Enhancement of tryptophan 2-monooxygenase thermostability by semi-rational enzyme engineering: A strategic design to minimize experimental investigation

Sirus Kongjaroon,[a] Narin Lawan,[b] Duangthip Trisivirat,[a] and Pimchai Chaiyen*[a]

Supplementary Information

Content

No.	Mutation	FoldX (kcal/mol)	Not within 8 Å	No.	Mutation	FoldX (kcal/mol)	Not within 8 Å
1	M25W	-1.22	✓	26	A333W	-1.80	
2	S33P	-1.02	✓	27	T358L	-1.36	
3	T36L	-1.02	✓	28	T361M	-1.66	
4	T38P	-2.27	✓	29	T372R	-0.52	
5	T55Y	-0.15		30	K377R	-1.23	✓
6	S70A	-1.04		31	T382Q	-1.27	✓
7	Q85W	-1.01	✓	32	Q385L	-2.13	✓
8	T120P	-0.20		33	Q399G	0.31	
9	D123N	-1.44		34	H406P	0.23	
10	K183R	-0.15		35	A419S	0.48	
11	R186D	-0.14		36	Q420L	-0.81	
12	A191M	-1.24	✓	37	A424P	-1.08	✓
13	G196L	-1.75	✓	38	T430L	-1.33	✓
14	C204Y	-1.10	✓	39	H443Y	-0.95	
15	S236T	-1.26		40	A473W	-1.38	✓
16	V247L	-1.10		41	Y480L	0.31	
17	G251S	-1.28		42	V485R	-0.07	
18	Q264D	0.43		43	Q488F	-1.63	✓
19	S265F	-1.00		44	T496W	-1.98	✓
20	D272G	0.34		45	S499I	-1.07	✓
21	A290P	-1.19	✓	46	Y507F	0.42	
22	S298L	-1.06		47	C511D	-1.54	
23	S304L	-1.57	✓	48	S512D	2.46	
24	A307S	-1.46	✓	49	S530A	-1.16	✓
25	N331P	-1.55	✓	50	K544P	-0.88	
Total number					50	34	22

Table S1 A list of all 50 predicted mutations from the FireProt web base with FoldX energy calculations listed in kcal.mol^{-1}. The last column indicates whether the position is located 8 Å away from the active site.

Chain A				Chain B			
Rank	Residue	No.	B-factor value	Rank	Residue	No.	B-factor value
1	Ala	316	43.07	1	Gly	169	44.24
2	Glu	315	38.04	2	Glu	168	41.92
3	Glu	168	34.84	3	Glu	318	40.47
4	Glu	308	33.94	4	Gly	170	38.76
5	Lys	379	33.46	5	Arg	217	36.16
6	Gly	317	32.48	6	Arg	320	35.95
7	Glu	404	30.99	7	Met	25	35.67
8	Gly	169	30.14	8	Arg	87	33.41
9	Glu	318	30.21	9	Glu	308	32.72
10	Ala	34	29.73	10	Ala	316	32.44
11	Ala	307	28.53	11	Leu	167	32.03
12	Asp	342	26.94	12	Gln	319	31.6
13	Thr	86	26.92	13	Thr	86	31.58
14	Asn	6	26.68	14	Lys	379	30.43
15	Leu	167	26.2	15	Asp	342	29.92
16	Arg	217	26.03	16	Gly	317	29.52
17	Glu	402	26.03	17	Met	425	28.91
18	Gly	170	25.82	18	Glu	315	28.3
19	Gln	85	25.68	19	Ser	171	27.94
20	Lys	502	25.65	20	Pro	426	27.52

Table S2 A list of residues with the top 20 highest B-factor values calculated from the B-FITTER software. The results of different chains are presented as Chain A and Chain B.

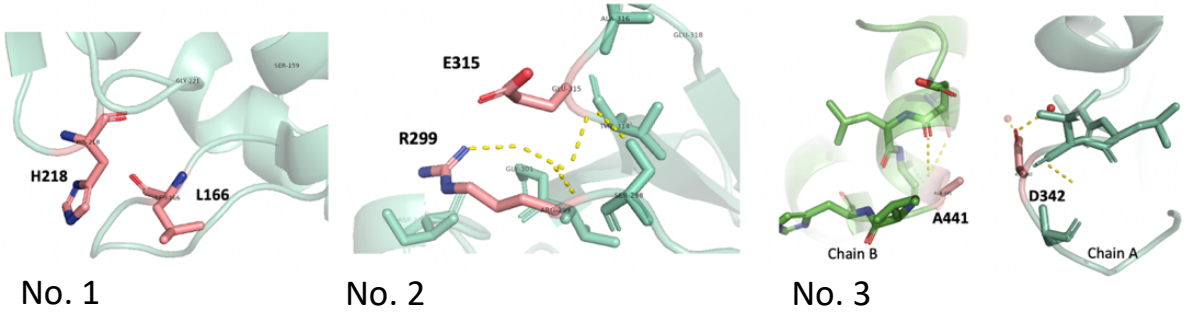
Design Name	# Mutations	% Mutations
4iv9_design_1	25	4.53
4iv9_design_2	34	6.16
4iv9_design_3	38	6.88
4iv9_design_4	41	7.43
4iv9_design_5	50	9.06
4iv9_design_6	62	11.23
4iv9_design_7	75	13.59
4iv9_design_8	84	15.22
4iv9_design_9	123	22.28

