

Center for Structural Genomics of Infectious Diseases

Contract No.: HHSN272200700057C

Seattle Structural Genomics Center for Infectious Disease

Principal Investigator: Peter J. Myler, Ph.D.

Seattle Biomedical Research Institute

307 Westlake Ave N

Suite 500

Seattle, WA 98109-5219

Tel.: (206) 256-7332

FAX: (206) 256-7229

e-mail: peter.myler@sbri.org

First report of Targets Proposed for Structure Determination

Author: Peter J. Myler

Date: March 07, 2007

Amended:

Target Selection process

For the second round (Batch02) of Target Selection, we focused our efforts on identifying additional drug candidates from bacterial genera on the Biodefense and Emerging Infections (BEI) list. A list of 42 bacterial drug targets (see Table 1) being actively pursued by pharmaceutical and academic researchers was compiled by literature survey. Orthologues of these drug targets were identified in representative sequenced genomes within the Actinobacteria (*Mycobacterium tuberculosis* H37Rv), Alphaproteobacteria (*Bartonella henselae* str. Houston-1, *Brucella melitensis* biovar Abortus 2308, and *Rickettsia prowazekii* str. Madrid E), Betaproteobacteria (*Burkholderia pseudomallei* 1710b), and Spirochaetes (*Borrelia burgdorferi* B31).

Table 1 – Manually selected bacterial drug targets

2,4- cyclodiphosphate synthase	fatty_acid_synthase
2-c-methyl-d-erythritol 2,4-cyclodiphosphate synthase	glucosamine-6-phosphate synthase
7-cyano-7-deazaguanine reductase	glutamate racemase
acetylmuramoyl:l-alanine ligase	GTP cyclohydrase II
acyl carrier protein synthase	integrase/recombinase
adenide deaminase	integration host factor
alpha riboswitch	leucine transporter
aminodeoxychorismate synthase	malonyl-coa-acyl carrier protein transacylase
anthranilate synthase	penicillin-binding proteins
beta-ketoacyl-acyl carrier protein reductase	peptidoglycan glycosyltransferases
chorismate mutase	phosphopantetheine adenyltransferase
chorismate synthase	site-specific tyrosine recombinase
conjugative DNA relaxase	tmRNA-small protein b
conjugative DNA transferase	topoisomerase 1a
dephosphocoenzyme a kinase	topoisomerase 1b
DNA gyrase	tRNA synthetase
DNA helicase	tyrosine recombinases
DNA primase	UDP-3-o-[3-hydroxymyristoyl] n-acetylglucosamine deacetylase
DNA relaxase	UDP-n-acetylglucosamine enolpyruvyltransferase
enoyl-(acyl carrier protein) reductase	UDP-n-acetylglucosamine pyrophosphorylase
f like relaxase	uridine diphospho-n-acetylenolpyruvylglucosamine reductase

The first step in identifying the corresponding orthologues consisted of text searching the *Burkholderia pseudomallei* 1710a genome via ENTREZ. No orthologues were found for 16 targets (highlighted in yellow in Table 1), while the

other 26 targets revealed 1-11 orthologues each. These were subsequently used for BLAST searches of each reference genome to identify a total of 316 candidates. These candidates were then screened through a number of filters to eliminate proteins longer than 750 amino acids; containing more than 8 cysteine residues; and/or containing any transmembrane spanning domains (predicted using TMPRED and/or TMHMM/PHOBIUS), except for N-terminal signal sequences (Table 2). The 204 remaining candidate proteins were BLASTED against TARGETDB and those that showed >95% similarity over >80% of their length eliminated. This ensured that none of the targets had been selected for structure determination at any other center. Thirty-six proteins (mostly from *M. tuberculosis*) were eliminated at this stage. The remaining protein sequences were BLASTED against those in Batch01 to remove 11 targets previously selected, and then against each other to determine (with a cut-off of 75% similarity over 75% of their length) whether there were any closely related paralogues. This resulted in removal of 14 *Borrelia* sequences, for a final total of 143 Target Sequences. These are included in Table 3 and the attached spreadsheet.

