Figure S1. Alternative approach for supplying marine natural products (MNPs). Left, the traditional method for obtaining MNPs: large amounts of the biological sample are collected and chemically extracted. New or known metabolites are then subjected to structural elucidation studies. Right, our alternative strategy for supplying MNPs: a much smaller sample is required (milligrams). This sample is divided in two portions; one is used for chemical extraction and subjected to HRMS analysis while the other is used for DNA extraction and PCR. After structural prediction, the pathway is cloned and heterologously expressed in *E. coli* to produce the natural product. Additionally, rational genetic engineering in the same system can be used to produce natural derivatives with unknown genes or unnatural derivatives for lead optimization.

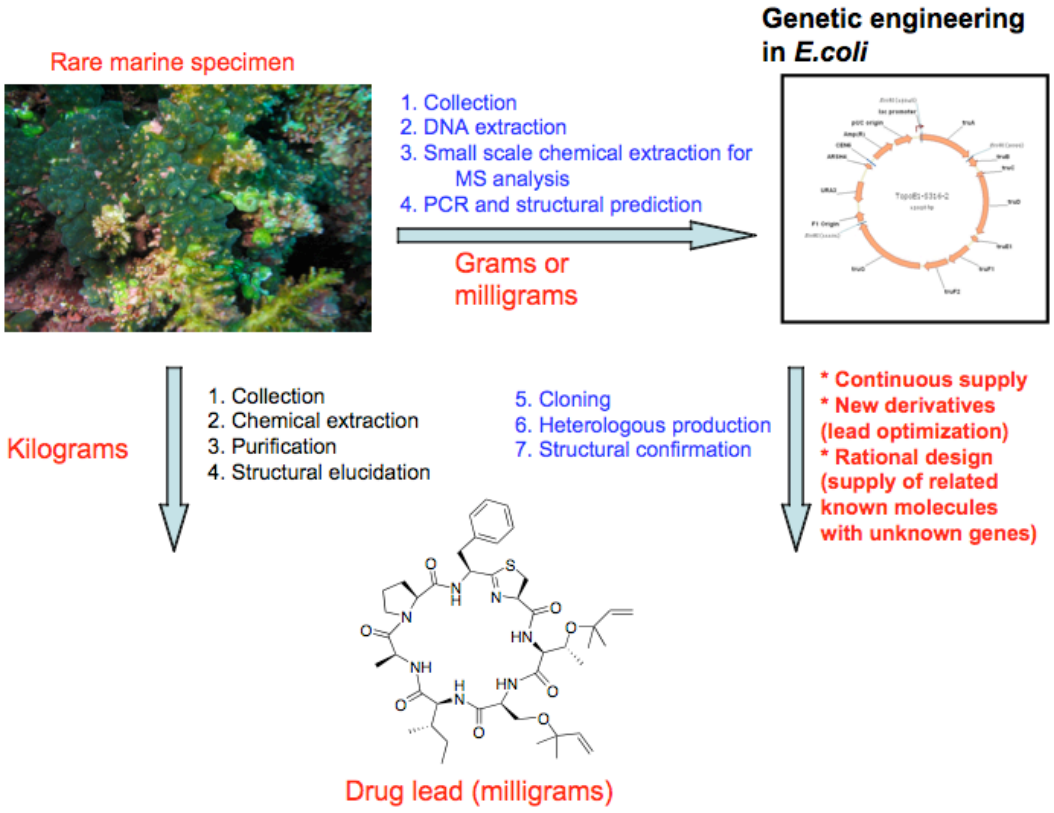


Figure S2. Compounds produced by *tru*. Schematic representations of some of the 6, 7, and 8-amino acid cyanobactins synthesized by *tru* are shown, indicating the broad substrate nature of the enzymes. Identical enzymes synthesize derivatives with all of the mutations shown.