#	Pos...	WT	des 1	des 2	des 3	des 4	des 5	des 6	des 7	des 8	des 9
9	55	T	Y	Y	Y	Y	Y	Y	Y	Y	Y
9	473	A	Y	Y	Y	Y	Y	Y	Y	Y	Y
9	258	M	W	W	W	W	W	W	W	W	W
9	204	C	V	I	I	L	L	V	L	L	I
6	223	D	V	V		I		I	I		I
8	426	P	Q	Q	Q	Q		Q	Q	Q	Q
9	33	S	P	P	P	P	P	P	P	P	P
9	34	A	P	P	P	P	P	P	P	P	P
9	544	K	P	P	P	P	P	P	P	P	P
9	402	E	N	N	N	N	N	N	N	N	N
9	236	S	L	L	L	L	L	L	L	L	L
9	314	T	L	L	L	L	L	L	L	L	L
9	331	N	L	L	L	L	L	L	L	L	L
9	543	S	L	L	L	I	I	I	I	A	L
9	311	I	K	K	K	V	K	K	K	K	K
9	285	S	I	I	I	I	I	I	I	I	I
9	535	L	I	I	I	I	I	I	I	I	I
9	307	A	G	G	G	G	G	G	G	G	G
9	160	L	F	F	F	F	F	F	F	F	F
9	319	Q	F	F	F	F	F	F	T	T	T
9	494	P	F	F	F	F	F	F	F	F	F
9	416	E	D	D	D	G	D	D	G	G	Q
9	499	S	D	D	D	D	D	D	D	D	D
9	78	V	A	A	A	A	L	L	L	L	L
9	530	S	A	A	A	A	A	A	A	A	A
1	18	P									A
2	21	K								E	E
1	22	K									R
5	25	M					S	S	S	S	S
1	26	T									H
1	31	S									Y
1	36	T									V
1	40	R									K
6	44	V				I	I	I	I	I	I
5	49	S					A	A	A	A	A
1	62	V									I
3	66	V							T	T	T
3	68	Y							F	F	F
1	70	S									A
3	85	Q							P	P	P
3	86	T							R	R	R
4	87	R						H	Y	Y	Y
1	89	R									H
1	111	K									N
5	116	S					P	P	P	P	P
4	119	T						Q	Q	Q	Q
1	120	T									R

3	139	H							E	E	E
7	144	K			Q	Q	Q	Q	Q	Q	Q
8	145	K		P	P	P	P	P	P	P	P
4	148	E			P			P	P		P
8	159	S		A	A	A	A	A	A	A	A
4	162	S						R	R	R	R
1	166	L									T
1	176	L									M
1	177	D									Q
1	188	E									D
7	191	A			R	R	R	R	R	R	H
5	192	I					E	Q	Q	Q	E
8	196	G		Q	Q	Q	A	A	A	A	A
5	199	N					D	D	D	D	D
8	208	N		E	E	E	E	E	E	E	E
1	210	I									L
8	212	C		R	R	R	R	R	R	R	R
1	217	R									P
7	226	A			R	R	R	R	R	R	R
3	247	V							L	L	L
7	251	G			A	A	S	A	A	A	A
4	253	T						L	L	L	L
1	259	V									I
3	263	Y							W	W	W
6	264	Q				E	E	E	E	E	E
4	265	S						E	E	E	E
1	279	A									Q
2	289	K								I	V
8	290	A		P	P	P	P	P	P	P	P
1	293	D									E
4	298	S						V	V	V	V
3	299	R							Q	Q	Q
3	308	E							D	D	D
4	316	A						Q	N	N	N
1	320	R									Y
1	336	M									L
2	337	I								V	V
2	339	C								R	R
8	345	S		T	T	T	T	T	T	T	T
2	352	A								C	C
2	355	V								I	I
2	360	L							M		M
1	362	G									A
1	371	R									E
1	372	T									R
5	376	I					L	L	L	L	L
1	378	N									H
3	379	K							N	N	N
8	382	T		Q	Q	Q	Q	Q	Q	Q	V
4	386	S						T	T	T	T
3	388	G							T	T	T
5	391	R					K	K	K	K	K
2	399	Q								E	E
1	404	E									D
1	411	L									I
3	420	Q							H	H	H
1	422	M									L
1	424	A									S
1	425	M					F				
2	430	T								E	E
5	443	H					N	N	N	N	N
8	445	T		E	E	E	E	E	E	E	E
1	449	Y									H
5	451	L					V	V	V	V	V
2	454	D								N	N
1	458	E									D
6	472	S				A	A	A	A	A	A
2	483	E								Q	Q
4	501	N						E	E	E	S
1	502	K									T
4	511	C						D	D	D	D
1	512	S									C
2	515	F								W	W
1	520	I									L
2	528	L								I	I
5	550	C					Q	Q	Q	Q	Q
1	554	S									R

