

Molecular structure of a 5,10-methylenetetrahydrofolate dehydrogenase from the silkworm *Bombyx mori*

Mohammad R. Haque¹, Akifumi Higashiura^{2,†}, Atsushi Nakagawa², Aiko Hirowatari¹, Shigeki Furuya¹ and Kohji Yamamoto¹

¹ Department of Bioscience and Biotechnology, Kyushu University Graduate School, Fukuoka, Japan

² Institute for Protein Research, Osaka University, Suita, Japan

Keywords

5,10-methylenetetrahydrofolate dehydrogenase; *Bombyx mori*; crystal structure; NADP; serine

Correspondence

K. Yamamoto, Department of Bioscience and Biotechnology, Kyushu University Graduate School, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan
Tel: +81-92-621-4991
E-mail: yamamok@agr.kyushu-u.ac.jp

†Present address

Department of Virology, Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan

(Received 16 December 2018, revised 27 December 2018, accepted 14 January 2019)

doi:10.1002/2211-5463.12595

The enzyme 5,10-methylenetetrahydrofolate dehydrogenase (MTHFD) is essential for the production of certain amino acids (glycine, serine, and methionine) and nucleic acids (thymidylate and purine). Here, we identified a cDNA encoding this enzyme from the silkworm *Bombyx mori*. The recombinant *B. mori* MTHFD (bmMTHFD) expressed in *Escherichia coli* recognized 5,10-methylenetetrahydrofolate and 5,10-methenyltetrahydrofolate as substrate in the presence of NADP⁺ as well as NAD⁺. The bmMTHFD structure was determined at a resolution of 1.75 Å by X-ray crystallography. Site-directed mutagenesis indicated that the amino acid residue Tyr49 contributed to its catalytic activity. Our findings provide insight into the mechanism underlying the activity of MTHFD from *B. mori* and potentially other insects and may therefore facilitate the development of inhibitors specific to MTHFD as insecticides.

One-carbon metabolism is involved in the synthesis of amino acids like alanine, glycine, and serine, as well as purine and pyrimidine bases [1,2]. 5,10-methylenetetrahydrofolate (MTHF) dehydrogenase (MTHFD; EC 1.5.1.5) is an enzyme involved in one-carbon metabolism in eukaryotes [2]. To date, cytosolic and mitochondrial versions of this protein have been reported. For example, MTHFD1 is a cytoplasmic protein [3], whereas MTHFD1L, MTHFD2, and MTHFD2L are mitochondrial proteins [4–6]. Human MTHFD1 is a trifunctional enzyme with dehydrogenase (D), cyclohydrolase (C), and synthetase (S) activities that catalyze the oxidation of MTHF to 5,10-methenyl-THF, which is then

hydrolyzed to 10-formyl-THF, and finally converted to THF and formate [3]. The 3D structure of the D/C domain of MTHFD1, referred to as DC301, has been reported [3]. MTHFD2 and MTHFD2L are bifunctional enzymes [7,8], whereas MTHFD1L is a monofunctional enzyme [9]. MTHFD often requires NADP⁺ or NAD⁺ as the cofactor for their activity. MTHFD1 requires NADP⁺ [3], MTHFD2 and MTHFD2L use either NADP⁺ or NAD⁺ [7,8], whereas MTHFD1L is monofunctional with only S activity and does not use either cofactors [9]. Likewise, the prokaryotic MTHFD of *Escherichia coli* is a bifunctional enzyme that uses NADP⁺ [10], and the monofunctional enzyme of

Abbreviations

bmMTHFD, *Bombyx mori* 5,10-methylenetetrahydrofolate dehydrogenase; C, cyclohydrolase; D, dehydrogenase; MTHF, 5,10-methylenetetrahydrofolate; MTHFD, 5,10-methylenetetrahydrofolate dehydrogenase; S, synthetase; WT, wild-type.

Peptostreptococcus productus requires NADP⁺ as the cofactor [11]. Although one-carbon metabolism has been studied in vertebrates, there are no reports from invertebrates, including silkworm and other insects.

To characterize one-carbon metabolism in insects, we isolated mRNA encoding an MTHFD of the silkworm *Bombyx mori* MTHF dehydrogenase (bmMTHFD), which is an important lepidopteran insect model. The structure–function relationships of insect MTHFDs have not been studied in detail. Since many agricultural pests are lepidopteran insects, it is useful to investigate the amino acid residues present in the active site of bmMTHFD. Further, because MTHFD is involved in the synthesis of important biomolecules such as amino acids and purine and pyrimidine bases, the inhibitors could be effective insecticides against agricultural pests. Here, we determined the three-dimensional structure of bmMTHFD to identify the amino acid residues important for bmMTHFD activity and conducted mutation analysis of bmMTHFD to determine the role of the amino acids lining the substrate-binding site. Examination of bmMTHFD catalytic activity indicated that it participates in the D and C activities. The active site in bmMTHFD was then determined to better understand the structural basis for this conversion. As described, mammalian MTHFDs are key enzymes involved in the synthesis of amino acids and purine and pyrimidine bases, which are crucial biomaterials for survival. Analysis of inhibition of insect MTHFDs would aid in the design of pesticides and insecticides. The crystal structure of bmMTHFD and the identification of the amino acid residues involved in catalytic function in the current study may provide insights into the mechanism underlying MTHFD activity and could facilitate the development of inhibitors specific to MTHFD as insecticides. To the best of our knowledge, this study is the first to report on MTHFD in insects.

Materials and methods

Insects

Bombyx mori larvae (p50T strain) were reared at the Kyushu University Graduate School (Fukuoka, Japan) and fed mulberry leaves. Day-3 fifth-instar larvae were dissected on ice, and fat body was stored at –80 °C until use.

RNA extraction, cloning, and sequencing of cDNA encoding bmMTHFD

Total RNA was isolated from the fat body using RNeasy Plus Mini Kit (Qiagen, Hilden, Germany) and was

analyzed by reverse transcription–PCR. First-strand cDNA was obtained using SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA, USA) and an oligo-dT primer. The resulting cDNA was used as a PCR template with the following oligonucleotide primers: 5'-CAACAGCCATATG GCGCGTATCCTCGATGG-3' (sense) and 5'-CCGGATC CTTAATTGGATTTGTTTGCTTGA-3' (antisense). The primer designs were based on a partial sequence obtained from the SilkBase database (<http://silkbases.ab.a.u-tokyo.ac.jp/cgi-bin/index.cgi>). The underlined and double-underlined regions indicate NdeI and BamHI restriction enzyme sites, respectively, which were used for insertion of the PCR product into an expression vector. The PCR program was as follows: 94 °C for 2 min, 35 cycles of 94 °C for 1 min, 59 °C for 1 min, and 72 °C for 2 min, and 72 °C for 10 min. The bmMTHFD cDNA (*bmmthfd*) was ligated into the pGEM-T Easy Vector (Promega, Madison, WI, USA) and transformed into *E. coli* DH5 α .

