**In silico identification of metabolic enzyme drug targets in Burkholderia pseudomallei**

Jean F. Challacombe  
Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, 87545

Correspondence: jchalla@lanl.gov

**Abstract**

The intracellular pathogen *Burkholderia pseudomallei*, which is endemic to parts of southeast Asia and northern Australia, causes the disease melioidosis. Although acute infections can be treated with antibiotics, melioidosis is difficult to cure, and some patients develop chronic infections or a recrudescence of the disease months or years after treatment of the initial infection. *B. pseudomallei* strains have a high level of natural resistance to a variety of antibiotics, and with limited options for new antibiotics on the horizon, alternatives are needed.

The aim of the present study was to characterize the metabolic capabilities of *B. pseudomallei*, identify metabolites and aspects of the metabolic network crucial for pathogen survival, understand the metabolic interactions that occur between pathogen and host cells, and determine if any of the metabolic enzymes produced by the pathogen might be potential antibacterial targets.

This aim was accomplished through genome scale metabolic modeling of *B. pseudomallei* under different external conditions, including all nutrients that could be consumed by the model and only the nutrients available in culture media. Using this approach, candidate chokepoints were identified, then knocked out in silico under the different nutrient conditions. The effect of each knockout on the metabolic network was examined. When four of the candidate chokepoints were knocked out in silico, the flux through the *B. pseudomallei* network was decreased, depending on the nutrient conditions. These results demonstrate the utility of genome-scale metabolic modeling methods for in silico studies of pathogen metabolism and for drug target identification in *B. pseudomallei*. 
Introduction

The intracellular pathogen *Burkholderia pseudomallei*, which causes the disease melioidosis, is acquired from the environment in parts of southeast Asia and northern Australia (1, 2). Although acute infections can be treated with antibiotics, melioidosis is difficult to cure, requiring lengthy treatment in two phases for a duration of ~20 weeks (3, 4). Despite antibiotic therapy, some patients have persistent cases that develop into chronic infections, and others experience a recrudescence of the disease months or years after treatment of the initial infection with antibiotics (5). *B. pseudomallei* strains have a high level of natural resistance to a variety of antibiotics (6-8), and with limited options for new antibiotics on the horizon, alternatives are desperately needed.

The availability of many *B. pseudomallei* genomes and advances in computational analysis methods make possible the rapid identification of novel antibacterial targets by selecting the best targets from complete sets of protein coding genes. Previous studies comparing the genomes of bacterial pathogens have demonstrated that essential genes present in pathogen genomes, but not in the host, make the best therapeutic targets (9). Many known antibacterial compounds are enzyme inhibitors (10, 11), so metabolic enzymes specific to bacteria represent promising drug targets (9).

Enzyme targets present in key metabolic pathways have been identified in *B. pseudomallei* and other bacterial pathogens; these pathways include fatty acid biosynthesis (12-14), the glyoxalate shunt (15, 16), the chorismate pathway for biosynthesis of aromatic amino acids (17), purine, histidine, 4-aminobenzoate, and lipoate biosynthesis (18, 19), leucine, threonine, p-aminobenzoic acid, aromatic compound biosynthesis (20), branched chain amino acid biosynthesis (21), purine metabolism (22). Other enzyme targets have been identified that are not in pathways - alanine racemase (interconverts L- and D-alanine) (23), superoxide dismutase (24) and cyclic di-GMP phosphodiesterase (25).

Drugs acting on pathogen targets that are not present in the host should not cause significant side effects. However, before the human genome was available, the process of identifying bacterial pathogen-specific drug targets was labor intensive, involving comparison of candidate pathogen targets against all known eukaryotic sequences to filter out targets likely to occur in the human (9). Since then, various software tools have made the process of *in silico* target identification in pathogen genomes easier. Available *in silico* tools encompass various cheminformatic (26) and bioinformatic (27, 28) approaches to identify new protein targets. Among the bioinformatic tools, metabolic pathway/metabolic network analysis has emerged as an efficient *in silico* method to identify candidate metabolic enzyme targets in pathogen genomes.

Several software packages are available to facilitate genome-scale metabolic network analyses (29). Starting with an annotated pathogen genome, the components of metabolic pathways are identified, curated and refined (30). The resulting genome-scale metabolic model can be used to integrate omics datasets and to perform various analyses to determine the most likely drug targets (10). Perhaps the most important task with respect to finding good candidate targets is metabolic chokepoint identification. By definition, a chokepoint enzyme either consumes a unique substrate or produces a unique product in the pathogen metabolic network (31).
Inhibition of chokepoint enzymes may disrupt crucial metabolic processes in the pathogen, so chokepoints that are essential to the pathogen represent good potential drug targets (32-34).

The aim of the present study was to characterize the metabolic capabilities of *B. pseudomallei*, identify metabolites and aspects of the metabolic network crucial for pathogen survival, understand the metabolic interactions that occur between pathogen and host cells, and determine if any of the metabolic enzymes produced by the pathogen might be potential antibacterial targets.

This aim was accomplished through genome scale metabolic modeling of *B. pseudomallei* under different external conditions, including all nutrients that could be consumed by the model and only the nutrients available in culture media. Using this approach, we identified candidate chokepoints, then knocked out the genes encoding chokepoint enzymes *in silico* under the different nutrient conditions, and examined the effect of each knockout on the metabolic network. These results demonstrate the utility of genome-scale metabolic modeling methods for *in silico* studies of pathogen metabolism and for drug target identification in *B. pseudomallei*.

**Methods**

**Metabolic pathway reconstructions and annotation curation**

Pathway genome databases (PGDBs) for *B. pseudomallei* strains MSHR668 and K96243 were obtained through the Pathway Tools software (version 18.5) from the PGDB registry (35). We found that the original annotation of *B. pseudomallei* K96243 identified many fewer coding sequences than that of MSHR668, so we re-annotated the original complete genome sequences of both strains using the RAST (36) resource to better compare them (37). The RAST-annotated genome sequences were loaded into Pathway Tools, using the PathoLogic component to predict the metabolic pathways (38).

For each genome, the set of protein coding sequences from the original annotation was compared to those from the RAST annotation, using blast to identify coding sequences in common between the two annotations, and to identify coding sequences that were missing from each annotation. For the MSHR668 genome, RAST annotation identified 247 fewer coding sequences than the original annotation. However there were some predicted coding sequences with annotated functions in the RAST annotation that were not present in the original, so these were added to the original PGDB for MSHR668. The RAST annotation of the K96243 genome identified 1,317 more protein coding sequences than the original annotation, so the PGDB that was created from the RAST annotation was used as the starting point. Coding sequences with annotated functions from the original annotation that were not present in the RAST annotation were added to the PGDB that was created from the RAST annotated genome.

