Global identification of CobB interactors by an Escherichia coli proteome microarray

Cheng-xi Liu1,2, Fan-lin Wu1,2, He-wei Jiang1,2, Xiang He1,2, Shu-juan Guo1,2, and Sheng-ce Tao1,2,3*

1Shanghai Center for Systems Biomedicine, Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Jiao Tong University, Shanghai 200240, China
2State Key Laboratory of Oncogenes and Related Genes, Shanghai Jiao Tong University, Shanghai 200240, China
3School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
*Correspondence address. Tel/Fax: +86-21-34207069; E-mail: taosc@sjtu.edu.cn

Protein acetylation is one of the most abundant post-translational modifications and plays critical roles in many important biological processes. Based on the recent advances in mass spectrometry technology, in bacteria, such as Escherichia coli, tremendous acetylated proteins and acetylation sites have been identified. However, only one protein deacetylase, i.e. CobB, has been identified in E. coli so far. How CobB is regulated is still elusive. One right strategy to study the regulation of CobB is to globally identify its interacting proteins. In this study, we used a proteome microarray containing ~4000 affinity-purified E. coli proteins to globally identify CobB interactors, and finally identified 183 binding proteins of high stringency. Bioinformatics analysis showed that these interacting proteins play a variety of roles in a wide range of cellular functions and are highly enriched in carboxylic acid metabolic process and hexose catabolic process, and also enriched in transferase and hydrolase. We further used bio-layer interferometry to analyze the interaction and quantify the kinetic parameters of putative CobB interactors, and clearly showed that CobB could strongly interact with TopA and AccC. The novel CobB interactors that we identified could serve as a start point for further functional analysis.

Keywords CobB; E. coli; proteome microarray; protein–protein interaction; acetylation

Introduction

Protein acetylation is abundant in both eukaryotes and prokaryotes, it plays critical roles in many biological processes (BPs), e.g. metabolism [1,2], cell cycle [3], apoptosis [4], stress responses [5], aging [6], and DNA repair [7]. Protein acetylation is reversible, dynamic, and highly conserved. With recent advances in mass spectrometry, tremendous acetylated proteins and acetylation sites have been identified in bacteria, such as the model organism Escherichia coli. The acetylated proteins include many functionally important proteins, such as acetyl-coenzyme A synthetase (ACS), chemotaxis response regulator protein (CheY), arylamine acetyltransferase (NhoA) [8–10]. For example, Zhang et al. [11] recently identified 349 acetylated proteins in E. coli, in a striking contrast, there is only one protein deacetylase CobB which has been identified. CobB, a homolog of eukaryotic Sir2 [12], was first reported as a deacetylase which could remove lysine-609 acetylation of ACS and then regulated the enzyme activity in Salmonella enterica [10]. Although several substrates of CobB have been identified individually [8–10,13], how CobB could efficiently deacetylate so many proteins in a dynamic fashion is still elusive, the regulation of CobB may fundamentally lie in its interacting proteins. We need a powerful tool to globally identify the CobB interactors.

Protein microarrays are miniaturized, parallel assay systems that are constructed by spotting hundreds to thousands of affinity-purified proteins on a solid surface at high density [14–17]. Proteome microarrays have been successfully applied in a variety of studies, such as biomarker discovery, cell surface marker/glycosylation profiling, post-translational modifications (PTMs), protein–protein, protein–lipid, protein–DNA, protein–small compound, and protein–peptide interactions, and enzyme substrates identification [16–20]. Recently, a proteome microarray carrying ~4000 E. coli proteins has been constructed. It could serve as an ideal platform for global identification of CobB-interacting proteins.

Herein, we used the E. coli proteome microarray for global identification of CobB-interacting proteins. We have successfully identified 183 CobB interactors. Bioinformatics analysis showed that these interacting proteins play roles in a wide range of BPs. They were highly enriched in carboxylic acid metabolic process and hexose catabolic process, purine ribonucleotide binding and ATP binding. Further kinetic analysis clearly validated the strong interactions between CobB, TopA, and AccC.
Materials and Methods

Chemicals and reagents
All chemicals were purchased from Sigma-Aldrich (St Louis, USA) unless otherwise stated. Ni-NTA Super Flow was purchased from QIAGEN (Hilden, Germany). Biotin, EZ-Link® Sulfo-NHS-LC-Biotin, Zeba™ Spin Desalting Columns, 7K MWCO, and Piece® BCA Protein Assay Kit were from Thermo Scientific (Rockford, USA). The GenePix 4200A slide scanner and GenePix Pro 6.0 software were obtained from Axon Instruments (Elk Grove Village, USA). The SD buffer (10 mM PBS + 0.1% BSA + 0.02% Tween 20, pH 7.4) were provided by Fortebio (Menlo Park, USA). Fast Slides were from Whatman (Kent, UK). The E. coli clones for protein over-expression were from the ASKA library.

