THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 287, NO. 49, pp. 41392–41404, November 30, 2012

Structural Insights into the Substrate Specificity of the *Rhodopseudomonas palustris* Protein Acetyltransferase *Rp*Pat

IDENTIFICATION OF A LOOP CRITICAL FOR RECOGNITION BY RpPat[*S]◆

Received for publication, September 6, 2012, and in revised form, October 16, 2012 Published, JBC Papers in Press, October 17, 2012, DOI 10.1074/jbc.M112.417360

Heidi A. Crosby[‡1], **Katherine C. Rank**[§2], **Ivan Rayment**[§], **and Jorge C. Escalante-Semerena**[¶3]

From the Departments of [‡]*Bacteriology and* [§]*Biochemistry, University of Wisconsin, Madison, Wisconsin 53706 and the* [¶]*Department of Microbiology, University of Georgia, Athens, Georgia 30602*

Background: *Rp*Pat acetylates many acyl-CoA synthetase enzymes.
Results: *Rp*Pat does not acetylate *Rp*MatB but can acetylate chimeric versions of it that differ from the wild-type enzyme in a loop region >20 Å away from the acetylated lysine.
Conclusion: *Rp*Pat likely interacts with a large surface area of its substrates.
Significance: *Rp*Pat substrates cannot be predicted by a short acetylation motif alone.

Lysine acetylation is a post-translational modification that is important for the regulation of metabolism in both prokaryotes and eukaryotes. In bacteria, the best studied protein acetyltransferase is Pat. In the purple photosynthetic bacterium *Rhodopseudomonas palustris,* at least 10 AMP-forming acyl-CoA synthetase enzymes are acetylated by the Pat homologue *Rp*Pat. All *bona fide Rp*Pat substrates contain the conserved motif PX_4GK. Here, we show that the presence of such a motif is necessary but not sufficient for recognition by *Rp*Pat. *Rp*Pat failed to acetylate the methylmalonyl-CoA synthetase of this bacterium (hereafter *Rp*MatB) *in vivo* and *in vitro*, despite the homology of *Rp*MatB to known *Rp*Pat substrates. We used *Rp*MatB to identify structural determinants that are recognized by *Rp*Pat. To do this, we constructed a series of *Rp*MatB chimeras that became substrates of *Rp*Pat. In such chimeras, a short region (11–25 residues) of *Rp*MatB located >20 residues N-terminal to the acetylation site was replaced with the corresponding sequences from other AMP-forming acyl-CoA synthetases that were known *Rp*Pat substrates. Strikingly, the enzymatic activity of *Rp*MatB chimeras was regulated by acetylation both *in vitro* and *in vivo*. Crystal structures of two of these chimeras showed that the major difference between them and wild-type *Rp*MatB was within a loop region ~23 Å from the acetylation site. On the basis of these results, we suggest that *Rp*Pat likely interacts with a relatively large surface of its substrates, in addition to the PX_4GK motif, and that *Rp*Pat probably has relatively narrow substrate specificity.

Lysine acetylation is a post-translational modification that occurs in all three domains of life (1, 2). In eukaryotes, histone tails are subject to many types of modifications, including lysine acetylation, and such modifications play an important role in gene expression and chromatin stability (3, 4). Because histone acetylation has been extensively studied, the modifying enzymes are often referred to as histone acetyltransferases, despite the fact that they can acetylate non-histone proteins as well. There are three groups of histone acetyltransferases as follows: the Gcn5-related *N*-acetyltransferases, the MYST family, and the p300 family (5, 6). Only Gcn5-related *N*-acetyltransferases are found in bacteria, and different family members acetylate a wide range of substrates containing primary amines, including small molecules, protein N-terminal amino groups, and lysine side chains (7).

The bacterial protein acetyltransferase Pat[4] was first discovered as a regulator of the AMP-forming acetyl-CoA synthetase (Acs) enzyme in *Salmonella enterica* (8). In *Escherichia coli*, the homologue of *Se*Pat is known as Pka or YfiQ (9). Acs (EC 6.2.1.1; pfam 00501) activates acetate to acetyl-CoA (10). In *S. enterica*, Acs activity is abolished when residue Lys-609 is acetylated. Acs acetylation is reversed by the action of a sirtuin-type NAD^+-dependent protein deacetylase called CobB (11, 12). In other bacteria, Pat homologues have been shown to acetylate different acyl-CoA synthetases and to propionylate the enzyme propionyl-CoA synthetase (13–15). A feature shared by all these protein substrates is the acetylation site, which is located within a conserved PX_4GK motif located near the C terminus of these proteins. At present, it is not known which components of this motif are important for recognition or whether this motif is sufficient for acetylation by Pat.

There are reports of protein substrates of Pat that are not acyl-CoA synthetases and do not contain the PX_4GK motif. For example, the *E. coli* Pat homologue Pka acetylates RNase R, and acetylation affects the stability of the protein rather than its

* This work was supported, in whole or in part, by National Institutes of Health Grants R01 GM62203 (to J. C. E.-S.) and R01 GM83987 (to I. R.).
◆ This article was selected as a Paper of the Week.
[S] This article contains supplemental Experimental Procedures, Tables S1–S3, and additional references.
The atomic coordinates and structure factors (codes 4GXQ and 4GXR) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[1] Supported by National Institutes of Health Biotechnology Training Grant T32-GM08349 from USPHS.
[2] Supported by National Institutes of Health Molecular Biosciences Training Grant T32-GM07215 from USPHS.
[3] To whom correspondence should be addressed: Dept. of Microbiology, University of Georgia, 527 Biological Sciences Bldg., 120 Cedar St., Athens, GA 30602. Tel.: 706-542-2651; Fax: 706-542-2815; E-mail: jcescala@uga.edu.

[4] The abbreviations used are: Pat, protein acetyltransferase; Acs, acetyl-CoA synthetase; MatB, methylmalonyl-CoA synthetase; rTEV, recombinant tobacco etch virus; TCEP, tris(2-(17)carboxyethyl)phosphine; r.m.s.d., root mean square deviation; PDB, Protein Data Bank.

enzymatic activity (9, 16). *S. enterica* Pat has been reported to acetylate a wide range of metabolic enzymes in addition to Acs, including glyceraldehyde-3-phosphate dehydrogenase (GapA), isocitrate dehydrogenase phosphatase/kinase (AceK), and isocitrate lyase (AceA) (17). We note that our group was unable to repeat these findings (14). Thus, whether Pat has broad or narrow substrate specificity remains an open question. Because it has been suggested that Pat may be a master regulator of bacterial metabolism (17), it is important to define the substrate range of Pat.

In this work, we investigated the substrate specificity of *Rp*Pat, the Pat homologue in the purple photosynthetic α-proteobacterium *Rhodopseudomonas palustris*. In this bacterium, *Rp*Pat acetylates 10 different acyl-CoA synthetases involved in the activation of short, medium, and long chain fatty acids and aromatic acids (13, 14). To date, all known substrates of *Rp*Pat are acyl-CoA synthetases, and all contain the PX_4GK motif at the site of acetylation. *R. palustris* has two protein deacetylases, *Rp*SrtN and *Rp*LdaA, both of which can deacetylate and thus reactivate acyl-CoA synthetases (13, 14). We examined two acyl-CoA synthetases involved in the catabolism of dicarboxylic acids in *R. palustris*, namely pimeloyl-CoA synthetase (*Rp*PimA) and methylmalonyl-CoA synthetase (*Rp*MatB). For reasons explained below, we also studied benzoate:CoA ligase (BclM) from *Burkholderia xenovorans* LB400 (*Bx*BclM).

*Rp*PimA activates medium chain mono- and dicarboxylic acids (18), and *Rp*MatB activates malonate (a C_3 dicarboxylate) and methyl malonate (a related C_4 dicarboxylate) (19). Here, we show that wild-type *Rp*MatB is not a substrate of *Rp*Pat, despite the fact that *Rp*MatB contains the catalytic lysine residue located within the PX_4GK motif.

Using a combination of *in vivo*, *in vitro*, and structural approaches, we determined that in *Rp*MatB the sequence of a structural loop in the C-terminal domain recognized by *Rp*Pat in its *bona fide* substrates is different. As a result of relatively minor changes in this sequence, *Rp*Pat does not modify *Rp*MatB. In substrates of *Rp*Pat, the alluded region is surprisingly far away from the acetylation site, strongly suggesting that it may be difficult to identify *Rp*Pat substrates exclusively on the presence of the PX_4GK motif. It is noteworthy that *Rp*Pat (and probably its homologues in other organisms) does not acetylate all AMP-forming acyl-CoA synthetases. These results and others reported elsewhere (14) continue to question the idea proposed by others that *S. enterica* Pat is not a very specific enzyme (17).

EXPERIMENTAL PROCEDURES

Bacterial Strains and Growth Conditions—All strains and plasmids used in this study are listed in supplemental Tables S1 and S2. *E. coli* strains were grown at 37 °C in lysogenic broth (LB, Difco) (18, 19). *B. xenovorans* LB400 was obtained from the Agricultural Research Service culture collection, United States Department of Agriculture, and was cultured at 30 °C on LB devoid of NaCl. All *R. palustris* strains were derivatives of *R. palustris* CGA009 (20) and were cultured at 30 °C in photosynthetic medium (PM) (21) supplemented with succinate (10 mM), benzoate (3 mM), and $NaHCO_3$ (10 mM) or methyl malonate (10 mM) and vitamin B_{12} (150 nM). When used, ampicillin was at 100 μg ml^{-1}; chloramphenicol was at 20 μg ml^{-1}, and kanamycin was at 50 μg ml^{-1} (*E. coli*) or 75 μg ml^{-1} (*R. palustris*). Radiolabeled [1-^{14}C]acetyl-CoA (54 mCi/mmol) was purchased from Moravek, and all other chemicals were from Sigma. For growth curves, *R. palustris* starter cultures were grown to full density photosynthetically on PM containing succinate and kanamycin in Balch tubes (22) flushed with N_2. Cultures were diluted 1:16 into fresh medium supplemented with the appropriate carbon source and kanamycin. Cultures were incubated at 30 °C in light without shaking, and growth was monitored at 660 nm using a Spectronic Spec20D (Thermo). Three biological replicates of each strain were monitored, and each experiment was repeated twice.

Molecular Techniques—DNA manipulations were performed using standard techniques (23). Restriction endonucleases were purchased from Fermentas. DNA was amplified using PfuUltra II Fusion DNA polymerase (Stratagene), and site-directed mutagenesis was performed using the QuikChange kit (Stratagene). Plasmid DNA was purified using the Wizard Plus SV miniprep kit (Promega), and PCR products were purified using the Wizard SV gel and PCR clean-up kit (Promega). DNA sequencing reactions were resolved and analyzed at the University of Wisconsin-Madison Biotechnology Center. Oligonucleotide primer sequences are listed in supplemental Table S3. Plasmids were moved into *R. palustris* by electroporation.

