Supplementary Information (SI) for Chemical Communications.
This journal is © The Royal Society of Chemistry 2026

Ancestral Sequence Reconstruction Reveals New Functional Fluorinases and Mechanistic Insights into Enzymatic Fluorination

Shreyas Supekar[a‡*], Wan Lin Yeo[b‡] , Elaine Tiong[c‡], Juliana Rizal[d], Ee Lui Ang [b,e*], Fong Tian Wong[c,d*], Yee Hwee Lim[c,e*], Hao Fan[a,e,f*]

[a] Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, #07-01 Matrix Building, Singapore 138671, Republic of Singapore.
[b] Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), 31 Biopolis Way, Nanos, Singapore 138669, Republic of Singapore.
[c] Institute of Sustainability for Chemicals, Energy and Environment (ISCE^2), Agency for Science, Technology and Research (A*STAR), 8 Biomedical Grove, #07-01 Neuros Building, Singapore 138665, Republic of Singapore.
[d] Institute of Molecular and Cell Biology (IMCB), Agency for Science, Technology and Research (A*STAR), 61 Biopolis Drive, Proteos #07-06, Singapore, 138673, Republic of Singapore.
[e] Synthetic Biology Translational Research Program, Yong Loo Lin School of Medicine, National University of Singapore, 10 Medical Drive, Singapore 117597, Republic of Singapore.
[f] Institute for Molecular and Cellular Therapeutics, Chinese Institutes for Medical Research, Beijing 100069, China

‡ Equal contributions

*Corresponding authors

Materials and methods

Chemicals
All chemicals were purchased from Sigma-Aldrich and used as received unless otherwise stated. S-Adenosyl-L-methionine (SAM) was purchased from Santa Cruz Biotechnology. 5'-fluoro-5'-deoxyadenosine (5'-FDA) was synthesized as previously mentioned.[1]

Construction of expression plasmids for expression of fluorinases in *E. coli*
Codon-optimized DNA sequences for *Escherichia coli* (*E. coli*) were synthesized by Twist Bioscience. The inserts were assembled into the pET28a(+) expression vector via Golden Gate cloning and initially transformed into *E. coli* OmniMAX competent cells (Thermo Fisher Scientific) for plasmid propagation. The verified constructs were subsequently transformed into *E. coli* XJb (Zymo Research) for protein expression.

New variants construction
Site-directed mutagenesis was performed on the FL135 construct to generate new protein variants. Mutation-specific primers were synthesized by Integrated DNA Technologies (IDT) (Table S9). The mutagenesis reactions were carried out using FL135 as the plasmid template. Mutations were verified by nanopore sequencing of the plasmids prior to downstream protein expression.

High throughput (HTP) growth and lysis
The plasmids harbouring the fluorinase genes were transformed into *E. coli* BL21(DE3) ΔPNP strain,[1] 84 colonies of the various constructs, 6 colonies of FLA_MA37 (as positive control),[2] and 6 colonies of pET28a (as negative control) were inoculated into a 96-well plate. Individual colony was inoculated in 180 µL LB medium containing kanamycin (50 µg/mL) and cultured overnight at 37°C. 20 µL of cultures were transferred to 380 µL 2xYT medium containing 50 µg/mL kanamycin and grown until an absorbance of 0.5 at 600 nm was reached. Overexpression was induced by adding 0.2 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG) and incubation was continued at 22°C for 20 h. Cells were harvested by centrifugation (2,182 x *g* for 10 min at 4°C). The cell pellets were lysed with 250 µL lysis buffer (50 mM sodium phosphate buffer, pH 7.8, 10% glycerol, 0.1% Triton-X) at room temperature for 2 h at speed 6.5 on titer plate shaker (Barnstead). The lysates were obtained by centrifugation (2,182 x *g* for 20 min at 4°C).

Expression and purification of fluorinases
A single colony from the transformed *E. coli* XJb plate was inoculated into 5 mL of LB medium supplemented with 50 μg/mL kanamycin and cultured overnight at 37℃ with shaking at 250 rpm. The overnight culture was used to inoculate 500 mL of fresh LB medium containing 50 μg/mL kanamycin and grown at 37℃, 250 rpm until the optical density at 600 nm (OD_{600}) reached 0.6. Protein expression was induced by the addition of IPTG to a final concentration of 0.2 mM. and the culture was incubated overnight at 22℃ with shaking at 250 rpm. Cells were harvested by centrifugation and resuspended in binding buffer (50 mM sodium phosphate, pH 7.8; 300 mM NaCl; 5 mM imidazole; 10% v/v glycerol). The cell suspension was lysed by sonication, and the lysate was clarified by centrifugation for 1 h at 4 °C. The supernatant was incubated with HisPur Ni-NTA resin (Thermo Fisher Scientific) overnight at 4℃. The resin was then washed with 40 mL of binding buffer. Bound proteins were eluted using elution buffer (50 mM sodium phosphate, pH 7.8; 300 mM NaCl; 250 mM imidazole; 10% v/v glycerol). The eluate was buffer-exchanged and concentrated using Amicon Ultra centrifugal filters (Merck) and stored in 50 mM sodium phosphate buffer (pH 7.8) containing 10% v/v glycerol at -80℃. The protein concentration was measured using a NanoDrop 2000 spectrophotometer (Thermo Scientific). The extinction coefficient was determined using the ExPAsy ProtParam tool.

Enzymatic reaction
Enzymatic reactions were conducted at two different scales: HTP 96-well plates and 1.5 mL tube formats.

HTP enzymatic reaction conditions (plates) were performed with SAM (1 mM), NaF or NaCl (200 mM), and fluorinase lysates (182 μL) in a final volume of 260 μL. HTP enzymatic reaction was carried out in a 96-well deepwell plate at 37℃ for 4 h and reaction was stopped by quenching with 1:1 volume of methanol. The quenched reaction mixtures were centrifuged (2,182 x g for 20 min) and 10 μL was used for HPLC-UV analysis.

Enzymatic reaction conditions (tubes) were performed with SAM (1 mM), NaF or NaCl (200 mM), and fluorinase purified proteins (20 μM) in a final volume of 100 μL. Purified proteins reactions were carried out in triplicates in 1.5 mL tubes in shaking incubator at 250 rpm at 37℃ for 1.5 h and 24 h. Reactions were stopped by heating the samples at 95℃ for 1 min (using heatblock). The precipitated protein was then removed by centrifugation (20,238 x g for 2 min). 10 μL of the reaction mixture was used for HPLC-UV analysis.

HPLC analysis
HPLC analysis was performed using an Agilent 1200 Infinity series system coupled with a CTC Analytics HTS PAL autosampler. Mobile phase A: 0.1% formic acid in water; B: 0.1% formic acid in methanol, Phenomenex, Kinetex® 2.6 μm Biphenyl 100 Å, LC Column 150 x 4.6 mm, 0.6 mL/min flow rate, isocratic elution, 23% B for 15 min. The retention time of 5'-FDA is approximately 3.75 min. The retention time of 5'-ClDA is approximately 5.53 min.

Protein thermal shift assay
Protein thermal shift assay conditions: 0.25 μg/μL Purified proteins, 1x GloMelt dye, 20 μL total volume, SYBR Green channel, 25℃ for 1 min, melting curve of 25℃ – 95℃ with increment of 0.1℃/s, 95℃ for 1.5 min. Protein thermal shift assay were carried out using Biorad CFX96™ Real-Time System and melting temperatures were obtained from analysis using its respective software.

Dynamics of key catalytic components

Variants ancestral_16, M2, M3, M12, and M23 exhibit elevated SAM and SBS dynamics (Figs S3–S8; Table S8), likely contributing to reduced activity. To discern the individual contributions of mutations at positions 50 (SBS), 157 (IBS), and 252 (SCL), we examined key intermolecular distances: (1) W50/F50 to SAM purine ring, (2) Y157/F157 to fluoride ion, and (3) P252/R252 to SAM purine ring (Fig S9). FLA_MA37 maintains the most native-like contact geometry, closely followed by ancestral_22 and M13. Position 157 (Y/F, IBS) exhibits minimal conformational variability across all variants, likely due to the chemical similarity of tyrosine and phenylalanine. In contrast, position 50 (W/F, SBS) shows divergent behaviour: variants containing the F50W mutation (M1, M12) fail to establish native-like SAM contacts, whereas M13 and M123 restore proper SAM positioning, explaining their improved activities.

In simulations of variants containing the R252P mutation (M3, M13, M23, M123), the R252 sidechain reorients away from SAM (Fig S9, bottom panel) and forms a strong cation-π interaction with Y232 (Fig S9-S11). This interaction destabilizes the local β-sheet structure spanning residues 232–252 (Fig S12). We hypothesize that the P252R mutation increases backbone flexibility in the SAM capping loop, potentially disrupting the configurations necessary for maintaining the catalytically competent active site geometry. Secondary structure quantification shows a significant loss of β-sheet content in this region for R252-containing variants (Fig S12), consistent with this hypothesis.

