to about 0.1 equivalents, about 0.001 equivalents to about 0.05 equivalents, about 0.001 equivalents to about 0.01 equivalents, about 0.001 equivalents to about 0.008 equivalents, about 0.004 equivalents to about 0.008 equivalents, about 0.004 equivalents to about 0.006 equivalents. In other embodiments, the concentration is about 0.005 equivalents.

In some embodiments, the co-enzyme regeneration conditions comprises glucose dehydrogenase (GdHi) at about 0.1 mol% to about 2 mol% relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration. In other embodiments, the concentration is about 0.2 mol% to about 2 mol%, about 0.4 mol% to about 2 mol%, about 0.5 mol% to about 2 mol%, about 0.5 mol% to about 1.8 mol%, about 0.5 mol% to about 1.6 mol%, about 0.5 mol% to about 1.4 mol%, about 0.5 mol% to about 1.2 mol%, or about 0.5 mol% to about 1 mol%. In other embodiments, the concentration is about 0.5 mol%.

In some embodiments, a ratio of flavin adenine dinucleotide (FAD): *E. coli* flavin reductase (Fre) : glucose dehydrogenase (GDH) is about 1: 2.5 : 2.5.

In some embodiments, a ratio of FAD: Fre : NADH : GDH is about 1: 2.5 : 2500: 2.5.

In some embodiments, the co-enzyme regeneration conditions further comprises a halogen at about 10 equivalents to about 30 equivalents relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration. In other embodiments, the concentration is about 10 equivalents to about 25 equivalents, about 10 equivalents to about 20 equivalents, about 10 equivalents to about 15 equivalents, or about 10 equivalents to about 12 equivalents.

In some embodiments, the halogen is derived from a halide, the halide selected from Cl^- or Br^-. For example an inorganic halide salt. The inorganic halide salt may be $MgBr_2$ or $MgCl_2$. Sodium and potassium salts may also be used.

In some embodiments, the co-enzyme regeneration conditions further comprises a buffer. The buffer may be phosphate buffer or Tris HCl buffer. The concentration of the buffer can be from about 10mM to about 50mM.

In some embodiments, flavin-dependent halogenases (or PrnC biocatalyst) is at about 2 mol% to about 10 mol% relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration. In other embodiments, the concentration is about 2 mol% to about 8 mol%, about 2 mol% to about 6 mol%, or about 2 mol% to about 4 mol%.

In some embodiments, the flavin-dependent halogenases is monodechloroaminopyrrolnitrin halogenase (PrnC). In other embodiments, the enzyme is selected from halB and tryptophan 7-halogenase (RebH). halB is a halogenase gene isolated from a cosmid library of the pentachloropseudilin producer Actinoplanes sp. ATCC 33002.

In order to further improve the yield of the pyrrole product, the inventors also explored the overexpression of the Flavin-dependent halogenase gene. Halogenating activity can only be detected after the halogenase gene is highly overexpressed and soluble halogenase is produced. Initial attempts at heterologous expression of PrnC yielded <1 mg/L as the majority of the protein were insoluble. After subsequent optimization of the expression construct, we discovered that addition of a N-terminal 11 amino acid solubility tag (NT11), together with a C-terminal His_6-tag (6His), allowed for the

production of sufficient quantities of soluble functional PrnC, and purification by metal affinity chromatography (~9 mg/L culture).

Specifically, engineering of a NT11 soluble tag into the PrnC gene enabled the improvement in heterologous expression yield of the PrnC in E. coli by at least 18-fold, whilst keeping the protein functional. It is believed that the tag allowed the protein to fold better and thus perform more consistently.

Accordingly, in some embodiments, the flavin-dependent halogenases comprises a N-terminal 11 amino acid solubility tag. In some embodiments, the flavin-dependent halogenases comprises a N-terminal 11 amino acid solubility tag and a C-terminal His6-tag. The His6 tag may be used for affinity purification.

In some embodiments, the N-terminal 11 amino acid solubility tag is derived from a first 11 amino acid residues within a N-terminal N-half domain of a duplicated carbonic anhydrase (dCA) from Dunaliella species. For example, the N-terminal 11 amino acid solubility tag may be as described in Nguyen, Thi Khoa My, et al. Applied microbiology and biotechnology 103.5 (2019): 2205-2216), the reference of which is incorporated herein.

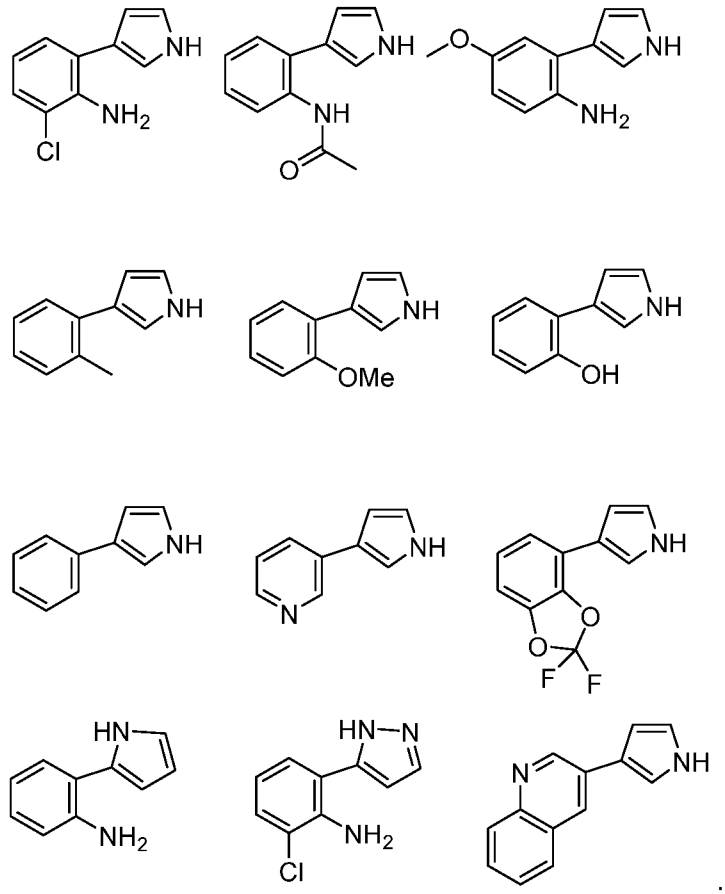
In some embodiments, the 11 amino acid solubility tag has an amino acid sequence of VSEPHDYNYEK. In some embodiments, the solubility tag is selected from Human influenza hemagglutinin (HA) tag, Small Ubiquitin-like Modifier (SUMO), maltose-binding protein (MBP) and Glutathione-S-transferase (GST) tag. The HA-tag is derived from the HA-molecule corresponding to amino acids 98-106.

In some embodiments, the method was performed for at least 4 h. In other embodiments, the method was performed for at least 6 h, 8 h, 10 h, 12 h, 14 h, 16 h or 18 h.

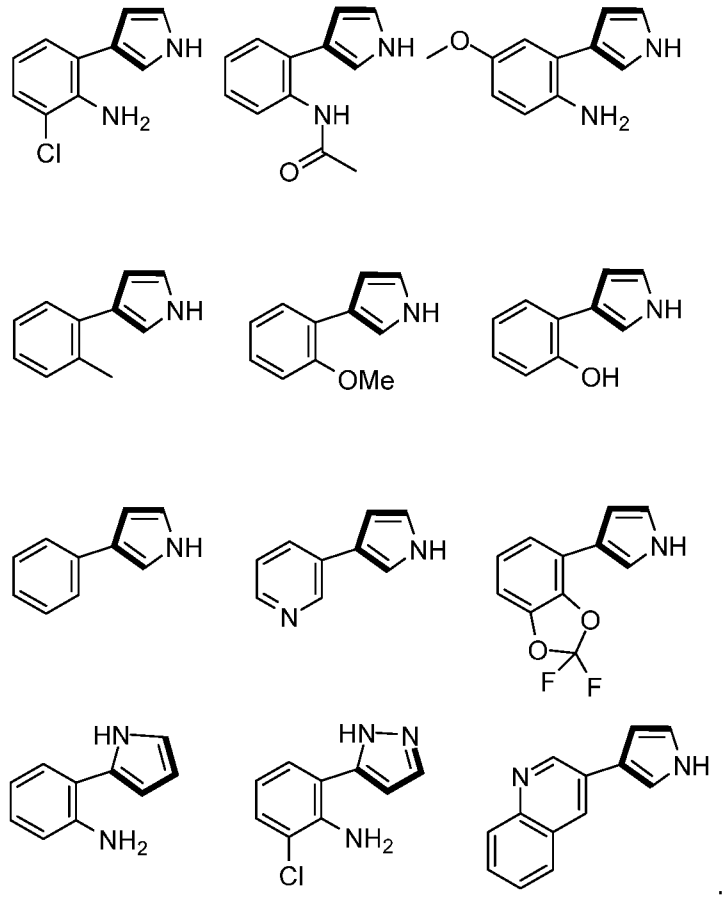
In some embodiments, the method was performed at a temperature of about 25 °C to about 45 °C. In other embodiments, the temperature is about 30 °C.

In some embodiments, method was performed under constant mixing. The mixing may be by orbital shaking at about 300 rpm.

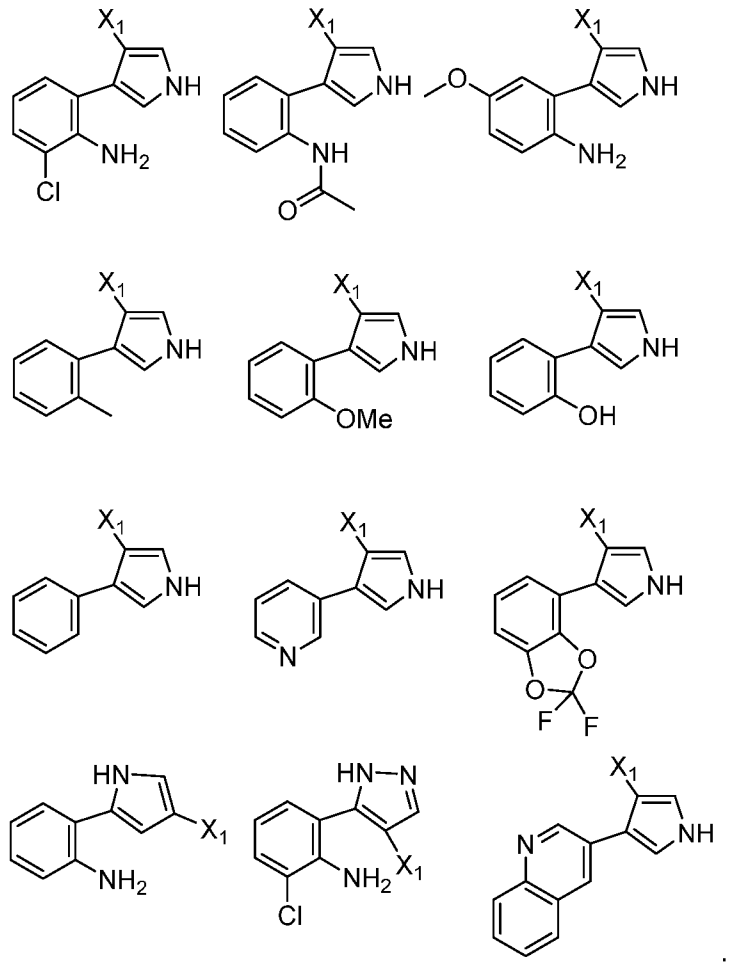
In some embodiments, the compound of Formula (I) is selected from:



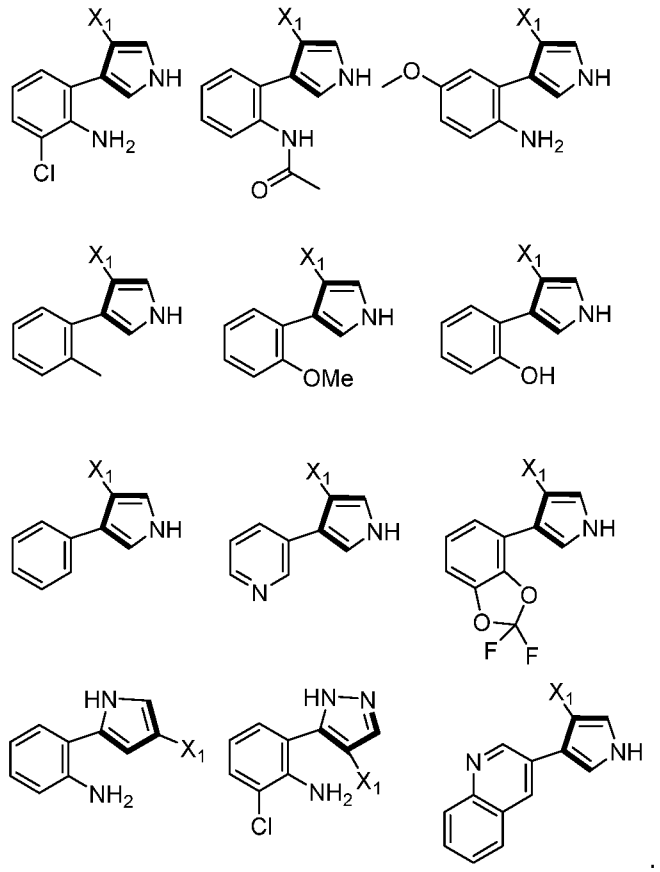
In some embodiments, the compound of Formula (I) is selected from:



In some embodiments, the halogenated compound of Formula (I) is (where X_1 represents halo):



In some embodiments, the halogenated compound of Formula (I) is (where X_1 represents halo):



In some embodiments, the method further comprises a step of purifying the halogenated compound of Formula (I).

In some embodiments, the method is characterised by a K_{cat} of about 7 x 10^{-3} s^{-1} to about 8 x 10^{-3} s^{-1}.

In some embodiments, the method is characterised by a K_m of about 1 x 10^{-5} M to about 2 x 10^{-5} M.

In some embodiments, the method is characterised by a K_{cat}/K_m of about 4 x 10^{2} s^{-1} M^{-1} to about 5 x 10^{2} s^{-1} M^{-1}.

The present disclosure also concerns a flavin-dependent halogenase and its clonal construct thereof, comprising:

a) a 11 amino acid solubility tag at a N-terminus of the flavin-dependent halogenase; and

b) a His6 tag at a C-terminus of the flavin-dependent halogenase.

The present disclosure also concerns a method of producing a flavin-dependent halogenase, comprising:

a) tagging a 11 amino acid solubility tag at a N-terminus of the flavin-dependent halogenase; and

b) tagging a His6 tag at a C-terminus of the flavin-dependent halogenase.

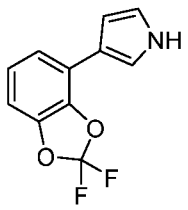
In some embodiments, the flavin-dependent halogenase is expressed in a cell. The cell may be a bacterium, such as E. coli.

Accordingly, the method comprises expressing the flavin-dependent halogenase in a cell, wherein the flavin-dependent halogenase comprises an 11 amino acid solubility tag at a N-terminus thereof and a His6 tag at a C-terminus thereof.

In some embodiments, the method further comprises a step of purifying the flavin-dependent halogenase by metal affinity chromatography. In some embodiments, the method further comprises a step of purifying the flavin-dependent halogenase from the cell. The purification step may involve binding the enzyme to a resin and eluting the enzyme from the resin.

The present disclosure also concerns a method of synthesising Fludioxonil, comprising:

a) contacting a compound of formula (II) or derivative thereof with a halogen and a flavin-dependent halogenase under co-enzyme regeneration conditions in order for the halogen to be substituted at a 4' position on the compound or derivative thereof;



(II)

wherein the co-enzyme regeneration conditions comprises flavin adenine dinucleotide (FAD) at about 0.01 equivalents to about 0.2 equivalents relative to a flavin-dependent halogenase concentration;

E. coli flavin reductase (Fre) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration;

nicotinamide adenine dinucleotide hydrogen (NADH) at about 2 equivalents to about 250 equivalents relative to a flavin-dependent halogenase concentration;

glucose 1-dehydrogenase (GdHi) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration;

monosaccharide at about 5 equivalents to about 500 equivalents relative to a flavin-dependent halogenase concentration;

b) contacting the 4' halogenated compound of step a) with a palladium precatalyst and a cyanide precursor.

