Steady-State Kinetics and Molecular Evolution of *Escherichia coli* MenD [(1*R*,6*R*)-2-Succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate Synthase], an Anomalous Thiamin Diphosphate-Dependent Decarboxylase−Carboligase†

Milan Bhasin, Jennifer L. Billinsky, and David R. J. Palmer*

Department of Chemistry, University of Saskatchewan, 110 Science Place, Saskatoon, Saskatchewan S7N 5C9, Canada

Received July 11, 2003; Revised Manuscript Received September 22, 2003

ABSTRACT: (1*R*,6*R*)-2-Succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) synthase, or MenD, catalyzes the thiamin diphosphate- (ThDP-) dependent decarboxylation of 2-oxoglutarate, the subsequent addition of the resulting succinyl-ThDP moiety to isochorismate, and the δ-elimination of pyruvate to yield SHCHC, pyruvate, and carbon dioxide. The enzyme is part of a superfamily of ThDP-dependent 2-oxo acid decarboxylases that includes pyruvate decarboxylase, benzoylformate decarboxylase, and acetohydroxy acid synthase, among others. However, this is the only enzyme known to catalyze a Stetter-like 1,4-addition of a ThDP adduct to the β-carbon of an unsaturated carboxylate. Herein we report properties of the MenD protein from *Escherichia coli*, including the results of the first steady-state kinetic studies of the SHCHC synthase reaction. The protein is a dimer and shows cooperativity with respect to both substrates. The enzyme prefers divalent manganese as its metal ion cofactor and shows no dependence on FAD. MenD, required for biosynthesis of menaquinone and phylloquinone, is found in the genomes of a wide range of bacteria, as well as that of the archaeon *Halobacterium* sp. NRC-1 and the eukaryote *Arabidopsis thaliana*. Sequence alignments with other members of the superfamily are used to predict amino acid residues likely to be important in the binding and activation of ThDP. A site-directed mutant that replaces the conserved glutamic acid residue (E55), predicted to interact with N1′ of the aminopyrimidine ring, with glutamine was generated, with catastrophic results for catalysis. There is no evidence for the release of succinate semialdehyde as a product; therefore, EC 4.1.1.71 should not be used for this enzyme.

Thiamin diphosphate (ThDP),[1] a coenzyme derived from vitamin B_1, is also known as cocarboxylase because of its role in the enzymatic decarboxylation of 2-oxo acids. Although the widely accepted mechanism by which ThDP promotes these reactions was proposed over 50 years ago (*1*), the details of this mechanism remain an area of strong interest (*2*, *3*). There is a superfamily of ThDP-dependent 2-oxo acid decarboxylases which includes benzoylformate decarboxylase (BFD) (*4*), pyruvate decarboxylase (PDC) (*5*), indolepyruvate decarboxylase (IPDC) (*6*), pyruvate oxidase (POX) (*7*), benzaldehyde lyase (*8*), glyoxalate carboligase (*9*), and acetohydroxy acid synthase (AHAS) (*10*). Most of these enzymes have been the subject of extensive kinetic and structural studies. All apparently activate ThDP in approximately the same way, via enforcing a "V-conformation" on the cofactor and by a conserved array of residues involved in hydrogen bonding and coordination of a divalent metal ion. A review of ThDP-dependent enzymatic reactions has appeared recently (*11*).

The most mysterious member of this superfamily is (1*R*,6*R*)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) synthase, or MenD, named for the fourth gene identified in the menaquinone biosynthetic pathway. This enzyme converts isochorismate and 2-oxoglutarate to SHCHC, pyruvate, and carbon dioxide as shown in Scheme 1. This reaction is the first committed step in the biosynthesis of menaquinone, also called vitamin K_2, and phylloquinone (*12*). The absolute configuration of the product of this reaction was only recently determined in the laboratory of Ivan Rayment (*13*).

The likely mechanism of the reaction catalyzed by MenD, proposed in essentially this form almost 20 years ago (*14*), is shown in Scheme 1. In this mechanism, ThDP is deprotonated at carbon 2 of the thiazolium ring, and this ylide then attacks the carbonyl carbon of 2-oxoglutarate. The resulting tetrahedral intermediate decarboxylates to form a resonance-stabilized carbanion/enamine, which then acts as a nucleophile in attacking the β-carbon of isochorismate. The resulting adduct breaks down by loss of ThDP to give the ketone. The α-proton is now activated for abstraction by some active site base, resulting in the δ-elimination of pyruvate to give the product SHCHC.

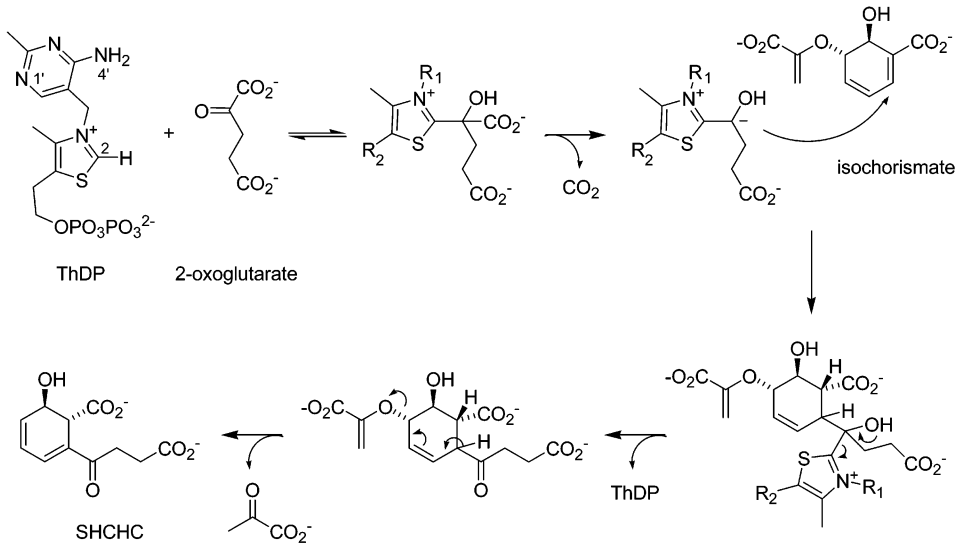
† This work was supported by the Department of Chemistry, University of Saskatchewan, and by an operating grant to D.R.J.P. from National Sciences and Engineering Research Council of Canada (NSERC).

* To whom correspondence should be addressed. E-mail: palmer@sask.usask.ca.

[1] Abbreviations: ThDP, thiamin diphosphate; SHCHC, (1*R*,6*R*)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate; AHAS, acetohydroxy acid synthase; BFD, benzoylformate decarboxylase from *Pseudomonas putida*; PDC, pyruvate decarboxylase; yPDC, pyruvate decarboxylase from *Saccharomyces cerevisiae*; POX, pyruvate oxidase; PSI-BLAST, position-specific iterative basic local alignment search tool.