**Chokepoint Identification**

Chokepoints were identified in each PGDB using the chokepoint reaction finder in Pathway Tools. All reactions were included except those found in human.
Flux balance analysis

For each metabolic network reconstruction, flux balance analysis (FBA) was performed using the MetaFlux module within the Pathway Tools (39). Development FBA models were constructed iteratively to determine the compounds that each model could use and produce. This was accomplished by trying all compounds in the PGDB as biomass metabolites, nutrients and secretions in the various compartments (cytosol, periplasmic space and extracellular). Each model was refined iteratively, first identifying the specific biomass components that could be produced, then trying all compounds as nutrients and secretions, then specifying the biomass metabolites and nutrients and trying all compounds as secretions. Once the nutrients, secretions and biomass components that could be consumed or produced by the metabolic networks were determined, the log file produced by MetaFlux was examined and problematic reactions were fixed, if possible. Most of the problematic reactions were unbalanced due to missing chemical formulas of one or more metabolites. A few of these reactions were corrected by copying the missing structures from other PGDBs. For example, the structures of D-ribose, D-glucuronate, D-glucose and some other compounds were copied from the more highly curated Escherichia coli K12 substr. MG1655 PGDB. Many of the reactions with compounds that were lacking chemical formulas were generic so no suitable chemical structure could be found or created. For example, reactions involving starch, glycogen, and glucans with variable lengths and non-numeric stoichiometries, could not be balanced. Other reactions were missing H\(^+\) or H\(_2\)O on one side or the other, and the addition of these balanced the reaction. However, there were a small number of reactions that could not be fixed and these were marked as unbalanced. Once all reactions that could be fixed were corrected in the PGDB, MetaFlux was run in development mode again to identify additional biomass metabolites, nutrients, and secretions. Once the set of biomass metabolites was constant, the nutrients that could be consumed by the model were determined, and then the secretions. The result of this process was a final unconstrained FBA model utilizing all nutrients that could be consumed by the model.

To mimic the nutrient conditions in culture, only the estimated set of ingredients present in LB medium were included as nutrients in the FBA model. LB medium includes as its main ingredients tryptone (40) and yeast extract (41). Tryptone provides peptides and peptones, which are good sources of amino acids (http://www.sigmaaldrich.com/analytical-chromatography/microbiology/basic-ingredients/protein-sources.html); yeast extract provides vitamins, nitrogenous compounds, carbon, sulfur, trace elements and minerals (42). The LB media composition reported in a Bacillus subtilis modeling study (43) was used as a starting point for this study. Starting with the nutrient set that mimicked LB media (43), all compounds were included in the try-biomass set for the cytosol and periplasmic cellular compartments to see which biomass metabolites could be produced given only the nutrients present in LB media. Once a stable set of nutrients was determined, iterations were performed to determine the stable set of biomass metabolites that could be produced. Then all compounds were tried as secretions in the cytosol and periplasmic compartments to determine the compounds that could be secreted by the model.

A list of nutrients that might be available to B. pseudomallei while residing inside host cells during infection was compiled by searching the literature for infection studies involving B. pseudomallei and host cells. Gene expression studies of other intracellular pathogens and host
cells during infection were also considered. A development FBA model mimicking infection conditions was constructed as described above for the LB media model.

**Essential gene and candidate drug target identification**

Essential genes in the MSHR668 genome were determined by blasting all protein coding sequences (amino acid format) against the Database of Essential Genes (44) using the blastp program with an E-value cut-off of $1 \times 10^{-10}$ and 70% identity as thresholds.

Candidate drug target sequences were determined by blasting all protein coding sequences in the MSHR668 genome (as both nucleotide and amino acid formats) against the complete set of DrugBank targets (45), using an E-value cutoff of 0.005 and a threshold identity of 70% to select likely targets.

**In silico knockout experiments**

To reduce the set of potential drug targets, candidate chokepoint enzymes were compared against the list of essential genes and all of the drug targets in DrugBank to identify chokepoint sequences representing essential genes and DrugBank targets. Knockout experiments were performed *in silico* through the Pathway Tools MetaFlux module (39); the chokepoint genes were knocked out one at a time and the effects on total biomass flux through the metabolic network were noted.

**Network visualization**

For each PGDB, an .sbml file was exported from the Pathway Tools and loaded into Cytoscape (46) version 3.1.1 for visualization and comparison of network features.

**Results**

**General features of genomes and metabolic networks**

The general characteristics of each PGDB and metabolic network for *B. pseudomallei* MSHR668 and K96243 are listed in Table 1.

| Table 1. Features of *B. pseudomallei* PGDBs and metabolic networks |
|-----------------------------------------------|-----------------|-----------------|
| **Genome (Pathway Tools PGDB)**               | curated MSHR668 (original + RAST) | curated K96243 (RAST + original) |
| Coding sequences                              | 7133            | 7045            |
| Pathways                                      | 339             | 387             |
| Enzymatic reactions                           | 1870            | 1995            |
| Transport reactions                            | 144             | 82              |
| Enzymes                                       | 1666            | 1685            |
| Transporters                                  | 292             | 38              |
| Compounds                                     | 1397            | 1578            |
| **Metabolic Network**                         |                 |                 |
| Nodes (metabolites)                           | 4295            | 4219            |
| Edges (reaction steps)                        | 10860           | 9747            |
| **Chokepoints**                               |                 |                 |
| Producing                                     | 444             | 506             |
| Consuming                                     | 419             | 498             |
Candidate  479  348
Dead-end reactants  166  217
Dead-end products  72  118

The *B. pseudomallei* MSHR668 and K96243 genomes contained similar numbers of coding sequences, pathways, enzymatic reactions and enzymes. Differences between the PGDBs were noted in the numbers of transporters, transport reactions and compounds. The K96243 PGDB contained fewer transporters and transport reactions and more compounds than MSHR668. In terms of the metabolic network characteristics, both networks contained similar numbers of nodes (representing metabolites), while the MSHR668 network had more edges (reaction steps) than K96243. This is likely due to the more extensive curation of the MSHR668 network that was performed during refinement of the metabolic network models (see Methods).

In a metabolic network, chokepoint reactions either uniquely produce or uniquely consume a metabolite. Inhibiting an enzyme that consumes a unique substrate may cause that metabolite to accumulate, and it may be toxic to the cell; inhibiting an enzyme that produces a unique product may starve the cell of an essential metabolite (31). Identifying chokepoint enzymes in pathogens is a promising *in silico* approach to recognize potential metabolic drug targets. For example, analysis of *Plasmodium falciparum* metabolism revealed that 87.5% of proposed drug targets supported by evidence are chokepoint reactions. However, to be a valid chokepoint, the metabolite in question must be balanced by a producing or consuming reaction and not be a dead-end metabolite (31). Table 1 compares the numbers of chokepoint reactions and dead-end metabolites that were identified in each PGDB. Overall the numbers were similar between the two PGDBs, and the lower numbers of dead-ends in the MSHR668 database were likely due to the more extensive curation performed on the MSHR668 PGDB (see Methods).

