Three-Dimensional Structure of Methionyl-tRNA Synthetase from *Pyrococcus abyssi*[†,‡]

Thibaut Crepin, Emmanuelle Schmitt,* Sylvain Blanquet, and Yves Mechulam

Laboratoire de Biochimie, Unité Mixte de Recherche 7654, CNRS-Ecole Polytechnique, F-91128 Palaiseau Cedex, France

Received September 9, 2003; Revised Manuscript Received December 19, 2003

ABSTRACT: In class 1 aminoacyl-tRNA synthetases, methionyl-tRNA synthetases (MetRS) are homodimers or monomers depending on the presence or absence of a domain appended at the C-side of the polypeptide chain. Beyond this C-domain, all MetRS display a highly conserved catalytic core with a Rossmann fold, the two halves of which are linked by a connective peptide (CP). Three-dimensional folding of CP and its putative zinc content have served as a basis to propose a division of the MetRS family into four subgroups. All subgroups but one, which is predicted to display two zincs per MetRS polypeptide, have been characterized. In the present study, the 3D structure of MetRS from *Pyrococcus abyssi* could be solved at 2.9 Å resolution. The data obtained and atomic absorption spectroscopic measurements establish the presence of two metal ions per polypeptide chain. This finding brings strong support to the above classification. In the crystal, the C-terminal dimerization domain is disordered. This observation is thought to reflect marked flexibility of the two core moieties with respect to the C-domains in the dimer. Gel shift experiments were performed with the isolated C-terminal dimerization domain and a core monomeric MetRS, both derived from the *P. abyssi* enzyme. Complex formation between the C-domain and the core enzyme could not be evidenced. Moreover, association of $tRNA^{Met}$ to the core enzyme is enhanced in the presence of the C-domain. Together, these experiments suggest positive control in trans by the C-domain on recognition of tRNA by the core moiety of MetRS.

Enzymes belonging to the aminoacyl-tRNA synthetase (aaRS) family catalyze the esterification of amino acids on their specific tRNAs. aaRS are divided into two classes depending on the structure of their catalytic center (*1*, *2*). Class 1 synthetases are characterized by a catalytic center built around a Rossmann fold harboring the two signature sequences HIGH and KMSKS, while class 2 enzymes exhibit a central antiparallel β-sheet containing three conserved sequence stretches.

In class 1, methionyl-tRNA synthetase (MetRS)[1] is particular since depending on the living species from which it is isolated, this enzyme behaves as a monomer or as a homodimer (*3*). Dimer formation of MetRS is always related to the presence, at the C-side of the core enzyme, of an appended domain. Interestingly, some organisms express paralogues of this dimerization domain. Such paralogues are, for example, Trbp111 (tRNA-binding protein) of *Aquifex aeolicus* (*9*), the cold-shock protein CsaA from *Bacillus subtilis* (*10*), yeast Arc1p [aminoacyl-tRNA cofactor protein (*11*)], and the mammalian cytokine endothelial monocyte activating polypeptide II [EMAP II (*12*)].

At a functional level, the existence of dimeric MetRS is intriguing. Indeed, it was early demonstrated that dimeric *Escherichia coli*, *Bacillus stearothermophilus*, and *Thermus thermophilus* MetRS could each be converted by C-terminal truncation into a monomeric minimal core of nearly 550 residues. The obtained monomers retain activity in the reaction of aminoacylation of $tRNA^{Met}$ (*4*−*6*). The 3D structures of such shortened units derived from *E. coli* and *T. thermophilus* MetRS (MetRSec and MetRStt, respectively) have been solved (*7*, *8*). Four distinct domains are encountered: a Rossmann fold, a connective polypeptide (CP) linking the two halves of the Rossmann fold, the KMSKS domain, and an α-helix bundle responsible for the binding of the anticodon of $tRNA^{Met}$.

The dimerization domain of methionyl-tRNA synthetase from *Pyrococcus abyssi* (MetRSpa) could be recently isolated and crystallized (*13*). This domain (MetRS-C12K) shows an OB-fold such as that found in several proteins having RNA binding properties (*9*, *14*−*16*). Indeed, when present at the C-end of the enzyme, MetRS-C12K improves $tRNA^{Met}$ affinity (*17*). To possibly explain this behavior, data on the relative 3D locations of the dimerization domain and of the remainder of the enzyme have to be obtained.

In the case of *E. coli* MetRS, a tightly bound zinc ion was shown to play a role in the folding of the CP region (*18*). The functional importance of this zinc has been highlighted by site-directed mutagenesis experiments (*18*). These studies and the crystal structure of MetRSec have suggested that, to participate to catalysis, the zinc binding

[†] T.C. was a recipient of a doctoral fellowship from the Association pour la Recherche sur le Cancer.

[‡] PDB accession number 1RQG.

* Corresponding author. Tel: +33 1 69334885. Fax: +33 1 69333013. E-mail: emma@botrytis.polytechnique.fr.

[1] Abbreviations: MetRS, methionyl-tRNA synthetase; MetRSpa, MetRS from *Pyrococcus abyssi*; MetRSec, MetRS from *Escherichia coli*; MetRStt, MetRS from *Thermus thermophilus*; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

10.1021/bi0356247 CCC: $27.50
Published on Web 02/10/2004

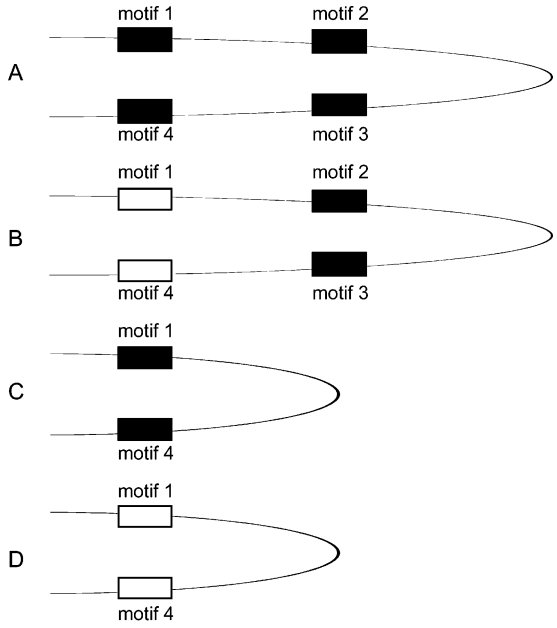

FIGURE 1: Schematic structural organization of the zinc binding regions of MetRS from various origins as deduced from structural and sequence alignments (*7*). A maximum of two structural knuckles can be formed. The proximal knuckle is formed by motifs 1 and 4, and the distal knuckle is formed by motifs 2 and 3. Filled boxes represent motifs containing putative zinc binding ligands whereas open boxes represent motifs unable to bind zinc. Family A contains two knuckles with two putative zinc ions. Family B contains two knuckles, but only the distal knuckle can accommodate the metal. Family C contains one knuckle with one zinc ion, and family D contains one knuckle devoid of metal.

motif should move with respect to the active site (*7*, *19*). Moreover, several lysines at the vicinity of the metal were shown to react with a 3′-oxidized $tRNA^{Met}$ derivative (*20*). This result indicates an involvement of the zinc region of *E. coli* MetRS in the binding of the acceptor stem of tRNA.