To obtain the complete sequence of *bmmthfd* and to deduce its amino acid sequence, the GENETYX-MAC software (ver. 14.0.12; GENETYX Corporation, Tokyo, Japan) was used. Homology alignment was performed using CLUSTALW (ver. 1.83; DNA Data Bank of Japan, Shizuoka, Japan), with 10 and 0.2 as the gap creation penalty and gap extension, respectively. A phylogenetic tree was generated using neighbor-joining plot software (<http://d-oua.prabi.fr/software/njplot>).

Overexpression and purification of recombinant protein

The *bmmthfd* clone was digested with NdeI and BamHI, subcloned into the expression vector pET-15b (Merck Millipore, Darmstadt, Germany), and transformed into competent *E. coli* Rosetta (DE3) pLysS cells (Merck Millipore). The cells were then grown at 37 °C in Luria–Bertani media containing 100 $\mu\text{g}\cdot\text{mL}^{-1}$ ampicillin. After the cell density reached an OD₆₀₀ of 0.7, isopropyl-1-thio- β -D-galactoside was added to a final concentration of 1 mM to induce recombinant protein production. The culture was further incubated for 3 h, and the cells were harvested by centrifugation at 10 000 *g* for 15 min. Bacteria were resuspended in phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 1.76 mM KH₂PO₄, pH 7.4) containing 4 mg·mL⁻¹ lysozyme and were subsequently disrupted by sonication. Sonication was performed three times for 1 min each on ice using a 3-mm tapered microtip probe. The supernatant containing the recombinant protein was clarified by centrifugation at 10 000 *g* for 15 min and subjected to Ni²⁺-affinity chromatography equilibrated with the same buffer. After washing with the same buffer, the samples were eluted with a linear gradient of 0–0.5 M imidazole. The enzyme-containing fractions, assayed as described below, were pooled, concentrated using a centrifugal filter (10 000 molecular weight cutoff; Millipore Corp., Billerica, MA, USA), and

applied to a Superdex 200 column (GE Healthcare Bio-Sciences, Buckinghamshire, UK) equilibrated with 20 mM Tris/HCl buffer (pH 8.0), with the addition of 0.2 M NaCl. The purity of the pooled material was analyzed by sodium dodecyl sulfate (SDS)/PAGE performed using a 15% polyacrylamide slab gel containing 0.1% SDS, according to the method reported by Laemmli [12]. Protein bands were visualized by Coomassie Brilliant Blue R-250 (Sigma-Aldrich, St. Louis, MO, USA) staining.

Protein crystallization and X-ray diffraction data analysis

Recombinant bmMTHFD was purified as described in 'Overexpression and purification of recombinant protein' and then concentrated using a centrifugal filter (Millipore) to 10 mg·mL⁻¹ in 20 mM Tris/HCl buffer, pH 8.0, containing 0.2 M NaCl. Crystallization was performed using the sitting-drop vapor diffusion method at 20 °C with Crystal Screen Kits (Hampton Research, Aliso Viejo, CA, USA). Each drop was formed by mixing an equal (1 µL) or twofold greater volume (1 : 2 µL) of protein and reservoir solutions, respectively. Crystals suitable for X-ray analysis were grown for 4 weeks in 0.2 M ammonium sulfate, 0.1 M sodium cacodylate trihydrate, and 30% PEG 8000 (w/v). The crystal was transferred to reservoir solution with 25% (v/v) ethylene glycol using a cryoloop and flash frozen with liquid nitrogen before data collection. X-ray diffraction data collection was performed using synchrotron radiation on SPring-8 beamline BL44XU with a wavelength of 0.9 Å, in a nitrogen vapor stream at 100 K [13]. The data set was integrated and scaled using the DENZO and SCALEPACK programs, as implemented in the HKL2000 program package [14].

Determination of structure

The initial structure was determined using a molecular replacement method with the program PHENIX PHASER-MR [15], using the polyalanine model of *E. coli* MTHFD (PDB ID: 1B0A) as a search model. The ARP/WARP program [16] was used to build missing side-chain atoms and to add water molecules. After manual adjustment to the electron density map with the program COOT [17], refinement was performed using the program PHENIX.REFINE [18]. Figures were prepared using the PyMOL software (<http://py.mol.sourceforge.net>). The atomic coordinates and structure factor of bmMTHFD have been deposited in the Protein Data Bank (PDB ID: 5ZF1). The 3D structure alignment was performed using the MATRAS server [19].

Site-directed mutagenesis

Amino acid-substituted mutants of bmMTHFD were constructed using the QuikChange Site-Directed Mutagenesis

Kit (Agilent, Santa Clara, CA, USA), according to the previous method [20]. The mutagenesis primers were as follows: 5'-TAATTGTTGGCGATGACCCAGCGGCTCACACCTATGTCAGGAACAA-3' (sense) and 5'-TTGTTCTGACATAGGTGTGAGCCGCTGGGTCATCGCCAACAATTA-3' (antisense) for S46A, 5'-GATGACCCAGCGAGCCACACCGCTGTCAGGAACAAAGTTGAAGCCGCA-3' (sense) and 5'-TGCGGCTTCAACTTTGTTCTGACAGCGGTGTGGCTCGCTGGGTCATC-3' (antisense) for Y49A, 5'-GCCACACCTATGTCAGGAACGCTGTTGAAGCCGCAAATTTGT-3' (sense) and 5'-ACAAATTTTCGCGCTTCAACAGCGTTCCTGACATAGGTGTGGC-3' (antisense) for K53A, 5'-GTGTTGATGGAATTCTCGTCGCTTTACCTGTACCTGATACAAT-3' (sense) and 5'-ATTGTATCAGGTACAGGTAAAGCGACGAGAATGCCATCAACAC-3' (antisense) for Q97A, and 5'-ATTGCTCCTGAAAAAGATATCGCTGGTTTTCCACATAATCAACATC-3' (sense) and 5'-GATGTTGATTATGTGAAACCAGCGATATCTTTTCAGGAGCAAT-3' (antisense) for D120A. An expression plasmid containing *bmmthfd* cDNA was used as a template, and full-length mutated cDNAs were verified by DNA sequencing with 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA).

Measurements of enzyme activity

Dehydrogenase activity of bmMTHFD was spectrophotometrically measured using MTHF (Toronto Research Chemicals, North York, ON, Canada) as substrate and NADP⁺ or NAD⁺ as cofactor. MTHF, NADP⁺, and NAD⁺ were stored at -80 °C until use. The reaction mixture (100 µL) included 20 mM sodium phosphate buffer (pH 6.5), 30 mM 2-mercaptoethanol, and various concentrations of substrate, NADP⁺ or NAD⁺. The assay was initiated by addition of the enzyme and then incubated at 28 °C for 10 min. Production of methenyltetrahydrofolate was measured at 350 nm by an endpoint assay and a background rate determined in enzyme-free assays. The substrate concentrations were calculated using the extinction coefficient of 24.9 mm⁻¹·cm⁻¹ [21]. Kinetic parameters were measured with a nonlinear least squares data fit to the Michaelis-Menten equation using KALEIDAGRAPH (Synergy Software; HULINKS Inc., Tokyo, Japan).