**Flux balance analysis**

**Unconstrained FBA model**
Given a set of nutrients for consumption, along with secretions and metabolites that can be produced, a FBA model predicts the steady-state flux rates of the metabolic reactions in an organism, and provides an estimate of the overall biomass flux. FBA was conducted on each metabolic network as described in the Methods section. Following several cycles of refinement, an initial unconstrained MSHR668 model was solved. The K96243 development model did not reach a stable solution, likely because there were missing transporter-encoding genes (and possibly other genes) in the PGDB, indicating that the annotation needed more curation. Since the unconstrained MSHR668 model reached a stable solution after performing the initial network curation steps suggested by MetaFlux, only this model was analyzed further.

The FBA model of the unconstrained network in MSHR668, included all possible biomass compounds that could be produced, all nutrients that could be consumed, had no weights imposed on the biomass metabolites and no constraints imposed on the nutrients. [Supplemental material, Final_unconstrained_model_inputs.pdf and Final_unconstrained_model_solution.pdf].
LB media FBA model
To mimic the conditions that *B. pseudomallei* experiences in culture, only the nutrients present in LB media plus glycerol were included in the FBA model (43). Constraints were included on some of the nutrients (ADP, Pi, proton and glycerol). [Supplemental material, Final_LB_model_inputs.pdf and Final_LB_model_solution.pdf].

Table 2. Characteristics of the MSHR668 unconstrained and LB media models

<table>
<thead>
<tr>
<th>Model</th>
<th>total # rxns</th>
<th># rxns carrying flux</th>
<th>biomass metabolites produced</th>
<th>nutrients consumed</th>
<th>secretions</th>
<th>biomass flux</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconstrained</td>
<td>3193</td>
<td>1619</td>
<td>1403</td>
<td>667</td>
<td>10</td>
<td>15000.00</td>
</tr>
<tr>
<td>LB media</td>
<td>3213</td>
<td>1025</td>
<td>282</td>
<td>47</td>
<td>0</td>
<td>0.079412</td>
</tr>
</tbody>
</table>

In addition to the unconstrained and LB media models, the original plan for this study included the development of a model of *B. pseudomallei* that mimics infection conditions. However, there was very little information available on the growth requirements of *B. pseudomallei* inside human macrophages. Most studies of the nutritional requirements of intracellular pathogens growing inside host cells have been performed on *Legionella pneumophila* (47, 48), which can grow and replicate similarly in human macrophages and amoebae (49). Growing in both human macrophages and amoebae, *L. pneumophila* utilizes amino acids as its main sources of carbon, nitrogen and energy; *L. pneumophila* obtains amino acids from the host through proteasomal degradation (48). However, glucose is also used to feed central metabolism under both culture and infection conditions (50).

Comparing the nutrients provided by the LB media model to the nutrients used by *L. pneumophila* in culture and during infection of amoebae (47), the only difference was the carbohydrate carbon source: glycerol (*B. pseudomallei* LB media (51)) vs. glucose (*L. pneumophila* AYE media (50) and in amoebae (47, 50))

When glycerol was replaced by glucose in the LB media model of *B. pseudomallei*, no FBA solution was found [data not shown]. Several possible explanations for this result are presented in the Discussion.

Host cell infection model
To construct a model representing infection conditions, a list of nutrients present in the host cell cytosol is a required input. However, no comprehensive studies have been performed to identify the complete list of host cell nutrients that are available to *B. pseudomallei* during infection. Also, the specific carbon requirements of *B. pseudomallei* in either human macrophages or amoebae have not been determined. However, a recent study produced whole-genome tiling array expression data to assess *B. pseudomallei* transcriptional responses under 82 different conditions, including infection (52). From their supplemental table S2, a list of metabolic genes expressed in the infection conditions was used to infer the potential nutrients consumed by *B. pseudomallei* during infection. Additional candidate host cell nutrients were identified from the literature, focusing on studies of intracellular pathogen-mammalian host infections. The nutrients identified as potential carbon sources for intracellular survival of various pathogens included
aromatic compounds, such as benzoate and phenylacetic acid and related derivatives (53), sugar acids like gluconate, galactonate, glucuronate, and galacturonate (54), ribo- and deoxyribonucleosides, hexuronates (55), glutathione (56), glucose 6-phosphate (57, 58), glycerol-3-phosphate (59). The complete list of potential host cell nutrients is in the [Supplemental material, Nutrients_infection_model.pdf] file.

When a flux balance analysis was performed using the list of candidate host cell nutrients in a developmental model and trying to identify the biomass metabolites that could be produced, some of the nutrients were not present in any reactions in the MSHR668 PGDB. When the rest of the compounds were included as nutrients in addition to glycerol, none of them were consumed by the model, but glycerol was consumed and biomass was produced. When glycerol was excluded from the nutrient list, none of the other nutrients were consumed and the model was not solvable [data not shown].

**Metabolic chokepoints and candidate drug targets**

To narrow down the list of metabolic chokepoints, which represented candidate drug targets, essential genes and genes with sequence homology to existing DrugBank targets were identified in the MSHR668 genome. This analysis identified 34 chokepoint genes that were also essential genes and DrugBank targets (Table 3).