Proteome microarray construction
Briefly, ~4000 proteins representing the majority of the E. coli proteome were affinity purified with Ni-NTA Beads. After the quality of the purified proteins were assessed by western blot analysis with an anti-6xHis antibody, each protein was spotted twice on a FAST slide using a SmartArrayer 48 microarrayer from CapitalBio Co. (Beijing, China). The printed microarrays were stored at −80°C prior to use.

Candidates purification and CobB biotinylation
The E. coli strains were from the ASKA library [21]. All strains were cultured in Luria–Bertani medium and induced by 1 mM isopropyl-β-D-thiogalactopyranoside when OD600 reached 0.6, and cultured for another 4 h. Cells were harvested by centrifugation and washed with phosphate-buffered saline (PBS). Lysis buffer (50 mM NaH2PO4, 300 mM NaCl, 10 mM imidazole, 1 × CelLytic B, 50 U/ml of Benzonase, 1 mM phenylmethanesulfonyl fluoride, and pH = 8) was added. The supernatants were collected by centrifugation. The recombinant protein was captured by Ni-NTA beads and incubated at 4°C for 2 h with vigorous shaking. After several rounds of washes, the proteins were eluted and then desalted with Zeba™ Spin Desalting Column, 7K MWCO. Protein concentration was determined by BCA Protein Assay. The purified CobB was biotinylated by adding EZ-Link® Sulfo-NHS-LC-Biotin at mol/mol 1 : 20, at 4°C for 2 h. The excess of biotin was neutralized by 20 mM glycine at 4°C for 30 min. The biotinylated protein was then desalted.

CobB-interacting protein identification using the E. coli proteome microarray
Proteome microarrays were blocked with blocking buffer [3% bovine serum albumin (BSA) in 0.1% Tween 20, PBST] for 1 h at room temperature with gentle agitation. The protein microarray was incubated with 50 ng/μl biotin-CobB in PBST. A control microarray was also included with 50 ng/μl biotin. The reactions were carried out at 4°C overnight. The microarrays were washed with PBST for three times, 5 min each. The microarray was incubated with 1 : 1000 diluted Cy3-streptavidin in PBST buffer for 2 h at room temperature, followed by three times (10 min each) washes in PBST and rinsed with ddH2O for 3 min. The microarrays were spun dry at 300 g for 3 min and subject to scanning at 532 nm with GenePix 4200A microarray scanner (Axon Instruments) to visualize and record the results. The GenePix Pro 6.0 (Axon Instruments) was used for data extraction.

Microarray data processing
To generate the list of CobB-interacting proteins, the signal-to-noise ratio (SNR) was defined as F532 Median/B532 Median.
Median and was calculated for all the spots on the E. coli proteome microarray. The SNR of a protein was averaged from the two duplicated spots and then calculated for the two microarrays incubated with biotin-CobB and biotin as SNR (+) and SNR (−), respectively. The Calling_Score was defined as SNR (+)/SNR (−) and was calculated for all the proteins. CobB-interacting protein candidates were identified by examining the SNR (+) and Calling_Scores for each protein. For the final candidates, the cutoff was set as SNR (+) ≥ 3 and Calling_Score ≥ 3.

Bioinformatics analysis
The newly identified CobB interactors (Supplementary Table S1) were classified using the PANTHER (Protein Analysis through Evolutionary Relationships) classification system with default settings. Gene symbols were used as input for the classification system [22]. The interacting proteins were then analyzed to determine if any types of proteins are over-represented. Gene ontology (GO) analysis [23] was performed using the web-accessible program DAVID Bioinformatics Resources 6.7 [24]. The default E. coli proteome was used as the background list. The significance of the enrichments was statistically evaluated with a modified Fisher’s exact test (EASE score), and a P-value for each term was calculated by applying a Benjamini–Hochberg false discovery rate correction [25]. The BPs and molecular functions (MFs) of the CobB interactome were analyzed. Protein interaction networks for the CobB interactome were generated by STRING [26].