Cyclohydrolase activity of bmMTHFD was assayed with 5,10-methenyl-THF (Santa Cruz Biotechnology, Inc., Dallas, TX, USA), that was stored at -80 °C until use. Reaction mixture (100 µL) contained 0.1 M potassium maleate (pH 7.4), 0.02 M 2-mercaptoethanol, and 0.1 mM 5,10-methenyl-THF. The reaction was started by addition of the enzyme and then incubated at 30 °C for 10 min. The assay was performed by measuring the decrease in absorption at 355 nm [22], and the values were compared to those from an enzyme-free assay.

Results

Sequence of cDNA encoding bmMTHFD

We isolated *bmmthfd* cDNA from *B. mori*, and the nucleotide sequence has been deposited in DDBJ/EMBL/GenBank databases under accession no. LC366023. The sequence contained an open reading frame of 906 base pairs encoding 301 amino acid residues (Fig. 1). The theoretical molecular mass and isoelectric point of bmMTHFD are 32 733 Da and 6.1, respectively. Screening the predicted 3D model of MTHFD in Protein Data Bank (<https://www.rcsb.org>) showed highest similarity to the human MTHFD1 with *E*-value of 1.57E-92. The amino acid sequence of bmMTHFD showed 41% identity to DC301, 49% to human MTHFD2 (hMTHFD2), 49% to mouse MTHFD2 (mMTHFD2), 23% to *Saccharomyces cerevisiae* MTHFD, 13% to mouse MTHFD1L (mMTHFD1L), 13% to human MTHFD1L (hMTHFD1L), 14% to mouse MTHFD1 (mMTHFD1), 51% to human MTHFD2L (hMTHFD2L), 51% to mouse MTHFD2L (mMTHFD2L), 45% to *E. coli* MTHFD, 72% to *Papilio xuthus* MTHFD, 61% to *Culex quinquefasciatus* MTHFD, and 55% to *Aedes aegypti* MTHFD. The DC301 structure has NADP⁺- and substrate-binding sites along with the dimer-interaction site [3], all of which were conserved among these protein sequences (Fig. 1). Based on the phylogenetic tree generated from the aligned amino acid sequences of the MTHFDs, the bmMTHFD cloned in this study belonged to an independent group including putative MTHFDs of insects (Fig. 2).

Structural description of bmMTHFD

The crystal structure of bmMTHFD was solved by the molecular replacement method, and the structure was refined against 1.75 Å resolution data. The data collection and refinement statistics are summarized in Table 1. Two molecules were included in the crystallographic asymmetric unit and showed similar structures with root-mean-square deviation (rmsd) between the corresponding atoms of 1.05 Å. The residues 44–46 and 246–253 of the A chain and 244–252 of the B chain were not modeled because of poor electron density. Figure S1 shows that bmMTHFD eluted at 205 mL, corresponding to a molecular size of 63 kDa, and appears to be a homodimeric protein based on the elution profile. Dali search for protein structure comparison between the determined structure and structures already deposited in the Protein

Data Bank [23] revealed that DC301 (PDB ID: 1A4I) with rmsd of 1.40 Å and hMTHFD2 (PDB ID: 5TC4) with rmsd of 1.10 Å were most similar. The sequence alignment of bmMTHFD to that of DC301 showed that the α -helices and β -strands of bmMTHFD were conserved across the structures (Fig. 1). bmMTHFD is composed of two domains containing N-terminal and C-terminal domains with an α/β fold. In DC301, hydrogen bonds are crucial for stabilizing interactions of the DC301 monomers [3]. Thr191, Val192, Thr193, and Thr194 of one subunit of DC301 interact with Thr199, Thr194, Thr193, and Val192 of the other subunit via hydrogen bonds [3]. The sequence alignment of DC301 to bmMTHFD (Fig. 1) indicated that Thr193, Val194, Thr195, and Ile196 of the *B. mori* residues interact with Thr201, Ile196, Thr195, and Val194 of the other subunit in a similar manner to DC301. These amino acid residues form β -strand in the C-terminal region and provide bridging hydrogen bonds in bmMTHFD that contribute to stabilizing the dimer of bmMTHFD. There are hydrogen bonds observed between the side chain of Thr193 in one subunit and side chain of Thr201 in the other subunit, main chain of Val194 of one subunit and main chain of Ile196 in the other subunit, side chain of Thr195 in one subunit and the counterpart in the other subunit, and main chain of Ile196 in one subunit and main chain of Val194 in the other subunit (Fig. 1).

Cofactor-binding site

The active site of DC301 contains a cofactor-binding site and a substrate-binding site [3]. Superposition of structures between bmMTHFD and DC301 (PDB ID: 1A4I) revealed that the NADP⁺-binding site of bmMTHFD could be composed of Arg166, Ser167, Arg199, and Ile242 (Fig. 4A). By using superposition of structures between bmMTHFD and hMTHFD2 (PDB ID: 5TC4) [24], we found that the amino acid residues Arg166, Ser167, Asn169, Arg199, Ile242, and Thr281 could be involved in NAD⁺ binding (Fig. 4B).

Putative substrate-binding sites

In DC301, Tyr52-XXX-Lys56 was identified as the YXXXK motif that plays a role in catalysis [3]. The motif is also conserved in bmMTHFD as Tyr49-XXX-Lys53. Ser49, Gln100, and Pro102 may also play a role in catalysis [3]. Superposition of DC301 with bmMTHFD revealed that the equivalent bmMTHFD residues were Ser46, Gln97, and Pro99 (Fig. 3A). Furthermore, the putative substrate-binding site of *E. coli*

