


Functional dissection and modulation of the BirA protein for improved autotrophic growth of gas-fermenting *Clostridium ljungdahlii*

Can Zhang,^{1,2} Xiaoqun Nie,¹ Huan Zhang,^{1,2} Yuwei Wu,^{1,2} Huiqi He,¹ Chen Yang,¹ Weihong Jiang^{1,*} and Yang Gu^{1,*} 

¹Key Laboratory of Synthetic Biology, The State Key Laboratory of Plant Carbon-Nitrogen Assimilation, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, 200032, China.

²University of Chinese Academy of Sciences, Beijing, China.

Summary

Gas-fermenting *Clostridium* species can convert one-carbon gases (CO₂/CO) into a variety of chemicals and fuels, showing excellent application prospects in green biological manufacturing. The discovery of crucial genes and proteins with novel functions is important for understanding and further optimization of these autotrophic bacteria. Here, we report that the *Clostridium ljungdahlii* BirA protein (CBirA) plays a pleiotropic regulator role, which, together with its biotin protein ligase (BPL) activity, enables an effective control of autotrophic growth of *C. ljungdahlii*. The structural modulation of CBirA, combined with the *in vivo* and *in vitro* analyses, further reveals the action mechanism of CBirA's dual roles as well as their interaction in *C. ljungdahlii*. Importantly, an atypical, flexible architecture of the binding site was found to be employed by CBirA in the regulation of a lot of essential pathway genes, thereby expanding BirA's target genes to a broader range in clostridia. Based on these findings, molecular modification of CBirA was performed, and an improved cellular performance of *C. ljungdahlii* was achieved in gas fermentation. This work reveals a previously unknown potent role of BirA in gas-fermenting clostridia, providing new perspective for

understanding and engineering these autotrophic bacteria.

Introduction

Gas-fermenting *Clostridium* species, an important class of autotrophic bacteria, can fix and assimilate C1 gases (CO₂/CO) via the Wood–Ljungdahl pathway (WLP) to produce multiple organic acids and alcohols, thereby exhibiting huge potential for the sustainable carbon-neutral production of chemicals and biofuels (Berg *et al.*, 2010). To effectively use C1 gases, these bacteria have evolved their metabolic pathways and regulatory systems, allowing them to adapt to autotrophic growth conditions. For example, a large number of crucial genes involved in carbon fixation and energy supply in gas-fermenting *Clostridium* species showed significant transcriptional changes under different fermentation conditions (C1 gases vs. sugars) (Tan *et al.*, 2013; Aklujkar *et al.*, 2017), indicating a unique metabolic regulation model in these autotrophic bacteria in gas fermentation.

To unlock the full potential of gas-fermenting *Clostridium* species, the discovery of more crucial regulatory elements and a comprehensive understanding of underlying regulatory mechanism are necessary. However, to date, this aspect remains minimally explored due to the challenges in genetic tools. Specific to transcriptional regulators (TFs), to our knowledge, only a global regulator CcpA and a TetR-family regulator (CAETHG_0459) have been identified and characterized in these autotrophic bacteria (Lemgruber *et al.*, 2019; Zhang *et al.*, 2020a,2020b). CcpA was found to directly regulate the expression of some WLP genes in *Clostridium ljungdahlii*, and the elimination of CcpA's inhibition on these genes could significantly enhance CO₂ utilization (Zhang *et al.*, 2020a,2020b). CAETHG_0459, as a transcription activator in *Clostridium autoethanogenum*, can directly bind to the core enzyme of RNA polymerase and subsequently activate the expression of multiple WLP genes (Lemgruber *et al.*, 2019). These studies have indicated the importance of TFs in gas-fermenting clostridia, thereby strongly supporting a continued investigation.

Recently, based on the Pfam database (<http://pfam.xfam.org/>), we used the similarity search and domain prediction for mining for TFs in *C. ljungdahlii*. Over 400 TFs were

Received 26 February, 2021; accepted 22 June, 2021.

For correspondence. *E-mail wjiang@cemps.ac.cn; Tel. 86 21 54924172; Fax 86 21 54924015. **E-mail ygu@cemps.ac.cn; Tel. 86 21 54924284; Fax 86 21 54924015.

Microbial Biotechnology (2021) 14(5), 2072–2089
doi:10.1111/1751-7915.13884

© 2021 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

predicted and subsequently subjected to functional screening through gene deletion or transcriptional repression using the CRISPR-Cas-based genome editing tools established in our lab (Huang *et al.*, 2016; Zhao *et al.*, 2019). Among these TFs, BirA has been known to be an important bifunctional protein with both the biotin protein ligase (BPL) activity and biotin operon repressor function in some model bacteria, such as *Escherichia coli*, *Bacillus subtilis* and *Staphylococcus aureus* (Weaver *et al.*, 2001; da Costa *et al.*, 2012; Feng *et al.*, 2014). As the ligase, BirA is responsible for the biotinylation of acetyl-CoA carboxylase (ACC) and pyruvate carboxylase (PYC), thereby significantly influencing the activity of these two enzymes (Li and Sousa, 2012; Peters-Wendisch *et al.*, 2012); as the biotin operon repressor, BirA regulates biotin synthesis and uptake and thus enables the *in vivo* biotin homeostasis (Ye *et al.*, 2016; Zhang *et al.*, 2016). Additionally, a recent study revealed that BirA can interact with other regulators to realize a wide regulatory scope in enterohaemorrhagic *E. coli* (Yang *et al.*, 2015). However, specific to autotrophic bacteria, the function of BirA remains unexplored.

In this study, we identified a functional BirA (CBirA) protein that greatly affected the performance of *C. ljungdahlii* in gas fermentation. Next, we dissected the action mechanisms of CBirA through a combination of omics (ChIP-seq), biochemical and genetic approaches. The data from these experiments revealed the pleiotropic regulator role of CBirA as well as the interaction between its dual functions in *C. ljungdahlii*. Furthermore, a new atypical flexible architecture of CBirA-binding sites that are widespread in *C. ljungdahlii* was found, revealing a wider regulatory scope of the BirA protein in this bacterium. Based on these findings, we achieved an improved autotrophic growth and product synthesis of *C. ljungdahlii* in gas fermentation through the functional modulation of CBirA.

Results

CBirA is crucial for the autotrophic growth of C. ljungdahlii in gas fermentation

As mentioned above, among the TFs that have been examined by gene deletion or transcriptional repression in *C. ljungdahlii*, CBirA was found to be crucial for strain growth in gas fermentation. We compared the growth and product synthesis of the *birA*-deleted (Clju $\Delta birA/p$) and wild-type (Clju WT/p) *C. ljungdahlii* strains in gas fermentation. Here, because BirA has been known to specifically regulate biotin synthesis in bacteria (Weaver *et al.*, 2001; da Costa *et al.*, 2012), the media with and without biotin addition were both adopted.

The results showed that the Clju $\Delta birA/p$ strain could not grow in either medium (Fig. 1A and B), and

consequently, almost no product (acetic acid or ethanol) was synthesized by this mutant (Fig. 1C and D). However, after *birA* was reintroduced into the *birA*-deleted mutant strain *via* an expression plasmid, the growth and product synthesis (acetic acid and ethanol) of the resulting strain (Clju $\Delta birA/pbirA$) was restored to the level of the wild-type strain (Clju WT/p) (Fig. 1). Such a deficiency of the Clju $\Delta birA/p$ strain was also found in the fermentation using fructose as the carbon source (Fig. S1). All these findings suggest that the *birA* gene is essential for the growth and product formation of *C. ljungdahlii* in fermenting both C1 gases and sugars.

In addition, we found that extra biotin supplementation in the medium could promote the growth of both the Clju WT/p and Clju $\Delta birA/pbirA$ strains, whereas no promotion effect was observed for the Clju $\Delta birA/p$ strain (Fig. 1A and B). Therefore, it seems that the deletion of *birA* greatly impaired biotin metabolism in *C. ljungdahlii* that may be crucial for cell growth. Furthermore, in light of the BPL activity of BirA, the deletion of *birA* will destroy the *in vivo* biotinylation of ACC and PYC, two crucial enzymes for the basic metabolism of cells. The above results also suggest the possibility that CBirA is involved in the regulation of crucial metabolic processes other than biotin synthesis in *C. ljungdahlii*. In other words, CBirA may play a pleiotropic regulatory role in *C. ljungdahlii*, and the deletion of *birA* would result in a broad impact on this anaerobic bacterium.

CBirA mediates a pleiotropic regulation in C. ljungdahlii

To prove the pleiotropic regulator role of CBirA in *C. ljungdahlii*, we used the chromatin immunoprecipitation sequencing (ChIP-Seq) analysis to screen the potential target genes that are directly regulated by CBirA. The results showed that 24 118 060 and 26 872 496 raw reads were generated from the ChIP sample and mock ChIP sample (negative control, without the addition of antibodies in ChIP) respectively. After removing low-quality reads, the remaining 22 815 380 and 25 796 238 clean reads were aligned to the reference genome of *C. ljungdahlii*, yielding 20 640 883 and 24 853 697 uniquely mapped reads (with < 9.5% and 3.7% mismatches), respectively (Table S1), which reached 668 \times and 805 \times coverage of the *C. ljungdahlii* reference genome (Table S1). As expected, 439 potential CBirA-binding sites (fold enrichment ≥ 1.5) linked with 439 genes were found (Fig. 2A), including *bioY*, a well-known target gene directly controlled by BirA in *E. coli* and *B. subtilis* (Weaver *et al.*, 2001; Feng *et al.*, 2014). Among these potential CBirA's target genes, many were associated with crucial physiological and metabolic processes in *C. ljungdahlii* (Fig. 2B). All these

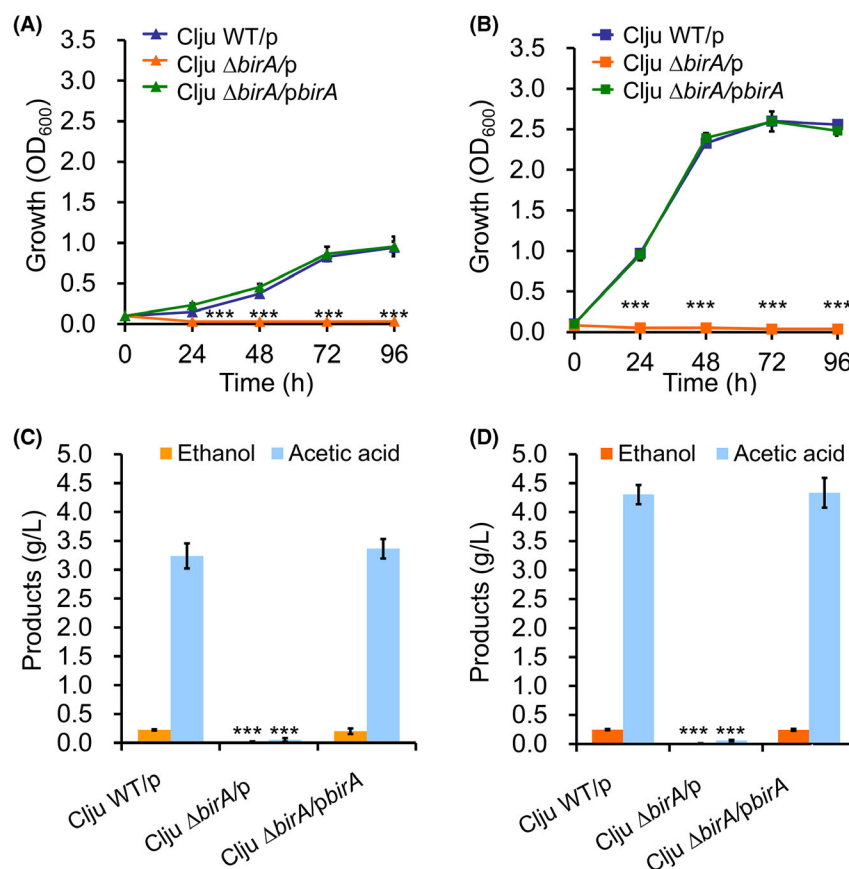


Fig. 1. The influence of the deletion of the *C*BirA-encoding gene (*birA*) on the autotrophic growth and product formation of *C. ljungdahliae* in gas fermentation.

