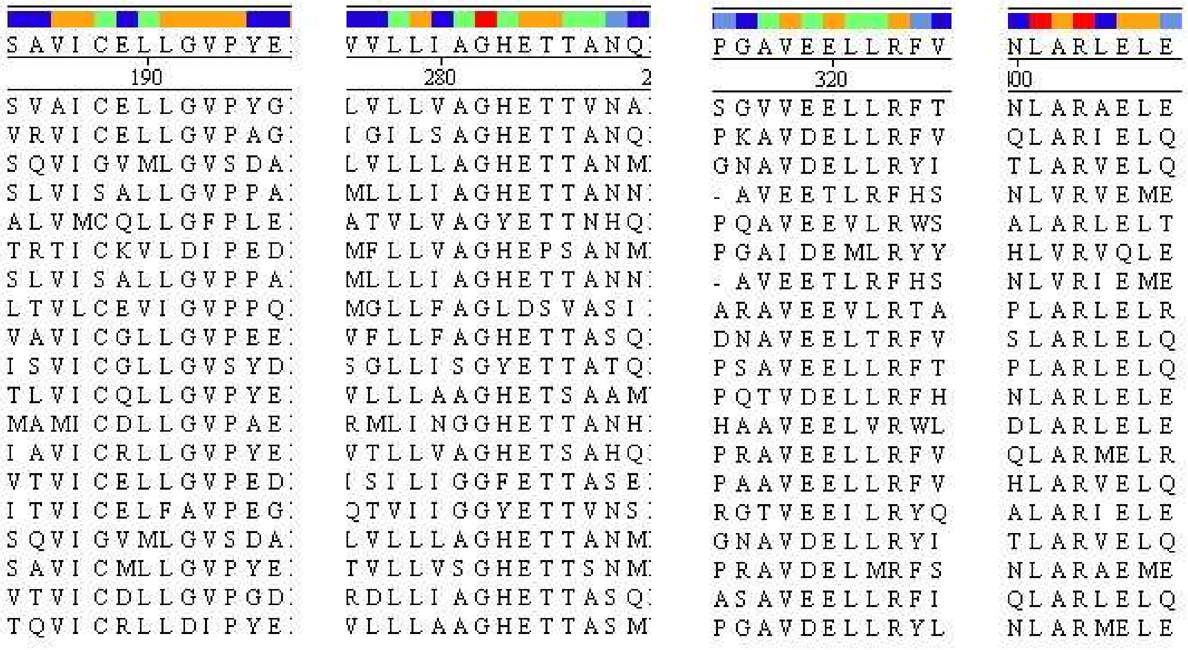
CHEMBIOCHEM

Supporting Information

Identification of the Tirandamycin Biosynthetic Gene Cluster from *Streptomyces* sp. 307-9

Jacob C. Carlson, J. L. Fortman, Yojiro Anzai, Shengying Li, Douglas A. Burr, and
David H. Sherman*[a]