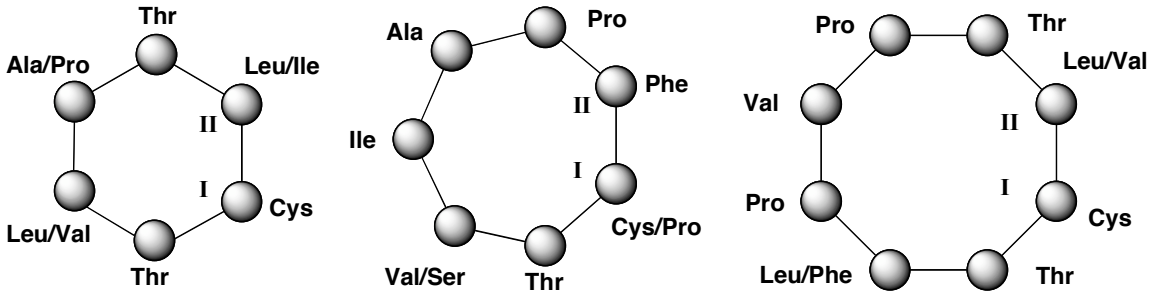
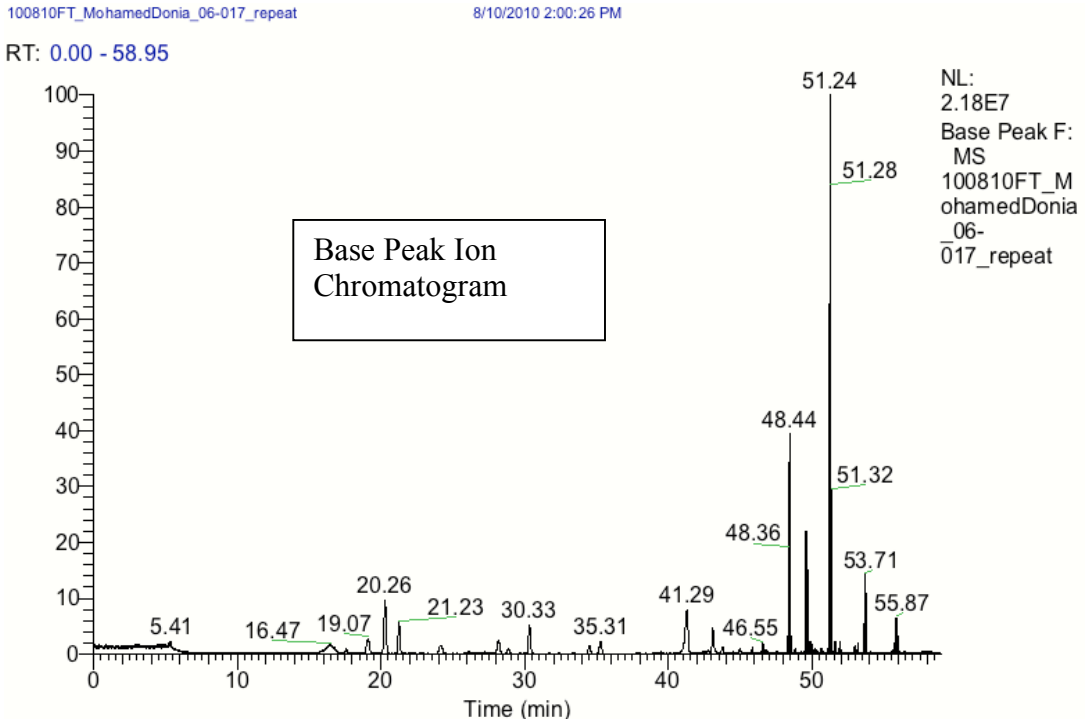
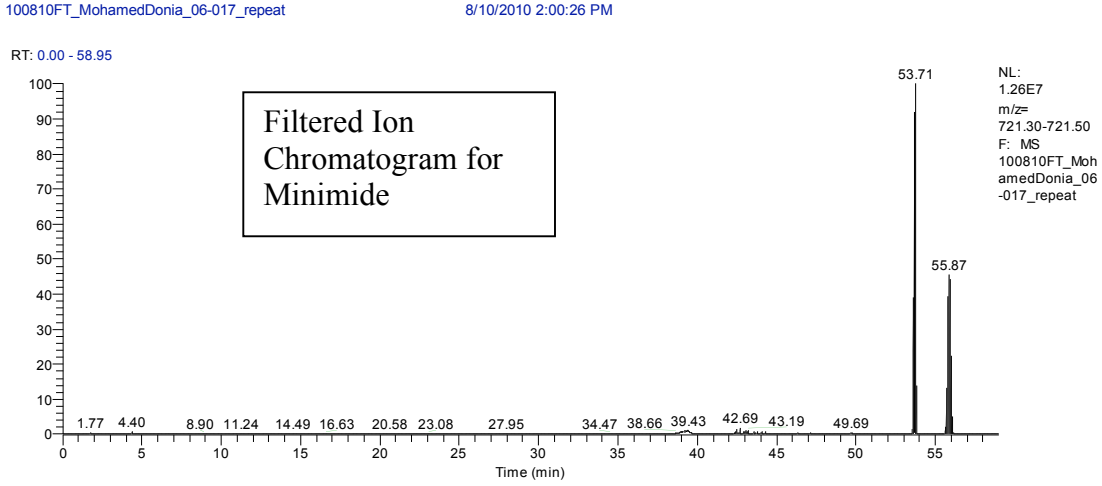
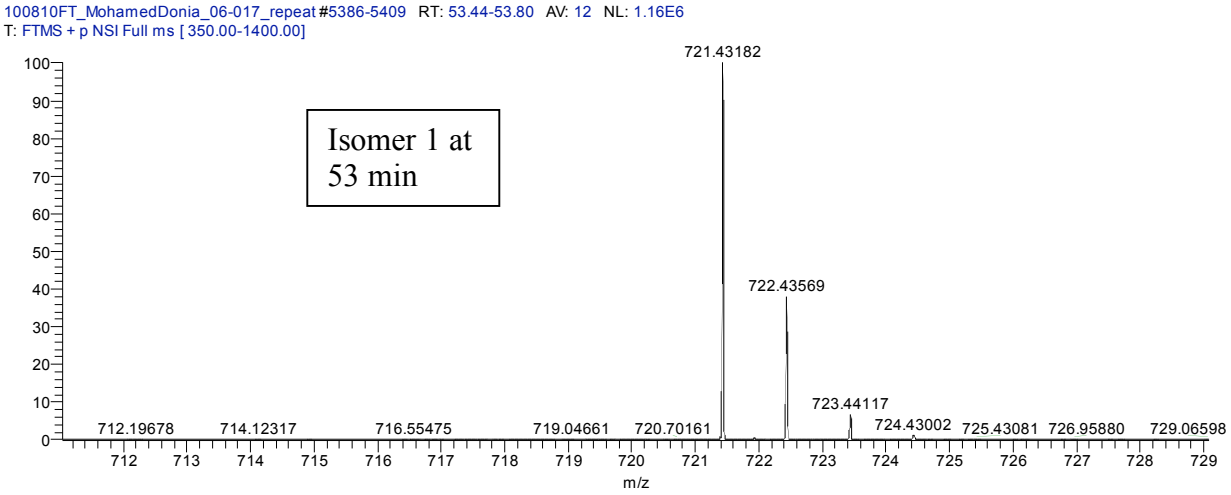
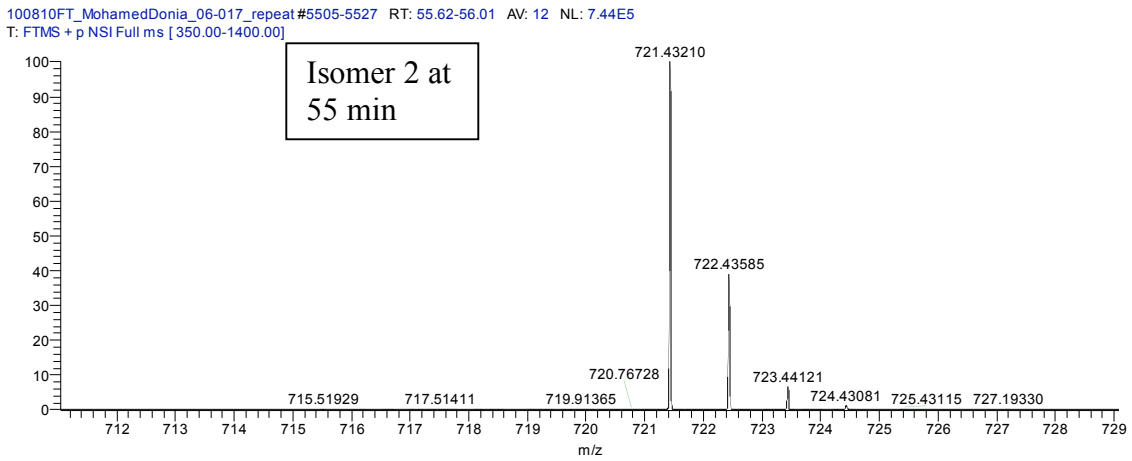