Single mutants M1 (F50W), M2 (F157Y), and M3 (R252P) all fail to rescue activity despite addressing individual defects. M1 exhibits unfavourable positioning of catalytic residues relative to SAM/F^- (Figure 4), while M2 and M3 show elevated SBS flexibility and high SAM fluctuations (Table S8; Figures S3-S4), indicating that single-site corrections cannot overcome the complex interplay of catalytic defects in ancestral_16. Among double mutants, M13 uniquely succeeds by simultaneously addressing two critical factors: the F50W mutation provides stronger π-stacking with SAM while R252P removes the disruptive R252-Y232 cation-π interaction (Figures S9-S12), restoring native-like SAM positioning and substantially reducing water permeation to levels comparable with FLA_MA37 and ancestral_22 (Table 2). In contrast, M12 retains the unfavourable residue-SAM contact geometry despite moderate water exclusion (Figure S9; Table 2), while M23 exhibits high water permeation (78%) and SAM fluctuations despite improved SCL structure (Tables 2, S8). Remarkably, triple mutant M123 shows native-like protein and cofactor dynamics (Figure 4; Table S8; Figures S3-S9) but exhibits the highest water penetration at the IBS (Table 2), demonstrating that active site dehydration is the dominant factor controlling fluorination activity, that even optimal protein dynamics cannot overcome.

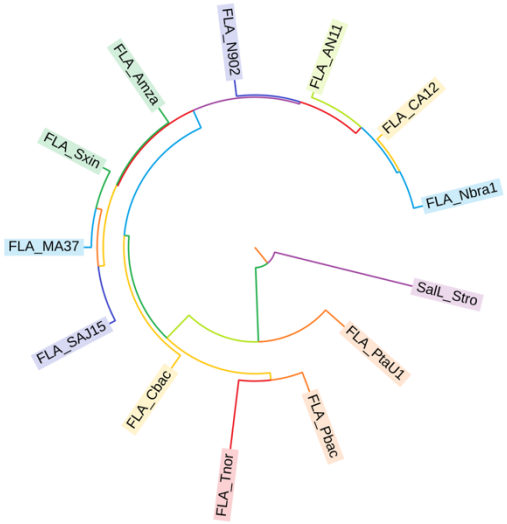

Figure S1. Phylogenetic tree of the candidate sequences emplyed for ancestral sequence reconstruction.

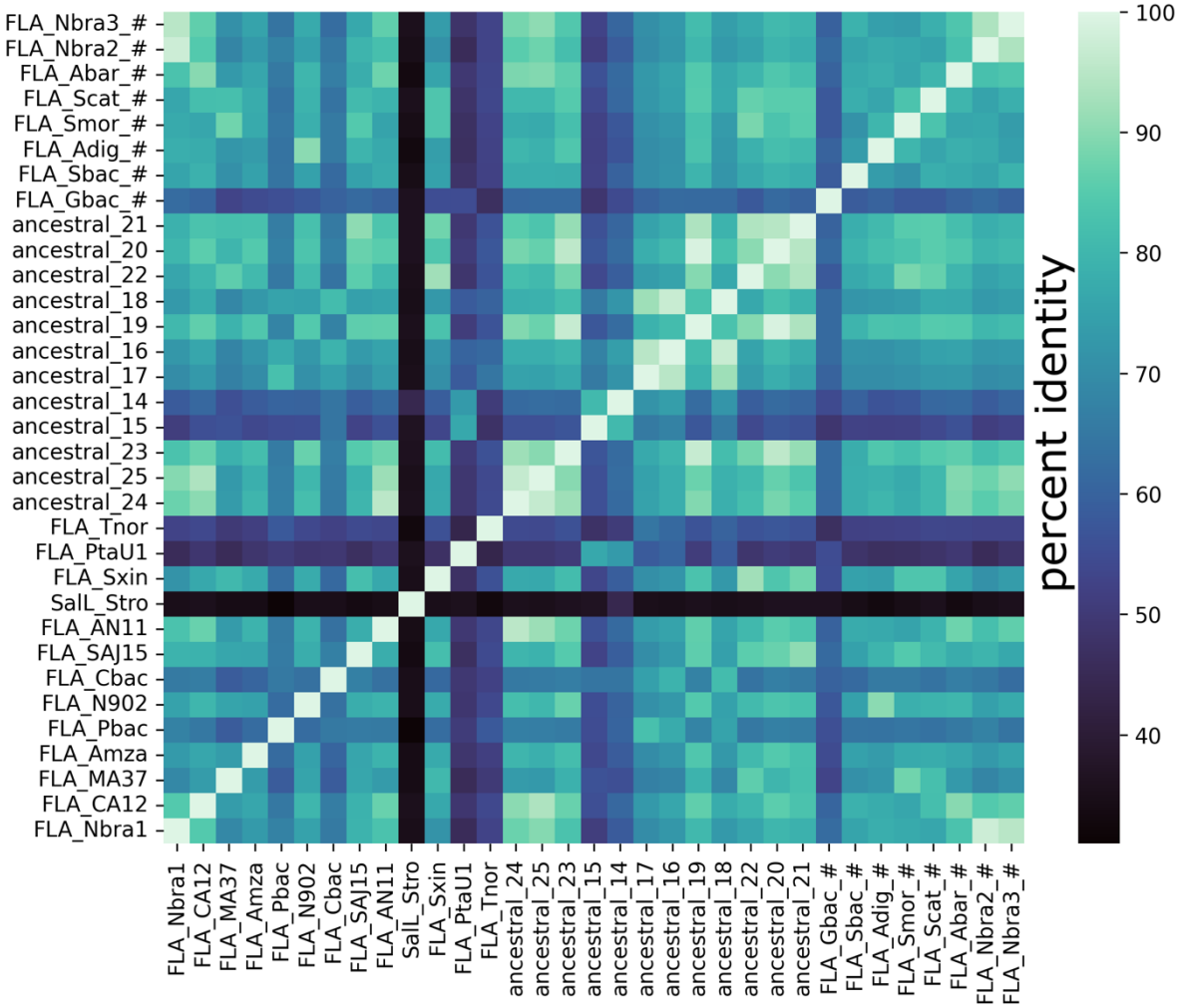

Figure S2. sequence identity matrix of reconstructed fluorinases, candidate fluorinases and other newly identified fluorinases (suffix: #) that are not in candidate set employed here for ancestral reconstruction. See Table S3 for details.

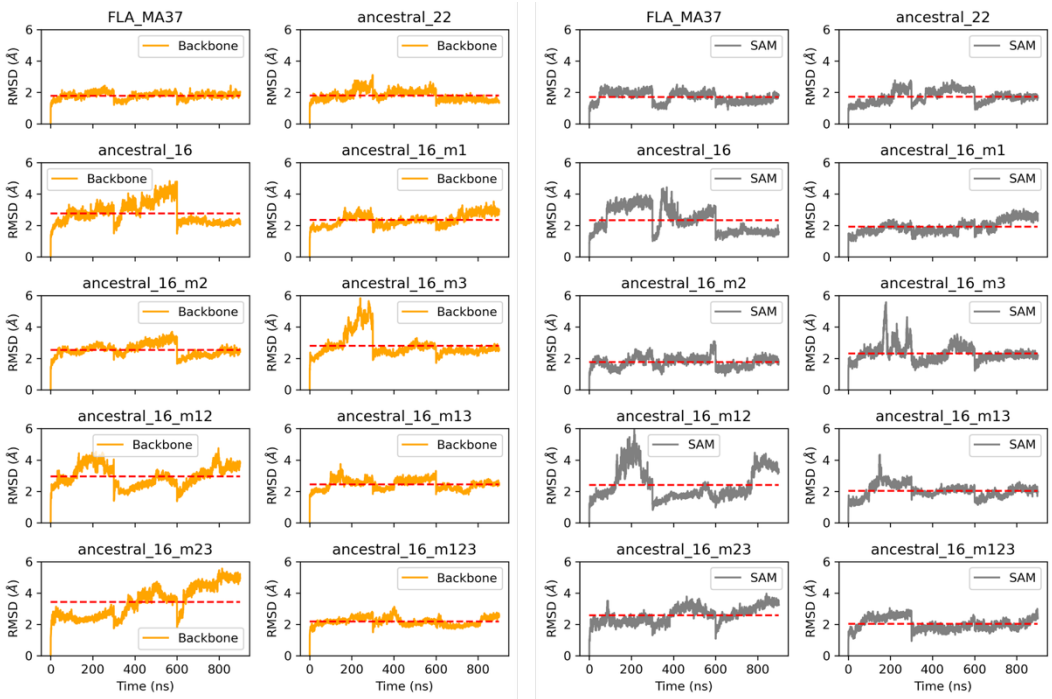

Figure S3. Root-mean-squared-deviation (RMSD) for backbone atoms and SAM cofactor from MD simulations of FLA_MA37, ancestral_22, ancestral_16 and its single, double and triple mutants. Dotted red lines indicate mean RMSDs.