Kinetic assay
The ForteBio Octet™ system was used to measure the binding kinetics of CobB and its interactors. The biotinylated CobB was tethered on the tip of a streptavidin-coated sensor. The binding partners in SD buffer was then exposed to the tethered one, and the binding was measured by coincident change in the interference pattern [27,28].

Results
Identification of CobB-interacting proteins on proteome microarray
To globally identify the interactors for CobB, an E. coli proteome microarray with ~4000 individually affinity-purified E. coli proteins was incubated with biotinylated CobB. A parallelly processed microarray with biotin was carried out as a negative control. The microarrays were further incubated with a Cy3-labeled streptavidin (Fig. 1A,B). This experiment was repeated twice. The SNR for each spot was defined as the median foreground/median background, and the SNR associated with each protein was the average from two duplicated spots. The SNR of the experiment with

Table 1. The top 20 list of the CobB interactors identified on the E. coli proteome microarray

<table>
<thead>
<tr>
<th>No.</th>
<th>UniProt</th>
<th>Gene name</th>
<th>Lias/description</th>
<th>SNR</th>
<th>Calling_Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P69797</td>
<td>manX</td>
<td>Fused mannose-specific PTS enzyme IIAB components</td>
<td>19.87574349</td>
<td>20.10804043</td>
</tr>
<tr>
<td>2</td>
<td>P75791</td>
<td>ybiU</td>
<td>Hypothetical protein</td>
<td>18.95806163</td>
<td>19.09674859</td>
</tr>
<tr>
<td>3</td>
<td>P00962</td>
<td>glnS</td>
<td>Glutamyl-tRNA synthetase</td>
<td>17.48503496</td>
<td>17.70621774</td>
</tr>
<tr>
<td>4</td>
<td>P0A8F4</td>
<td>udk</td>
<td>Uridine/cytidine kinase</td>
<td>12.59949791</td>
<td>12.74921225</td>
</tr>
<tr>
<td>5</td>
<td>P0AC53</td>
<td>zwf</td>
<td>Glucose 6-phosphate-1-dehydrogenase</td>
<td>12.64482579</td>
<td>12.63753878</td>
</tr>
<tr>
<td>6</td>
<td>P21170</td>
<td>speA</td>
<td>Biosynthetic arginine decarboxylase, PLP binding</td>
<td>11.45036124</td>
<td>11.63951783</td>
</tr>
<tr>
<td>7</td>
<td>P42596</td>
<td>ygjO</td>
<td>23S rRNA m2G1835 methyltransferase</td>
<td>11.2378174</td>
<td>11.4484926</td>
</tr>
<tr>
<td>8</td>
<td>P22259</td>
<td>pckA</td>
<td>Phosphoenolpyruvate carboxykinase</td>
<td>11.30765451</td>
<td>11.3194389</td>
</tr>
<tr>
<td>9</td>
<td>P0A6E4</td>
<td>argG</td>
<td>Argininosuccinate synthetase</td>
<td>10.20435984</td>
<td>10.3611884</td>
</tr>
<tr>
<td>10</td>
<td>P29208</td>
<td>menC</td>
<td>o-Succinylbenzoate synthase</td>
<td>9.073751976</td>
<td>9.15188335</td>
</tr>
<tr>
<td>11</td>
<td>P05804</td>
<td>uidA</td>
<td>Beta-6-glucuronidase</td>
<td>8.78233065</td>
<td>9.054370215</td>
</tr>
<tr>
<td>12</td>
<td>P37146</td>
<td>nrdF</td>
<td>Ribonucleoside-diphosphate reductase 2, β subunit dimer</td>
<td>8.859965439</td>
<td>8.999049175</td>
</tr>
<tr>
<td>13</td>
<td>P29745</td>
<td>pepT</td>
<td>Peptidase T</td>
<td>8.816453404</td>
<td>8.920128387</td>
</tr>
<tr>
<td>14</td>
<td>P27830</td>
<td>rffG</td>
<td>dTDP-Glucose 4,6-dehydratase 2</td>
<td>8.725653687</td>
<td>8.869724601</td>
</tr>
<tr>
<td>15</td>
<td>P24182</td>
<td>accC</td>
<td>Acetyl-CoA carboxylase, biotin carboxylase subunit</td>
<td>8.610630533</td>
<td>8.477415095</td>
</tr>
<tr>
<td>16</td>
<td>P0A759</td>
<td>nagB</td>
<td>Glucosamine-6-phosphate deaminase</td>
<td>8.264939999</td>
<td>8.286098278</td>
</tr>
<tr>
<td>17</td>
<td>P32138</td>
<td>yibQ</td>
<td>α-Glucosidase</td>
<td>7.98904906</td>
<td>8.076718758</td>
</tr>
<tr>
<td>18</td>
<td>P09026</td>
<td>dsdA</td>
<td>β-Serine ammonia-lyase</td>
<td>7.68824271</td>
<td>7.741425553</td>
</tr>
<tr>
<td>19</td>
<td>P15286</td>
<td>div</td>
<td>Predicted flagella assembly protein</td>
<td>7.5883823</td>
<td>7.716503073</td>
</tr>
<tr>
<td>20</td>
<td>P77243</td>
<td>prpD</td>
<td>2-Methylcitrate dehydratase</td>
<td>7.670180926</td>
<td>7.707573723</td>
</tr>
</tbody>
</table>
biotinylated CobB and the SNR of the experiment with biotin were designated SNR (+) and SNR (−), respectively, and a Calling_Score, SNR (+)/SNR (−), was calculated for each protein. One hundred eighty-three proteins (Table 1 and Supplementary Table S1) with SNR (+) ≥3 and Calling_Score ≥3 were identified as potential CobB interactors. Eight representatives of potential interactors are shown in Fig. 1C along with their SNR (+) values and Calling_Scores.