**Table 3. Chokepoint genes that encode candidate metabolic enzyme drug targets in B. pseudomallei MSHR668**

<table>
<thead>
<tr>
<th>locus_tag</th>
<th>Gene symbol</th>
<th>Enzyme name</th>
<th>E.C. number</th>
<th>DrugBank target (drug IDs)</th>
<th>Target in human?</th>
<th>Chokepoint in K96243?</th>
<th>Bp Mutant(s)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>BURPS668_0305</td>
<td>argG⁺</td>
<td>argininosuccinate synthase</td>
<td>6.3.4.5</td>
<td>P0A6E4 Argininosuccinate synthase (DB00536; DB04077)</td>
<td>no</td>
<td>yes</td>
<td>no*</td>
</tr>
<tr>
<td>BURPS668_0328</td>
<td>folB⁺</td>
<td>dihydroweopterin aldolase</td>
<td>4.1.2.25</td>
<td>P56740 Dihydroweopterin aldolase (DB01778; DB01906; DB02119; DB02489; DB03231; DB03571; DB04168; DB04400; DB04425; DB06906)</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_0567</td>
<td>pth⁺</td>
<td>peptidyl-tRNA hydrolase</td>
<td>3.1.1.29</td>
<td>No</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_0675</td>
<td>aspS⁺</td>
<td>aspartyl-tRNA synthase</td>
<td>6.1.1.12</td>
<td>P36419 Aspartate- tRNA ligase (DB01895)</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_0810</td>
<td>RecA⁺</td>
<td>RecA protein</td>
<td>5.99.1.-</td>
<td>P62219 Protein RecA (DB01660; DB03222; DB04444; DB06906)</td>
<td>no</td>
<td>yes</td>
<td>no*</td>
</tr>
<tr>
<td>BURPS668_0964</td>
<td>dut&lt;sup&gt;+&lt;/sup&gt;</td>
<td>deoxyuridine 5'-triphosphate nucleotidohydrolase</td>
<td>3.6.1.19/3.6.123</td>
<td>DB04395</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_0967</td>
<td>ileS&lt;sup&gt;+&lt;/sup&gt;</td>
<td>isoleucyl-tRNA synthetase</td>
<td>6.1.1.5</td>
<td>Q09SE4</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_1446</td>
<td>pckG&lt;sup&gt;+&lt;/sup&gt;</td>
<td>phosphoenolpyruvate carboxykinase</td>
<td>4.1.1.32</td>
<td>P35558</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_1465</td>
<td>dnaQ&lt;sup&gt;+&lt;/sup&gt;</td>
<td>DNA polymerase III subunit epsilon</td>
<td>2.7.7.7</td>
<td>P03007</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_1544</td>
<td>NA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>alpha-glucosidase</td>
<td>3.2.1.20</td>
<td>O33830</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_1669</td>
<td>valS&lt;sup&gt;+&lt;/sup&gt;</td>
<td>valyl-tRNA synthetase</td>
<td>6.1.1.9</td>
<td>P26640</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_1712</td>
<td>thrS&lt;sup&gt;+&lt;/sup&gt;</td>
<td>threonyl-tRNA synthetase</td>
<td>6.1.1.3</td>
<td>P0A8M5</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_1750</td>
<td>sucA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>2-oxoglutarate dehydrogenase E1</td>
<td>1.2.4.2</td>
<td>Q02218</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_1752</td>
<td>lpdA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>dihydrolipoamide dehydrogenase</td>
<td>1.8.1.4</td>
<td>P14218</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_2178</td>
<td>purA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>adenylosuccinate synthetase</td>
<td>6.3.4.4</td>
<td>Q83P33</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_2189</td>
<td>hisS&lt;sup&gt;+&lt;/sup&gt;</td>
<td>histidyl-tRNA synthetase</td>
<td>6.1.1.21</td>
<td>P60906</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_2308</td>
<td>gpi&lt;sup&gt;+&lt;/sup&gt;</td>
<td>glucose-6-phosphate isomerase</td>
<td>5.3.1.9</td>
<td>P06744</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Gene Symbol</td>
<td>Gene Product</td>
<td>EC Number</td>
<td>Regulator</td>
<td>Regulation</td>
<td>Genes Known to Be Regulated</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>BURPS668</td>
<td>lpxD</td>
<td>UDP-3-O-{3-hydroxymyristoyl] glucosamine N-acyltransferase</td>
<td>2.3.1.191</td>
<td>no</td>
<td>yes</td>
<td>DB02093; DB02548; DB03042; DB03581; DB03937; DB04493; O67648 UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase (DB01991; DB04257; DB04399; DB07355; DB07536; DB08231; DB07861) P60472 Ditrans, polycis- undecaprenyl-diphosphate synthase ((2E,6E)-farnesyl-diphosphate specific) (DB04695; DB04714; DB07404; DB07409; DB07410; DB07426; DB07780) P10802 Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex (DB01846; DB01992; DB08120) Q02054 Actinorhodin polyketide synthase acyl carrier protein (DB08585; DB08586) Q9HU22 Glucose-1-phosphate thymidylyltransferase (DB01643; DB02452; DB02843; DB03723; DB03751; DB04272; DB04485;</td>
<td></td>
</tr>
<tr>
<td>BURPS668</td>
<td>uppS</td>
<td>undecaprenyl diphosphate synthase</td>
<td>2.5.1.31</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>BURPS668</td>
<td>lpdA</td>
<td>pyruvate dehydrogenase complex E3 component, dihydrolipoamide dehydrogenase</td>
<td>1.8.1.4</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>BURPS668</td>
<td>fabF</td>
<td>3-oxoacyl-ACP synthase</td>
<td>2.3.1.41</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>BURPS668</td>
<td>rfbA</td>
<td>glucose-1-phosphate thymidylyltransferase</td>
<td>2.7.7.24</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Strain</td>
<td>Gene</td>
<td>Description</td>
<td>GO ID</td>
<td>PubMed ID</td>
<td>Expression</td>
<td>Localization</td>
<td>Function</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>--------------------------------------------------</td>
<td>-------</td>
<td>------------</td>
<td>------------</td>
<td>--------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>BURPS668</td>
<td>tgt</td>
<td>Queuine tRNA-ribosyltransferase</td>
<td>2.4.2.29</td>
<td>P28720</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>BURPS668</td>
<td>ruvB</td>
<td>Holliday junction DNA helicase RuvB</td>
<td>3.1.22.4</td>
<td>Q5SL87</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>BURPS668</td>
<td>aroQ</td>
<td>3-dehydroquinate dehydratase</td>
<td>4.2.1.10</td>
<td>P15474</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>BURPS668</td>
<td>murG</td>
<td>Undecaprenylidiphosphomuramoylpentapeptide beta-N-acetylglucosaminyltransferase</td>
<td>2.4.1.227</td>
<td>P17443</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>BURPS668</td>
<td>ung</td>
<td>Uracil-DNA glycosylase</td>
<td>3.2.2.27</td>
<td>Q8X444</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>BURPS668</td>
<td>murA</td>
<td>UDP-N-acetylglucosamine 1-carboxyvinyltransferase</td>
<td></td>
<td>P33038</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Accession</td>
<td>Gene</td>
<td>Enzyme Name</td>
<td>EC Number</td>
<td>DrugBank ID</td>
<td>Inhibition</td>
<td>Essential</td>
<td>Occurs in other bacteria</td>
</tr>
<tr>
<td>-------------</td>
<td>--------</td>
<td>------------------------------------</td>
<td>-----------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>BURPS668_A0384</td>
<td>acnD</td>
<td>aconitate hydratase</td>
<td>4.2.1.3</td>
<td>Q99798</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BURPS668_A1869</td>
<td>ileS</td>
<td>isoleucyl-tRNA synthetase</td>
<td>6.1.1.5</td>
<td>P41972</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_A2053</td>
<td>NA</td>
<td>putative acetyl-CoA carboxylase, biotin carboxylase</td>
<td>6.3.4.14</td>
<td>P24182</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_A2451</td>
<td>leuB</td>
<td>3-isopropylmalate dehydrogenase</td>
<td>1.1.1.85</td>
<td>Q56268</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_A2546</td>
<td>polA</td>
<td>DNA polymerase I</td>
<td>2.7.7.7</td>
<td>P00582</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>BURPS668_A3190</td>
<td>lpdA</td>
<td>dihydrolipoamide dehydrogenase</td>
<td>1.8.1.4</td>
<td>P09063</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

* mutants exist in other *Burkholderia* species, +essential gene in *B. pseudomallei* K96243 and MSHR668

Eight of the targets in Table 3 were also DrugBank targets in human, so the remaining twenty-six targets were searched in DrugBank, which revealed that they also occur in other bacteria.