A. The effect of the *birA* deletion on the growth of *C. ljungdahliae* grown in the medium without biotin.

B. The effect of the *birA* deletion on the growth of *C. ljungdahliae* grown in the medium supplemented with biotin.

C. The effect of the *birA* deletion on product formation of *C. ljungdahliae* grown in the medium without biotin.

D. The effect of the *birA* deletion on product formation of *C. ljungdahliae* grown in the medium supplemented with biotin. All the fermentations were performed in the modified ATCC medium 1754. Clju WT/p: the WT strain containing an empty plasmid. Clju Δ birA/p: the *birA*-deleted strain containing an empty plasmid. Clju Δ birA/pbirA: the *birA*-deleted strain complemented with the plasmid-carried *birA* expression. Data shown are means \pm standard deviations ($n = 3$) calculated from triplicate individual experiments. Error bars show standard deviations. Statistical analysis was performed by a two-tailed Student's *t*-test. *** $P < 0.001$ vs. the Clju WT/p strain.

data strongly suggest that *C*BirA has a wide regulatory scope in *C. ljungdahliae*.

To examine the quality of the ChIP-seq results, we selected 42 potential *C*BirA-binding sites that are associated with the genes responsible for oxidation reduction/energy metabolism and one-carbon source metabolism for the *in vitro* EMSA analysis. Finally, 24 sites (associated with 24 genes) were confirmed to be bound by *C*BirA (Fig. S2), including a site in the coding sequence region of an essential WLP gene, *meTr* (Fig. 2C). These data suggest a high credibility of the ChIP-seq data. The discovery of the *C*BirA-binding to *meTr* indicates that *C*BirA can regulate WLP genes in *C. ljungdahliae*. To test this anticipation, we examined the interactions between *C*BirA to the other WLP genes (*cooS1*, *cooS2*, *codH*, Clju_c06990, Clju_c08930, Clju_c20040, *fhs*, *folD*, *metF*,

acsA, *acsC*, *acsD*) via EMSAs. As expected, a new *C*BirA-binding site was found in the coding sequence region of *fdh* (Clju_c06990) (Fig. 2C), although it was not detected in the ChIP-seq analysis. Therefore, these results suggest that *C*BirA can directly regulate carbon fixation and assimilation in *C. ljungdahliae*.

Next, we employed the *in vivo* approach to further verify the pleiotropic regulation of *C*BirA in *C. ljungdahliae*. The structure of *C*BirA was first evaluated by a homology modelling using the BirA protein from *Staphylococcus aureus* subsp. *aureus* ECT-R 2 as the template. As shown in Fig. 3A, the modelled ribbon structure of *C*BirA included a typical N-terminal DNA-binding domain and a C-terminal with enzymatic activity, thus defining *C*BirA as a type-II BirA (Satiaputra *et al.*, 2016). Based on this model, the DNA sequence

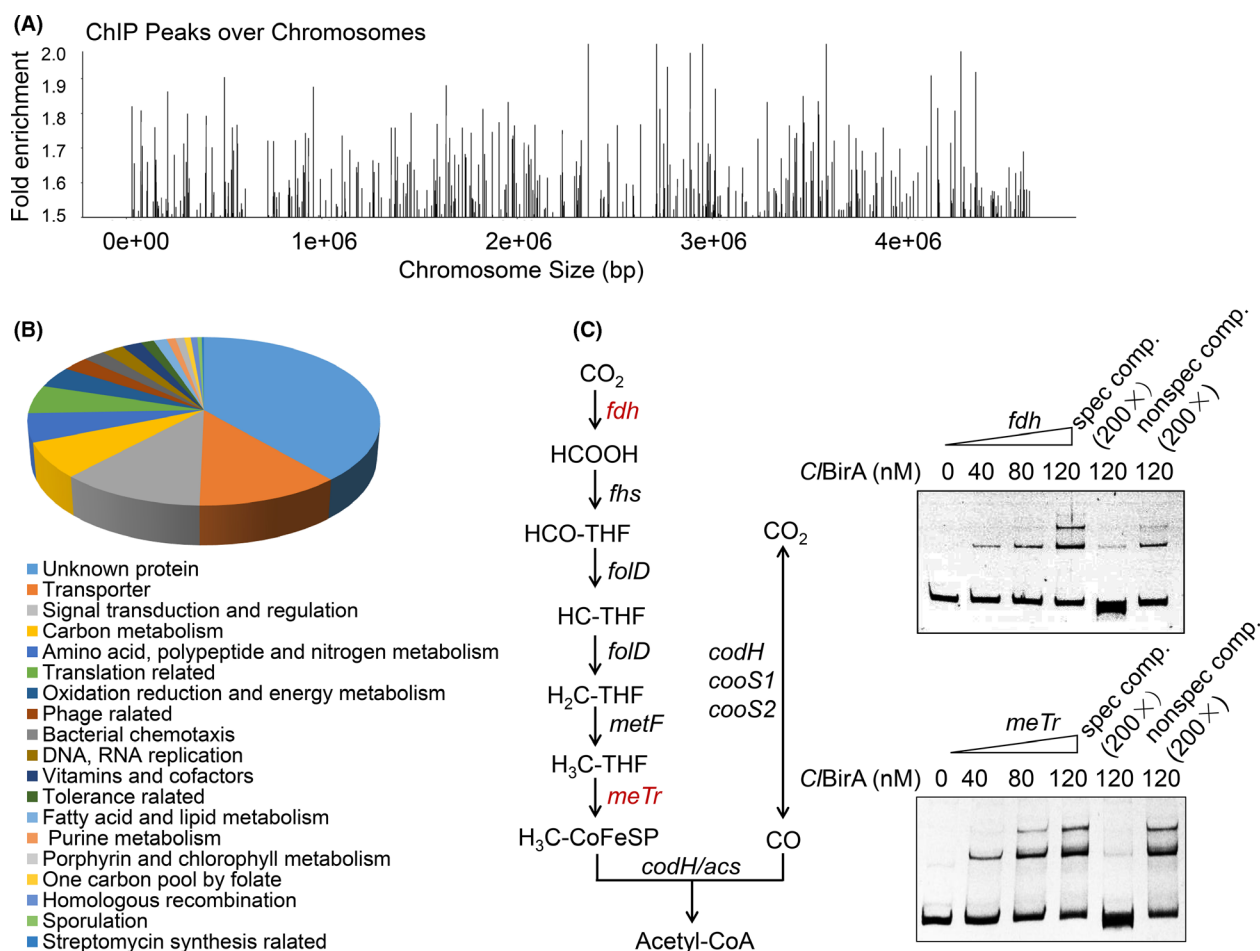


Fig. 2. ChIP-seq analysis of C/BirA-binding sites in the *C. ljungdahliae* genome.

A. Genome wide mapping of C/BirA-binding sites. On the y-axis is the fold enrichment of the peaks.

B. A pie chart presenting the percentage of C/BirA's targets with functional categories defined in the *C. ljungdahliae* database.

C. The *in vitro* confirmation of C/BirA binding to the coding sequence regions of two essential WLP (Wood-Ljungdahl pathway) genes (*fdh* and *meTr*). The coding sequence regions of *fdh* and *meTr* were used as the DNA probes for EMSAs. Fdh, formate dehydrogenase; Fhs, formyl-THF synthetase; FolD, Formyl-THF cyclohydrolase/methylene-THF dehydrogenase; MetF, methylene-THF reductase; MeTr, methyltransferase; CodH/Acs, carbon monoxide dehydrogenase/acetyl-CoA synthase.

coding for the N-terminal 64 amino acid residues (residue 2 to 65) of C/BirA was deleted in the chromosome, yielding a mutant strain (Clju $\Delta birA$ -N) lacking the regulatory function (Fig. S3). By comparing the transcriptional levels of the aforementioned 24 C/BirA's target genes (Fig. S2) between the Clju $\Delta birA$ -N mutant and WT strain (Clju WT), we found that most genes were up-regulated in Clju $\Delta birA$ -N (Fig. 3B and C). These data confirmed the pleiotropic regulator role of C/BirA. Moreover, it seems likely that C/BirA mainly acts as a repressor in *C. ljungdahliae*.

Discovery of an atypical architecture of C/BirA-binding sites that is wide occurrence in *C. ljungdahliae*

Interestingly, a further examination of the above 24 binding sites of C/BirA (Fig. S2) revealed that only 13

contained a typical BirA-binding motif (Rodionov *et al.*, 2002) (Fig. 4A). This finding indicates the existence of atypical motifs for the remaining 11 sites. To verify this hypothesis, two C/BirA-binding sequences (within the promoter regions of Clju_c04680 and Clju_c23900, respectively) were selected from the above 11 candidates for DNase I footprinting analysis. As expected, an obvious protection region was observed in either of the two DNA probes in the presence of BirA protein (Fig. 4B). A common feature of these two protection regions was the two 5-nt repeats (quasi-palindromic sequence) separated by a variable intervening spacer region, namely 'TTTAC-N₁₆-GTAA' and 'TTGTC-N₁₉-GGTAA' (Fig. 4C). The architecture of these two sites, to our knowledge, is different from all the known BirA-binding motifs, which are normally 14–16 bp long and include several highly conserved nucleotides (Rodionov *et al.*, 2002).

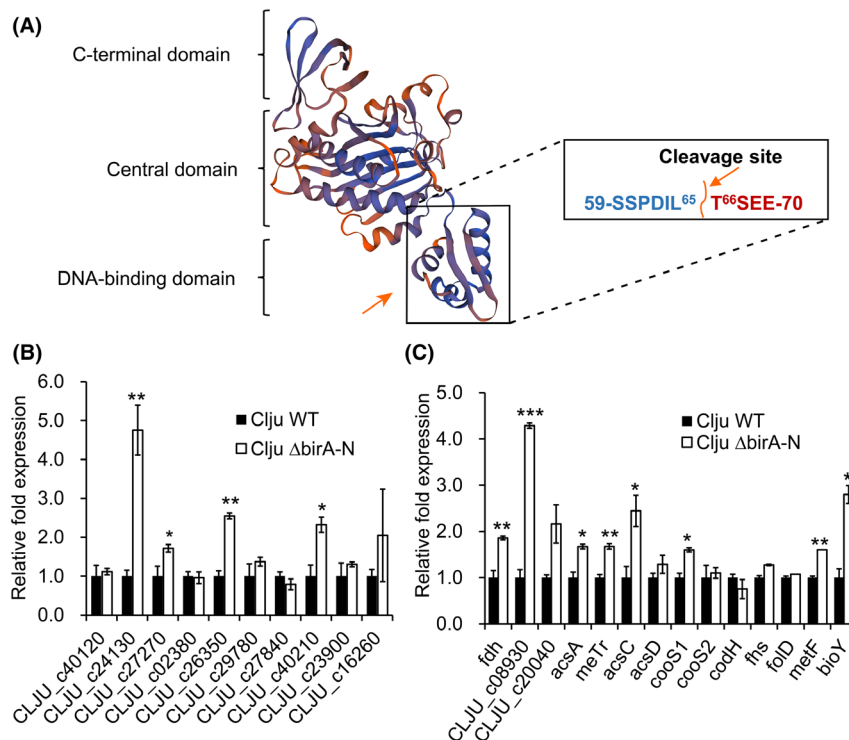


Fig. 3. Identification of the pleiotropic regulator role of C/BirA in *C. ljungdahliae*.