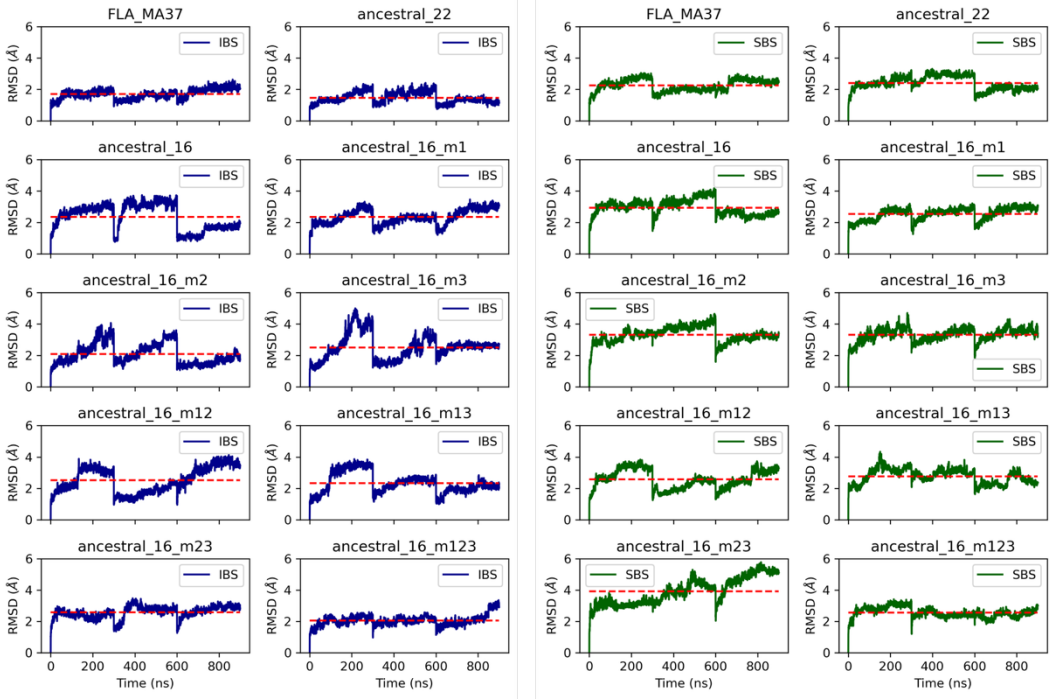

Figure S4. Root-mean-squared-deviation (RMSD) for IBS residues and SBS residues from MD simulations of FLA_MA37, ancestral_22, ancestral_16 and its single, double and triple mutants. Dotted red lines indicate mean RMSDs.

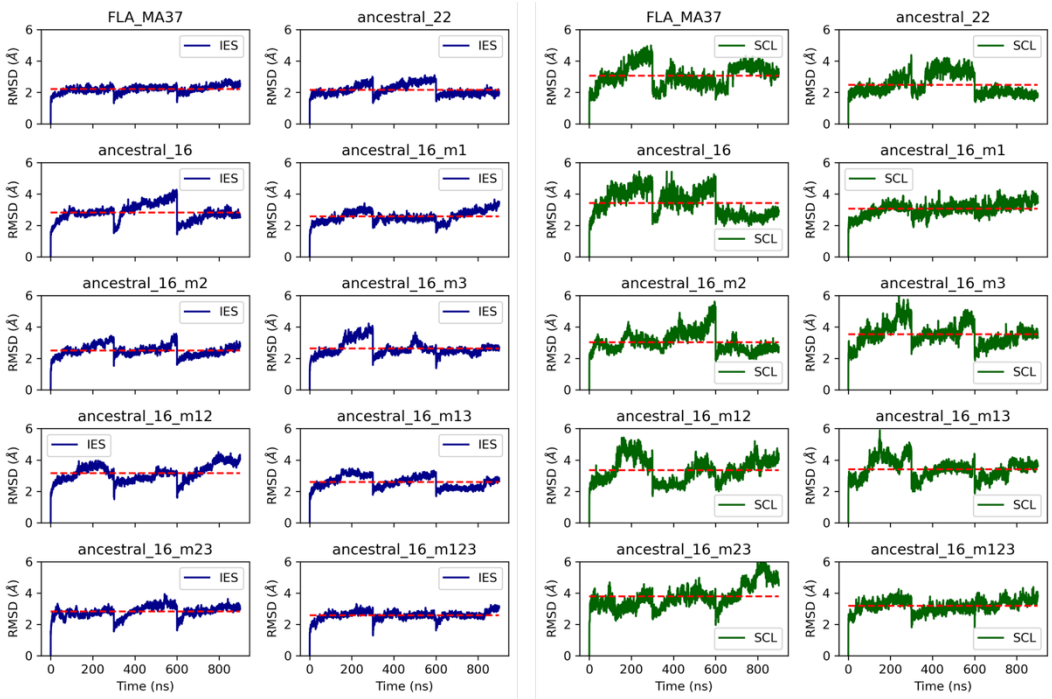

Figure S5. Root-mean-squared-deviation (RMSD) for IES residues and SCL residues from MD simulations of FLA_MA37, ancestral_22, ancestral_16 and its single, double and triple mutants. Dotted red lines indicate mean RMSDs.

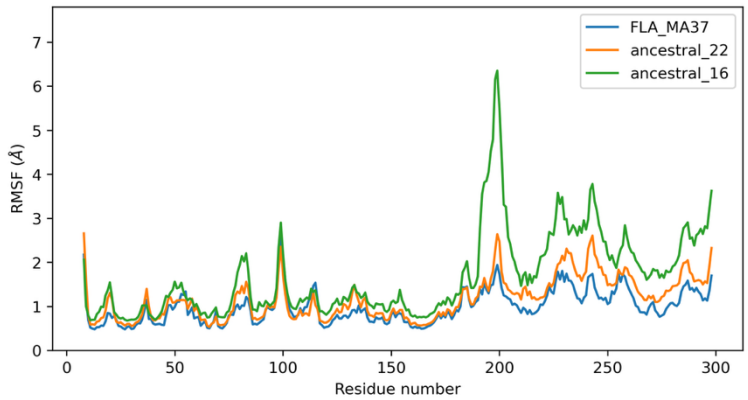

Figure S6 Root-mean-squared-fluctuations (RMSF) for FLA_MA37, high-performing ancestral_22 and low-performing ancestral_16 fluorinases.

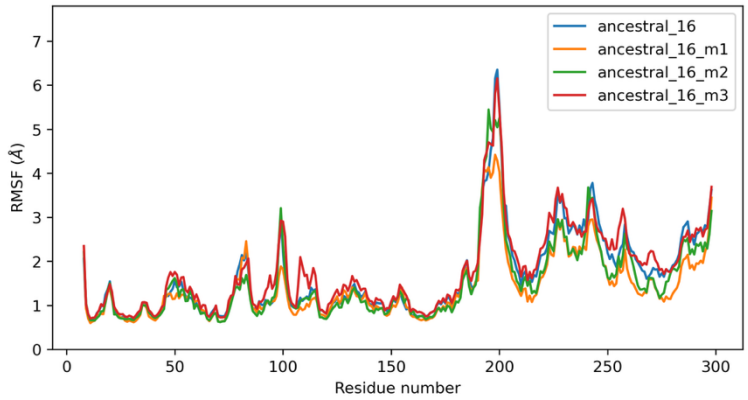

Figure S7 Root-mean-squared-fluctuations (RMSF) for ancestral_16 and ancestral_16 single mutants.

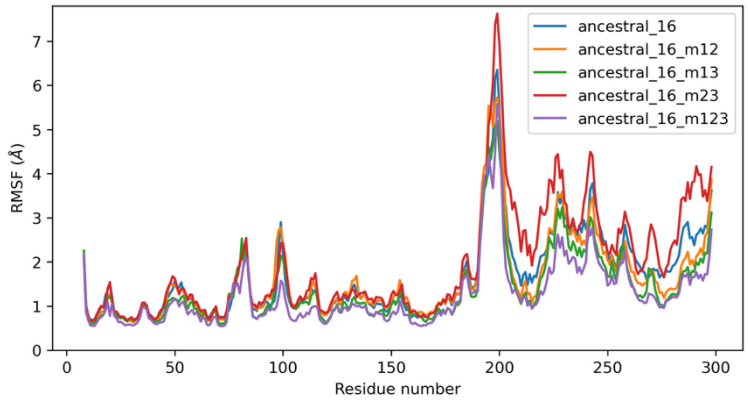

Figure S8 Root-mean-squared-fluctuations (RMSF) for ancestral_16 and ancestral_16 double and triple mutants.

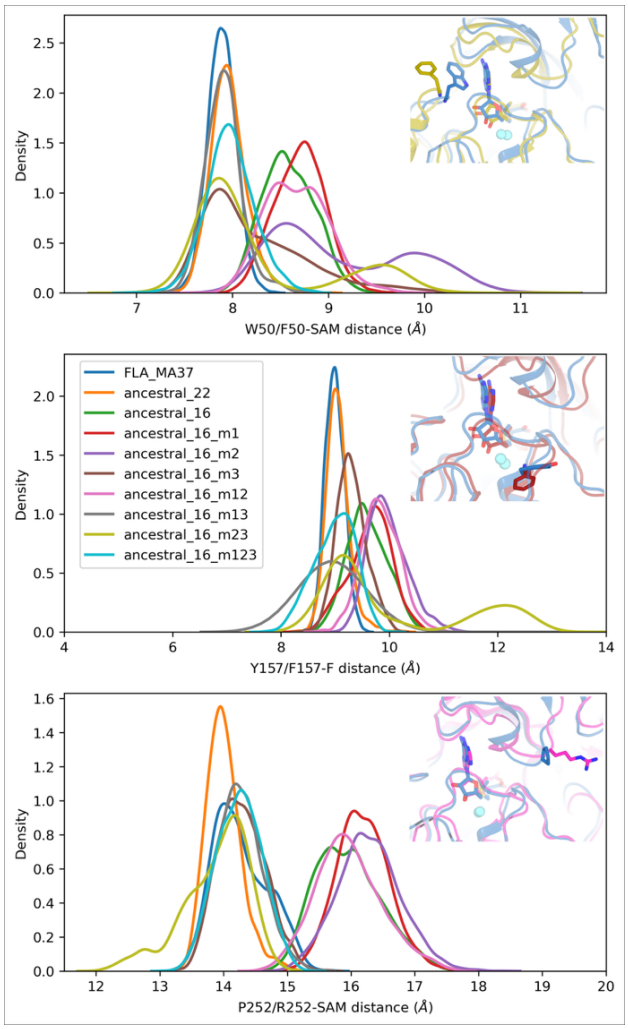

Figure S9. Distributions of distances between residues 50, 157 and 252 to SAM/F from MD simulations. Insets: representative averaged structures from MD trajectories showing interacting components, with FLA_MA37 in blue and mutants M23, M1 and M12 in olive, red and pink, respectively.