Plasmid Construction—*R. palustris pat*, *pimA*, and *matB* were amplified from *R. palustris* genomic DNA using the primers listed in supplemental Tables S2 and S3. The PCR products were cut with the specified restriction enzymes and ligated into the overexpression vector pTEV5 cut with the same enzymes (24). The resulting plasmids directed the synthesis of proteins with an N-terminal hexahistidine (His_6) tag, which was removed using recombinant tobacco etch virus (rTEV) protease (25). *B. xenovorans bclM* (locus tag Bxe_A1419) was amplified from a single colony and cloned into pTEV5 using PIPE cloning (26). The chimeric genes were generated using overlap-extension PCR (27) or site-directed mutagenesis and were subsequently cloned into pTEV5. A detailed description of chimera construction can be found in the supplemental material.

Protein Overexpression—The plasmid encoding *Rp*Pat (pRpPAT13) was transformed into strain BL21/pLysS. The resulting strain was grown until early stationary phase and subcultured 1:100 in 3 liters of super broth supplemented with ampicillin (100 μg ml^{-1}) and chloramphenicol (10 μg ml^{-1}). Cultures were grown with shaking to an A_{600} ~0.5, and protein expression was induced with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside. Cultures were grown overnight at 30 °C; cells were harvested by centrifugation at 8000 × *g* for 12 min in a Beckman Coulter Avanti J-20 XOI refrigerated centrifuge with a JLA-8.1000 rotor, and cell pellets were frozen at −80 °C for later use. All other overexpression plasmids were transformed into strain JE9314, a *pka*-deficient derivative of *E. coli* C41(λDE3) (28), to ensure that proteins were purified in their nonacetylated state. The resulting strains were grown until early stationary phase and subcultured 1:100 in 2 liters of lysogenic broth supplemented with ampicillin (100 μg ml^{-1}). Protein expression was induced with isopropyl 1-thio-β-D-galactopyranoside as

described above, and cell pellets were frozen at −80 °C for later use.

Protein Purification—Protein concentrations were determined using a NanoDrop 1000 spectrophotometer (ThermoFisher) and A_{280} molar extinction coefficients for each protein. Extinction coefficients were obtained from the Integrated Microbial Genomes website (29) or calculated using the ProtParam tool (30).

rTEV Protease—rTEV protease was purified as described previously (25).

R. palustris Pat—*Rp*Pat was overproduced as an rTEV-cleavable His_6 N-terminal fusion protein. The cell pellet was resuspended in 30 ml of *Rp*Pat binding buffer (sodium phosphate buffer (50 mM, pH 7.5), NaCl (500 mM), imidazole (20 mM), tris(2-carboxyethyl)phosphine (TCEP, 0.5 mM)) and supplemented with lysozyme and deoxyribonuclease I (1 mg ml^{-1} each). Cells were lysed using a French pressure cell (two passes, 10 MPa). Debris was removed by centrifugation at 39,000 × *g* for 30 min at 4 °C, and the soluble fraction was passed through a 0.45-μm filter. His_6-*Rp*Pat was purified by nickel-affinity purification using an FPLC (ÄKTA) equipped with a 5-ml HisTrap FF column. After binding to the column, the column was washed with 6% *Rp*Pat elution buffer (sodium phosphate buffer (50 mM, pH 7.5), NaCl (500 mM), imidazole (500 mM), TCEP (0.5 mM)) and eluted in a linear gradient up to 100% *Rp*Pat elution buffer. Fractions containing *Rp*Pat were pooled and incubated with rTEV (1:10 rTEV to protein ratio) for 3 h at room temperature to cleave the tag. *Rp*Pat was dialyzed three times against *Rp*Pat binding buffer lacking imidazole; the first dialysis buffer contained EDTA (1 mM). The His_6 tag and His_7-tagged rTEV protease were removed by passage over the HisTrap column again, using the same buffers as the first purification except that the binding buffer did not contain imidazole. Untagged *Rp*Pat eluted in the wash step containing 6% *Rp*Pat elution buffer. Fractions were pooled and dialyzed three times against *Rp*Pat storage buffer containing HEPES (50 mM, pH 7.5), glycerol (20%, v/v), and NaCl (500, 300, and 150 mM, successively), before drop freezing in liquid nitrogen and storage at −80 °C.

Other Proteins—All other proteins were purified in a similar manner, with the following modifications. Cell pellets were resuspended in MatB binding buffer (sodium phosphate buffer (50 mM, pH 8.0), NaCl (300 mM), imidazole (20 mM)) and lysed with a Sonic Dismembrator (Fisher) at power level 9, with 2-s pulses separated by 5-s breaks, for 1.5 min. His_6-tagged proteins were purified from the soluble fraction by nickel-affinity purification using a 1.5-ml bed volume of nickel-nitrilotriacetic acid Superflow resin (Qiagen). After binding to the column, the column was washed with *Rp*MatB washing buffer (sodium phosphate buffer (50 mM, pH 8.0), NaCl (300 mM), imidazole (30 mM)), and proteins were eluted with *Rp*MatB elution buffer (sodium phosphate buffer (50 mM, pH 8.0), NaCl (300 mM), imidazole (250 mM)). rTEV was added (1:50 to 1:100 rTEV to protein ratio), and after a 3-h cleavage reaction at room temperature proteins were dialyzed against *Rp*MatB dialysis buffer 1 (sodium phosphate buffer (50 mM, pH 8.0), NaCl (300 mM), EDTA (0.5 mM)), then *Rp*MatB dialysis buffer 2 (sodium phosphate buffer (50 mM, pH 8.0), NaCl (300 mM)), and finally *Rp*MatB binding buffer. The second purification step employed the same buffers as the first, and untagged proteins were eluted in the flow-through. Fractions were pooled and dialyzed against *Rp*MatB storage buffer 1 (Tris (50 mM, pH 7.5), NaCl (100 mM), EDTA (0.5 mM), glycerol (20%, v/v)) and then *Rp*MatB storage buffer 2 (Tris (50 mM, pH 7.5), NaCl (100 mM), glycerol (20%, v/v)) before drop-freezing in liquid N_2.

Chimeras—The purification protocols for chimeras *Rp*MatB-*Bx*BclM $B1^{K491A}$ and $B3^{K488A}$ for crystallography were similar to the ones described above, except that the purified, untagged protein was dialyzed into HEPES buffer (50 mM, pH 7.5) before concentration to 10–12 mg ml^{-1} using Ultracel 10,000 molecular weight cutoff centrifugal filters (Amicon Ultra). Accessible lysine side chains were reductively methylated using formaldehyde and dimethylamine borane complex as described previously (31). The reaction was quenched by addition of Tris-HCl buffer (pH 7.5) to a final concentration of 500 mM, and the protein was dialyzed overnight against 2 liters of Tris-HCl buffer (50 mM, pH 7.5) containing dithiothreitol (DTT, 2 mM), before dialysis into Tris-HCl buffer (10 mM, pH 7.5). The protein concentration was adjusted to 10 mg ml^{-1} before drop-freezing in liquid nitrogen.

In Vitro Protein Acetylation Assay—Proteins were acetylated using radiolabeled acetyl-CoA as described previously (8). Reactions contained HEPES buffer (50 mM, pH 7.0), TCEP (1 mM), [1-^{14}C]Ac-CoA (19 μM), protein substrate (3 μM), and *Rp*Pat (0.06 μM). Reactions (25 μl total volume) were incubated for 60 min at 30 °C. Aliquots (10 μl each) were resolved using SDS-PAGE and visualized by staining with Coomassie Blue R-250. Gels were dried and exposed overnight to a MultiPurpose Phosphor Screen (Packard Instrument Co.). Radioactivity associated with protein was quantified using a Typhoon FLA9000 Biomolecular Imager (GE Healthcare) and OptiQuant version 04.00 software (Packard Instrument Co.).

In Vitro Acyl-CoA Synthetase Activity Assay—Acyl-CoA synthetases (1.5 μM each) were individually incubated with *Rp*Pat (0.5 μM) plus or minus 50 μM acetyl-CoA for 1 h at 30 °C using the same buffer system described above for ^{14}C acetylation assays. Acyl-CoA synthetase activity was quantified using an NADH consumption assay (32). Reactions contained HEPES (50 mM, pH 7.5), TCEP (1 mM), ATP (2.5 mM), CoASH (0.4 mM), $MgCl_2$ (5 mM), phosphoenolpyruvate (3 mM), NADH (0.1 mM), pyruvate kinase (1 unit), myokinase (5 units), lactate dehydrogenase (1.5 units), and an organic acid substrate (0.2 mM). Reactions were started by the addition of acyl-CoA synthetase (30 nM), and changes in the absorbance at 340 nm were monitored for 8 min in a 96-well plate format using a Spectramax Plus UV-visible spectrophotometer (Molecular Devices).

Crystallization of RpMatB—Methylated *Rp*MatB chimera $B1^{K491A}$ crystals were grown by mixing 2 μl of 10 mg/ml protein in Tris-HCl buffer ((10 mM, pH 7.5) containing TCEP (1 mM) and MgATP (10 mM)) with 2 μl of reservoir solution ((HEPES buffer (100 mM, pH 7.5), 200 mM potassium glutamate, polyethylene glycol 8000 (26% (w/v), propylene glycol (5%, w/v)) and hanging the resulting mixture over 500 μl of reservoir solution in Linbro plates at 25 °C. After 24 h, crystals were nucleated by streak seeding with previously grown crystals and allowed to grow for an additional week. For freezing, crystals were trans-

TABLE 1
Crystal structure data collection and refinement statistics

Data collection	Chimera B1	Chimera B3
Space group	$P6_1$	$P6_1$
Cell dimensions		
a, *b*, *c*	299.3, 299.3, 47.9 Å	173.2, 173.2, 47.9 Å
α, β, γ	90, 90, 120°	90, 90, 120°
Wavelength	0.98	0.98
Resolution[a]	25 to 2.0 Å (2.03 to 2.0 Å)	25 to 2.0 Å (2.03 to 2.0 Å)
R_{merge}[a]	6.3 (38.5)	5.8 (30.1)
$I/\sigma I$[a]	11.9 (3.8)	15.8 (6.1)
Completeness[a]	99.8% (99.7%)	99.9% (100%)
Redundancy[a]	5.1 (4.3)	6.2 (6.1)
Refinement		
Resolution[a]	25 to 2.0 Å (2.03 to 2.0 Å)	25 to 2.0 Å (2.03 to 2.0 Å)
No. reflections[b]	158,213 (8364)	52,933 (2833)
R_{work} / R_{free}[c]	0.219/0.254	0.148/0.178
No. of atoms		
Protein	11,832	3958
Water	1586	535
Ligand	139	63
Wilson *B*-factor	23.0	25.3
Average *B*-factor		
Protein	14.6 Å^2	21.2 Å^2
Water	40.8 Å^2	41.4 Å^2
Ligand	26.9 Å^2	43.5 Å^2
Ramachandran		
Most favored	99.02%	98.95%
Allowed	1.95%	1.87%
Disallowed	0.14%	0.0%
r.m.s.d.		
Bond lengths	0.008 Å	0.005 Å
Bond angles	1.325°	1.169°
PDB accession no.	4GXQ	4GXR

[a] Data in parentheses represent highest resolution shell.
[b] Data in parentheses represent the number of reflections used during refinement.
[c] $R_{factor} = \sum|F_{obs} - F_{calc}|/\sum|F_{obs}|$, where R_{work} refers to the R_{factor} for the data utilized in the refinement and R_{free} refers to the R_{factor} for 5% of the data that were excluded from the refinement.

ferred to 100 μl of reservoir solution with the addition of MgATP (10 mM) and then transferred stepwise to a solution of polyethylene glycol 8000 (25%, w/v), propylene glycol (15%, w/v), ethylene glycol, HEPES buffer (100 mM, pH 7.5), potassium glutamate (200 mM), and MgATP (10 mM). Crystals were flash-frozen in liquid nitrogen.