10.1021/bi035224j CCC: $25.00
Published on Web 10/28/2003

Scheme 1: Reaction and Proposed Mechanism Catalyzed by MenD[a]



[a] Putative active site acid/base catalysts are not shown.

This reaction is remarkable in many ways. The ThDP-dependent decarboxylation reaction itself is expected to proceed in a manner similar to those of homologous enzymes such PDC. The details of ThDP deprotonation and 2-oxo acid decarboxylation mechanisms have been discussed at length elsewhere (*2*, *11*, *15*). However, MenD is the only enzyme known to catalyze the addition of a ThDP adduct to the β-carbon of a second substrate. The nonenzymatic chemical reaction that most resembles this process is the Stetter reaction (*16*, *17*), a Michael-type variation of the benzoin condensation. As one might expect, Stetter reactions only occur efficiently in solution when the acceptor molecule is an electron-deficient alkene such as an α,β-unsaturated ester or nitrile.

Previous research, primarily in the laboratory of R. Meganathan, has demonstrated the overall reaction and its role in the synthesis of menaquinone. The gene encoding the protein responsible for activity has been identified experimentally in *Escherichia coli* (*18*), *Bacillus subtilis* (*19*), *Synechocytis* sp. 6803 (*20*), and *Staphylococcus aureus* (*21*), and involvement of the 2-oxoglutarate dehydrogenase complex has been ruled out. There are, however, many other aspects of this enzyme which have remained the subject of speculation. The National Center for Biotechnology Information (NCBI) database contains many genes identified by sequence similarity as MenD. These database entries are annotated in a variety of ways, ascribing a variety of characteristics to the enzyme without benefit of experiment. The enzyme has been called "bifunctional" and "multifunctional" and is usually called "2-oxoglutarate decarboxylase/SHCHC synthase". This has led several databases, including BRENDA (*22*) and ExPASy (*23*), to assign to MenD the EC 4.1.1.71, which refers to the decarboxylation of 2-oxoglutarate to succinate semialdehyde and carbon dioxide. Enzyme Commission numbers describe only the reaction catalyzed, i.e., the reactants and the products, and bear no relation to sequence, structure, or phylogeny. However, succinate semialdehyde has never been reported as a product of this enzyme. *There is no EC number* associated with the enzymatic synthesis of SHCHC.

Our laboratory has purified recombinant *E. coli* K12 MG1655 MenD bearing an N-terminal polyhistidine tag and developed the first continuous assay of enzyme activity. Here we report the properties of this enzyme and the reaction it catalyzes with the intent of clearing up the misconceptions about this enzyme which are currently propagating in databases. A phylogenetic analysis of MenD and a comparison with other members of the ThDP-dependent 2-oxo acid decarboxylase superfamily provide considerable insight into this enzyme and should aid in the study of related 2-oxo acid decarboxylases.

MATERIALS AND METHODS

Chemical reagents, including buffers, salts, 2-oxoglutarate, succinate semialdehyde, NADH, and FAD, were obtained from Aldrich, Sigma, or VWR CanLab and were categorized as molecular biology grade or were of the highest grade available. Succinate semialdehyde dehydrogenase was purchased from Sigma and used without further purification. DNA-manipulating biochemicals, including restriction enzymes, T4 DNA ligase, and associated reagents, were obtained from New England Biolabs. *Pfu* DNA polymerase and the QuickChange mutagenesis kit were obtained from Stratagene.

Centrifugation was performed using a Beckman J2-HS refrigerated centrifuge with a JLA-10.5 or JA-25.5 rotor. Protein chromatography was performed using an AP Bioscience AKTAfplc in a refrigerated chamber. UV−visible absorbance was measured using a Beckman DU-640 spectrophotometer with a circulating-bath-controlled temperature block. PCR amplifications were performed using an Eppendorf gradient thermocycler.

Molecular Cloning and Heterologous Gene Expression. Standard molecular biology techniques were followed (*24*). The gene *entC*, encoding an enterobactin-related isochorismate synthase, and the gene *menD* were cloned previously (*25*). Briefly, each gene was PCR-amplified from *E. coli* K12 MG1655 genomic DNA using primers which incorporated *Xho*I and *Bam*HI restriction sites at the N- and C-terminus, respectively, and ligated into the multiple cloning site of pET-

16b (Novagen) such that the gene was in-frame with the polyhistidine tag encoded on the vector. All cloned genes were completely sequenced. Sequencing reactions were carried out by the DNA Technologies Unit of the NRC, Plant Biotechnology Institute. These vectors were used to transform *E. coli* BL-21(DE3), and the target genes were typically overexpressed without the necessity of IPTG induction by incubation at 37 °C with shaking at 250 rpm. A 5 mL LB−ampicillin culture was grown for 6 h, and then a larger culture was inoculated, using 1 mL of culture for every 50 mL of fresh media. MenD protein was harvested from sonicated cells of the second culture after 16 h of growth. EntC protein was harvested after no more than 7 h growth.

Site-Directed Mutagenesis. The mutant of the polyhistidine-tagged MenD protein was generated using the QuickChange kit from Stratagene according to manufacturer's directions. A mutagenic primer (and its reverse complement) was designed to incorporate a change in the nucleotide sequence coding for glutamic acid-55 to a glutamine, as well as a silent mutation which introduced a new restriction site in the mutant. The primer sequence was 5′-CCCATTTC-GAT*cAG*CGTGGccTAGGGCATCTGGCGC-3′, where the lower case letters represent mismatches to the wild-type sequence, the italicized codon is that of glutamine-55, and the underlined sequence is the newly incorporated *Avr*II site. Transformants were screened using *Avr*II restriction digests, and selected clones were tested for expression. All selected clones showed expression similar to the that of the wild-type enzyme. A clone was selected, and the complete gene was sequenced to ensure no other mutations had been introduced.

Protein Purification. The protein expression cell culture was pelleted, and the resulting cell pellet was lysed by sonication on ice in a volume of chilled buffer equivalent to 0.04 × the culture volume. The lysis buffer (buffer 1) contained 50 mM Tris-HCl, 10 mM imidazole, 500 mM NaCl, and 12.5% (v/v) glycerol, pH 8.0. The cellular debris was pelleted at 15000 rpm for 10 min using a JA-25.5 rotor. The supernatant was passed directly onto a column of chelating Sepharose FF (AP Biosciences) which had been charged with 4 column volumes of 50 mM $NiSO_4$, followed by 3 column volumes of buffer 1. The column was washed with 6 column volumes of buffer 1, followed by 4 volumes of buffer 2, which was buffer 1 with 40 mM imidazole. The column was reequilibrated with 1 column volume of buffer 1, and then eluted with 4 column volumes of buffer 3, which was buffer 1 containing 25 mM EDTA. Further chromatography did not result in higher specific activity. The eluate was fractionated, and the desired fractions were pooled and stored after dialysis into buffer containing 50 mM Tris-HCl, pH 8.5, 1 mM $MnCl_2$, and 50% glycerol.