Examples

Halogenase cloning, expression and purification. NT11[1]-PrnC-6His construct in pET-28a(+) was ordered from Twist Biosciences as a clonal construct and transformed into *T7 Express E. coli* (NEB). The resulting strain was cultured in 1 L LB media at 37 °C. When OD_{600} reached 0.4, 0.1 mM IPTG was used to induce for overnight expression at 16 °C. After expression, the cultures were centrifuged at 10, 000 *g* for 10 minutes at 4°C. The resulting pellets were resuspended in 20 mL of 100mM sodium phosphate pH 7, 10 mM imidazole, 150 mM sodium chloride before sonication. After sonication, the resulting lysate was then centrifuged at 19,000 *g* for 1 hour at 4°C. The supernatant was incubated with Ni-NTA agarose for 1 hour at 4°C. The resin was washed with 20 mL 100mM sodium phosphate pH 7-7.4, 80 mM imidazole, 150 mM sodium chloride, 10% glycerol and the bound protein was eluted with 5 mL of 100mM sodium phosphate pH 7-7.4, 150-500 mM imidazole, 50 mM sodium chloride, 10% glycerol. The elution was buffer exchanged and concentrated with 50 mM sodium phosphate pH 7-7.4, 10% glycerol. The final yields of NT11-PrnC were an estimated 9 mg/L culture. In comparison, PrnC alone (or with various placements of his-tags) were expressed at yields of <0.5mg/L culture and give inconsistent assay results.

Glucose dehydrogenase (GdHi or GDH). Purchased from Sigma-aldrich with activity units ≥200 U/mg.

Method for Fre cloning, expression and purification: The nucleic acid sequence that encodes for Flavin reductase Fre was purchased as a gBlock from Integrated DNA Technologies. Fre sequence was cloned into pET-28a(+) vector via NEBuilder® HiFi DNA

Assembly method and transformed into *E. coli Acella* (EdgeBio). *E. Coli* strain expressing Fre was cultured in 1 L of LB Kan^{50} media at 37 °C. At OD_{600} 0.4-0.6, 0.1 mM IPTG was used to induce protein expression at 16 °C over 18 h. Cell culture was harvested by centrifugation at 4000 rcf for 10 min at 4 °C. After media was decanted, cell pellet was resuspended in 30 ml of 50 mM tris pH 7.4, 300 mM sodium chloride, 10 mM imidazole and lysed by cell disruption. Cell lysate was centrifuged at 33,600 rcf for 45 min at 4 °C to differentiate supernatant from insoluble debris. Fre proteins from lysate supernatant were purified using immobilized metal affinity chromatography via TALON resins interaction with N-terminus His-tag Fre. After lysate supernatant was applied, 10 ml of 50 mM tris pH 7.4, 300 mM sodium chloride, 10 mM imidazole was used to wash the resins. Fre proteins were eluted from the resins using 5 ml of 50 mM tris pH 7.4, 300 mM sodium chloride, 200 mM imidazole. Eluted samples were buffer exchanged and concentrated with 50 mM tris pH 7.4, 100 mM NaCl, 10% glycerol

Method for PrnC mutants construction in mutagenesis studies. Single-site mutations of the *prnC* gene were constructed via overlap extension PCR, using PrimeSTAR Max DNA Polymerase, on the pET-28a(+) NT11-PrnC-6His plasmid template. The mutagenesis primers used are as shown:

E60A forward primer 5'-CGAAAGTTCCATCCCGGCGACTTCGTTGATGAATC-3' (SEQ ID NO: 1);

E60A reverse primer 5'-GATTCATCAACGAAGTCGCCGGGATGGAACTTTCG-3' (SEQ ID NO: 2);

K97A forward primer 5'-CATCGTCAACCGGAATTGCGCGTAATTTCGGCTTTG-3' (SEQ ID NO: 3);

K97A reverse primer 5'-CAAAGCCGAAATTACGCGCAATTCCGGTTGACGATG-3' (SEQ ID NO: 4);

E129A forward primer 5'-GCTTCCCTGGGGACCTGCGTCACATTATTATCGTC-3' (SEQ ID NO: 5);

E129A reverse primer 5'-GACGATAATAATGTGACGCAGGTCCCCAGGGAAGC-3' (SEQ ID NO: 6).

Standard flanking primers for *prnC* gene were used for amplification of the respective PCR segments in each mutant construct:

prnC forward primer 5'-GGAGATATACCATGGTAAGTGAACCCCACGACTATAATTATG-3' (SEQ ID NO: 7);

prnC reverse primer 5'-GTGGTGGTGCTCGAGTTTTTTCAGCGCTAATCCAATACGC-3' (SEQ ID NO: 8).

PCR segments and linearized plasmids were ligated using NEBuilder Hifi DNA assembly protocol and the mutations were confirmed by DNA sequencing.

Sequences – NT11-PrnC - nt

ATGGTAAGTGAACCCCACGACTATAATTATGAGAAAGCTAGCGCGTCTATGACACAAAAATCT
CCCGCCAACGAGCACGATTCAAATCACTTTGATGTCATTATTCTGGGATCGGGCATGTCAGGT
ACGCAGATGGGCGCTATTTTGGCTAAACAACAATTTCGTGTGTTAATTATTGAAGAGTCCAGC
CACCCACGTTTTACCATTGGCGAAAGTTCCATCCCGGAAACTTCGTTGATGAATCGTATCATCG
CAGACCGCTATGGTATCCCAGAACTGGACCATATTACCTCCTTTTATAGCACACAACGTTACGT
CGCATCGTCAACCGGAATTAAACGTAATTTCGGCTTTGTATTTCACAAACCAGGTCAAGAGCA
CGATCCTAAGGAATTCACACAGTGTGTGATTCCGGAGCTTCCCTGGGGACCTGAATCACATTA
TTATCGTCAGGACGTAGATGCATATTTATTGCAGGCGGCGATCAAGTACGGGTGTAAAGTGCA
CCAAAAGACAACTGTGACGGAGTATCATGCTGATAAGGATGGCGTAGCGGTCACAACAGCTC
AAGGGGAACGCTTCACAGGTCGCTACATGATCGATTGCGGTGGACCCCGCGCGCCCTTGGCT
ACTAAATTCAAGTTACGTGAAGAGCCATGCCGTTTTAAAACGCACAGCCGTTCTCTGTATACTC
ACATGTTAGGGGTCAAGCCATTCGACGACATTTTTAAAGTAAAGGGGCAGCGCTGGCGTTGG
CATGAAGGAACCCTTCATCACATGTTCGAAGGAGGATGGCTGTGGGTGATTCCCTTTAATAAC
CATCCCCGTAGCACAAATAACTTAGTCTCGGTTGGTTTGCAACTTGACCCACGTGTCTACCCC
AAGACCGACATCTCCGCTCAACAAGAATTCGACGAGTTCCTTGCCCGTTTTCCGTCAATCGGA
GCCCAGTTTCGCGACGCCGTTCCAGTTCGCGATTGGGTGAAAACTGATCGCCTTCAATTTTCT
AGTAATGCATGCGTCGGAGACCGTTACTGTTTAATGTTACACGCGAACGGATTTATTGATCCG
TTGTTTTCGCGTGGGCTTGAGAATACTGCGGTCACGATCCACGCCTTAGCTGCGCGTCTTATC
AAGGCCCTGCGCGACGATGACTTCTCTCCAGAACGCTTCGAGTATATCGAGCGTTTGCAACAG
AAGTTGTTGGACCACAATGACGACTTCGTGTCATGTTGCTACACAGCTTTTTCGGATTTTCGTC
TTTGGGATGCCTTTCACCGCCTTTGGGCCGTGGGGACTATCCTTGGACAATTTCGTCTTGTGC
AGGCCCACGCACGCTTTCGCGCATCGCGTAACGAGGGTGATCTGGATCATTTAGATAACGAC
CCACCCTATTTGGGGTATCTGTGCGCTGATATGGAAGAGTACTATCAGCTTTTCAATGACGCC
AAGGCTGAGGTGGAAGCAGTATCAGCGGGACGTAAGCCTGCCGACGAGGCTGCCGCGCGCA
TTCATGCCTTGATCGACGAACGTGACTTCGCCAAACCTATGTTCGGGTTCGGTTACTGCATTA
CAGGAGATAAACCACAGTTGAATAATTCCAAGTACTCTTTATTGCCTGCGATGCGCTTGATGTA
CTGGACGCAGACCCGTGCCCCAGCGGAAGTCAAGAAGTATTTCGATTATAACCCTATGTTTGC
GTTGCTTAAGGCATATATTACTACGCGTATTGGATTAGCGCTGAAAAAACTCGAGCACCACCA
CCACCACCACTGA (SEQ ID NO: 9)

Sequences – NT11-PrnC - aa

MVSEPHDYNYEKASASMTQKSPANEHDSNHFDVIILGSGMSGTQMGAILAKQQFRVLIIEESSH
PRFTIGESSIPETSLMNRIIADRYGIPELDHITSFYSTQRYVASSTGIKRNFGFVFHKPGQEHDPKE
FTQCVIPELPWGPESHYYRQDVDAYLLQAAIKYGCKVHQKTTVTEYHADKDGVAVTTAQGERFT
GRYMIDCGGPRAPLATKFKLREEPCRFKTHSRSLYTHMLGVKPFDDIFKVKGQRWRWHEGTLHH
MFEGGWLWVIPFNNHPRSTNNLVSVGLQLDPRVYPKTDISAQQEFDEFLARFPSIGAQFRDAVPV
RDWVKTDRLQFSSNACVGDRYCLMLHANGFIDPLFSRGLENTAVTIHALAARLIKALRDDDFSPE
RFEYIERLQQKLLDHNDDFVSCCYTAFSDFRLWDAFHRLWAVGTILGQFRLVQAHARFRASRNE
GDLDHLDNDPPYLGYLCADMEEYYQLFNDAKAEVEAVSAGRKPADEAAARIHALIDERDFAKPMF
GFGYCITGDKPQLNNSKYSLLPAMRLMYWTQTRAPAEVKKYFDYNPMFALLKAYITTRIGLALKKL
EHHHHHH (SEQ ID NO: 10)

Homology modelling and substrate docking of NT11-PrnC. Homology modelling was performed in Modeller v10 program and generated 100,000 homology models based on the template crystal structure of halogenase PltM (PDB code: 6BZA) whose sequence was aligned with that of PrnC (Figure 10) with the sequence identity of 36%. Subsequently, all the models were subjected to the backbone, sidechain and loop optimization respectively, and then the top 50 optimized models with the lowest DOPE scores were used for the subsequent molecular docking with the native substrate (**1**) by using GOLD v2018 program with the optimal docking parameters (the binding pocket is defined by the ligand copied from the crystal structure of 6BZA with the spherical radius of 8.0 Angstrom; scoring function is GoldScore; population size is 500; the number of operations is 500000; number of island is 10; crossover frequency is 95%; mutation frequency is 95%; migration frequency is 20%; the number of output docking solutions is 3), which afforded 150 docking solutions in total.

In order to achieve an optimal model for the PrnC/**1** complex, 150 docking solutions were subjected to the further inspection based on the following two criteria: (1) there must be a nearby lysine residue stretching towards the pyrrole ring in **1**, because a lysine is required for the catalysis; (2) the hydrogen on the pyrrolic nitrogen of **1** should form a hydrogen bond with the hydrogen acceptor of a residue to stabilize the intermediate during the catalytic reaction. After the manual and visual examination, a reasonable model of PrnC/**1** complex**,** which satisfied both criteria above, was harvested and shown in Figure 2 and **Figure 11**.

To determine which residues are important to the catalytic reaction of PrnC enzyme, we conducted a systematic computational modeling of the enzyme and constructed a three-dimensional model for the PrnC/substrate complex of PrnC (Figure 8). The PrnC/substrate model (Figure 5) indicates that the lysine residue K97 is close to **1** and may result in the initiation of catalysis; E129 forms a salt bridge with K97 to stabilize its sidechain conformation and E60 may form a hydrogen bond with **1**. To further validate the importance of our computationally proposed key residues, an experimental mutagenesis study was conducted, and its results are shown in **Figure 9**. The mutations K97 and E60 completely abolished enzyme activity, indicating that K97 is a key catalytic residue with E60 playing a critical role in substrate stabilization by H-bonding. As a consequence of the E129A mutation the enzyme activity was significantly reduced, revealing a preference for a spatial orientation of the K97 residue. It appears that the catalytic Glu residue E129 plays a key role in improving HOX's electrophilicity by interacting with K97 (Figure 9) or serving as a general base to deprotonate a Wheland-type pyrrolic intermediate. These mutagenesis data show that our concept is plausible from this point of view.

Analytical Scale Biotransformations:

NT-11 PrnC-Catalyzed halogenation. In a solution containing the pyrrolic derivative starting material (0.5 mM), $MgCl_2$/ $MgBr_2$ (10 mM), glucose (5.0 mM), FAD (1.0 µM), NT-11 PrnC (12.5 µM), Fre (2.5 µM) and Gdhi (2.5 µM) in 10mM potassium phosphate buffer, NADH (2.5 mM) was added to a total volume of 200 µL. After an overnight incubation of 30°C and orbital shaking at 350 rpm, reactions were quenched with an equivalent volume of MeOH, pelleted by centrifugation (15000 rpm for 10 min) and the supernatant analyzed by HPLC-MS using the analytical HPLC method.

Determination of Kinetic Parameters for NT-11 PrnC and PrnC mutant assay. Kinetic analysis of PrnC (2.5 µM) activity against MDA was performed over a 5-250 µM substrate concentration range. The assay reaction was supplemented with Fre (2.5 µM), FAD (1 µM) and $MgCl_2$ (10 mM) in 20 mM Tris buffer, pH 7.4. NADH (2.5 mM) was added last for reaction initiation. The amount of products formed were measured at 120, 300 and 600 seconds via a Kinetex XB-C18 reversed-phased column (2.6 µm, 150 x 4.6 mm) on a Shimadzu LC-20AD HPLC. Absorbance at λ = 254 nm was used to monitor product formation during an isocratic flow rate of 0.6 mL/min (50% MeCN/H_2O + 0.1%

TFA) over 10 min. Kinetic parameters were determined by nonlinear fitting of a Michaelis-Menten curve using the GraphPad Prism software. Activity assays for PrnC mutants were performed in lysates with the respective over-expressed protein variant. Wild-type PrnC enzyme was used as a positive control and reaction conditions were similar to the 18 h assay method described above. The amount of products formed by PrnC mutants were quantified using the analytical HPLC method.

Preparative Scale Cell Lysate Biotransformations:

Method for PrnC lysate preparation:

E. Coli strain expressing PrnC was cultured in 1 L of LB Kan^{50} media at 37 °C. At OD_{600} 0.4-0.6, 0.1 mM IPTG was used to induce protein expression at 16 °C over 18 h. Cell culture was harvested by centrifugation at 4000 rcf for 10 min at 4 °C. After media was decanted, cell pellet was resuspended in 30 ml of 50 mM tris pH 7.4, 300 mM sodium chloride, 10 mM imidazole and lysed by cell disruption. Cell lysate was centrifuged at 33,600 rcf for 45 min at 4 °C to differentiate supernatant from insoluble debris. Lysate supernatant was buffer exchanged into 50 mM tris pH 7.4 in preparation for activity assay.

PrnC-Cell lysate Chlorination. In a solution of the PrnC cell lysate (0.01 mM, enzyme loading ~2.0 mol%, total volume 20 mL), the pyrrolic derivative starting material (0.5 mM), $MgCl_2$ (10 mM), glucose (5.0 mM), FAD (1.0 μM), Fre (2.5 μM), GDH2 (2.5 μM), NADH (2.5 mM) was added and allowed to stir at 30°C with a stir bar in a petri dish at 30 rpm.

After an overnight incubation of 25°C and stir-bar shaking at 30 rpm, reactions were quenched with an equivalent volume of (1:1) MeOH-brine solution. The aqueous layer was 3x extracted with 20 mL of ethyl acetate pelleted by centrifugation (40 rpm for 10 min). The combined organic layers were dried over Na_2SO_4, filtered, and concentrated before purification by semi-preparative HPLC.

General HPLC and LC-MS methods

Analytical methods: Spectroscopic grade solvents were purchased from Sigma Aldrich. Low-resolution LC-MS spectra were recorded on an Agilent LCMS machine with dual MM-APCI-ES. High-resolution mass spectra (HRMS) were recorded on an Agilent

ESI-TOF mass spectrometer at 3500 V emitter voltage. Exact m/z values are reported in Daltons.