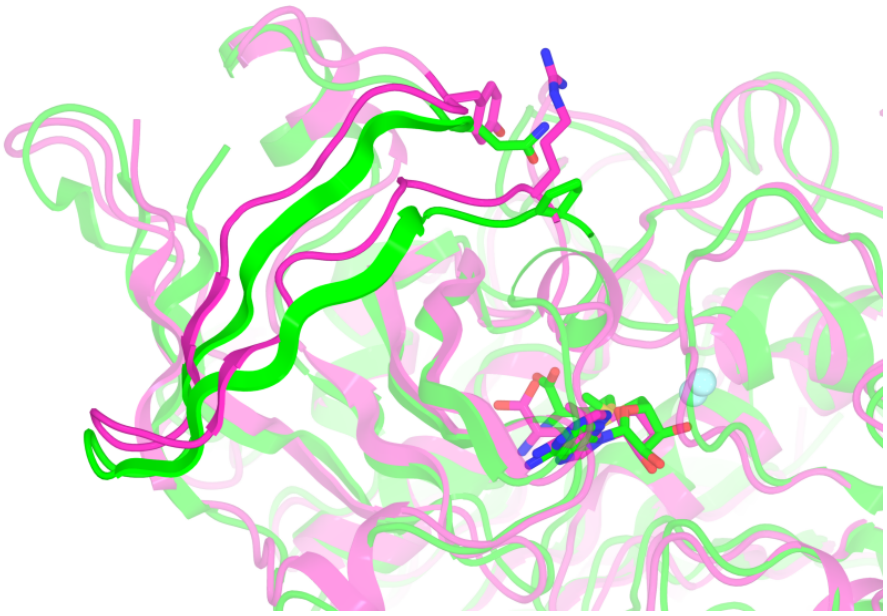

Figure S10. Representative structures of FLA_MA37 (P252, Q232; green) and ancestral_16_M12 (R252, Y232; magenta) from MD simulations showing R252-Y232 cation- π interaction and a partial loss of secondary structure in the residue range 232-252 for M12 variant (and other variants with R252).

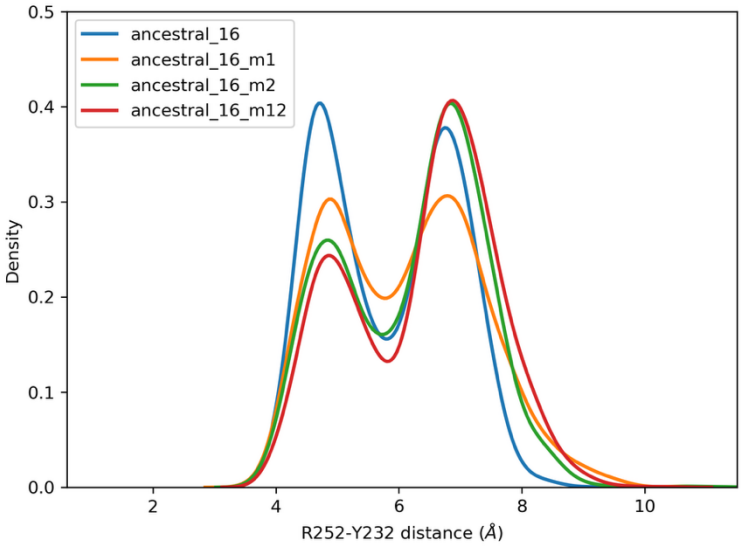

Figure S11 Distance between R252 guanidinium group and Y232 ring (cation- π interaction frequency) for ancestral_16 variants with arginine at position 252. Variants with R252 exhibit a strong interaction with R252-Y232 distance < 6Å for *ca.* 50% of the respective MD simulation frames.

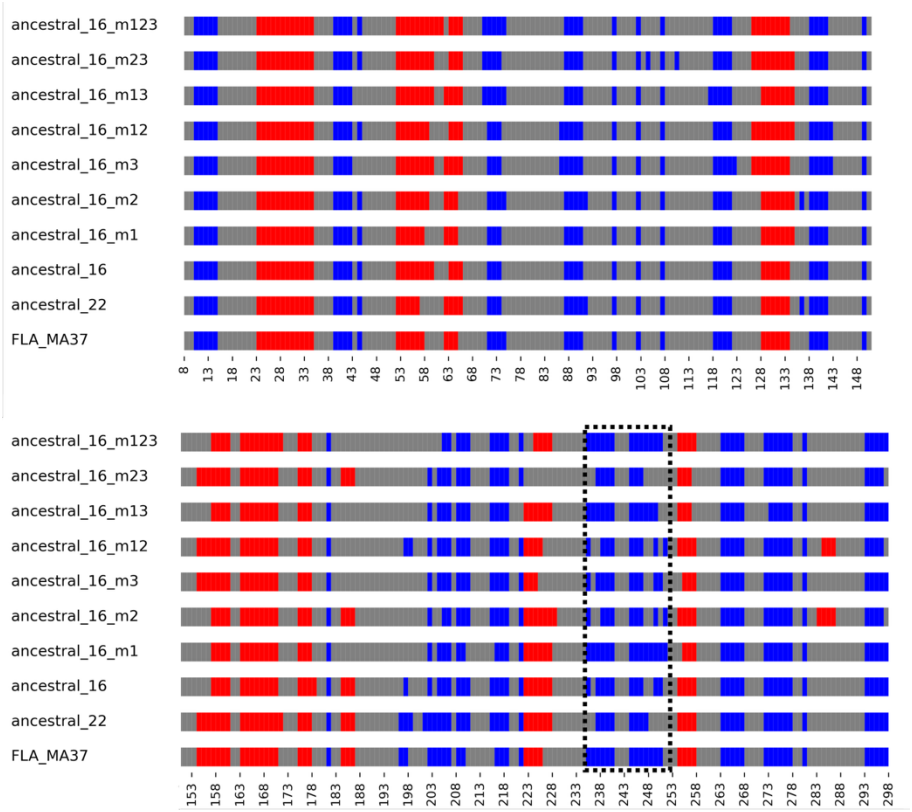

Figure S12 Averaged secondary structure (dssp) for each residue from MD simulations of FLA_MA37, ancestral_22, ancestral_16 and its single, double and triple mutants. Red, blue and gray indicate helices, β-strands and loops, respectively. Highlighted box shows β -sheet disruption in the region comprising residues 232-252 for variants containing P252R substitution.

Table S1. Known fluorinases with their respective amino acid sequences.