Gel Filtration-Based Determination of Oligomerization. Purified MenD was passed over a 2.6 × 100 cm column of Sephacryl S-300HR (AP Biosciences) at a flow rate of 0.32 mL/min. The column was calibrated using the following standards: carbonic anhydrase (29 kDa), bovine serum albumin (67 kDa), β-amylase (200 kDa), apoferritin (443 kDa), and thyroglobulin (667 kDa). The void volume was determined using blue dextran. The molecular mass determination, including column packing and calibration, was performed in duplicate.

Preparation of Isochorismate. Chorismic acid was biosynthesized from glucose using *Klebsiella pneumoniae* 62-1 and isolated following the method of Turnbull (*26*). Isochorismate was obtained from the action of purified isochorismate synthase on chorismate at pH 8.0 and isolated by preparative reverse-phase chromatography using a prepacked C-8 column eluted with a gradient of water and methanol following the method of Schmidt and Leistner (*27*). Baseline separation of chorismic and isochorismic acid fractions was achieved, and isochorismic acid fractions were concentrated in vacuo at room temperature to remove methanol and lyophilized to dryness. The residue was dissolved in buffer at pH 8.5 and stored at −80 °C. Thawed samples were kept on ice and used for experiments within 48 h.

Assays of Enzymatic Activity. MenD activity was measured in a 1.0 mL volume at 25 °C in Tris-HCl buffer (100 mM), KCl (100 mM), DTT (1 mM), and appropriate concentrations of $MnCl_2$, ThDP, 2-oxoglutarate, and isochorismate. Activity was measured at varying pHs under saturating concentrations of substrates and cofactors to determine the optimum, which was pH 8.5. All rates reported are at this pH. The steady-state rate of product (pyruvate) formation was measured routinely using a lactate dehydrogenase- (LDH-) coupled assay by inclusion of rabbit muscle LDH (2 μL of a 12.5 unit/μL stock) and NADH (40 μM) and monitoring the disappearance of absorbance due to NADH at 340 nm. Kinetics reported are in the presence of 90 nM MenD protein, as measured by Bradford assay. All reaction rates were measured relative to a blank sample containing no MenD. Increasing the concentration of LDH did not affect the kinetics. All data points represent the average of at least two experiments.

Kinetic constants were determined using the program Leonora (*28*) by fitting the data to the equation:

$$v = V_{max}[S]^n/K' + [S]^n$$

where V_{max} = the maximal velocity, S = substrate, n = the Hill coefficient (n_H), $K' = (S_{0.5})^n$, and $S_{0.5}$ = the substrate concentration necessary for half-maximal velocity. In cases where $n = 1$, i.e., when no cooperativity is observed, this equation becomes the Michaelis−Menten equation, and $S_{0.5}$ becomes the Michaelis constant K_M. The maximal velocity was interpreted as a direct reflection of the turnover number, i.e., $V_{max} = k_{cat}[E]_0$, where $[E]_0$ is the initial concentration of enzyme. When the dependence of rate on cofactor is determined, the same equation is used, but S represents cofactor rather than substrate. (We have used the term $S_{0.5}$ or K_M when discussing cofactors as well as substrates, rather than K_c or K_{act} as is sometimes seen.) The Hill coefficient was determined from the slope of a plot of $\log[v/(V_{max} - v)]$ vs log [S], using values of v between 10% and 90% V_{max}.

Assay for Succinate Semialdehyde Formation. The enzyme was assayed as above, with the exception that LDH was replaced by 2 units of succinate semialdehyde dehydrogenase (EC 1.2.1.16) and $NADP^+$ was used in place of NADH. The reaction was performed in saturating conditions of cofactors and 2-oxoglutarate and in the presence and absence of isochorismate. A control reaction with added succinate semialdehyde was performed to demonstrate the activity of the dehydrogenase under these conditions.

Sequence Analyses. Protein sequences were obtained from GenBank, which was exhaustively searched using the program PSI-BLAST (*29*). Sequence alignments, amino acid identities and similarities, evolutionary distances, and phylogenetic trees were generated using the programs CLUSTAL W version 1.82 (*30*) and PHYLIP (*31*), including protdist, seqboot, neighbor, consense, and fitch, which were accessed through the Canadian Bioinformatics Resource (www.cbr.nrc.ca). Alignments were created using the Gonnet matrix (gap open penalty = 10, gap extension penalty = 0.2 for family members and 1.0 for superfamily members), and the results were compared with those using BLOSUM62 and PAM matrices. Evolutionary distances were calculated using the Dayhoff PAM matrix. Sequence alignments were visualized, colored, and edited using Genedoc 2.6 (http://www.psc.edu/biomed/genedoc). Phylogenetic trees were visualized using TreeView (http://taxonomy.zoology.gla.ac.uk/rod/rod.html).

RESULTS AND DISCUSSION

The MenD Protein. The *E. coli* K12 MenD protein has been reported as at least three distinct amino acid sequences, first by Popp (*32*), then by Meganathan (*19*), and then as the result of two *E. coli* K12 genome sequencing projects conducted independently at the University of Tokyo (*33*) and the University of Wisconsin (*34*). The last of these reported sequences matched that which resulted from PCR amplification using primers corresponding to the N- and C-terminal regions of the sequence matched.

The gene encoding the polyhistidine-tagged MenD was overexpressed in *E. coli*, resulting in the isolation of about 50 mg/L of liquid culture. Purification of the protein by Ni^{2+} affinity chromatography resulted in protein that was homogeneous by SDS−PAGE analysis. Further chromatographic steps did not result in higher specific activity. Removal of the N-terminal His tag resulted in no change in activity but considerable loss of protein; therefore, experiments were performed in the presence of this tag. Gel filtration indicated an apparent molecular mass of 1.4×10^5 Da, consistent with a dimer (predicted molecular mass of His-tagged monomer = 64261 Da). This differs from an earlier report that the protein is a trimer in solution (*35*).

The SHCHC Synthase Reaction. The enzyme is absolutely dependent on the presence of ThDP and a divalent metal ion for activity. EDTA abolishes activity. Although previous work has only reported the use of Mg^{2+} to activate the enzyme, we have found that addition of Mn^{2+} instead results in slightly higher activity with a significantly lower half-maximum concentration (K_M or $S_{0.5}$). The maximum activity observed was at pH = 8.5. Added KCl, up to 300 mM, resulted in activity 1.5 times greater, with no increase in activity for concentrations higher than 100 mM.