A. Homology modelling of C/BirA. A truncation of the N-terminal DNA-binding domain of C/BirA (for blocking the regulatory function of C/BirA) was also shown.

B. Transcriptional changes of the genes responsible for oxidation reduction and energy metabolism after the truncation of the N-terminal DNA-binding domain of C/BirA.

C. Transcriptional changes of the WLP genes after the truncation of the N-terminal DNA-binding domain of C/BirA. The cultivation of *C. ljungdahliae* cells for comparative transcriptomic analysis was performed in the modified ATCC medium. The data are presented as the means \pm standard deviations calculated from two independent experiments. Statistical analysis was performed by a two-tailed Student's *t*-test. **P* < 0.05, ***P* < 0.01, and ****P* < 0.001 vs. Clju WT.

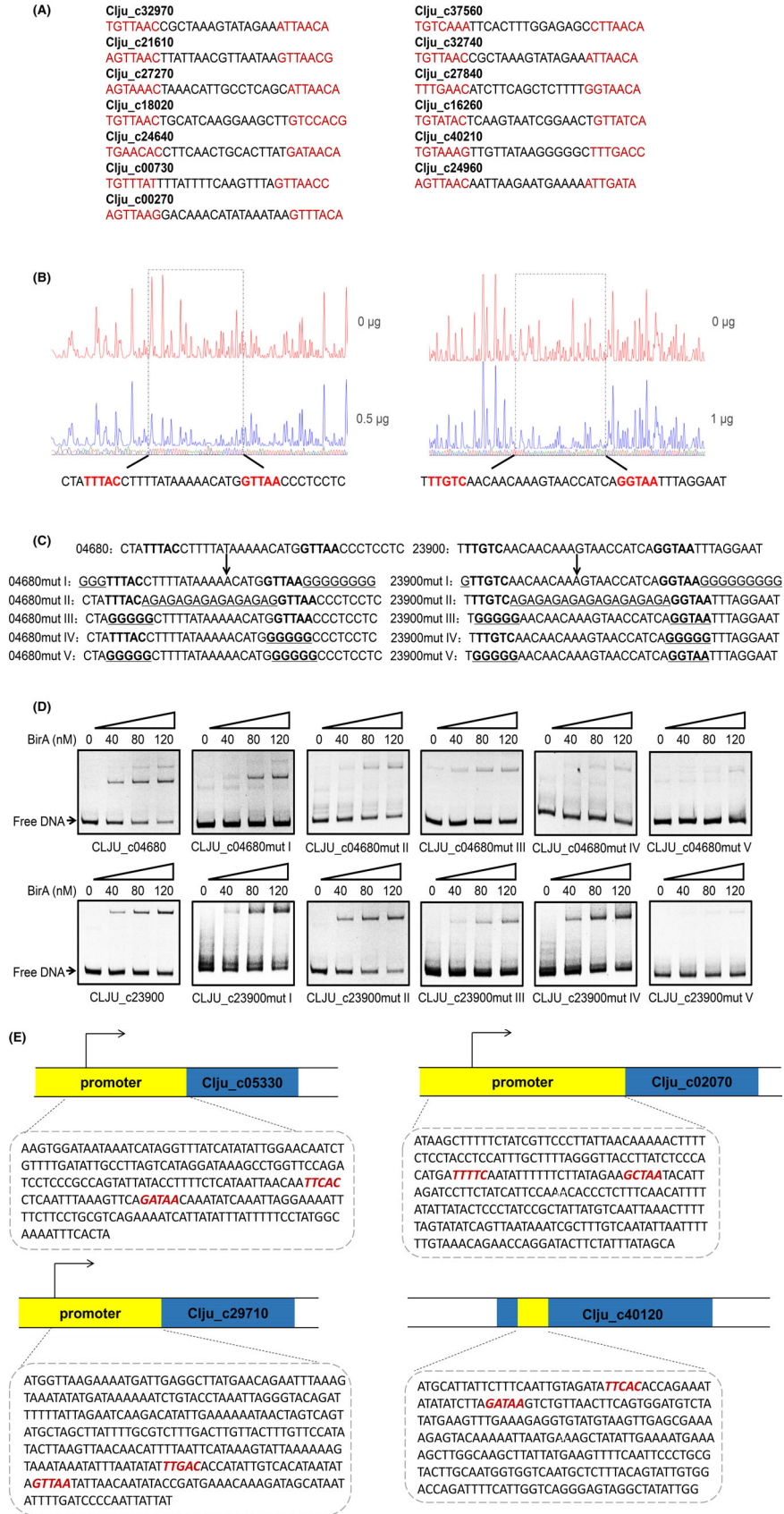
To further confirm the C/BirA binding to these two sites, we separately mutated their palindromic sequences, intervening spacers or flanking regions (Fig. 4D), and the yielding DNA fragments were used as probes for EMSAs. The results showed that the affinity of the probes for C/BirA was almost completely abolished with the mutation of the two palindromic sequences (04680 mut V and 23900 mut V), while the mutation of the other regions remained the affinity to different extents (Fig. 4D). These data verified that the 'TTTAC-N₁₆-GTTAA' and 'TTGTC-N₁₉-GGTAA' sequences are the C/BirA-binding sites, in which the palindromic sequences are crucial for the binding activity. Next, we searched for more C/BirA-binding sites within the remaining nine candidates (Clju_c02070, Clju_c05330, Clju_c15080, Clju_c16810, Clju_c24130, Clju_c26350, Clju_c29710, Clju_c29780 and Clju_c40120) (Fig. S2) via gene sequence alignment (SnapGene 2.3.2) using a template 'TTNWC-N_{16/19}-GNTAA' (N represents any bases; W represents A or T; the length of the intervening spacer was 16 or 19 nt). As

expected, one and three potential C/BirA-binding sites were found in the coding region and upstream non-coding region of their corresponding genes, respectively (Fig. 4E).

To investigate the universality of this atypical motif, we further searched the *C. ljungdahliae* genome via gene sequence alignment (SnapGene 2.3.2) using the template 'TTNAC-N_{16/19}-GNTAA'. The result revealed total 519 potential C/BirA-binding sites. This finding suggests that this new binding site architecture with a variable intervening spacer may play a broad role in C/BirA-mediated metabolic regulation.

Phenotypic changes of *C. ljungdahliae* caused by functional modulation of C/BirA

Since C/BirA belongs to the type-II BirA, which has both regulatory function and BPL activity, a derived question is how the two functions of C/BirA influence the performance of *C. ljungdahliae* in gas fermentation. To answer this question, we constructed two C/BirA variants, in



which the N-terminal DNA-binding domain and the enzymatically active domain were deleted and inactivated, respectively (Fig. 3A), for phenotypic analysis.

We first purified a *CBirA* variant, *CBirA^E*, lacking the N-terminal DNA-binding domain (from the 2nd to 65th amino acid residue). *CBirA^E* was then tested for its BPL activity. The results showed that the biotinylation degree of the substrate BCCP (the biotin carboxyl carrier protein that is the subunit *accB* of ACC) by *CBirA* and *CBirA^E* was 96.7% and 97.6%, respectively (Fig. 5A), indicating that the deletion of the N-terminal portion did not impair the BPL activity of *CBirA*. We further constructed the other *CBirA* variants without BPL activity by mutating essential amino acid residues in its enzymatically active domain. Based on a homology modelling using the *Pyrococcus horikoshii* OT3 BirA as the template (Bagautdinov *et al.*, 2008), nine potentially important amino acid residues (Pro¹⁷⁶, Asn¹⁷⁷, Asp¹⁷⁸, Leu¹⁹⁰, Glu¹⁹², Ile³¹⁶, Ser³¹⁷, Gly³¹⁸, Glu³¹⁹) of *CBirA* were picked out for individual or combined mutations (Fig. 5B), and the yielding mutants were used for BPL activity assays. As shown in Fig. 5C, the *CBirA^{N177A}*, *CBirA^{D178A}* and *CBirA^M* (triple mutations, P176A, N177A and D178A) mutants exhibited significantly decreased BPL activity compared to the WT *CBirA*, whereas the other two mutants, *CBirA^N* (double mutations, L190A, E192A) and *CBirA^P* (quadruple mutations, I316A, S317A, G318A, E319A), had no changes in BPL activity. Therefore, the *CBirA^{N177A}*, *CBirA^{D178A}* and *CBirA^M* mutants were adopted for DNA-binding activity assays (EMSAs) using the promoter region of the *bioY* gene (an identified direct target of *CBirA*, shown in Fig. S2) as the probe. The results showed that only *CBirA^{N177A}* still remained the binding activity to *bioY* (Fig. 5D), indicating that the N177A mutation did not impair the regulatory function of *CBirA^{N177A}*. Here, the *CBirA^{N177A}* protein was named *CBirA^R*.

Next, we began to examine the phenotypic changes of *C. ljungdahlii* caused by *CBirA^E* and *CBirA^R*. These two unfunctional *CBirA* variants were transferred into a *birA*-deleted *C. ljungdahlii* chassis (Clju Δ *birA*), solely or together, via an expression plasmid (Fig. 6A). The resulting three mutant strains, Clju Δ *birA*-*CBirA^E*, Clju Δ *birA*-*CBirA^R* and Clju Δ *birA*-*CBirA^{R+E}*, were compared to a control strain (Clju Δ *birA*-*CBirA^{WT}*, the Clju Δ *birA* strain containing a plasmid-expressed WT *CBirA*) to determine

their differences in autotrophic growth in gas fermentation. As shown in Fig. 6B, Clju Δ *birA*-*CBirA^E* and Clju Δ *birA*-*CBirA^{R+E}* exhibited increased and decreased growth rate, respectively, compared to Clju Δ *birA*-*CBirA^{WT}*, while Clju Δ *birA*-*CBirA^R* lost the growth capability. Obviously, the differences between the growth rate and biomass of Clju Δ *birA*-*CBirA^{R+E}* and Clju Δ *birA*-*CBirA^E* indicate a negative effect of the pleiotropic regulatory function of *CBirA* on *C. ljungdahlii*.

To better understand the improved autotrophic growth of Clju Δ *birA*-*CBirA^E* over Clju Δ *birA*-*CBirA^{WT}* (Fig. 6B), the regulatory property and BPL activity of *CBirA* in these two strains were compared. For the regulatory property analysis, the transcriptional levels of the aforementioned three *CBirA*'s target genes, *fdh*, Clju_c26350 and *bioY* (Fig. S2 and Fig. 2C), in Clju Δ *birA*-*CBirA^E* and Clju Δ *birA*-*CBirA^{WT}*, were examined. The results showed that these genes, compared to those in Clju Δ *birA*-*CBirA^{WT}*, were significantly up-regulated at most time points in Clju Δ *birA*-*CBirA^E* (Fig. 6C). Next, the activities of PYC and ACC, the two known biotinylated enzymes by BirA, in Clju Δ *birA*-*CBirA^E* and Clju Δ *birA*-*CBirA^{WT}*, were also compared. It was found that, compared to Clju Δ *birA*-*CBirA^{WT}*, Clju Δ *birA*-*CBirA^E* had much higher PYC activity at 48 h and higher ACC activity at 48 h and 72 h, while no obvious difference was observed at the other time points (Fig. 6D). Therefore, it seems likely that the deletion of the DNA-binding domain of *CBirA* (*CBirA^E*) destroys its transcriptional repression on target genes, and consequently, *CBirA^E* can fully function as a BPL to activate PYC and ACC by biotinylation, leading to the improved cell growth of *C. ljungdahlii* in gas fermentation (Fig. 6B).