Fig. S1 Suggestion for mutations by PROSS. PROSS suggested 9 designs for the variants. Design 9 contains the largest number of mutations, while the first design contains the lowest number of suggested mutations. The data were taken from the PROSS website.

No.	Chain	Seq No.	Residue	Chain	Seq No.	Residue
1	A	166	LEU	A	218	HIS
2	A	299	ARG	A	315	GLU
3	A	342	ASP	B	441	ALA



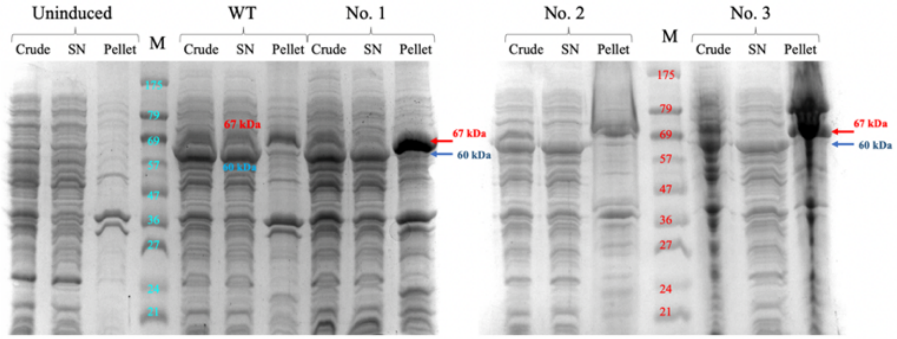


Fig. S2 Expression of DbD variants analyzed by SDS-PAGE. Putative interactions of DbD variants No. 1, 2, and 3 are displayed from left to right. Crude lysate, supernatant (SN), and pellet fractions of all DbD variants and the wild type (WT) enzyme under uninduced and induced conditions were analyzed using SDS-PAGE. The molecular weight markers (M) indicate their sizes in the figure. The chaperone GroEL band is around 60 kDa.