Using succinate semialdehyde dehydrogenase, we sought to detect the formation of succinate semialdehyde, including in the absence of isochorismate. The high-energy carbanion formed after decarboxylation (see Scheme 1) under such conditions could conceivably collapse via protonation, and subsequent expulsion of ThDP would yield succinate semialdehyde. However, with this assay we observed no change in absorbance at 340 nm due to the reduction of $NADP^+$ over the course of 1 h, indicating that the carbanion/enamine is protected from alternative reaction pathways. Given the extinction coefficient of NADPH, submicromolar concentrations of product should be detectable. The idea of a ThDP-dependent enzyme protecting the cofactor−substrate adduct from alternative pathways has recently been discussed explicitly by Kluger (*36*). There remains no evidence for succinate semialdehyde production by MenD. From a physiological perspective, the (essentially irreversible) decarboxylation of 2-oxoglutarate in the absence of isochorismate would deplete the substrate pool needed for converting isochorismate to SHCHC.

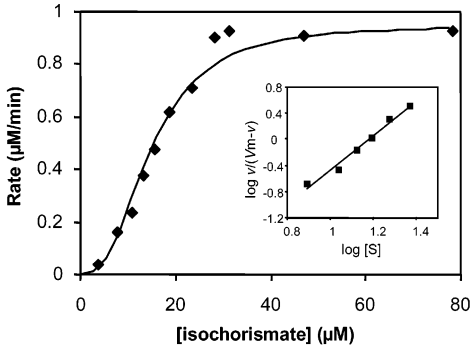


FIGURE 1: Variation of rate with isochorismate concentration in the presence of saturating concentrations of 2-oxoglutarate, ThDP, and $MnCl_2$ at pH 8.5. The line is generated using the Hill equation. Inset: Hill plot from which n_H was determined.

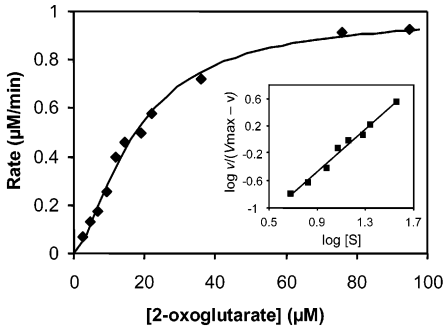


FIGURE 2: Variation of rate with 2-oxoglutarate concentration in the presence of saturating concentrations of isochorismate, ThDP, and $MnCl_2$ at pH 8.5. The line is generated using the Hill equation. Inset: Hill plot from which n_H was determined.

Table 1

substrate or cofactor	Hill coeff (n_H)	$S_{0.5}$ or K_M (μM)	$k_{cat}/S_{0.5}$ ($s^{-1}\,M^{-1}$)
isochorismate	2.6 ± 0.1	15.5 ± 0.3	$(1.1 \pm 0.1) \times 10^4$
2-oxoglutarate	1.6 ± 0.1	17.1 ± 0.2	$(1.0 \pm 0.1) \times 10^4$
ThDP	1	4.28 ± 0.09	$(4.0 \pm 0.3) \times 10^4$
Mn^{2+}	2.0 ± 0.1	$(9.5 \pm 0.3) \times 10^2$	180 ± 11

Some 2-oxo acid dehydrogenases, including AHAS (*37*) and PDC (*38*), are known to exhibit a kinetic "lag phase", or a slow approach to the zero-order steady state. These enzymes also show cooperativity with respect to substrate. This behavior is not a "rule" in the family: BFD and IPDC show normal initial rate behavior. MenD does exhibit a lag phase, and as shown in Figures 1 and 2, MenD shows cooperativity with respect to both substrates, isochorismate and 2-oxoglutarate. Analysis of the data using Hill plots gave the Hill coefficients shown in Table 1. Cooperativity makes the rate of SHCHC production much more sensitive to substrate concentration at levels near $S_{0.5}$. Both substrates

have other roles in the cell: 2-oxoglutarate is found in the TCA cycle and is the direct precursor of glutamic acid; isochorismate is the substrate for the first enzyme in the enterobactin biosynthetic pathway (*39*). Isochorismate is relatively unstable in aqueous solution, due to facile elimination to form aromatic compounds such as salicylate (*40*), and therefore sharp response to its presence would be advantageous. Tight control of the partitioning of these substrates under differing conditions within the cell may be optimized by this cooperativity. Nonhyperbolic kinetic behavior is well established for some ThDP-dependent enzymes, including PDC from yeast (*38*) and AHAS from *Aerobacter aerogenes* (*41*) and *Serratia maracens* (*42*).

While the rate dependence on ThDP fits a normal hyperbolic relation, as is typical of ThDP-dependent enzymes, the dependence on Mn^{2+} also showed cooperativity, as indicated in Table 1. A more complete study of the effects of varying metal ions on the activation of the enzyme is underway in our laboratory. The maximal velocity observed was 0.924 μM min^{-1}, corresponding to a $k_{cat} = 10$ min^{-1}.

Phylogenetics of MenD. Because menaquinone, and therefore presumably MenD, is used for anaerobic growth of bacteria, one might expect this protein to be present in many bacterial genomes. Indeed, if one considers that life on this planet likely arose in an oxygen-poor atmosphere, the menD gene would therefore be expected to be widespread and ancient. The presence of the sequence in only about one-fourth of the complete, annotated genomes reflects the reliance of many organisms, for example, aerobic Gram-negative bacteria, on ubiquinone rather than menaquinone in respiratory electron transport (*43*).

Careful examination of the protein databases revealed 40 complete sequences that can be identified as MenD proteins (and a few incomplete sequences). The accession numbers and complete species names (as annotated) are provided in a table in the Supporting Information. Of these, only two are not bacterial: one archaeal protein from the *Halobacterium* sp. NRC-1 and one plant protein from *Arabidopsis thaliana*. The *Arabidopsis* sequence is that of a large protein, of which only the C-terminal 564 residues resemble MenD from other sources. The 330-residue N-terminal region is apparently the fusion of another domain with a distinct function. PSI-BLAST searching using this 330 amino acid fragment suggests that this domain is an isochorismate synthase, consistent with fusion to the subsequent enzyme required for conversion of isochorismate toward vitamin K. The presence of a gene that encodes MenD in *Arabidopsis* has also been observed recently by others (*44*); however, we are the first to suggest that this gene encodes both isochorismate synthase and SHCHC synthase functions.