Figure S3. FT-MS analysis of minimide. Full FT-MS data are shown for the natural minimide. Two isomers are apparent (at 53 and 55 min) and have essentially identical exact masses and fragmentation patterns. The IRMPD fragmentation indicates loss of 2 prenyl groups.



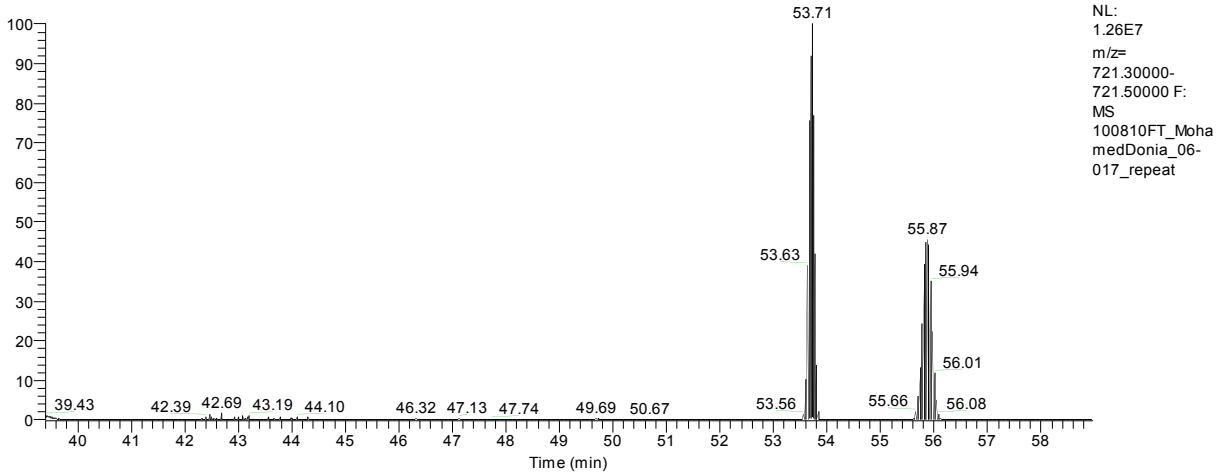



100810FT_MohamedDonia_06-017_repeat #5386-5409 RT: 53.44-53.80 AV: 12 NL: 1.16E6
T: FTMS + p NSI Full ms [350.00-1400.00]
Isomer 1 at 53 min
721.43182
722.43569
723.44117
712.19678
714.12317
716.55475
719.04661
720.70161
724.43002
725.43081
726.95880
729.06598
m/z