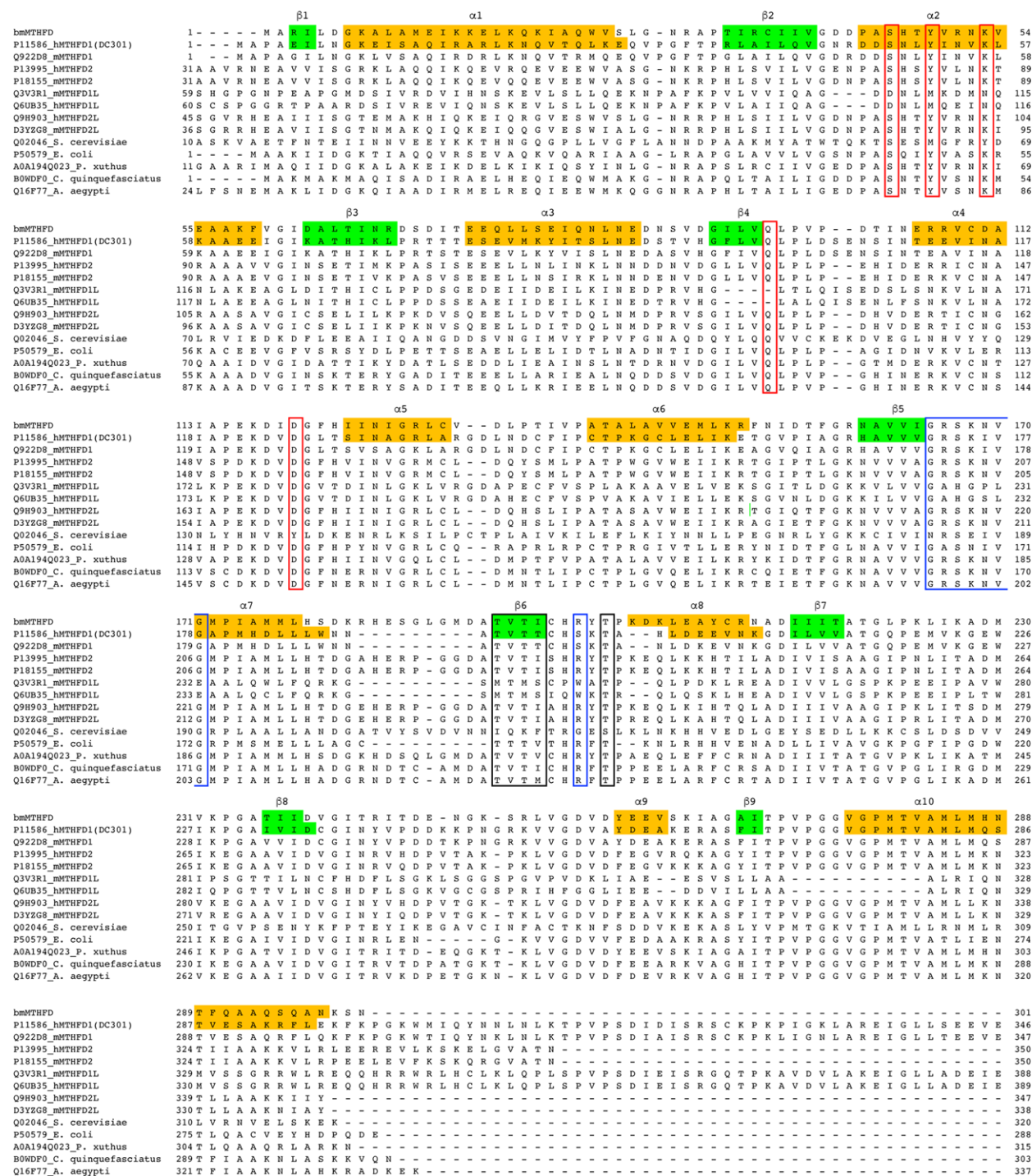


Fig. 1. Alignment of amino acid sequences of MTHFDs. Sequences were obtained from the Swiss-Prot database (<http://expasy.org/sprot/>). Each entry contains an accession number. DC301 is the D/C domain of hMTHFD1. The NAD(P)⁺-binding, substrate-binding, and dimer-interaction sites are boxed in blue, red, and black, respectively. α -Helices (shaded in orange) and β -strands (shaded in green) are labeled by α and β , respectively.

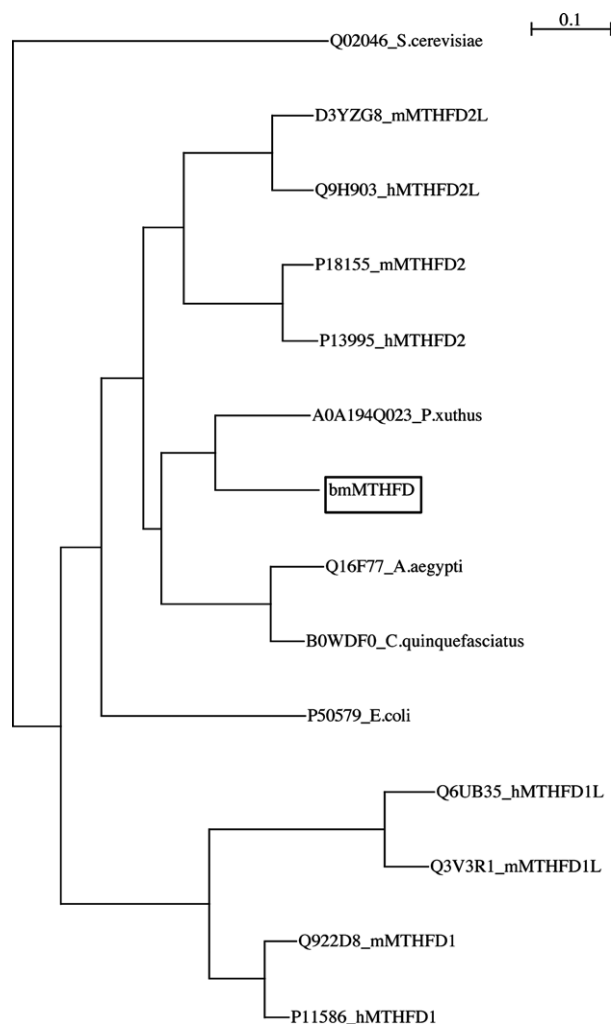


Fig. 2. Phylogenetic analysis of MTHFD amino acid sequences. The phylogenetic tree was constructed with a neighbor-joining plot software using different GST sequences from the Swiss-Prot database (https://web.expasy.org/docs/swiss-prot_guideline.html). Each entry contains the species name or protein name and the accession number. The numbers attached to the nodes indicate the bootstrap values. bmMTHFD is boxed. The scale bar corresponds to a 0.1 distance.

MTHFD (PDB ID: 1B0A) is composed of Lys9, Tyr50, Lys54, Lys56, Gln58, Asp121, and Arg234 [10]. We identified four residues of bmMTHFD, Tyr49, Lys53, Gln97, and Asp120 that correspond to these in *E. coli* MTHFD [10] (Fig. 3B).

Amino acid residues involved in substrate and cofactor binding

Bombyx mori MTHFD produced in bacteria was purified to homogeneity, yielding a single band in SDS/PAGE with a molecular size of ~ 32 000 Da. This was

Table 1. Data collection and refinement statistics. Values in parentheses indicate the highest-resolution shell.