Sequence alignments indicate that MenD paralogues do not, in general, show high levels of sequence conservation. Closely related species (such as those of the same genus) do share high amino acid identities; e.g., MenD from *Salmonella typhimurium* shares 99% amino acid identity with MenD of *Salmonella enterica* and 90% with MenD of *E. coli* K12. But this value drops to 68% when compared to the MenD from *Yersinia pestis*, although this is also an Enterobacteriacea. To contrast this result with that of a well-conserved protein family, enolases from *E. coli* K12 and *Y. pestis* are 94% identical. When one compares the *E. coli* MenD sequence with those of *Haemophilis influenzae* and

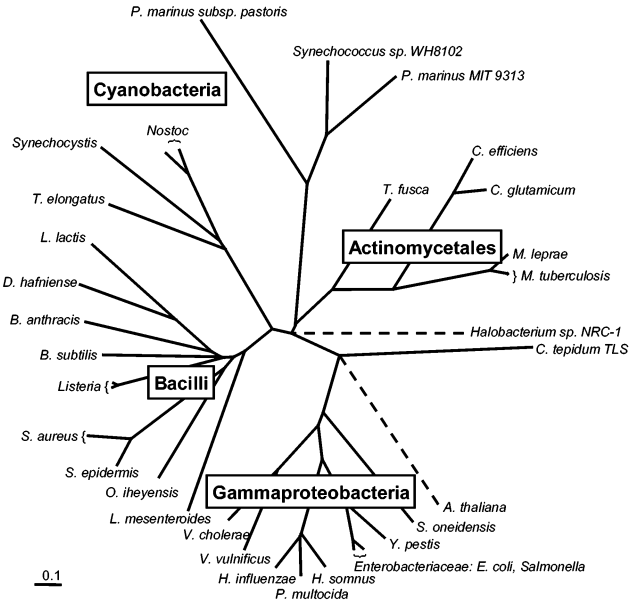


FIGURE 3: Unrooted phylogenetic tree of 40 MenD sequences. The tree represents the consensus of 100 neighbor-joined trees, with branch lengths indicating distance, measured in substitutions per site. The dotted branches are used to highlight the nonbacterial sequences. Because of space constraints, some species are denoted as part of a larger category, as indicated by the braces. Complete names of all species are listed in the Supporting Information.

Bacillus subtilis, the pairwise identities drop to 46% and 35%, respectively, compared to values for enolase of 86% and 60%. MenD protein sequences typically share between 20% and 40% identity.

When the evolutionary distance of the MenD sequences is calculated, the resulting phylogenetic tree, shown in Figure 3, indicates that MenD phylogeny corresponds well to taxonomy; i.e., species judged to be closely related by other criteria contain MenD protein sequences which group together. For example, the archaeal sequence does not group with any bacteria. An exception is the seven sequences available from cyanobacteria, which (although adjacent in Figure 3) form two distinct groups. This grouping of sequence in accordance with inferred taxonomy is supported by high bootstrap values within each clade. Interestingly, the *Arabidopsis* sequence does not group with either cluster of photosynthetic bacteria. The overall agreement of phylogeny and taxonomy lends confidence to the assigned function of these largely uncharacterized proteins.

Despite the overall divergence of sequence, there are many residues which are strictly conserved among MenDs. Alignment of MenD sequences with other ThDP-dependent decarboxylases allows a comparison with structurally characterized proteins such as AHAS, PDC, and BFD. The MenD sequence can thus be divided into three domains: a ThDP/pyrimidine-binding domain corresponding to the N-terminal ~180 residues (amino acids 1−180 in the *E. coli* sequence), a central domain of ~160 residues (181−341), and a ThDP/pyrophosphate-binding domain made up of the C-terminal ~220 residues (342−556). (The names of these domains are not intended to indicate their sole function in the protein.) Notable is that the central domain has only a few residues which are conserved in 50% or more of the proteins, much less conservation than in the N- and C-terminal domains. The sequence alignment in this region is much less reliable