Methylated *Rp*MatB chimera B3^{K488A} crystals were grown by mixing 2 μl of 10 mg/ml protein in Tris-HCl buffer (10 mM, pH 7.5) containing TCEP (1 mM), and MgATP (10 mM) with 2 μl of reservoir solution (triethanolamine (100 mM, pH 8.0), with magnesium sulfate (50 mM), monomethyl polyethylene glycol 5000 (21%, w/v), and glycerol (4%, w/v)) and hanging the resulting mixture over 500 μl of reservoir solution in Linbro plates at 25 °C. After 24 h, crystals were nucleated by streak seeding with previously grown crystals and allowed to grow for an additional week. For freezing, crystals were transferred to 100 μl of reservoir solution with the addition of MgATP (10 mM) and then transferred stepwise to a solution of monomethyl polyethylene glycol 5000 (25%, w/v), glycerol (18%, w/v), triethanolamine (100 mM, pH 8.0), magnesium sulfate (50 mM), and MgATP (10 mM). Crystals were flash-frozen in liquid nitrogen.

X-Ray Data Collection and Structural Refinement—X-ray diffraction data for all crystals were collected at the SBC 19-ID beam line (Advanced Photon Source, Argonne, IL). The data sets were integrated and scaled with the program HKL2000 (33). X-ray data collection statistics are given in Table 1.

Structural Determination—The structures of both methylated *Rp*MatB chimeras were solved by molecular replacement using the program Phaser (34, 35) and the methylated *Rp*MatB^{K488A} structure (PDB accession number 4FUT) (36) as the search model. This was followed by iterative cycles of manual model building in Coot (37) and TLS and restrained refinement in Refmac 5.6 and Phenix.refine: 1.6_289 (38, 39). For both structures, TLS groups were defined such that each polypeptide chain within the asymmetric unit was divided into the N-terminal (residues 1–400) and C-terminal domains (residues 401–503 or 506). These two domains were naturally flexible during the catalytic cycle, and introduction of these TLS groups led to a 2% decrease in R_{free} for both structures (21.88/25.44% R/R_{free} with TLS compared with 23.10/27.61% R/R_{free} for 4GXQ and 14.83/17.79% R/R_{free} with TLS compared with 16.06/19.48% R/R_{free} for 4GXR). Data processing and refinement statistics are presented in Table 1. All structural alignments were done with the program Superpose (40).

Structures have been deposited in the RCSB Protein Data Bank under accession codes 4GXQ and 4GXR for ATP-*Rp*MatB chimera B1^{K491A} and ATP-*Rp*MatB chimera B3^{K488A}, respectively.

RESULTS

In R. palustris, MatB Is Not Regulated by Acetylation—We tested whether *Rp*Pat could acetylate the acyl-CoA synthetases *Rp*MatB, *Rp*PimA, and *Bx*BclM. *Rp*MatB is 39% identical to *Rp*PimA over 376 residues and 28% identical to *Bx*BclM over 485 residues, suggesting that all three enzymes share many structural features, although they do exhibit different substrate specificities. *Rp*PimA, *Rp*MatB, and *Bx*BclM contain the conserved lysine residue and surrounding motif PX_4GK found in other AMP-forming acyl-CoA synthetases that are acetylated, such as the *R. palustris* benzoate:CoA ligase (*Rp*BadA, Fig. 1*A*). We purified each enzyme, and each was tested as a substrate for *Rp*Pat using ^{14}C-labeled acetyl-CoA as co-substrate (Fig. 1*B*). *Rp*Pat efficiently acetylated *Rp*PimA and *Rp*BclM, and when the conserved lysine residues (Lys-534 in *Rp*PimA or Lys-520 in *Bx*BclM) were changed to alanine, acetylation was no longer detectable, consistent with the expectation that this lysine was modified in both proteins. Surprisingly, *Rp*Pat did not acetylate *Rp*MatB, despite its similarity to other acyl-CoA synthetases that are substrates of *Rp*Pat.

Growth of R. palustris on Methyl Malonate Is Not under Acetylation Control—We recently showed that *matB* is required for photoheterotrophic growth of *R. palustris* on methyl malonate as a carbon source (36). If *Rp*MatB activity were regulated by acetylation *in vivo*, we would expect that a deacetylase-deficient strain (Δ*srtN* Δ*ldaA*) would have a growth defect, as is observed during growth on benzoate (Fig. 1*C*). This growth defect on benzoate occurs because *Rp*BadA (benzoate:CoA ligase) is acetylated and inactivated by *Rp*Pat, and the deacetylases *Rp*SrtN and *Rp*LdaA are not present to reactivate *Rp*BadA. As shown in Fig. 1*D*, an *R. palustris* Δ*srtN* Δ*ldaA* strain (JE11616, supplemental Table S1) grew as well as the wild-type strain on methyl malonate, a result that was consistent with our *in vitro* data, which suggested that *Rp*MatB was not a substrate of *Rp*Pat.

We further confirmed that *Rp*Pat did not acetylate *Rp*MatB. To do this, we quantified *Rp*MatB activity after incubation with *Rp*Pat and acetyl-CoA (Table 2). Results of positive control

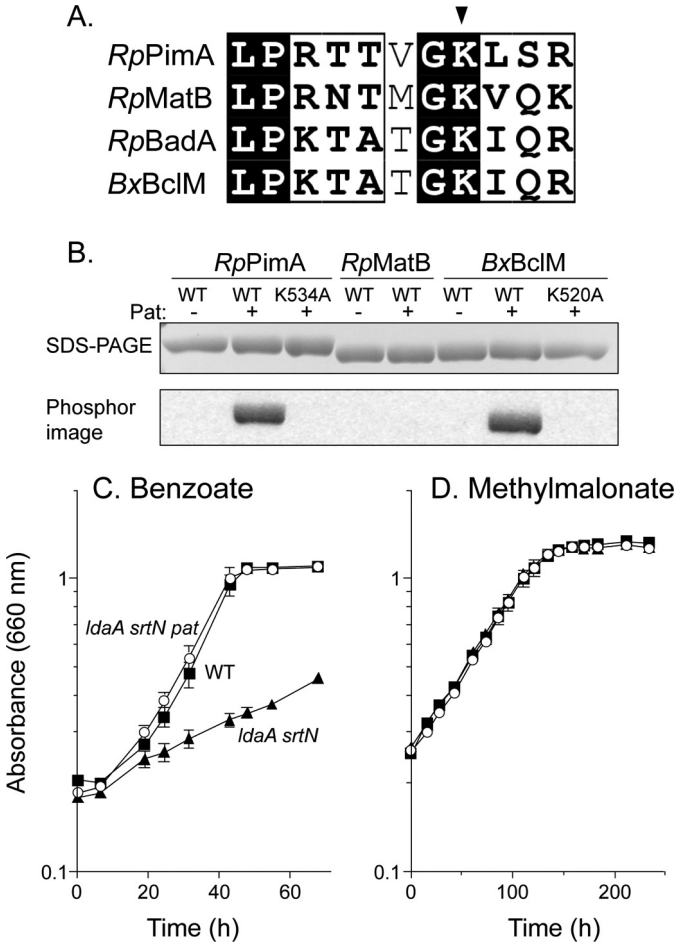

FIGURE 1. ***Rp*Pat acetylates *Rp*PimA and *Bx*BclM but not *Rp*MatB.** *A,* alignment of the region around the lysine residue that is acetylated in *Rp*BadA. *Arrow* indicates Lys-512 in *Rp*BadA, corresponding to Lys-534 in *Rp*PimA, Lys-488 in *Rp*MatB, and Lys-520 in *Bx*BclM. *B, Rp*PimA, *Rp*MatB, and *Bx*BclM were incubated with [1-^{14}C]acetyl-CoA and with or without *Rp*Pat. Changing Lys-534 of *Rp*PimA or Lys-520 of *Bx*BclM to alanine abolished acetylation. *Top panel* shows the acyl-CoA synthetases separated by SDS-PAGE, and *bottom panel* is the phosphorimage of the same gel. *C,* photosynthetic growth of *R. palustris* on benzoate is regulated by acetylation. A deacetylase-deficient strain (Δ*ldaA* Δ*srtN*, *triangles*) grew poorly because *Rp*BadA was acetylated and rendered inactive, and deletion of *pat* (Δ*ldaA* Δ*srtN* Δ*pat*, *circles*) restored growth. *D,* photosynthetic growth of *R. palustris* on methyl malonate was not regulated by acetylation. Wild-type (*squares*), deacetylase-deficient (Δ*ldaA* Δ*srtN*, *triangles*), and deacetylase- and *pat*-deficient (Δ*ldaA* Δ*srtN* Δ*pat*, *circles*) strains all grew equally well.

experiments, where *Rp*PimA or *Bx*BclM substituted for *Rp*MatB in the reaction mixture, showed a ≥99% loss of acyl-CoA synthetase activity upon acetylation. In contrast, *Rp*MatB activity was unaffected by incubation with *Rp*Pat and acetyl-CoA.

Generation of Chimeric Variants of RpMatB That Can Be Acetylated by RpPat—All AMP-forming acyl-CoA synthetases consist of a large N-terminal domain and a smaller C-terminal domain connected by a flexible hinge region (41, 42). These enzymes activate fatty acids in two steps, an adenylation half-reaction and a thioester-forming half-reaction; to accomplish this, the C-terminal domain rotates ~140° between two catalytic conformations (43, 44). The catalytic lysine residue that is acetylated in members of this family of enzymes is located within the C-terminal domain (Fig. 2*A*). When the enzyme is in the adenylation conformation, the catalytic lysine residue is buried in the active site in the cleft between the N- and C-terminal domains, although this lysine residue is surface-exposed when the enzyme rotates into the thioester-forming conformation. Thus, this lysine is likely only available for modification when the enzyme adopts the thioester-forming conformation.