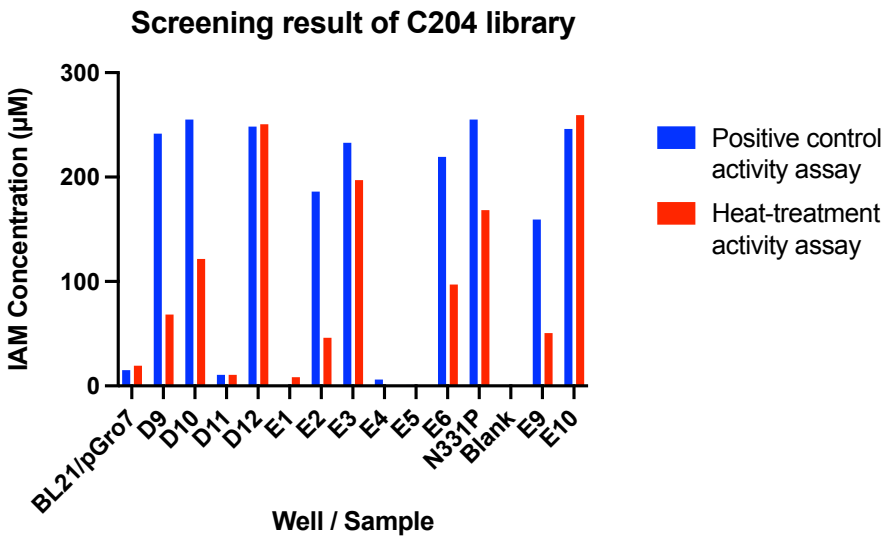


Fig. S3 Example of screening data obtained from high throughput screening. The C204 library was screened during the second round of screening, where N331P was used as the template. The controls for the screening assays were *E. coli* BL21 (DE3) with pGro7 plasmid (BL21/pGro7), blank media, and *E. coli* BL21 (DE3) with pGro7 and pET15b-His-GFP-TMO:N331P. The supernatant of the cell lysate was used in the reactions, and the assays were carried out twice as described in Experimental Procedures. The first assay was done before heat treatment (positive control activity assay), and the second assay was done after the heat treatment (heat-treatment activity assay). For the second assay, the sample was pre-incubated at 55 ˚C for one hour before measuring the enzyme activity. The samples that could generate higher or equal IAM production as TMO-N331P (TMO-P), i.e., D12, E3, and E10, were selected for DNA sequencing.

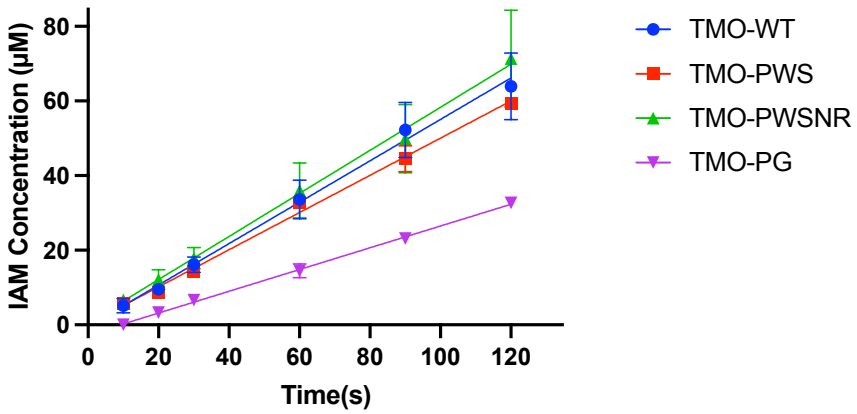


Fig. S4 Initial rates for IAM formation by TMO-variants. Activity assays under initial rate conditions indicate a drop in activity for TMO-PG. Therefore, the mutation of Q85G was neglected for further investigation.

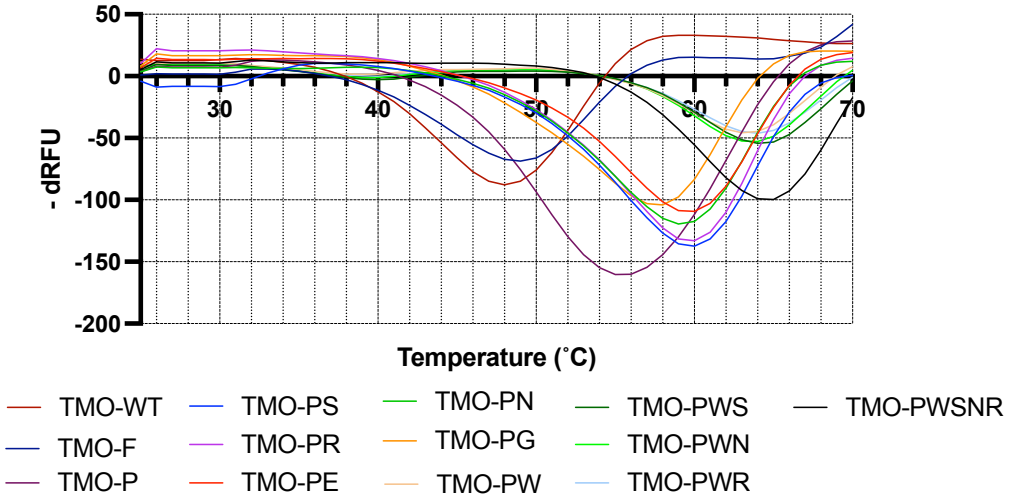
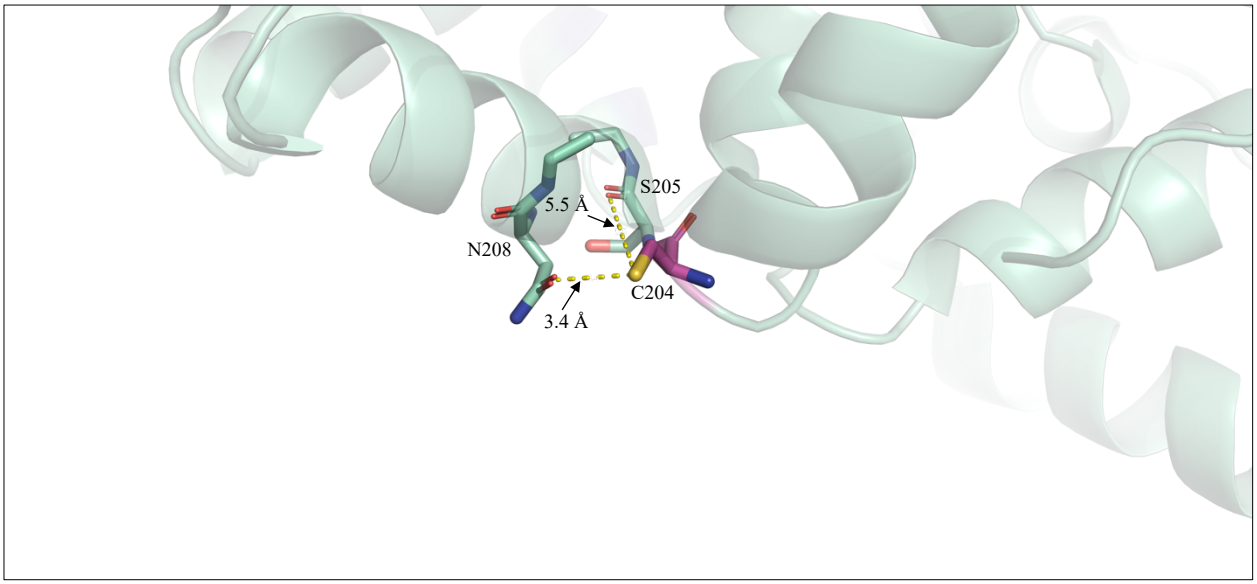