Space group	$P4_3$
Unit cell parameters (Å)	$a = b = 89.79, c = 127.31$
Resolution range (Å)	38.3–1.75 (1.78–1.75)
Total number of reflections	735 590
Number of unique reflections	100 729 (5028)
Multiplicity	7.3 (7.6)
R_{merge}^a (%)	5.7 (48.7)
$\langle I \rangle / \langle \sigma(I) \rangle$	52.3 (6.15)
Completeness (%)	99.3 (100.0)
Refinement statistics	
Resolution range (Å)	38.3–1.75
Number of reflections	100 664
R_{work}^b (%) / R_{free}^c (%)	21.1/24.1
Root-mean-square deviations	
Bond lengths (Å)/bond angles (°)	0.007/0.869
Average B -factors (Å ²)/number of atoms	
Protein (Chain A, B)	32.4/2230, 33.5/2244
Small molecules ^d	37.5/32
Water	37.7/315
Ramachandran plot	
Favored region (%)	98.4
Allowed region (%)	1.23
Outliers (%)	0.35

^a $R_{\text{merge}} = \sum(I - \langle I \rangle) / \sum \langle I \rangle$, where I is the intensity measurement for a given reflection and $\langle I \rangle$ is the average intensity for multiple measurements of this reflection.

^b $R_{\text{work}} = \sum |F_{\text{obs}} - F_{\text{cal}}| / \sum F_{\text{obs}}$, where F_{obs} and F_{cal} are the observed and calculated structure-factor amplitudes.

^c The R_{free} value was calculated using only an unrefined, randomly chosen subset of reflection data (5%) that were excluded from refinement.

^d Small molecules include sulfate ion and ethylene glycol.

close to the estimated size based on the deduced amino acid sequence.

We identified Ser46, Tyr49, Lys53, Gln97, and Asp120 as amino acid residues crucial for catalysis. These amino acid residues are highly conserved within MTHFD as the substrate-binding site (boxed in red, Fig. 1). To determine whether these residues were important for catalytic activity, we performed site-directed mutagenesis. The resulting mutants were named S46A, Y49A, K53A, Q97A, and D120A and were purified from *E. coli* clones using the same method as for the wild-type (WT) enzyme. A single band was obtained for each preparation of the mutant enzyme in SDS/PAGE. We examined optimum condition for bmMTHFD assay (Fig. S2). We determined the kinetic parameters of the mutant enzymes under the optimum condition (Fig. S2) and compared these with the kinetic parameters of the WT enzyme (Table 2). The catalytic efficiencies represented by $k_{\text{cat}}/K_{\text{m}}$ for MTHF with Y49A and D120A were insignificant (Table 2). The $k_{\text{cat}}/K_{\text{m}}$ for NADP⁺ was

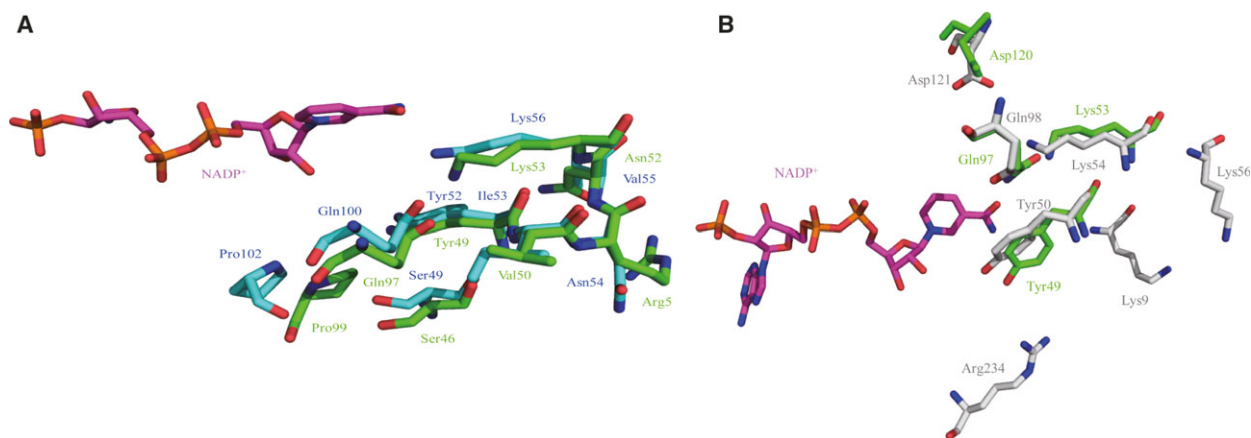


Fig. 3. Superimposed structures of putative substrate-binding site of bmMTHFD with DC301 (A) and *Escherichia coli* MTHFD (B). Structures of bmMTHFD, DC301, *E. coli* MTHFD, and NADP⁺ are shown in green, cyan, gray, and magenta, respectively. Atoms for oxygen, nitrogen, and phosphorus are red, blue, and orange, respectively. Amino acid residues for bmMTHFD, DC301, and *E. coli* MTHFD are described in green, blue, and gray, respectively.

Table 2. Comparison of kinetic data for D activity (A) and C activity (B) of bmMTHFD. The values of K_m and V_{max} are expressed as the mean with \pm SD from three independent experiments. K_m , V_{max} , and k_{cat}/K_m are expressed as mM, mM·min⁻¹, and mM⁻¹·s⁻¹, respectively. Kinetic parameters were measured with a nonlinear least squares data fit to the Michaelis–Menten equation under the assay conditions described in Materials and methods. Kinetic data of MTHF were measured in the presence of NADP⁺. ND means not detected.

(A) D activity

	WT	S46A	Y49A	K53A	Q97A	D120A
MTHF						
K_m	0.51 ($\pm 3.3 \times 10^{-2}$)	0.19 ($\pm 4.8 \times 10^{-3}$)	ND	1.1 ($\pm 1.8 \times 10^{-2}$)	0.81 ($\pm 2.8 \times 10^{-2}$)	ND
V_{max}	1.6×10^{-3} ($\pm 3.7 \times 10^{-5}$)	6.9×10^{-4} ($\pm 6.2 \times 10^{-6}$)	ND	4.1×10^{-3} ($\pm 6.9 \times 10^{-5}$)	5.8×10^{-3} ($\pm 1.8 \times 10^{-4}$)	ND
k_{cat}/K_m	0.043	0.049	ND	0.051	0.097	ND
NADP ⁺						
K_m	0.22 ($\pm 1.0 \times 10^{-2}$)	0.98 ($\pm 1.3 \times 10^{-1}$)	ND	0.62 ($\pm 6.5 \times 10^{-2}$)	0.33 ($\pm 5.8 \times 10^{-2}$)	2.0 ($\pm 1.4 \times 10^{-1}$)
V_{max}	6.9×10^{-4} ($\pm 1.4 \times 10^{-5}$)	1.1×10^{-3} ($\pm 2.0 \times 10^{-5}$)	ND	9.8×10^{-4} ($\pm 1.4 \times 10^{-4}$)	6.9×10^{-4} ($\pm 3.8 \times 10^{-5}$)	1.0×10^{-3} ($\pm 2.5 \times 10^{-5}$)
k_{cat}/K_m	8.6×10^{-3}	3.3×10^{-3}	ND	6.2×10^{-3}	5.2×10^{-3}	2.1×10^{-3}
NAD ⁺						
K_m	0.27 ($\pm 2.0 \times 10^{-2}$)	0.45 ($\pm 8.5 \times 10^{-2}$)	ND	0.30 ($\pm 3.0 \times 10^{-2}$)	0.47 ($\pm 8.6 \times 10^{-2}$)	2.0 ($\pm 9.0 \times 10^{-2}$)
V_{max}	6.2×10^{-4} ($\pm 1.7 \times 10^{-5}$)	8.4×10^{-4} ($\pm 3.0 \times 10^{-5}$)	ND	7.3×10^{-4} ($\pm 3.0 \times 10^{-6}$)	9.7×10^{-4} ($\pm 4.1 \times 10^{-5}$)	1.0×10^{-3} ($\pm 5.9 \times 10^{-5}$)
k_{cat}/K_m	7.0×10^{-3}	5.6×10^{-3}	ND	6.6×10^{-3}	5.1×10^{-3}	1.9×10^{-3}