Phylogenetic analysis on two types of BirAs in bacteria and archaea

As aforementioned, *CBirA* belongs to the type II BirA because it contains both the DNA-binding and enzymatically active domains (Satiaputra *et al.*, 2016); but its variant, *CBirA^E*, has lost the regulatory function and, thus, is more similar to the type I BirA. The different performance of *C. ljungdahlii* strains in gas fermentation caused by *CBirA^E* and *CBirA^R* (Fig. 6B) indicates the different roles of the type I and II BirAs in bacteria.

Fig. 4. Identification of the novel BirA-binding site consensus in *C. ljungdahlii*.

A. The genes associated with the typical conserved motif of BirA-binding sites.

B. DNase I footprinting analysis of the *CBirA* binding to Clju_c04680 and Clju_c23900.

C. Mutational analysis of the BirA-binding sites.

D. EMSAs of the *CBirA* binding with various mutated DNA fragments of Clju_c04680 and Clju_c23900.

E. The genes associated with the novel BirA-binding sites identified in this study. The sequences highlighted with red indicate the putative binding sites. Reactions were performed with 0.04 pmol of Cy5-labelled probes in the presence of different concentrations of BirA (40, 80 and 120 nM). Free probe is shown by an arrow.

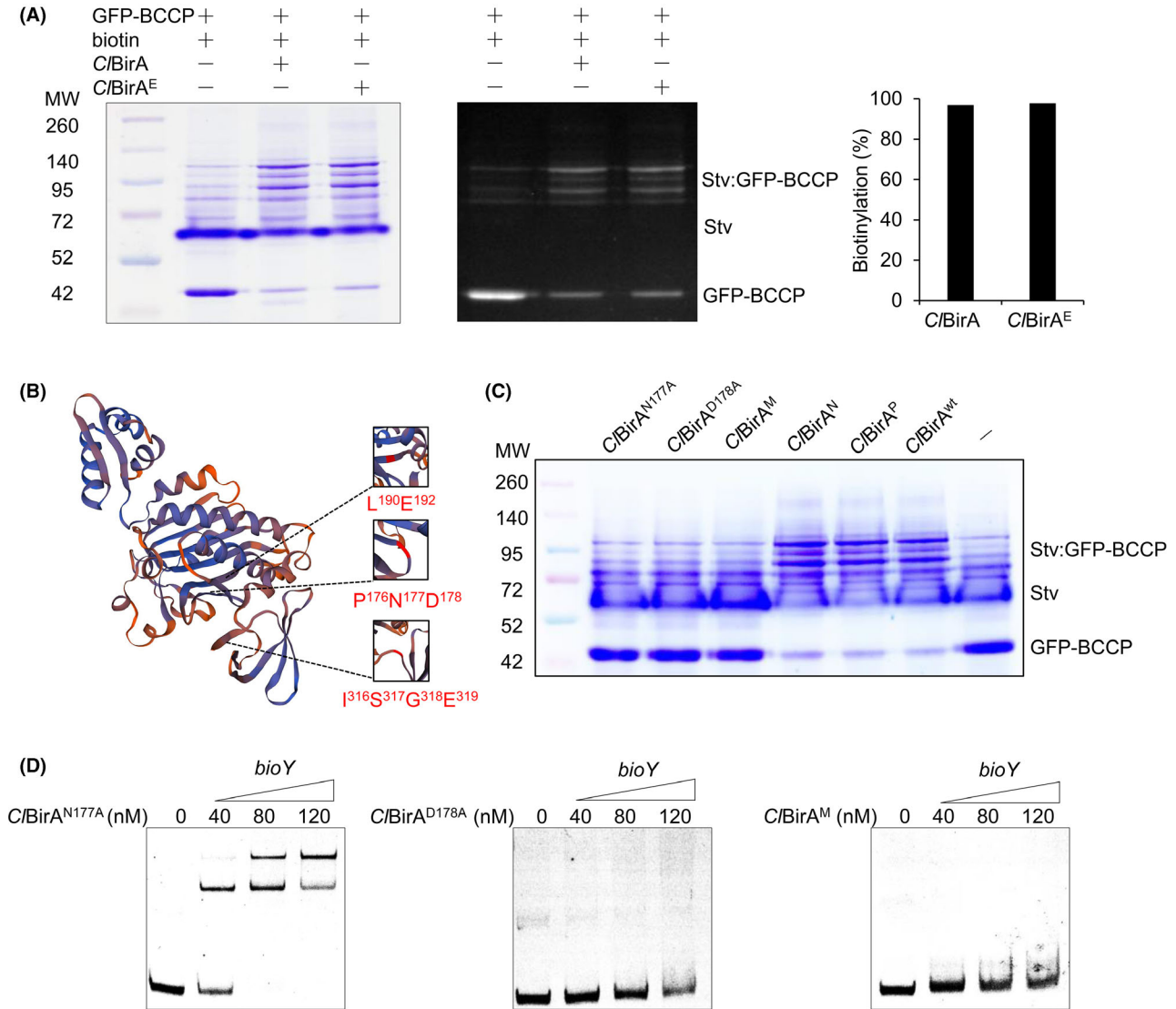


Fig. 5. Identification of the crucial amino acid residues for the BPL (biotin protein ligase) activity of C/BirA.

A. Detection of the biotinylation of GFP-BCCP by SDS-PAGE and Stv-induced band-shift analysis.

B. The predicted crucial amino acid residues for the BPL activity of C/BirA.

C. Detection of the GFP-BCCP biotinylation by C/BirA mutants using SDS-PAGE.

D. The comparison of the binding activities of C/BirA^{N177A}, C/BirA^{D178A} and C/BirA^M to the DNA probe of *bioY*. Biotinylation reactions were performed and processed as described in method. Control reactions were performed in the absence of the C/BirA protein. Fluorescent proteins are analysed by direct UV-irradiation of the gel and integration of GFP fluorescence.

Fig. 6. Dissection of the interaction between C/BirA's dual functions.

A. Construction of the *C. ljungdahlii* mutants, Clju Δ *birA*-C/BirA^{WT}, Clju Δ *birA*-C/BirA^E, Clju Δ *birA*-C/BirA^R and Clju Δ *birA*-C/BirA^{R+E}.

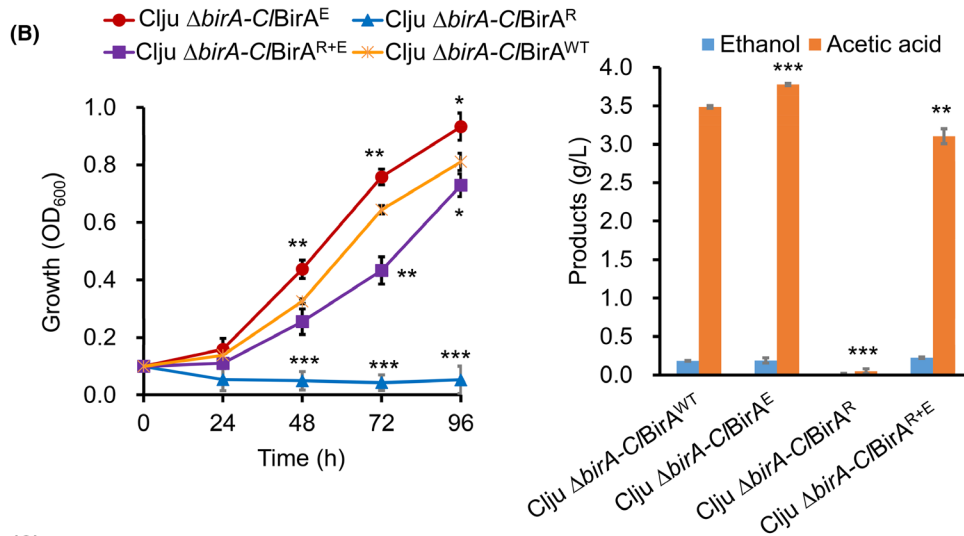
B. Comparison of the growth and product formation of the Clju Δ *birA*-C/BirA^{WT}, Clju Δ *birA*-C/BirA^E, Clju Δ *birA*-C/BirA^R and Clju Δ *birA*-C/BirA^{R+E} strains in gas fermentation.

C. Real-time qRT-PCR assay for detecting the transcriptional differences of the *fdh*, Clju_c26350, and *bioY* genes between Clju Δ *birA*-C/BirA^{WT} and Clju Δ *birA*-C/BirA^E.

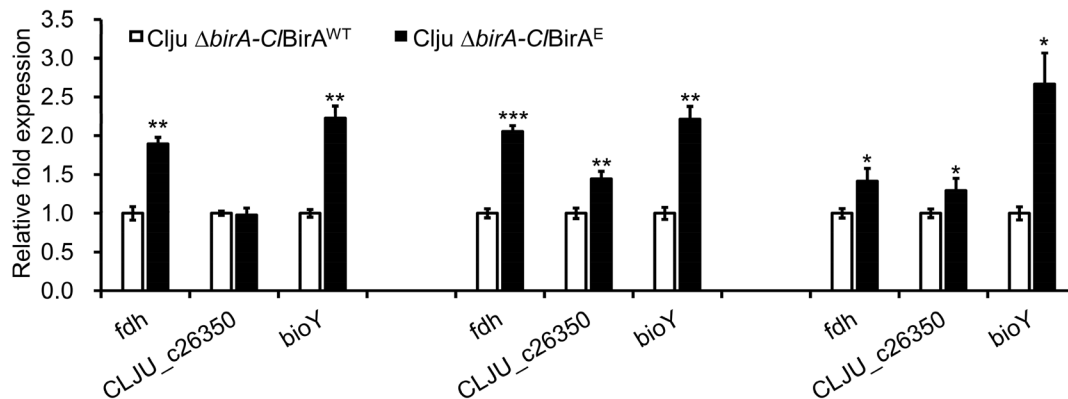
D. The *in vivo* PYC and ACC activities of Clju Δ *birA*-C/BirA^{WT} and Clju Δ *birA*-C/BirA^E strains. All the fermentations were performed in the modified ATCC medium 1754. Clju Δ *birA*-C/BirA^{WT}: the *birA*-deleted strain with a plasmid expressing the WT C/BirA; Clju Δ *birA*-C/BirA^E: the *birA*-deleted strain with a plasmid expressing C/BirA^E; Clju Δ *birA*-C/BirA^R: the *birA*-deleted strain with a plasmid expressing C/BirA^R; Clju Δ *birA*-C/BirA^{R+E}: the *birA*-deleted strain with a plasmid expressing both C/BirA^E and C/BirA^R. The data are presented as the means \pm standard deviations calculated from two independent experiments. Statistical analysis was performed by a two-tailed Student's *t*-test. **P* < 0.05, ***P* < 0.01, and ****P* < 0.001 vs. Clju Δ *birA*-C/BirA^{WT}.