Actually, MetRS from various sources differ in their zinc content. Sequence analyses of MetRS could be made on the basis of the structural alignment of MetRSec and MetRStt. Four families were recognized according to the structural organization of the CP domain (*7*) (Figure 1). Moreover, each family is predicted to harbor a fixed number of zinc ions. Representatives of three families have been characterized (*7*, *8*, *21*, *22*). Yeast cytoplasmic MetRS belongs to a fourth family. However, this enzyme was reported to contain one zinc instead of two as predicted by the sequence comparisons (*22*). In this context, it is of interest to analyze another putative member of the fourth MetRS family to check the validity of the classification.

MetRS from *P. abyssi* is a suitable target to address unsolved questions regarding the properties of MetRS. Under its native form, this enzyme is homodimeric. Its C-terminal dimerization domain (MetRS-C12K), of known 3D structure, is available for biochemical studies (*13*). Moreover, this enzyme is predicted to belong to the family of MetRS harboring two knuckles with two zinc atoms. The present work describes the biochemical characterization and crystallization of dimeric MetRSpa. A 3D model of the protein is drawn from X-ray analysis of the crystals. Two zinc ions per protomer are evidenced, therefore validating the proposed existence of MetRS polypeptides carrying two structural zinc ions. In the 3D model, the dimerization domain is disordered. This suggests that, in each MetRS protomer, the core enzyme and the dimerization domain are articulated around a flexible linker region. Finally, it is demonstrated that the isolated MetRS-C12K domain can form a complex with tRNA and a monomeric truncated MetRS equivalent to the core enzyme region. This observation suggests a role of the C-domain in the adjustment of the tRNA conformation toward tight recognition by the core moiety of MetRS.

MATERIALS AND METHODS

Production and Purification of P. abyssi Methionyl-tRNA Synthetase. The *metS* gene of *P. abyssi* was previously cloned between the *Nco*I and *Sac*II sites of pET15blpa to yield pET15bMetpa (*13*). This plasmid produces both the expected full-length MetRS plus its C-terminal domain. Translation of the C-domain starts at the ATG codon at position 616 in the ORF. To increase the expression level of native MetRSpa, an AGGAGG Shine−Dalgarno sequence internal to the *metS* reading frame and located upstream from the ATG at position 616 was changed into AAGAAG by site-directed mutagenesis. The two nucleotide changes do not modify the amino acid sequence of MetRSpa. The resulting plasmid was called pET15bSDMetpa.

E. coli BL21(De3) (Novagen Inc.) competent cells were transformed with pET15bSDMetpa plasmid for protein production. Transformed cells were cultured overnight at 37 °C in 1 L of 2 × TY medium supplemented with 50 $\mu g \cdot mL^{-1}$ ampicillin plus 1 mM isopropyl thiogalactopyranoside (IPTG). Cells were harvested by centrifugation. The pellet was resuspended in buffer A containing 10 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS), pH 6.7, 10 mM 2-mercaptoethanol, 50 mM NaCl, and 0.1 mM phenylmethanesulfonyl fluoride (PMSF) and disrupted by sonication. The cell extract was centrifuged (20 min, 20000*g*, 4 °C). The supernatant was heated (20 min, 80 °C), and the precipitated material was removed by centrifugation. Nucleic acids were precipitated with streptomycin (3% w/v). After precipitation with ammonium sulfate (80% saturation), proteins were resuspended in buffer A, dialyzed, and loaded on a Q-Hiload column (16 × 100 mm, Amersham Biosciences) equilibrated in the same buffer. MetRSpa was eluted with a linear gradient of 0.05−0.6 M NaCl (2 mL/min; 0.55 M/h) in buffer A. Fractions containing MetRS were pooled, dialyzed against buffer A, and loaded on a S-Sepharose column (11 × 40 mm; Amersham Biosciences) equilibrated in buffer A. The enzyme was eluted with a linear gradient of 0.05−0.6 M NaCl in buffer A (1 mL/min; 0.95 M/h). Pooled fractions were loaded at 0.4 mL/min on a Superdex 200 Hiload column (16 × 600 mm, Amersham Biosciences) equilibrated with 10 mM MOPS, pH 6.7, 10 mM 2-mercaptoethanol, and 500 mM NaCl. The purified protein (3 mg) was dialyzed against buffer A and concentrated to about 7 $mg \cdot mL^{-1}$ using a Vivaspin 30000 MWCO concentrator.

For the preparation of the M606 enzyme, the TTA codon for L607 was replaced by a stop TGA codon by using site-directed mutagenesis. The truncated enzyme expressed in *E. coli* BL21(De3) was purified as described above for the full-length enzyme. At the last stage of its purification on the molecular sieving Superdex 200 column, the M606 enzyme (5 μM) eluted as a single sharp peak at a volume of 72.8 mL, which markedly differs from the elution volume

Table 1: Native Data Used in the Structure Determination of MetRSpa[a]

X-ray source	ESRF-ID14-EH1
wavelength (Å)	0.934
unique reflections	20834
resolution (Å)	2.9
completeness (%)	99 (99.8)
redundancy	5.4
$R_{sym}(I)$ (%)	8.7 (35.8)
R/free *R* (%)	24.3/29.6
rms deviation	
bonds (Å)	0.008
angles (deg)	1.5
mean *B* values (Å^2)	
protein atoms	56
water	44

[a] The data set was collected at 100 K with a single crystal. The values in parentheses correspond to the highest shell of resolution (2.98−2.9 Å). $R_{sym}(I) = \sum_{hkl}\sum_i|\langle I_{hkl}\rangle - I_{hkl,i}|/\sum_{hkl}\sum_i|I_{hkl}|$, where *i* is the number of reflection *hkl*.

of native dimeric MetRSpa (64.8 mL). This shows that M606 behaves as a monomer.

Crystallization and X-ray Analysis. Initial crystallization screening experiments were carried out at 297 K by the hanging-drop vapor diffusion technique using the commercially available Crystal Screen I and Crystal Screen II kits (Hampton Research). Small crystals could be observed after 48 h under two conditions containing PEG 20000. Optimized crystallization conditions were defined as 0.1 M Mes, pH 6.5, 3.5% PEG 6000, and 2% dioxane. Small crystals could be observed after less than 2 h, and growth continued during about 2 weeks. Crystals belong to space group $I4_122$ with unit cell parameters $a = b = 113.2$ Å, $c = 287.0$ Å, and $\alpha = \beta = \gamma = 90°$.