Table 2 – Summary of Target Selection screens in Batch02

organism_id	code	organism	total genes	total homologs	size filter	cysteine filter	transmembrane	targetdb	same gene in batch01	homolog screen
Totals										
56	56	BobuA	Borrelia burgdorferi B31	3323	42	34	32	28	28	14
14	14	BrabA	Brucella melitensis biovar Abortus 230	3036	56	45	45	40	38	33
13	13	BupsA	Burkholderia pseudomallei 1710b	6349	82	65	59	51	51	46
12	12	RiprA	Rickettsia prowazekii str. Madrid E	838	33	26	24	21	21	20
33	33	BaheA	Bartonella henselae str. Houston-1	1489	37	29	27	23	22	22
32	32	MytuD	Mycobacterium tuberculosis H37Rv	3990	66	42	41	40	8	8
				19025	316	241	228	204	168	143

Table 3 – Batch02 Targets Selected

target_id	cluster	gid	refseq_id	annotation
BaheA.00010.a	10	49474945	YP_032986.1	enoyl-(acyl carrier protein) reductase
BaheA.00010.b	10	49475228	YP_033269.1	enoyl-(acyl carrier protein) reductase
BaheA.00010.c	10	49475329	YP_033370.1	3-ketoacyl-(acyl-carrier-protein) reductase
BaheA.00013.a	13	49476146	YP_034187.1	Holliday junction DNA helicase B
BaheA.00057.a	57	49474848	YP_032889.1	ABC transporter/ATP-binding protein
BaheA.00057.b	57	49475591	YP_033632.1	High-affinity branched-chain amino acid transport, ATP-binding protein
BaheA.00113.a	113	49474946	YP_032987.1	3-oxoacyl-(acyl carrier protein) synthase I
BaheA.00113.b	113	49475331	YP_033372.1	3-oxoacyl-(acyl carrier protein) synthase II
BaheA.00113.c	113	49475515	YP_033556.1	3-oxoacyl-(acyl carrier protein) synthase II
BaheA.00133.a	133	49475313	YP_033354.1	cysteinyI-tRNA synthetase
BaheA.00145.a	145	49475729	YP_033770.1	aspartyl-tRNA synthetase
BaheA.00148.a	148	49476240	YP_034281.1	site-specific tyrosine recombinase XerC
BaheA.00148.b	148	49476286	YP_034327.1	Integrase/recombinase xerD
BaheA.00150.a	150	49475738	YP_033779.1	UDP-N-acetylglucosamine pyrophosphorylase
BaheA.00154.a	154	49475229	YP_033270.1	chorismate synthase
BaheA.00157.a	157	49475728	YP_033769.1	DNA topoisomerase IV subunit A
BaheA.00159.a	159	49475736	YP_033777.1	ATP-dependent DNA helicase recG
BaheA.00162.a	162	49475554	YP_033595.1	DNA topoisomerase IV subunit B
BaheA.00163.a	163	49474906	YP_032947.1	Phenylalanyl-tRNA synthetase alpha chain
BaheA.00166.a	166	49475848	YP_033889.1	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase
BaheA.00168.a	168	49475375	YP_033416.1	hypothetical protein
BaheA.00171.a	171	49475534	YP_033575.1	3-oxoacyl-(acyl carrier protein) synthase III
BobuA.00013.a	13	15594368	NP_212156.1	Holliday junction DNA helicase B
BobuA.00057.a	57	15594811	NP_212600.1	ABC transporter, ATP-binding protein
BobuA.00133.a	133	15594944	NP_212733.1	cysteinyI-tRNA synthetase (cysS)
BobuA.00137.a	137	15595162	NP_212951.1	UDP-N-acetylmuramate-L-alanine ligase
BobuA.00145.a	145	161511079	NP_212580.2	aspartyl-tRNA synthetase
BobuA.00153.a	153	15595055	NP_212844.1	DNA primase (dnaG)
BobuA.00155.a	155	15594689	NP_212478.1	DNA helicase (uvrD)
BobuA.00155.b	155	15594952	NP_212741.1	rep helicase, single-stranded DNA-dependent ATPase (rep)