(A)

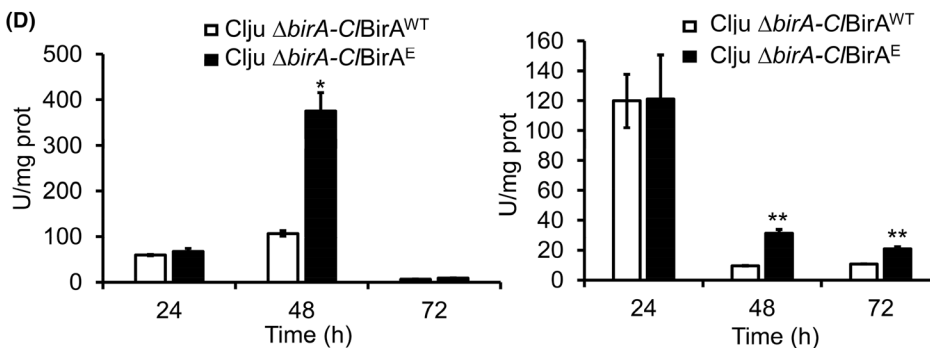
Entry	Recombinant <i>C. ljungdahliae</i> strains ($\Delta birA$)	Plasmids
1	Clju $\Delta birA$ -C/BirA ^{WT}	
2	Clju $\Delta birA$ -C/BirA ^E	
3	Clju $\Delta birA$ -C/BirA ^R	
4	Clju $\Delta birA$ -C/BirA ^{R+E}	

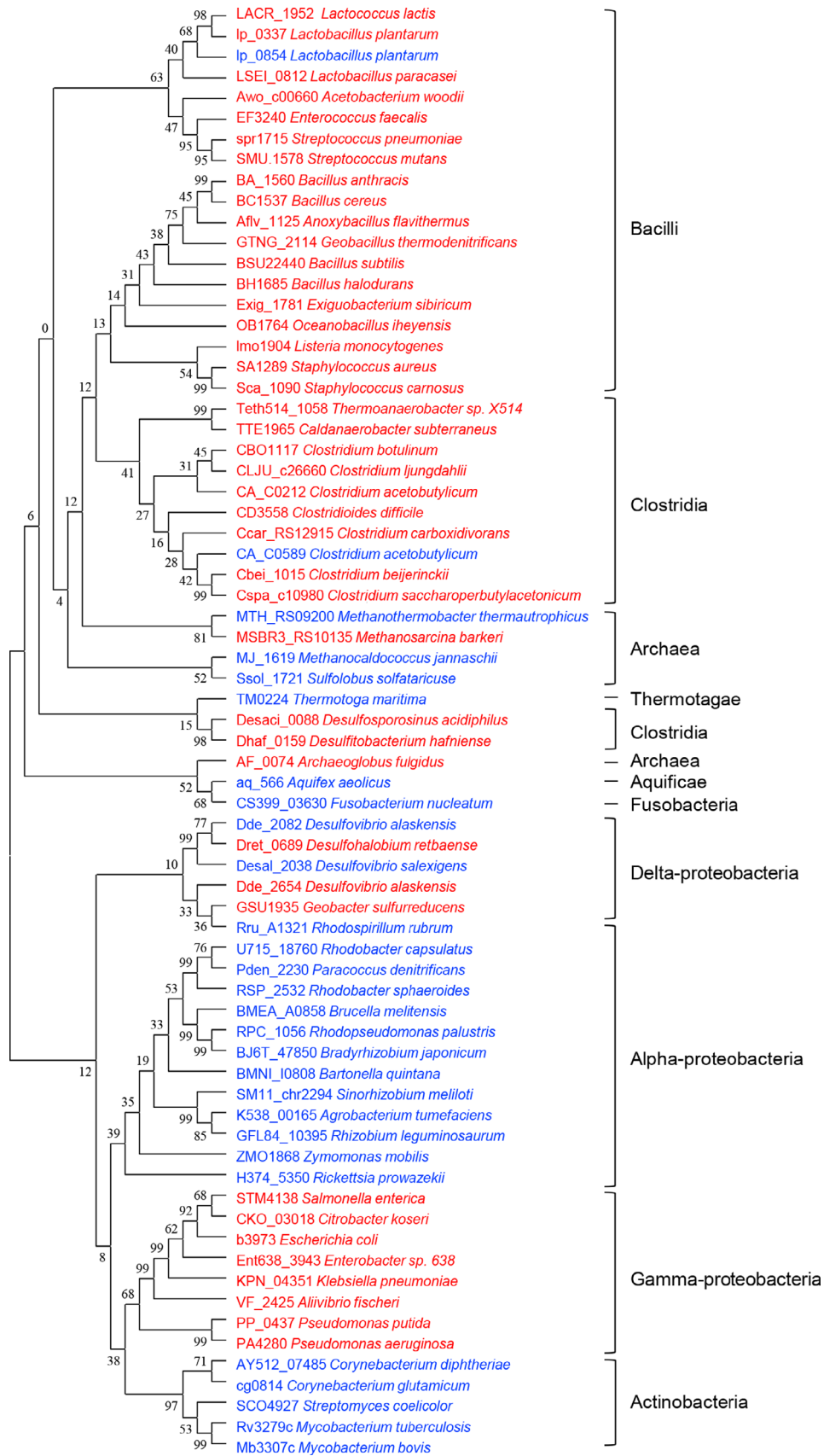


(C)



(D)





Here, we carried out a maximum likelihood phylogenetic analysis of 70 BirA proteins (28 type I and 42 type II) from 67 bacterial and archaeal strains that have complete genome data in KEGG database. Because type I BirA proteins lack the DNA-binding domain, the two conservative domains (BPL_ligase domain, PF03099; BPL_ligase C-terminal domain, PF02237), rather than the whole protein, of these BirAs were used for alignment. The BPL_ligase domains and BPL_ligase C-terminal domains from five BirAs that have determined structures in the PDB database (PDB number: 3FJP, 2EJ9, 4OP0, 1WNL, 2CGH) were picked out for the amino acid sequence alignment, and the result was used as an initial seed for a further alignment of the other BirA proteins. The results showed that most bacteria have only one type of BirA, in which the type II BirA often occurred in the class Bacilli, Clostridia and γ -Proteobacteria, while the type I BirA was normally found in the class α -Proteobacteria, Antibacteria, Thermotogae, Aquificae and Fusobacteria (Fig. 7). Moreover, the BirAs from same species are often clustered in the same clades and subclades (Fig. 7). Interestingly, the type I and type II BirAs were found to coexist in some bacterial hosts, including *Clostridium acetobutylicum* ATCC 824 (CA_C0589 and CA_C0212), *Lactococcus lactis* II1403 (L0191 and L0192) and *Francisella novicida* U112 (FTN_0811 and FTN_0568).

Function modulation of C/BirA improved autotrophic growth of *C. ljungdahlii*

The above data not only defined the pleiotropic functions of C/BirA but also provided clues for further strain improvement. It seems likely that a modified C/BirA with only enzymatically active domain is better for *C. ljungdahlii*, based on the advantage seen in the growth of Clju Δ birA-C/BirA^E over Clju Δ birA-C/BirA^{WT} in gas fermentation (Fig. 6B).

Thus, we overexpressed the C/BirA^E-coding gene (*clbirA^E*) in the aforementioned Clju Δ birA-N strain (chromosomal deletion of the N-terminal DNA-binding domain of C/BirA) (Fig. 8A); simultaneously, the *acc* (Clju_c42100-42140) and *pyc* (Clju_c37390) genes, coding for the *C. ljungdahlii* ACC and PYC enzymes, respectively, were also overexpressed separately (Fig. 8A). Such an engineering strategy was expected to specifically block the regulatory function, and simultaneously, reinforce the BPL activity of C/BirA in *C. ljungdahlii*. The yielding Clju Δ birA-N-p (*birA^E-acc*) and Clju

Δ birA-N-p (*birA^E-pyc*) strains were compared to the control strain (*C. ljungdahlii* Clju WT/p containing an empty plasmid) to determine their difference in growth and product synthesis in gas fermentation. As expected, the growth rate, biomass and final titer of acetic acid (a major product) of Clju Δ birA-N-p (*birA^E-acc*) were much higher than those of Clju WT/p (Fig. 8B and C); in contrast, no obvious phenotypic change was found for Clju Δ birA-N-p (*birA^E-pyc*) (Fig. 8B and C). Besides, the overexpression of the *acc* or *pyc* genes without *clbirA^E* in the Clju Δ birA-N chassis did not cause phenotypic changes (Fig. S4). Collectively, these findings strongly suggest that the reinforcement of the biotinylation of ACC contributes to the performance of *C. ljungdahlii* in gas fermentation.

Discussion

Autotrophic bacteria have evolved unique metabolic regulatory approaches for the fixation and conversion of C1 gases. The BirA protein, despite its physiological importance for heterotrophic bacteria, remains unexplored in autotrophic bacteria. This study has elucidated the regulatory scope and action mechanism of C/BirA in the autotrophic gas-fermenting *C. ljungdahlii*, providing new insights into the function of this protein in bacteria.

It has been known that biotin synthesis is tightly controlled by BirA in response to extra biotin supply and demand in bacteria (Feng *et al.*, 2014). When the intracellular biotin level exceeds the physiological requirement or biotin-requiring enzyme supply declines, BirA, together with biotin, will bind directly to the biotin operator and then repress biotin synthesis (Feng *et al.*, 2014). The biotin synthesis normally requires a large amount of energy in bacteria (Satiaputra *et al.*, 2016). In view of that biotin synthesis is a high energy-consuming process in cells (Feng *et al.*, 2013), such a BirA-mediated feedback regulation on biotin synthesis may be crucial for the control of energy expenditure in bacteria. Theoretically, this regulatory mechanism is especially important for gas-fermenting clostridia because the carbon fixation and assimilation in these autotrophic bacteria are a net energy consumption process (Zhang *et al.*, 2020a,2020b), and excessive energy consumption in biotin synthesis will cause metabolic burden. But interestingly, a modified C/BirA with only enzymatically active domain led to a better performance of *C. ljungdahlii* in gas fermentation (Figs 6 and 8), indicating that an uncontrolled overproduction of biotin is not the key point

Fig. 7. Maximum likelihood phylogenetic tree of BirA homologues in bacteria and archaea. The representative strains of Gram-positive/negative bacteria and archaea are included. Type I and II BirAs are marked in blue and red respectively.

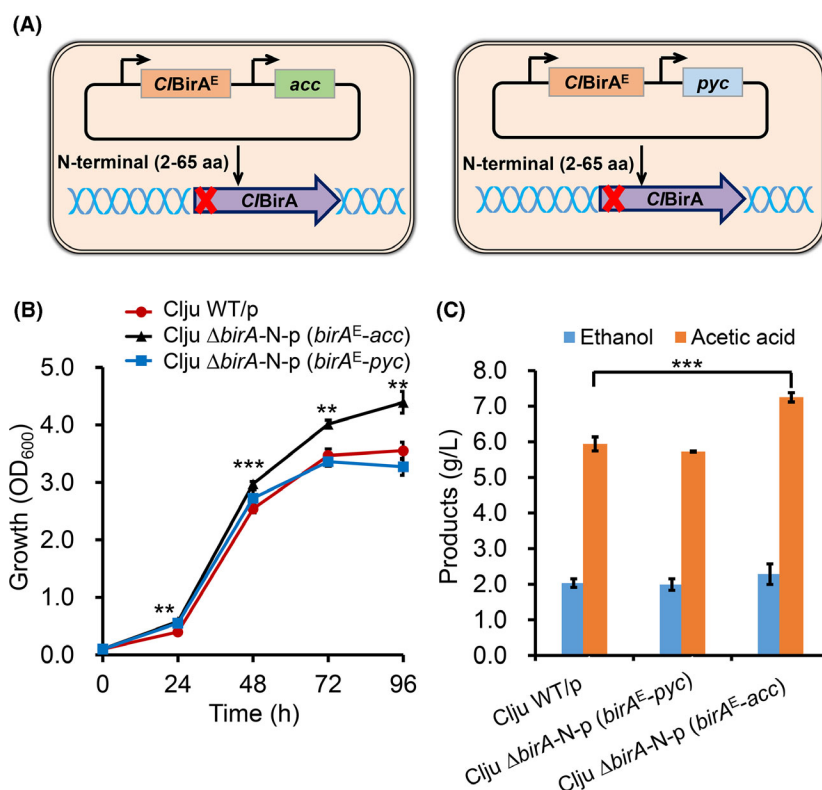


Fig. 8. Improved performance of *C. ljungdahlii* by the functional modulation of *C/BirA*.