For data collection, crystals were transferred in three steps into mother liquor solutions containing increasing concentrations of PEG 400 (from 10% to 30%) and flash-cooled in a nitrogen gas cryostream at 100 K. Data to 2.9 Å resolution were collected from a single crystal at the ID14-EH1 beamline of the European Synchrotron Radiation Facility (Grenoble) using an ADSC Quantum 4 CCD detector. Diffraction images were integrated using MOSFLM (*23*) and data reduced using the CCP4 suite (*24*) (Table 1).

Structure Determination and Refinement. The structure was solved by molecular replacement [AMORE program (*25*)] by using as a starting model *T. thermophilus* MetRS [PDB code 1A8H (*8*)], from which the zinc binding region (residues 114−165) has been removed. The structure was refined by cycles of manual model building using O (*26*) and conjugate gradient minimization using CNS (*27*). Final *R* and free *R* factors are 24.3% and 29.6%, respectively, for all data to 2.9 Å resolution. The final model shows good geometry (Table 1) and contains 606 residues out of the 722 composing the MetRSpa protein, 51 water molecules, and 2 zinc ions. The mean *B*-value of 56 Å^2 is consistent with the deduced value from the Wilson plot (74 Å^2) and reflects the presence of a disordered part within the structure. Only highly ordered structural waters were added.

Fluorescence at Equilibrium and Atomic Absorption Spectroscopy. Variations of the intrinsic fluorescence of MetRSpa (1 μM) upon titration with methioninyl adenylate [laboratory stocks (*28*)] were followed as described (*29*), at 54 °C, in 20 mM Tris-HCl (pH 7.6), 10 mM 2-mercaptoethanol, and 0.1 mM EDTA. Fluorescence showed a linear increase up to a plateau level of 125%. Concentration of active sites in MetRSpa was deduced from the total concentration of added methioninyl adenylate at which the plateau was reached.

Zinc atomic absorbancies were followed at a wavelength of 213.9 Å, using a Varian SpectrAA220 spectrophotometer as described (*30*). Prior to metal analysis, MetRSpa was extensively dialyzed against a buffer containing 10 mM potassium phosphate, pH 7.1, 10 mM 2-mercaptoethanol, 150 mM KCl, and either 0.1 mM EDTA or 6 μM $ZnCl_2$.

Enzyme Activity Measurements. Methionine-dependent [^{32}P]PP_i−ATP exchange activity was measured in standard buffer (20 mM Tris-HCl, 7 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, pH 7.6) containing 2 mM [^{32}P]-PP_i, 2 mM ATP, and 2 mM L-methionine (*29*). Aminoacylation assays (*31*) were performed in the same standard buffer plus 150 mM KCl, in the presence of 2 mM ATP, 50 μM L-[^{14}C]methionine (Amersham; 1.6 TBq/mol), and either 5.2 μM *E. coli* $tRNA_f^{Met}$ or 7 μM *P. abyssi* $tRNA_i^{Met}$. Both tRNAs were produced in *E. coli* from synthetic genes and purified, as described (*32*). Methionine acceptances of these two tRNAs were 1250 pmol per A_{260} unit and 800 pmol per A_{260} unit, respectively.

Gel Shift Assay for tRNA Binding. The gel shift assay was performed as described in ref *14*. Electrophoresis was on native gels containing 10% acrylamide in 50 mM Tris−borate, pH 8.3, 5 mM $MgCl_2$, and 0.0001% Triton X-100. Samples (10 μL) in 50 mM Tris−borate, pH 8.3, 5 mM $MgCl_2$, and 0.001% Triton X-100 were incubated for 20 min at 4 °C. Before being loaded onto the gel, the sample was supplemented with 3 μL of 40% saccharose. The gels were run at 100 V for 3 h at 4 °C.

RESULTS AND DISCUSSION

Production and Characterization of Dimeric MetRS from P. abyssi. The *metS* gene encoding MetRS from the hyperthermophilic archaeon *P. abyssi* has been previously cloned in the T7-based expression vector pet15blpa (*13*). The resulting plasmid led to the expression of two polypeptides. One is dimeric MetRSpa, while the other exactly corresponds to the C-terminal dimerization domain (MetRS-C12K) (*13*). The unexpected high-level expression of the latter domain results from the presence of a Shine−Dalgarno sequence upstream to an internal ATG corresponding to the 616th codon in *metS*. To favor overexpression of full-length MetRS and prevent that of the C-domain, this Shine−Dalgarno sequence was mutated without modifying the amino acid sequence of the protein (see Materials and Methods). As expected, the mutations increased the production of native MetRS and canceled that of the C-terminal dimerization domain (Figure 2). MetRSpa was purified to homogeneity (see Materials and Methods). According to molecular sieving and native polyacrylamide gel analysis, it behaves as a dimer. Activities of the purified enzyme in the tRNA aminoacylation and the [^{32}P]PP_i−ATP isotopic exchange reactions were measured (Table 2). The rate of the [^{32}P]PP_i−ATP isotopic exchange increases by a factor of 34 upon raising the temperature of the assay from 25 to 60 °C. MetRSpa can transfer methionine on either initiator $tRNA_f^{Met}$ from *E. coli* or initiator $tRNA_i^{Met}$ from *P. abyssi* produced in *E. coli*. The

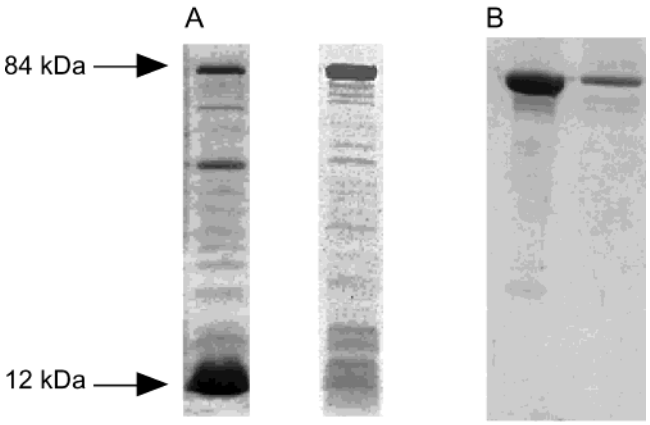

FIGURE 2: SDS−PAGE analysis of MetRSpa. Panel A shows the consequence of the presence of the internal Shine−Dalgarno sequence on overexpression of full-length MetRS of *P. abyssi*. Heated cell extracts were loaded on a 12.5% SDS−PAGE gel. Proteins were stained with Coomassie Blue. Left: *E. coli* BL21 transformed with pet15bmpa. Right: *E. coli* BL21 transformed with pet15bSDmpa carrying the modified *metS* gene. Panel B shows that the protein has remained intact in the crystals. Left: purified MetRSpa. Right: protein from washed and dissolved crystals.