cbic_200900658_sm_miscellaneous_information.pdf



SAVICELLGVPYE	VVLLIAGHETTANQ	PGAVEELLRFV	NLARLELE
190	280 2	320	00
SVAICELLGVPYG	LVLLVAGHETTVNA	SGVVEELLRFT	NLARAELE
VRVICELLGVPAG	GILSAGHETTANQ	PKAVDELLRFV	QLARIELQ
SQVIGVMLGVSDA	LVLLLAGHETTANM	GNAVDELLRYI	TLARVELQ
SLVISALLGVPPA	MLLLIAGHETTANN	-AVEETLRFHS	NLVRVEME
ALVMCQLLGFPLE	ATVLVAGYETTNHQ	PQAVEEVLRWS	ALARLELT
TRTICKVLDIPED	MFLLVAGHEPSANM	PGAIDEMLRYY	HLVRVQLE
SLVISALLGVPPA	MLLLIAGHETTANN	-AVEETLRFHS	NLVRIEME
LTVLCEVIGVPPQ	MGLLFAGLDSVASI	ARAVEEVLRTA	PLARLELR
VAVICGLLGVPEE	VFLLFAGHETTASQ	DNAVEELTRFV	SLARLELQ
ISVICGLLGVSYD	GLLISGYETTATQ	PSAVEELLRFT	PLARLELQ
TLVICQLLGVPYE	VLLLAAGHETSAAM	PQTVDELLRFH	NLARLELE
MAMICDLLGVPAE	RMLINGGHETTANH	HAAVEELVRWL	DLARLELE
IAVICRLLGVPYE	VTLLVAGHETSAHQ	PRAVEELLRFV	QLARMELR
VTVICELLGVPED	SILIGGFETTASE	PAAVEELLRFV	HLARVELQ
ITVICELFAVPEG	QTVIIGGYETTVNS	RGTVEEILRYQ	ALARIELE
SQVIGVMLGVSDA	LVLLLAGHETTANM	GNAVDELLRYI	TLARVELQ
SAVICMLLGVPYE	TVLLVSGHETTSNM	PRAVDELMRFS	NLARAEME
VTVICDLLGVPGD	RDLLIAGHETTASQ	ASAVEELLRFI	QLARLELQ
TQVICRLLDIPYE	VLLLAAGHETTASM	PGAVDELLRYL	NLARMELE

Design of degenerate primers for two center conserved regions:

P450 Forward

Conserved amino acids: I C E/D L L/V G V/I P

Corresponding codons: A,T,C | T,G,C | G,A,I | C,T,G/C | C,T,G/C | G,G,I | A/G,T,G/C | C,C,G/C

Final primer sequence: ATCTGCGAICTSCTSGGIRTSCCS (32-fold degeneracy)

P450 Reverse

Conserved amino acids: R L/M L/V E E/D V A/V

Corresponding codons: I,C,G | G/C,A,G/T | G/C,A,I | C/T,T,C | I,T,C | G/C,A,C | G/C,I,C

Final primer sequence: ICGSAKSAIYTCITCSACSIC (64-fold degeneracy)

KS Forward (previously reported)

Primer 4UU: MGIGARGCIYTICARATGGAYCCICARCARMG

KS Reverse (previously reported)

Primer 5LL: GGRTCNCCIARITGIGTICCIGTICCRTGIGC

Figure S1. Alignment of bacterial biosynthetic cytochrome P450 enzymes (top) and the degenerate primers designed from the center two conserved regions (bottom).