A. Construction of the Clju $\Delta birA$ -N-p (*birA^E-acc*) and Clju $\Delta birA$ -N-p (*birA^E-pyc*) strains.

B. Comparison of the autotrophic growth of the Clju $\Delta birA$ -N-p (*birA^E-acc*), Clju $\Delta birA$ -N-p (*birA^E-pyc*) and Clju WT/p (control) strains in gas fermentation.

C. Comparison of the production of ethanol and acetic acid of the Clju $\Delta birA$ -N-p (*birA^E-acc*), Clju $\Delta birA$ -N-p (*birA^E-pyc*) and Clju WT/p strains in gas fermentation. All the fermentations were performed in the YT (YTF medium without fructose) medium with the supplement of a mixture of CO-CO₂-H₂-N₂. Clju $\Delta birA$ -N-p (*birA^E-acc*): the Clju $\Delta birA$ -N strain containing the pMTL83151-*birA^E-acc* plasmid. Clju $\Delta birA$ -N-p (*birA^E-pyc*): the Clju $\Delta birA$ -N strain containing the pMTL83151-*birA^E-pyc* plasmid. Data shown are means \pm standard deviations ($n = 3$) calculated from triplicate individual experiments. Error bars show standard deviations. Statistical analysis was performed by a two-tailed Student's *t*-test.

** $P < 0.01$ and *** $P < 0.001$ vs. Clju WT/p.

for the autotrophic growth of *C. ljungdahlii*. The molecular mechanism of this phenomenon remains to be explored. Taken together, these findings suggest that *C. ljungdahlii*, as a clostridial chassis for gas fermentation, needs to be further optimized in terms of the *in vivo* biotin supply.

Additionally, given the fact that *C/BirA* acts as a pleiotropic regulator in *C. ljungdahlii* (Figs 2 and 3), the above *C/BirA*-mediated feedback regulation in response to the *in vivo* biotin level is supposed to achieve a large-scale dynamic metabolic regulation. Of note, *C/BirA* seems to mainly play a repressor role because the specific deletion of the DNA-binding domain of *C/BirA* caused transcriptional up-regulation of many target genes, whereas only very few were down-regulated (Fig. 3B and C). This finding offered the possibility for the functional improvement of *C/BirA*. For example, because *C/BirA* has been found to directly control the expression of some WLP genes, the functional optimization of this protein may

contribute to the metabolism of C1 carbon sources (e.g. CO₂, CO and formate) via WLP in *C. ljungdahlii*. Such an engineering strategy, to our knowledge, has never been reported for strain improvement.

Another interesting finding in this study is that *C/BirA* employs a 'flexible' binding motif (TTNAC...N₁₆/19...GN-TAA) in its pleiotropic regulation. To date, most of the reported binding motifs of TFs are reverse palindromes or direct repeats, which are normally between 16 and 18 bp (Wang *et al.*, 2013; Antunes *et al.*, 2016; Bouillaut *et al.*, 2019). Only very few TFs have been found capable of recognizing such a 'flexible' binding structure. For example, the *E. coli* cyclic AMP (cAMP) receptor protein involved in regulating carbon metabolism recognizes the binding sites containing a 6- or 8-bp spacer (Barber *et al.*, 1993); the *Clostridium acetobutylicum* CcpA exerts its global regulation via binding the *cre_{var}* sites that consists of two inverted repeats separated by a highly variable intervening spacer (Yang *et al.*, 2017a, 2017b).

Such a variable binding structure is considered to create more flexibility to TFs' regulation (Yang *et al.*, 2017a,2017b). Therefore, it can be speculated that the binding motif 'TTNAC...N_{16/19}...GNTAA' will give the flexibility in both the regulatory scope and strength to C/BirA, thereby enabling diverse regulatory outputs.

In previous studies, researchers have proposed that BirA-like bifunctional proteins that possess both enzymatic and regulatory activities belong to an intermediate state between specific enzymes and transcription factors in evolution (Commichau and Stulke, 2008). For example, the PutA protein in α -Proteobacteria (e.g. *rhizobia*) that appeared relatively early in microbial evolution has only enzymatic activity; but this protein in γ -Proteobacteria (e.g. *Escherichia coli* and *Salmonella*) that appeared later has contained a ribbon-helix-helix (RHH) DNA-binding domain and thus can play a TF role (Krishnan and Becker, 2005; Liu *et al.*, 2017). Such an evolutionary process may also happen to BirA, namely that type I BirAs appeared relatively early and then became type II BirAs by capturing a DNA-binding domain during their evolutionary process. As mentioned above, some microbial strains were found to contain two BirA proteins. The study on the two BirAs in *Francisella novicida* U112 has revealed that one mainly acts as a BPL, while the other plays the regulatory function with very low BPL activity (Feng *et al.*, 2015). Therefore, it seems that functional differentiation has occurred in the BirA proteins in these strains. Additionally, the group II BirA can be found in multiple autotrophic bacteria, such as *C. ljungdahlii*, *Clostridium carboxidivorans* and *Acetobacterium woodii* (Fig. 7), indicating a broad regulatory role of BirA in autotrophs. Considering the pleiotropic regulation of BirA, it is highly possible that BirA controls some crucial metabolic pathways in autotrophic bacteria, thereby playing an important regulatory role in these bacteria.

In summary, our results elucidated the pleiotropic functions of the C/BirA protein in gas-fermenting *C. ljungdahlii*, which has renewed our understanding on the role of BirA in bacteria. More importantly, the functional modulation of C/BirA not only revealed the interaction between its dual functions but also achieved the genetic improvement of *C. ljungdahlii*, thereby showcasing the potential application value of this bifunctional protein in autotrophic bacteria.

Experimental procedures

Bacterial strains, plasmids, media and growth conditions

All strains and plasmids used in this work are listed in Table S2. *E. coli* Top10 was used for all cloning experiments, and BL21 (DE3) was used for protein expression. *E. coli* cells were grown in LB (lysogeny broth) medium (Bertani, 2004) supplemented with 12.5 $\mu\text{g ml}^{-1}$

chloramphenicol or 50 $\mu\text{g ml}^{-1}$ kanamycin antibiotics when needed. *C. ljungdahlii* DSM 13528 was grown anaerobically at 37°C in the YTF medium (Humphreys *et al.*, 2015) or a modified ATCC medium 1754 (Huang *et al.*, 2016), in which 5 $\mu\text{g ml}^{-1}$ of thiamphenicol was added for plasmid stability when needed.

Plasmid construction

The primers used in this work are listed in Table S3.

The pMTLcas-*birA* plasmid was constructed as follows: the pMTLcas-*pta* plasmid was digested with *SalI* and *XhoI*, yielding a linear pMTLcas vector. The sgRNA fragment targeting the *birA* gene was obtained from the pMTLcas-*pta* plasmid by PCR amplification using primers *birA*-gRNA-for/gRNA-rev. The two homologous arms (HAs) that flank the coding region of *birA* were obtained by PCR amplification using the *C. ljungdahlii* genomic DNA as the template and primers *birA*-up-for/*birA*-up-rev and *birA*-dn-for/*birA*-dn-rev and then linked to the sgRNA fragment via overlapping PCR using primers *birA*-gRNA-for/*birA*-dn-rev, yielding a DNA fragment, sgRNA-HA1. Finally, the linear pMTLcas vector and the sgRNA-HA1 fragment were assembled in one step using the ClonExpress MultiS One Step Cloning Kit (Vazyme Biotech, Nanjing, China), yielding the pMTLcas-*birA* plasmid for the deletion of *birA* in *C. ljungdahlii*.

The pMTLcas-*birA*-N plasmid was constructed as follows: The sgRNA fragment targeting the DNA-binding region (4–195 nt) of *birA* was obtained by PCR amplification using the pMTLcas-*pta* plasmid as the template and primers *birA*-gRNA-for/gRNA-rev. The two HAs that flank the DNA-binding region (4–195 nt) of *birA* were obtained by PCR amplification using the *C. ljungdahlii* genomic DNA as the template and primers *birA*-N-up-for/*birA*-N-up-rev and *birA*-N-dn-for/*birA*-N-dn-rev and then linked to the sgRNA fragment via overlapping PCR, yielding the sgRNA-HA2 fragment. Next, the aforementioned linear pMTLcas vector and the sgRNA-HA2 fragment were assembled in one step using the ClonExpress MultiS One Step Cloning Kit.

The pMTL83151-P₁₄₄₀-*birA*-3 \times FLAG plasmid for the ChIP-seq experiment was constructed as follows: the DNA fragments of the promoter P₁₄₄₀ and *birA* were obtained from the *C. ljungdahlii* genome by PCR amplification using primers 83151-P₁₄₄₀-o-for/P₁₄₄₀-*birA*-o-rev and *birA*-for/*birA*-FLAG-o-rev respectively. The 3 \times FLAG fragment was obtained by PCR amplification using primers FLAG-for/FLAG-83151-o-rev. The P₁₄₄₀-*birA*-3 \times FLAG fragment was obtained by overlapping PCR using the P₁₄₄₀, *birA*, 3 \times FLAG fragments as the templates and primers 83151-P₁₄₄₀-o-for/FLAG-83151-o-rev. Finally, the linear pMTL83151 plasmid and the P₁₄₄₀-*birA*-3 \times FLAG fragment were assembled in one step using the ClonExpress MultiS

One Step Cloning Kit, yielding the pMTL83151-P₁₄₄₀-birA-3×FLAG plasmid.

The construction of the pMTL83151-birA plasmid was performed as follows: the pMTL83151 plasmid was digested with *Nde*I and *Hind*III. The P_{birA}-birA fragment was obtained by PCR amplification using primers 83151-P_{birA}-o-for/birA-83151-o-rev and the *C. ljungdahlii* genomic DNA as the template. Next, the P-birA fragment and the linear pMTL83151 were assembled in one step using the ClonExpress MultiS One Step Cloning Kit, yielding the pMTL83151-birA plasmid. The pMTL83151-birA^{N177A} and pMTL83151-birA^E plasmids were constructed with the same step except primers used for PCR amplification (83151-P_{birA}-o-for/177-1-rev and 177-2-for/birA-83151-o-rev for pMTL83151-birA^{N177A}; P_{birA}-for/P_{birA}-rev and birA-N-for/birA-N-rev for pMTL83151-birA^E).

The pMTL83151-birA^{N177A}-birA^E plasmid was constructed as follows: the pMTL-8315-birA^{N177A} plasmid was digested with *Nhe*I and *Hind*III. The fragments of P_{birA} and birA^E were obtained from the *C. ljungdahlii* genomic DNA by PCR amplification using primers P_{birA}-for/P_{birA}-rev and birA-N-for/birA-N-rev, respectively, and then linked via overlapping PCR using primers P_{birA}-for/birA-N-rev, yielding the P_{birA}-birA^E fragment. The P_{birA}-birA^E fragment and the linear pMTL83151-birA^{N177A} were assembled in one step using the ClonExpress MultiS One Step Cloning Kit.

The pMTL83151-birA^E-acc plasmid was constructed as follows: the pMTL83151-birA^E plasmid was digested with *Nhe*I and *Hind*III. The DNA fragments of P₁₄₄₀ and acc were obtained by PCR amplifications using the *C. ljungdahlii* genomic DNA as the template and primers birA^E-P₁₄₄₀-o-fw/P₁₄₄₀-rev and acc-for/acc-83151-o-rev respectively. Finally, the linear pMTL83151-birA^E was assembled with the DNA fragments of P₁₄₄₀ and acc in one step using the ClonExpress MultiS One Step Cloning Kit. The pMTL83151-birA^E-pyc plasmid was constructed with the same steps except that primers pyc-for/pyc-83151-o-rev were used.