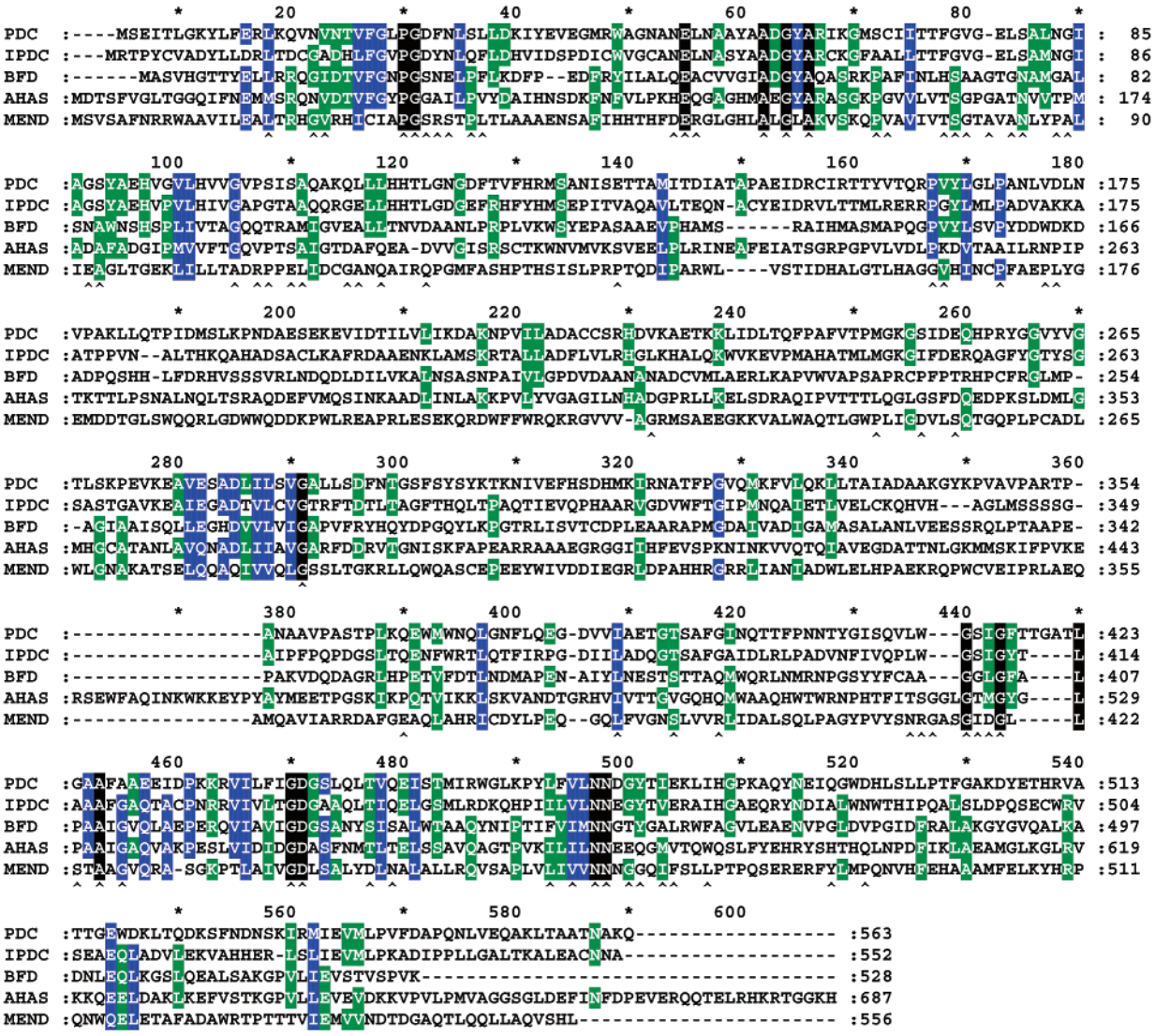


FIGURE 4: Sequence alignment of MenD from *E. coli* K12 with the structurally characterized proteins pyruvate decarboxylase from yeast (PDC), indolepyruvate decarboxylase from *Enterobacter cloacae* (IPDC), benzoylformate decarboxylase from *Pseudomonas putida* (BFD), and acetohydroxy acid synthase from yeast (AHAS). To save space, the first 84 residues of AHAS, which do not align with the other sequences, were removed. The sequence numbering nevertheless corresponds to the complete AHAS sequence. Color code: black, identical residues; blue, at least four of five residues are similar; green, three of five residues are similar. The chevron (∧) indicates a residue conserved among at least 80% of the 40 MenD sequences.

(i.e., alignments vary perceptibly with the choice of weight matrix). The alignment indicates that there are just over 40 absolutely conserved residues among the complete MenD sequences, but only two of these, glycine-227 and glycine-286, are found in the central domain. This divergence of sequence in the central domain mirrors the relationship of MenD to other members of the ThDP-dependent decarboxylase superfamily, as discussed below. Residues which are conserved in 80% or more of the MenD sequences examined are labeled in Figure 4.

Comparison of the MenD Sequence with Homologous Enzymes. An alignment of MenD with four structurally characterized members of the superfamily is shown in Figure 4. Typically, members of an enzyme superfamily maintain crucial catalytic residues despite an overall divergence in sequence (*45*). It has been noted before (*11*, *46*, *47*) that ThDP-dependent decarboxylases show little sequence conservation beyond those elements which bind and/or activate ThDP. Certain residues which are found at the active site, or have been shown by mutagenesis to be critical for catalysis, in ThDP-dependent decarboxylases are conserved in MenD, and it is not surprising that these are all residues which interact with the cofactors. For example, a glutamic acid residue is believed to donate a hydrogen bond to N1′ of the aminopyrimidine ring of ThDP in all superfamily members. There is considerable evidence that proton transfer to N1′, resulting in a 4′-imino tautomer, is a necessary part of ThDP catalysis; this imino group is the candidate base for abstraction of the proton attached to C2 of the thiazole ring (*11*, *15*, *48*). In MenD, an absolutely conserved glutamic acid residue, E55 in MenD *E. coli* numbering, aligns with this residue. E55 is part of a strictly conserved Asp-Glu-Arg sequence which can be used as a diagnostic feature of MenD sequences when comparing BLAST hits with low scores. We propose that E55 donates a hydrogen bond to N1′ of ThDP.

MenD is apparently most closely related to AHAS; i.e., alignment search tools such as PSI-BLAST return AHAS

sequences with higher scores than other superfamily members when MenD sequences are used as the query. These enzymes also share the carboligase activity for which other superfamily members have not evolved [although BFD has been used to catalyze benzoin condensations (*49*)]. It is therefore most instructive to compare and contrast features of the MenD sequence primarily with those of AHAS. The crystal structure of AHAS from yeast (*10*) shows that residues G115, E139, and T162 interact with the aminopyrimidine portion of ThDP. These residues are all conserved in MenD, corresponding to G31, E55 (discussed above), and T78 in the *E. coli* enzyme. The position occupied by T78 is sometimes a serine residue in MenD sequences, but the glycine and glutamic acid residues are invariant.

The structure of AHAS indicates that the divalent metal ion, which anchors the diphosphate moiety of ThDP, is coordinated by the side chains of D550 and N577. These residues are both conserved in MenD, aligning with D442 and N469. These residues are part of the "ThDP-binding motif" (*50*) which stretches from G441 to N469 in MenD. In the superfamily, this motif invariably begins with Gly-Asp and ends with Asn-Asn residues. In the family of MenD proteins, one of these two asparagine residues (but never both) is replaced by aspartate about 10% of the time. Interacting directly with the diphosphate of the cofactor in AHAS is the backbone of S552, which aligns with S444 in MenD. While not strictly conserved, this position is always occupied by serine, alanine, or in two cases threonine. The residue at this position is a serine or alanine in PDC, IPDC, and BFD also.

AHAS apparently uses M525 to hold ThDP in the crucial V-conformation. This role is always fulfilled by a bulky hydrophobic residue among the superfamily members; e.g., I415 in yPDC, L403 in BFD, and I410 in IPDC. These enzymes all house this residue in a Gly-Ser-Xxx-Gly motif, where Xxx is the large hydrophobic residue. MenD, however, does not share this motif, having instead the stretch of residues Gly-Xxx-Yyy-Gly-Ile-Asp-Gly, where Xxx = Ala or Val, and Yyy = Ser or Asn. This causes a dilemma with regard to the appropriate alignment of glycine residues, since there are two GXXG motifs present; however, the absolutely conserved I418 may fill the role analogous to that of M525 in AHAS.

Chipman, Barak, and Schloss have suggested (*51*) that a conserved glutamine residue in AHAS, Q501 in the yeast protein, could be involved in binding of the second "acceptor" substrate in the AHAS reaction. This residue is not conserved across the superfamily. Interestingly, this residue does align with a strictly conserved arginine residue (R395 in *E. coli*) in MenD. This residue is found in a short stretch of hydrophobic residues which contains the mentioned arginine and a strictly conserved serine residue (S391). It may be that these residues are involved in binding isochorismate during the MenD reaction. The same researchers proposed another glutamine residue (Q202 in the yeast protein) to be involved in AHAS catalysis. This residue may also interact with the second substrate or with N4′ of the aminopyrimidine portion of ThDP. This residue aligns with a strictly conserved glutamine residue in MenD, Q118 in the *E. coli* protein. However, the alignment of the superfamily proteins in this region is poor and cannot be interpreted with confidence. Another strictly conserved glutamine residue, Q122, is nearby in MenD. The corresponding region of PDC is disordered in some structures, suggesting that there is a role in substrate binding (*52*).