Semi-Preparative HPLC method. 900 µL of the crude mixture dissolved in H_2O/MeCN was injected onto a Phenomenex Jupiter® semi-preparative C18 HPLC column (90Å, 5 µm packing, 250 x 10 mm) and purified using reverse phase chromatography. Gradient starting conditions of 5% MeCN/H_2O (+0.1% Formic acid) to 25% MeCN/H_2O over 10 min, followed by 25% MeCN/H_2O into 50% MeCN/H_2O over 20 min, followed by 50% MeCN/H_2O into 75% MeCN/H_2O over 10 min, followed by 75% MeCN/H_2O into 95% MeCN/H_2O over 5 min, followed by a hold at 95% MeCN/H_2O for 5 min. Column condition was equilibrated back to starting conditions over 2 mins post-run. Flow rates were kept constant at at 3 mL/ min. UV absorbance was monitored at 220 nm, 254 nm and 280 nm.

Analytical HPLC Method. 10 µL of the supernatant injected onto SecurityGuard™ column (KJ0-4282) with a (4.0 mm x 3.0 mm) guard cartridge before separation using a Phenomenex Gemini® C18 analytical column (5 µm packing, 150 mm x 4.6 mm). Gradient starting conditions of 5% MeCN/H_2O (+0.1% Formic acid) were held for 1 min before development into 50% MeCN/H_2O over 3 min, followed by development into 95% MeCN/H_2O over 3 min. 95% MeCN/H_2O was held for 1 min before equilibration back to starting conditions over 1 min. Starting conditions was held for 1 min followed by another 2 min post-run. Flow rates were kept constant at 1 mL/ min. Column temperature was kept constant at 30 °C. UV absorbance was detected at 220 nm, 254 nm and 210 nm throughout the run.

General LC-MS Method. 10 µL of the supernatant was separated using the appropriate analytical HPLC method described above. Detection was performed using an Agilent® single quadrupole LC/MSD system.

Synthesis of Substrates & Standards

A substrate scope of the enzyme by the application of the Prnc enzyme on a list of structural diverse pyrrolic substrates is determined. An optimized co-factors regeneration protocol was used to regulate the generation of the activated halogen which is captured by the Prnc enzyme into the active site for the bioconversion. Reaction is conducted at ambient temperature of 30 °C, in an aqueous non-toxic buffer solution

with an environmentally friendly chloride source.

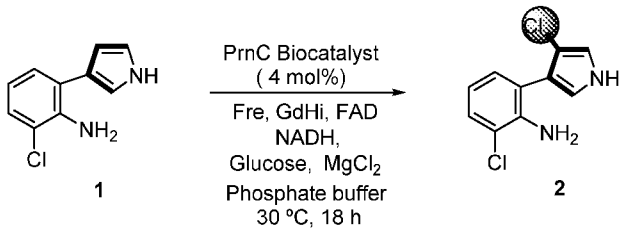
Table 1 Optimization.

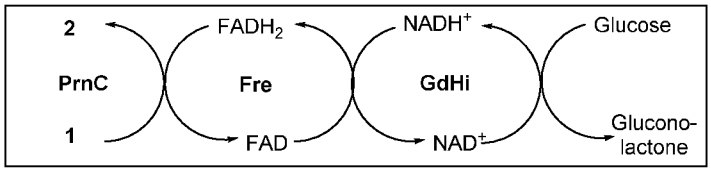
			No catalase		NADH		Tris buffer		Fre	
Reaction component	**1**		**2**		**3**		**4**		**5**	
FAD (µM)	1	1	1	1	1	1	1	1	1	1
NADH (mM)	2.5	2.5	2.5	2.5	25 µM	25 µM	2.5	2.5	2.5	2.5
MgCl2 (mM)	10	10	10	10	10	10	10	10	10	10
PrnC substrate (mM)	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Catalase (U/ml)	33	33			33	33	33	33	33	33
Glucose (mM)	5	5	5	5	5	5	5	5	5	5
GDH (µM)	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
Fre (µM)	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	5	5
Prnc (µM)	10		10		10		10		10	
10 mM sodium phosphate pH 7.0	Top up to 200 µl								Top up to 200 µl	
50 mM Tris pH 7.4							Top up to 200 µl			

Table 2 Optimization.

Reaction component	**Iodide**	**5x NadH, No Gdh**	**0.05 NADH**	**No PrnC**	**Std Conditions**
FAD (µM)	1	1	1	1	1
	2.5	2.5	0.05	2.5	2.5
MgCl2 (mM)		10	10	10	10
PrnC substrate (mM)	0.5	0.5	0.5	0.5	0.5
Iodide (mM)	20				
Glucose (mM)	5		5	5	5
GDH (µM)	2.5		2.5	2.5	2.5
Fre (µM)	2.5	2.5	2.5	2.5	2.5
Prnc (µM)	12.5	12.5	12.5	0	12.5

Table 3 Optimization study of PrnC biocatalyst on the native substrate 1





Entry	Variation from standard conditions	conversion[a] of **2** (%)
1.	2 mol% FAD (10x), no GdHi & glucose	18
2.	2 mol% FAD (10x), GdHi regen	16
3.	None	91[b]
4.	1 mol% of Fre (2x)	72
5.	0.05 Equiv NADH (Less 100x)	32
6.	0.05 Equiv NADH (Less 100x) + 8 mol% PrnC (2x)	53
7.	0.05 Equiv NADH (Less 100x) in Tris HCl buffer	49

[a]Conversion is calculated based on calibration curves of product (**2**) and starting material (**1**).[b]Conditions: PrnC Biocatalyst (4 mol%), GdHi (0.5 mol%), FAD (0.2 mol%), Fre (0.5 mol%), NADH (5.0 equiv), Glucose (10.0 equiv), $MgCl_2$ (20.0 equiv), phosphate buffer (10 mM, pH 7.4)

The conditions in the table above would be referred to as the standard conditions of the present invention; i.e. flavin-dependent halogenase (4 mol%), flavin adenine dinucleotide (FAD) (0.2 mol%), *E. coli* flavin reductase (Fre) (0.5 mol%), nicotinamide adenine dinucleotide hydrogen (NADH) (5.0 equiv), glucose 1-dehydrogenase (GdHi) (0.5 mol%), monosaccharide (10.0 equiv), halogen (20.0 equiv), buffer (10 mM, pH 7.4).

With the purified biocatalyst PrnC enzyme in hand, a preliminary optimization study of the enzymatic assay conditions was performed on the native substrate monodechloroaminopyrrolnitrin (**1**). The enzyme assays were carried out for 18 h with orbital shaking at 300 rpm and a temperature of 30 °C. In addition to the substrate and PrnC, the reaction mixture contained FAD (0.1 equivalents relative to PrnC), NADH and *E. coli* flavin reductase (Fre), needed to generate $FADH_2$ for the halogenation reaction. In initial assays, even using a super stoichiometric amount of NADH (5.0 equiv) (entry 1) only yielded 18% conversion to the desired product. Addition of a glucose

dehydrogenase (GdHi)-NADH regeneration system also did not lead to discernable improvement (entry 2), indicating that other parameters should be evaluated.

Surprisingly, a tenfold decrease in FAD concentration (entry 3) had a major effect on the conversion yield, increasing it to 91%. The reasoning is that is that while the rate of $FADH_2$ formation is slower, it parallels the progressive generation of hypohalous acid that is necessary for the formation of the chloroamine intermediate involved in site-specific halogenation. Conversely, increasing the Fre concentration to maximize $FADH_2$ turnover (entry 4) did not produce results that were comparable to the standard conditions. A potential reason for this could be the accumulation of uncoupled $FADH_2$ that reacts with molecule oxygen to form H_2O_2, causing detrimental effects on the enzyme.

Despite already being present in excess (relative to FAD), conservation of relatively expensive NADH (entry 5) did not yield comparable results to standard conditions. This data indicates that a threshold amount of NADH is necessary to initiate the cascade catalytic cycle. An increase of two-fold in the biocatalyst loading (c.f. entry 5 vs entry 6) led to an almost corresponding increase in **2**, possibly due to the higher coupling probability between the halogenase and $FADH_2$. The switch from phosphate to Tris HCl buffer (c.f. entry 5 vs entry 7) appears to provide an incremental improvement but was not pursued because of conflicts with other potential halide sources (i.e., Br-).

In-vitro halogenation of the native MDA substrate (2'-pyrroles) using our optimized conditions gave 55% of product **2**. The catalytic efficiency ($\boldsymbol{k_{cat}/K_m}$) of PrnC enzyme against MDA substrate was determined to be 4.83×10^2 $\mathbf{s^{-1}\,M^{-1}}$ (Figure 8). A curated list of biaryl-pyrroles was assembled (Figure 4) with respect to a) the different substituents on the adjacent phenyl group, b) the structural positional isomers of the pyrrolic fragment and c) other heterocyclic 5-membered rings. This regioselectivity of this bio-halogenation has been analogously assigned to be on the backbone of the pyrrolic fragment based on our computation model with further collaboration by 2D NMR structural confirmation on selected substrates **6**, **9** and **11**.

The capping of the free amino group with acetanilide in product **3** seems to significantly reduce the conversion yield (18%), which may be due to the limited space for an extra acetyl group (**Figure 12**). In products **5**, **6**, and **7**, the simultaneous substitution of the

ortho-phenyl -NH_2 group and removal of the phenyl 3-Cl atom had a moderately negative impact on the conversion yields (22%, 41% and 21%). These results also confirm that prior installation of the 3-Cl atom by PrnA has no bearing on the subsequent step of halogenation.

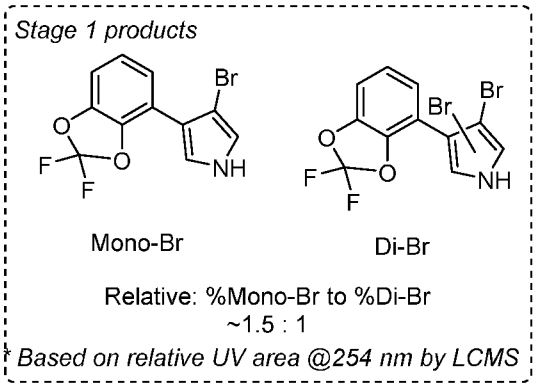
The total removal of substituents from the aryl fragment **8** (77%) or substitution with a heterocycle pyridine **9** (73%) or quinoline **14** (63%) yielded good conversion, indicating that the hydrophobic interaction of these phenyl groups embedded in the hydrophobic patch is highly relevant (Figure 12). These findings further suggest that the $-NH_2$ moiety does not appear to play an active directing function in the halogenation step.

Other bio-halogenation attempts on isosteric alternatives in the pyrrolic fragment were met with moderate success with 2'-pyrrole **11** (68 %) and 5'-pyrazole **12** (14 %). Other 5-membered heterocycles such as furan **15** and thiophene **16**, on the other hand, were not tolerated by the PrnC enzyme. This may be hypothesized that these substrates are unable to form a hydrogen bond with the crucial residue E60, as illustrated in Figure 2, despite having electron-rich rings that are highly susceptible to electrophilic substitution.

It is believed that in **13**, the highly electron deactivated pyrazole ring coupled with unfavorable binding interaction of the Nitrogen-atom with the active residues within the enzyme binding pocket may hinder the halogenation reaction at 3' and 5' position.

To further demonstrate the utility of the Prnc biocatalyst on a practical target, the enzymatic halogenation sequence was applied en route to the synthesis of Fludioxonil (Figure 6 and 7).

Stage 1



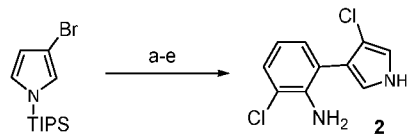
Stage 2: (Done in 1 pot)

Upon completion of stage 1, crude mixture was directly topped up with reagents and submitted to conditions from stage 2.

%Conversion has not been determined but a strong mass signal on -ve mode LCMS has been detected which correspond to the mass of the CN-product.

NMR spectra were recorded on Bruker Avance III 400 MHz spectrometer in CDCl3 or MeOD-d4. Data is reported in the following order: chemical shifts are given(δ); multiplicities are indicated as s (singlet), d (doublet), t (triplet), q (quartet) and m (multiplet).

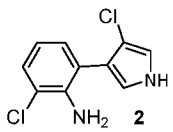
Chemicals and anhydrous solvents were obtained from Sigma Aldrich and were used without further purification.



Reagents and conditions

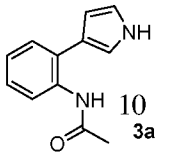
a) t-BuLi, THF, -78 °C, than C_2Cl_6/ THF, -78 °C to rt, 68%. b) NIS, Acetone, rt, 76%. c) HBPin, $PdCl_2(CH_3CN)_2$,Sphos, PhMe, 90 °C. d) 2-bromo-6-chloroaniline, $PdAc_2$, K_3PO_4, Sphos, $BuOH/H_2O$, 40 °C, 38% (2 steps).e) TBAF, THF, rt, 68%

Scheme 1. Synthesis of **2** [Note: Precursors to compounds **3**, **4**, **5**, **6**, **7**, **8**, **10 and 14** were prepared through similar sequences].



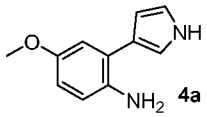
2-chloro-6-(4-chloro-1H-pyrrol-3-yl)aniline (2)

^{1}H NMR (400 MHz, CDCl3) δ 8.30 (s, 1H), 7.24 (dd, J = 8.0, 1.5 Hz, 1H), 7.09 (dd, J = 7.6, 1.5 Hz, 1H), 6.86 (ddd, J = 15.4, 2.9, 2.3 Hz, 2H), 6.71 (t, J = 7.8 Hz, 1H), 4.16 (s, 2H). ^{13}C NMR (100 MHz, CDCl3) δ 141.7, 130.0, 128.5, 120.1, 119.4, 119.3, 117.92, 117.2, 116.2, 112.2. This is in accordance with literature data.



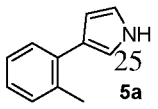
N-(2-(1H-pyrrol-3-yl)phenyl)acetamide (3a)

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.52 (s, 1H), 8.31 (d, *J* = 8.2 Hz, 1H), 7.71 (s, 1H), 7.34 – 7.23 (m, 2H), 7.14 – 7.06 (m, 1H), 6.96 – 6.88 (m, 2H), 6.36 (q, *J* = 2.4 Hz, 1H), 2.10 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ=168.2, 135.2, 129.9, 127.3, 126.2, 124.0, 120.7, 120.7, 119.2, 116.7, 108.8, 24.8. IR (neat) 3245, 1624, 1602, 1527, 1373, 1298, 1084, 759 cm^{-1}. HRMS (ESI) calcd. for $C_{12}H_{13}N_2O$ m/z $(M+H)^+$: 201.1028, found: 201.1016 .



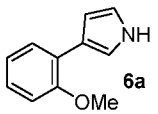
4-methoxy-2-(1H-pyrrol-3-yl)aniline (4a)

^{1}H NMR (400 MHz, $CDCl_3$) δ=8.36 (s, 1H), 7.02 (dt, *J* = 2.6, 1.8 Hz, 1H), 6.94 – 6.84 (m, 2H), 6.75 – 6.64 (m, 2H), 6.46 (td, *J* = 2.7, 1.6 Hz, 1H), 3.77 (s, 3H), 3.70 (s, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ=152.7, 137.5, 123.5, 122.0, 118.5, 116.7, 116.4, 115.3, 112.9, 108.7, 55.8. IR (neat) 3351, 3284, 2963, 1615, 1508, 1475, 1284, 1241, 1207, 871, 806, 746, 726 cm^{-1}. HRMS (ESI) calcd. for $C_{11}H_{13}N_2O$ m/z $(M+H)^+$: 189.1028, found: 189.1087.



2-(o-tolyl)-1H-pyrrole (5a)[3]

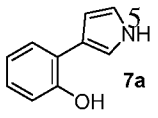
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.29 (s, 1H), 7.39 (dd, *J* = 7.5, 1.7 Hz, 1H), 7.27 – 7.11 (m, 3H), 6.90 (dt, *J* = 2.6, 1.8 Hz, 1H), 6.85 (td, *J* = 2.7, 2.0 Hz, 1H), 6.42 (td, *J* = 2.7, 1.6 Hz, 1H), 2.44 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ=135.8, 135.3, 130.5, 129.3, 125.9, 125.8, 124.4, 117.6, 116.6, 109.5, 21.4. HRMS (ESI) calcd. for $C_{11}H_{12}N$ m/z $(M+H)^+$: 158.0961, found: 158.0970. This is in accordance with literature data.