*In silico* knockout experiments were performed with the MetaFlux module of Pathway Tools to test the effect of inhibiting each chokepoint enzyme on *B. pseudomallei* growth under different conditions. These experiments were performed in both the unconstrained and LB media models. Results (Table 4) show that knockout of BURPS668_3328 (*tgt*) and BURPS668_A2451 (*leuB*), eliminated the biomass flux in the unconstrained model, while knockout of BURPS668_2426 (*lpxD*), BURPS668_2433 (*upps*), and BURPS668_3525 (*murG*) eliminated the biomass flux in the LB media model. Knockout of BURPS668_2433 (*upps*), and BURPS668_3525 (*murG*) decreased the biomass flux in the unconstrained model, but did not eliminate it.
Table 4. Results of *in silico* knockout experiments on twenty-six chokepoint reactions in *B. pseudomallei*

<table>
<thead>
<tr>
<th>locus_tag</th>
<th>Gene name</th>
<th>Reaction(s)</th>
<th>Biomass flux (unconstrained model)</th>
<th>Biomass flux (LB media model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BURPS668_0305</td>
<td>argG</td>
<td>ARGSUCCINSYN-RXN</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_0328</td>
<td>folB</td>
<td>H2NEOPTERINALDOL-RXN</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_0675</td>
<td>aspS</td>
<td>ASPARTATE--TRNA-LIGASE-RXN-ASP-tRNAs/L-ASPARTATE/ATP/PROTON//Charged-ASP-tRNAs/AMP/PPI.60</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_0810</td>
<td>recA</td>
<td>RXN0-5100</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_0964</td>
<td>dut</td>
<td>DUTP-PYROP-RXN</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_1465</td>
<td>dnaQ</td>
<td>DNA-DIRECTED-DNA-POLYMERASE-RXN</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_1712</td>
<td>thrS</td>
<td>THREONINE--TRNA-LIGASE-RXN-THR-tRNAs/THR/ATP/PROTON//Charged-THR-tRNAs/AMP/PPI.52</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_1752</td>
<td>lpdA</td>
<td>RXN-9718</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_2178</td>
<td>purA</td>
<td>ADENYLOSUCCINATE-SYNTASE-RXN</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_2189</td>
<td>hisS</td>
<td>HISTIDINE--TRNA-LIGASE-RXN-HIS-tRNAs/HIS/ATP/PROTON//Charged-HIS-tRNAs/AMP/PPI.52</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_2426</td>
<td>lpxD</td>
<td>UDPHYDROXYMYRGLUCOSAMNECTYLTRANS-RXN-R-3-hydroxymyristoyl-ACPs/UDP-OHMYRGLUCOSAMINE/OHMYRISTOYL/ACP/PROTON.73</td>
<td>15000.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>BURPS668_2433</td>
<td>uppS</td>
<td>RXN-8999</td>
<td>14545.454545</td>
<td>0.000000</td>
</tr>
<tr>
<td>BURPS668_2610</td>
<td>lpdA</td>
<td>1.8.1.4-RXN</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_2788</td>
<td>fabF</td>
<td>3-OXOACYL-ACP-SYNTH-RXN / 3-OXOACYL-ACP-SYNTH-BASE-RXN / 2.3.1.41-RXN / many other reactions</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_3103</td>
<td>rfbA</td>
<td>DTDPGLUCOSEPP-RXN</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_3328</td>
<td>tgt</td>
<td>QUEOSINE-TRNA-RIBOSYLTRANSFERASE-RXN; RXN0-1321</td>
<td>0.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_3366</td>
<td>ruvB</td>
<td>3.1.22.4-RXN</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_3464</td>
<td>aroQ</td>
<td>3-DEHYDROQUINATE-DEHYDRATASE-RXN</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_3525</td>
<td>murG</td>
<td>RXN-11346, RXN-8976, NACGLTRANS-RXN, RXN-11029</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_3561</td>
<td>ung</td>
<td>RXN0-2584</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
<tr>
<td>BURPS668_3668</td>
<td>murA</td>
<td>UDPNACETYGLUCOSAMENOLPYRTRANS-RXN</td>
<td>15000.000000</td>
<td>0.079412</td>
</tr>
</tbody>
</table>
The overall biomass fluxes were different in the two models. The unconstrained model had a much greater total biomass flux than the LB media model.

**Discussion**

Inside host cells, pathogenic bacteria must acquire nutrients to survive, grow and reproduce. Intracellular bacteria are heterotrophic, so they require various sources of carbon, such as carbohydrates, lipids, glycolipids, dicarboxylic acids and amino acids, from their host environment (59, 60). Preferred carbon sources vary among intracellular pathogens, and the types of nutrients available in the host cell cytosol may determine the cell-type specificities of different intracellular pathogens (61). Obviously and unsurprisingly, without adequate nutrition, intracellular pathogens cannot cause disease (62).

It has been suggested that bacterial pathogens survive in the host cell cytosol because they have adapted metabolically to maximize their ability to utilize the available nutrients and grow in that niche (59, 63-65). For example, many bacteria prefer hexoses, like glucose, for carbon and energy sources. These sugars are catabolized through glycolysis, the pentose phosphate and Entner-Doudoroff pathways (60). Some bacteria with different lifestyles lack the glycolysis pathway and preferentially metabolize glucose via the Entner-Doudoroff pathway (66), and some intracellular bacteria lack both glycolysis and Entner-Doudoroff pathways and live on pyruvate that they obtain from the host cell cytosol (67). This diversity in the metabolic capabilities of intracellular pathogens is a fascinating topic, especially because metabolic pathways in the host and pathogen are important to virulence and disease processes. In order to colonize the host, proliferate, establish an infection, and spread to new hosts, pathogens employ various virulence and survival strategies, often involving links between metabolic pathways and virulence genes. Although current knowledge regarding the links between metabolism and virulence is limited (65), this topic is becoming an increasing focus for host-pathogen studies. Some of the general links between metabolism and virulence include regulatory connections between specific metabolites and virulence gene expression (68-72), metabolic requirements for adaptation of the pathogen to the host niche (59, 63-65), and carbon catabolite repression (73). More detailed information on the topic of metabolism and virulence can be found in these recent reviews (59, 62-65, 74-76).

Regardless of the specific metabolic capabilities possessed by a pathogen, essential nutrient acquisition and utilization mechanisms are proving to be good potential therapeutic targets, as
inhibition of these targets would deprive the pathogen of needed substrates for growth and replication inside host cells.

We previously reported that *B. pseudomallei* MSHR668, K96243 and 1106a have abundant capabilities to metabolize hexoses, including the complete sets of genes encoding the glycolysis, pentose phosphate cycle, and Entner-Doudoroff pathways. They also have several pathways for metabolism of pyruvate to acetyl-CoA, acetate and ethanol (37). In general, *B. pseudomallei* as a species has a very diverse set of metabolic capabilities, likely a reflection of its ability to live both in the natural environment and in hosts.