**Categorization and GO analysis of the CobB interactors**

To understand the biological relevance of the potential CobB interactors, we applied the online protein classification tool, PANTHER, to the 183 identified CobB interactors to identify enrichment for specific BPs, FNs, and protein classes. We found that the CobB interactors could be classified into 10 BP groups (Fig. 2A and Supplementary Table S2), the largest of which was metabolic processes group (GO: 0008152) (67%). In terms of FN, the candidates could be classified into nine groups (Fig. 2B), the larger of which was catalytic activity (GO: 0003824) (75%), followed by binding (GO: 0005488) (12%). Finally, the candidates could be classified into 16 protein class groups (Fig. 2C), the top 3 of which were transferase (PC00220) (17%), hydrolase (PC00121) (16%), and oxidoreductase (PC00176) (16%). These data indicate that CobB is a critical regulator of diverse cellular functions involving widespread BPs.

To gain insight into possible functional roles of the identified CobB binding proteins, the enrichment of ontology terms and components of molecular pathways of the candidates were analyzed using DAVID in comparison with their occurrence in the *E. coli* proteome. The candidate proteins were examined for enrichment in BP and FN. For BP, the top 11 GO terms of *P* < 0.05 were shown in Fig. 3A and Supplementary Table S3. The candidate list is substantially enriched in carboxylic acid metabolic process (*P* = 1.30 × 10^-14), hexose catabolic process (*P* = 2.00 × 10^-5), and tRNA aminocacylation (*P* = 1.90 × 10^-3). The candidate list is also significantly enriched for fatty acid biosynthetic process, and pyrimidine nucleotide biosynthetic process. For MF, the GO terms of *P* < 0.05 were shown in Fig. 3B. There is a significant amount of the candidates localizing in purine ribonucleotide binding (*P* = 7.50 × 10^-6). Besides the ATP binding (*P* = 7.60 × 10^-5), other significantly enriched MF terms were 2, 3-bisphosphoglycerate-dependent phosphoglycerate mutase activity (*P* = 7.50 × 10^-6) and aminoaeryl-tRNA ligase activity (*P* = 3.30 × 10^-4).

To create significance out of the newly identified CobB interactors, biological protein–protein interaction networks of these proteins were constructed. The CobB-interacting proteins were imported into STRING [26] to build the network (Fig. 4A). Two enriched clusters centered on AccC and TopA were also identified (Fig. 4B,C).

**Validation of CobB-interacting proteins by bio-layer interferometry**

To further validate the microarray experiments, we used bio-layer interferometry (BLI) to analyze the interaction and determine the kinetic parameters of putative CobB interactors. In this work, we evaluated the interaction between CobB and YccW, YdaL, TopA, AccC, NhoA, and YcdC. BSA was included as a negative control. The results clearly showed that CobB could strongly interact with TopA and AccC, but not with BSA under the same conditions. Furthermore, we determined the equilibrium dissociation constants (*K_D*) for the interactions between CobB, TopA, and AccC (Fig. 5).
Hence, as might be expected from their identification in the microarray experiment, these interactions are relatively with high affinity.