Name	Sequence
FLA_Gbac	MSDLGTVDDSVGICKGLMLSICPDATIVDICHAMTPWDIKQGARLIVDLPRYYPEWTVFATTTYPDTGTAMRSVAIRVPKGQVYV APNNGLLTSVIEDHGFVEAYEVTSTEVIPEEPEPTFFSREMVAVPSAHLAAGFPLDQVGRPLKDEEIVRFENQHPVVVEDGALRGIL TNVDWPYGNVWVNIRQSVLEANGIGYGTKVKITIDDVLPFELQLTRTFGDVPLGAPVTYLSSRGYLGIARNRANLADTYNLRQG MSVTVKAC
FLA_Sbac	MTANGSRRPTIAFMSDLGVTDDSVAQCKGLMLSICPDVNIIDICHTMTPFDVVEGARYIVDLPRYFPEGTVFATTTYPATGTTARSV AIRLKRAALGGARGQLAGSGKGFERAEGAYIYIAPNNGLLTRVIEEHGYLEAYEVSSTDVIPERPEPTFYSREMVAIPSAHLAAGFP LEKVGRKLEDHEIVRFEQKKGAVAGEALVGEVSAIDHPFGNVWTNLHRTDLEKAGIKYGTPMKIVVDGVLPFELPLSPTFADAEG PGAPVAYLNSRGYLSVARNAASLAYPYNLNAGMSVRVTTA
FLA_Adig	MSDLGTTDDSVAQCKGLMLSICPSVTIVDVCHSMTPWDVVEGARYIVDLPRFFPEGTVFATTTYPATGTGTRSVALRIKQAALGGA RGQWAGSGAGLERAEGSYIYIAPNNGLLTTVIEEHGYIEAYEVSNTKVIPAEPEPTFYSREMVAIPSAHLAAGFPLNEVGRQLSDDE IVRFERPKASTVSGGVLSGTITNVDHPFGNLWTNIHRTDLEKAGIGYQTQLKLVLDGVLTFDLPLVPTFADAGKIGDPVIYINSRGY LALARNAAPLAYPYNLKAGISVAVTKA
FLA_Smor	MSDLGTTDDSVAQCKGLMHSICPGVTVVDVCHSMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTTTRSVAVRIRQAAKGG ARGQWAGSGDGFERADGSYIYIAPNNGLLTTVLEEHGYIEAYEVTSTKVIPANPEPTFYSREMVAIPSAHLAAGFPLAEVGRRLDD SEIVRFHRPAVEISGDTLSGVVTAIDHPFGNIWTNIHRVDLEKAGIGHGKHLRIILDDVLPFEAPLTPTFADAGAIGNIAFYLNSRGYL SLARNAASLAYPYNLKAGLKVRVETR
FLA_MA37	MAANGSQRPIIAFMSDLGTTDDSVAQCKGLMHSICPGVTVVDVCHSMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTTTR SVAVRIRQAAKGGARGQWAGSGDGFERADGSYIYIAPNNGLLTTVLEEHGYIEAYEVTSTKVIPANPEPTFYSREMVAIPSAHLAA GFPLAEVGRRLDDSEIVRFHRPAVEISGEALSGVVTAIDHPFGNIWTNIHRTDLEKAGIGQGKHLKIILDDVLPFEAPLTPTFADAGAI GNIAFYLNSRGYLSLARNAASLAYPYNLKAGLKVRVEAR
FLA_Scat	MAANSTRRPIIAFMSDLGTTDDSVAQCKGLMYSICPDVTVVDVCHSMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTTTRS VAVRIKQAAKGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTTVLEEHGYLEAYEVTSPKVIPEQPEPTFYSREMVAIPSAHLAAG FPLSEVGRPLEDHEIVRFNRPAVEQDGEALVGVVSAIDHPFGNVWTNIHRTDLEKAGIGYGARLRLTLDGVLPFEAPLTPTFADAGE IGNIAIYLNSRGYLSIARNAASLAYPYHLKEGMSARVEAR
FLA_Sxin	MSADPTQRPIIGFMSDLGTTDDSVAQCKGLMHSICPGVTVIDVCHSMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTETRS VAVRIKQAAKGGARGQWAGSAGGFERAEGSYIYVAPNNGLLTTVLEEHGYIEAYEVSSTKVIPERPEPTFYSREMVAIPAAHLAAG FPLSEVGRPLEDSEIVRYQPPQVEISGDTLTGVVSAIDHPFGNVWTNIHRTHLEKAGIGYGKRIKIILDDVLPFEQTLVPTFADAGEIG GVAAYLNSRGYLSLARNAASLAYPFNLKAGLKVRVETN
FLA_SAJ15	MSDLGTTDDSVAQCKGLMLSICPDVTIVDVCHSMTPFDVEEGSRYIVDLPRFFPEGTVFATTTYPATGTETRSVAVRIRQAAQGGA RGQWAGSGAGFERQEGSYIYIAPNNGLLTPVLEEHGYIEAYEVSSTEVIPERPEPTFYSREMVAIPSAHLAAGFPLDQVGRPLKDSEI VRFSRPAVETAGTELTGVVSAIDHPFGNIWTNIHRTDLEKAGIGYGRQIRITLDDVLPFELTLVPTFADAGEVGNVVAYLNSRGYLSL ARNAASLAYPYNLKAGLSVRVDAH
FLA_N902	MPANGNPIIAFMSDLGTTDDSVAQCKGLMLSICPGVTIVDVNHSMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTATRSVAL RIKQAAQGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTTVIEEHGYIEAYEVSNTKVIPAEPEPTFYSREMVAIPSAHLAAGFPLN EVGRALSDDEIVRFAKPKPSTVSGGVLSGVITNIDHPFGNLWTNIHRTDLEKAGIGYQTQLRLLLDGVLTFDLPLVPTFADAGQIGD PVIYINSRGYLALARNAAPLAYPYNLKAGLTVTVTKA
FLA_Amza	MSDSYSRPIIAFMSDLGTTDDSVAQCKGLMMSICQDVTVVDVCHSMEPWNVEEGARYIVDLPRFFPEGTVFATTTYPATGTTARS VAVRIKYPAKGGARGQWAGSGEGFERSEGSYIYIAPNNGLLTTVLQEHGYTEAYEVSSTDVVPARPEPTFYSREMVAIPSAHLAAG YPLEKVGRKLQDSEIVRFTPPQATVSPEGDLSGVVTAIDHPFGNIWTSIHRDNLESAGVGYGTNLKIVLDDVFPFELPLSPTFADAG EVGDPVVYVNSRGYLSLARNAASLAYPYNLKEGMSVRVTRS
FLA_Abar	MATTKRPVLAFMSDLGITDDSVAQCKGLMLSICPDVTIVDVCHTMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTGTRSVA LRIKQAAKGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTSVIEEHGYLEAYEVSSTEVIPEQPEPTFYSREMVAVPSAHLAAGFPL EKVGRKLADDEIVRFDRPKPVQDEDGDLVGVVTTIDHPFGNVWTNIHREDLEKAGVGYGTELTITLDEVLPFRLPLSPTFADAGPI GSPVAYLSSRGYLALARNAASLAYPYNIQAGIPVRVHVA
FLA_CA12	MAKPSRPIIAFMSDLGITDDSVAQCKGLMLSVCPDVTIVDVCHTMKPWDVEEGARYIVDLPRLFPEGTVFATTTYPATGTTTRSVA LRIKQAAKGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTSVIEEHGYVEAYEVSSTEVIPEQPEPTFYSREMVALPSAHLAAGFP LEKVGRPLADDEIVRFERAKPAQNDDGELVGVVTAIDHPFGNVWTNIHREDLEKLGAGYGTRLRITLDEVLPFDLPLSPTFADAGP IGTPVAYLSSRGYLALARNAASLAYPYNLNAGISVRVVAA
FLA_AN11	MASSRGNRPIIAFMSDLGITDDSVAQCKGLMLSVCPDVTIVDVCHTMKPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTTARS VAIRIKQAAKGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTSVIEEHGYLEAYEVSSTEVIPELPEPTFYSREMVALPSGHLAAGF PLNKVGRPLADDEIVRFTRPQPSVETDGDLVGVVTNIDHPFGNIWTNIHRTDLERAGVGYGTKLKVVLDEVLPFELPLSPTFADAG PVGTPVAYLSSRGYLALARNAASLAYPYNLEAGIPVRVKVG
FLA_Nbra2	MSDLGITDDSVAQCKGLMLSVCPDVTIVDICHTMQPWDVEEGARYIVDLPRLFPEGTVFATTTYPATGTTARSVALRIAHASKGGA RGQWAGSGAGFERKEGSYIYIAPNNGLLTTVIKEHGYLEAYEVSSPEVIPEQPEPTFYSREMVALPSAHLAAGFPLEKVGRRLADD EIVRFERKDPELVADNELLGYVTNIDHPFGNVWTNIHRTDLEKLGVGYGTELRITLDGVLPFELPLSPTFADAGEVGAAVAYLSSRG YLALARNAASLAYPYNLNAGISVQVKVD
FLA_Nbra3	MTTANGRRPIIAFMSDLGITDDSVAQCKGLMLSVCPDVTIVDICHTMQPWDVEEGARYIVDLPRLFPEGTVFATTTYPATGTTARS VALRIAHASKGGARGQWAGSGAGFERKEGSYIYIAPNNGLLTTVIKEHGYLEAYEVSSPEVIPEQPEPTFYSREMVALPSAHLAAG FPLEKVGRRLADDEIVRFERKDPELVADTDLVGYVTNIDHPFGNVWTNIHRTDLEKLGVGYGTKLRITLDGVLPFELPLSPTFADA GEVGAAVAYLSSRGYLALARNAASLAYPYNLKAGISVQVKVG
FLA_Nbra1	MTTTNGRRPIIAFMSDLGITDDSVAQCKGLMLSVCPDVTIVDICHTMQPWDVEEGARYIVDLPRLFPEGTVFATTTYPATGTTARSV ALRIAHASKGGARGQWAGSGAGFERKEGSYIYIAPNNGLLTTVIKEHGYLEAYEVSSPEVIPEQPEPTFYSREMALPSAHLAAGFP LEKVGRRLADDEIVRFERKDPELVADHDLVGYVTNIDHPFGNVWTNIHRTDLEKLGVGYGTKLRITLDGVLPFELPLSPTFADAGE IGAAVAYLSSRGYLALARNAASLAYPYNLKAGISVQVKVG
FLA_Cbac	MSDLGTFDDSVGICKGLMLSVCPDVVIVDICHAMTPFDIEEGARLIVDLPRFFPEGTVFATTTYPATGTSTRSVALRIRQAAVGGAR GQWAGAGEGIQRAEGAYIYIAPNNGLLTSVIEEHGYLEAYEVTSTSVIPARPEPTFFSREMVAVPAAHLAAGFPLSQVGRPLQDSEIV RFDRRRPASLGDGGFAGVITVVDRPYGNVWTNISRKDLANQGIAYGTRLRITVDNVLPFDLTLTPTFADAGEIGAPVCYVNSRGYL SLARNASNLADTYNIRRHMPVNVQVVSRTDGEPHRSESSELASVKGT
FLA_Pbac	MEKVDNKTVFQRPIIAFMSDLGAFDDSVGICKGLMLSVCPEAQIIDICHSMTPFDVEQGARLIVDLPRFFPEGRTVFATTTYPATGT MARSLALRIKRPAKGGALGQWAGAGFGIERGVGGYIYVAPNNGLLTDVIEEHGYLEAYEITSTNVIPENPEPTFFSREMVALPSAYL AAGYPLSEVGQPLGDSEIVRFKKVLPRKMSESELVGVVAAIDRPFGNVWTNISRRDLDKIGVTYGSQLKVVLDNALMFELPLSQT FADAREIGAAVAYINSRGHLSLGRYAANLADRYNINRGMPIRLKVITG
FLA_Tnor	MDIGPIIGFASDLGLKDDSVALCKGLMISICPEVYIVDICHTMTPFDIEEGAWLALDLPRFFPEGRTIFAVTTYPATGTEARSIAVRIKK AVPGGSLEKWEGPGGGMERTLEGGYIYVAPNNGLLTFVLERYGYIEAYEIISTEFIPENPEPTFFSREMVAIRAACIAKKVVSQGVPL SKVGPPITEDKLARIKLSLPEKIAHNEIRGKIIRIDYPYGNVWTNISFNDLKSMGINYGSRLTVVIGDILSFNVYLTRTFADAGGIGDV ISYINSRGYFSLGGYAANLADLCNLRRGMNVVVIKV
FLA_PtaU1	MLQNTKNEYKPRDGGGHRPIIGFMSDLGTEDDSVGICKGVMLGLCPDAVIVDISHSMTPWDIDQGSRLIVDLPKFFPNWTTFATTS YRETGTSARSVAIKLPSGHVYVAPNNGLLTRVIEDHGYVEAYEVTTVGAIPAEPEPTWFSRDMVAYPAAAIAAGFPLENVGRPLND SEIVRADLPRYTQAEDGIIQGIVTTIDRPFGNIWTNIPRRIIENELRIAYGTNIRVVLDNLLVLEVPFMRTFGDVGLKKPMCYINSRGY FSLAYYGGNLADPYNIRRGMPVKIESITR

Table S2. Predicted ancestral fluorinases with their respective amino acid sequences.