The metabolic power of *B. pseudomallei* has been targeted in mutation studies. Mutation of various metabolic genes results in attenuation of *B. pseudomallei* virulence; these genes include phosphoribosylformylglycinamidine cyclo-ligase (*purM*) (77), aspartate-β-semialdehyde dehydrogenase (*asd*) (78), acetolactate synthase (*ilvI*) (21), dehydroquinate synthase (*aroB*) (20), chorismate synthase (*aroC*) (17), phosphoserine aminotransferase (*serC*) (79), phosphoribosylglycinamide formyltransferase 1 (*purN*) and phosphoribosylformylglycinamidase cyclo-ligase (*purM*) (18), two phospholipase C enzymes (80), disulfide oxidoreductase (*dsbA*) (81).

Targeted mutation of *purM*, which encodes aminimidazole ribotide, a precursor of de novo adenine and thiamine biosynthesis, predictably causes a deficiency in adenine and thiamine biosynthesis (77). Aspartate-β-semialdehyde dehydrogenase (*asd*) mutants cannot synthesize diaminopimelate for cell wall biosynthesis (78), while acetolactate synthase (*ilvI*) mutants cannot synthesize the branched chain amino acids isoleucine, valine, and leucine (21). Dehydroquinate synthase (*aroB*) mutants are defective in the shikimate pathway for chorismate biosynthesis, and addition of the aromatic compounds tyrosine, tryptophan, phenylalanine, PABA, and 2,3-dihydroxybenzoate is required to restore the growth in minimal medium (20). Chorismate synthase (*aroC*) mutants are also defective in aromatic compound synthesis and cannot grow without the addition of aromatic compounds to the media (17). Phosphoserine aminotransferase (*serC*) mutants are defective in serine and pyridoxal 5-phosphate biosynthesis and require minimal medium supplemented with serine for normal growth (79).

Using live attenuated bacteria as vaccines can be effective in preventing disease, as attenuated bacteria may replicate in the host and contain immune-stimulatory epitopes that are not found in subunit or heat-inactivated vaccines (82). Many of the *B. pseudomallei* mutants described above have already been tried as attenuated vaccines with mixed results. Vaccination with the attenuated *asd* mutant protected BALB/c mice against acute melioidosis, but did not protect against chronic melioidosis (78). Vaccination with the attenuated *ilvI* mutant of *B. pseudomallei* protected BALB/c mice against a challenge with a virulent strain (21). In mice vaccinated with the *aroB* mutant, the time to death following challenge with the virulent K96243 strain was a bit longer than in unvaccinated mice, but all mice eventually died (20). The *aroC* mutant was unable to persist in vaccinated BALB/c mice long enough to elicit protective immunity, however C57BL/6 mice were protected against challenge with a virulent strain (17). Intraperitoneal vaccination of BALB/c mice with a *serC* mutant resulted in higher levels of survival after challenge with K96243 virulent strain (79).
While immunization of mice with attenuated *B. pseudomallei* mutants has resulted in the induction of protective immunity in some cases, sterile immunity was rarely reported (reviewed by (83)). Also, the live attenuated vaccine model may not be the best solution for the prevention of melioidosis, because an attenuated mutant might revert to virulence, and might also establish a latent infection (83).

An alternative avenue to combat melioidosis is through antimicrobial therapeutics. With genomic sequencing technologies becoming more affordable, to date there are 393 *B. pseudomallei* genomes available at the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/genome/476). While some of these entries represent re-annotations of previous submissions, and there are some genomes representing colony morphology variants of the same strain, an impressive number of individual genomes are available to use with computational approaches to identify new therapeutic targets.

*In silico* methods for the identification of therapeutic targets in bacterial pathogens include comparative genomics-based approaches, such as identification of essential genes specific to the pathogen, and techniques based on metabolic pathway analysis and metabolic network modeling. The more robust approaches use a combination of comparative genomics methods and metabolic pathway analysis. Two studies computationally identified essential gene targets in genomes of the bacterial pathogens *Mycoplasma genitalium* (84) and *Mycobacterium ulcerans* (85). Another method, subtractive target identification, involves identification of enzymes in the metabolic pathways of the pathogen, and comparing them to human proteins to identify pathogen enzymes that are not found in human. A list of likely targets is compiled by focusing on enzymes in pathways that are usually essential for pathogen growth and survival, like lipid metabolism, carbohydrate metabolism, amino acid metabolism, energy metabolism, vitamin and cofactor biosynthetic pathways and nucleotide metabolism. This approach has been used to identify putative targets in *M. tuberculosis* (86), MRSA (87-89) and a collection of other bacterial pathogens (90).

Methods for therapeutic target discovery based on metabolic pathway analysis and metabolic network modeling have become very popular in the last ten years. Numerous studies have identified candidate drug targets in various bacterial (27, 34, 91-102), fungal (102) and protoscoan (31, 102-105) pathogens using a variety of methods to analyze metabolic pathways and networks. One method employs chokepoint analysis to identify metabolic enzymes that are critical to the pathogen, because they uniquely consume and/or produce certain metabolites. Chokepoint analysis has been used to identify candidate metabolic enzyme targets in various pathogen (27, 31, 34, 95, 103, 104, 106, 107). However, no studies to date have identified potential drug targets using a chokepoint analysis of *B. pseudomallei* metabolic networks.

For this study, a chokepoint analysis was performed on the curated *B. pseudomallei* MSHR668 metabolic network using the Pathway Tools software (35). Table 3 lists the chokepoint enzymes identified in the metabolic networks of *B. pseudomallei* MSHR668 and K96243. Twenty-four of the *B. pseudomallei* chokepoints were not indicated as human targets in DrugBank, and therefore represented good candidate therapeutic targets against melioidosis. Six of the chokepoints in Table 3 were aminoacyl-tRNA synthetases, which are likely good targets as they are critical enzymes involved in protein translation. These chokepoints included aspartyl-, threonyl-,...
have been identified by high throughput screening of the antibiotic ramoplanin in carboxyvinyltransferase (compounds that inhibit Upp) required for bacterial survival, and is therefore essential for bacterial growth. Holiday junction processing components were previously identified as targets for antimicrobials in E. coli and Neisseria gonorrhoeae. The chokepoint enzyme uracil-DNA glycosylase (ung) has a role in uracil excision repair, and is a candidate anti-malarial drug target (127), as well as a potential target to control growth of GC-rich bacteria such as Pseudomonas aeruginosa and Mycobacterium smegmatis (128). DNA polymerase I was identified as a chokepoint in B. pseudomallei. Putative inhibitors of DNA polymerase I (polA), and subsequently DNA synthesis, have been explored as possible antimicrobials (129, 130).