(B) C activity

	WT	S46A	Y49A	K53A	Q97A	D120A
K_m	0.13 ($\pm 6.0 \times 10^{-3}$)	0.16 ($\pm 1.0 \times 10^{-2}$)	ND	0.16 ($\pm 1.5 \times 10^{-2}$)	0.019 ($\pm 2.5 \times 10^{-3}$)	0.23 ($\pm 2.5 \times 10^{-2}$)
V_{max}	6.4×10^{-4} ($\pm 2.5 \times 10^{-5}$)	6.3×10^{-4} ($\pm 1.5 \times 10^{-5}$)	ND	5.7×10^{-4} ($\pm 2.5 \times 10^{-5}$)	1.5×10^{-4} ($\pm 5.0 \times 10^{-6}$)	8.2×10^{-4} ($\pm 4.5 \times 10^{-5}$)
k_{cat}/K_m	0.13	0.11	ND	0.097	0.22	0.097

8.6×10^{-3} mM⁻¹·s⁻¹ for the WT bmMTHFD, which was 2.6-, 1.4-, 1.7-, and 4.1-fold higher than that for the S46A, K53A, Q97A, and D120A mutants,

respectively (Table 2). The values of Y49A toward MTHF and NADP⁺ were not detected. In the case of NAD⁺, the k_{cat}/K_m values for the mutants were

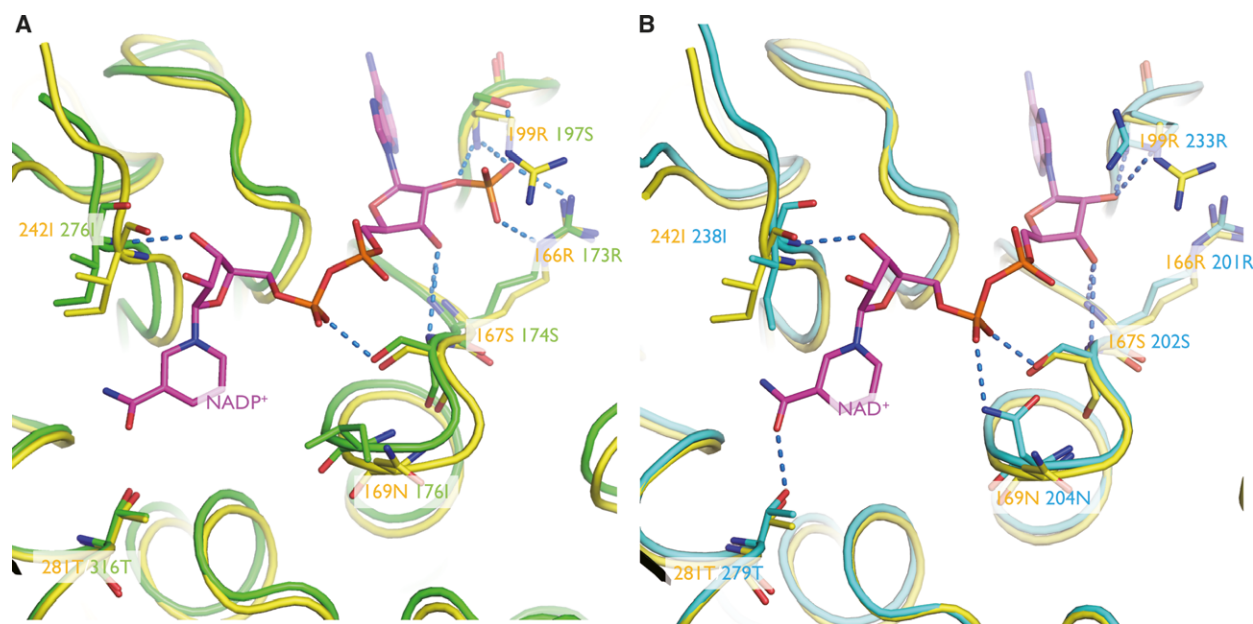


Fig. 4. Cofactor-binding site. (A) The structures of bmMTHFD, DC301, and NADP⁺ are shown in yellow, green, and magenta, respectively. Residue labels for bmMTHFD and DC301 are in yellow and green, respectively. (B) The structures of bmMTHFD, hMTHFD2, and NADP⁺ are shown in yellow, cyan, and magenta, respectively. Residue labels for bmMTHFD and hMTHFD2 are in yellow and cyan, respectively. The sulfate ion is indicated in yellow and labeled in black. Color scheme: oxygen, red; nitrogen, blue; sulfur, yellow. The blue dots show hydrogen bonds.

similar to that of WT, although we were unable to detect the $k_{\text{cat}}/K_{\text{m}}$ of Y49A for 5,10-methenyl-THF, unlike the other mutants (Table 2).

Discussion

This study aimed to detect the presence of one-carbon metabolism in the silkworm *B. mori* and to analyze whether MTHFD is relevant for insecticide design. The amino acid sequence of bmMTHFD revealed high homologies to those of mitochondrial MTHFD2 (49%) and MTHFD2L (51%). However, bmMTHFD did not possess a predicted N-terminal mitochondrial targeting sequence (Fig. 1).

According to substrate specificity, MTHFD proteins are divided into mono-, bi-, and trifunctional proteins. For instance, MTHFD1 is a trifunctional enzyme [25], MTHFD2 is a bifunctional enzyme [26], and MTHFD1L is a monofunctional enzyme [9]. Our experiments revealed that bmMTHFD is bifunctional, exhibiting D and C activities (Table 2). In DC301, Tyr52-XXX-Lys56 and/or Ser49, Gln100, and Pro102 were proposed as C active site. Since we found the corresponding residues in bmMTHFD to be Tyr49-XXX-Lys53 and/or Ser46, Gln97, and Pro99 [3], we predicted that bmMTHFD possesses the C domain (Fig. 1). The amino acid residues required for S

activity are located in the S domain of the 70 kDa MTHFD1. However, bmMTHFD lacks the S domain for S activity (Fig. 1).