Name	Sequence
ancestral_14	MLQNTKNEYKPRDGGGRRPIIGFMSDLGTADDSVGICKGVMLGVCPDAVIVDISHSMTPFDIEEGARLIVDLPRFFPEWTVFATTT YPETGTAARSVAIRLKRAARGGARGQWAGAGAGIERAESGHIYVAPNNGLLTRVIEEHGYVEAYEVTSPGVIPEEPEPTWFSRDMV AVPAAHLAAGFPLEGVGRPLDDSEIVRLDLPRPAEVADGGIRGVVTRIDRPFGNVWTNIPRRLIENELRVGYGTEIRVVLDNVLVLE LPFVRTFGDVGELGRPLVYINSRGYLSLAYYAGNLADPYNIRRGMPVRVEPVTRDSELGPVLG
ancestral_15	MGSSHMLQNTKNGLVPREYKPRDGGGHRPIIGFMSDLGTEDDSVGICKGVMLGLCPDAVIVDISHSMTPWDIEEGARLIVDLPRFF PNWRTVFATTTYPETGTSARSVAIRLKQAAKGGARGQWAGAGAGIERALESGHIYVAPNNGLLTRVIEDHGYVEAYEVTSTGVIPA EPEPTWFSRDMVAYPAAHLAKKVVAAGFPLENVGRPLNDSEIVRFDLPRPTQAEDGVIQGVVTTIDRPFGNIWTNIPRRIIENELRIA YGTNIRVVLDNLLVLELPFMRTFGDVGELGKPMCYINSRGYLSLAYYAGNLADPYNIRRGMPVKIESITRTDGEPHRSESSELASV KGT
ancestral_16	GGSRPIIGFMSDLGTFDDSVGICKGLMLSVCPDVVIVDICHSMTPFDIEEGARLIVDLPRFFPEGTVFATTTYPATGTTARSVALRIKQ AAKGGARGQWAGAGAGIERAEGGYIYVAPNNGLLTSVIEEHGYLEAYEVTSTNVIPERPEPTFFSREMVAIPAAHLAAGFPLSQVG RPLQDSEIVRFDRPRPAKLSDGELAGVVTAIDRPFGNVWTNISRRDLEKMGIAYGTQLRIVLDNVLPFELPLTRTFADAGEIGAPVC YINSRGYLSLARYAANLADPYNIRRGMPVRVKVVTRT
ancestral_17	GGSRPIIGFMSDLGTFDDSVGICKGLMLSVCPEVVIVDICHSMTPFDIEEGARLIVDLPRFFPEGRTVFATTTYPATGTTARSVALRIK RAAKGGALGQWAGAGAGIERAEGGYIYVAPNNGLLTSVIEEHGYLEAYEITSTNVIPENPEPTFFSREMVAIPAAHLAAGFPLSEVG RPLQDSEIVRFKRPLPRKMSDGELAGVVTAIDRPFGNVWTNISRRDLEKMGIAYGSQLKVVLDNVLPFELPLTRTFADAGEIGAPV CYINSRGYLSLGRYAANLADPYNIRRGMPVRVKVVTRT
ancestral_18	ANSRPIIGFMSDLGTFDDSVGICKGLMLSVCPDVVIVDICHSMTPFDIEEGARLIVDLPRFFPEGTVFATTTYPATGTTTRSVALRIKQ AAKGGARGQWAGAGAGIERAEGGYIYIAPNNGLLTSVIEEHGYLEAYEVTSTSVIPERPEPTFFSREMVAIPAAHLAAGFPLSQVGR PLQDSEIVRFDRPRPASLSDGELAGVVTAIDRPFGNVWTNISRKDLEKMGIAYGTQLRIVLDNVLPFELPLTPTFADAGEIGAPVCYI NSRGYLSLARNAANLADPYNIRRGMPVRVKVVTRT
ancestral_19	DSNSRPIIAFMSDLGTTDDSVAQCKGLMLSICPDVTIVDVCHSMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTTTRSVALRI KQAAKGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTTVIEEHGYIEAYEVSSTEVIPERPEPTFYSREMVAIPSAHLAAGFPLNKV GRPLQDSEIVRFTRPQPSTTSDGELTGVVTAIDHPFGNIWTNIHRTDLEKAGIGYGTQLKIVLDDVLPFELPLTPTFADAGEIGNPVA YLNSRGYLSLARNAASLAYPYNLKAGMSVRVKAA
ancestral_20	DSNSRPIIAFMSDLGTTDDSVAQCKGLMLSICPDVTIVDVCHSMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTTTRSVAVRI KQAAKGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTTVLEEHGYIEAYEVSSTEVIPERPEPTFYSREMVAIPSAHLAAGFPLDK VGRPLQDSEIVRFTRPQPSTTSDGELTGVVTAIDHPFGNIWTNIHRTDLEKAGIGYGTQLKIVLDDVLPFELPLTPTFADAGEIGNPV AYLNSRGYLSLARNAASLAYPYNLKAGMSVRVKAA
ancestral_21	DSSQRPIIAFMSDLGTTDDSVAQCKGLMLSICPDVTIVDVCHSMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTETRSVAVRI KQAAKGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTTVLEEHGYIEAYEVSSTEVIPERPEPTFYSREMVAIPSAHLAAGFPLDQ VGRPLQDSEIVRFSRPQVETSGDELTGVVTAIDHPFGNIWTNIHRTDLEKAGIGYGKQLKIVLDDVLPFELPLVPTFADAGEIGNVV AYLNSRGYLSLARNAASLAYPYNLKAGLSVRVEAH
ancestral_22	DGSQRPIIAFMSDLGTTDDSVAQCKGLMHSICPGVTVVDVCHSMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTETRSVAV RIKQAAKGGARGQWAGSGGGFERAEGSYIYIAPNNGLLTTVLEEHGYIEAYEVSSTKVIPERPEPTFYSREMVAIPSAHLAAGFPLS EVGRPLEDSEIVRFQRPQVEISGDALTGVVTAIDHPFGNIWTNIHRTDLEKAGIGYGKQLKIILDDVLPFEQPLTPTFADAGEIGNVA AYLNSRGYLSLARNAASLAYPYNLKAGLKVRVEAN
ancestral_23	SANSRPIIAFMSDLGTTDDSVAQCKGLMLSICPDVTIVDVCHSMTPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTTTRSVALRI KQAAKGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTTVIEEHGYIEAYEVSSTEVIPEQPEPTFYSREMVAIPSAHLAAGFPLNKV GRPLQDDEIVRFTRPQPSTVSDGDLTGVVTNIDHPFGNIWTNIHRTDLEKAGIGYGTQLRIVLDDVLPFELPLTPTFADAGEIGNPVA YLNSRGYLALARNAASLAYPYNLKAGLSVRVKAA
ancestral_24	SRGNRPIIAFMSDLGITDDSVAQCKGLMLSVCPDVTIVDVCHTMKPWDVEEGARYIVDLPRFFPEGTVFATTTYPATGTTARSVALR IKQAAKGGARGQWAGSGAGFERAEGSYIYIAPNNGLLTSVIEEHGYLEAYEVSSTEVIPEQPEPTFYSREMVALPSAHLAAGFPLN KVGRPLADDEIVRFTRPQPSVVSDGDLVGVVTNIDHPFGNIWTNIHRTDLEKAGVGYGTKLRIVLDEVLPFELPLSPTFADAGPIGT PVAYLSSRGYLALARNAASLAYPYNLKAGISVRVKVG

Table S3. Pair-wise percentage identity of reconstructed fluorinases, candidate fluorinases used for reconstruction, and other newly identified fluorinases (indicated by # suffix) that are not in candidate set employed here for ancestral reconstruction.