TABLE 2
Acyl-CoA synthetase enzyme activity

Enzyme	Substrate	Unacetylated activity[a]	Acetylated activity	Decrease
				%
PimA	Octanoate	6.9 ± 0.2	0.04 ± 0.02	99
MatB	Methyl malonate	8.6 ± 0.8	8.8 ± 1.01	0
BclM	Benzoate	6.2 ± 0.7	0.04 ± 0.03	99
P9 chimera	Methyl malonate	3.7 ± 0.2	0.05 ± 0.03	98
B1 chimera	Methyl malonate	6.1 ± 0.1	1.8 ± 0.32	71
B2 chimera	Methyl malonate	6.0 ± 0.2	2.1 ± 0.56	65
B3 chimera	Methyl malonate	6.8 ± 0.2	4.1 ± 0.39	40

[a] Activity was reported as micromoles of AMP min^{-1} mg^{-1}.

To determine why *Rp*Pat could not acetylate *Rp*MatB, we made a series of fusion proteins in which a portion of the C-terminal domain of *Rp*MatB was replaced with the corresponding sequence of *Rp*PimA (Fig. 2*C*). *Rp*Pat acetylated *Rp*MatB-*Rp*PimA chimeras that contained at least the last 62 residues of *Rp*PimA (Fig. 2*B*, annotated as *chimeras P1, P2,* and *P3*). These chimeras included at least 43 *Rp*PimA-derived residues upstream of the acetylated lysine, demonstrating that acetylation depended on more than just the motif immediately surrounding the acetylation site.

To determine whether this entire region upstream of the acetylated lysine was required to turn *Rp*MatB into an *Rp*Pat substrate, we made a second set of chimeras in which an internal piece of *Rp*MatB was replaced with the corresponding sequence from *Rp*PimA (Fig. 2*E*). Strong acetylation was observed when residues 440–465 of *Rp*MatB were replaced with the corresponding region of *Rp*PimA (Fig. 2*D*, *chimera P9*). Acetylation still occurred at the conserved lysine residue (Lys-489 in the P9 chimera) because substitution of an alanine at this position abolished acetylation (Fig. 2*F*). These results showed that the primary sequence surrounding Lys-488 in *Rp*MatB (Fig. 2*H*) would not block the modification of Lys-488 by *Rp*Pat, but a region spanning 49 to 23 residues N-terminal to this lysine was critical for recognition by *Rp*Pat.

*Rp*Pat also acetylated a similar *Rp*MatB-*Bx*BclM chimera (Fig. 2, *F* and *G*), although the segment of *Rp*MatB replaced by *Bx*BclM had to be adjusted slightly to yield a stable protein. In this chimera, referred to as *Rp*MatB-*Bx*BclM B1, residues 444–465 of *Rp*MatB were replaced with the corresponding sequence of *Bx*BclM (Fig. 2*H*). Substitution of the conserved active site lysine residue, Lys-491 in the *Rp*MatB-*Bx*BclM B1 chimera, abolished acetylation (Fig. 2*F*), demonstrating that acetylation occurred at the expected site.

RpMatB Chimeras Are Regulated by Acetylation in Vitro and in Vivo—Both the *Rp*MatB-*Rp*PimA P9 and *Rp*MatB-*Bx*BclM B1 chimeras retained enzymatic activity with methyl malonate, although the *Rp*MatB-*Rp*PimA P9 chimera was only ~40% as active as wild-type *Rp*MatB, and the *Rp*MatB-*Bx*BclM B1 chimera was ~70% as active as wild-type *Rp*MatB (Table 2). The *Rp*MatB-*Rp*PimA P9 chimera was efficiently acetylated by Pat and lost ~99% of its activity upon acetylation, whereas the

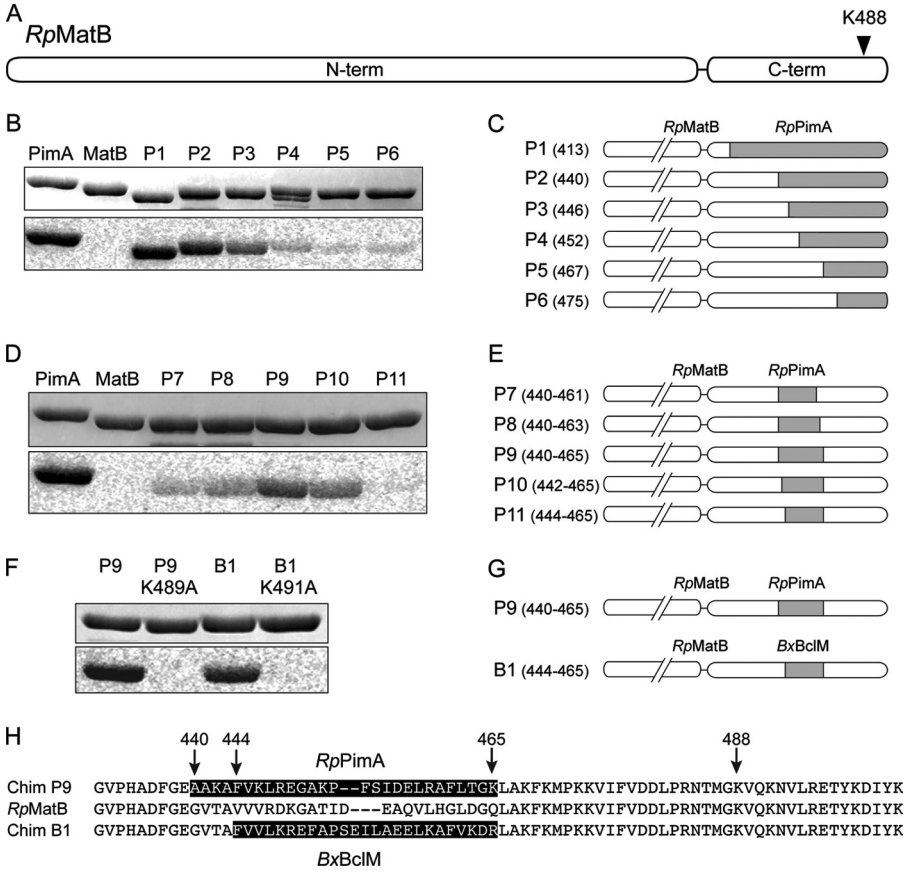

FIGURE 2. **Construction of *Rp*MatB chimeric proteins that can be acetylated.** *A,* scheme of *Rp*MatB N- and C-terminal domains drawn to scale. *Arrowhead* indicates location of conserved lysine residue Lys-488. *B,* acetylation of six *Rp*MatB-*Rp*PimA chimeras by *Rp*Pat using [1-^{14}C]acetyl-CoA. *Top panel* shows SDS-polyacrylamide gel, and *bottom panel* shows the phosphorimage of the same gel. *C,* schematic representation of the six *Rp*MatB-*Rp*PimA chimeras *P1–P6*, in which a section of the C-terminal domain of *Rp*MatB was replaced with the corresponding sequence from *Rp*PimA (*gray*). C-terminal domains are drawn to match scale of *A*. *Numbers* in *parentheses* indicate the residue of *Rp*MatB at which the fusion begins (*i.e.* first residue replaced by PimA). *D,* acetylation by *Rp*Pat of *Rp*MatB chimeras containing an internal fragment of *Rp*PimA. *E,* schematic representation of *Rp*MatB-*Rp*PimA chimeras *P7–P11*, containing an internal fragment of *Rp*PimA (*gray*). *Numbers* in *parentheses* indicate residues of *Rp*MatB that were replaced by the corresponding sequence of *Rp*PimA. *F,* acetylation by *Rp*Pat of the *Rp*MatB-*Bx*BclM chimera B1. Changing the conserved lysine residue (Lys-488 in wild-type *Rp*MatB numbering) in either chimera resulted in no acetylation by *Rp*Pat. *G,* schematic representation of the *Rp*MatB-*Rp*PimA chimera P9 and *Rp*MatB-*Bx*BclM chimera B1 that were the best identified substrates of *Rp*Pat. *H,* alignment of the C-terminal ends of *Rp*MatB and the *Rp*MatB-*Rp*PimA P9 and *Rp*MatB-*Bx*BclM B1 chimeras, starting at Gly-431. The sequences derived from *Rp*PimA and *Bx*BclM are *shaded*, and *Rp*MatB residue numbers are indicated with *arrows*.

*Rp*MatB-*Bx*BclM B1 chimera lost ~70% of its activity upon acetylation, under our assay conditions (Table 2).

The *Rp*MatB-*Rp*PimA P9 chimera, which retained 40% of the wild-type *Rp*MatB activity, did not complement a Δ*matB* strain growing photoheterotrophically on methyl malonate as a carbon source (Fig. 3). However, the same chimera did complement a Δ*matB* Δ*pat* strain (Fig. 3, *circles*), suggesting that in the Δ*matB pat*$^+$ strain the chimera was acetylated and rendered inactive. If this were true, we predicted that a Δ*matB* strain co-expressing the genes encoding the chimera and the *Rp*LdaA deacetylase would also grow, and this was the case (Fig. 3, *open diamonds*). These results indicated that the *Rp*MatB-*Rp*PimA P9 chimera could be regulated by acetylation *in vivo*.

Three-dimensional Crystal Structure of RpMatB-RpBclM Chimera B1—Because acyl-CoA synthetases show little sequence conservation in the region between residues 440 and 465 (*Rp*MatB numbering, Fig. 2*H*), it was unclear why this region was important for recognition or acetylation by *Rp*Pat. We hypothesized that there was a structural difference between wild-type *Rp*MatB and its chimeras, and we attempted to crystallize both the *Rp*MatB-*Rp*PimA P9 and *Rp*MatB-*Bx*BclM B1 chimeras so that we could compare them with *Rp*MatB and *Bx*BclM, whose structures have been previously published (36, 45). Attempts to crystallize either *Rp*PimA or the *Rp*MatB-*Rp*PimA P9 chimera were not successful; therefore, only the *Rp*MatB-*Bx*BclM B1 chimera was pursued.

We obtained crystals of the *Rp*MatB-*Bx*BclM B1 chimera with ATP bound only after we chemically methylated the surface-exposed lysine residues. Methylation of lysine residues is known to affect packing within the crystal lattice without significantly changing the three-dimensional structure of the protein (46–48). This method has been used to obtain crystals of recalcitrant proteins, including *Rp*MatB (36, 49–53). To avoid predictable problems stemming from chemical methylation of the active site lysine residue, this lysine (Lys-491) was changed to alanine.