Fig. S5 The plot of -dRFU over temperature (˚C) showing raw data from thermal shift assays for determination of T_m values of different variants.

The C204 measurement of the TMO-WT crystal structure (PDB:4IV9)



The S204 measurement of the TMO-PWS structure from the MD structural model

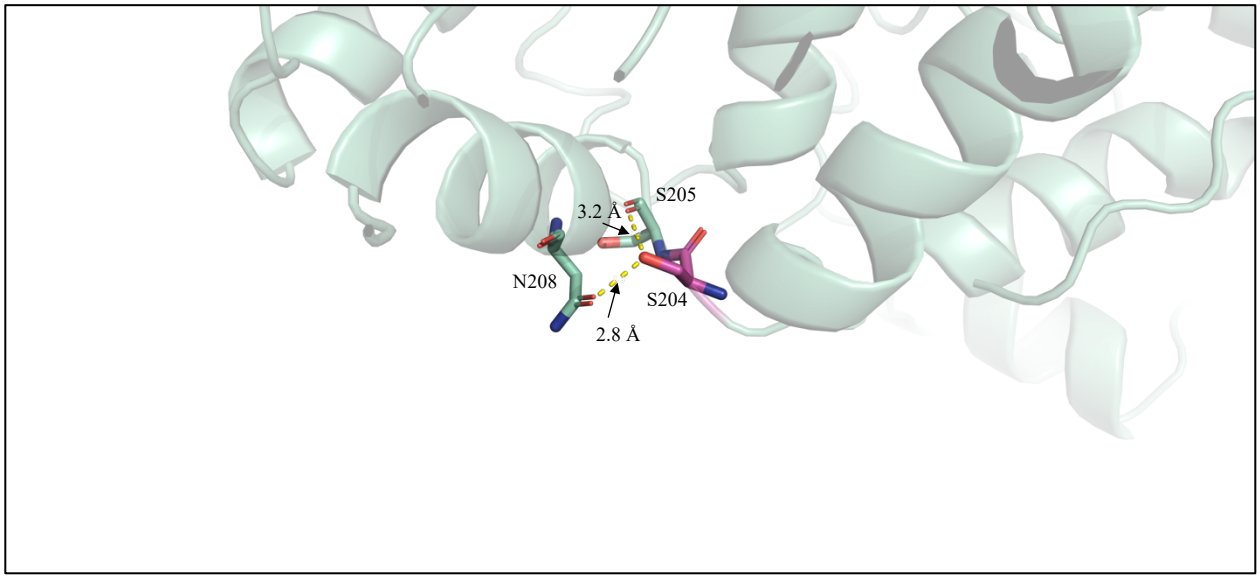
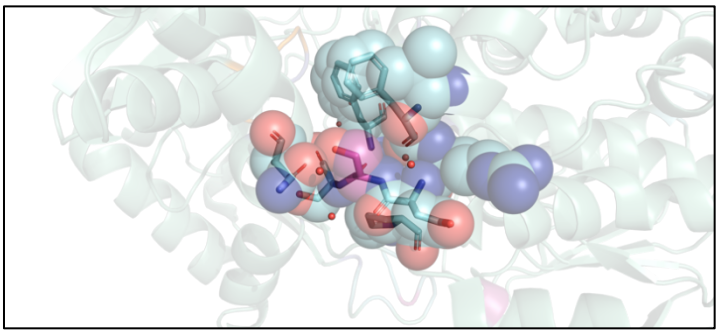
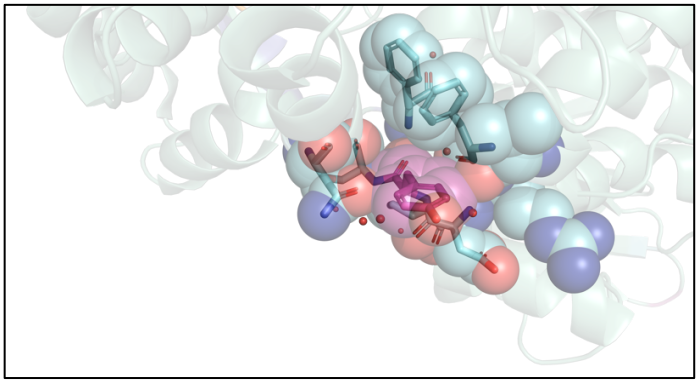