Some chokepoint enzymes in B. pseudomallei have functions in the biosynthesis of cell wall components. One of these chokepoints, UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acetyltransferase (pxd), catalyzes the third step in the lipid A biosynthesis pathway (131). The lipid A component of bacterial LPS is of particular interest because it is essential for cell viability and is highly conserved (132). This pathway is a target for new antibacterial therapeutics in Escherichia coli (133). Another chokepoint involved in cell wall synthesis is undecaprenyl diphosphate synthase (upps), which catalyzes the synthesis of a polyisoprenoid essential for both peptidoglycan and cell wall teichoic acid synthesis. Upps is a critical enzyme required for bacterial survival, and is an antibacterial target in Staphylococcus aureus (134), Bacteroides fragilis, Vibrio vulnificus, E. coli (135) and H. pylori (136). Several classes of compounds that inhibit Upps function have been discovered (134, 137, 138). Two additional B. pseudomallei chokepoints, undecaprenyl diphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase (murG), and UDP-N-acetylglucosamine 1-carboxyvinyltransferase (murA) are involved in peptidoglycan biosynthesis. MurG is the target of the antibiotic ramosplanin in Staphylococcus aureus (139). Other potential inhibitors of MurG have been identified by high throughput screening (140). A small molecule inhibitor of MurG
that augments the activity of β-lactams against methicillin-resistant *Staphylococcus aureus* was recently identified (141). MurA has been a popular target for the design of novel antibiotics, and several inhibitors of MurA have been identified that are active against various bacterial species (87, 142-147). The chokepoint enzyme glucose-1-phosphate thymidylate transferase (*rfba/rmlA*), involved in O antigen biosynthesis, is also a target in *Streptococcus pneumoniae* (148) and *Pseudomonas aeruginosa* (149).

The rest of the chokepoint enzymes in Table 3 were components of various biosynthesis pathways. These enzymes included argininosuccinate synthase (*argG*), which catalyzes the second to last step in L-arginine biosynthesis, and is associated with pathogenesis in the parasite *Leishmania donovani* (150), *Streptococcus pneumoniae* (151), and *B. cenocepacia* (152). The chokepoint enzyme 3-isopropylmalate dehydrogenase (*leuB*) is the third enzyme specific to leucine biosynthesis in microorganisms (153), and has been investigated as an antibacterial target in *M. tuberculosis* (154). The *B. pseudomallei* chokepoint dihydronicotinamide aldolase (*folB*) is part of the tetrahydrofolate biosynthesis process, and is essential for growth and biomass production in *Acinetobacter baylyi*, *Bacillus anthracis*, *Francisella tularensis*, *F. tularensis* subsp. novicida strain U112, *Mycobacterium tuberculosis*, *Helicobacter pylori*, *Pseudomonas aeruginosa*, *Salmonella enterica* serovar Typhi and *Yersinia pestis* (91). Three of the chokepoint genes identified in *B. pseudomallei* encoded lipoamide dehydrogenase, a component of the pyruvate dehydrogenase complex, which converts pyruvate to acetyl-CoA as part of central metabolism (155). Lipoamide dehydrogenase is also a target in *M. tuberculosis*, where deletion drastically impaired the pathogen’s ability to establish infection in the mouse (156). The mycobacterial version has only 36% identity with the human homolog. Lipoamide dehydrogenase is a target of drugs against trypanosomal infections (157). Adenylosuccinate synthetase (*purA*) is also a chokepoint enzyme and potential therapeutic target that is involved in purine salvage in *Leishmania donovani* (158). Chokepoint enzyme 3-dehydroquinate dehydratase (*aroQ*) is a component of the shikimate pathway for chorismate biosynthesis and is a target of known inhibitors in *M. tuberculosis*, *Enterococcus faecalis* and *Streptomyces coelicolor* (159-163).

The *B. pseudomallei* chokepoint enzyme 3-oxoacyl-ACP synthase (*fabF*), involved in fatty acid synthesis, is already an antibacterial target in *E. coli*, and a specific inhibitor, cerulenin, has been identified (164, 165). Another fatty acid synthesis chokepoint enzyme in *B. pseudomallei* was a biotin-dependent acetyl-CoA carboxylase. Biotin dependent carboxylases comprise a large group of enzymes that participate in a variety of cellular processes, including fatty acid metabolism, amino acid metabolism, carbohydrate metabolism, polyketide biosynthesis, urea utilization, etc. (reviewed by (166)). Acetyl-CoA carboxylase is comprised of two enzymes, biotin carboxylase and carboxyltransferase, and catalyzes the first committed step in fatty acid synthesis (167). Acetyl-CoA carboxylase is an antimicrobial target in *M. tuberculosis* (168), *E. coli* (169, 170), other bacteria and most living organisms (reviewed by (171)). All of the chokepoints in Table 3 are essential genes in *B. pseudomallei* MSHR668 and K96243, as determined by blasting the chokepoint enzyme sequences against essential gene sequences in the Database of Essential Genes (172) and by comparing to the list of essential genes previously identified in K96243 (173).

To determine if *B. pseudomallei* deletion mutants were available for each of the chokepoints in
Table 3, searches of the internet, PubMed, and the Burkholderia Genome Database (http://burkholderia.com) were performed. Based on these searches, none of the chokepoint enzymes in Table 3 had a mutant available; however, a \textit{B. cepacia} \textit{argG} mutant has attenuated virulence (174), and \textit{recA} mutants have been identified in \textit{B. cepacia} (175).

Additional metabolic enzymes, not identified as chokepoints in this study, and pathways critical for bacterial growth and survival have been called out previously with respect to target identification. These include anaplerotic pathways that turned on by limiting carbon sources (176), the glyoxylate shunt enzyme isocitrate lyase (16, 177, 178), involved in the metabolism of fatty acids (15, 177), enoyl-ACP reductase (FabI) in the type II fatty acid biosynthesis pathway (179), and alanine racemase (23).

To gain an understanding of the metabolic processes in \textit{B. pseudomallei} MSHR668 that are active under different environmental conditions, and to test the effect of deletion of each chokepoint enzyme on the growth of \textit{B. pseudomallei in silico}, metabolic network models were constructed and FBA was performed. The first FBA model, of the unconstrained network in MSHR668, included all possible biomass compounds that could be produced, all nutrients that could be consumed, had no weights imposed on the biomass metabolites and no constraints imposed on the nutrients. [Supplemental material, Final_unconstrained_model_inputs.pdf and Final_unconstrained_model_solution.pdf]. This model likely represents the metabolic potential of \textit{B. pseudomallei} in a soil or water environment where abundant carbon and nitrogen sources are available. To mimic the conditions that \textit{B. pseudomallei} experiences in culture, a separate model was constructed that provided only the nutrients present in LB media plus glycerol (43). Constraints were included on some of the nutrients in this model (ADP, Pi, proton and glycerol). [Supplemental material, Final_LB_model_inputs.pdf and Final_LB_model_solution.pdf].