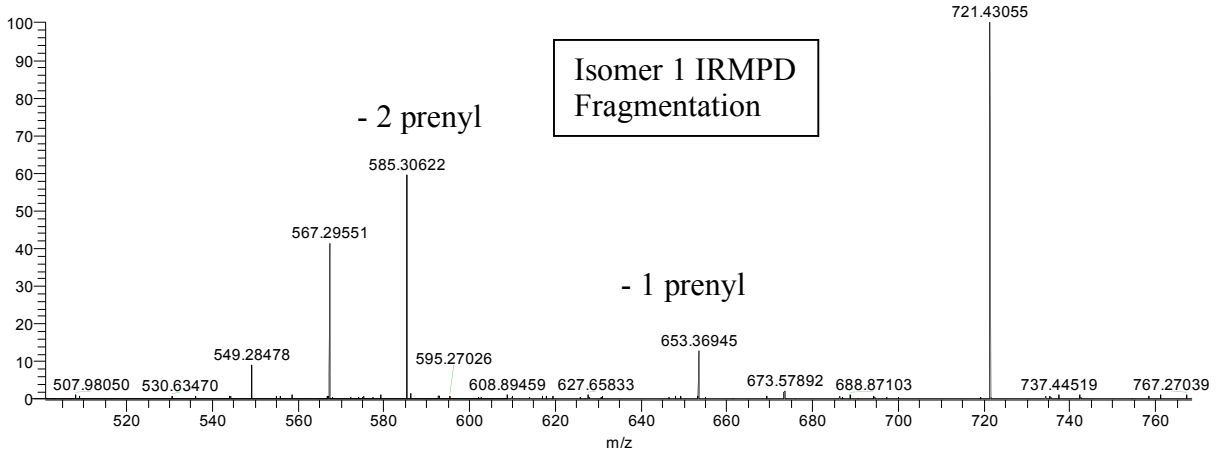

100810FT_MohamedDonia_06-017_repeat #5505-5527 RT: 55.62-56.01 AV: 12 NL: 7.44E5
T: FTMS + p NSI Full ms [350.00-1400.00]
Isomer 2 at 55 min
721.43210
722.43585
720.76728
723.44121
724.43081
715.51929
717.51411
719.91365
725.43115
727.19330
m/z

100810FT_MohamedDonia_06-017_repeat 8/10/2010 2:00:26 PM

RT: 39.39 - 58.95

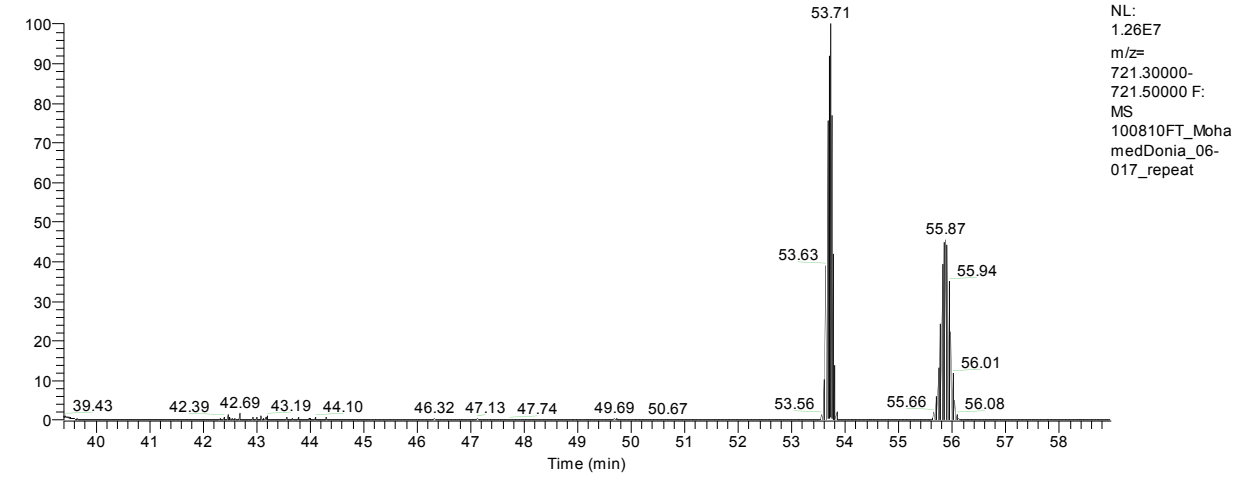


100810FT_MohamedDonia_06-017_repeat #5396-5407 RT: 53.60-53.75 AV: 6 NL: 1.11E4
T: FTMS + p NSI Full ms2 721.40@0.00 mpd@60.00 [mpd@100.00 -mpd@1500.00]