The structure of the *Rp*MatB-*Bx*BclM B1 chimera was solved by molecular replacement, using *Rp*MatB^{K488A} with ATP bound as a search model. Three *Rp*MatB-*Bx*BclM B1 chimera monomers were found within the unit cell each having substantially the same structure, including that of the chimera region. Within the *Rp*MatB-*Bx*BclM B1 chimera structure, chain A

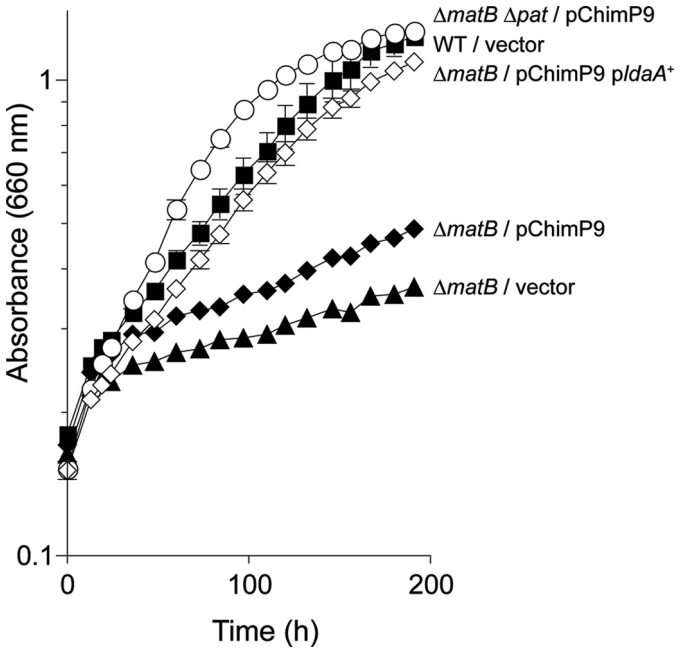

FIGURE 3. **Activity of the *Rp*MatB-*Rp*PimA P9 chimera is controlled by acetylation *in vivo*.** Absorbance was monitored during photosynthetic growth of *R. palustris* on methyl malonate (10 mM). Data points are averages of four replicates, and *error bars* represent standard deviations. *Squares,* wild-type (CGA009) harboring plasmid pBBR1MCS-2 as a vector control; *filled triangles,* strain JE13568 (Δ*matB*/pBBR1MCS-2); *filled diamonds,* strain JE13912 (Δ*matB*/pRpMATB25 (encodes the *Rp*MatB-*Rp*PimA P9 chimera)); *circles,* strain JE14288 (Δ*matB* Δ*pat*/pRpMATB25 (encodes the *Rp*MatB-*Rp*PimA P9 chimera)); *empty diamonds,* strain JE14957 (Δ*matB*/pRpMATB45 (encodes the *Rp*MatB-*Rp*PimA P9 chimera and *ldaA*$^+$).

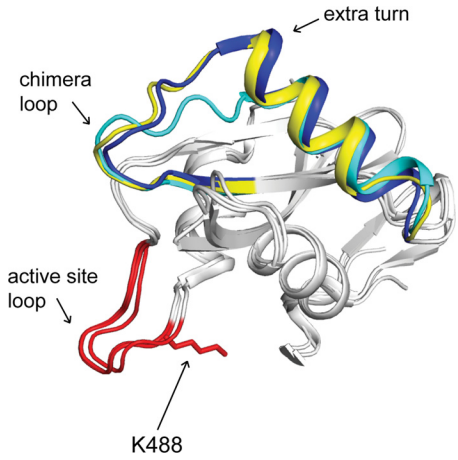

FIGURE 4. **Crystal structure of *Rp*MatB-*Bx*BclM chimera B1.** The C-terminal domain of the *Rp*MatB-*Bx*BclM B1 chimera is shown aligned with the C-terminal domains of *Rp*MatB (PDB accession number 4FUT (36)) and *Bx*BclM (PDB accession number 2V7B (45)). The *Bx*BclM-derived residues of the chimera are colored *yellow*, and the corresponding residues of *Rp*MatB and *Bx*BclM are shown in *cyan* and *dark blue*, respectively. The active site lysine residue of *Rp*MatB was modeled from the *Rp*MatB apo structure (PDB accession number 4FUQ (36)), and the PX_4GK region surrounding this residue (active site loop) is shown in *red*. *Arrows* indicate the chimera loop and the extra turn in the α-helix of *Bx*BclM and the B1 chimera compared with *Rp*MatB. Figure was prepared with PyMOL (60).

compared with chain B and chain A compared with chain C have root mean square deviations (r.m.s.d.) of 0.46 Å over 3948 eq atoms and 0.50 Å over 3949 eq atoms, respectively. However, chain C does show some disorder in the C-terminal domain due to crystal packing. Fig. 4 shows an alignment of the C-terminal domains of the *Rp*MatB-*Bx*BclM B1 chimera, *Rp*MatB^{K488A}, and *Bx*BclM. All three structures are very similar, having an r.m.s.d. of 0.46 Å over 501 eq α-carbons and 1.88 Å over 456 eq α-carbons when comparing *Rp*MatB-*Bx*BclM B1 chimera and *Bx*BclM, respectively, with *Rp*MatB^{K488A} and differ significantly only within the chimeric region.

The *Bx*BclM-derived residues in the *Rp*MatB-*Bx*BclM B1 chimera encompass the end of a β-strand, a connecting loop (referred to here as the chimera loop), and an α-helix that is located on the opposite side of the domain from the acetylated lysine residue and active site loop (Fig. 4). Electron density maps for this region are well defined as shown in Fig. 5. Although this chimeric region is very similar to the *Bx*BclM structure from which it is derived (having an r.m.s.d. of 1.56 Å over 207 eq atoms), it is distinct from the corresponding region of *Rp*MatB. The most notable difference is that there is one extra turn in the α-helix of *Bx*BclM and the chimera compared with *Rp*MatB (four turns instead of three), and the chimera loop has a different spatial orientation. Based on this structural comparison, it seemed possible that *Rp*Pat made crucial contacts with either the extra turn in the α-helix and/or the chimera loop.

We generated a sequence alignment of *Rp*MatB and the *Rp*MatB-*Bx*BclM B1 chimera based on the structural alignment shown in Fig. 4, which included a three-amino acid gap in the *Rp*MatB sequence where the extra turn in the α-helix would be (Fig. 6*A*). To identify which parts of this chimera were important for acetylation by *Rp*Pat, we first focused on the last three turns in the α-helix, which were structurally very similar in *Rp*MatB and the chimera. We constructed chimera *Rp*MatB-*Bx*BclM B2, in which the last 11 residues in the chimeric region were changed back to the corresponding sequence in *Rp*MatB (Fig. 6*A*). Chimera *Rp*MatB-*Bx*BclM B2 was acetylated as well as the *Rp*MatB-*Bx*BclM B1 chimera (Fig. 6, *B* and *C*; Table 2), suggesting that the amino acid side chains on the last three turns of the α-helix of the chimera were not important for acetylation. Next, we eliminated the three residues (Glu-Ile-Leu) that comprised the extra turn of the α-helix to generate chimera *Rp*MatB-*Bx*BclM B3 (Fig. 6*A*). The latter was also acetylated by *Rp*Pat (Fig. 6, *B* and *C*), albeit not as efficiently as the *Rp*MatB-*Bx*BclM B1 chimera (40% decrease in activity upon acetylation compared with a 71% decrease in activity for the *Rp*MatB-*Bx*BclM B1 chimera, Table 2). These data suggested that *Rp*Pat recognized proteins that have either a three-turn or a four-turn α-helix in this position.

Crystal Structure of RpMatB-RpBclM Chimera B3—We determined the crystal structure of the *Rp*MatB-*Bx*BclM B3 chimera to confirm that it did have a shortened α-helix. The *Rp*MatB-*Bx*BclM B3 chimera K488A variant was methylated prior to crystallization and crystallized with the same unit cell and space group as the previously reported ATP-*Rp*MatB^{K488A} (19). As expected, the *Rp*MatB-*Bx*BclM B3 chimera did have a shorter α-helix within the chimeric region, similar to that of wild-type *Rp*MatB (Fig. 6*D*). Again electron density for this region was well defined as shown in Fig. 5. The chimera loop had a different conformation than both wild-type *Rp*MatB and the *Rp*MatB-*Bx*BclM B1 chimera suggesting that it was not the

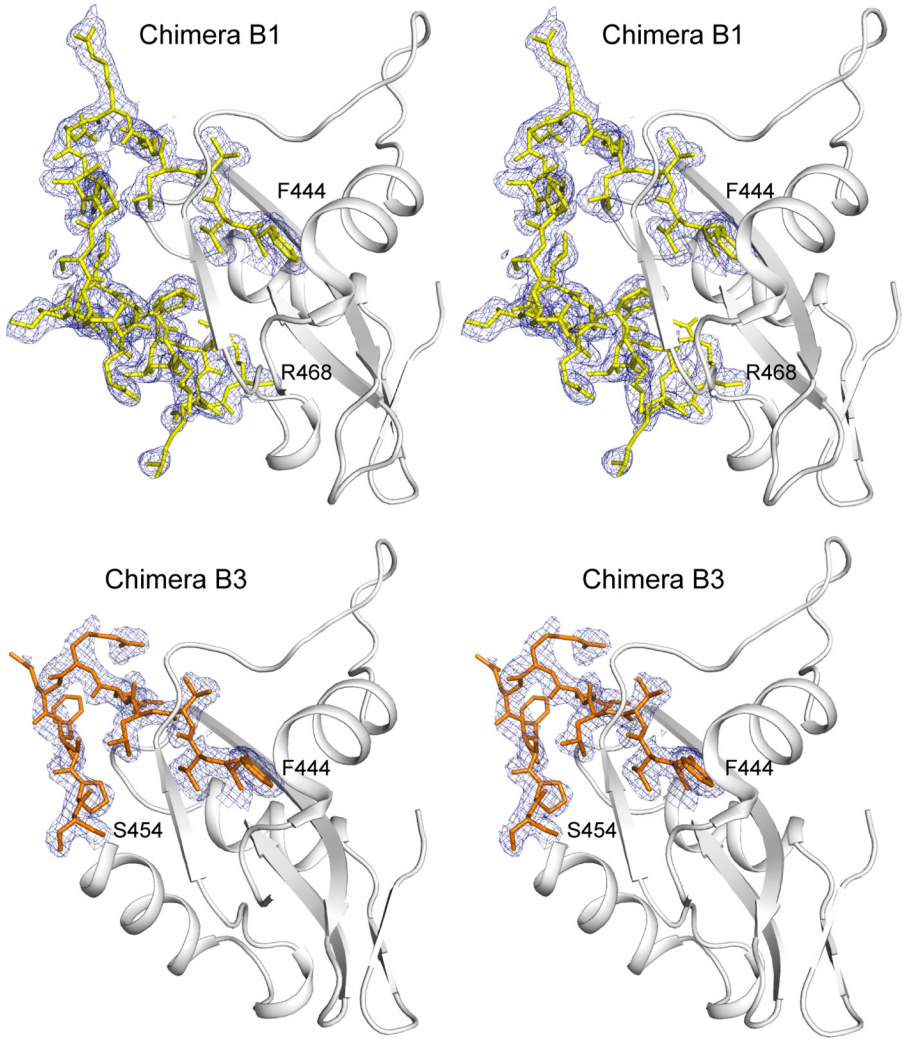

FIGURE 5. **Stereo view of the C-terminal domains of *Rp*MatB-*Bx*BclM B1 chimera and *Rp*MatB-*Bx*BclM B3 chimera.** The chimeric regions are depicted as *sticks* superimposed on an electron density map calculated with coefficients of the form $2F_o - F_c$ contoured at 1σ. Density for the backbone residues is well defined, although density for side chains that do not make crystal contacts is weak (Val-465 and MLY 466 in the B1 chimera and Arg-449, Glu-450, and Phe-451 in the B3 chimera). Figure was prepared with PyMOL (60).

backbone conformation of the loop but rather the disposition of amino acid side chains that is important for *Rp*Pat recognition.