2-(2-methoxyphenyl)-1H-pyrrole (6a)[4]

^{1}H NMR (400 MHz, $CDCl_3$) δ=8.27 (s, 1H), 7.56 (dd, *J* = 7.6, 1.7 Hz, 1H), 7.36 (dt, *J* = 2.7, 1.8 Hz, 1H), 7.17 (ddd, *J* = 8.2, 7.4, 1.8 Hz, 1H), 7.03 – 6.91 (m, 2H), 6.84 (td, *J* = 2.8, 2.0 Hz, 1H), 6.64 (td, *J* = 2.8, 1.5 Hz, 1H), 3.90 (s,

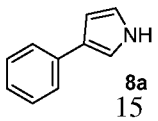
3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 156.1, 128.0, 126.3, 124.6, 120.8, 120.4, 117.9, 117.7, 111.1, 108.0, 55.4. HRMS (ESI) calcd. for $C_{11}H_{12}NO$ m/z $(M+H)^+$: 174.0913, found: 174.0919. This is in accordance with literature data.



7a

2-(1H-pyrrol-2-yl)phenol (7a)

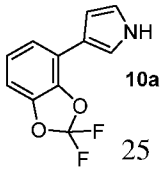
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.44 (s, 1H), 7.33 (dd, *J* = 7.5, 1.8 Hz, 1H), 7.17 (ddd, *J* = 8.1, 7.3, 1.7 Hz, 1H), 7.05 (q, *J* = 2.0 Hz, 1H), 7.00 – 6.89 (m, 3H), 6.44 (td, *J* = 2.7, 1.6 Hz, 1H), 5.60 (s, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ=152.7, 129.4, 127.7, 122.6, 120.5, 119.5, 119.4, 116.4, 115.2, 108.4. IR (neat) 3418, 1452, 753 cm^{-1}. HRMS (ESI) calcd. for $C_{10}H_{10}NO$ m/z $(M+H)^+$: 160.0755, found: 160.0762.



8a

3-phenyl-1H-pyrrole (8a)[5]

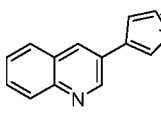
^{1}H NMR (400 MHz, $CDCl_3$) δ=8.25 (s, 1H), 7.60 – 7.52 (m, 2H), 7.40 – 7.31 (m, 2H), 7.24 – 7.15 (m, 1H), 7.10 (dt, *J* = 2.6, 1.8 Hz, 1H), 6.84 (td, *J* = 2.7, 2.0 Hz, 1H), 6.57 (td, *J* = 2.7, 1.6 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 135.8, 128.6, 125.5, 125.3, 125.0, 118.9, 114.6, 106.6. HRMS (ESI) calcd. for $C_{10}H_{10}N$ m/z $(M+H)^+$: 144.0813, found: 144.0804. This is in accordance with literature data.



10a

3-(2,2-difluorobenzo[d][1,3]dioxol-4-yl)-1H-pyrrole (10a)

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.41 (s, 1H), 7.34 (dt, J = 2.8, 1.8 Hz, 1H), 7.29 (dd, J = 8.2, 1.2 Hz, 1H), 7.05 (t, J = 8.0 Hz, 1H), 6.92 – 6.81 (m, 2H), 6.66 (td, J = 2.8, 1.6 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 144.0, 139.6, 131.5 (^{1}JC-F = 254.5 Hz), 123.6, 120.6, 119.4, 118.9, 117.6, 117.4, 106.9, 105.9. IR (neat) 3410 3392 1654 1453 1231 1130 1082 880 767 723 712 cm^{-1}. HRMS (ESI) calcd. for $C_{11}H_8F_2NO_2$ m/z $(M+H)^+$: 224.0523, found: 224.0514.

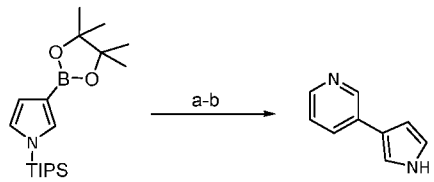


14a

3-(1H-pyrrol-3-yl)quinolone (14a)

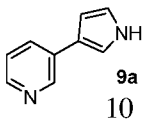
^{1}H NMR (400 MHz, $CDCl_3$) δ 9.16 (d, J = 2.3 Hz, 1H), 8.59 (s, 1H), 8.24 – 8.14 (m, 1H), 8.07 (dq, J = 8.5, 0.9 Hz, 1H), 7.81 (ddt, J = 8.1, 1.4, 0.5 Hz, 1H), 7.63 (ddd, J = 8.4, 6.9, 1.5 Hz, 1H), 7.52 (ddd, J = 8.1, 6.9, 1.2 Hz, 1H), 7.29 (dt, J = 2.8, 1.8 Hz, 1H), 6.94 (td, J = 2.8, 2.0 Hz, 1H), 6.70 (td, J = 2.7, 1.6 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 149.42, 146.55, 129.80, 129.15, 128.97,

128.55, 128.21, 127.53, 126.72, 121.71, 119.59, 115.39, 106.70. IR (neat) 3168 2924 2854 1602 1489 1339 1165 1083 902 789 748 cm^{-1}. HRMS (ESI) calcd. for $C_{13}H_{11}N_2$ m/z $(M+H)^+$: 195.0922, found: 195.0913.



Reagents and conditions
a) 3-iodopyridine, $Pd(PPh_3)_4$, Na_2CO_3, MeOH/H_2O/PhMe (1:1:5), 100 °C, 83%.
b) TBAF, THF, rt, 98%

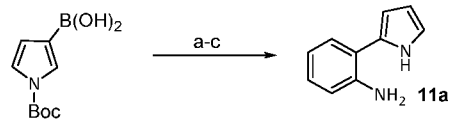
Scheme 2. Synthesis of **9a**



9a

3-(1H-pyrrol-3-yl)pyridine (9a)[6]

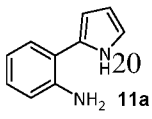
^1H NMR (400 MHz, $CDCl_3$) δ=9.08 (s, 1H), 8.81 (dd, *J* = 2.3, 0.9 Hz, 1H), 8.40 (dd, *J* = 4.8, 1.6 Hz, 1H), 7.80 (ddd, *J* = 7.9, 2.3, 1.6 Hz, 1H), 7.30 – 7.22 (m, 1H), 7.15 (dt, *J* = 2.7, 1.8 Hz, 1H), 6.88 (td, *J* = 2.7, 1.9 Hz, 1H), 6.55 (td, *J* = 2.7, 1.6 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 146.6, 146.4, 132.4, 131.9, 123.6, 121.3, 119.5, 115.2, 106.3, 77.3. HRMS (ESI) calcd. for $C_9H_9N_2$ m/z $(M+H)^+$: 145.0766, found: 145.0754. This is in accordance with literature data.



11a

Reagents and conditions
a) 3-iodo-2-nitrobenzene, $Pd(OAc)_2$, SPhos, K_3PO_4, Dioxane/H_2O, 60 °C, 95%.
b) NaOMe, MeOH, 78 °C, 83%. c) Pd/C, MeOH, rt, 50%.

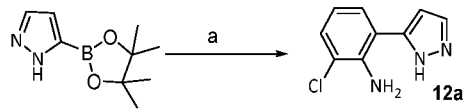
Scheme 3. Synthesis of **11a**



11a

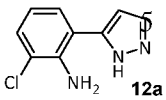
2-(1H-Pyrrol-2-yl)aniline (11a)[7]

^1H NMR (400 MHz, $CDCl_3$) δ=8.60 (s, 1H), 7.26 (dd, *J* = 7.7, 1.5 Hz, 1H), 7.10 (ddd, *J* = 7.9, 7.4, 1.6 Hz, 1H), 6.88 (td, *J* = 2.7, 1.5 Hz, 1H), 6.86 – 6.74 (m, 2H), 6.43 (ddd, *J* = 3.4, 2.6, 1.5 Hz, 1H), 6.33 (dt, *J* = 3.4, 2.6 Hz, 1H), 3.96 (s, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 143.4, 129.6, 128.5, 127.9, 119.7, 119.2, 118.0, 116.5, 109.4, 107.4. HRMS (ESI) calcd. for $C_{10}H_{11}N_2$ m/z $(M+H)^+$: 159.0922, found: 159.0911. This is in accordance with literature data.



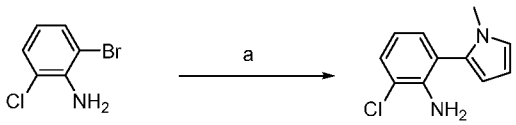
Reagents and conditions
a) 2-bromo-6-chloroaniline, $Pd(dppf)Cl_2.CH_2Cl_2$, Na_2CO_3, Dioxane/H_2O, Reflux, 30%.

Scheme 3. Synthesis of **12a**



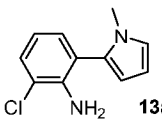
2-chloro-6-(1H-pyrazole-5-yl)aniline (12a)

^{1}H NMR (400 MHz, $CDCl_3$) δ=7.62 (d, *J* = 2.5 Hz, 1H), 7.46 (dd, *J* = 7.8, 1.5 Hz, 1H), 7.24 (dd, *J* = 7.9, 1.5 Hz, 1H), 6.74 – 6.62 (m, 2H), 5.96 (s, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ=151.4, 141.4, 129.3, 128.6, 126.8, 120.1, 117.3, 116.9, 103.6. IR (neat) 3426 3310 2924 1611 1458 1074 1045 747 cm^{-1}. HRMS (ESI) calcd. for $C_9H_9ClN_3$ m/z $(M+H)^+$: 194.0485, found: 194.0485.



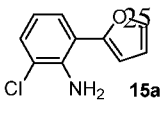
Reagents and conditions
a) (1*H*-pyrrol-2-yl)boronic acid, $Pd(OAc)_2$, SPhos, $K_3PO_4.H_2O$, *n*-butanol /H_2O, 40 °C, 82%.

Scheme 4. Synthesis of **13a** [Note: compounds **15a** and **16a** were prepared through similar sequences with the corresponding boronic acid/ester partner].



2-chloro-6-(1-methyl-1H-pyrrol-2-yl)aniline (13a)

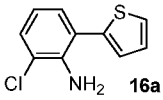
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 – 7.27 (m, 1H), 7.02 (dd, *J* = 7.6, 1.5 Hz, 1H), 6.75 (dd, *J* = 2.7, 1.7 Hz, 1H), 6.70 (dd, *J* = 8.0, 7.5 Hz, 1H), 6.23 (dd, *J* = 3.5, 2.7 Hz, 1H), 6.18 (dd, *J* = 3.5, 1.8 Hz, 1H), 4.19 (s, 2H), 3.48 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 142.7, 130.2, 129.8, 129.2, 123.1, 119.9, 119.3, 117.8, 109.1, 107.9, 34.4. IR (neat) 3476 3374 3103 2941 2295 1858 1610 1586 1486 1451 1425 1408 1358 1309 1250 10731054 989 888 776 755 736 715 cm^{-1}. HRMS (ESI) calcd. for $C_{11}H_{11}ClN_2$ m/z $(M+H)^+$: 207.0689, found: 207.0682.



2-chloro-6-(furan-2-yl)aniline (15a)

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.52 (dd, *J* = 1.9, 0.8 Hz, 1H), 7.37 (dd, *J* = 7.8, 1.5 Hz, 1H), 7.23 (dd, *J* = 7.9, 1.5 Hz, 1H), 6.71 (t, *J* = 7.9 Hz, 1H), 6.62 (dd, *J* = 3.4, 0.8 Hz, 1H), 6.52 (dd, *J* = 3.4, 1.9 Hz, 1H), 4.85 (s, 2H). ^{13}C

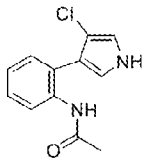
NMR (101 MHz, $CDCl_3$) δ 152.8, 141.8, 140.0, 128.9, 126.3, 120.6, 118.2, 117.4, 111.6, 107.4. IR (neat) 3468 3374 1615 1499 1463 1447 1372 1317 1271 1244 1213 1152 1075 1014 905 885 802 772 725 cm^{-1}. HRMS (ESI) calcd. for $C_{10}H_8ClNO$ m/z $(M+H)^+$: 194.0373, found: 194.0367.



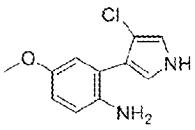
16a

2-chloro-6-(thiophen-2-yl)aniline (16a)

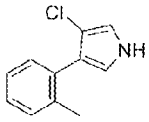
^1H NMR (400 MHz, $CDCl_3$) δ 7.37 (dd, *J* = 5.2, 1.2 Hz, 1H), 7.26 – 7.20 (m, 2H), 7.17 (dd, *J* = 7.7, 1.5 Hz, 1H), 7.13 (dd, *J* = 5.2, 3.5 Hz, 1H), 6.71 (t, *J* = 7.8 Hz, 1H), 4.43 (s, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 141.1, 140.3, 129.6, 129.2, 127.8, 126.4, 125.9, 121.1, 119.9, 118.3. IR (neat) 3421 3325 1893 1843 1793 1614 1452 1348 1304 1268 1233 1197 1149 1069 1044 972 949 850 825 765 725 736 688 cm^{-1}. HRMS (ESI) calcd. for $C_{10}H_8ClNS$ m/z $(M+H)^+$: 210.0144, found: 210.0138.



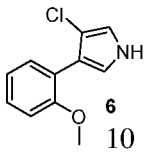
N-(2-(4-chloro-1H-pyrrol-3-yl)phenyl)acetamide (3): ^1H NMR (400 MHz, Methanol-*d*4) δ 8.55 (s, 1H), 7.70 (d, *J* = 8.0 Hz, 1H), 7.31 (d, *J* = 6.8 Hz, 1H), 7.29 (t, *J* = 7.0 Hz, 1H), 7.19 (t, *J* = 7.5 Hz, 1H), 6.85 (d, *J* = 2.4 Hz, 1H), 6.79 (d, *J* = 2.3 Hz, 1H), 2.04 (s, 3H). ^{13}C NMR (101 MHz, Methanol-*D*4) δ 172.1, 136.7, 132.5, 129.2, 128.3, 126.3, 125.8, 119.6, 118.7, 117.3, 112.1, 23.5. HRMS *m/z* (ESI) cald for $C_{12}H_{12}ClN_2O$ ($[M+H]^+$) 235.0638, found: 235.0637.



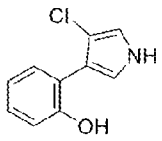
2-(4-chloro-1H-pyrrol-3-yl)-4-methoxyaniline (4): ^1H NMR (400 MHz, Methanol-*d*4) δ 7.40 – 7.32 (m, 1H), 7.09 – 7.02 (m, 2H), 6.98 (d, *J* = 1.4 Hz, 2H), 3.88 (s, 3H). ^{13}C NMR (101 MHz, Methanol-*D*4) δ 161.1, 131.7, 125.5, 123.3, 119.3, 119.1, 118.5, 117.5, 115.3, 112.1, 56.5. HRMS *m/z* (ESI) cald for $C^{11}H^{12}N_2OCl$ ($[M+H]^+$) 223.0638, found: 223.0669.



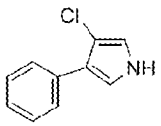
3-chloro-4-(o-tolyl)-1H-pyrrole (5): ^{1}H NMR (400 MHz, Methylene Chloride-d2) δ 8.36 (s, 1H), 7.31 – 7.14 (m, 4H), 6.85 (q, J = 2.8 Hz, 1H), 6.72 (q, J = 2.9 Hz, 1H), 2.28 (d, J = 2.8 Hz, 3H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 137.5, 133.2, 131.1, 130.0, 127.3, 125.3, 122.4, 117.0, 115.4, 112.1, 20.1. HRMS *m/z* (ESI) cald for $C_{11}H_{11}ClN$ ($[M+H]^+$) 192.0580, found: 192.0574.