Table 2: Initial Velocities (v_i) of MetRSpa in tRNA Aminoacylation and Isotopic Exchange [^{32}P]PP$_i$−ATP Reactions Measured at Various Temperatures[a]

	v_i (10^{-2} s^{-1})		v_i (s^{-1})
T (°C)	*E. coli* tRNA$_f^{Met}$ aminoacylation	*P. abyssi* tRNA$_i^{Met}$ aminoacylation	isotopic exchange [^{32}P]PP$_i$−ATP
25	<0.1	<0.1	0.10 ± 0.01
42	1.0 ± 0.2	4.0 ± 0.3	0.60 ± 0.03
60	4.2 ± 0.2	18.4 ± 0.5	3.4 ± 0.2

[a] Rates were determined in the presence of 12.5 nM dimeric MetRSpa at 42 and 60 °C or in the presence of 250 nM dimeric MetRSpa at 25 °C. tRNA concentrations used were 5.2 μM for *E. coli* tRNA$_f^{Met}$ and 7 μM for *P. abyssi* tRNA$_i^{Met}$.

tRNA aminoacylation rate increases by a factor greater than 150 upon changing the temperature from 25 to 60 °C.

Determination of Zinc Associated to MetRS. Structural alignment as well as sequence analyses of MetRS from various sources allowed us to propose a partition of MetRS enzymes into four families according to the structural organization of their zinc binding domain (*7*) (Figure 1). The MetRSpa polypeptide can be predicted to belong to a family with two knuckle structures, each liganding one zinc atom (Figure 1A). The two knuckles display clusters of cysteines (C127, C173, C176, D130 and C143, C146, C156, C159) for the coordination of the metal. Notably, no MetRS possessing two zinc ions per protomer could yet be evidenced experimentally.

To measure the zinc content of MetRSpa, we needed to determine accurately the molar absorbance coefficient of this enzyme. To achieve this, the active sites allowing methionyl adenylate formation were titrated. We took advantage of the early observation that methioninyl adenylate (an analogue of methionyl adenylate) binds *E. coli* MetRS with a very high affinity while causing an increase of the intrinsic protein fluorescence by 40% (*29*). As in the case of the *E. coli* enzyme, addition of methioninyl adenylate to MetRSpa promotes an increase of the intrinsic fluorescence up to a saturation level of 125%. Therefore, by titrating the enzyme with methioninyl adenylate and following the fluorescence signal, a concentration of active sites could be deduced. This experiment was performed at 54 °C. By assuming two active sites per dimer, an absorbance coefficient at 280 nm of 15.75 $\times$ 10^4 L·mol^{-1}·cm^{-1} was calculated and used in the rest of this study. This coefficient differs from the theoretical value calculated using the DNAid program [11.83 $\times$ 10^4 L·mol^{-1}·cm^{-1} (*33*)]. The reason for this discrepancy is likely to reflect the folding of aromatic residues in the MetRS 3D structure. Indeed, a similar discrepancy occurs between the measured (10.57 $\times$ 10^4 L·mol^{-1}·cm^{-1}) and the calculated (8.96 $\times$ 10^4 L·mol^{-1}·cm^{-1}) absorbance coefficients of MetRSec (*17*).

Table 3: Zinc Content in MetRS Protomers[a]

	+EDTA	+$ZnCl_2$
M547	1.0 ± 0.1	0.8 ± 0.2
M676	1.1 ± 0.1	0.8 ± 0.2
MetRSpa	2.0 ± 0.2	1.9 ± 0.1

[a] M547 and M676 are the truncated monomeric and the native dimeric forms of *E. coli* MetRS, respectively. Zinc content was measured by atomic absorption spectroscopy after dialysis against a buffer containing either 0.1 mM EDTA or 6 μM $ZnCl_2$. The calculated numbers of zinc ions per each protomer are indicated. In all cases, experimentally measured absorption molar coefficients were used for enzyme concentration calculations (ref *29* and this study).

The zinc content of MetRSpa was determined by atomic absorption spectroscopy after dialysis against a buffer containing either 0.1 mM EDTA or 6 μM $ZnCl_2$. Monomeric M547 and dimeric M676 *E. coli* MetRS were used as controls. As shown in Table 3, a stoichiometry of two zinc ions per protomer of MetRSpa was found whatever the dialysis buffer used. This result establishes that each protomer tightly binds two zinc ions and confirms the prediction of the proposed classification (Figure 1).

Notably, although monomeric yeast cytoplasmic MetRS can also be predicted to possess two zinc ions, only one tightly bound zinc atom was detected in this enzyme by X-ray fluorescence spectrometry (*22*). In view of the present results, the possibility that yeast cytoplasmic MetRS carries a second weaker zinc binding site deserves consideration. Interestingly, all archaeal and eukaryal cytoplasmic MetRS belong to the family with two zinc ions. Taking into consideration the structural role of the zinc ion evidenced in the case of MetRSec (*18*), it is likely that the presence of two zinc ions reinforces the stability of these MetRS.

Crystal Structure of MetRSpa. (A) Structure Solution. *I*-centered tetragonal crystals of MetRSpa were obtained by using 5% PEG 6000 as the precipitating agent in the presence of 2% dioxane and 0.1 M Mes−NaOH, pH 6.5 at 24 °C. Crystals belong to space group $I4_122$ with cell parameters $a = b = 113.2$ Å and $c = 287$ Å. Cryoprotection of the crystals could be obtained, though in an irreproducible manner, by short-time soaking in mother liquor containing increasing concentrations of PEG 400 from 10% to 30%. A complete data set to 2.9 Å resolution was recorded at 100 K at the ESRF ID14-EH1 beamline. The structure was solved by molecular replacement by using the coordinates of the monomeric *T. thermophilus* MetRS (*8*) deprived of its zinc binding domain as the starting model. After manual adjustment and refinement, a 3D model corresponding to residues 1−606 of MetRSpa could be traced in the electronic density. Unfortunately, despite the fact that its 3D structure is already

FIGURE 3: Structure of the 1−606 portion of *P. abyssi* methionyl-tRNA synthetase. The Rossmann fold (residues 1−97 and 251−336) is in dark gray, the CP domain (residues 98−250) is in light gray with the positions of the two zinc atoms indicated, the KMSKS domain (residues 337−396 and 548−562) is in gray, and the anticodon binding domain (residues 397−548) is in black. The three additional helices at the C-terminal extremity (residues 562−601) are in middle gray. Some relevant side chains of residues discussed in the text are shown. This figure as well as Figures 5−7 was drawn using Setor (*44*).

known (*13*), the C-terminal domain could not be associated to any electronic density. However, the crystal packing should have allowed its placement. The integrity of the length of the crystalline protein was verified by dissolution of crystals and analysis on SDS−PAGE. As shown in Figure 2, crystals of MetRSpa contain intact native protein. Therefore, the absence of electronic density for the C-domain must be explained by abnormal mobility inside the crystal lattice.

The structure of the 1−606 part of MetRSpa is represented in Figure 3. As in the case of *E. coli* and *T. thermophilus* MetRS, the enzyme displays four structural domains: the Rossmann fold with the catalytic center, the connective peptide (CP1) linking the two halves of the Rossmann fold and carrying the zinc binding knuckles, the domain containing the KMSKS signature sequence, and the α-helix bundle responsible for the binding of the anticodon of tRNA. The core of MetRSpa (excluding the zinc binding region) could be superimposed to those of *E. coli* and *T. thermophilus* with rms values of 1.7 Å for 481 Cα atoms aligned and 1.6 Å for 406 Cα atoms aligned, respectively.