ORF	**Start**	**End**	**Size**[a]	**Closest Homologue and Origin**	**%I**[b]	**%S**[c]	**Proposed Function**
ORF31	2	154 r	50	BAG17552.1 *Streptomyces griseus* subsp. griseus NBRC 13350	86	93	Two component response regulator
ORF30	878	1606	242	BAC75001.1 *Streptomyces avermitilis* MA-4680	41	55	Conserved hypothetical membrane protein
ORF29	1599	2462 r	287	CAJ88687.1 *Streptomyces ambofaciens* ATCC 23877	79	87	Formamidopyridine glycosylase
ORF28	2487	3896 r	469	YP_001822237.1 *Streptomyces griseus* subsp. griseus NBRC 13350	71	77	Conserved hypothetical protein
ORF27	3995	6067 r	690	YP_001822238.1 *Streptomyces griseus* subsp. griseus NBRC 13350	79	89	Protein phosphatase
ORF26	6120	7397 r	425	YP_001822239.1 *Streptomyces griseus* subsp. griseus NBRC 13350	44	57	Conserved hypothetical protein
ORF25	8416	9123	235	YP_001822241.1 *Streptomyces griseus* subsp. griseus NBRC 13350	69	74	Conserved hypothetical protein
ORF24	9165	10523	452	NP_828463.1 *Streptomyces avermitilis* MA-4680	73	82	Conserved hypothetical transmembrane protein
ORF23	11126	12589 r	487	YP_001822244.1 *Streptomyces griseus* subsp.griseus NBRC 13350	87	92	Short-chain dehydrogenase
ORF22	13624	15042 r	472	YP_001822257.1 *Streptomyces griseus* subsp. griseus NBRC 13350	63	73	Peptidase
ORF21	15108	16550	480	YP_001106590.1 *Saccharopolyspora erythraea* NRRL 2338	81	87	6-phosphogluconate dehydrogenase
ORF20	18418	18852 r	144	YP_001822287.1 *Streptomyces griseus* subsp. griseus NBRC 13350	89	92	Aspartate decarboxylase
ORF19	19639	20754	371	AF126281_1 *Rhodococcus erythropolis*	40	57	Transposase
ORF18	21155	22642 r	495	YP_480214.1 *Frankia* sp. CcI3	63	74	Conserved hypothetical protein
TamM	23201	23836 r	211	YP_001106197.1 *Saccharopolyspora erythraea* NRRL 2338	52	65	Phosphopantetheinyl transferase
TamL	23970	25472	500	CAJ42334.1 *Streptomyces steffisburgensis*	58	71	FAD-dependent oxidase
TamAI	25678	43164	5828	NP822114 *Streptomyces avermitilis* MA-4680	45	57	PKS (4 modules)
TamAII	43304	53401	3365	AAZ94390 *Streptomyces neyagawaensis*	55	67	PKS (2 modules)
TamAIII	53406	61811	2801	AAX98191 *Streptomyces aizunensis*	57	68	PKS (2 modules)
TamB	61808	62584	258	NP_821582.1 *Streptomyces avermitilis* MA-4680	53	68	Type II thioesterase
TamC	62776	63606	276	CAN89637.1 *Streptomyces collinus*	44	56	Conserved hypothetical protein (LipX2 homologue)
TamD	63723	66878	1051	CAN89638.1 *Streptomyces collinus*	49	62	NRPS
TamE	66902	67732	276	YP_946579.1 *Arthrobacter aurescens* TC1	35	49	Glycoside hydrolase
TamF	67729	69267	512	BAD86797.1 *Streptomyces* sp. KO-3988	39	55	Terpene synthase
TamG	69758	71851	697	NP_825193.1 *Streptomyces avermitilis* MA-4680	63	74	DNA helicase
TamH	72117	74930	937	AAX98177.1 *Streptomyces aizunensis*	37	51	LuxR-type transcriptional activator
TamI	74884	76125 r	413	YP_001107923.1 *Saccharopolyspora erythraea* NRRL 2338	55	66	Cytochrome P450 monooxygenase
TamJ	76432	77928 r	498	NP_825134.1 *Streptomyces avermitilis* MA-4680	56	69	Efflux pump
TamK	78090	78722	210	CAJ89338.1 *Streptomyces ambofaciens* ATCC 23877	59	73	TetR-type transcriptional repressor
ORF1	79330	79755	141	AAZ23099.1 *Streptomyces fradiae*	88	92	Transposase N-terminal fragment
ORF2	79806	80204	132	AAZ23099.1 *Streptomyces fradiae*	95	98	Transposase fragment
ORF3	80298	80672	124	YP_001821630.1 *Streptomyces griseus* subsp. griseus NBRC 13350	78	82	Transposase fragment
ORF4	80771	81127	118	YP_001821630.1 *Streptomyces griseus* subsp. griseus NBRC 13350	85	89	Transposase fragment
ORF5	81607	82227 r	206	NP_625296.1 *Streptomyces coelicolor* A3(2)	95	97	Phage integrase C-terminal fragment
ORF6	82451	83122	223	NP_627673.1 *Streptomyces coelicolor* A3(2)	44	59	Transposase
ORF7	83126	83728 r	200	AAD33129.1 *Streptomyces coelicolo*r A3(2)	92	95	Phage integrase
ORF8	84983	85483 r	166	CAJ88710.1 *Streptomyces ambofaciens* ATCC 23877	97	98	Conserved hypothetical protein
ORF9	85639	86034 r	131	CAJ88711.1 *Streptomyces ambofaciens* ATCC 23877	93	96	Conserved hypothetical protein
ORF10	86149	86337 r	62	CAJ88711.1 *Streptomyces ambofaciens* ATCC 23877	91	93	Conserved hypothetical protein
ORF11	86565	86735	56	CAJ89311.1 *Streptomyces ambofaciens* ATCC 23877	87	92	50S ribosomal protein
ORF12	86742	87929	395	YP_001159829.1 *Salinispora tropica* CNB-440	60	69	Cobalamin biosynthesis
ORF13	88273	88578 r	101	YP_001822056.1 *Streptomyces griseus* subsp. griseus NBRC 13350	80	91	30S ribosomal protein
ORF14	88578	88910 r	110	YP_001822057.1 *Streptomyces griseus* subsp. griseus NBRC 13350	87	94	50S ribosomal protein
ORF15	90021	90371	116	NP_627722.1 *Streptomyces coelicolor* A3(2)	86	91	ARS regulatory protein
ORF16	90374	91147	257	YP_001825281.1 *Streptomyces griseus* subsp. griseus NBRC 13350	84	92	Cation efflux pump
ORF17	91555	92124	189	CAJ89573.1 *Streptomyces ambofaciens* ATCC 23877	66	79	GCN5 acetyltransferase