Fig. S6 The measured distances of C204S mutation. (A) For C204, the distance was measured from the sulfur atom of the side chain of C204. The measurement was made from the crystal structure of TMO-WT (4IV9). (B) For S204, the distance was measured from the oxygen atom of the side chain of S204. The measurement was made from the MD structural model. The shorter distances were found in the S204 mutation.



A) C204S mutation showing S204 in pink



B) C204Y mutation showing Y204 in pink

Fig. S7 The predicted structural models of (A) C204S and (B) C204Y. The pink spheres show the overlapping regions with the nearby spheres. Unlike in the C204S variant, a hydroxyl group from the side chain of tyrosine is not suitable for forming hydrogen bonds with nearby residues.

Nucleotide sequence of His-GFP11-TMO coding fragment
ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGcgtga ccacatggtgctgcacgagtacgttaacgcggcgggcattaccctgattggtagcgatggtggcagcggtggcggtagcaccagcG CTAGCatgtatgatcactttaacagcccgagcattgacatcctgtacgattatggtccgttcctgaagaaatgcgaaatgaccggcg gtattggcagctatagcgcgggcaccccgaccccgcgtgtggcgattgttggtgcgggtatcagcggcctggtggcggcgaccgagct gctgcgtgcgggtgttaaggacgtggttctgtatgaaagccgtgatcgtatcggcggtcgtgtgtggagccaggtttttgatcaaaccc gtccgcgttacattgcggagatgggtgcgatgcgtttcccgccgagcgcgaccggcctgtttcactatctgaagaaattcggtatcagc accagcaccacctttccggacccgggcgtggttgataccgaactgcactaccgtggcaagcgttatcactggccggcgggtaagaaa ccgccggagctgttccgtcgtgtgtacgaaggctggcagagcctgctgagcgagggttatctgctggaaggcggtagcctggttgcgc cgctggacatcaccgcgatgctgaaaagcggccgtctggaggaagcggcgattgcgtggcaaggttggctgaacgtgtttcgtgatt gcagcttctacaacgcgattgtttgcatctttaccggtcgtcacccgccgggcggtgaccgttgggcgcgtccggaggatttcgagctg tttggcagcctgggcatcggtagcggcggtttcctgccggtgtttcaggcgggcttcaccgagattctgcgtatggttatcaacggttac cagagcgaccaacgtctgattccggatggcatcagcagcctggcggcgcgtctggcggaccagagcttcgatggcaaggcgctgcgt gaccgtgtgtgcttcagccgtgttggccgtattagccgtgaggcggaaaaaatcattatccagaccgaggcgggtgaacaacgtgtgt ttgatcgtgtgatcgttaccagcagcaaccgtgcgatgcagatgattcactgcctgaccgacagcgagagcttcctgagccgtgatgt ggcgcgtgcggttcgtgaaacccacctgaccggtagcagcaagctgtttattctgacccgtaccaaattctggatcaagaacaaactg ccgaccaccattcagagcgacggcctggtgcgtggtgtttactgcctggactatcaaccggatgagccggaaggccacggtgtggttc tgctgagctacacctgggaggacgatgcgcagaagatgctggcgatgccggataagaaaacccgttgccaagtgctggttgacgatc tggcggcgatccacccgacctttgcgagctatctgctgccggtggacggtgattacgagcgttatgttctgcaccacgattggctgacc gatccgcacagcgcgggtgcgttcaagctgaactacccgggcgaggacgtttatagccagcgtctgttctttcaaccgatgaccgcga acagcccgaacaaagataccggcctgtacctggcgggttgcagctgcagctttgcgggcggttggatcgagggtgcggtgcagaccg cgctgaacagcgcgtgcgcggttctgcgtagcaccggcggtcaactgagcaaaggcaacccgctggactgcattaacgcgagctacc gttattaa
Amino acid sequence of His-GFP11-TMO (67 kDa)
MGSSHHHHHHSSGLVPRGSHMRDHMVLHEYVNAAGITLIGSDGGSGGGSTSASMYDHFNSPSIDILYDY GPFLKKCEMTGGIGSYSAGTPTPRVAIVGAGISGLVAATELLRAGVKDVVLYESRDRIGGRVWSQVFDQTR PRYIAEMGAMRFPPSATGLFHYLKKFGISTSTTFPDPGVVDTELHYRGKRYHWPAGKKPPELFRRVYEGW QSLLSEGYLLEGGSLVAPLDITAMLKSGRLEEAAIAWQGWLNVFRDCSFYNAIVCIFTGRHPPGGDRWAR PEDFELFGSLGIGSGGFLPVFQAGFTEILRMVINGYQSDQRLIPDGISSLAARLADQSFDGKALRDRVCFSR VGRISREAEKIIIQTEAGEQRVFDRVIVTSSNRAMQMIHCLTDSESFLSRDVARAVRETHLTGSSKLFILTRTK FWIKNKLPTTIQSDGLVRGVYCLDYQPDEPEGHGVVLLSYTWEDDAQKMLAMPDKKTRCQVLVDDLAAI HPTFASYLLPVDGDYERYVLHHDWLTDPHSAGAFKLNYPGEDVYSQRLFFQPMTANSPNKDTGLYLAGC SCSFAGGWIEGAVQTALNSACAVLRSTGGQLSKGNPLDCINASYRY