The *Rp*MatB-*Bx*BclM B3 chimera has nine residues that differ from wild-type *Rp*MatB. The first is a phenylalanine in place of valine at position 444, within the β-strand immediately preceding the chimera loop, which appears to be required for stability of the chimera. Two conserved valine residues and eight nonconserved residues follow this phenylalanine within the chimera loop that differs between *Rp*MatB and the *Rp*MatB-*Bx*BclM B3 chimera (Fig. 6*A*). When the eight nonconserved residues were changed, two at a time, to the corresponding *Rp*MatB sequence, the amount of acetylation decreased substantially (Fig. 6, *B* and *C*). The largest effect was seen when the first two pairs of amino acids were changed (chimeras *Rp*MatB-*Bx*BclM B4 and B5, changes in residues Leu-447 to Glu-450 of the chimera), which were each acetylated about 5-fold less than the *Rp*MatB-*Bx*BclM B1 chimera. These results showed that changes within the chimera loop strongly affected how well *Rp*Pat recognized the *Rp*MatB-*Bx*BclM chimeras.

Both the *Rp*MatB-*Bx*BclM B1 and B3 chimeras crystallized in the adenylation conformation, in which the acetylated lysine residue (Lys-488) is buried in the active site cleft between the N- and C-terminal domains. We have previously reported structures of *Rp*MatB in the adenylation and apo conformations, and we generated a model of the thioester-forming conformation based on the structure of *Streptomyces coelicolor* MatB (36, 54). Fig. 7 shows a stereo view of the B3 chimera modeled in all three conformations, with the Lys-488 side chain modeled from the MatB apo structure. It is likely that *Rp*Pat interacts with acyl-CoA synthetases when they are in the thioester-forming conformation, as the active site lysine residue is most accessible in this state; however, modification may also be possible in the apo conformation.

From the model of the *Rp*MatB-*Bx*BclM B3 chimera in the thioester-forming conformation (Fig. 8*A*), it seems plausible that *Rp*Pat would interact with both Lys-488 in the active site loop and the tip of the chimera loop, because both loops protrude from the same face of the C-terminal domain. This information also suggested that *Rp*Pat recognized a large section of *Rp*MatB as the two most extreme points of these loops (the hydroxyl of Ser-454 and methylene group of Met-486) were 35 Å apart within the crystal structure.

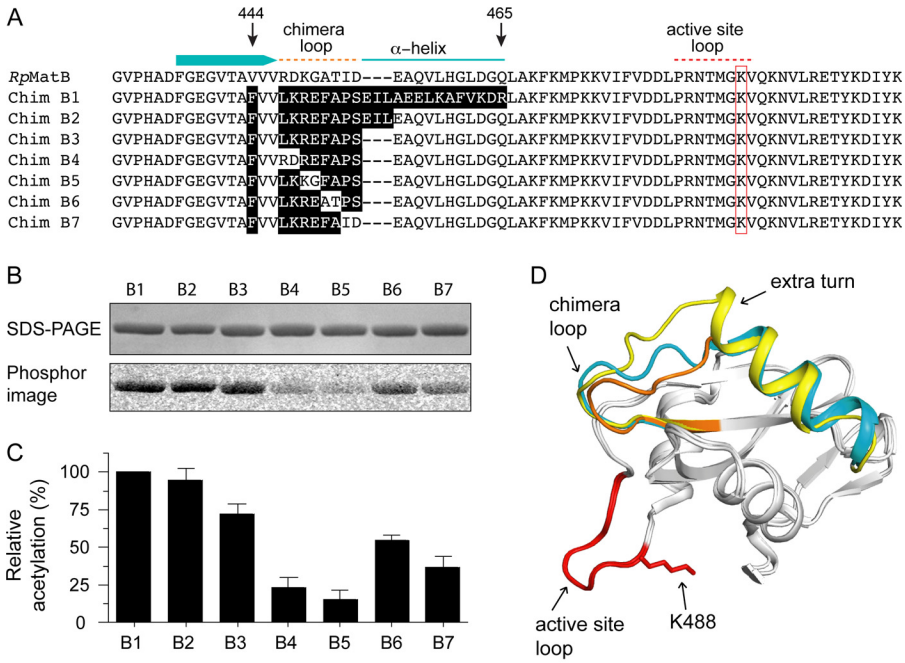

FIGURE 6. **Construction of a minimal *Rp*MatB-*Bx*BclM chimera.** *A,* alignment of the C-terminal end of *Rp*MatB, starting at residues 431, with the *Rp*MatB-*Bx*BclM chimeras (labeled *B1–B7*). Relevant secondary structure elements of *Rp*MatB are shown *above* the alignment, and the *red box* indicates the active site lysine (Lys-488). Residues in the chimeras derived from *Bx*BclM are highlighted in *black*. *B,* acetylation of each *Bx*BclM chimera by *Rp*Pat using [1-^{14}C]acetyl-CoA. *C,* amount of acetylation in *B* was quantified and normalized to the total acetylation of the *Rp*MatB-*Bx*BclM B1 chimera. Values represent averages and standard deviations of three experiments. *D,* alignment of the C-terminal domains of *Rp*MatB and the *Rp*MatB-*Bx*BclM B1 and B3 chimeras. *Bx*BclM-derived residues in the B1 chimera are indicated in *yellow*, *Bx*BclM-derived residues in the B3 chimera are shown in *orange*, and the corresponding residues in wild-type *Rp*MatB are shown in *cyan*. In all three structures, the catalytic loop and active site lysine are shown in *red* (lysine residue is modeled from the *Rp*MatB apo structure). *D* was prepared with PyMOL (60).

Fig. 8*B* shows calculated surface electrostatic potentials for the C-terminal domains of MatB and the *Rp*MatB-*Bx*BclM B3 chimera. Both proteins have positively charged regions surrounding the active site lysine residue and then a negatively charged band in the upper part of the C-terminal domain. Interestingly, the *Rp*MatB-*Bx*BclM B3 chimera has another pronounced positively charged region consisting of Lys-448 and Arg-449 within the chimera loop, whereas the wild-type *Rp*MatB structure is predominantly electronegative throughout the chimera loop region. These surface representations show substantial differences in both surface shape and charge in the chimera loop that may explain how *Rp*Pat can recognize the *Rp*MatB-*Bx*BclM B3 chimera but not wild-type *Rp*MatB.

DISCUSSION

A Loop ~20 Å Away from the Acetylation Site Is Critical for RpPat Recognition of Its Protein Substrate—Previous studies have shown that *Rp*Pat acetylates at least 10 different acyl-CoA synthetases in *R. palustris*, all of which contain a conserved PX_4GK motif within the active site loop (13, 14). Notably, this motif is not sufficient for acetylation. For example, *Rp*MatB is not a substrate of *Rp*Pat even though it shares extensive homology with other acyl-CoA synthetases and does contain the PX_4GK motif preceding the catalytic lysine residue. The key difference in *Rp*MatB appears to be the shape and electrostatic potential of a loop region located ~20 Å away from the conserved active site lysine residue and >20 amino acids away in the protein sequence. Such a structure appears crucial for acetylation (Figs. 2 and 6). This loop region found in *bona fide* substrates of *Rp*Pat can replace the corresponding sequence of *Rp*MatB to yield catalytically active *Rp*MatB variants that are now subject to acetylation control (Figs. 2 and 3).

At present, the three-dimensional structure of *Rp*Pat or its homologues has not been reported. Hence, we can only speculate how *Rp*Pat may interact with its substrates. Results reported herein suggest that *Rp*Pat interacts with a large surface area of its substrates (Fig. 8). It seems possible that Pat simultaneously interacts with the active site lysine residue and the loop region identified by this work and that *Rp*Pat may require these two positively charged loop regions for substrate recognition. These results demonstrate that *Rp*Pat, and perhaps other Pat homologues, shows a great deal of specificity among acyl-CoA synthetases. These results further support recently reported data (14) that challenged suggestions by others that Pat homologues in *E. coli* and *S. enterica* acetylate a wide range of structurally diverse substrates (17, 55).

Insights into the Physiological Role of RpPat—*Rp*MatB appears to function in methyl malonate metabolism in *R. palustris*, where the ultimate product of methyl malonate catabolism is likely the tricarboxylic acid cycle intermediate succinyl-CoA (36). Many, if not most, of the AMP-forming acyl-CoA synthetases that are regulated by *Rp*Pat either activate acetate or propionate directly to their corresponding acyl-CoA derivative or initiate the breakdown of longer chain fatty acids or aromatic acids to acetyl-CoA. It is plausible that *Rp*Pat is a sensor of acetyl-CoA or propionyl-CoA concentrations in the cell and uses these compounds to turn off the acyl-CoA synthetases at