BobuA.00156.a	156	15595065	NP_212854.1	threonyI-tRNA synthetase (thrZ)
BobuA.00157.a	157	15594381	NP_212169.1	DNA topoisomerase IV subunit A
BobuA.00159.a	159	15594926	NP_212715.1	DNA recombinase (recG)
BobuA.00162.a	162	15594382	NP_212170.1	DNA topoisomerase IV subunit B
BobuA.00162.b	162	15594781	NP_212570.1	DNA gyrase, subunit B (gyrB)
BobuA.00169.a	169	15594457	NP_212245.1	replicative DNA helicase (dnaB)
BrabA.00010.d	10	83268985	YP_418276.1	Short-chain dehydrogenase/reductase SDR:Glucose/ribitol dehydrogenase
BrabA.00010.e	10	82700418	YP_414992.1	7-alpha-hydroxysteroid dehydrogenase
BrabA.00010.f	10	82700795	YP_415369.1	Short-chain dehydrogenase/reductase SDR:Glucose/ribitol dehydrogenase
BrabA.00010.g	10	83268982	YP_419113.1	3-ketoacyl-(acyl-carrier-protein) reductase
BrabA.00057.b	57	82699072	YP_413646.1	ATP/GTP-binding site motif A (P-loop):ABC transporter:AAA ATPase
BrabA.00057.c	57	83268982	YP_418273.1	ATP/GTP-binding site motif A (P-loop):ABC transporter:AAA ATPase
BrabA.00057.d	57	83269203	YP_418494.1	ATP/GTP-binding site motif A (P-loop):ABC transporter:AAA ATPase
BrabA.00057.e	57	83269482	YP_418773.1	ATP/GTP-binding site motif A (P-loop):Ras GTPase superfamily:ABC transporter:AAA ATPase
BrabA.00057.f	57	83269739	YP_419030.1	Chaperonin clpA/B:Glycoside hydrolase, family 18/2:ATP/GTP-binding site motif A (P-loop):ABC transporter:AAA ATPase:TOBE doma...
BrabA.00057.g	57	83269986	YP_419277.1	ATP/GTP-binding site motif A (P-loop):ABC transporter:AAA ATPase
BrabA.00113.b	113	82699372	YP_413946.1	3-oxoacyl-(acyl carrier protein) synthase II
BrabA.00113.c	113	82699726	YP_414300.1	3-oxoacyl-(acyl carrier protein) synthase II
BrabA.00133.a	133	82699565	YP_414139.1	cysteinyI-tRNA synthetase
BrabA.00138.a	138	82699903	YP_414477.1	Alanyl-transfer RNA synthetase
BrabA.00140.a	140	83269033	YP_418324.1	para-aminobenzoate synthase component I
BrabA.00143.a	143	83269364	YP_418655.1	3-ketoacyl-(acyl-carrier-protein) reductase
BrabA.00145.a	145	82699634	YP_414208.1	aspartyl-tRNA synthetase
BrabA.00146.a	146	82698963	YP_413537.1	prephenate dehydratase
BrabA.00148.a	148	82700682	YP_415256.1	site-specific tyrosine recombinase XerC
BrabA.00148.b	148	82700788	YP_415362.1	site-specific tyrosine recombinase XerD
BrabA.00149.a	149	82699362	YP_413936.1	SAM (and some other nucleotide) binding motif:Generic methyltransferase:Cyclopropane-fatty-acyl-phospholipid synthase
BrabA.00150.a	150	83269541	YP_418832.1	Bacterial transferase hexapeptide repeat:ATP/GTP-binding site motif A (P-loop):Nucleotidyl transferase
BrabA.00153.a	153	82700293	YP_414867.1	DNA primase
BrabA.00156.a	156	82699915	YP_414489.1	threonyI-tRNA synthetase
BrabA.00159.a	159	83269543	YP_418834.1	DEAD/DEAH box helicase:RecA bacterial DNA recombination protein:Helicase, C-terminal:ATP/GTP-binding site motif A (P-loop):Ha...
BrabA.00162.a	162	83269533	YP_418824.1	DNA topoisomerase IV subunit B