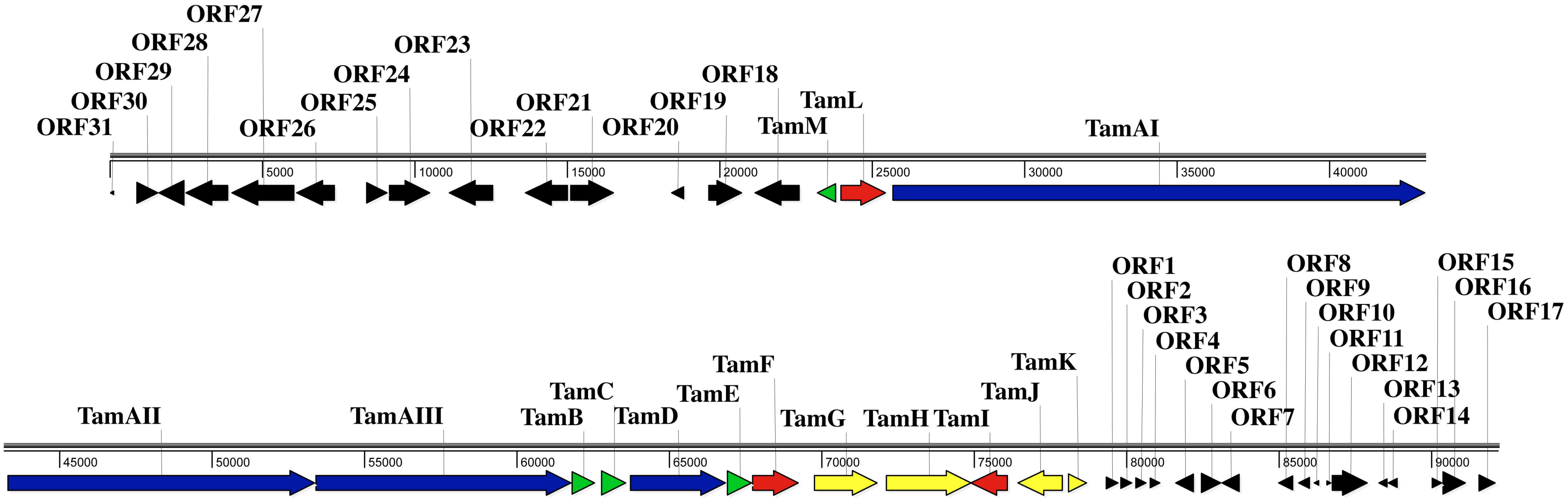


Figure S2. Complete list of ORF assignments within the 92 kb of sequence derived from *Streptomyces* sp. 307-9 gDNA (top) and a map of these ORFS (bottom) showing their relationship to the tirandamycin biosynthetic gene cluster (in color).