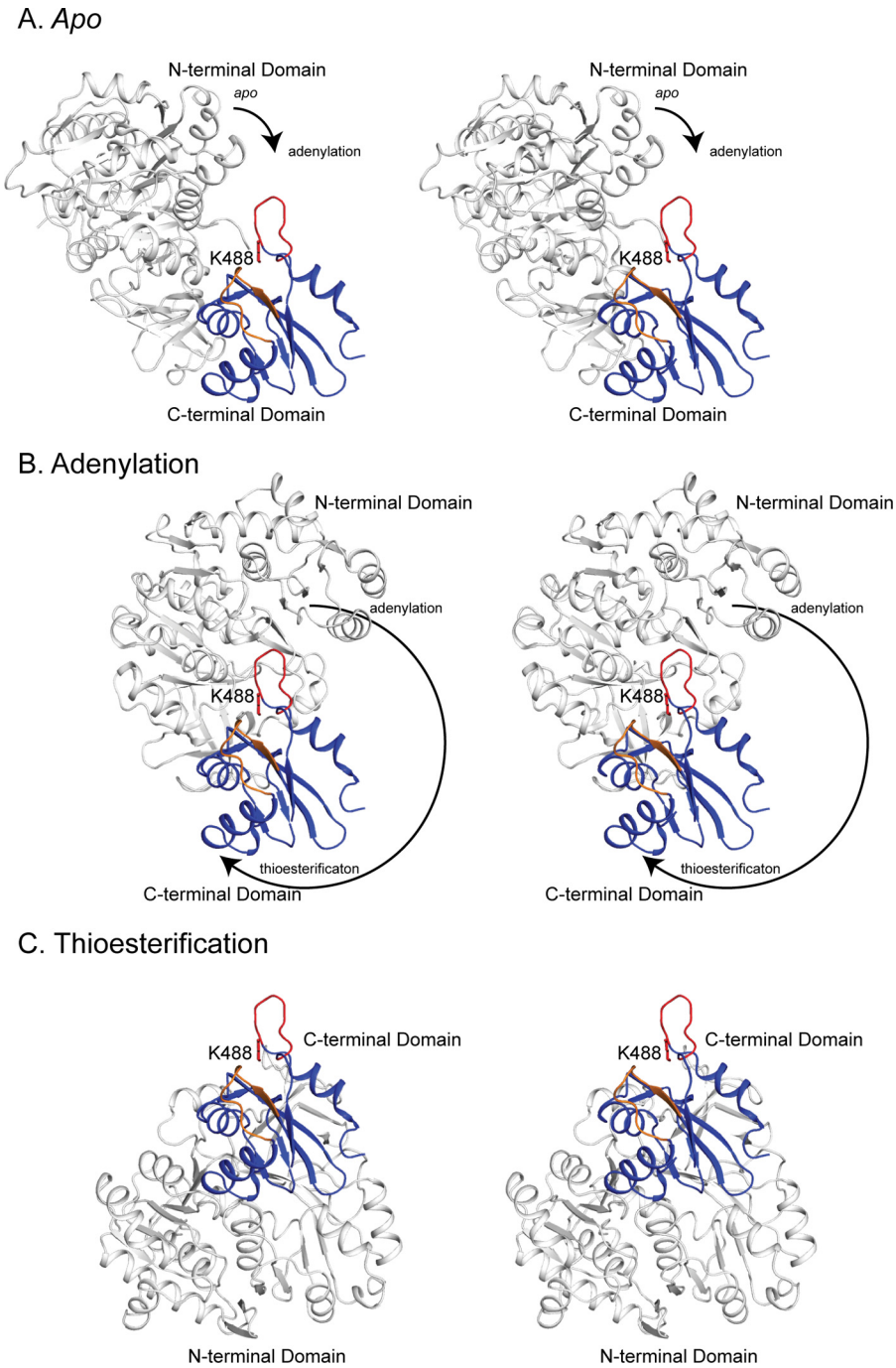

FIGURE 7. **Stereo view of the accessibility of Lys-488 in the apo (*A*), adenylation (*B*), and thioesterification (*C*) conformations.** The N-terminal domain is shown in *white* and the C-terminal domain in *blue*. The active site loop and Lys-488 are shown in *red,* and the chimera loop is shown in *orange*. The Lys-488 side chain is modeled from the MatB apo structure. The C-terminal domains are shown in the same orientation in all three conformations. The *arrows* indicate the direction and magnitude of rotation of the N-terminal domain relative to the C-terminal domain during the transition from one conformation to the next. In all cases, the C-terminal domain is from the *Rp*MatB-*Bx*BclM B3 chimera structure, and the N-terminal domains are either from wild-type MatB (apo and thioesterification, PDB accession code 4FUQ) or MatB^{K488A} (adenylation, PDB accession code 4FUT). In each case the N- and C-terminal domains are aligned to *Rp*MatB structures (apo, 4FUQ; adenylation, 4FUT) or to the *S. coelicolor* MatB structure (thioesterification, 3NYQ (54)). Figure was prepared with PyMOL (60).

the beginning of these catabolic pathways. If this were the case, then we would hypothesize that *Rp*Pat only modifies acyl-CoA synthetases involved in acetyl-CoA or propionyl-CoA generation. One example of an acyl-CoA synthetase not involved in acetyl- or propionyl-CoA generation in *R. palustris* is *Rp*MatB, which we have shown is not acetylated. Similarly, acyl-CoA synthetases involved in biosynthesis of natural products such as siderophores and antibiotics would not be regulated by *Rp*Pat. This hypothesis awaits further experimental work.

In R. palustris MatB May Have Evolved to Evade Inactivation by RpPat—*Rp*MatB is the first example to date of an acyl-CoA synthetase that is not regulated by *Rp*Pat in *R. palustris*. Changes in a relatively nonconserved loop region distant from the acetylation site greatly affect is susceptibility to acetylation,

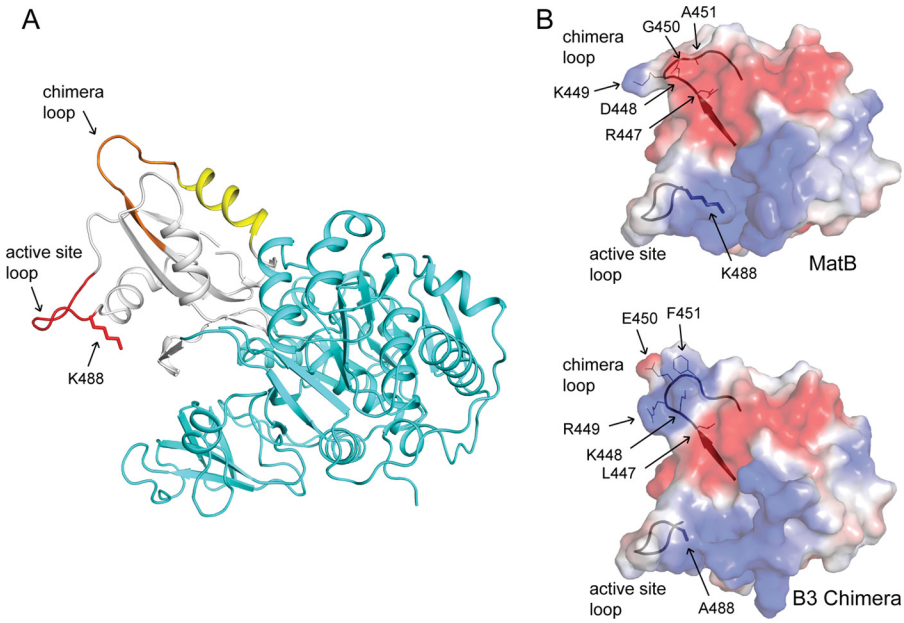

FIGURE 8. ***Rp*Pat recognition elements of the *Rp*MatB-*Bx*BclM B3 chimera.** *A, Rp*MatB-*Bx*BclM B3 chimera modeled in the thioester-forming conformation. The N-terminal domain is shown in *cyan*, the C-terminal domain in *white*, the active site loop in *red*, the residues derived from *Bx*BclM in *orange*, and additional residues changed in the *Rp*MatB-*Bx*BclM B1 chimera are shown in *yellow*. The acetylation site, Lys-488, was modeled from the *Rp*MatB apo structure. *B,* surface electrostatic calculations for the *Rp*MatB and *Rp*MatB-*Bx*BclM B3 chimera C-terminal domains were generated using PyMOL (60), with electropositive regions shown in *blue* and electronegative regions shown in *red*. *Ribbon* diagrams of the active site and chimera loop backbones are shown for reference, and residues within the chimera loop are labeled. The active site lysine residue (Lys-488) was mutated to an alanine in the *Rp*MatB-*Bx*BclM B3 chimera structure to aid crystallization. Figure was prepared with PyMOL (60).

suggesting that *Rp*MatB evolved to escape regulation by *Rp*Pat. A variant of *S. enterica* acetyl-CoA synthetase that is not susceptible to acetylation has also been generated by changing one residue near the C terminus of the enzyme from leucine to proline (56). Together, these findings suggest that there may be many ways for these enzymes to evolve such that they retain their enzymatic activity but cannot be modified by Pat. In addition, some microbes do not have Pat homologues, suggesting that they either have alternative acetyltransferases that have the same role (such as *Bacillus subtilis* (57), *Mycobacterium tuberculosis* (58), and *Sulfolobus solfataricus* (59)) or that these microbes do not regulate acyl-CoA synthetases by acetylation.

Closing Remarks—From a practical perspective, our results indicate that it is difficult to predict, on the basis of primary sequence alone, whether or not an AMP-forming acyl-CoA synthetase is regulated by acetylation. *Rp*MatB is the first example of an acyl-CoA synthetase that contains the PX_4GK motif within the active site loop that is not acetylated by *Rp*Pat, although there may be many others. To make the computational identification of *Rp*Pat substrates more difficult, there appears to be little sequence conservation within the "chimera loop" identified herein as needed for acetylation. Thus, the newly discovered region does not provide needed homology to facilitate the identification of acyl-CoA synthetases that are under acetylation control. As shown by this work, more research is needed to define the true extent of the bacterial lysine acetylome and to establish how many proteins are true substrates of *Rp*Pat and its homologues in other bacteria. Evidence is mounting that *Rp*Pat specializes in acetylating a subset of acyl-CoA synthetases in *R. palustris*, and it will be interesting to see if this holds true for Pat homologues in other organisms.

Acknowledgments—Use of the Advanced Photon Source, an Office of Science User Facility operated for the United States Department of Energy Office of Science by Argonne National Laboratory, was supported by the United States Department of Energy under Contract DE-AC02-06CH11357.