100810FT_MohamedDonia_06-017_repeat 8/10/2010 2:00:26 PM

RT: 39.39 - 58.95



100810FT_MohamedDonia_06-017_repeat #5510-5520 RT: 55.71-55.88 AV: 6 NL: 2.34E3
T: FTMS + p NSI Full ms2 721.40@0.00 mpd@60.00 [mpd@100.00 -mpd@1500.00]

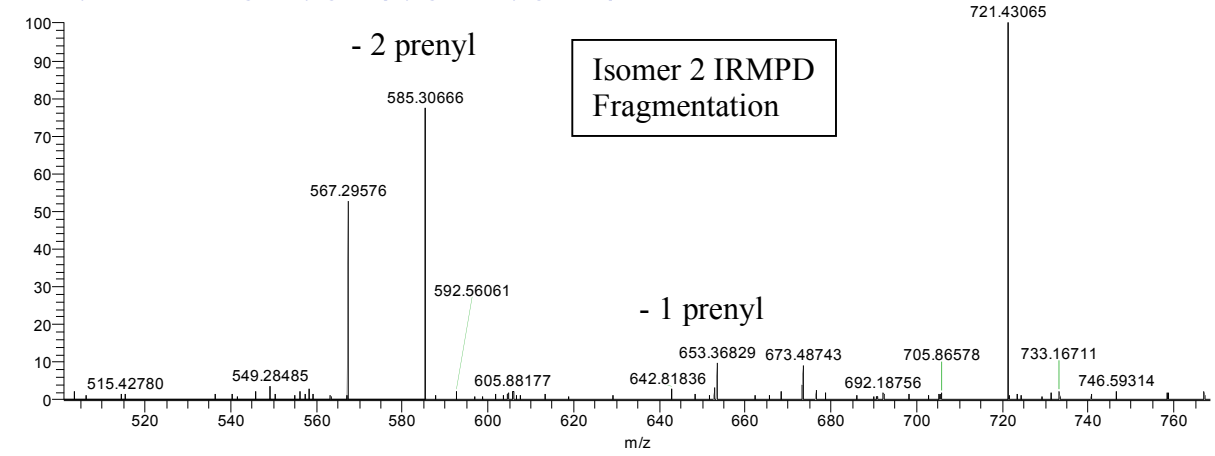


Figure S4. ***E. coli*** **expression systems for cyanobactins.** Black arrows represent the T7 promoter while red arrows represent the *lac* promoter. A) Every gene is under the control of the inducible T7 promoter. This system did not yield products with the *tru* pathway and was variable with the *pat* pathway. B) Half of the pathway is on one plasmid while the other half is on another, also under the control of the inducible T7 promoter. This system also failed to produce compounds. C) The whole pathway is on one operon under the control of the constitutive *lac* promoter. This pathway was functional and was optimized in these studies. D) The *tru* operon was co-expressed with *truE* precursor peptide gene to yield natural products and variants in the optimized system.