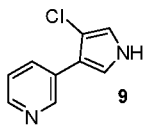
3-chloro-4-(2-methoxyphenyl)-1H-pyrrole (6): Biotransformation was prepared as per the general preparative PrnC-Cell lysate chlorination method detailed above to overall volume of 20 mL. ^{1}H NMR (400 MHz, Methanol-d_4) δ 7.58 (ddd, *J* = 7.6, 1.8, 0.4 Hz, 1H), 7.44 – 7.35 (m, 1H), 7.21 – 7.14 (m, 1H), 7.11 (td, *J* = 7.5, 1.2 Hz, 1H), 7.04 (d, *J* = 2.4 Hz, 1H), 6.93 (d, *J* = 2.4 Hz, 1H), 3.97 (s, 3H). HRMS *m/z* (ESI) cald for $C_{11}H_{11}ClNO$ ($[M+H]^+$) 208.0529, found: 208.0523.



2-(4-chloro-1H-pyrrol-3-yl)phenol (7): ^{1}H NMR (400 MHz, Methanol-d4) δ 7.41 (dd, J = 7.6, 1.7 Hz, 1H), 7.11 (ddd, J = 8.1, 7.3, 1.7 Hz, 1H), 6.97 (d, J = 2.1 Hz, 1H), 6.88 (dd, J = 8.2, 1.2 Hz, 1H), 6.86 (td, J = 7.4, 1.3 Hz, 1H), 6.81 (d, J = 2.4 Hz, 1H). ^{13}C NMR (101 MHz, Methanol-D4) δ 155.9, 132.4, 128.7, 122.8, 120.6, 119.6, 118.6, 116.9, 116.7, 112.2, 49.3. HRMS *m/z* (ESI) cald for $C_{10}H_9ClNO$ ($[M+H]^+$) 194.0373, found: 194.0365.

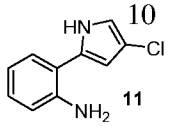


3-chloro-4-phenyl-1H-pyrrole (8): ^{1}H NMR (400 MHz, Methylene Chloride-*d*2) δ 8.42 (s, 1H), 7.61 – 7.54 (m, 2H), 7.41 – 7.33 (m, 2H), 7.29 – 7.22 (m, 1H), 6.95 – 6.89 (m, 1H), 6.85 (t, *J* = 2.6 Hz, 1H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 133.9, 128.4, 128.4, 127.6, 127.6, 126.4, 122.4, 116.7, 116.3, 110.6. HRMS *m/z* (ESI) cald for $C_{10}H_9ClN$ ($[M+H]^+$) 178.0424, found: 178.0417.



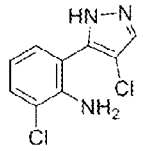
9

3-(4-chloro-1H-pyrrol-3-yl)pyridine (9): Biotransformation was prepared as per the general preparative PrnC-Cell lysate chlorination method detailed above to overall volume of 20 mL. ^{1}H NMR (400 MHz, Methanol-*d*$_4$) δ 8.92 (dd, *J* = 2.3, 0.9 Hz, 1H), 8.55 (dd, *J* = 4.9, 1.6 Hz, 1H), 8.23 (ddd, *J* = 8.0, 2.3, 1.6 Hz, 1H), 7.60 (ddd, *J* = 8.0, 4.9, 0.9 Hz, 1H), 7.23 (d, *J* = 2.3 Hz, 1H), 7.04 (d, *J* = 2.4 Hz, 1H). ^{13}C NMR (101 MHz, Methanol-*d*$_4$) δ 149.2, 147.8, 137.4, 133.6, 125.9, 119.7, 119.3, 118.9, 111.6, 49.9. HRMS *m/z* (ESI) cald for $C_9H_8N_2Cl$ $[M+H]^+$ 179.0371, found: 179.0390.

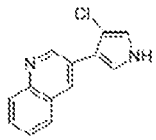


10 11

2-(4-chloro-1H-pyrrol-2-yl)aniline (11): Biotransformation was prepared as per the general preparative PrnC-Cell lysate chlorination method detailed above to overall volume of 20 mL.^{1}H NMR (400 MHz, Methanol-*d*$_4$) δ 7.34 – 7.25 (m, 2H), 6.99 (dd, *J* = 8.5, 1.2 Hz, 1H), 6.95 (d, *J* = 3.0 Hz, 1H), 6.91 (td, *J* = 7.6, 1.2 Hz, 1H), 6.32 (d, *J* = 3.0 Hz, 1H), 5.67 (s, 1H), 4.74 (s, 2H). ^{13}C NMR (101 MHz, Methanol-*D*$_4$) δ 147.7, 133.0, 130.9, 127.3, 119.8, 119.7, 119.5, 117.9, 111.5, 110.5. HRMS m/z (ESI) cald for $C_{10}H_{10}ClN_2$ ($[M+H]^+$) 193.527, found: 193.0535.



2-chloro-6-(4-chloro-1H-pyrazol-5-yl)aniline (12): 1H NMR (400 MHz, Methylene Chloride-*d*2) δ 7.76 (s, 1H), 7.59 (dd, *J* = 7.8, 1.5 Hz, 1H), 7.33 (dd, *J* = 7.9, 1.4 Hz, 1H), 6.77 (t, *J* = 7.8 Hz, 1H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 144.5, 142.0, 130.2, 129.6, 128.6, 119.8, 117.1, 115.7, 109.1. HRMS *m/z* (ESI) cald for $C_9H_8Cl_2N_3$ ($[M^+H]^+$) 228.0095, found: 228.0098.



3-(4-chloro-1H-pyrrol-3-yl)quinoline: ^{1}H NMR (400 MHz, Methylene Chloride-*d*2) δ 9.10 (d, *J* = 2.3 Hz, 1H), 8.77 (s, 2H), 8.43 (d, *J* = 2.3 Hz, 1H), 8.07 (d, *J* = 8.5 Hz, 1H), 7.87 (dd, *J* = 8.2, 1.4 Hz, 1H), 7.68 (ddd, *J* = 8.5, 6.9, 1.5 Hz, 1H), 7.56 (ddd, *J* = 8.1, 6.8, 1.3 Hz, 1H), 7.13 (t, *J* = 2.8 Hz, 1H), 6.95 (t, *J* = 2.6 Hz, 1H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 151.0, 147.2, 133.2, 129.6, 129.3, 128.6,

128.4, 127.7, 127.3, 119.6, 117.8, 117.4, 111.5. HRMS *m/z* (ESI) cald for $C_{13}H_{10}ClN_2$ ($[M+H]^+$) 229.0533, found: 227.0572.

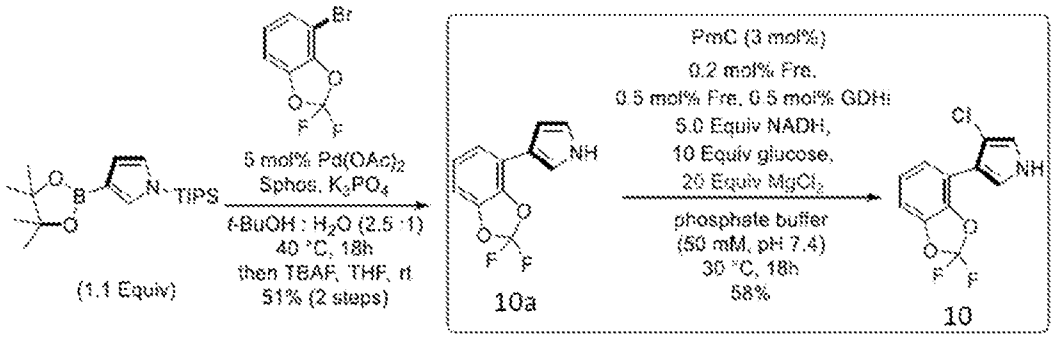
Comparison of Enzymatic chlorination vs Synthetic methods (Reactivity and Selectivity):

Synthetic Chlorination Procedures: To demonstrate and contrast the enzymatic efficiency with commonly routinely employed protocols, chlorination was attempted using these synthetic reagents.

Three substrates **1**, **11a** and **14a** were selected based on structural diversity in their ring sizes, electronics and substitution position with respect to the pyrrolic fragment. The chlorination procedures tested were a) N-chlorosuccinimide (NCS) in MeCN and b) sodium hypochlorite (NaOCl) prepared from Sodium Hypochlorite Pentahydrate ($NaOCl.5H_2O$) in aqueous buffer (50 mM sodium phosphate buffer, pH 7.4). Initial attempts to promote mono-chlorination reactions were attempted using benign conditions with only 1.2 equivalent of the chlorinating source at room temperature, stirring at 600 rpm for 18 hours. Reactions were quenched with an equal volume of MeOH and analyzed by LCMS.

Additional investigation was conducted under more accelerating conditions in light of the initial low conversion observed with only 10 equivalents of the chlorinating source at 50°C, stirring at 600 rpm for 2 days. Compared to the presently disclosed method, synthetic reactions require heating at elevated temperature (vs room temperature), the using of organic solvent (vs water), and the formation of a mixture of isomers (vs 1 major isomer).

Synthesis of substrate 10a and chlorinated Fludioxonil analog 10



4-bromo-2,2-difluorobenzo[*d*][1,3]dioxole (237 mg, 0.14 mL, 1 mmol), *N*-TIPS pyrrole-3-boronic acid pinacol ester (384 mg, 1.1 mmol, 1.1 eq), Palladium acetate (12 mg, 5 mol%), SPhos (33 mg, 8 mol%) and K_3PO_4 (0.38 g, 1.8 mmol) were added to a schlenk tube containing degassed *n*-BuOH : H_2O (2.5 :1, 3 ml). The mixture was stirred at 40 °C for 18 hours. Upon complete conversion of starting material by TLC, water (5 mL) was added to the mixture and extracted with (5 x 3mL) of ethyl acetate. The organic layers were combined, filtered through a plug of celite and dried over anhydrous sodium sulphate. Concentration over reduced pressure afforded a crude colorless oil (341 mg, 90%) which was used directly in the next step without purification. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 – 7.26 (m, 2H), 7.03 (t, *J* = 8.0, 1H), 6.88 – 6.77 (m, 2H), 6.77 – 6.66 (m, 1H), 1.54 – 1.45 (m, 3H), 1.14 (d, *J* = 7.5, 18H).

3-(2,2-difluorobenzo[d][1,3]dioxol-4-yl)-1H-pyrrole (341 mg, 0.9 mmol) was dissolved in THF (5.0 ml) under nitrogen. TBAF (1.0 M in THF, 1.6 ml, 1.8 eq) was added and the mixture was stirred at room temperature for 1 hour. Upon complete conversion of starting material by TLC, the mixture was evaporated to dryness. The resultant residue was directly subjected to purification by silica chromatography. Side product silanol and unreacted starting material were removed by 10:1 of (hexane: ethyl acetate) and the product **10a** (113 mg, 57%) was eluted using 3:1 of (hexane: ethyl acetate) as a red oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.41 (s, 1H), 7.34 (dt, *J* = 2.8, 1.8 Hz, 1H), 7.29 (dd, *J* = 8.2, 1.2 Hz, 1H), 7.05 (t, *J* = 8.0 Hz, 1H), 6.92 – 6.81 (m, 2H), 6.66 (td, *J* = 2.8, 1.6 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 144.0, 139.6, 131.5 (1*J*C-F = 254.5 Hz), 123.6, 120.6, 119.4, 118.9, 117.6, 117.4, 106.9, 105.9. IR (neat) 3410 3392 1654 1453 1231 1130 1082 880 767 723 712 cm^{-1}. HRMS (ESI) calcd. for $C_{11}H_8F_2NO_2$ m/z $(M+H)^+$: 224.0523, found: 224.0514.

3-chloro-4-(2,2-difluorobenzo[d][1,3]dioxol-4-yl)-1H-pyrrole (10): In a solution containing **10a** (0.5 mM), MgCl2 (10 mM), glucose (5.0 mM), FAD (1.0 μM), NT-11 PrnC (12.5 μM), Fre (2.5 μM) and Gdhi (2.5 μM) in 50mM potassium phosphate buffer, NADH (2.5 mM) was added to a total volume of 8.0 mL. After an overnight incubation of 30°C and orbital shaking at 350 rpm, reactions were quenched with an equivalent volume of MeOH, pelleted by centrifugation (15000 rpm for 10 min) and the supernatant extracted with (3 x 10mL) of ethyl acetate. The precipitated solids were vortex, sonicated and extracted with additional (3 x 10mL) of MeOH. The combined

organic layers were filtered through a plug of celite and dried over anhydrous sodium sulphate. Concentration over reduced pressure afforded a crude brown oil which was subjected over preparatory thin-layer chromatography using 3:1 of (hexane: ethyl acetate) to afford (0.6 mg, 58%) of **10** as a reddish-brown oil. ^{1}H NMR (400 MHz, Methanol-*d*4) δ 7.57 (dd, *J* = 8.1, 1.2 Hz, 1H), 7.14 (t, *J* = 8.1 Hz, 1H), 7.09 (d, *J* = 2.3 Hz, 1H), 7.02 (dd, *J* = 8.0, 1.1 Hz, 1H), 6.86 (d, *J* = 2.3 Hz, 1H). ^{13}C NMR (101 MHz, Methanol-*D*4) δ 144.9, 141.4, 132.7 (1*J*C-F = 251.6 Hz), 124.7, 124.3, 119.9, 119.6, 118.2, 114.4, 111.0, 107.8. HRMS *m/z* (ESI) cald for $C_{11}H_5ClF_2NO_2$ ([M+H]$^+$) 255.9982, found: 255.9989.

In conclusion, we have successfully enhanced the heterologously expressed of PrnC as soluble proteins by attaching an 11-amino acid long NT-11 tag. This modification enables the smooth scale up and easy isolation of the purified PrnC enzyme which has been a re-occurring issue with traditional standard protocols. Yields up to 9mg/L can be routinely obtained as compared to the <0.5 mg/L without the tag (>18-fold improvement) To the best of our knowledge, this is the first time the FAD-dependent halogenase PrnC has been isolated and applied to a mini library of free-standing pyrrolic analogs. This is a mild, green approach to access a unique C-3 halogenation of the pyrrolic substrate scaffold under aqueous buffer conditions. The bio halogenation reaction was found to be highly selective for the backbone position of the pyrrole fragment, with a broad tolerance for its appended aryl group. The key residues that are responsible for PrnC's catalytic activity have been identified and proposed by molecular docking studies and mutagenesis experiments. This ability to introduce a halide handle enables further late-stage functionalization of the resulting product to more complex molecules.

It will be appreciated that many further modifications and permutations of various aspects of the described embodiments are possible. Accordingly, the described aspects are intended to embrace all such alterations, modifications, and variations that fall within the spirit and scope of the appended claims.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

Throughout this specification and the claims which follow, unless the context requires otherwise, the phrase "consisting essentially of", and variations such as "consists essentially of" will be understood to indicate that the recited element(s) is/are essential i.e. necessary elements of the invention. The phrase allows for the presence of other non-recited elements which do not materially affect the characteristics of the invention but excludes additional unspecified elements which would affect the basic and novel characteristics of the method defined.

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

Claims

1. A method of halogenating a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof, comprising:

a) contacting the N-heteroaryl or derivative thereof with a halogen and a flavin-dependent halogenase under co-enzyme regeneration conditions in order for the halogen to be substituted at a 4' position on the N-heteroaryl or derivative thereof;

wherein the co-enzyme regeneration conditions comprises flavin adenine dinucleotide (FAD) at about 0.01 equivalents to about 0.2 equivalents relative to a flavin-dependent halogenase concentration;

E. coli flavin reductase (Fre) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration;

nicotinamide adenine dinucleotide hydrogen (NADH) at about 2 equivalents to about 250 equivalents relative to a flavin-dependent halogenase concentration;

glucose 1-dehydrogenase (GdHi) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration;

monosaccharide at about 5 equivalents to about 500 equivalents relative to a flavin-dependent halogenase concentration.

2. The method according to claim 1, wherein the co-enzyme regeneration conditions comprises NADH at about 0.05 equivalents to about 5 equivalents relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration.

3. The method according to claim 1 or 2, wherein the co-enzyme regeneration conditions comprises monosaccharide at about 5 equivalents to about 20 equivalents relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration.

4. The method according to any one of claims 1 to 3, wherein a ratio of FAD: Fre : GDH is about 1: 2.5 : 2.5.

5. The method according to any one of claims 1 to 4, wherein a ratio of FAD: Fre : NADH : GDH is about 1: 2.5 : 2500: 2.5.

6. The method according to any one of claims 1 to 5, wherein the co-enzyme

regeneration conditions further comprises a halogen at about 10 equivalents to about 30 equivalents relative to a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof concentration.