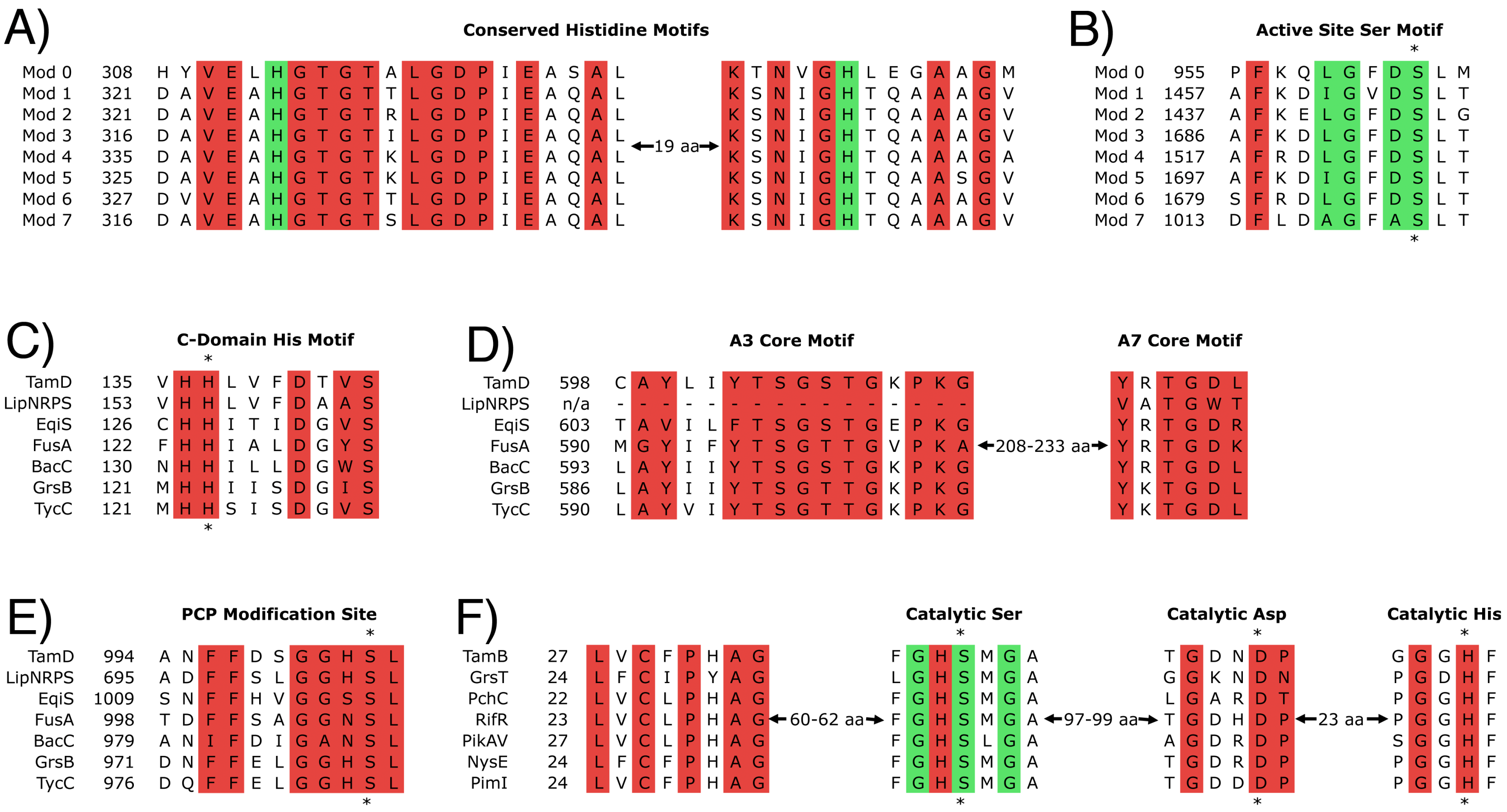


Figure S3. Sequence motifs consistent with fully active assembly line enzymes. A) Conserved KS domain motifs; residues in green denote presumed active site histidine residues, residues in red are conserved in all known active domains. B) ACP residues; LGxDS motif (green) and otherwise conserved residues (red). C) Conserved condensation domain residues (red) and the active site histidine (asterisk). D) Adenylation domain residues from the A3 and A7 core motifs; residues in red are well conserved in active domains. E) Conserved PCP domain residues. F) Type II thioesterase motifs, including the conserved GxSxG motif of a/b hydrolases (green) and other conserved residues (red).