Upon alignment of the sequence of MetRSpa with those of other dimeric MetRS, an obvious difference is the presence of an insertion (residues 561−606) between the C-end of the active core common to all MetRS and the N-end of the dimerization domain. This insertion is specific of MetRS from organisms belonging to the *Pyrococcus* subgroup. In MetRSec, the C-terminus of the active core is made by the α15 helix, which protrudes out of the α-helix bundle domain and contacts the KMSKS domain. The α15 helix is absent from MetRStt. In MetRSpa, the inserted region makes three additional helices, α16, α17, and α18. Together with α15, they form two groups of antiparallel helices, with α15−α16 perpendicular to α17−α18. This group of four α-helices can be considered as a part of the KMSKS domain. Unfortunately, no density accounts for residues beyond α18. Likely, residues 607−615 form a flexible region, linking α18 to the C-terminal dimerization domain. This linker region is rich in K and E residues. Proteolytic cleavage in this region generates the well-known active truncated monomeric form of MetRS (*4*−*6*).

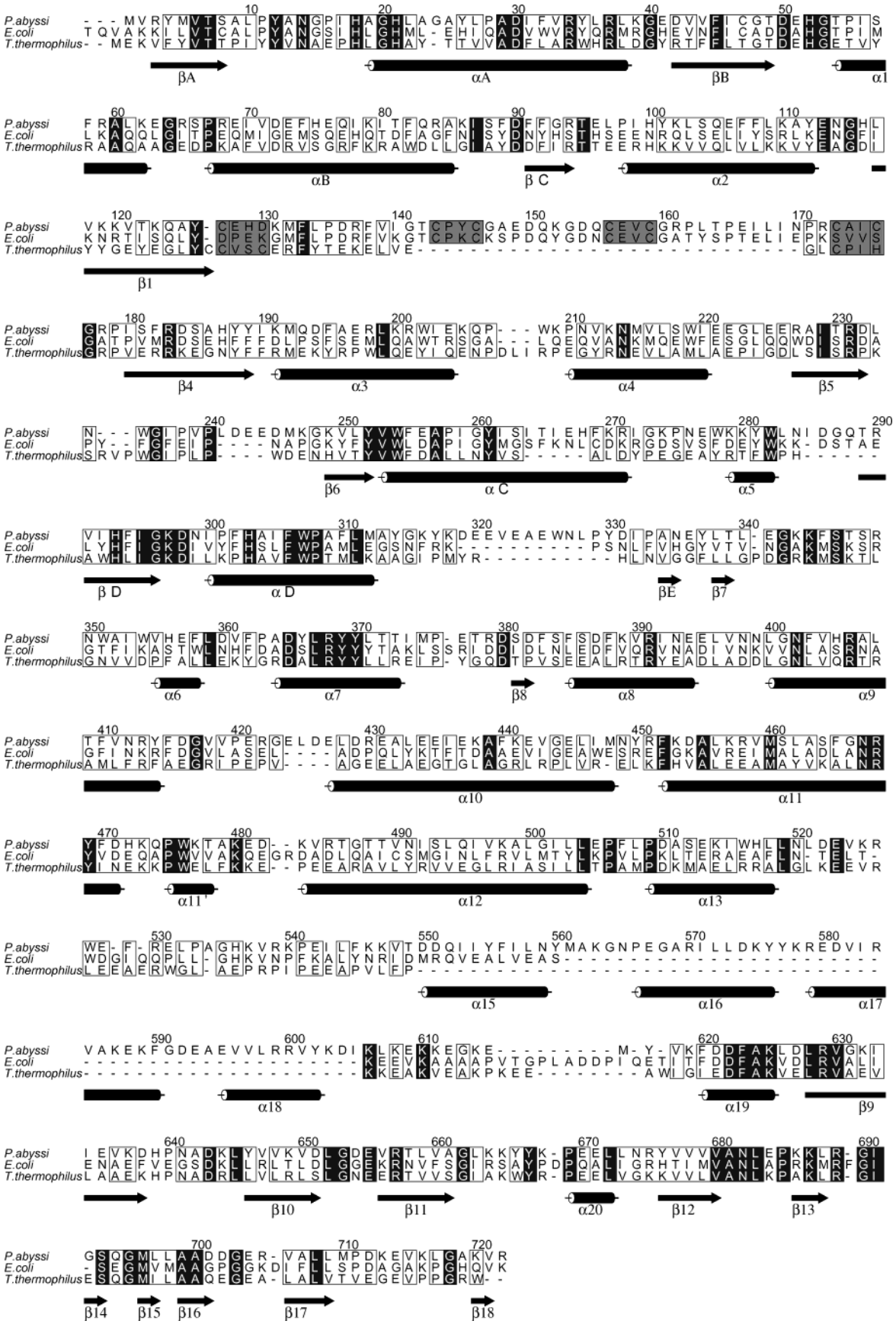

FIGURE 4: Sequence alignment of the MetRS from *P. abyssi*, *E. coli*, and *T. thermophilus* based on the 3D structures of these three enzymes (residues 1−606; this work and refs *7* and *8*). The secondary structure elements of MetRSpa are indicated below the aligned sequences. For the dimerization domain (residues 616−751), the secondary structure elements are based on the 3D structure of the P12K domain of MetRSpa (*13*). Numbering refers to the MetRSpa sequence. Residues strictly conserved in all proteins are boxed in black whereas conservative replacements are simply boxed. Motifs 1−4 comprising the two structural knuckles (Figure 1) are boxed in gray. This figure was drawn using Alscript (*45*).

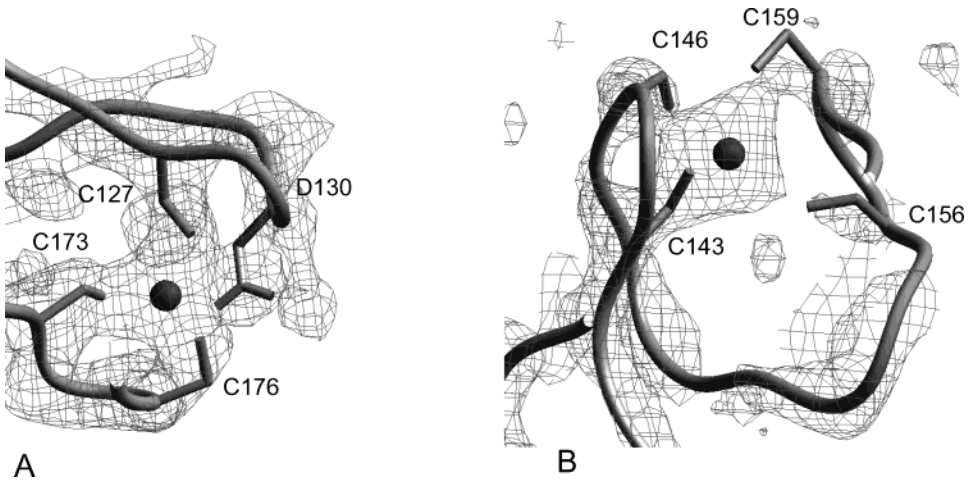

FIGURE 5: Views of the final $2F_o - F_c$ electronic density map contoured at 1σ in the region of the two zinc ions. (A) View encompassing the proximal knuckle. (B) View encompassing the distal knuckle.