Variant	FLA_Nbra1	FLA_CA12	FLA_MA37	FLA_Amza	FLA_Pbac	FLA_N902	FLA_Cbac	FLA_SAJ15	FLA_AN11	SalL_Stro	FLA_Sxin	FLA_PtaU1	FLA_Tnor	ancestral_24	ancestral_25	ancestral_23	ancestral_15	ancestral_14	ancestral_17	ancestral_16	ancestral_19	ancestral_18	ancestral_22	ancestral_20	ancestral_21	FLA_Gbac_#	FLA_Sbac_#	FLA_Adig_#	FLA_Smor_#	FLA_Scat_#	FLA_Abar_#	FLA_Nbra2_#	FLA_Nbra3_#
FLA_Nbra1		84.6	68.8	73.2	67.2	75.2	65.3	79.1	83	34.9	71.7	45.7	52.6	87.2	89.6	82.8	51	58.1	69.7	72.2	80.5	72.9	75.8	79.8	78.5	62	75.7	78	77	76.1	82.9	97.6	95
FLA_CA12	84.6		73.8	76.3	64.8	79.9	66.1	78.9	87	35.6	76.4	48.9	53.8	90.3	93.6	86.9	55.3	61.2	72.8	76.1	86.2	77.4	79.9	85.6	82.9	60.4	78.7	77.5	76.3	81.7	89.9	84.6	86
FLA_MA37	68.8	73.8		73.8	58.3	75.6	58.5	76.2	73.5	34	80.6	45.3	50	73.4	72.9	78.4	56.2	55.4	67	67.8	79.1	68.7	85.9	79.7	82.1	53	70.8	72.2	87.8	82.1	72.8	68.1	70.6
FLA_Amza	73.2	76.3	73.8		61.9	74.9	60.6	75.9	79	34.1	77.1	48.7	51.6	79.6	78.3	81.6	53.9	58.4	69.6	71.9	83.9	72.5	80.7	84.6	82.3	54.5	74.7	73.6	77.1	77.4	76.3	72.9	75.3
FLA_Pbac	67.2	64.8	58.3	62		62.1	65	65.6	65.3	31.5	62.5	50.6	57.8	66.7	66.3	65.7	54.7	60.4	82.3	77.3	67.7	75.3	64.8	67	66.1	55.5	64	63.8	64.6	63.4	65.1	67.2	64
FLA_N902	75.2	79.9	75.6	74.9	62.1		63.3	77.9	79	35	77.3	48.9	53.8	81.9	80.3	86.9	55	59.9	70.9	72.8	84.6	74.8	80.9	83.6	82.2	57.2	74.7	90.3	78.2	79	79.9	74.5	77
FLA_Cbac	65.3	66.1	58.5	60.6	65	63.3		66.9	63.2	34.8	61.4	49.4	52.2	65.1	65.7	66.4	64.2	64.2	74.5	79.2	67.3	81.4	63.8	66.7	65.4	59.7	61.7	65.3	65.3	63.7	66.5	64.9	62.3
FLA_SAJ15	79.1	78.9	76.2	75.9	65.6	77.9	66.9		77.7	33.4	81.9	46.6	54.2	79.5	78.9	84.8	52.6	58.8	70	73.2	85.9	74.6	85.5	86.9	90.5	61.5	74.9	79.4	84.6	80.9	77.7	79.4	76.3
FLA_AN11	83	87	73.5	79	65.3	79	63.2	77.7		34.4	76.4	49.5	53.6	95.3	91.7	87.3	55.9	61.5	73.7	75.2	86	75.6	80.4	85.3	83.4	59.7	77.3	76.7	75.4	79.1	87.3	82	86
SalL_Stro	34.9	35.6	34	34.1	31.5	35	34.8	33.4	34.4		34.9	35.8	32.7	35.4	35.2	35.9	36.6	44.1	35.2	34.7	35.5	34.7	35.3	36.1	36	36.1	34.4	33.1	34.4	35.4	33.2	34.6	35.5
FLA_Sxin	71.7	76.4	80.6	77.1	62.5	77.3	61.4	81.9	76.4	34.9		47.3	55.8	76.7	76.3	81.3	53.3	59	70.4	71.9	83	72.6	92.3	83.7	87.6	55.1	74.9	74	83.6	83.6	75.7	71	73.7
FLA_PtaU1	45.7	48.9	45.3	48.7	50.6	49	49.4	46.6	49.5	35.8	47.3		43.7	49.2	49.2	50.2	76.5	73.3	58.1	60.2	50.8	58	48.6	50.5	49.5	54.8	48.2	46.8	47	48.2	49.5	45.7	48.2
FLA_Tnor	52.6	53.8	50	51.6	57.8	53.8	52.2	54.2	53.6	32.7	55.8	43.7		54.4	54.1	55.9	47.4	50.6	64.5	61.3	56.6	59.9	56.4	57.2	56.2	46.9	52.3	52.5	52.6	53.5	53.1	53	52.9
ancestral_24	87.2	90.3	73.4	79.6	66.7	81.9	65.1	79.5	95.3	35.4	76.7	49.2	54.4		96.3	91.9	55.6	61.5	75.1	78	89.2	78.3	81.8	88.2	85.2	60.9	77.7	79.8	77.4	80.4	88.9	85.2	88.3
ancestral_25	89.6	93.6	72.9	78.3	66.3	80.3	65.7	78.9	91.7	35.2	76.3	49.2	54.1	96.3		89.9	55.6	62.4	74.8	78	87.2	78.7	80.8	86.5	83.5	61.6	78	78.5	76.8	80.4	89.3	87.5	90.3
ancestral_23	82.8	86.9	78.4	81.6	65.7	86.9	66.4	84.8	87.3	35.9	81.3	50.2	55.9	91.9	89.9		56.7	61.8	76.7	79.7	96.6	81.7	87.9	95.6	91.2	61.6	77.7	83.8	82.2	84.7	85.9	81.8	84.3
ancestral_15	51	55.3	56.2	53.9	54.7	55	64.2	52.6	55.9	36.6	53.3	76.5	47.4	55.6	55.6	56.7		80.8	65.3	67.3	57.3	65.3	54.2	57	55.9	49	52.4	52.1	52.4	54	55.9	51.3	52.7
ancestral_14	58.1	61.2	55.4	58.4	60.4	59.9	64.2	58.8	61.5	44.1	59	73.3	50.6	61.5	62.4	61.8	80.8		71.5	73.9	62.1	71.7	59	61.8	60.9	54.3	59.6	56.5	57.3	61	61.5	58.4	60.9
ancestral_17	69.7	72.8	67	69.6	82.3	70.9	74.5	70	73.7	35.2	70.4	58.1	64.5	75.1	74.8	76.7	65.3	71.5		95	78.7	92	73.8	78.1	75.1	60	69.8	70	70.7	72.5	72.5	69.7	70.4
ancestral_16	72.2	76.1	67.8	71.9	77.3	72.8	79.2	73.2	75.2	34.7	71.9	60.2	61.3	78	78	79.7	67.3	73.9	95		81.3	97	75.3	80.7	77.7	61.9	71.6	71.2	72.2	74.7	75.4	72.2	72.9
ancestral_19	80.5	86.2	79.1	83.9	67.7	84.6	67.3	85.9	86	35.5	83	50.8	56.6	89.2	87.2	96.6	57.3	62.1	78.7	81.3		83	89.6	99	93.6	61.6	79.3	82.8	82.5	84.7	84.6	80.1	81.3
ancestral_18	72.9	77.4	68.7	72.5	75.3	74.8	81.4	74.6	75.6	34.7	72.6	58	59.9	78.3	78.7	81.7	65.3	71.7	92	97	83		76.3	82.3	79	61.5	71.9	72.6	73.6	77	76.7	72.9	73.9
ancestral_22	75.8	79.9	85.9	80.7	64.8	80.9	63.8	85.5	80.4	35.3	92.3	48.6	56.4	81.8	80.8	87.9	54.2	59	73.8	75.3	89.6	76.3		90.2	93.9	57.7	76.6	78.5	88.5	86.6	79.5	75.1	76.7
ancestral_20	79.8	85.6	79.7	84.6	67	83.6	66.7	86.9	85.3	36.1	83.7	50.5	57.2	88.2	86.5	95.6	57	61.8	78.1	80.7	99	82.3	90.2		94.6	61.6	79	81.8	83.2	85.4	83.9	79.5	80.7
ancestral_21	78.5	82.9	82.1	82.3	66.1	82.2	65.4	90.5	83.4	36.1	87.6	49.5	56.2	85.2	83.5	91.2	55.9	60.9	75.1	77.7	93.6	79	93.9	94.6		60.1	77.9	80.8	85.1	85.3	81.9	78.1	79.3
FLA_Gbac_#	62	60.4	53	54.5	55.5	57.2	59.7	61.5	59.7	36.1	55.1	54.8	46.9	60.9	61.6	61.6	49	54.3	60	61.9	61.6	61.5	57.7	61.6	60.1		58.5	60.1	57.5	57.5	60.1	61.7	59.3
FLA_Sbac_#	75.7	78.7	70.8	74.7	64	74.7	61.7	74.9	77.3	34.4	74.9	48.2	52.3	77.7	78	77.7	52.4	59.6	69.8	71.6	79.3	71.9	76.6	79	77.9	58.5		73.7	71.2	77.9	77	75.7	78.7
FLA_Adig_#	78	77.5	72.2	73.6	63.8	90.3	65.3	79.4	76.7	33.1	74	46.6	52.5	79.8	78.5	83.8	52.1	56.5	70	71.2	82.8	72.6	78.5	81.8	80.8	60.1	73.7		79.8	76	78.2	77.4	74.3
FLA_Smor_#	77	76.3	87.8	77.1	64.6	78.2	65.3	84.6	75.4	34.4	83.6	47	52.6	77.4	76.8	82.2	52.4	57.3	70.7	72.2	82.5	73.6	88.5	83.2	85.1	57.5	71.2	79.8		83.3	76.2	76.3	73.7
FLA_Scat_#	76.1	81.7	82.1	77.4	63.4	79	63.7	80.9	79.1	35.4	83.6	48.2	53.5	80.4	80.4	84.7	54	61	72.5	74.7	84.7	77	86.6	85.4	85.3	57.5	77.9	76	83.3		79.7	75.4	78.4
FLA_Abar_#	82.9	89.9	72.8	76.3	65.1	79.9	66.5	77.7	87.3	33.2	75.7	49.5	53.1	88.9	89.3	85.9	55.9	61.5	72.5	75.4	84.6	76.7	79.5	83.9	81.9	60.1	77	78.2	76.2	79.7		82.2	83.7
FLA_Nbra2_#	97.6	84.6	68.1	72.9	67.2	74.5	64.9	79.4	82	34.6	71	45.7	53	85.2	87.5	81.8	51.3	58.4	69.7	72.2	80.1	72.9	75.1	79.5	78.1	61.7	75.7	77.4	76.3	75.4	82.2		93.7
FLA_Nbra3_#	95	86	70.6	75.3	64	77	62.3	76.3	86	35.5	73.7	48.2	52.9	88.3	90.3	84.3	52.7	60.9	70.4	72.9	81.3	73.9	76.7	80.7	79.3	59.3	78.7	74.3	73.7	78.4	83.7	93.7	