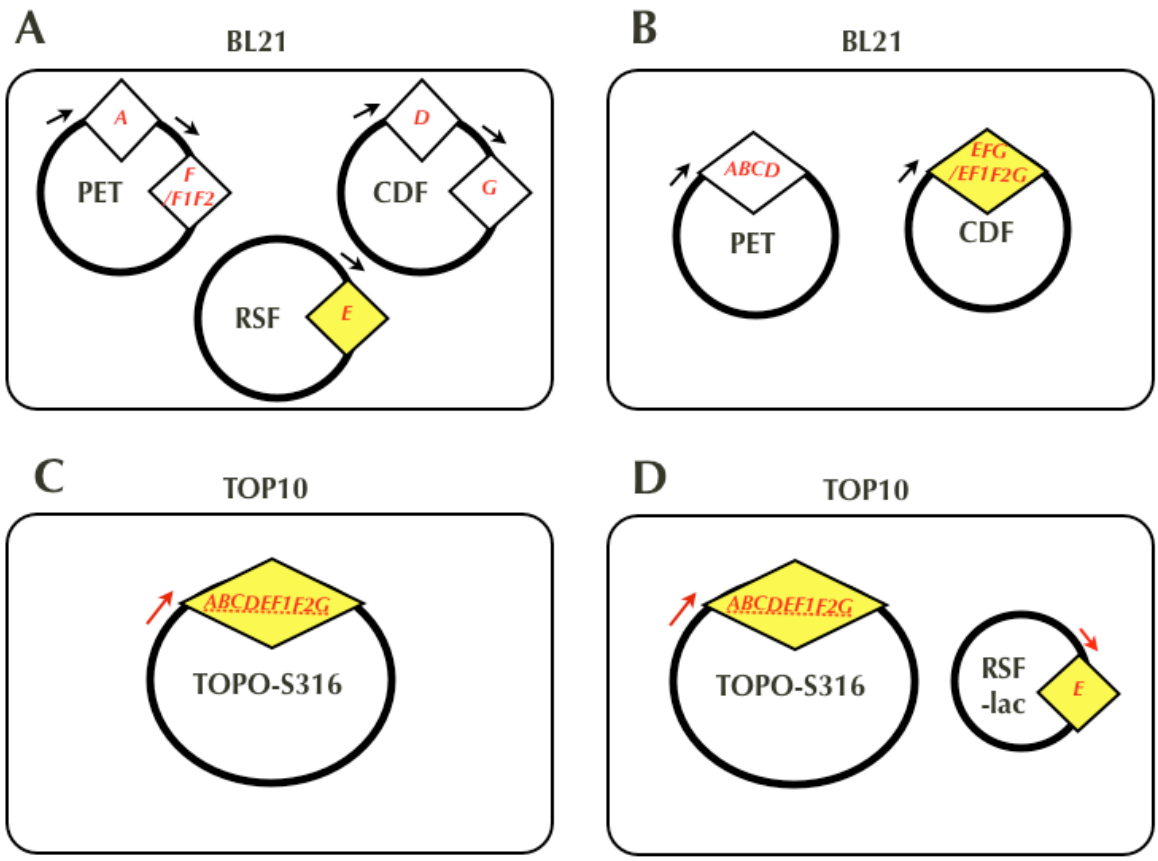


Figure S5. Genetic engineering of a fluorescent fusion construct. Middle, *gfp* was amplified from the pGFP-UV with primers homologous to the final gene of the *tru* pathway (*truG*). Homologous recombination in yeast was then used to construct the fusion construct. Left, restriction digests with NcoI showing the correct insertion. Right, two colonies of Top10 *E. coli* harboring the fusion construct show less fluorescence than the positive control (*gfp* only) and more than the non-fluorescent initial construct (negative control).

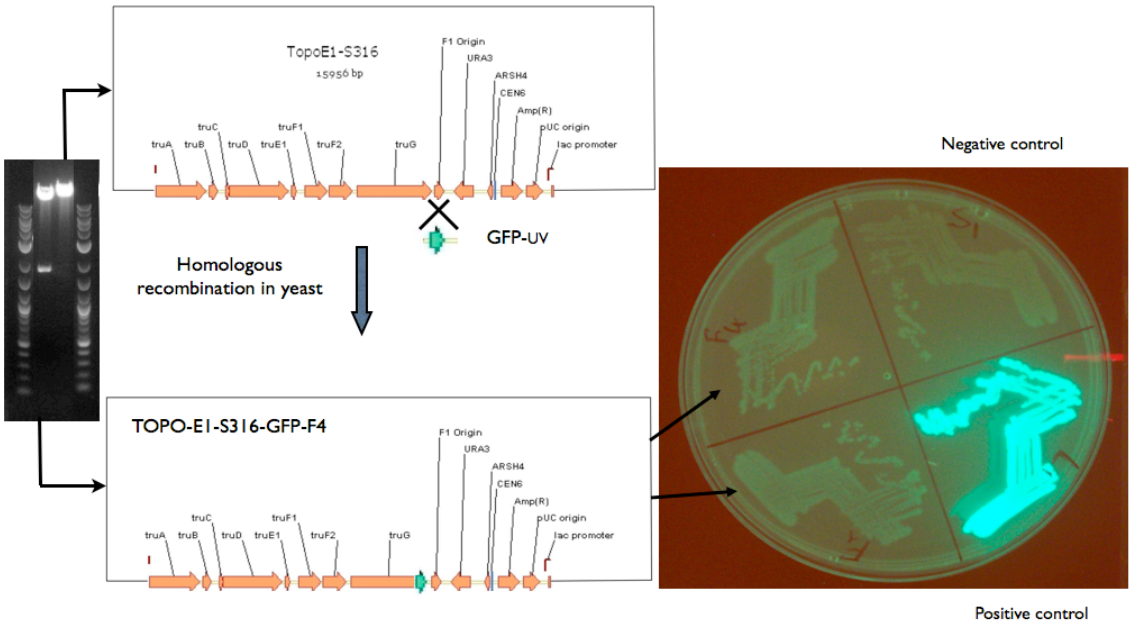


Figure S6. Effect of media on expression levels. A) Relative production levels of patellin 3 assessed by UPLCMS. B) Relative production levels of patellin 2 assessed by UPLCMS. The top panel represents the expression in LB and the bottom represents the expression in 2XYT. The same cell line, colonies, expression conditions and extraction procedures were used for these two experiments. The yield in 2XYT is higher than that in LB.

