

## SUPPLEMENTAL DATA

The Biosynthesis of Capuramycin-Type Antibiotics: Identification of the A-102395 Biosynthetic Gene Cluster, Mechanism of Self-Resistance, and Formation of Uridine-5'-Carboxamide\*

**Wenlong Cai<sup>1</sup>, Anwasha Goswami<sup>1</sup>, Zhaoyong Yang<sup>2</sup>, Xiaodong Liu<sup>1</sup>, Keith D. Green<sup>1</sup>, Sandra Barnard-Britson<sup>1</sup>, Satoshi Baba<sup>3</sup>, Masanori Funabashi<sup>4</sup>, Koichi Nonaka<sup>5</sup>, Manjula Sunkara<sup>6</sup>, Andrew J. Morris<sup>6</sup>, Anatol P. Spork<sup>7</sup>, Christian Ducho<sup>7</sup>, Sylvie Garneau-Tsodikova<sup>1</sup>, Jon S. Thorson<sup>1</sup>, and Steven G. Van Lanen<sup>1</sup>**

<sup>1</sup>Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky

<sup>2</sup>Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences and Peking Union Medical College

<sup>3</sup>New Modality Research Laboratories, R&D Division, Daiichi Sankyo Co., Ltd.

<sup>4</sup>Drug Discovery and Biomedical Technology Unit, Daiichi Sankyo RD Novare Co., Ltd.

<sup>5</sup>Biologics Technology Research Laboratories, R&D Division, Daiichi Sankyo Co., Ltd.

<sup>6</sup>Division of Cardiovascular Medicine and Gill Heart Institute, College of Medicine, University of Kentucky

<sup>7</sup>Department of Pharmacy, Pharmaceutical and Medicinal Chemistry, Saarland University

\*Running Title: *Biosynthesis of capuramycin-type antibiotics*

To whom correspondence should be addressed: Steven G. Van Lanen, Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, 789 S. Limestone, Lexington, KY, USA, Tel.: (859) 323-6271; Email: [svanlanen@uky.edu](mailto:svanlanen@uky.edu)

Keywords: Antibiotics, antibiotic resistance, biosynthesis, translocase 1, dioxygenase, transaldolase, phosphotransferase, natural product biosynthesis

**SUPPLEMENTAL TABLE** Annotation of ORFs within the 3 gene cluster.

<b>Protein<sup>a</sup></b>	<b>Size<sup>b</sup></b>	<b>Proposed function</b>	<b>Homolog (accession no.)<sup>c</sup></b>	<b>Identity / Similarity<sup>d</sup> (%)</b>
ORF1	319	Cell-wall associated hydrolase	Svir_07310 (YP_003132627)	44/58
ORF2	363	Hypothetical protein	Francci3_2715 (YP_481805)	49/59
ORF3	855	Hypothetical protein	Franean1_2549 (YP_001506885)	58/69
ORF4	639	Type IV secretory pathway VirB4	Franean1_2548 (YP_001506884)	60/73
ORF5	398	Hypothetical protein	Franean1_2547 (YP_001506883)	57/66
ORF6	774	Hypothetical protein	Caci_4083 (YP_003114789)	46/64
ORF7	107	Hypothetical protein	Franean1_2741 (YP_001507073)	58/77
ORF8	322	Hypothetical protein	Franean1_2740 (YP_001507072)	45/56
ORF9	247	Hypothetical protein	SACE_2308 (YP_001104534)	47/56
ORF10	350	Hypothetical protein	SACE_2307 (YP_001104533)	70/80
ORF11	233	Putative cellulase	SCAB_16431 (CBG68778)	24/41
Cpr12	259	Aminotransferase class IV	Caci_2606 (YP_003113364)	45/56
ORF13	162	Putative short chain dehydrogenase	MAB_3919c (YP_001704647)	48/62
ORF14	373	Transposase	RHA1_ro08502 (YP_707704)	76/85
ORF15	122	Transposase	Svir_17380 (YP_003133592)	75/85
ORF16	123	Transposase	SSAG_01038 (ZP_04996736)	68/77
Cpr17	305	Capuramycin 3"-phosphotransferase	CapP <sup>e</sup>	53/64
Cpr18	331	Fe <sup>2+</sup> -dependent, $\alpha$ KG dioxygenase	CapD	63/74
Cpr19	275	Fe <sup>2+</sup> -dependent, $\alpha$ KG:UMP dioxygenase	CapA	76/88
Cpr20	317	putative 3-ketoreductase	CapC	73/81
Cpr21	457	putative 2,3-dehydratase	CapE	73/83
Cpr22	313	putative 4-epimerase	CapF	70/79
Cpr23	378	SelA-related PLP-dependent enzyme	ORF15	47/65
Cpr24	384	Putative glycosyltransferase	CapG	77/86
Cpr25	412	L-Thr:uridine-5'-aldehyde transaldolase	CapH	79/90
Cpr26	87	Putative pyrophosphatase	CapI	67/81
Cpr27	261	Putative carboxyl methyltransferase	CapS	65/79
Cpr28	774	Putative CO dehydrogenase	CapJ	75/84
Cpr29	244	Putative O-methyltransferase	CapK	71/84
Cpr30	165	Putative CO dehydrogenase	CapL	76/89
Cpr31	289	Putative CO dehydrogenase	CapM	60/72
Cpr32	350	Luciferase-like monooxygenase	LLB_0745 (ZP_06185947)	43/56
Cpr33	234	3-oxoacyl-[acyl-carrier-protein] reductase	Noca_2571 (YP_923762)	57/71
Cpr34	391	$\beta$ -Ketoacyl synthase	ShygA5_010100009994 (ZP_05513707)	53/63
Cpr35	257	Hypothetical protein	SCAB_69811 (CBG73975)	44/63
Cpr36	84	Aryl carrier protein	AcmD (AAD30112)	35/55
Cpr37	466	Actinomycin synthetase I	AcmA (AAD30111)	49/62
Cpr38	663	4-amino-4-deoxychorismate synthase	PabAB (P32483)	55/64
Cpr39	521	ATP-dependent synthetase and ligase	SrosN15_010100025847 (ZP_04696386)	50/61
Cpr40	311	Dehydrogenase E1, TPP-dependent	Caci_5069 (YP_003115770)	55/70