The pUCm-T-04680 plasmid was constructed as follows: the DNA fragment of the Clju_c04680 gene was obtained by PCR amplification using the *C. ljungdahlii* genomic DNA as the template and primers E-04680-fw/E-04680-rev and then linked to pUCm-T using TaqDNA Polymerase, yielding the pUCm-T-04680 plasmid. The pUCm-T-23900 plasmid was constructed with the same steps, except that primers E-23900-fw/E-23900-rev were used for PCR amplification.

Gene expression and deletion in *C. ljungdahlii*

The plasmids were delivered into *C. ljungdahlii* by electroporation. The preparation of electrocompetent cells and electroporation procedure were same as previously

described (Huang *et al.*, 2016). Gene deletion in *C. ljungdahlii* was performed using the CRISPR-Cas9-based genome editing method, and the procedure was the same as previously described (Huang *et al.*, 2016).

Fermentation

Inoculum preparation of the wild type *C. ljungdahlii* strain and its derivatives were performed anaerobically in the YTF medium. Fermentations were performed anaerobically in the modified ATCC medium 1754 or YT (YTF medium without fructose), in which 5 µg ml⁻¹ of thiamphenicol was added when needed. Briefly, 300 µl frozen stock was inoculated into 4 ml liquid YTF medium and incubated anaerobically at 37°C for 24 h. When the optical density (OD₆₀₀) of grown cells reached 1.0–1.2, 1.5 ml of the grown cells was transferred into 30 ml of the aforementioned modified ATCC medium 1754 or YT medium for gas fermentation. The fermentation was carried out in 125 ml Wheaton serum bottles (Sigma-Aldrich, Saint Louis, MO, USA) with a headspace of CO₂-H₂-N₂ (56%/20%/9%/15%; pressurized to 0.2 MPa). The gases were added into the headspace (again at 0.2 MPa) for every 24 h.

Protein purification

The genes coding for CBirA (Clju_c26660) and the derived CBirA mutants, including CBirA^E, CBirA^{N177A}, CBirA^{D178A}, CBirA^M, CBirA^N, CBirA^P and BCCP (Clju_c42140), were obtained by PCR amplification using corresponding primers (Table S3) and the *C. ljungdahlii* genomic DNA as the template. The DNA fragment of *gfp* was obtained by PCR amplification using primers *gfp*-o-pET-28a-fw/*gfp*-o-bccp-rev and the pIMP1-thi9-gfp plasmid as the template (Yang *et al.*, 2017a, 2017b). These genes were inserted into the pET-28a vector (Invitrogen, Carlsbad, CA, USA), and the yielding plasmids were transferred into *E. coli* BL21 (DE3) for expression. The subsequent protein purification procedure was same as previously described (Ren *et al.*, 2012).

ChIP-Seq library construction and sequencing

The pMTL83151-P₁₄₄₀-birA-3×FLAG plasmid carrying the 3×FLAG-tagged birA was transformed into the Clju ΔbirA strain. Cells were cross-linked with 1% formaldehyde for 10 min at room temperature. The reaction was quenched by adding glycine to a final concentration of 0.125 M. Formaldehyde cross-linked cells were washed with pre-chilled phosphate buffer saline and then sonicated to generate 200–500 bp DNA fragments. Immunoprecipitation was performed using monoclonal anti-3×FLAG antibody (Cat#F1804; Sigma) and protein A

magnetic beads (Cat#10002D; Invitrogen) according to the manufacturer's instructions. Sonicated extracts, without the addition of antibodies, was performed as a negative control. Proteins were then removed by incubation with proteinase K for 2 h at 55°C. The libraries were prepared using the VAHTS Universal DNA Library Prep Kit for Illumina V3 (Catalog NO. ND607; Vazyme) and sequenced using the Novaseq 6000 system (Illumina, San Diego, CA, USA).

Electrophoretic mobility shift assay (EMSA)

The DNA probes for EMSAs were prepared by two-step PCR amplification. First, the unlabelled DNA fragments were obtained by PCR amplification using the *C. ljungdahlii* genomic DNA as the template and primers E-fw/E-rev that contained a universal sequence (5'-AGCCAGTGGCGATAAG-3') at the 5' terminus. Next, the labelled DNA fragments were generated by PCR amplification using a 5'-Cy5-labelled universal primer (5'-AGCCAGTGGCGATAAG-3') and then recovered by agarose gel electrophoresis using the PCR purification Kit (Cat#AP-GX-250; Axygen, Hangzhou, China). The obtained DNA fragments were used as probes for EMSAs, which were carried out as previously described (Ren *et al.*, 2012).

RT-qPCR

The WT *C. ljungdahlii* strain and the derived mutants were grown anaerobically at 37°C in the modified ATCC medium 1754 using syngas (CO-CO₂-H₂-N₂). The grown cells were harvested by centrifugation at 8600 *g* for 10 min. Total RNA was extracted using the kit (Cat#cw0581; CWBIO, Beijing, China) according to the manufacturer's instructions. Contaminating DNA in the extracted RNA was eliminated by DNase I (TaKaRa, Kyoto, Japan) digestion. The RNA concentration was determined by a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). cDNA was obtained by reverse transcription using the PrimeScript RT reagent kit (TaKaRa). The subsequent procedure of qPCR was the same as described previously (Zhang *et al.*, 2018). The *rho* gene (Clju_c02220) was used as the internal control according to the previous report (Zhao *et al.*, 2019).

Biotinylation activity assay

The biotinylation activity of *CBirA* was measured as previously described (Sorenson *et al.*, 2015), using BCCP (an biotin carboxyl carrier protein that is the acetyl-coenzyme A carboxylase subunit *accB*) as the substrate. Biotinylation of GFP-BCCP was performed in 50 μ l

reactions containing 0.3 μ M BirA protein, 0.5 mM biotin, 2.5 mM ATP, 1 mM MgCl₂, 30 μ M GFP-BCCP and 25 mM Tris (pH 8.0). The reaction mixtures were incubated for 1 h at room temperature. After incubation, free biotin was removed from the reactions by Ni-affinity purification using a Promity IMAC nickel resin column (Cat#732-6008; Bio-Rad, Hercules, CA, USA). The resin was then washed twice with 500 μ l of 25 mM Tris (pH 8.0) supplemented with 5% glycerol and 150 mM NaCl by centrifugation at 1000 *g* for 1 min. Fusion proteins were eluted from the column with 50 μ l of the same buffer supplemented with 200 mM imidazole and centrifugation at 1000 *g* for 1 min. Next, 7.5 μ l of protein elution and 0.75 μ l Stv (streptavidin, 5 mg ml⁻¹, Invitrogen) were incubated at room temperature for 15 min. The reactions were subjected to SDS-PAGE without a heat-denaturation step to visualize the protein complexes. Gels were first photographed under a UV transilluminator before Coomassie staining. The fluorescence of GFP-BCCP was analyzed by integration of protein bands (ImageJ; NIH, Bethesda, MD, USA) to determine the extent of biotinylation. Control reactions (without *CBirA*) were performed in parallel.

DNase I footprinting assays

To prepare the probes labelled with fluorescent FAM (6-carboxyfluorescein), the noncoding regions upstream of Clju_c04680 and Clju_c23900 were amplified by PCR from the pUCm-T-04680 and pUCm-T-23900 plasmids, respectively, using primers M13F/M13R(FAM). The FAM-labelled probes were purified and quantified with NanoDrop 2000C (Thermo).

The DNase I footprinting assay was performed as follows. First, 250 ng probes were incubated with indicated amounts of purified *CBirA* proteins in 40 μ l of binding buffer containing 20 mM Tris-HCl (pH 7.9), 1 mM DTT, 10 mM MgCl₂, 0.5 mg ml⁻¹ calf BSA and 5% (vol./vol.) glycerol. After incubation at 25°C for 30 min, 10 μ l of RNase-free DNase I buffer, 0.015 unit of DNase I (Promega, Madison, WI, USA) and 100 nmol of fresh CaCl₂ were added to digest the DNA probe at 37°C for 1 min. The reaction was stopped by adding 140 μ l stop solution (200 mM sodium acetate, 30 mM EDTA and 0.15% SDS). The DNA was recovered by phenol/chloroform extraction and ethanol precipitation and then dissolved in 30 μ l MiniQ water. Finally, 1 μ l of sample was added to 8.5 μ l of HiDi formamide and 0.5 μ l of GeneScan-LIZ600 size standards (Applied Biosystems, Foster City, CA, USA) was sequenced on a 3130xl DNA analyzer (Applied Biosystems). The preparation of the DNA ladder and data analysis was same as previously described (Wang *et al.*, 2012). Genescan results were evaluated with Peak Scanner software v1.0 (Applied Biosystems).

Pyruvate carboxylase and acetyl-CoA carboxylase activity assays

The *C. ljungdahlii* strains were grown in the modified ATCC medium 1754 with the supplement of a mixture of CO₂-CO₂-H₂-N₂. When the the OD₆₀₀ of grown cells reached 1.0, the cell pellets were harvested by centrifugation (7000 g, 4°C, 10 min), dissolved in the reagents from the following assay kits (as shown below) and vortexed for 1 min for cell lysis. Finally, the cell lysate was centrifuged at 12 000 g for 30 min, and the supernatant was used for PYC and ACC activity assays. The preparation of cell lysis was performed under aerobic conditions.

The PC assay kit (Cat#BC0730; Solarbio, Beijing, China) was used to measure the activity of PYC. The PYC activity was measured following the decrease in absorbance at 340 nm induced by oxidized oxaloacetate (the product from the carboxylation of pyruvate catalyzed by PYC) in the presence of NADH.

The ACC assay kit (Cat#BC0410; Solarbio) was used to measure the activity of ACC. The determination of the ACC was based on the reaction of acetyl-CoA being catalyzed to malonyl-CoA and inorganic phosphorus. The inorganic phosphorus was further determined by a colour reaction with molybdenum blue, and the absorbance was read at 660 nm. Standard curve was generated by using known amounts of phosphate in combination with the molybdenum blue reagent as described in the manufacture's instructions.

Phylogenetic tree construction

Gene annotations were derived from SEED (Overbeek *et al.*, 2005) and KEGG (Kanehisa *et al.*, 2012) database. Comparative genomic analysis was performed to predict genes with unknown functions by using GenomeExplorer (Mironov *et al.*, 2000). Multiple protein alignments were done using MUSCLE (Edgar, 2004) and MUFFT (Kato and Standley, 2013). The phylogenetic tree was constructed via the maximum likelihood method implemented in PhyML (Guindon *et al.*, 2010) and MEGA (Tamura *et al.*, 2011) and further edited with iTOL (Letunic and Bork, 2011).

Acknowledgements

This work was supported by the National Key R&D Program of China (2018YFA0901500), the National Natural Science Foundation of China (31970067, 31921006) and Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-KJGG-016).

Funding Information

This work was supported by the National Key R&D Program of China (2018YFA0901500), the National Natural

Science Foundation of China (31970067, 31921006) and Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-KJGG-016).

Conflict of interest

The authors declare no competing financial interest.