**Discussion**

Protein lysine acetylation is one of the most abundant PTMs in both eukaryotes and prokaryotes. Weinert et al. [29] showed that acetyl-phosphate (AcP) can chemically acetylate proteins non-enzymatically in bacteria in a global, uniform manner, which may explain the abundance of protein acetylation. However, how protein acetylation is reversed is still not clear.

The *E. coli* proteome microarray is well suited for global identification of protein–protein interactions for CobB for the following reasons: (i) the proteins on the microarray were over-expressed and purified from *E. coli*, which should maintain their designated biological functions and binding activities; (ii) each protein is physically addressable on the microarray, thus facilitating the identification of novel protein–protein interactions.

In this study, we used the *E. coli* proteome microarray to globally screen the interactors of CobB. We have identified 183 proteins that could potentially interact with CobB. The CobB interactors (Table 1 and Supplementary Table S1) were significantly enriched for carboxylic acid metabolic process, hexose catabolic process, and tRNA aminoacylation, which suggested that CobB might be involved in the regulation of metabolism and biosynthesis. By protein class analysis, we found that most proteins were enriched in functionally important enzymes, such as transferase, hydrolase, oxidoreductase, ligase, isomerase, and so on, which reflected...
that CobB influence metabolism and biosynthesis through regulating a series of key enzyme by protein deacetylation. We selected nine candidates, i.e. YcdC, YdaL, YccW, NhoA, TopA, AccC, ManX, YbiU, and FabD to validate their interaction with CobB in vitro by kinetics analysis. These candidates were chosen because some were on the top list of Supplementary Table S1 (ManX, YbiU, YdaL, AccC, FabD) or because of their functional relevance (NhoA, YccW, YcdC, TopA) [8]. The results clearly showed that CobB could strongly interact with TopA and AccC.

TopA is responsible for relaxing negative supercoils in DNA and function by causing single-strand DNA breaks [30], so TopA is essential for growth of E. coli. Excess negative supercoiling that results from mutations in the topA gene promotes the formation of R-loops during transcription, inhibiting RNA synthesis, and inhibits the growth [31]. Yu et al. [32] reported that the lysine 484 of TopA was acetylated. The acetylation of TopA may be related to DNA damage response through CobB regulation.

AccC is a subunit composition of acetyl-CoA carboxylase, which is one of the key enzymes in the biosynthesis of fatty acids. This enzyme belongs to the family of enzymes that catalyze the intermolecular transfer of carboxyl groups via the transient formation of a carboxyphosphate intermediate covalently linked to a biotin prosthetic group [33]. Galdieri and Vancura [34] reported that acetyl-CoA carboxylase regulated global histone acetylation. In yeast, Acc1p (homolog of AccC) regulates the availability of acetyl-CoA
for histone acetyltransferases, thus representing a link between metabolism and epigenetic for transcriptional regulation.

To our knowledge, prior to this study, there are no experimentally validated CobB interactors from *E. coli*. It is known that enzymes usually interact with its corresponding protein substrates. There are a few well-studied CobB substrates, e.g. NhoA, ACS, and CheY; however, the Calling_Scores of these are only 1.8, 1.1, and 1.3, which reflected the weak binding of CobB to these substrates. By STRING analysis, we identified several possible functional partners of CobB, i.e. CheY, NagA, ManX, GlmM, ACS, NagB, and NagK. Among these proteins, there are strong interactions between NagA, NagB, and CobB, with Calling_Scores of 4.6 and 8.3, respectively. These results suggested the reliability of the interactions identified on the *E. coli* proteome microarray.

Taken together, we have performed the global protein–protein interaction analysis of *E. coli* CobB using the *E. coli* proteome microarray. We have identified 183 CobB interactors, and most of them are novel and involved in a wide range of BPs. We believe that the CobB-interacting proteins we have identified in this study will serve as a starting point for a more comprehensive exploration of the role of acetylation.

**Supplementary Data**

Supplementary Data are available at *ABBS* online.

**Funding**

This work was supported by the grants from the National Natural Science Foundation of China (31370813), the State Key Development Program for Basic Research of China (2010CB529205), the National High Technology Research and Development Program of China (2012AA020103 and 2012AA020203).

**References**


**Figure 5.** Validation the interaction between CobB and the candidate interactors by BLI. (A) Binding kinetics of CobB immobilized on SA biosensors and the CobB interactors TopA. (B) Binding kinetics of CobB immobilized on SA biosensors and the CobB interactors AccC. (C) Parameters calculated from (A) and (B).