G286, one of only two strictly conserved residues in the central domain of MenD, aligns with G374 of AHAS, G286 of yPDC, G284 of IPDC, and G274 of BFD. This residue was first shown in the structure of pyruvate oxidase from *Lactobacillus plantarum* (POX) (*53*) to introduce a 90° turn in the structure of this domain. This residue is conserved across the superfamily, despite the fact that the central domain seems to be used for widely differing purposes. In POX and AHAS, this domain is involved in binding FAD, required for catalysis by both enzymes, but not involved in the chemistry of the AHAS reaction.

In yPDC, D28 has been shown by site-directed mutagenesis to affect a postdecarboxylation step of the reaction, and D28A and D28N mutants show significant acetolactate synthase activity (*54*). This residue is conserved in IPDC, but in BFD the residue at this position is S26, proposed to donate a hydrogen bond to the carboxylate group of the substrate. Consistent with this residue mediating decarboxylase/carboligase activity, a glycine residue is found in this position in AHAS (and in glyoxalate carboligase; data not shown). However, in MenD a serine residue (S32) is found in this position, as in BFD, suggesting that the carboligase activity of MenD is not mediated in the same way.

Of equal interest are residues conserved in MenD which differ from those found in other enzymes, since the reaction catalyzed by MenD differs in many ways from those of all other superfamily members. R395, mentioned above, is one such residue. K292 is not strictly conserved in MenD, but a positively charged residue appears in this or the adjacent position in every sequence. This residue aligns with residues at the active site of PDC (F292) and BFD (Y280). Another arginine residue, R33, is adjacent to residues suggested to be at the active site, but no other enzymes in the superfamily have an arginine residue in this position. A conserved Arg-Pro-Xxx-Leu sequence beginning at position 107 bears no resemblance to other superfamily members. Given the number of negative charges found in the substrates and intermediates of the SHCHC synthase reaction, the number of conserved arginine and lysine residues is particularly intriguing.

Mutagenesis of Glutamate-55. The above discussion is predicated on the widely demonstrated concept that a sequence alignment can be used to predict the function of particular amino acids in the absence of structural or kinetic data, even when the aligned proteins share very little overall sequence identity. This case is an extreme one: MenD only shares between 8% and 14% sequence identity (23−28% similarity) with the proteins shown in Figure 4. As proof of principle, we chose E55 to test our understanding of MenD as a member of this superfamily. As mentioned above, the glutamate residue with which E55 aligns in each protein has been shown in crystallographic and mutagenesis experiments to be a key residue for catalysis. We generated the mutant enzyme E55Q, on the basis that the isosteric glutamine residue would allow binding of the ThDP residue, but the drastically different pK_a relative to glutamic acid would result in vastly lowered activity. The mutant enzyme could be expressed to the same extent as the wild-type enzyme under the same conditions, but the resulting protein showed no

detectable activity. Because isochorismate slowly decomposes in aqueous solution at 25 °C, assays could not be run for extended lengths of time to detect trace activity. We can say with confidence that E55Q shows less than 1% activity of the wild-type enzyme.

CONCLUSIONS

MenD, the second enzyme in the menaquinone biosynthetic pathway, catalyzes the ThDP- and Mn^{2+}-dependent formation of SHCHC without release of succinate semialdehyde and no requirement for FAD. We recommend that the use of EC 4.1.1.71 be discontinued for this enzyme, since the reaction described by this number is not apparently catalyzed by MenD. The substrate dependence of rate of product formation by this homodimeric protein is cooperative with respect to both isochorismate and 2-oxoglutarate. This enzyme has relatively low overall sequence conservation, particularly in the central domain of the protein. The sequence is found in a variety of bacteria and in at least one archaeon and one eukaryote, *A. thaliana*, although this plant sequence exists as a fusion protein. MenD phylogeny corresponds to inferred taxonomy.

MenD shares many of the common features of the homologous 2-oxo acid decarboxylases but catalyzes a carbon−carbon bond forming reaction quite unlike that of any other enzyme. This functional divergence is reflected in many aspects of the sequence which are well-conserved within the MenD family but not found in other members of the superfamily. Mutation of the glutamic acid residue predicted to be necessary for activation of ThDP resulted in an enzyme with no detectable activity.

Considerably more remains to be learned about this enzyme. In particular, isochorismate must be activated by the enzyme by Brønsted or Lewis acid catalysis in order to act as a Michael acceptor. We plan to pursue a course of research involving the crystallographic characterization of MenD, site-directed mutagenesis of many of the residues discussed above, the effects of varying divalent metal ions on enzyme activation, and the scope of carbon−carbon bond forming reactions that might be catalyzed by this enzyme.

ACKNOWLEDGMENT

The authors thank Dr. Janet Hill, National Research Council, Plant Biotechnology Institute, for helpful discussions regarding phylogenetic analyses; Dr. Joanne Turnbull, Concordia University, for the gift of chorismate and input regarding its biosynthesis in our laboratory; and Dr. John Gerlt, University of Illinois, for support and encouragement in the initiation of this work.