References

- Aklujkar, M., Leang, C., Shrestha, P.M., Shrestha, M., and Lovley, D.R. (2017) Transcriptomic profiles of *Clostridium ljungdahlii* during lithotrophic growth with syngas or H₂ and CO₂ compared to organotrophic growth with fructose. *Sci Rep-Uk* **7**: 13135.
- Antunes, A., Golfieri, G., Ferlicca, F., Giuliani, M.M., Scarlato, V., and Delany, I. (2016) HexR controls glucose-responsive genes and central carbon metabolism in *Neisseria meningitidis*. *J Bacteriol* **198**: 644–654.
- Bagautdinov, B., Matsuura, Y., Bagautdinova, S., and Kunishima, N. (2008) Protein biotinylation visualized by a complex structure of biotin protein ligase with a substrate. *J Biol Chem* **283**: 14739–14750.
- Barber, A.M., Zhurkin, V.B., and Adhya, S. (1993) CRP-binding sites: evidence for two structural classes with 6-bp and 8-bp spacers. *Gene* **130**: 1–8.
- Berg, I.A., Kockelkorn, D., Ramos-Vera, W.H., Say, R.F., Zarzycki, J., Hügler, M., *et al.* (2010) Autotrophic carbon fixation in archaea. *Nat Rev Microbiol* **8**: 447–460.
- Bertani, G. (2004) Lysogeny at mid-twentieth century: P1, P2, and other experimental, systems. *J Bacteriol* **186**: 595–600.
- Bouillaut, L., Dubois, T., Francis, M.B., Daou, N., Monot, M., Sorg, J.A., *et al.* (2019) Role of the global regulator Rex in control of NAD(+)-regeneration in *Clostridioides (Clostridium) difficile*. *Mol Microbiol* **111**: 1671–1688.
- Commichau, F.M., and Stulke, J. (2008) Trigger enzymes: bifunctional proteins active in metabolism and in controlling gene expression. *Mol Microbiol* **67**: 692–702.
- Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 1–19.
- Feng, Y.J., Chin, C.Y., Chakravarty, V., Gao, R.S., Crispell, E.K., Weiss, D.S., and Cronan, J.E. (2015) The atypical occurrence of two biotin protein ligases in *Francisella novicida* is due to distinct roles in virulence and biotin metabolism. *MBio* **6**: e00591–e615.
- Feng, Y.J., Napier, B.A., Manandhar, M., Henke, S.K., Weiss, D.S., and Cronan, J.E. (2014) A *Francisella* virulence factor catalyses an essential reaction of biotin synthesis. *Mol Microbiol* **91**: 300–314.
- Feng, Y.J., Zhang, H.M., and Cronan, J.E. (2013) Profliigate biotin synthesis in -proteobacteria a developing or degenerating regulatory system? *Mol Microbiol* **88**: 77–92.
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**: 307–321.

- Huang, H., Chai, C.S., Li, N., Rowe, P., Minton, N.P., Yang, S., *et al.* (2016) CRISPR/Cas9-based efficient genome editing in *Clostridium ljungdahlii*, an autotrophic gas-fermenting bacterium. *Acs Synth Biol* **5**: 1355–1361.
- Humphreys, C.M., McLean, S., Schatschneider, S., Millat, T., Henstra, A.M., Annan, F.J., *et al.* (2015) Whole genome sequence and manual annotation of *Clostridium autoethanogenum*, an industrially relevant bacterium. *BMC Genom* **16**: 1085.
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* **40**: D109–D114.
- Katoh, K., and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772–780.
- Krishnan, N., and Becker, D.F. (2005) Characterization of a bifunctional PutA homologue from *Bradyrhizobium japonicum* and identification of an active site residue that modulates proline reduction of the flavin adenine dinucleotide cofactor. *Biochemistry* **44**: 9130–9139.
- Lemgruber, R.D.P., Valgepea, K., Garcia, R.A.G., de Bakker, C., Palfreyman, R.W., Tappel, R., *et al.* (2019) A TetR-family protein (CAETHG_0459) activates transcription from a new promoter motif associated with essential genes for autotrophic growth in acetogens. *Front Microbiol* **10**: 2549.
- Letunic, I., and Bork, P. (2011) Interactive tree of life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* **39**: W475–W478.
- Li, Y.F., and Sousa, R. (2012) Expression and purification of *E. coli* BirA biotin ligase for in vitro biotinylation. *Protein Express Purif* **82**: 162–167.
- Liu, L.K., Becker, D.F., and Tanner, J.J. (2017) Structure, function, and mechanism of proline utilization A (PutA). *Arch Biochem Biophys* **632**: 142–157.
- Mironov, A.A., Vinokurova, N.P., and Gelfand, M.S. (2000) Software for analysis of bacterial genomes. *Mol Biol+* **34**: 222–231.
- Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.Y., Cohoon, M., *et al.* (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* **33**: 5691–5702.
- Peters-Wendisch, P., Stansen, K.C., Gotker, S., and Wendisch, V.F. (2012) Biotin protein ligase from *Corynebacterium glutamicum*: role for growth and L-lysine production. *Appl Microbiol Biot* **93**: 2493–2502.
- Ren, C., Gu, Y., Wu, Y., Zhang, W.W., Yang, C., Yang, S., and Jiang, W.H. (2012) Pleiotropic functions of catabolite control protein CcpA in butanol-producing *Clostridium acetobutylicum*. *BMC Genom* **13**: 349.
- Rodionov, D.A., Mironov, A.A., and Gelfand, M.S. (2002) Conservation of the biotin regulon and the BirA regulatory signal in eubacteria and archaea. *Genome Res* **12**: 1507–1516.
- Satiaputra, J., Shearwin, K.E., Booker, G.W., and Polyak, S.W. (2016) Mechanisms of biotin-regulated gene expression in microbes. *Syn Syst Biotechnol* **1**: 17–24.
- Soares da Costa, T.P., Tieu, W., Yap, M.Y., Pendini, N.R., Polyak, S.W., Sejer Pedersen, D., *et al.* (2012) Selective inhibition of biotin protein ligase from *Staphylococcus aureus*. *J Biol Chem* **287**: 17823–17832.
- Sorenson, A.E., Askin, S.P., and Schaeffer, P.M. (2015) In-gel detection of biotin–protein conjugates with a green fluorescent streptavidin probe. *Anal Methods* **7**: 2087–2092.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.
- Tan, Y., Liu, J.J., Chen, X.H., Zheng, H.J., and Li, F.L. (2013) RNA-seq-based comparative transcriptome analysis of the syngas-utilizing bacterium *Clostridium ljungdahlii* DSM 13528 grown autotrophically and heterotrophically. *Mol Biosyst* **9**: 2775–2784.
- Wang, R., Mast, Y., Wang, J., Zhang, W.W., Zhao, G.P., Wohlleben, W., *et al.* (2013) Identification of two-component system AfsQ1/Q2 regulon and its cross-regulation with GlnR in *Streptomyces coelicolor*. *Mol Microbiol* **87**: 30–48.
- Wang, Y., Cen, X.F., Zhao, G.P., and Wang, J. (2012) Characterization of a new GlnR binding box in the promoter of *amtB* in *Streptomyces coelicolor* inferred a PhoP/GlnR competitive binding mechanism for transcriptional regulation of *amtB*. *J Bacteriol* **194**: 5237–5244.
- Weaver, L.H., Kwon, K., Beckett, D., and Matthews, B.W. (2001) Corepressor-induced organization and assembly of the biotin repressor: a model for allosteric activation of a transcriptional regulator. *Proc Natl Acad Sci USA* **98**: 6045–6050.
- Yang, B., Feng, L., Wang, F., and Wang, L. (2015) *Enterohemorrhagic Escherichia coli* senses low biotin status in the large intestine for colonization and infection. *Nat Commun* **6**: 1–12.
- Yang, G.H., Jia, D.C., Jin, L., Jiang, Y.Q., Wang, Y., Jiang, W.H., and Gu, Y. (2017) Rapid generation of universal synthetic promoters for controlled gene expression in both gas-fermenting and saccharolytic *Clostridium* species. *Acs Synth Biol* **6**: 1672–1678.
- Yang, Y.P., Zhang, L., Huang, H., Yang, C., Yang, S., Gu, Y., and Jiang, W.H. (2017) A flexible binding site architecture provides new insights into CcpA global regulation in Gram-positive bacteria. *MBio* **8**: e02004–e2016.
- Ye, H.Y., Cai, M.Z., Zhang, H.M., Li, Z.C., Wen, R.H., and Feng, Y.J. (2016) Functional definition of BirA suggests a biotin utilization pathway in the zoonotic pathogen *Streptococcus suis*. *Sci Rep-Uk* **6**: 26479.
- Zhang, H., Wang, Q., Fisher, D.J., Cai, M., Chakravarty, V., Ye, H., *et al.* (2016) Deciphering a unique biotin scavenging pathway with redundant genes in the probiotic bacterium *Lactococcus lactis*. *Sci Rep-Uk* **6**: 25680.
- Zhang, L., Liu, Y.Q., Yang, Y.P., Jiang, W.H., and Gu, Y. (2018) A novel dual-cre motif enables two-way autoregulation of CcpA in *Clostridium acetobutylicum*. *Appl Environ Microb* **84**: e00114–e118.
- Zhang, L., Liu, Y.Q., Zhao, R., Zhang, C., Jiang, W.H., and Gu, Y. (2020) Interactive regulation of formate dehydrogenase during CO₂ fixation in gas-fermenting bacteria. *MBio* **11**: e00650–e720.
- Zhang, L., Zhao, R., Jia, D.C., Jiang, W.H., and Gu, Y. (2020) Engineering *Clostridium ljungdahlii* as the

gas-fermenting cell factory for the production of biofuels and biochemicals. *Curr Opin Chem Biol* **59**: 54–61.

Zhao, R., Liu, Y.Q., Zhang, H., Chai, C.S., Wang, J., Jiang, W.H., and Gu, Y. (2019) CRISPR-Cas12a-mediated gene deletion and regulation in *Clostridium ljungdahlii* and its application in carbon flux redirection in synthesis gas fermentation. *Acs Synth Biol* **8**: 2270–2279.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. The influence of the deletion of the *C*BirA-encoding gene (*birA*) on the growth and product formation of *C. ljungdahlii* in fermenting fructose. (A) The effect of the *birA* deletion on the growth of *C. ljungdahlii* grown in the medium supplemented with biotin. (B) The effect of the *birA* deletion on product formation of *C. ljungdahlii* grown in the medium supplemented with biotin. All the fermentations were performed in the modified ATCC medium 1754. Clju WT/p: the wild-type strain containing an empty plasmid. Clju $\Delta birA/p$: the *birA*-deleted strain containing an empty plasmid. Clju $\Delta birA/pbirA$: the *birA*-deleted strain complemented with the plasmid-carried *birA* expression. Data shown are means \pm standard deviations ($n = 3$) calculated from triplicate

individual experiments. Error bars show standard deviations. Statistical analysis was performed by a two-tailed Student's *t*-test. *** $P < 0.001$ vs. the Clju WT/p strain.

Fig. S2. EMSAs for determining the *C*BirA binding to the 24 potential target genes provided by ChIP-seq data. The coding sequence regions or promoter regions of genes were used as the DNA probes for EMSAs. The concentration of the BirA protein used in the experiment was 80 nM.

Fig. S3. EMSAs for determining the *C*BirA^E (lacking the N-terminal DNA-binding domain) binding to *bioY*. The promoter region of the *bioY* gene was used as the DNA probe. Reactions were performed with 0.04 pmol of Cy5-labelled probes in the presence of different concentrations of *C*BirA^E (40, 80, 120 nM).

Fig. S4. The influence of the overexpression of the genes coding for ACC (acetyl-CoA carboxylase) and PYC (pyruvate carboxylase) on the growth of *C. ljungdahlii* in gas fermentation. All the fermentations were performed in the YT (YTF medium without fructose) medium with the supplement of a mixture of CO-CO₂-H₂-N₂. Clju $\Delta birA$ -N-*pacc*: Clju $\Delta birA$ -N, carrying the pMTL83151-*acc* plasmid. Clju $\Delta birA$ -N-*ppyc*: Clju $\Delta birA$ -N, carrying the pMTL83151-*pyc* plasmid. Data shown are means \pm standard deviations ($n = 3$) calculated from triplicate individual experiments. Error bars show standard deviations. ns: no significance.

Table S1. Statistics for the ChIP-seq data sets.

Table S2. Strains and plasmids used in this study.

Table S3. Primers used in this study.