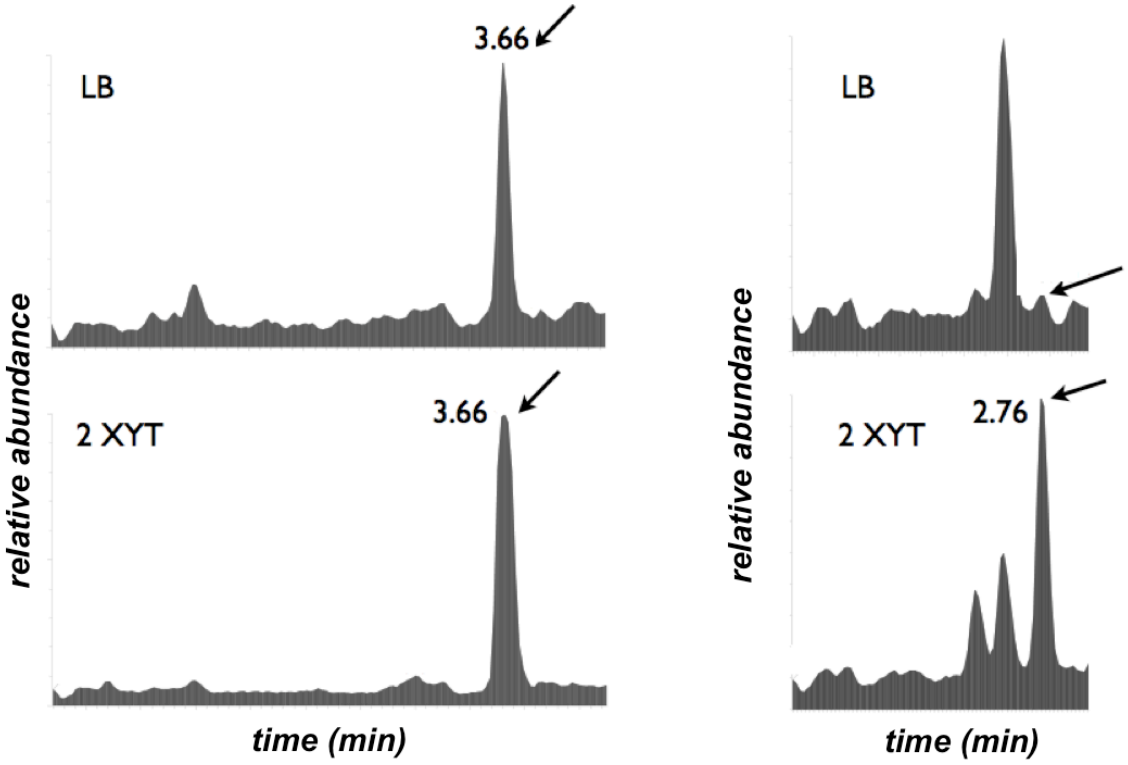


Table S1. Constructs reported in this study.

	Resulting vector	Gene content	Annealing temperature	Extension time	Forward primer	Reverse primer	Parent vector
1	pPET-patApatF	patA (MCS1)	52	2.5 min	patA-F-BspHI	patA-R-NotI	pPET-Duet (Novagen)
		patF (MCS2)	52	1 min	patF-F-NdeI	patF-R-KpnI	
2	pPET-truApatF	truA (MCS1)	52	2.5 min	patA-F-BspHI	truA-R-EagI	pPET-Duet (Novagen)
		patF (MCS2)	52	1 min	patF-F-NdeI	patF-R-KpnI	
3	pCDF-patDpatG	patD (MCS1)	55	2.5 min	patD-F-BsphI	patD-R-NotI	pCDF-Duet (Novagen)
		patG (MCS2)	54	4 min	patG-F-NdeI	patG-R-KpnI	
4	pACYC-patAtruF1F2	patA (MCS1)	52	2.5 min	patA-F-BspHI	patA-R-NotI	pACYC-Duet(Novagen)
		truF1F2 (MCS2)	52	2 min	truF1-F-KpnI	truF2-R-KpnI	
5	pACYC-patAtruF1	patA (MCS1)	52	2.5 min	patA-F-BspHI	patA-R-NotI	pACYC-Duet(Novagen)
		truF1 (MCS2)	52	1 min	truF1-F-KpnI	truF1-R-KpnI	
6	pACYC-patApatF	patA (MCS1)	52	2.5 min	patA-F-BspHI	patA-R-NotI	pACYC-Duet(Novagen)
		patF (MCS2)	52	1 min	patF-F-KpnI	patF-R-KpnI	
7	pRSF-patE2	patE2 (MCS2)	52	30 sec	patE-F-NdeI	patE-F-KpnI	pRSF-Duet (Novagen)
8	pPET-truAtruF1F2	truA (MCS1) X	52	2.5 min	patA-F-BspHI	truA-R-EagI	pPET-Duet (Novagen)
		truF1F2	52	2 min	truF1-F-KpnI	truF2-R-KpnI	
9	pCDF-truDtruG	truD (MSC1)	56	2.5 min	truD-F-BsphI	truD-R-EagI	pCDF-Duet (Novagen)
		truG (MCS2)	54	3.5 min	truG-F-NdeI	patG-R-KpnI	
10	pRSF-truE1	truE1	52	30 sec	patE-F-NdeI	truE-R-KpnI	pRSF-Duet (Novagen)
11	pRSF-truE2	truE2	52	30 sec	patE-F-NdeI	truE-R-KpnI	pRSF-Duet (Novagen)
12	pPET-patI	patApatBpatCpatD	52	6 min	patA-F-NcoI	patD-R-NotI	pPET-Duet (Novagen)
13	pCDF-patII	patE2patFpatG	52	6 min	patE-F-NdeI	patG-R-KpnI	pCDF-Duet (Novagen)