Elucidation of the bmMTHFD tertiary structure identified a globular shape that was similar to those of other known MTHFDs. In addition, the conserved amino acid residues are critical for stabilizing the interaction of bmMTHFD monomers. Our study indicates that Thr193, Val194, Thr195, Ile196, and Thr201 are involved in hydrogen bond interactions, suggesting that these amino acid residues are involved in maintaining the homodimer of bmMTHFD.

Our findings show that bmMTHFD contains the catalytic site where the substrate binds in the N-terminal domain and displays cofactor binding in the C-terminal domain, like other MTHFDs. bmMTHFD is composed of 10 α -helices and nine β -strands, and their location is highly conserved in DC301 (Fig. 1).

The K_{m} value of bmMTHFD was also found to be similar to those for *Pseudomonas aeruginosa* (26 μM for MTHF and 176 μM for NADP⁺) [27], and *Trypanosoma brucei* (35 μM for MTHF and 70 μM for NADP⁺) [28]. bmMTHFD uses either NADP⁺ or NAD⁺ for catalysis with similar $k_{\text{cat}}/K_{\text{m}}$ values (Table 2). Similar results were obtained with mammalian MTHFDs, where hMTHFD2 and mMTHFD2L were reported to utilize both NADP⁺

and NAD⁺ [7,8]. Superimposition between bmMTHFD and DC301 requiring NADP⁺ indicated that a single cofactor bound to the site contained Arg166, Ser167, Arg199, and Ile242 (Fig. 4A). Superimposition between bmMTHFD and hMTHFD2 requiring NAD⁺ showed that Arg166, Ser167, Asn169, Arg199, Ile242, and Thr281 were present in bmMTHFD in a dinucleotide-binding motif (Fig. 4B). In bmMTHFD, the motif (165GRSRQVG171) does not conform to the classical GXGXXG fingerprint pattern. The sequence comparisons did not suggest the amino acid residues that may be involved in NAD⁺ utilization. Cocrystallization with NADP⁺ or NAD⁺ would allow improved prediction of specificity.

Tyr52, Lys56, Gln100, and Asp125 in hMTHFD1 have been identified as critical for its activity [29]. Superimposition between DC301 and bmMTHFD revealed that the corresponding residues in bmMTHFD were Tyr49, Lys53, Gln97, and Asp120, respectively. Tyr52 (bmMTHFD Tyr49) was implicated in substrate binding, and Lys56 (bmMTHFD Lys53) and Gln100 (bmMTHFD Gln97) were proposed to work together in establishing an environment for hydride transfer [3,29,30]. Asp125 (bmMTHFD Asp120) serves in substrate binding. Mutations at Tyr49 eliminated affinities to MTHF and 5,10-methenyl-THF. Mutagenesis of Asp120 influences the D activity. The results are consistent with a key role for Tyr52, and Asp125 in hMTHFD1 to correctly bind the folate substrate. Since MTHFDs have an aspartate residue at the 125 position, the carboxylate is invariant. The side chain would then be in position to interact with the edge of the pterin ring at N3 and at the exocyclic amine of C2 of pterin [31]. In bmMTHFD, the electron density from Pro44 to Ser46 of A chain was disordered. In DC301, there is a hydrogen-bonding network near these amino acid residues [3]. To examine whether Ser46 contributes to bmMTHFD activity, we mutated the amino acid residue to Ala. Subsequent mutagenesis results showed that their kinetic parameters were not changed significantly (Table 2). In *E. coli* MTHFD, mutation of Asp121 (bmMTHFD Asp120) decreases the k_{cat}/K_m values toward MTHF to 0.17%, while the kinetic parameters of mutating Tyr50 (bmMTHFD Tyr49) could not be determined, due to very low activities [32]. The results indicate that the catalytic mechanism could be similar to that in *E. coli* MTHFD.

In conclusion, this study identified and biochemically characterized bmMTHFD, a MTHFD in the silkworm *B. mori*. To the best of our knowledge, this study is the first to report an MTHFD in insects. Furthermore, we identified bmMTHFD amino acid residues that are

likely to play roles in catalysis. We are currently attempting the cocrystallization of bmMTHFD with a suitable substrate analogue conjugate to aid in the determination of amino acid residues involved in bmMTHFD catalysis. Our findings provide insight into the mechanism underlying bmMTHFD activity and potentially that of other insects and may therefore facilitate the development of more effective and safe insecticides.

Acknowledgements

This study was supported in part by JSPS KAKENHI [grant numbers JP15H04611, 17K19272] and by a Research Grant for Young Investigators of the Department of Agriculture, Kyushu University. This work was performed in part under the Cooperative Research Program of Institute for Protein Research, Osaka University, CR -17- 05. This work was performed using a synchrotron beamline BL44XU at SPring-8 under the Cooperative Research Program of Institute for Protein Research, Osaka University. Diffraction data were collected at the Osaka University beamline BL44XU at SPring-8 (Harima, Japan) [2016B6654, 2017A6764, 2017B6764].

Conflict of interest

The authors declare no conflict of interest.

Author contributions

MRH, AH, and KY designed experiments. MRH, AH, AH, and KY performed experiments. MRH, AH, AN, AH, SF, and WC analyzed the data. AN, SF, and KY wrote the manuscript.

References

- 1 Fox JT and Stover PJ (2008) Folate-mediated one-carbon metabolism. *Vitam Horm* **79**, 1–44.
- 2 Burda P, Kuster A, Hjalmarsen O, Suormala T, Burer C, Lutz S, Roussey G, Christa L, Asin-Cayuela J, Kollberg G *et al.* (2015) Characterization and review of MTHFD1 deficiency: four new patients, cellular delineation and response to folic and folinic acid treatment. *J Inher Metab Dis* **38**, 863–872.
- 3 Allaire M, Li Y, MacKenzie RE and Cygler M (1998) The 3-D structure of a folate-dependent dehydrogenase/cyclohydrolase bifunctional enzyme at 1.5 Å resolution. *Structure (London, England: 1993)*, **6**, 173–182.
- 4 Belanger C and MacKenzie RE (1989) Isolation and characterization of cDNA clones encoding the murine NAD-dependent methylenetetrahydrofolate dehydrogenase-methylenetetrahydrofolate cyclohydrolase. *J Biol Chem* **264**, 4837–4843.