(B) Active Site. In the case of the *E. coli* enzyme, amino acid binding was deeply studied thanks to the 3D structures of complexes of the enzyme with methionine or with analogues (*34*, *35*). Recognition of the amino acid is triggered by an induced fit mechanism involving aromatic residues such as Y15, F300, F304, W229, and W253. In addition, Y15, D52, D296, and W253 play an important role in the amino acid activation (*35*). The catalytic centers of *E. coli* and *P. abyssi* free enzymes were superimposed. All important residues are conserved and have similar positions (see Y12, D50, D296, W254, F302, and F306 in Figure 4). The only exception in MetRSpa is R227, which replaces W229 of MetRSec. Moreover, the side chain of W254, the counterpart of W253 of MetRSec, is rotated by 90°. As a result, the position of the side chain of F306 of MetRSpa is different from that of the corresponding F304 in MetRSec. By analogy with amino acid recognition by MetRSec, it can be proposed that, in MetRSpa, binding of methionine triggers the following events: W254 flips first and is stabilized by Y12; F302 would then stack on W254 and R227 on F306.

(C) Binding Sites for the Two Zinc Ions. Another important region involved in catalysis is the connective peptide containing the zinc binding region (*19*, *36*). The connective peptide of MetRSpa has the same topology as that of MetRSec. Two subdomains can be distinguished. The first one is tightly associated with the Rossmann fold (residues 98−126 and residues 178−252). It is well-defined in the electronic density. The second domain corresponds to the zinc binding region (residues 127−180). As in the case of MetRSec, this region comprises two knuckles. The proximal knuckle in MetRSpa corresponds to the zinc-free knuckle in MetRSec. However, in MetRSpa, this knuckle clearly contains one metal ion liganded by three cysteines (C127, C173, C176) and one aspartate (D130) (Figure 5). The distal knuckle of MetRSpa appears poorly defined in the electronic density (mean *B*-value of 124 Å^2 for the distal site and 88 Å^2 for the proximal site). However, with the help of the MetRSec model, the main chain could be followed and density for a second zinc ion was evidenced (Figure 5). This metal is liganded by four cysteines (C143, C146, C156, C159). Together with atomic absorption data, these crystallographic data confirm the prediction made from structural alignment of MetRS sequences and show for the first time the presence of a zinc ion within the proximal knuckle of one member of the MetRS families displaying two structural knuckles (Figure 2).

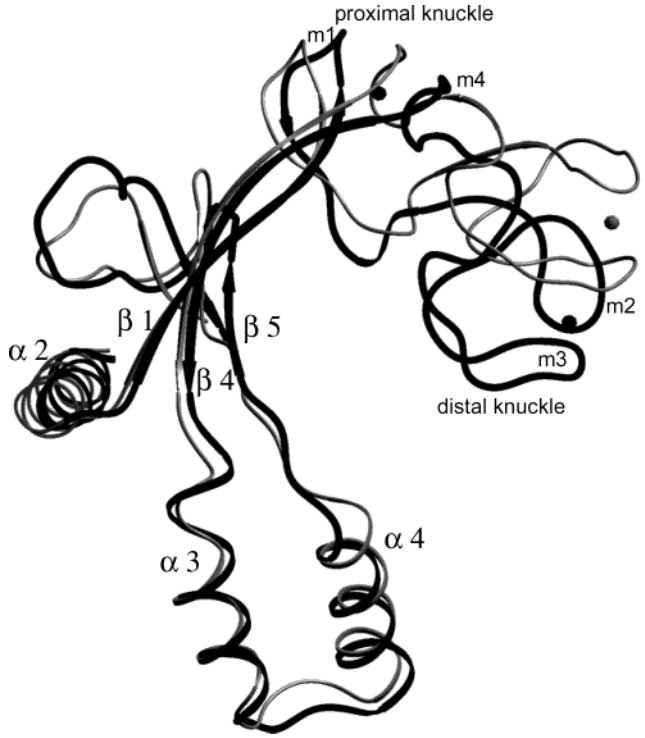

FIGURE 6: Comparison of the CP domains of *E. coli* (gray) and of *P. abyssi* (dark) MetRS. The Rossmann folds of the two enzymes were superimposed using O (*26*). The CP domains deduced from this superimposition are represented. The figure shows the different orientations of the zinc binding knuckles with respect to the active sites in the two enzymes. The two knuckles and the position of the four motifs (m1 to m4) are indicated according to Figure 2.

In MetRSec, two residues (D138, R139) belonging to the zinc binding region are involved in stabilization of the transition state, leading to methionyl adenylate formation (*19*). However, in the solved 3D structures, these residues are far away from the catalytic center. Therefore, movement of the CP domain during the course of the reaction was envisaged (*7*, *35*). In Figure 6, the Rossmann folds of MetRSec and MetRSpa are superimposed. The whole zinc binding domain of MetRSpa is rotated toward the active site if compared to that of MetRSec or MetRStt. As a consequence, in the zinc binding region of MetRSpa, D136 and