BrabA.00163.a	163	82700873	YP_415447.1	phenylalanyl-tRNA synthetase subunit alpha
BrabA.00164.a	164	82699747	YP_414321.1	arginyl-tRNA synthetase
BrabA.00167.a	167	82700376	YP_414950.1	Glutamine amidotransferase class-I:Carbamoyl-phosphate synthase, GATase domain:Anthranilate synthase component I and chorisma...
BrabA.00168.a	168	162002874	YP_414535.2	bifunctional 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase/2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase protein
BrabA.00169.a	169	82699361	YP_413935.1	replicative DNA helicase
BrabA.00171.a	171	82699657	YP_414231.1	3-oxoacyl-(acyl carrier protein) synthase III
BrabA.00171.b	171	83269629	YP_418920.1	3-oxoacyl-(acyl carrier protein) synthase II
BupsA.00010.c	10	76818323	YP_337448.1	enoyl-(acyl carrier protein) reductase
BupsA.00010.d	10	76811639	YP_332645.1	3-ketoacyl-(acyl-carrier-protein) reductase
BupsA.00010.e	10	76811242	YP_333724.1	acetylacetyl-CoA reductase
BupsA.00010.f	10	76809090	YP_334857.1	short chain dehydrogenase
BupsA.00010.g	10	76817363	YP_336172.1	acetoacetyl-CoA reductase
BupsA.00010.h	10	76818301	YP_337614.1	3-ketoacyl-(acyl-carrier-protein) reductase
BupsA.00057.b	57	76809899	YP_331616.1	branched-chain amino acid ABC transporter, ATP-binding protein
BupsA.00057.c	57	76809615	YP_331682.1	ABC transport system ATP-binding protein
BupsA.00057.d	57	76810319	YP_332179.1	ABC transporter system, ATP-binding protein
BupsA.00057.e	57	76809665	YP_332557.1	leucine/isoleucine/valine transport system ATP-binding protein
BupsA.00057.f	57	76811024	YP_332558.1	branched-chain amino acid ABC transporter, ATP-binding protein
BupsA.00057.g	57	76817634	YP_337292.1	leucine/isoleucine/valine transporter ATP-binding subunit
BupsA.00113.a	113	76811113	YP_334284.1	3-oxoacyl-(acyl carrier protein) synthase II
BupsA.00137.b	137	76808578	YP_334914.1	UDP-N-acetylmuramate-L-alanine ligase
BupsA.00138.a	138	76811885	YP_332812.1	alanyl-tRNA synthetase
BupsA.00138.b	138	76817280	YP_335495.1	putative alanyl-tRNA synthetase related protein
BupsA.00139.a	139	76808951	YP_334900.1	dephospho-CoA kinase
BupsA.00141.a	141	76811768	YP_334287.1	acyl-carrier-protein S-malonyltransferase
BupsA.00141.b	141	76818342	YP_337007.1	malonyl CoA-acyl carrier protein transacylase
BupsA.00141.c	141	76818102	YP_337765.1	malonyl CoA-acyl carrier protein transacylase
BupsA.00142.a	142	76810814	YP_332162.1	phosphopantetheine adenylyltransferase
BupsA.00144.a	144	76811001	YP_333297.1	integration host factor subunit alpha
BupsA.00144.b	144	76810828	YP_334371.1	integration host factor subunit beta
BupsA.00146.a	146	76809940	YP_334376.1	chorismate mutase/prephenate dehydratase
BupsA.00147.a	147	76811778	YP_334940.1	anthranilate synthase component I
BupsA.00148.a	148	76809899	YP_331806.1	site-specific tyrosine recombinase XerC
BupsA.00148.b	148	162210085	YP_332479.2	site-specific tyrosine recombinase XerD
BupsA.00149.a	149	76809915	YP_332472.1	cyclopropane fatty acid synthase family protein