- 5 Prasanna P, Pike S, Peng K, Shane B and Appling DR (2003) Human mitochondrial C1-tetrahydrofolate synthase: gene structure, tissue distribution of the mRNA, and immunolocalization in Chinese hamster ovary cells. *J Biol Chem* **278**, 43178–43187.
- 6 Bolusani S, Young BA, Cole NA, Tibbetts AS, Momb J, Bryant JD, Solmonson A and Appling DR (2011) Mammalian MTHFD2L encodes a mitochondrial methylenetetrahydrofolate dehydrogenase isozyme expressed in adult tissues. *J Biol Chem* **286**, 5166–5174.
- 7 Shin M, Bryant JD, Momb J and Appling DR (2014) Mitochondrial MTHFD2L is a dual redox cofactor-specific methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase expressed in both adult and embryonic tissues. *J Biol Chem* **289**, 15507–15517.
- 8 Shin M, Momb J and Appling DR (2017) Human mitochondrial MTHFD2 is a dual redox cofactor-specific methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase. *Cancer Metab* **5**, 11.
- 9 Walkup AS and Appling DR (2005) Enzymatic characterization of human mitochondrial C1-tetrahydrofolate synthase. *Arch Biochem Biophys* **442**, 196–205.
- 10 Shen BW, Dyer DH, Huang JY, D'Ari L, Rabinowitz J and Stoddard BL (1999) The crystal structure of a bacterial, bifunctional 5,10 methylene-tetrahydrofolate dehydrogenase/cyclohydrolase. *Protein Sci* **8**, 1342–1349.
- 11 Wohlfarth G, Geerligs G and Diekert G (1991) Purification and characterization of NADP(+)-dependent 5,10-methylenetetrahydrofolate dehydrogenase from *Peptostreptococcus productus* marburg. *J Bacteriol* **173**, 1414–1419.
- 12 Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- 13 Higashiura A, Yamashita E, Yoshimura M, Hasegawa K, Furukawa Y, Kumasaka T, Ueno G, Yamamoto M, Tsukihara T and Nakagawa A (2016) SPRing-8 BL44XU, beamline designed for structure analysis of large biological macromolecular assemblies. *AIP Conf Proc* **1741**, 030028.
- 14 Otwinowski Z and Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* **276**, 307–326.
- 15 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC and Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
- 16 Langer G, Cohen SX, Lamzin VS and Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* **3**, 1171–1179.
- 17 Emsley P, Lohkamp B, Scott WG and Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.
- 18 Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH and Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352–367.
- 19 Kawabata T (2003) MATRAS: a program for protein 3D structure comparison. *Nucleic Acids Res* **31**, 3367–3369.
- 20 Yamamoto K, Higashiura A, Suzuki M, Aritake K, Urade Y and Nakagawa A (2017) Molecular structure of a prostaglandin D synthase requiring glutathione from the brown planthopper, *Nilaparvata lugens*. *Biochem Biophys Res Commun* **492**, 166–171.
- 21 Tan LU, Drury EJ and MacKenzie RE (1977) Methylenetetrahydrofolate dehydrogenase-methenyltetrahydrofolate cyclohydrolase-formyltetrahydrofolate synthetase. A multifunctional protein from porcine liver. *J Biol Chem* **252**, 1117–1122.
- 22 Barlowe CK and Appling DR (1990) Isolation and characterization of a novel eukaryotic monofunctional NAD(+)-dependent 5,10-methylenetetrahydrofolate dehydrogenase. *Biochemistry* **29**, 7089–7094.
- 23 Holm L and Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545–W549.
- 24 Gustafsson R, Jemth AS, Gustafsson NM, Farnegardh K, Loseva O, Wiita E, Bonagas N, Dahllund L, Llona-Minguez S, Haggblad M *et al.* (2017) Crystal structure of the emerging cancer target MTHFD2 in complex with a substrate-based inhibitor. *Can Res* **77**, 937–948.
- 25 Thigpen AE, West MG and Appling DR (1990) Rat C1-tetrahydrofolate synthase. cDNA isolation, tissue-specific levels of the mRNA, and expression of the protein in yeast. *J Biol Chem* **265**, 7907–7913.
- 26 Mejia NR, Rios-Orlandi EM and MacKenzie RE (1986) NAD-dependent methylenetetrahydrofolate dehydrogenase-methenyltetrahydrofolate cyclohydrolase from ascites tumor cells, Purification and properties. *J Biol Chem* **261**, 9509–9513.
- 27 Eadsforth TC, Gardiner M, Maluf FV, McElroy S, James D, Frearson J, Gray D and Hunter WN (2012) Assessment of *Pseudomonas aeruginosa* N5, N10-methylenetetrahydrofolate dehydrogenase-cyclohydrolase as a potential antibacterial drug target. *PLoS ONE* **7**, e35973.
- 28 Eadsforth TC, Pinto A, Luciani R, Tamborini L, Cullia G, De Micheli C, Marinelli L, Cosconati S, Novellino E, Lo Presti L *et al.* (2015) Characterization of 2,4-diamino-6-oxo-1,6-dihydropyrimidin-5-yl ureido based inhibitors of *Trypanosoma brucei* FoD and testing for antiparasitic activity. *J Med Chem* **58**, 7938–7948.
- 29 Schmidt A, Wu H, MacKenzie RE, Chen VJ, Bewly JR, Ray JE, Toth JE and Cygler M (2000) Structures of three inhibitor complexes provide insight into the

- reaction mechanism of the human methylenetetrahydrofolate dehydrogenase/cyclohydrolase. *Biochemistry* **39**, 6325–6335.
- 30 Allaire M, Li Y, Mejia NR, Pelletier JN, MacKenzie RE and Cygler M (1996) Crystallization of the bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase domain of the human trifunctional enzyme. *Proteins* **26**, 479–480.
- 31 Reyes VM, Sawaya MR, Brown KA and Kraut J (1995) Isomorphous crystal structures of *Escherichia coli* dihydrofolate reductase complexed with folate, 5-deazafofolate, and 5,10-dideazatetrahydrofolate: mechanistic implications. *Biochemistry* **34**, 2710–2723.
- 32 Sah S and Varshney U (2015) Impact of mutating the key residues of a bifunctional 5,10-methylenetetrahydrofolate dehydrogenase-cyclohydrolase from *Escherichia coli* on its activities. *Biochemistry* **54**, 3504–3513.

Supporting information

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

Fig. S1. Superdex 200 analytical gel filtration. Protein standards (closed circle) include thyroglobulin (670 kDa), globulin (158 kDa), ovalbumin (44 kDa), and myoglobin (17 kDa). The plot of log molecular weight vs elution volume W_t with an R^2 value of 0.999. Closed triangle indicates position of bmMTHFD.

Fig. S2. Enzymatic properties of bmMTHFD. The enzymatic properties of bmMTHFD were analyzed using 5, 10-methylenetetrahydrofolate as substrate and NADP^+ as cofactor. The MTHFD activity was assayed under standard conditions, as described in Experimental Procedures, unless otherwise indicated. The maximum value obtained was set to 100%. Data represent the mean with \pm SD from three independent experiments. (A) pH stability was assessed by preincubation of the enzyme solution at various pH values at 4 °C for 24 h before the residual activity was assayed. (B) Thermostability was determined by preincubation of the enzyme solution at various temperatures at pH 6 for 30 min before the residual activity was assayed. (C) Optimum pH levels for the activities were assayed at 28 °C using citrate-phosphate-borate buffer at various pH value.