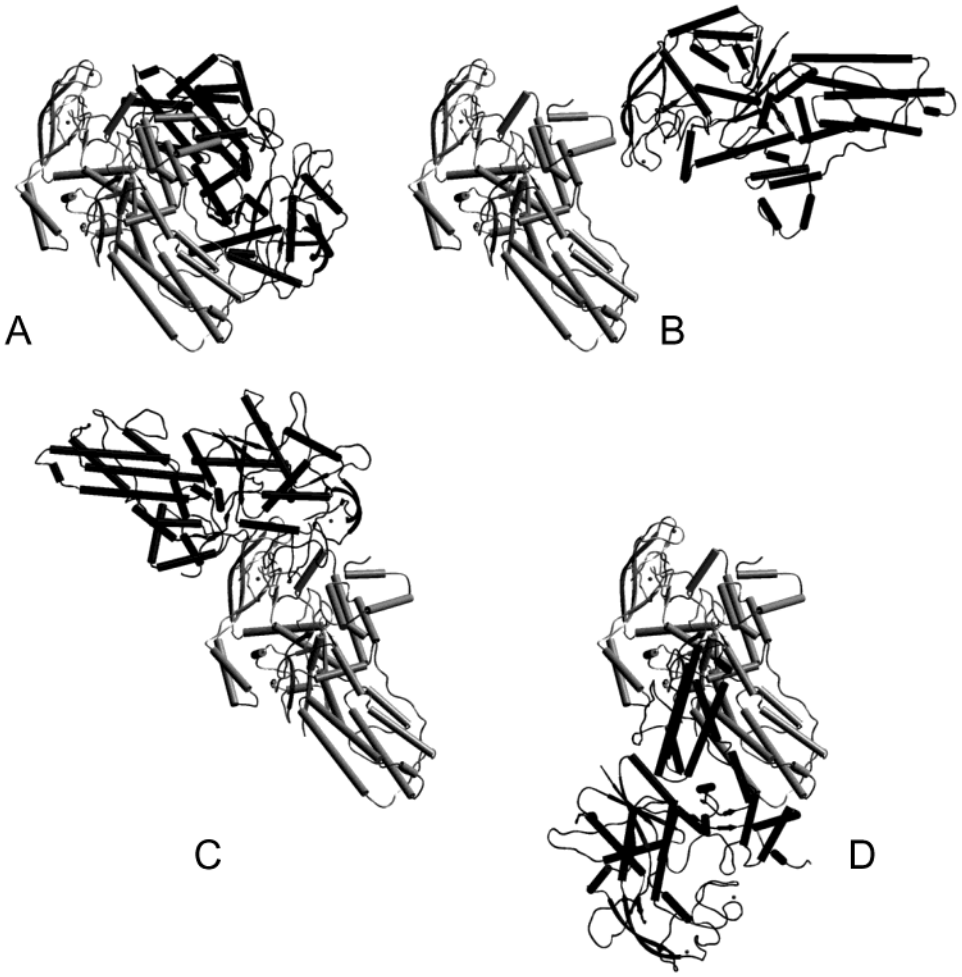

FIGURE 7: Four possible solutions for the biological dimer of MetRSpa (panels A−D). The reference molecule is shown in gray. The four closest neighbors related to the reference molecule by a 2-fold axis were generated. The corresponding operators are ($-x + {}^1/_2$, y, $-z + {}^3/_4$) for panel A, ($x + {}^1/_2$, $-y$, $-z + {}^3/_4$) for panel B, ($-x$, $-y$, z) for panel C, and ($-y + {}^1/_2$, $-x + {}^1/_2$, $-z + {}^1/_2$) for panel D. The deduced molecule is represented in black on each panel.

R137, which correspond to D138 and R139 in MetRSec, come closer to the methionyl adenylate binding pocket (Figure 3). The difference between the two synthetases possibly reflects the capacity of the CP domain to move quite easily. Notably, the mean *B*-value (106 Å^2) for the two zinc binding knuckles (residues 127−180) is significantly higher than that associated to the rest of the structure (52 Å^2). This also argues in favor of mobility of this enzyme region.

(D) tRNA Anticodon Binding Domain. There are very few differences between the α-helix bundle domain of MetRSpa and that of MetRSec or MetRStt. One additional turn occurs at the N-terminus of the α10 helix of MetRSpa. Region 530, which is helical in MetRSec and MetRStt, is elongated in MetRSpa because of the deletion of two residues. W461, N391, R395, and N452, the positive determinants of MetRSec (*37*−*40*) involved in recognition of the CAU anticodon, are spatially conserved in MetRSpa (W476, N402, R406, N467; Figure 3). Negative determinants (*40*) against non-methionine tRNAs are only partially conserved. D471 in MetRSpa corresponds to D456 of MetRSec. D449 of MetRSec is replaced by S464 in MetRSpa.

Dimerization Domain. During the building of the model and the refinement procedure, no electronic density could be evidenced beyond residue 606. Moreover, although the 3D structure of MetRS-C12K is known, no density allowed its placement. However, dissolution of the crystal warranted the presence of an intact dimeric enzyme. One possible explanation is that the dimerization domain does not obey all of the symmetry operators of the $I_1 22$ space group. This possibility was examined after P1 reindexation and comparison of R_{iso} between equivalent reflections for each symmetry operator. No obvious discrepancy could be evidenced for any operator. Moreover, an electron density map was calculated with a 95% complete data set at 4.8 Å resolution indexed in P1. Again, no density could be assigned to the C12K domain. Finally, another explanation for the lack of electronic density associated to the C12K domain must be a movement of this domain relative to the core enzyme in the crystal.

In the absence of data for the C12K domain, we analyzed the crystal packing in order to possibly identify the biological dimer. Four neighbors related to the reference molecule by a 2-fold axis are candidates (Figure 7). Molecule 4 is unlikely because it contacts the reference monomer through the anticodon binding domain and masks the tRNA binding site (buried surface of 2061 Å^2). Molecules 2 and 3 have only few contacts with the reference monomer (buried surfaces of 565 and 1495 Å^2, respectively). Therefore, by combining either of these two molecules with the reference monomer, the interface ensuring dimerization should mainly result from

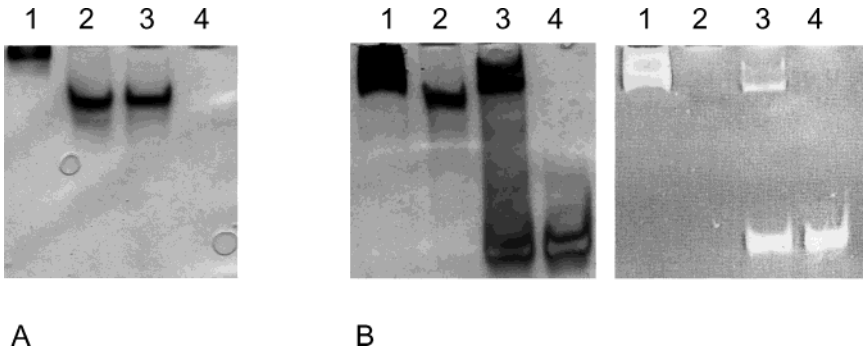

FIGURE 8: Formation of the ternary complex detected by gel electrophoresis. (A) Lanes: 1, MetRSpa; 2, M606; 3, M606 plus MetRS-C12K (*13*); 4, MetRS-C12K. Each protein was at a concentration of 10 μM during the incubation prior to the electrophoresis. The gel was stained with silver nitrate (Bio-Rad, Hercules, CA). In lanes 3 and 4, the band corresponding to free MetRS-C12K is absent, in agreement with the basic isolectric point of this protein (9.3). (B) Left and right: Same as part A but in the presence of 5 μM *P. abyssi* $tRNA_i^{Met}$ overexpressed in *E. coli*. The gel was stained first with ethidium bromide (right part) and second with silver nitrate (left part). Under the conditions of the experiment, the band corresponding to free tRNA has run out of the gel. Indeed, long running times were necessary to ensure separation of high molecular weight species. On lane 2 (right part), no shifted M606 band is visible, thus indicating low affinity of M606 for tRNA. With the addition of MetRS-C12K, a three-partner complex is formed (lane 3, right part). It was verified that, under the same conditions, *E. coli* $tRNA_1^{Val}$ did not shift the position of MetRSpa on the gel (not shown).