7. The method according to any one of claims 1 to 6, wherein the halogen is derived from a halide, the halide selected from Cl^- or Br^-.

8. The method according to any one of claims 1 to 7, wherein the co-enzyme regeneration conditions further comprises a phosphate buffer or a Tris HCl buffer.

9. The method according to any one of claims 1 to 8, wherein the flavin-dependent halogenases is monodechloroaminopyrrolnitrin halogenase (PrnC).

10. The method according to any one of claims 1 to 9, wherein the flavin-dependent halogenases comprises a N-terminal 11 amino acid solubility tag.

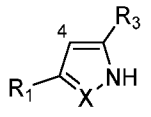
11. The method according to claim 10, wherein the N-terminal 11 amino acid solubility tag is derived from a first 11 amino acid residues within a N-terminal N-half domain of a duplicated carbonic anhydrase (dCA) from Dunaliella species.

12. The method according to claim 10 or 11, wherein the 11 amino acid solubility tag has an amino acid sequence of VSEPHDYNYEK.

13. The method according to any one of claims 1 to 12, wherein the N-heteroaryl or derivative thereof is a 5 membered N-heteroaryl or derivative thereof.

14. The method according to any one of claims 1 to 13, wherein the substituent on the 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof is each selected from optionally substituted aryl or optionally substituted heteroaryl.

15. The method according to any one of claims 1 to 14, wherein the N-heteroaryl or derivative thereof is a compound of formula (I)



(I)

wherein X is selected from CR_2 or N;

R_1 is H, optionally substituted aryl, or optionally substituted heteroaryl;

R_2 is H, optionally substituted aryl, or optionally substituted heteroaryl;

R_3 is H, optionally substituted aryl, or optionally substituted heteroaryl;

wherein at least one of R_1, R_2 and R_3 is optionally substituted aryl or optionally substituted heteroaryl;

or

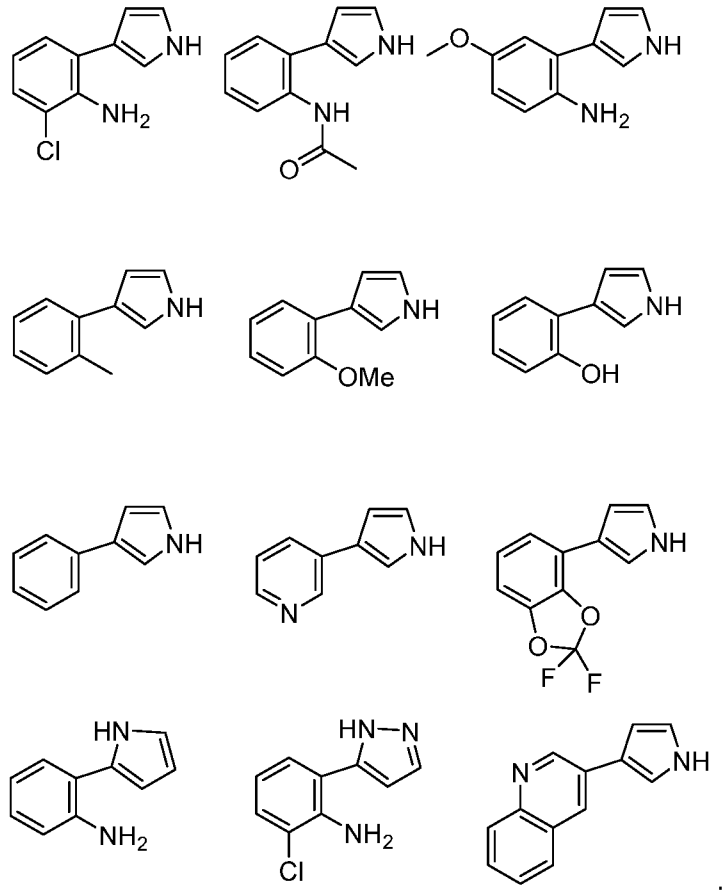
R_1 and R_2 are linked to form optionally substituted aryl, or optionally substituted heteroaryl.

16. The method according to claim 15, wherein when X is CR_2, R_3 is H.

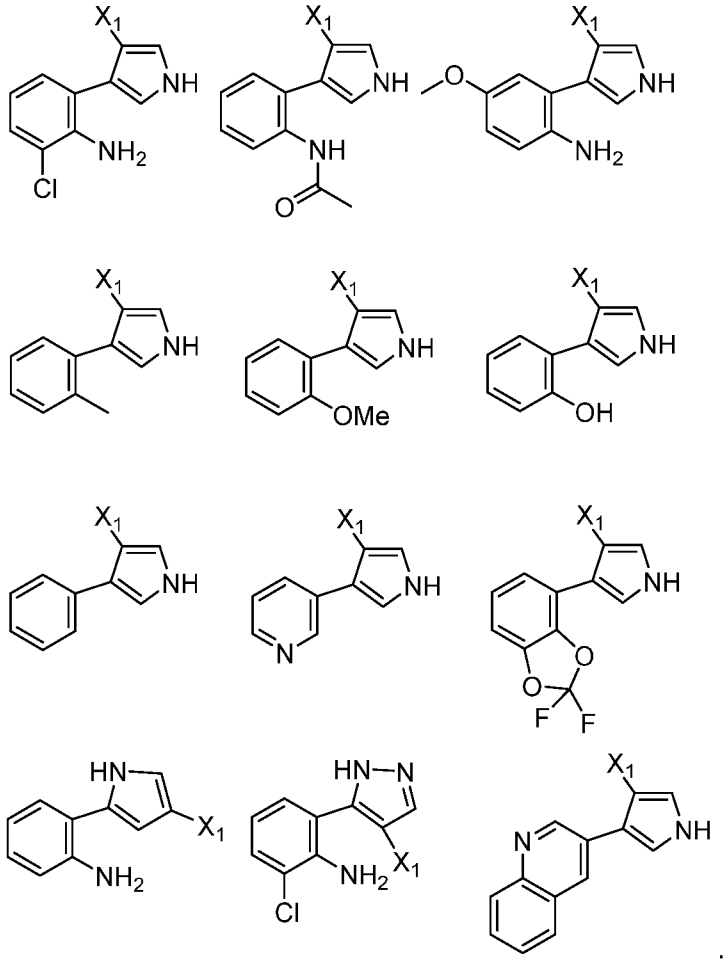
17. The method according to claim 15 or 16, wherein when X is CR_2, R_2 is H and R_3 is H.

18. The method according to any one of claims 15 to 17, wherein when X is N, R_1 is H.

19. The method according to any one of claims 1 to 18, wherein the N-heteroaryl or derivative thereof is selected from:



20. The method according to any one of claims 1 to 19, wherein the halogenated N-heteroaryl or derivative thereof is (where X_1 represents halo):



21. The method according to any one of claims 1 to 20, wherein the method is characterised by a regioisomeric ratio of 4' substitution to 2', 3' or 5' substitution is about 3:1 to about 1.1:1.

22. The method according to any one of claims 1 to 21, wherein the method is characterised by a ratio of 4' monohalogenation to 2', 4' dihalogenation, 3', 4' dihalogenation, and/or 4', 5' dihalogenation of about 20:1 to about 4:1.

23. The method according to any one of claims 1 to 22, wherein the method is characterised by a K_{cat} of about 7 x 10^{-3} s^{-1} to about 8 x 10^{-3} s^{-1}.

24. The method according to any one of claims 1 to 23, wherein the method is

characterised by a K_m of about 1 x 10^{-5} M to about 2 x 10^{-5} M.

25. The method according to any one of claims 1 to 24, wherein the method is characterised by a K_{cat}/K_m of about 4 x 10^2 s^{-1} M^{-1} to about 5 x 102 s^{-1} M^{-1}.

26. A method of producing a flavin-dependent halogenase, comprising:

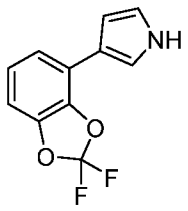
a) expressing the flavin-dependent halogenase in a cell;

wherein the flavin-dependent halogenase comprises an 11 amino acid solubility tag at a N-terminus thereof and a His6 tag at a C-terminus thereof.

27. The method according to claim 26, further comprising a step of purifying the flavin-dependent halogenase by metal affinity chromatography.

28. A method of synthesising Fludioxonil, comprising:

a) contacting a compound of formula (II) or derivative thereof with a halogen and a flavin-dependent halogenase under co-enzyme regeneration conditions in order for the halogen to be substituted at a 4' position on the compound or derivative thereof;



(II)

wherein the co-enzyme regeneration conditions comprises flavin adenine dinucleotide (FAD) at about 0.01 equivalents to about 0.2 equivalents relative to a flavin-dependent halogenase concentration;

E. coli flavin reductase (Fre) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration;

nicotinamide adenine dinucleotide hydrogen (NADH) at about 2 equivalents to about 250 equivalents relative to a flavin-dependent halogenase concentration;

glucose 1-dehydrogenase (GdHi) at about 0.1 equivalents to about 0.5 equivalents relative to a flavin-dependent halogenase concentration;

monosaccharide at about 5 equivalents to about 500 equivalents relative to a flavin-dependent halogenase concentration; and

b) contacting the 4' halogenated compound of step a) with a palladium precatalyst and a cyanide precursor.

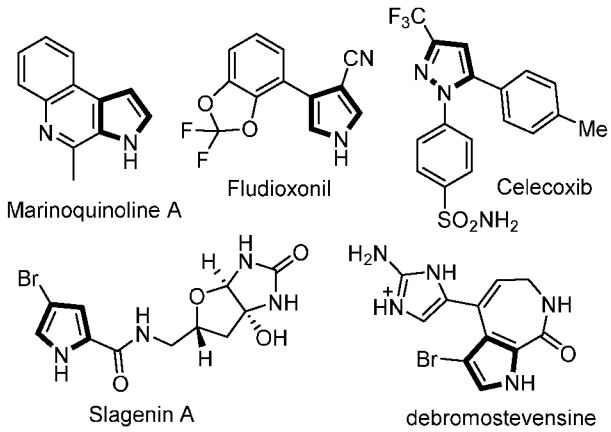
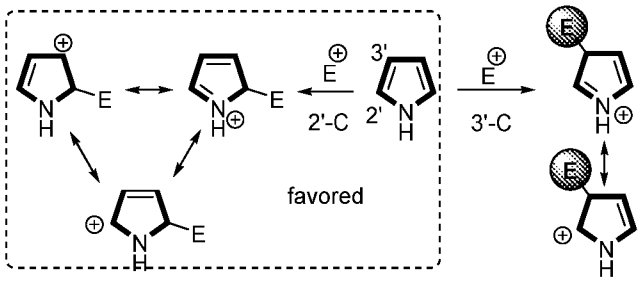