A third model was attempted, to mimic infection conditions using only the nutrients present in the host cell cytosol. However, no comprehensive studies have identified a complete list of host cell nutrients that are available to \textit{B. pseudomallei} during infection. Also, the specific carbon requirements of \textit{B. pseudomallei} in either human macrophages or amoebae have not been determined. After trying to compile a list of nutrients that mimic the content of the host cytosol, from the literature and from gene expression studies of \textit{B. pseudomallei} during infection, this model did not produce a solution so it was abandoned. However, the LB media model contained a similar set of nutrients, so it may in fact be somewhat representative of infection conditions.

In terms of carbon sources, glucose was utilized as a nutrient in the unconstrained model of MSHR668. However, when glycerol was replaced by glucose in the LB media model, no FBA solution was found [data not shown]. This was a somewhat unexpected result, as \textit{B. pseudomallei} can utilize glucose as a carbon source in culture (180). One possible explanation for this result is that additional nutrients required for glucose utilization were missing from the input nutrients list. Co-metabolism of more than one carbon substrate is a metabolic strategy employed by intracellular bacteria replicating inside host cells to provide carbon for energy and biosynthesis (59). For example, \textit{Listeria monocytogenes} can use both glycerol and lactate as carbon sources, (57, 181-183). It has been suggested that during infection by \textit{Listeria}, the host cell may not contain enough glucose to activate bacterial PTS glucose transporters, so alternative carbon sources are important for survival and virulence of the pathogen (57). \textit{M. tuberculosis} also relies
on glycerol and fatty acids as carbon sources in the macrophage environment (184) As the unconstrained *B. pseudomallei* model included a much longer list of nutrients than the LB model and also could use glucose as a nutrient, this could be the case. Another explanation for glucose not being utilized by the LB media model is that intracellular bacteria seem to prefer other substrates over glucose during infection (185), and glycerol may be a major carbon source for intracellular bacteria during infection (185, 186). This may also be the situation for *B. pseudomallei* inside host cells. Glycerol feeds into the second half of the glycolysis/gluconeogenesis pathway through its conversion to dihydroxyacetone phosphate [www.metacyc.org; (187)], bypassing the first four steps of glycolysis. We previously determined that the *B. pseudomallei* MSHR668 genome has the full set of genes to perform this conversion (37). Two studies of *Listeria monocytogenes* infection support the idea that intracellular pathogens generally may use glycerol rather than glucose as a main carbon source while inside host cells. Transcription profiles of *L. monocytogenes* grown in mouse macrophages showed reduced expression of genes encoding some of the enzymes involved in glycolysis, in particular phosphoglucone isomerase (*pgi*), which converts glucose-6-phosphate into fructose-6-phosphate, and the five steps involved in the conversion of glyceraldehyde-3-phosphate to pyruvate (185). Similar transcription profiles were seen during *L. monocytogenes* infection of Caco-2 epithelial cells (188). Both studies showed increased expression of genes involved in the uptake and utilization of glycerol (185, 188): these genes were *glpF, glpK, glpD* and *dhaK*.

To date, no study has determined precisely which carbon substrates are utilized by *B. pseudomallei* during infection of macrophages. In addition to glycerol, there is evidence that *B. pseudomallei* may utilize aromatic carbon compounds such as benzoate and phenylacetic acid as carbon sources for intracellular survival (53). One study showed that in *B. pseudomallei* 1026b, glycolytic pathway and TCA cycle genes were down-regulated during infection of hamster (189), supporting the idea that *B. pseudomallei* may prefer carbon sources other than glucose while inside host cells. Other studies examined genes induced by hypoxia, which is a condition present in infected macrophages (190) and changes in *B. pseudomallei* gene expression during infection of rat lungs (191), but these only provide further clues about *B. pseudomallei*’s carbon preferences during infection.

Complicating the situation even more, the complete nutrient content of a representative mammalian host cell cytosol has not been determined yet, so we are even farther from identifying a consensus set of nutrients present in the cytosol of different host cell types (186). This is largely due to the challenges in designing appropriate infection models and identifying robust analytical approaches to measure metabolic changes occurring in host cells during infection. Because of these limitations on both the pathogen and host sides, it is difficult to predict which carbon sources pathogens can use to grow inside host cells.

While we don’t know the exact biochemical composition of a mammalian cell cytosol, we do know some details about mammalian cells in general. For example, the cytosol of a typical cell has low magnesium, sodium and calcium concentrations, and a high potassium concentration at neutral pH (192). In addition, mammalian cells contain small amounts of amino acids, plus significant amounts of TCA cycle intermediates (193, 194). Once inside host cells, intracellular bacteria may stimulate host cell responses to produce needed nutrients (186). However, host-pathogen interactions during infection are complicated, as some host defense responses are aimed at inhibiting pathogen survival and proliferation, for instance by decreasing metabolic
activities that provide nutrients to the pathogen (176).

This work is the first to use genome scale metabolic modeling to address *B. pseudomallei* metabolism as a source of new drug targets. While identifying the nutrients available to *B. pseudomallei* inside host cells was difficult, the effort described here identified a set of 26 chokepoint enzyme drug targets; *in silico* deletion of 4 of these target enzymes reduced the total biomass flux through the *B. pseudomallei* metabolic network. While this genome-based approach can streamline the initial steps of antibacterial target identification, the true utility of this process will be demonstrated when the targets are experimentally verified by performing knockout experiments in culture, followed by efficacy testing of candidate drugs in culture and in animal models of infection.

**Acknowledgements**

This project was funded by the Defense Threat Reduction Agency

**References**


94. **Lee DS, Burd H, Liu J, Almaas E, Wiest O, Barabási AL, Oltvai ZN, Kapatral V.** 2009. Comparative genome-scale metabolic reconstruction and flux balance analysis of


Figure 1. Metabolic pathways in B. pseudomallei MSHR668 that show reduced flux when chokepoint enzymes (indicated by pink arrows) are deleted in silico. A. The mono-trans, poly-cis decaprenyl phosphate biosynthesis pathway that contains the chokepoint enzyme undecaprenyl diphosphate synthase (uppS). B. The two chokepoints UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-decaprenol N-acetylglucosamine transferase (murG) and UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase (lpxD), involved in peptidoglycan and lipid A biosynthesis, respectively. C. The 3-isopropylmalate dehydrogenase (leuB) chokepoint enzyme performs the third step in leucine biosynthesis. D. The chokepoint enzyme tRNA-guanine transglycosylase (tgt), involved in queosine biosynthesis. In silico deletion of UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase (lpxD) reduced flux through the B. pseudomallei metabolic network in the LB media model, deletion of undecaprenyl diphosphate synthase (uppS) reduced flux through both unconstrained and LB media models, and deletion of tRNA-guanine transglycosylase (tgt) and 3-isopropylmalate dehydrogenase (leuB) reduced flux in the unconstrained model. These pathways were rendered by the Cellular Overview feature of Pathway Tools.