BupsA.00150.a	150	76810169	YP_331932.1	UDP-N-acetylglucosamine pyrophosphorylase
BupsA.00152.a	152	76810520	YP_334941.1	anthranilate synthase component II
BupsA.00153.a	153	76819088	YP_335992.1	DNA primase
BupsA.00154.a	154	162210097	YP_333268.2	chorismate synthase
BupsA.00156.a	156	76810857	YP_333290.1	threonyl-tRNA synthetase
BupsA.00158.a	158	76818560	YP_337086.1	glutamate racemase
BupsA.00160.a	160	76809332	YP_334416.1	chorismate mutase
BupsA.00161.a	161	162210039	YP_334398.2	adenosine deaminase
BupsA.00162.a	162	76812112	YP_332710.1	DNA topoisomerase IV subunit B
BupsA.00163.a	163	162210099	YP_333295.2	phenylalanyl-tRNA synthetase subunit alpha
BupsA.00164.a	164	76812104	YP_332006.1	arginyl-tRNA synthetase
BupsA.00165.a	165	76810952	YP_332252.1	7-cyano-7-deazaguanine reductase
BupsA.00166.a	166	76810171	YP_334655.1	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase
BupsA.00166.b	166	76809675	YP_334907.1	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase
BupsA.00169.a	169	76809062	YP_333807.1	replicative DNA helicase
BupsA.00171.a	171	76810029	YP_331513.1	3-oxoacyl-(acyl-carrier-protein) synthase III
BupsA.00171.b	171	76808924	YP_331516.1	3-oxoacyl- synthase
BupsA.00171.c	171	76810561	YP_334288.1	3-oxoacyl-(acyl carrier protein) synthase III
MytD.00010.a	10	15607383	NP_214756.1	3-ketoacyl-(acyl-carrier-protein) reductase
MytD.00113.a	113	15610084	NP_217463.1	PROBABLE POLYKETIDE SYNTHASE PKS15
MytD.00145.a	145	15609709	NP_217088.1	aspartyl-tRNA synthetase
MytD.00148.a	148	15608839	NP_216217.1	site-specific tyrosine recombinase XerD
MytD.00148.b	148	15610031	NP_217410.1	site-specific tyrosine recombinase XerC
MytD.00151.a	151	15608801	NP_216179.1	Probable polyketide synthase pks17
MytD.00155.a	155	15610334	NP_217714.1	PROBABLE ATP-DEPENDENT DNA HELICASE II UVRD2
MytD.00163.a	163	15608787	NP_216165.1	phenylalanyl-tRNA synthetase subunit alpha
RiprA.00010.a	10	15603914	NP_220429.1	acetoacetyl-CoA reductase
RiprA.00010.b	10	15604596	NP_221114.1	3-ketoacyl-(acyl-carrier-protein) reductase
RiprA.00057.a	57	15604368	NP_220884.1	ABC TRANSPORTER ATP-BINDING PROTEIN (abcT2)
RiprA.00113.a	113	15604598	NP_221116.1	3-oxoacyl-(acyl carrier protein) synthase II
RiprA.00133.a	133	15603963	NP_220478.1	cysteinyl-tRNA synthetase
RiprA.00137.a	137	15604117	NP_220632.1	UDP-N-acetylmuramate-L-alanine ligase
RiprA.00141.a	141	15604570	NP_221088.1	acyl-carrier-protein S-malonyltransferase
RiprA.00145.a	145	15604021	NP_220536.1	aspartyl-tRNA synthetase
RiprA.00148.a	148	161729088	NP_220744.2	site-specific tyrosine recombinase XerD
RiprA.00148.b	148	15604648	NP_221166.1	site-specific tyrosine recombinase XerC
RiprA.00155.a	155	15604312	NP_220828.1	DNA HELICASE II (uvrD)
RiprA.00157.a	157	15603946	NP_220461.1	DNA topoisomerase IV subunit A
RiprA.00159.a	159	15604444	NP_220962.1	ATP-DEPENDENT DNA HELICASE RECG (recG)
RiprA.00162.a	162	15604098	NP_220613.1	DNA topoisomerase IV subunit B
RiprA.00163.a	163	15604282	NP_220798.1	phenylalanyl-tRNA synthetase subunit alpha
RiprA.00164.a	164	15603944	NP_220459.1	arginyl-tRNA synthetase
RiprA.00165.a	165	15603951	NP_220466.1	7-cyano-7-deazaguanine reductase
RiprA.00166.a	166	15604124	NP_220639.1	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase
RiprA.00169.a	169	15604399	NP_220915.1	replicative DNA helicase
RiprA.00170.a	170	15604232	NP_220748.1	enoyl-(acyl carrier protein) reductase