Figure 1

A. Preferred 2'-Carbon Site of Halogenation for Pyrroles



B. Site Specific 3'-Carbon halogenation via the PrnC Enzyme within the Pyrrolnitrin Biosynthesis

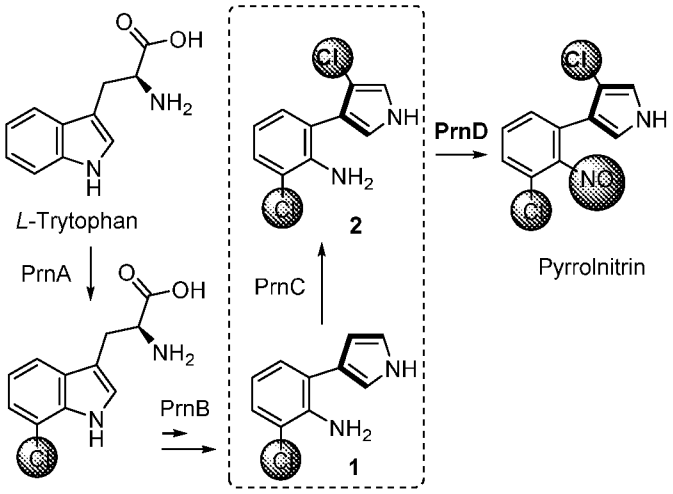


Figure 2A and 2B

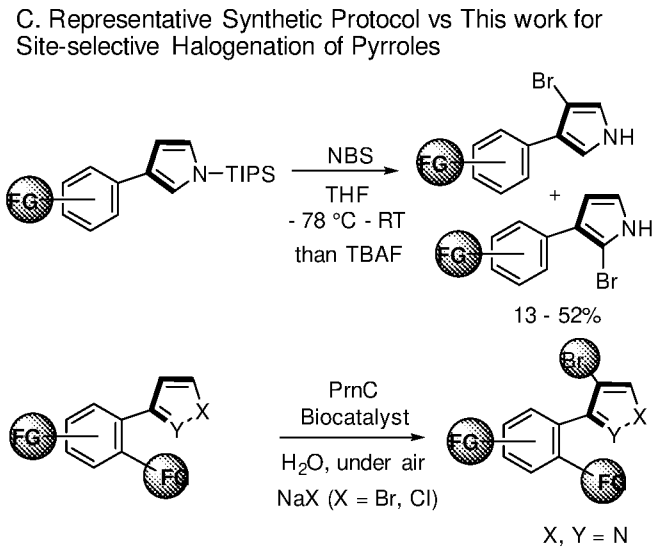


Figure 2C

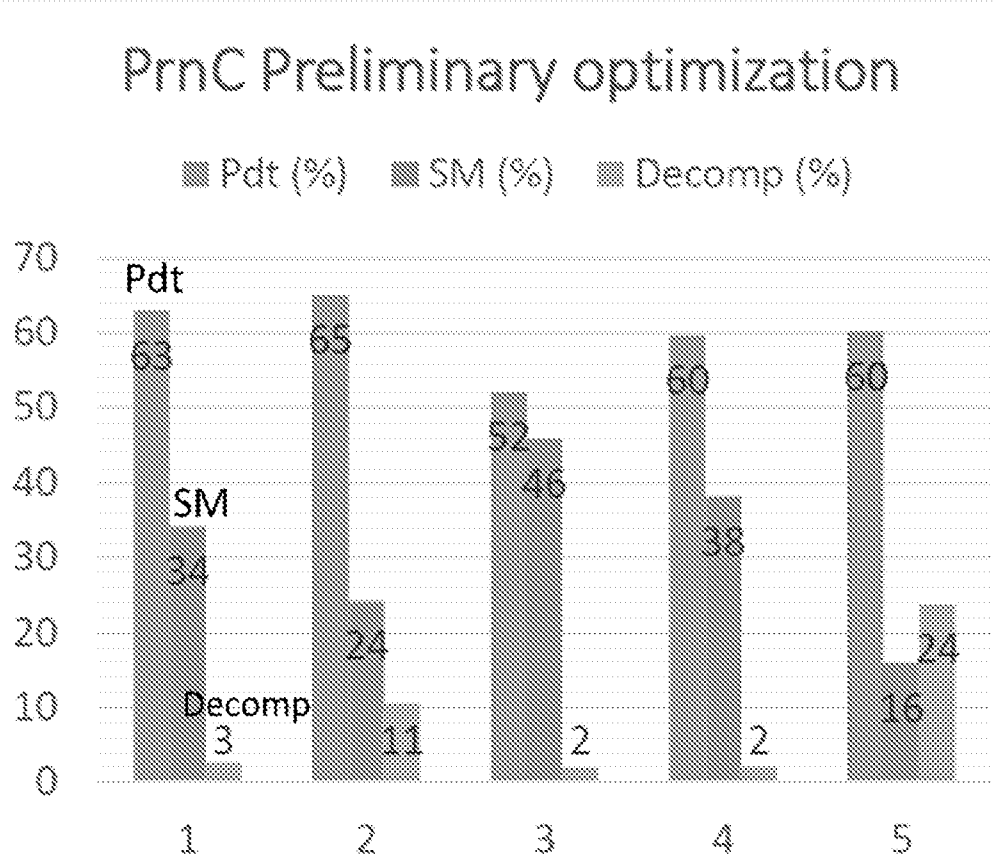


Figure 3A

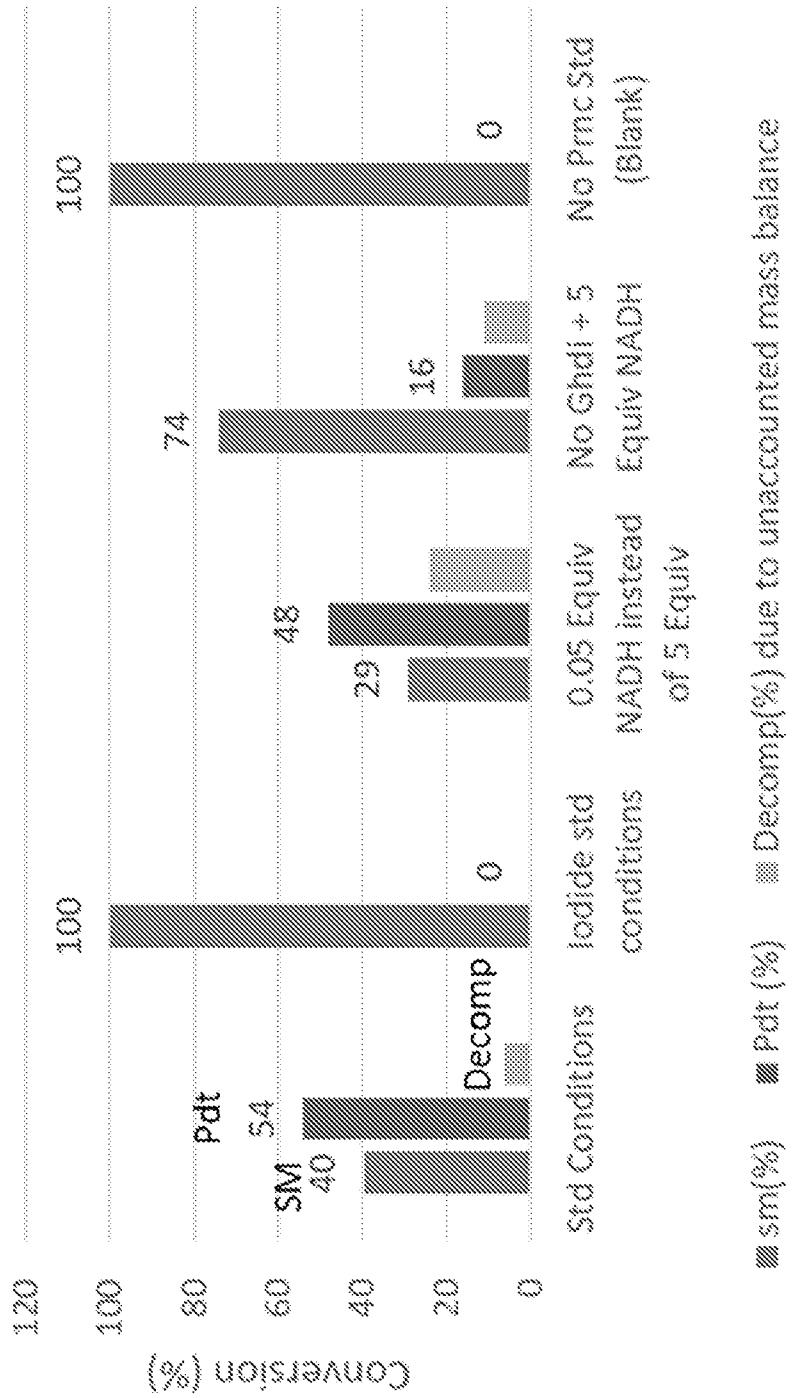
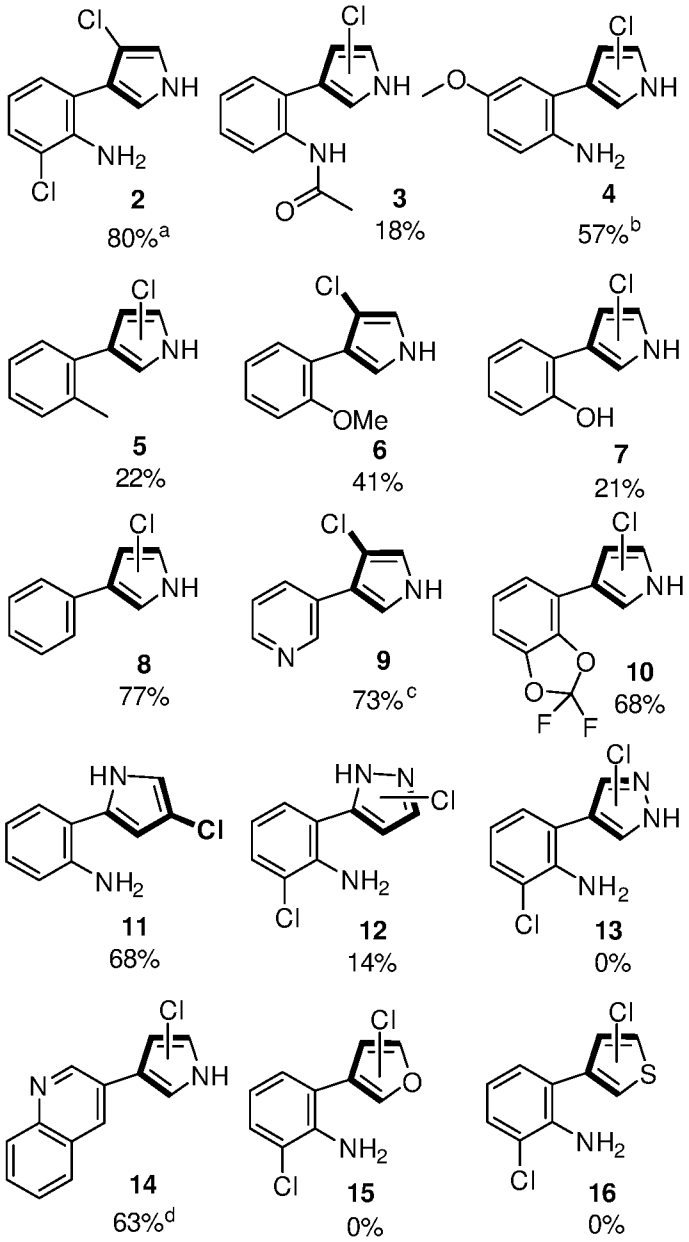

Preliminary Optimization Table 2
Conversion (%)
120
100
80
60
40
20
0
SM
Pdt
Decomp
40
54
100
0
29
48
74
16
100
0
Std Conditions
Iodide std conditions
0.05 Equiv NADH instead of 5 Equiv
No Ghdl + 5 Equiv NADH
No Prmc Std (Blank)
sm(%)
Pdt (%)
Decomp(%) due to unaccounted mass balance

Figure 3B



Conversion yields determined by integration of the HPLC peaks at 254 nm from LC-MS analyses of the crude reaction mixtures relative to control standard at 4 mol% loading of NT-11 PrnC. Duplicate data. [a]2 mol% loading. [b]Mono-Chlorinated: Di-Chlorinated mixture (7:1). [c]Mono-Cl : Di-Cl mixture (16:1). [d]Mono-Cl regioisomers (2:1)

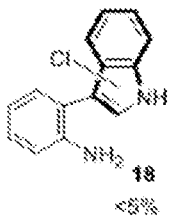


Figure 4

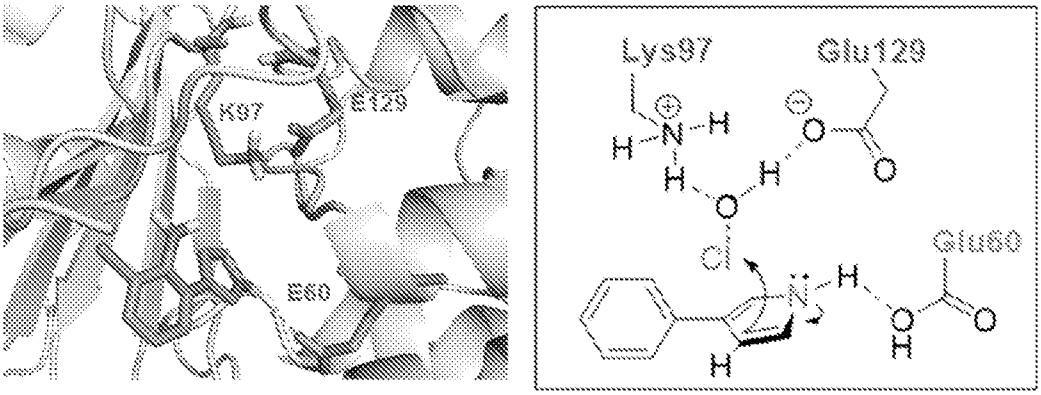


Figure 5

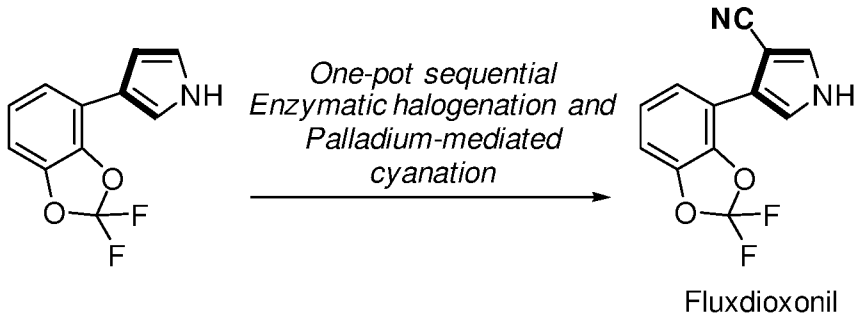


Figure 6

Telescoping enzymatic reaction with a subsequent Palladium catalyzed Cynation

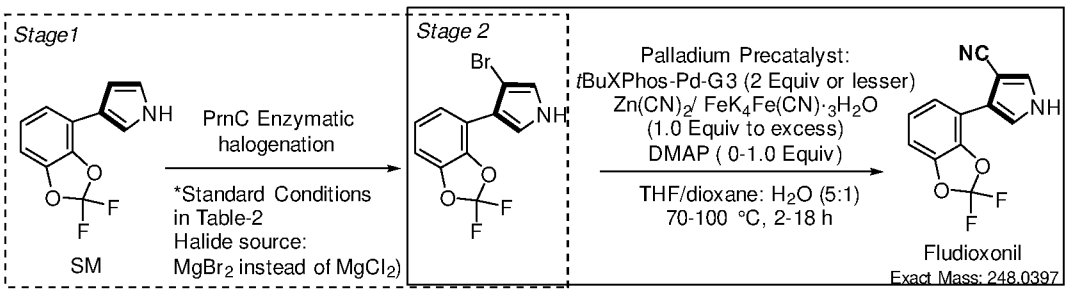
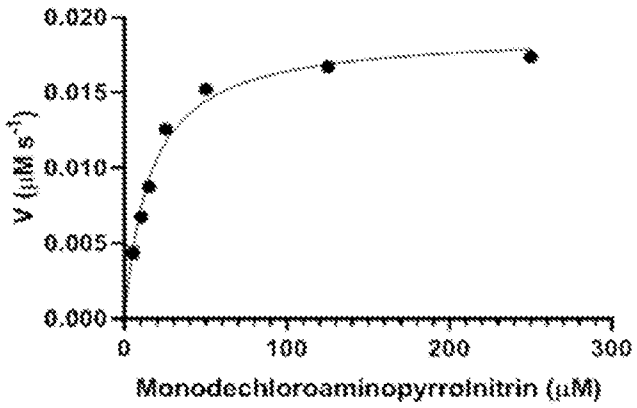
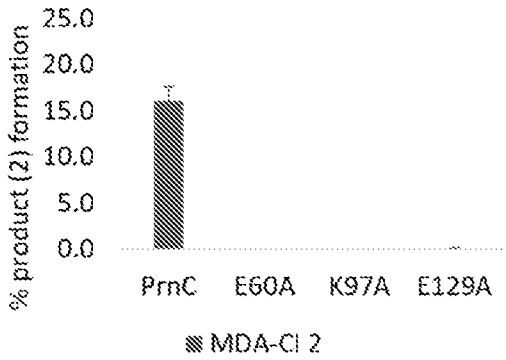


Figure 7



Kinetic parameters	Value
k_{cat} (s^{-1})	$7.60 \pm 0.08 \times 10^{-3}$
K_m (M)	$1.58 \pm 0.07 \times 10^{-5}$
k_{cat}/K_m ($s^{-1} M^{-1}$)	$4.83 \pm 0.24 \times 10^{2}$

Figure 8



PrnC mutants	% MDA-Cl (2)
Wild-type	16.01
E60A	N.D.
K97A	N.D.
E129A	0.07

Figure 9

```
PrnC   13   NHFDVIILGSGMSGTQMGAILAKQQFRVLIIEESSHPRFTIGESSIPETSLMNRIIADRY  72
6BZA   22   NQYDVIIIGSGIAGALTGAVLAKSGLNVLILDSAQHPRFSVGEAATPESGFLLRLLSKRF  81

PrnC   73   GIPELDHITSFYSTQRYVASST-GIKRNFGFVFHKPGQEHDPKEFTQCVIPELPWGPESH  131
6BZA   82   DIPEIAYLSHPDKIIQHVGSSACGIKLGFSFAWH---QENAPSSPDHLVAPPLKV-PEAH  137

PrnC   132  YYRQDVDAYLLQAAIKYGCKVHQKTTVTEYHADKDGVAVTTAQGERFTGRYMIDCGGPRA  191
6BZA   138  LFRQDIDYFALMIALKHGAESRQNIKIESISLNDDGVEVALSNAAPVKAAFIIDAAAQGS  197

PrnC   192  PLATKFKLREEPCRFKTHSRSLYTHMLGVKPFDDIFKVKGQR---WRWHEGTLHHMFEGG  248
6BZA   198  PLSRQLGLRTTEG-LATDTCSFFTHMLNVKSYEDALAPLSRTRSPIELFKSTLHHIFEEG  256

PrnC   249  WLWVIPFNNHPRSTNNLVSVGLQLDPRVYPKTDISAQQEFDEFLARFPSIGAQFRDAVPV  308
6BZA   257  WLWVIPFNNHPQGTNQLCSIGFQFNNAKYRPTE-APEIEFRKLLKKYPAIGEHFKDAVNA  315

PrnC   309  RDWVKTDRLQFSSNACVGDRYCLMLHANGFIDPLFSRGLENTAVTIHALAARLIKALRDD  368
6BZA   316  REWIYAPRINYRSVQNVGDRFCLLPQATGFIDPLFSRGLITTFESILRLAPKVLDAARSN  375

PrnC   369  DFSPERFEYIERLQQKLLDHNDDFVSCCYTAFSDFRLWDAFHRLWAVGTILGQFRLVQAH  428
6BZA   376  RWQREQFIEVERHCLNAVATNDQLVSCSYEAFSDFHLWNVWHRVWLSGSNLGSAFLQKLL  435

PrnC   429  ARFRASRNEGDLDHLDNDPPYLGYLCADMEEYYQLFNDAKAEVEAVSAGRKPADEAAARI  488
6BZA   436  HDLEHSGDARQFDAALEAVRFPGCLSLDSPAYESLFRQSCQVMQQAREQARPVAETANAL  495

PrnC   489  HALIDERD  496
6BZA   496  HELIKEHE  503
```