Table S4. High throughput fluorination/chlorination activities for reconstructed ancestral_15 and ancestral_25 activity with various solubility and fusion tags, including NT11[3], SUMO[4], PolyEK[5] and OmpA[6], as compared to FLA_MA37. All constructs have a C-terminal His-tag to facilitate purification.

	ID	Tag	Fluorination % conversion	Chlorination % conversion
ancestral_25	34	His Tag	2.64	0
	37	MBP	0.15	0
	39	NT11	4.17	0.28
	44	SUMO	2.28	0.05
	46	PolyEK	1.81	0
	48	OmpA	0	0
ancestral_15	66	His Tag	0	0
	69	MBP	0	0
	71	NT11	0	0
	76	SUMO	0	0
	78	PolyEK	0	0
	80	OmpA	0	0
FLA_MA37		His Tag	20.29	0.81

Table S5. High throughput fluorination/chlorination activities and selectivity for reconstructed ancestral fluorinases with two solubility tags (NT11 and SUMO), as compared to FLA_MA37.

Variant	ID	Tag	Fluorination	Chlorination	F/Cl selectivity
ancestral_14	119	NT11	0	0	NA
	120	SUMO	0	0	NA
ancestral_15	71	NT11	0	0	NA
	76	SUMO	0	0	NA
ancestral_16	135	NT11	0.09	0	100
	136	SUMO	0.1	0	100
ancestral_17	127	NT11	0.24	0	100
	128	SUMO	0	0	NA
ancestral_18	151	NT11	1.72	0	NA
	152	SUMO	2.07	0.13	15.92
ancestral_19	143	NT11	4.49	0.34	13.21
	144	SUMO	0	0	NA
ancestral_21	167	NT11	2.65	0.21	12.62
	169	SUMO	1.6	0.18	8.89
ancestral_22	159	NT11	4.87	0.1	48.7
	160	SUMO	0	0	NA
ancestral_23	111	NT11	6.36	0.27	23.56
	112	SUMO	0.19	0	100
ancestral_24	103	NT11	7.25	0.25	29
	104	SUMO	3.3	0.22	15
ancestral_25	39	NT11	4.17	0.28	14.89
	44	SUMO	2.28	0.05	45.6
FLA_MA37		His Tag	17.71	0.78	22.71

Table S6. NT11-tagged ancestral fluorinases vs. FLA_MA37. Fluorination and chlorination conversions determined via triplicate assays of purified proteins.

Variant	% conversion				F/Cl 24h selectivity	Protein melting temperature (T_m)
	Fluorination		Chlorination			
	1.5 h	24h	1.5 h	24h		
FLA_MA37	34.11 ± 1.75	82.20 ± 0.30	2.39 ± 0.06	15.21 ± 0.16	5.40	48.5 ± 0.1
ancestral_16	9.20 ± 1.80	28.57 ± 0.03	0.79 ± 0.15	3.67 ± 0.01	7.79	64.8 ± 0.1
ancestral_19	16.96 ± 3.48	41.34 ± 0.63	2.86 ± 0.52	11.37 ± 0.05	3.64	55.8 ± 0.0
ancestral_21	20.74 ± 3.46	51.76 ± 0.74	3.67 ± 0.06	19.01 ± 0.10	2.72	50.5 ± 0.4
ancestral_22	34.24 ± 1.28	79.17 ± 0.87	2.54 ± 0.02	12.49 ± 0.25	6.34	46.2 ± 0.2
ancestral_23	24.37 ± 1.72	66.07 ± 0.15	2.95 ± 0.10	14.68 ± 0.18	4.50	53.2 ± 0.0
ancestral_24	26.28 ± 2.66	66.88 ± 0.19	2.70 ± 0.04	15.16 ± 0.19	4.41	54.5 ± 0.1
ancestral_25	22.24 ± 3.83	52.49 ± 0.17	1.95 ± 0.09	8.44 ± 0.08	6.22	59.2 ± 0.2
ancestral_16_M13	20.07 ± 2.64	51.14 ± 0.61	2.19 ± 0.55	9.25 ± 0.09	5.53	62.0 ± 0.1

Table S7. High throughput (HTP) fluorination/chlorination and selectivity were performed for ancestral_16 fluorinase mutants as compared to the predicted ancestral_16 fluorinase (ID 135 with NT11 tag). Mutants M1, M2 and M3 correspond to F50W, F157Y and R253P, respectively. Similarly, the mutant M13 corresponds to the double mutant combining M1 and M3. There were no detectable conversion for chlorination for all constructs. The HTP assay were performed as described in the methods, the assays were performed in triplicates.

Variants	ID	Mutations	Fluorination % conversion	Average fold improvement (fluorination)
ancestral_16	135	WT	0.10 ± 0.01	1.00
ancestral_16_M1	302	F50W	0.12 ± 0.01	1.14
ancestral_16_M2	303	F157Y	0	0
ancestral_16_M3	304	R252P	0	0
ancestral_16_M12	305	F50W, F157Y	0.07 ± 0.01	0.66
ancestral_16_M23	306	F157Y, R252P	0.11 ± 0.01	1.04
ancestral_16_M13	307	F50W, R252P	0.43 ± 0.02	4.12
ancestral_16_M123	308	F50W, F157Y, R252P	0.12 ± 0.01	1.12

Table S8. Root-mean-squared fluctuations and deviations for the cofactor SAM from MD simulations of FLA_MA37, ancestral_22, ancestral_16 and its single, double and triple mutants from MD simulations.

Variant	SAM	
	RMSF (Å)	Mean RMSD (Å)
FLA_MA37	1.04	1.7
ancestral_22	1.18	1.73
ancestral_16	1.65	2.34
ancestral_16_m1	1.38	1.92
ancestral_16_m2	1.35	1.76
ancestral_16_m3	1.56	2.32
ancestral_16_m12	1.85	2.41
ancestral_16_m13	1.36	2.04
ancestral_16_m23	1.98	2.59
ancestral_16_m123	1.22	2.03

Table S9. Primers used in site directed mutagenesis.

Primer ID	**Description**	**DNA Sequence (5'-3')**
ppS1197	ancestral_16_F50W_F	CTCTATGACGCCCTGGGACATCGAAGAAG
ppS1198	ancestral_16_F50W_R	CTTCTTCGATGTCCCAGGGCGTCATAGAG
ppS1199	ancestral_16_F157Y_F	GAGCCTACGTTCTACTCCCGGGAGATG
ppS1200	ancestral_16_F157Y_R	CATCTCCCGGGAGTAGAACGTAGGCTC
ppS1201	ancestral_16_R253P_F	ACTTCCGTTAACTCCTACGTTTGCTGACG
ppS1202	ancestral_16_R253P_R	CGTCAGCAAACGTAGGAGTTAACGGAAGT

References

1 H. Sun, W. L. Yeo, Y. H. Lim, X. Chew, D. J. Smith, B. Xue, K. P. Chan, R. C. Robinson, E. G. Robins, H. Zhao and E. L. Ang, *Angew Chem Int Ed Engl*, 2016, **55**, 14277–14280.

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

3 Nguyen TKM, Ki MR, Son RG, Pack SP. The NT11, a novel fusion tag for enhancing protein expression in Escherichia coli. Appl Microbiol Biotechnol. 2019;103:2205–16.

4 Marblestone JG, Edavettal SC, Lim Y, Lim P, Zuo X, Butt TR. Comparison of SUMO fusion technology with traditional gene fusion systems: Enhanced expression and solubility with SUMO. Protein Sci. 2006;15(1):182–189. doi:10.1110/ps.051812706

5 Paraskevopoulou V, Falcone FH. Polyionic tags as enhancers of protein solubility in recombinant protein expression. Microorganisms. 2018;6(2):47. doi:10.3390/microorganisms6020047

6 Pechsrichuang P, Songsiriritthigul C, Haltrich D, Roytrakul S, Namvijtr P, Bonaparte N, Yamabhai M. OmpA signal peptide leads to heterogeneous secretion of *Bacillus subtilis*

chitosanase enzyme from *Escherichia coli* expression system. SpringerPlus. 2016;5(1):1200. doi:10.1186/s40064-016-2893-y.