Table S3 Nucleotide and amino acid sequences of His-GFP11-TMO (N-terminal 6xHis tag attached with GFP11 domain, linker sequence, and TMO-coding gene from *P. savastanoi*.

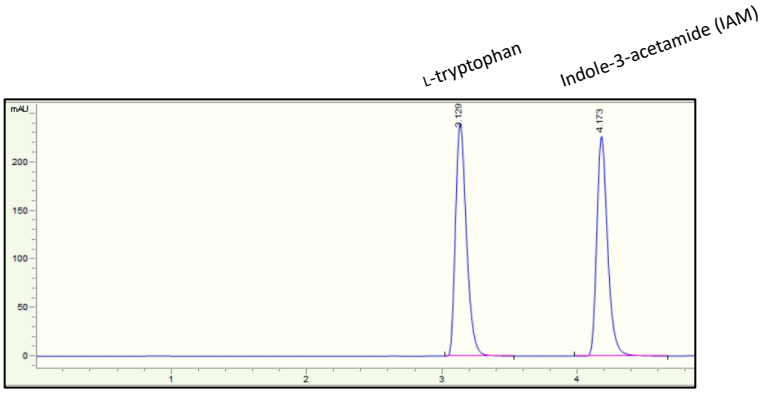


Fig. S8 The HPLC-UV chromatogram of standard L-tryptophan and IAM. The standard compounds were mixed in one solution and injected into the HPLC systems as described in Experimental Procedures. The retention time of L-tryptophan and IAM were 3.1 min and 4.2 min, respectively.