Table S5 continued.

Cpr41	338	Pyruvate dehydrog. (acetyl-transferring)	Sros_2206 (YP_003337932)	55/69
Cpr42	365	Dehydrogenase	Npun_F31621 (YP_001866563)	38/57
Cpr43	433	Glycosyl transferase family 2	MicauDRAFT 1168	34/49
Cpr44	612	Acetolactate synthase, large subunit	Vapar_2048 (YP_002943959)	26/42
Cpr45	227	Phosphatase (HAD superfamily)	Yaldo0001_22190 (ZP_04621838)	39/56
Cpr46	599	TPP-dependent synthase/transketolase	PlaT6 (ABB69756)	45/57
Cpr47	448	Condensation domain (NRPS)	MicauDRAFT_1165 (ZP_06216171)	29/42
Cpr48	97	Acyl carrier protein	SrosN15_010100019301 (ZP_04695086)	46/58
Cpr49	314	MitI transglutaminase	MitI (AAD28463)	29/48
Cpr50	395	MitI transglutaminase	MitI (AAD28463)	33/44
Cpr51	395	Putative transacylase	CapW	65/75
Cpr52	191	Hypothetical protein	SrosN15_010100025822 (ZP_04696381)	38/51
Cpr53	381	Acyl-CoA dehydrogenase	SrosN15_010100025827 (ZP_04696382)	45/57
Cpr54	515	Long-chain-fatty-acid-CoA ligase	MCAG_01785 (ZP_04605528)	75/82
Cpr55	84	Acyl carrier protein	VinL (BAD08369)	44/63
Cpr56	424	Orn/DAP/Arg decarboxylase 2	MCAG_01775 (ZP_04605518)	74/83
Cpr57	293	MitI transglutaminase	MitI (AAD28463)	29/45
ORF58	596	ABC transporter ATP-binding protein	MCAG_02642 (ZP_04606385)	37/52
ORF59	608	ABC transporter ATP-binding protein	SghaA1_010100034323 (ZP_04690305)	53/65
ORF60	594	ABC transporter related protein	Cfla_3590 (YP_003638660)	45/62
ORF61	663	ABC transporter related protein	Cfla_3591 (YP_003638661)	44/58
ORF62	441	MATE efflux family protein	Amir_4280 (YP_003101980)	57/71
ORF63	391	AstB/chuR/nirj-related protein	SghaA1_010100036797 (ZP_04690774)	74/83
ORF64	85	Hypothetical protein	SghaA1_010100036802 (ZP_04690775)	55/72
ORF65	352	Hypothetical protein	StAA4_010100023552 (ZP_05481114)	55/63
ORF66	320	Oxidoreductase domain protein	Amir_4534 (YP_003102223)	51/63
ORF67	219	Putative phosphotransferase	StAA4_010100023547 (ZP_05481113)	50/61
ORF68	358	Hypothetical protein	StAA4_010100023542 (ZP_05481112)	53/67
ORF69	420	Hypothetical protein	Sros_6837 (YP_003342286)	29/43
ORF70	179	Hypothetical protein	StAA4_010100029435 (ZP_05482273)	57/75

<sup>a</sup> Sequences deposited at NCBI (accession no. KP995196).

<sup>b</sup> Numbers are in amino acids.

<sup>c</sup> Given in brackets are accession numbers.

<sup>d</sup> % sequence identity/similarity for the entire protein.

<sup>e</sup> Putative proteins within shaded rows have homologues within the A-503083 gene cluster (accession no. AB538860) and A-500359 gene cluster (accession no. AB476988).