14	pCDF-truII	truE1truF1truF2truG	52	6 min	truF1-F-KpnI	patG-R-KpnI	pCDF-Duet (Novagen)
15	TOPO-E1	tru1	Constructed by TOPO-cloning				pCR2.1 (Invitrogen)
16	TOPO-E1-Dkan	tru1	Constructed by deleting the kanamycin resistance gene from TOPO-E1				NA
17	TOPO-E1-S316	tru1	Constructed by yeast recombination[12]				NA
18	TOPO-E1-S316-TK	tru2	Constructed by yeast recombination [12]				NA
19	TOPO-E1-S316-GFP-F4	tru1-GFP	Constructed by yeast recombination (explained above)				NA
20	pRSF-DlacI-truE1	truE1	Constructed by deleting *lacI* from pRSF-truE2				NA
21	pRSF-DlacI-lacp truE1	truE1	Constructed by cloning lac promoter into pRSF-DlacI-truE2				NA
22	pRSF-DlacI-lacp truE-TK	truE+trunkamide	Constructed by cloning lac promoter into pRSF-DlacI-truE2				NA
23	pRSF-DlacI-lacp-truEminim	TruE + minimide	Constructed by mutational PCR				NA

Table S2. Sequences of the primers used in this study. The underlined sequence is the homologous sequence designed to allow the recombination in yeast.

	Primer name	Sequence (5'-3')
1	patA-F-BspHI	ATCATGAATAGAGATATTTTGCGAAC
2	patA-F-NcoI	AACCATGGATAGAGATATTTTGCGAAC
3	patA-R-EagI	TCGGCCGTTCCTTAGTAAGAAGAAGACCAAG
4	patD-F-BspHI	TTCATGAACCCAACCGCGCTCCAAATTAAG
5	patD-R-NotI	GCCGCGGCCGCAAACTTGAAAATGCTTAAAACG
6	patE-F-NdeI	AACATATGGACTTAAATTGACAGGCTTC
7	patE-R-KpnI	ATGACTAGGTACCTGAGTCAATGCAAATG
8	patF-F-NdeI	AACATATGGACTTAAATTGACAGGCTTC
9	patF-R-KpnI	ATGACTAGGTACCTGAGTCAATGCAAATG
10	patG-F-NdeI	CCATATGATCACGATAGACTACCCTTTC
11	patG-R-KpnI	CGGTACCCCAATAACTACTTTGAGACGGTG
12	truA-R-EagI	AACGGCCGGTTAGTAAGAAGAAGACCAAGAACG
13	truD-F-BspHI	TTCATGCAACCAACCGCCCTCCAAATTAAG
14	truD-R-EagI	AACGGCCGTTAAAATGGCATCGGTGTAGGGTTC
15	truF1-F-Nde1	AGGACATATGATTATGACTACTACTTGGCCAG
16	truF1-R-KpnI	TTTGGTACCCTATACCGCCTTGCGATAATAG
17	truF2-R-KpnI	TTTGGTACCTCAGACTCTCATCATTCTTGACT
18	truG-F-NdeI	TTCATATGAGTCGTCATCCTTTTAATATTTGCC
19	truG-R-KpnI	CGGTACCCCAATAACTACTTTGAGACGGTG
20	truE-F-NdeI	AACATATGGACTTAAATTGACAGGCTTC
21	truE-R-KpnI	AATTCGGTACCTTAGTCGTCGTAAGAGCAGAG
22	TruG-LacP-GFP-F	<u>TTATTGGGGTACCGAAGGGC</u>GAAAGCGGGCAGT GAGCGCAA
23	TruG-Fusion-GFP-F	<u>CCTGGTAAAGGCACCGTCTCAAAG</u>ATGAGTAAA GGAGAAGAACTTTTCAC
24	GFP-F1Origin-R	<u>ACTGTTGGGAAGGGCGATCGGT</u>AGTTGGAATTC ATTATTTGTAGAGCTC

Table S3. Compounds synthesized in this study. Compounds obtained and validated from *E. coli* expression experiments are listed below.

Compound	Expression construct	Source	Sequence
Trunkamide	Operon and pRSF	Known ascidian natural product	TSIAPFC
Patellin 2	Operon	Known ascidian natural product	TVPTLC
Patellin 3	Operon and pRSF	Known ascidian natural product	TLPVPTLC
Patellin 6	Operon	Known ascidian natural product	TFPVPTVC
Minimide	pRSF	New ascidian natural product	TLATIC