REFERENCES

1. Kim, G. W., and Yang, X. J. (2011) Comprehensive lysine acetylomes emerging from bacteria to humans. *Trends. Biochem. Sci.* **36,** 211–220
2. Thao, S., and Escalante-Semerena, J. C. (2011) Control of protein function by reversible N^{ϵ}-lysine acetylation in bacteria. *Curr. Opin. Microbiol.* **14,** 200–204
3. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007) How chromatin-binding modules interpret histone modifications. Lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **14,** 1025–1040
4. Oberdoerffer, P., Michan, S., McVay, M., Mostoslavsky, R., Vann, J., Park, S. K., Hartlerode, A., Stegmuller, J., Hafner, A., Loerch, P., Wright, S. M., Mills, K. D., Bonni, A., Yankner, B. A., Scully, R., Prolla, T. A., Alt, F. W., and Sinclair, D. A. (2008) SIRT1 redistribution on chromatin promotes genomic stability but alters gene expression during aging. *Cell* **135,** 907–918
5. Wang, L., Tang, Y., Cole, P. A., and Marmorstein, R. (2008) Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases. Implications for histone acetyltransferase evolution and function. *Curr. Opin. Struct. Biol.* **18,** 741–747
6. Berndsen, C. E., and Denu, J. M. (2008) Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr. Opin. Struct. Biol.* **18,** 682–689
7. Vetting, M. W., S de Carvalho, L. P., Yu, M., Hegde, S. S., Magnet, S., Roderick, S. L., and Blanchard, J. S. (2005) Structure and functions of the

GNAT superfamily of acetyltransferases. *Arch. Biochem. Biophys.* **433,** 212–226
8. Starai, V. J., and Escalante-Semerena, J. C. (2004) Identification of the protein acetyltransferase (Pat) enzyme that acetylates acetyl-CoA synthetase in *Salmonella enterica. J. Mol. Biol.* **340,** 1005–1012
9. Liang, W., and Deutscher, M. P. (2012) Post-translational modification of RNase R is regulated by stress-dependent reduction in the acetylating enzyme Pka (YfiQ). *RNA* **18,** 37–41
10. Starai, V. J., and Escalante-Semerena, J. C. (2004) Acetyl-coenzyme A synthetase (AMP forming). *Cell. Mol. Life Sci.* **61,** 2020–2030
11. Starai, V. J., Celic, I., Cole, R. N., Boeke, J. D., and Escalante-Semerena, J. C. (2002) Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. *Science* **298,** 2390–2392
12. Starai, V. J., Takahashi, H., Boeke, J. D., and Escalante-Semerena, J. C. (2003) Short-chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in *Salmonella enterica* and *Saccharomyces cerevisiae. Genetics* **163,** 545–555
13. Crosby, H. A., Heiniger, E. K., Harwood, C. S., and Escalante-Semerena, J. C. (2010) Reversible N^{ϵ}-lysine acetylation regulates the activity of acyl-CoA synthetases involved in anaerobic benzoate catabolism in *Rhodopseudomonas palustris. Mol. Microbiol.* **76,** 874–888
14. Crosby, H. A., Pelletier, D. A., Hurst, G. B., and Escalante-Semerena, J. C. (2012) System-wide studies of *N*-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *J. Biol. Chem.* **287,** 15590–15601
15. Garrity, J., Gardner, J. G., Hawse, W., Wolberger, C., and Escalante-Semerena, J. C. (2007) *N*-Lysine propionylation controls the activity of propionyl-CoA synthetase. *J. Biol. Chem.* **282,** 30239–30245
16. Liang, W., Malhotra, A., and Deutscher, M. P. (2011) Acetylation regulates the stability of a bacterial protein. Growth stage-dependent modification of RNase R. *Mol. Cell* **44,** 160–166
17. Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., Li, H., Xie, L., Zhao, W., Yao, Y., Ning, Z. B., Zeng, R., Xiong, Y., Guan, K. L., Zhao, S., and Zhao, G. P. (2010) Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **327,** 1004–1007
18. Bertani, G. (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli. J. Bacteriol.* **62,** 293–300
19. Bertani, G. (2004) Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *J. Bacteriol.* **186,** 595–600
20. Larimer, F. W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L., Land, M. L., Pelletier, D. A., Beatty, J. T., Lang, A. S., Tabita, F. R., Gibson, J. L., Hanson, T. E., Bobst, C., Torres, J. L., Peres, C., Harrison, F. H., Gibson, J., and Harwood, C. S. (2004) Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris. Nat. Biotechnol.* **22,** 55–61
21. Kim, M.-K., and Harwood, C. S. (1991) Regulation of benzoate-CoA ligase in *Rhodopseudomonas palustris. FEMS Microbiol. Lett.* **83,** 199–203
22. Balch, W. E., and Wolfe, R. S. (1976) New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressurized atmosphere. *Appl. Environ. Microbiol.* **32,** 781–791
23. Moore, D. D., and Dowhan, D. (2002) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., eds) pp. 2.0.1–2.12.17, John Wiley & Sons, Inc., Hoboken, NJ
24. Rocco, C. J., Dennison, K. L., Klenchin, V. A., Rayment, I., and Escalante-Semerena, J. C. (2008) Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli. Plasmid* **59,** 231–237
25. Blommel, P. G., and Fox, B. G. (2007) A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expr. Purif.* **55,** 53–68
26. Klock, H. E., Koesema, E. J., Knuth, M. W., and Lesley, S. A. (2008) Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. *Proteins* **71,** 982–994
27. Horton, R. M., Ho, S. N., Pullen, J. K., Hunt, H. D., Cai, Z., and Pease, L. R. (1993) Gene splicing by overlap extension. *Methods Enzymol.* **217,** 270–279
28. Miroux, B., and Walker, J. E. (1996) Overproduction of proteins in *Escherichia coli.* Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260,** 289–298
29. Markowitz, V. M., Korzeniewski, F., Palaniappan, K., Szeto, E., Werner, G., Padki, A., Zhao, X., Dubchak, I., Hugenholtz, P., Anderson, I., Lykidis, A., Mavromatis, K., Ivanova, N., and Kyrpides, N. C. (2006) The integrated microbial genomes (IMG) system. *Nucleic Acids Res.* **34,** D344–D348
30. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003) ExPASy. The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31,** 3784–3788
31. Rayment, I. (1997) Reductive alkylation of lysine residues to alter crystallization properties of proteins. *Methods Enzymol.* **276,** 171–179
32. Reger, A. S., Carney, J. M., and Gulick, A. M. (2007) Biochemical and crystallographic analysis of substrate binding and conformational changes in acetyl-CoA synthetase. *Biochemistry* **46,** 6536–6546
33. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326
34. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674
35. Collaborative Computational Project No. 4 (1994) The CCP4 suite. Programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
36. Crosby, H. A., Rank, K. C., Rayment, I., and Escalante-Semerena, J. C. (2012) Structure-guided expansion of the substrate range of methylmalonyl-CoA synthetase (MatB) of *Rhodopseudomonas palustris. Appl. Environ. Microbiol.* **78,** 6619–6629
37. Emsley, P., and Cowtan, K. (2004) Coot. Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
38. Skubák, P., Murshudov, G. N., and Pannu, N. S. (2004) Direct incorporation of experimental phase information in model refinement. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2196–2201
39. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX. Building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1948–1954
40. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2256–2268
41. Gulick, A. M., Starai, V. J., Horswill, A. R., Homick, K. M., and Escalante-Semerena, J. C. (2003) The 1.75 Å crystal structure of acetyl-CoA synthetase bound to adenosine-5′-propylphosphate and coenzyme A. *Biochemistry* **42,** 2866–2873
42. Gulick, A. M. (2009) Conformational dynamics in the acyl-CoA synthetases, adenylation domains of nonribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol.* **4,** 811–827
43. Reger, A. S., Wu, R., Dunaway-Mariano, D., and Gulick, A. M. (2008) Structural characterization of a 140° domain movement in the two-step reaction catalyzed by 4-chlorobenzoate:CoA ligase. *Biochemistry* **47,** 8016–8025
44. Kochan, G., Pilka, E. S., von Delft, F., Oppermann, U., and Yue, W. W. (2009) Structural snapshots for the conformation-dependent catalysis by human medium-chain acyl-coenzyme A synthetase ACSM2A. *J. Mol. Biol.* **388,** 997–1008
45. Bains, J., and Boulanger, M. J. (2007) Biochemical and structural characterization of the paralogous benzoate CoA ligases from *Burkholderia xenovorans* LB400. Defining the entry point into the novel benzoate oxidation (box) pathway. *J. Mol. Biol.* **373,** 965–977
46. D'Arcy, A., Stihle, M., Kostrewa, D., and Dale, G. (1999) Crystal engineering. A case study using the 24-kDa fragment of the DNA gyrase B subunit from *Escherichia coli. Acta Crystallogr. D Biol. Crystallogr.* **55,** 1623–1625
47. McElroy, H. E., Sisson, G. W., Schoettlin, W. E., Aust, R. M., and Villafranca, J. E. (1992) Studies on engineering crystallizability by mutation of surface residues of human thymidylate synthase. *J. Cryst. Growth* **122,** 265–272
48. Rypniewski, W. R., Holden, H. M., and Rayment, I. (1993) Structural con-

sequences of reductive methylation of lysine residues in hen egg white lysozyme. An X-ray analysis at 1.8-Å resolution. *Biochemistry* **32,** 9851–9858
49. Kim, Y., Quartey, P., Li, H., Volkart, L., Hatzos, C., Chang, C., Nocek, B., Cuff, M., Osipiuk, J., Tan, K., Fan, Y., Bigelow, L., Maltseva, N., Wu, R., Borovilos, M., Duggan, E., Zhou, M., Binkowski, T. A., Zhang, R. G., and Joachimiak, A. (2008) Large-scale evaluation of protein reductive methylation for improving protein crystallization. *Nat. Methods* **5,** 853–854
50. Kobayashi, M., Kubota, M., and Matsuura, Y. (1999) Crystallization and improvement of crystal quality for x-ray diffraction of maltooligosyl trehalose synthase by reductive methylation of lysine residues. *Acta Crystallogr. D Biol. Crystallogr.* **55,** 931–933
51. Kurinov, I. V., and Uckun, F. M. (2003) High resolution x-ray structure of potent anti-HIV pokeweed antiviral protein-III. *Biochem. Pharmacol.* **65,** 1709–1717
52. Rayment, I., Rypniewski, W. R., Schmidt-Bäse, K., Smith, R., Tomchick, D. R., Benning, M. M., Winkelmann, D. A., Wesenberg, G., and Holden, H. M. (1993) Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science* **261,** 50–58
53. Schubot, F. D., and Waugh, D. S. (2004) A pivotal role for reductive methylation in the *de novo* crystallization of a ternary complex composed of *Yersinia pestis* virulence factors YopN, SycN, and YscB. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 1981–1986
54. Hughes, A. J., and Keatinge-Clay, A. (2011) Enzymatic extender unit generation for *in vitro* polyketide synthase reactions. Structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem. Biol.* **18,** 165–176
55. Lima, B. P., Antelmann, H., Gronau, K., Chi, B. K., Becher, D., Brinsmade, S. R., and Wolfe, A. J. (2011) Involvement of protein acetylation in glucose-induced transcription of a stress-responsive promoter. *Mol. Microbiol.* **81,** 1190–1204
56. Starai, V. J., Gardner, J. G., and Escalante-Semerena, J. C. (2005) Residue Leu-641 of acetyl-CoA synthetase is critical for the acetylation of residue Lys-609 by the protein acetyltransferase enzyme of *Salmonella enterica*. *J. Biol. Chem.* **280,** 26200–26205
57. Gardner, J. G., Grundy, F. J., Henkin, T. M., and Escalante-Semerena, J. C. (2006) Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD^+ involvement in *Bacillus subtilis*. *J. Bacteriol.* **188,** 5460–5468
58. Nambi, S., Basu, N., and Visweswariah, S. S. (2010) Cyclic AMP-regulated protein lysine acetylases in mycobacteria. *J. Biol. Chem.* **285,** 24313–24323
59. Brent, M. M., Iwata, A., Carten, J., Zhao, K., and Marmorstein, R. (2009) Structure and biochemical characterization of protein acetyltransferase from *Sulfolobus solfataricus*. *J. Biol. Chem.* **284,** 19412–19419
60. Delano, W. L. (2002) *The PyMOL Molecular Graphics System,* DeLano Scientific LLC, San Carlos, CA