SUPPORTING INFORMATION AVAILABLE

Table of MenD sequence accession numbers and organism names. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. (a) Ugai, T., Dokawa, T., and Tsubokawa, S. (1943) *J. Pharm. Soc. Jpn. 63*, 269; (b) Breslow, R. (1958) *J. Am. Chem. Soc. 80*, 3719.
2. Shellenberger, A. (1998) *Biochim. Biophys. Acta 1385*, 177−186.
3. Hu, Q., and Kluger, R. (2002) *J. Am. Chem. Soc. 124*, 14858−14859.
4. Polovnikova, E. S., McLeish, M. J., Sergienko, E. A., Burgner, J. T., Anderson, N. L., Bera, A. K., Jordan, F., Kenyon, G. L., and Hasson, M. S. (2003) *Biochemistry 42*, 1820−1830.
5. Arjunan, P., Umland, T., Dyda, F., Swaminathan, S., Furey, W., Sax, M., Farrenkopf, B., Gao, Y., Zhang, D., and Jordan, F. (1996) *J. Mol. Biol. 256*, 590−600.
6. (a) Schütz, A., Sandalova, T., Ricagno, S., Hübner, G., König, S., and Schneider, G. (2003) *Eur. J. Biochem. 270*, 2312−2321; (b) Schütz, A., Golbik, R., Tittmann, K., Svergun, D. I., Koch, M. H. J., Hübner, G., and König, S. (2003) *Eur. J. Biochem. 270*, 2322−2331.
7. Muller, Y. A., and Schultz, G. E. (1993) *Science 259*, 965−967.
8. Gonzalez, B., and Vicuna, R. (1989) *J. Bacteriol. 171*, 2401−2405.
9. Cromartie, T. H., and Walsh, C. T. (1976) *J. Biol. Chem. 251*, 329−333.
10. Pang, S. S., Duggleby, R. G., and Guddat, L. W. (2002) *J. Mol. Biol. 317*, 249−262.
11. Jordan, F. (2003) *Nat. Prod. Rep. 20*, 184−201.
12. Meganathan, R. (1996) Biosynthesis of the isoprenoid quinones menaquinone (vitamin K2) and ubiquinone (coenzyme Q), in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, 2nd ed., pp 642−656, American Society for Microbiology, Washington, DC.
13. I. Rayment, V. Klenchin, and J. A. Gerlt, personal communication.
14. Emmons, G. T., Campbell, I. M., and Bentley, R. (1985) *Biochem. Biophys. Res. Commun. 131*, 956−960.
15. Hubner, G., Tittmann, K., Killenberg-Jabs, M., Schaffner, J., Spinka, M., Neef, H., Kern, D., Kern, G., Schneider, G., Wikner, C., and Ghisla, S. (1998) *Biochim. Biophys. Acta 1385*, 221−228.
16. Kerr, M. S., and de Alaniz, J. R. (2002) *J. Am. Chem. Soc. 124*, 1028−1029.
17. Stetter, H. (1975) *Angew. Chem., Int. Ed. Engl. 15*, 639.
18. Palaniappan, C., Taber, H., and Meganathan, R. (1994) *J. Bacteriol. 176*, 2648−2653.
19. Palaniappan, C., Sharma, V., Hudspeth, M. E., and Meganathan, R. (1992) *J. Bacteriol. 174*, 8111−8118.
20. Johnson, T. W., Naithani, S., Stewart, C., Zybailov, B., Jones, A. D., Golbeck, J. H., and Chitnis, P. R. (2003) *Biochim. Biophys Acta 1557*, 67−76.
21. Bates, D. M., Eiff, C., McNamara, P. J., Peters, G., Yeaman, M. R., Bayer, A. S., and Proctor, R. A. (2003) *J. Infect. Dis. 187*, 1654−1661.
22. http://brenda.bc.uni-koeln.de/.
23. http://ca.expasy.org/enzyme/.
24. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
25. Palmer, D. R. J., Garrett, J. B., Sharma, V., Meganathan, R., Babbitt, P. C., and Gerlt, J. A. (1999) *Biochemistry 38*, 4252−5258.
26. Rieger, C. E., and Turnbull, J. E. (1996) *Prep. Biochem. Biotechnol. 26*, 67−76.
27. Schmidt, K., and Leistner, E. (1995) *Biotechnol. Bioeng. 45*, 285−291.
28. Cornish-Bowden, A. (1995) *Analysis of Enzyme Kinetic Data*, Oxford University Press, New York.
29. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res. 25*, 3389−3402.
30. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
31. Felsenstein, J. (1989) *Cladistics 5*, 164−166.
32. Popp, J. L. (1989) *J. Bacteriol. 171*, 4349−4354.
33. Yamamoto, Y., Aiba, H., Baba, T., Hayashi, K., Inada, T., Isono, K., Itoh, T., Kimura, S., Kitagawa, M., Makino, K., Miki, T., Mitsuhashi, N., Mizobuchi, K., Mori, H., Nakade, S., Nakamura, Y., Nashimoto, H., Oshima, T., Oyama, S., Saito, N., Sampei, G., Satoh, Y., Sivasundaram, S., Tagami, H., Takahashi, H., Takeda, J., Takemoto, K., Uehara, K., Wada, C., Yamagata, S., and Horiuchi, T. (1997) *DNA Res. 4*, 91−113.
34. Blattner, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science 277*, 1453−1474.

35. Kwon, O., Bhattacharyya, D. K., Palaniappan, C., and Meganathan, R. (1997) *Abstr. Gen. Meet. Am. Soc. Microbiol.*, 356.
36. Hu, Q., and Kluger, R. (2002) *J. Am. Chem. Soc. 124*, 14858−14859.
37. Bar-Ilan, A., Balan, V., Tittmann, K., Golbik, R., Vyazmensky, M., Hubner, G., Barak, Z., and Chipman, D. M. (2001) *Biochemistry 40*, 11946−11954.
38. Sergienko, E. A., and Jordan, F. (2002) *Biochemistry 41*, 3952−3967.
39. Gehring, A. M., Mori, I., and Walsh, C. T. (1998) *Biochemistry 37*, 2648−2659.
40. Young, I. G., Batterham, T. J., and Gibson, F. (1969) *Biochim. Biophys. Acta 177*, 389−400.
41. Störmer, F. C. (1968) *J. Biol. Chem. 243*, 3735−3739.
42. Yang, J. H., and Kim, S. S. (1993) *Biochim. Biophys. Acta 1157*, 178−184.
43. Meganathan, R. (2001) *Vitam. Horm. 61*, 173−218.
44. Lange, B. M., and Ghassemian, M. (2003) *Plant Mol. Biol. 51*, 925−948.
45. Babbitt, P. C., and Gerlt, J. A. (1997) *J. Biol. Chem. 272*, 30591−30594.
46. Hasson, M. S., Muscate, A., McLeish, M. J., Polovnikova, L. S., Gerlt, J. A., Kenyon, G. L., Petsko, G. A., and Ringe, D. (1998) *Biochemistry 37*, 9918−9930.
47. Muller, Y. A., Lindqvist, Y., Furey, W., Schulz, G. E., Jordan, F., and Schneider, G. (1993) *Structure 1*, 95−103.
48. Jordan, F., Zhang, Z., and Sergienko, E. (2002) *Bioorg. Chem. 30*, 188−198.
49. (a) Dünkelmann, P., Kolter-Jung, D., Nitsche, A., Demir, A. S., Siegert, P., Lingen, B., Baumann, M., Pohl, M., and Müller, M. (2002) *J. Am. Chem. Soc. 124*, 12084−12085; (b) Iding, H., Dünnwald, T., Greiner, L., Liese, A., Müller, M., Siegert, P., Grötzinger, J., Demir, A. S., and Pohl, M. (2000) *Chem. Eur. J. 6*, 1483−1495.
50. Hawkins, C. F., Borges, A., and Perham, R. N. (1989) *FEBS Lett. 255*, 77−82.
51. Chipman, D., Barak, Z., and Schloss, J. V. (1998) *Biochim. Biophys. Acta 1385*, 401−419.
52. Lu, G., Dobritzsch, D., Baumann, S., Schneider, G., and König, S. (2000) *Eur. J. Biochem. 267*, 861−868.
53. Muller, Y. A., Schumacher, G., Rudolph, R., and Schulz, G. E. (1994) *J. Mol. Biol. 237*, 315−335.
54. (a) Sergienko, E. A., and Jordan, F. (2001) *Biochemistry 40*, 7369−7381; (b) Wu, Y. G., Chang, A. K., Nixon, P. F., Li, W., and Duggleby, R. G. (2000) *Eur. J. Biochem. 267*, 6493−6500.

BI035224J