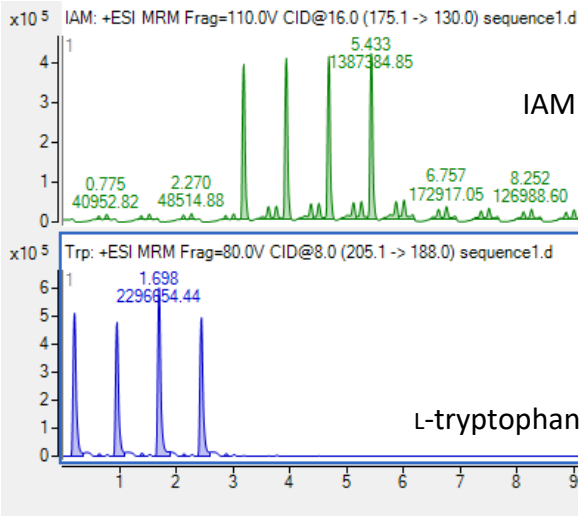
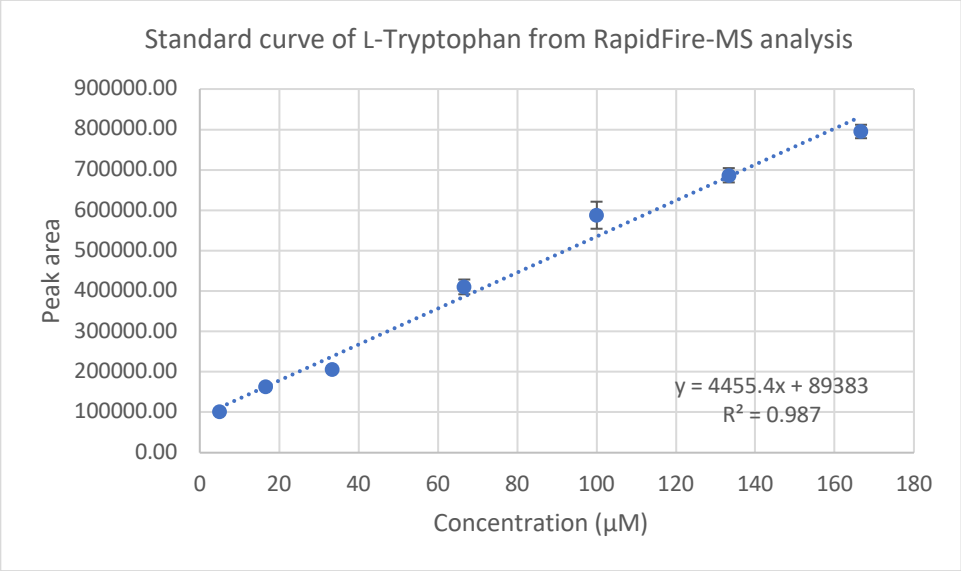


Fig. S9 The chromatograms of L-tryptophan and IAM standards analyzed on the RapidFire-Mass Spectrometry systems. The sequence of the injections included four injections of the L-tryptophan standard followed by another four injections of the IAM standard. The L-tryptophan peaks are in blue and the IAM peaks are in green.

A) Standard curve of L-tryptophan from the RapidFire-Mass Spectrometry systems



B) Standard curve of IAM from the RapidFire-Mass Spectrometry systems

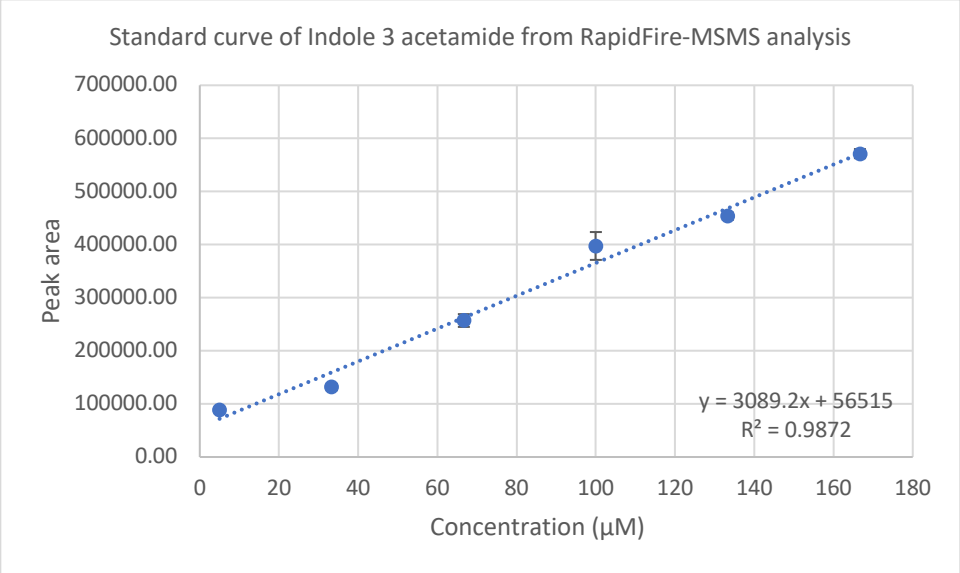


Fig. S10 The standard curves of L-tryptophan and IAM from the RapidFire-Mass Spectrometry systems. (A) Standard curve of L-tryptophan. (B) Standard curve of IAM.