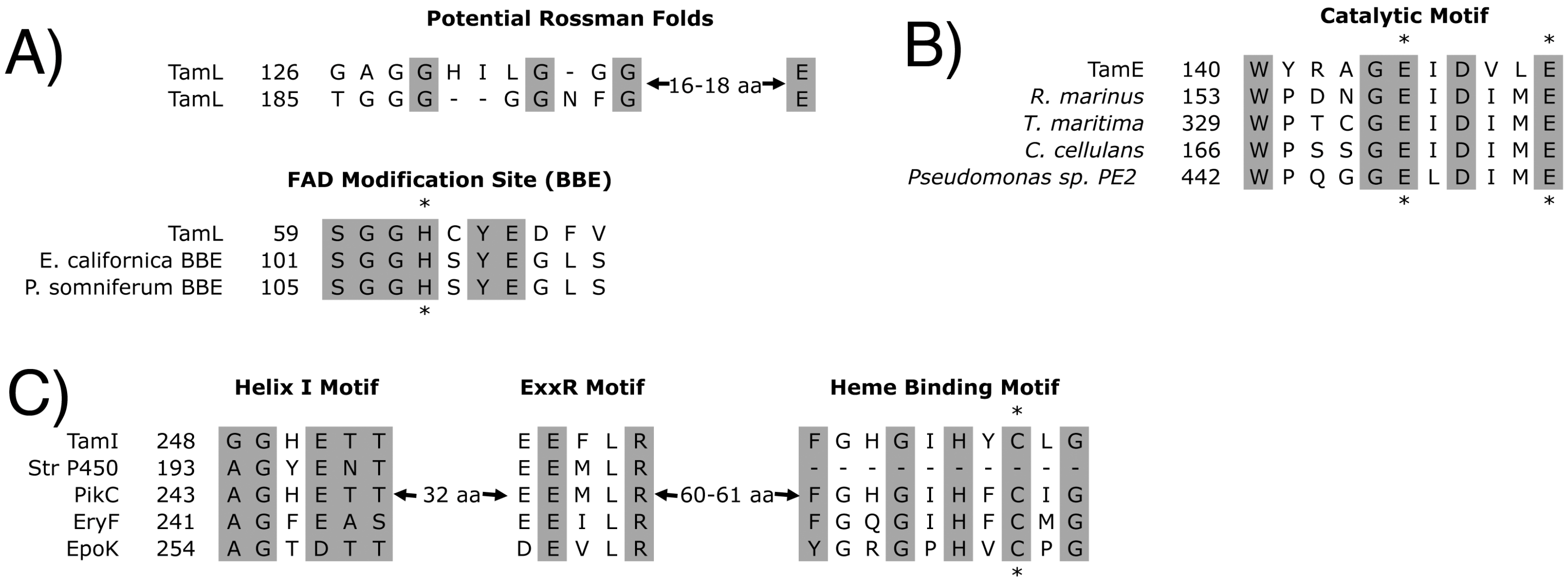


Figure S4. Sequence analysis of tirandamycin tailoring enzymes. A) TamL sequence includes potential Rossmann folds (top) and a berberine bridge enzyme motif for flavin cofactor attachment (45); aligned with GenBank IDs P30986 and P93479. B) TamE sequence motif containing conserved catalytic residues from bacterial glycoside hydrolases; aligned with GenBank IDs AAA60459, NP_227840, AAC38290, and BAC16332. C) TamI sequence motifs typical of cytochrome P450 enzymes; aligned with GenBank IDs ABB92562, AAC68886, AAA26496, and AAF62886.