Figure 10

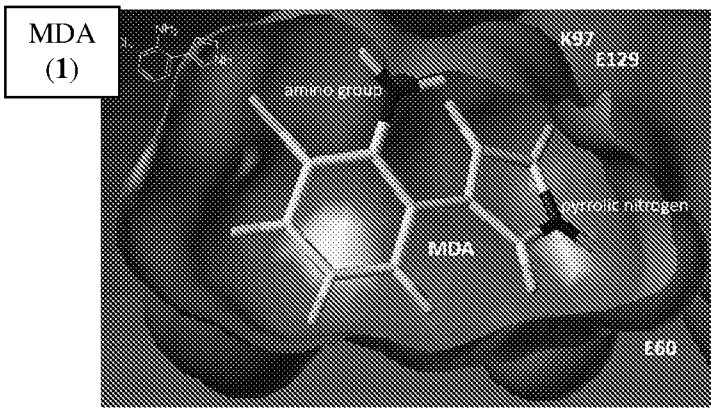


Figure 11

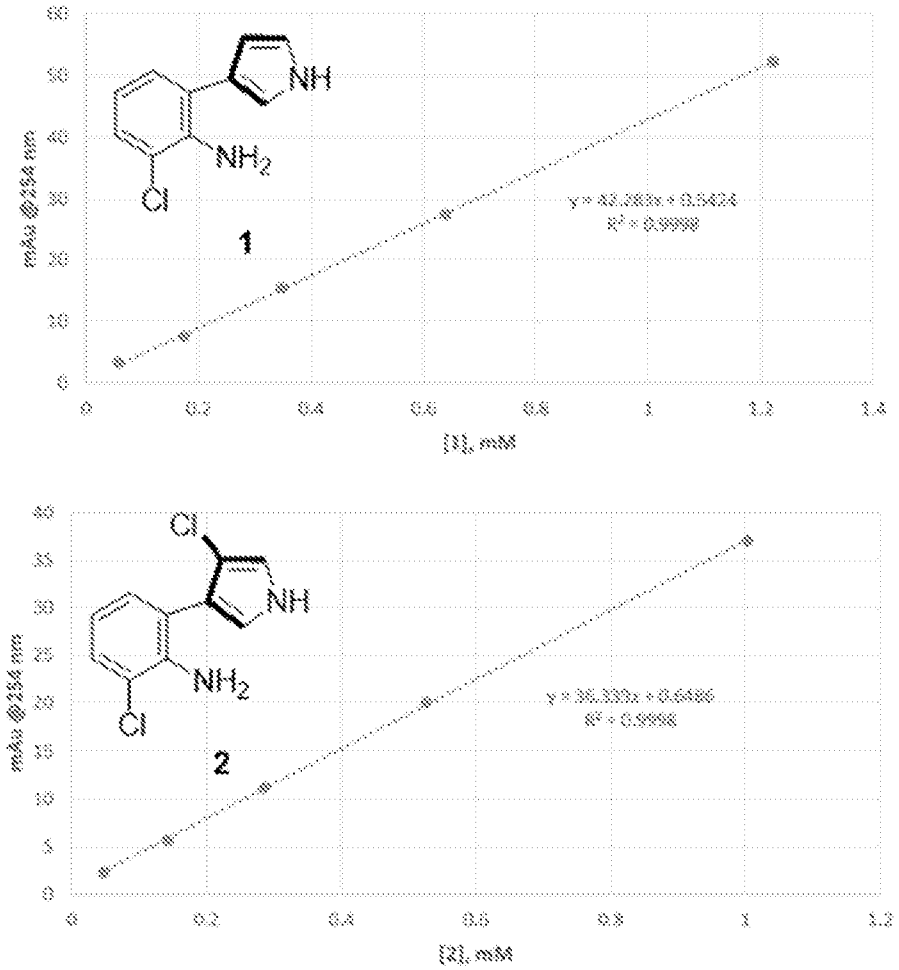

NH
NH2
Cl
1
mAu @254 nm
[1], mM
0
10
20
30
40
50
60
0
0.2
0.4
0.6
0.8
1
1.2
1.4
Cl
NH
NH2
Cl
2
mAu @254 nm
[2], mM
0
5
10
15
20
25
30
35
40
0
0.2
0.4
0.6
0.8
1
1.2

Figure 12

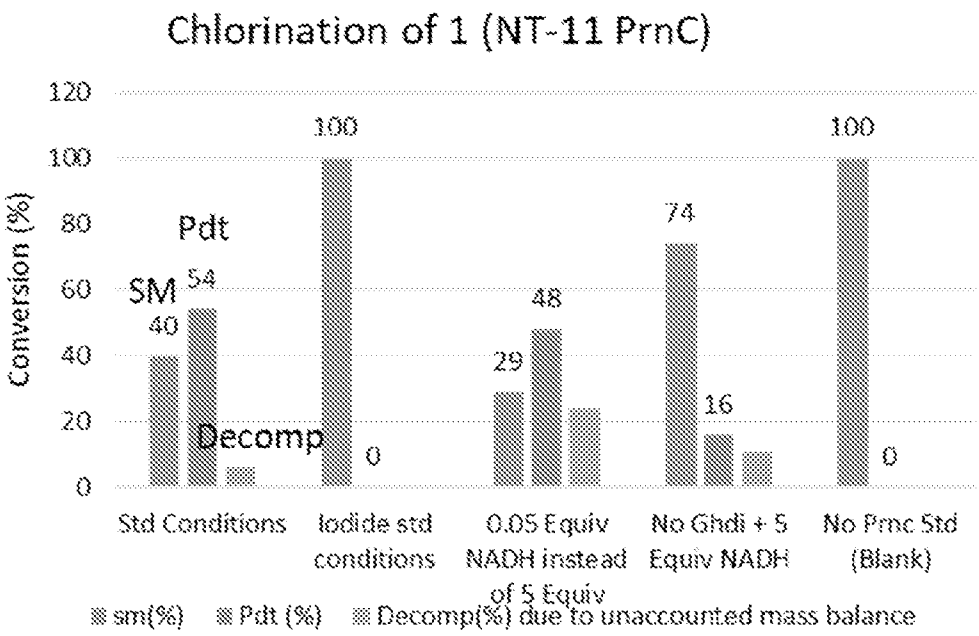

Chlorination of 1 (NT-11 PrnC)
Conversion (%)
120
100
80
60
40
20
0
SM
Pdt
Decomp
40
54
100
0
29
48
74
16
100
0
Std Conditions
Iodide std conditions
0.05 Equiv NADH instead of 5 Equiv
No Ghdi + 5 Equiv NADH
No Prnc Std (Blank)
sm(%)
Pdt (%)
Decomp(%) due to unaccounted mass balance

Figure 13

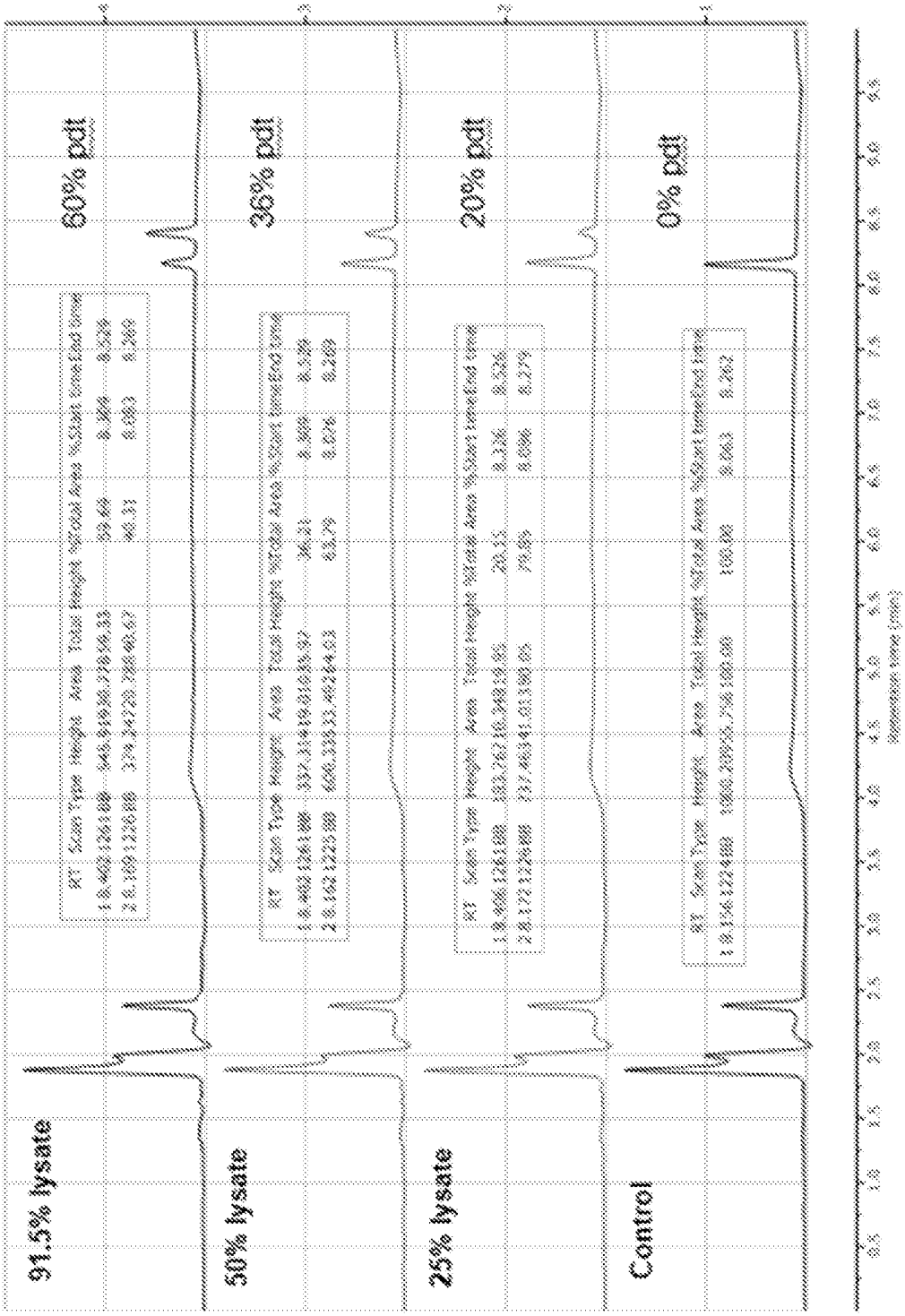

91.5% lysate
50% lysate
25% lysate
Control
60% pdt
36% pdt
20% pdt
0% pdt
Retention time (min)

Figure 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG2023/050347

A. CLASSIFICATION OF SUBJECT MATTER

C12N 9/02 (2006.01) C12N 15/70 (2006.01) C12P 13/00 (2006.01)

According to International Patent Classification (IPC)

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CAS REGISTRY, CAPLUS, CASREACT, FAMPAT and internet: enzymatic pyrrole/pyrazole halogenation, flavin-dependent halogenase, FDHs, monodechloroaminopyrrolnitrin halogenase, PrnC, pyrrole halogenases, 7-tryptophan halogenase, halB, RebH, chlorination, bromination, E. coli flavin reductase, Fre, reduced flavin adenine dinucleotide, FADH-, nicotinamide adenine dinucleotide hydrogen, NADH, glucose 1-dehydrogenase, GdHi, monosaccharide, glucose, NT11 soluble tag, Fludioxonil and related terms.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	MENON, N. ET AL., Versatile and Facile One-Pot Biosynthesis for Amides and Carboxylic Acids in *E. coli* by Engineering Auxin Pathways of Plant Microbiomes. *ACS Catalysis*, 31 January 2022, Vol. 12, No. 4, pages 2309-2319 [Retrieved on 2023-11-01] <DOI: 10.1021/ACSCATAL.1C04901> Figure 4, Table 3; page 7 section 1.6 of the supporting information	1-9, 13-25 10-12
X	WO 01/44447 A1 (SYNGENTA PARTICIPATIONS AG) 21 June 2001 page 14 the first full paragraph, examples 3 and 5	1-9, 13-25
X	XU, F. ET AL., Selective biochlorination of hydroxyquinolines by a flavin-dependent halogenase. *Tetrahedron Letters*, 1 January 2016, Vol. 57, No. 00404039, pages 5262-5265 [Retrieved on 2023-11-01] <DOI: 10.1016/J.TETLET.2016.10.044> Results and discussion, Figure 4; In vitro chlorination reactions of the supporting information	1-9, 13-25

☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

*Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
01/11/2023 *(day/month/year)*	04/11/2023 *(day/month/year)*

Name and mailing address of the ISA/SG	Authorized officer
Intellectual Property Office of Singapore 1 Paya Lebar Link, #11-03 PLQ 1, Paya Lebar Quarter Singapore 408533	Chen Wei (Dr)
Email: pct@ipos.gov.sg	IPOS Customer Service Tel. No.: (+65) 6339 8616

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG2023/050347

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☐ forming part of the international application as filed.

   b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)),

      ☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

   Although the sequence listing was not used for the purposes of search, the sequences mentioned in the claims have been taken into consideration for the purposes of this opinion.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG2023/050347

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:

   because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:

   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:

   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please refer to Supplemental Box (Continuation of Box No. III).

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

   1-25, 28

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG2023/050347

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MENON, B.R.K. ET AL., Structure and biocatalytic scope of thermophilic flavin-dependent halogenase and flavin reductase enzymes. *Organic & Biomolecular Chemistry*, 1 January 2016, Vol. 14, No. 39, pages 9354-9361 [Retrieved on 2023-11-01] <DOI: 10.1039/C6OB01861K> Fig 1-3	1-9, 13-25
Y	NGUYEN, T.K.M. ET AL., The NT11, a novel fusion tag for enhancing protein expression in Escherichia coli. *Applied Microbiology and Biotechnology*, 4 January 2019, Vol. 103, No. 5, pages 2205-2216 [Retrieved on 2023-11-01] <DOI: 10.1007/S00253-018-09595-W> The whole document	10-12
A	CN 111285854 A (BEIJING NUTRICHEM CO LTD) 16 June 2020 the whole document of the original non-English language document (a machine translation is enclosed **only** for your reference)	28
A	BÜCHLER, J. ET AL., Recent Advances in Flavin-Dependent Halogenase Biocatalysis: Sourcing, Engineering, and Application. *Catalysts*, 5 December 2019, Vol. 9, No. 12, 1030 [Retrieved on 2023-11-01] <DOI: 10.3390/CATAL9121030> Figure 4	-
A	FISHER, B.F. ET AL., Site-Selective C–H Halogenation Using Flavin-Dependent Halogenases Identified via Family-Wide Activity Profiling. *ACS Central Science*, 24 October 2019, Vol. 5, No. 11, pages 1844-1856 [Retrieved on 2023-11-01] <DOI: 10.1021/ACSCENTSCI.9B00835> Figures 6 and 7; supporting information: pages S36 -S43	-
A	GKOTSI, D.S. ET AL., A marine viral halogenase that iodinates diverse substrates. *Nature Chemistry*, 14 October 2019, Vol. 11, No. 12, pages 1091-1097 [Retrieved on 2023-11-01] <DOI: 10.1038/S41557-019-0349-Z> Fig 2	-
P,X	PEH, G. ET AL., Site-selective Chlorination of Pyrrolic Heterocycles by Flavin Dependent Enzyme PrnC. *ChenRxiv*, 1 November 2022 [Retrieved on 2023-10-30] <DOI: 10.26434/CHEMRXIV-2022-T1MRN> Scheme 1, Figure 5, Table 1, the last paragraph; Table S1 and pages S3-S5 section Method for PrnC mutants construction in mutagenesis studies in the supporting information	1-25, 28
P,A	JIANG, Y. ET AL., The Single-Component Flavin Reductase/Flavin-Dependent Halogenase AetF is a Versatile Catalyst for Selective Bromination and Iodination of Arenes and Olefins. *Angew. Chem. Int. Ed.*, 25 October 2022, Vol. 61, No. 51, e202214610 [Retrieved on 2023-10-30] <DOI: 10.1002/ANIE.202214610> Figures 1-3	1-25

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/SG2023/050347

Note: This Annex lists known patent family members relating to the patent documents cited in this International Search Report. This Authority is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2001/044447 A1	21/06/2001	EP 1238062 A1	11/09/2002
		AU 1707801 A	25/06/2001
		AR 026939 A1	05/03/2003
		CA 2393910 A1	21/06/2001
		EA 200200634 A1	26/12/2002
		JP 2003516749 A	20/05/2003
		MX PA02005868 A	23/10/2002
		HU 0203807 A2	28/02/2003
		CN 1409759 A	09/04/2003
		BR 0017024 A	07/01/2003
CN 111285854 A	16/06/2020	NONE	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG2023/050347

Supplemental Box
(Continuation of Box No. III)

This International Searching Authority found multiple inventions in this international application, as follows:

This International Application does not comply with the requirements of unity of invention under PCT Rule 13.1 because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept.

PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art. When there is no special technical feature common to all the claimed inventions there is no unity of invention.

This Authority has found that there are different inventions based on the following features that separate the claims into distinct groups:

Group 1: Claims 1-25 and 28 are directed to a method of halogenating a 2', 3' and/or 5' substituted N-heteroaryl or a derivative thereof with a halogen and a flavin-dependent halogenase under co-enzyme regeneration conditions.

Group 2: Claims 26-27 are directed to a method of producing a flavin-dependent halogenase.

In the above groups of claims, the only feature common to all the claimed inventions and which provides a technical relationship among them is a flavin-dependent halogenase. However, this feature is well-known in the art. Therefore, there is no special technical feature common to all the claimed inventions and the requirements for unity of invention are consequently not satisfied *a priori*.

Please refer to **Box No. IV** of Written Opinion of The International Searching Authority (Form PCT/ISA/237) for detailed explanation.