the C12K domain (4792 Å^2). The last candidate is monomer 1, which makes contacts with the reference molecule by a surface of 3308 Å^2. For this pair of protomers, the 2-fold axis in the crystal crosses a cavity large enough to accommodate the C12K domain. Notably, hydrogen bonds occur between monomer 1 and the reference monomer. They involve in particular the region from α16 to α18, far from the C12K region. To know whether such contacts might contribute to the dimerization of MetRSpa, a stop codon was introduced in place of the Leu codon at position 607. This mutation ensured deletion of the C12K domain. The resulting M606 enzyme was purified. The enzyme is active in the ATP−[^{32}P]PP_i isotopic exchange as well as in the $tRNA_f^{Met}$ aminoacylation reaction. It behaves as a monomer on a Superdex 200 molecular sieving column. We therefore disfavored the idea that intermolecular contacts such as those observed in the crystal lattice between the reference monomer and molecule 1 can occur in solution. In truth, as noted above, helices α16−α18 that are involved in these contacts are not conserved in the MetRS family. They are specific for the MetRS belonging to the *Pyrococcus* species.

Addition of the Isolated MetRS-C12K Reinforces the Binding of $tRNA^{Met}$ to Monomeric MetRS. A possible role of the MetRS-C12K domain in tRNA binding was investigated by using gel shift experiments. As shown in Figure 8 (panel B, lane 4), MetRS-C12K alone induces a mobility shift of $tRNA_i^{Met}$. A shift is also obtained with intact dimeric MetRSpa (panel B, lane 1). In contrast, with truncated monomeric M606, no shift is observed, indicating a lowered affinity of $tRNA_i^{Met}$ for the truncated enzyme (panel B, lane 2). Such a behavior is in keeping with the stronger binding of $tRNA^{Met}$ to native MetRSec if compared to its truncated derivative (*17*). When the experiment with truncated monomeric M606 enzyme was repeated in the presence of the MetRS-C12K domain, a significant amount of a complex involving M606, tRNA, and MetRS-C12K was observed (panel B, lane 3). However, in the absence of tRNA, M606 plus C12K do not form a complex (panel A, compare lanes 2 and 3). The latter behavior is in agreement with the view that the MetRS-C12K domain and the core enzyme moieties are not in contact in the crystalline enzyme. Since the MetRS-C12K domain acts in trans on the affinity of M606 for tRNA, it may be imagined that this domain adapts the conformation or the electric charge of $tRNA^{Met}$ toward better recognition by the core enzyme. Finally, in dimeric MetRSpa, the C-domain is physically linked to the rest of the enzyme. Consequently, the dimerization domain should contribute to tRNA binding affinity both directly, by establishing contacts with the polynucleotide, and indirectly, by adjusting the conformation of the polynucleotide. The formation of a complex where tRNA is sandwiched between M606 and MetRS-C12K is reminiscent of the ternary complex obtained from the combination of *E. coli* monomeric IleRS, $tRNA^{Ile}$, and Trbp111 (*41*). The latter protein, Trbp111, is a close homologue of the C-domain of dimeric MetRS (*9*).

Conclusion. This study indicates high flexibility of the assembly of the two domains composing each protomer of native MetRS: the core enzyme and the C12K domain. In the dimer, the two modules corresponding to the two core enzymes are likely to not interact (ref *42* and this study). Moreover, the two core enzymes are likely to not be in contact with the C12K domain, except through a flexible linker region. Possibly, in solution, they move freely around a pair of two strongly associated C12K domains. Relative mobility between the two protomers of *E. coli* native MetRS has indeed been already indicated by small angle neutron scattering studies (*43*). For instance, the radius of gyration of the dimer significantly varied as a function of ionic conditions or of the addition of ligands. Upon the binding of one $tRNA^{Met}$ per dimer, a contraction of the enzyme was observed, whereas upon anticooperative binding of a second tRNA, the radius of gyration of the enzyme moiety increased markedly (*43*). Possibly, new crystallization trials in the presence of various ratios of enzyme to tRNA will help to overcome the problem of the high mobility of the C-terminal domain in MetRS crystals.

ACKNOWLEDGMENT

We thank the staff of the LURE DW32 and ESRF-JSBG beamlines for assistance during data collection.

REFERENCES

1. Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) *Nature 347*, 203−206.

2. Cusack, S., Hartlein, M., and Leberman, R. (1991) *Nucleic Acids Res. 19*, 3489−3498.
3. Blanquet, S., Crepin, T., Mechulam, Y., and Schmitt, E. (2003) in *Aminoacyl-tRNA synthetase* (Ibba, M., Francklin, C., and Cusack, S., Eds.) Landes Bioscience, Georgetown, TX.
4. Cassio, D., and Waller, J. (1971) *Eur. J. Biochem. 20*, 283−300.
5. Kalogerakos, T., Dessen, P., Fayat, G., and Blanquet, S. (1980) *Biochemistry 19*, 3712−3723.
6. Kohda, D., Yokoyama, S., and Miyazawa, T. (1987) *J. Biol. Chem. 262*, 558−563.
7. Mechulam, Y., Schmitt, E., Maveyraud, L., Zelwer, C., Nureki, O., Yokoyama, S., Konno, M., and Blanquet, S. (1999) *J. Mol. Biol. 294*, 1287−1297.
8. Sugiura, I., Nureki, O., Ugaji-Yoshikawa, Y., Kuwabara, S., Shimada, A., Tateno, M., Lorber, B., Giege, R., Moras, D., Yokoyama, S., and Konno, M. (2000) *Struct. Folding Des. 8*, 197−208.
9. Swairjo, M. A., Morales, A. J., Wang, C. C., Ortiz, A. R., and Schimmel, P. (2000) *EMBO J. 19*, 6287−6298.
10. Muller, J. P., Bron, S., Venema, G., and van Dijl, J. M. (2000) *Microbiology 146*, 77−88.
11. Simos, G., Segref, A., Fasiolo, F., Hellmuth, K., Shevchenko, A., Mann, M., and Hurt, E. C. (1996) *EMBO J. 15*, 5437−5448.
12. Kao, J., Ryan, J., Brett, G., Chen, J., Shen, H., Fan, Y. G., Godman, G., Familletti, P. C., Wang, F., Pan, Y. C., et al. (1992) *J. Biol. Chem. 267*, 20239−20247.
13. Crepin, T., Schmitt, E., Blanquet, S., and Mechulam, Y. (2002) *Biochemistry 41*, 13003−13011.
14. Morales, A. J., Swairjo, M. A., and Schimmel, P. (1999) *EMBO J. 18*, 3475−3483.
15. Simos, G., Sauer, A., Fasiolo, F., and Hurt, E. C. (1998) *Mol. Cell 1*, 235−242.
16. Renault, L., Kerjan, P., Pasqualato, S., Menetrey, J., Robinson, J. C., Kawaguchi, S., Vassylyev, D. G., Yokoyama, S., Mirande, M., and Cherfils, J. (2001) *EMBO J. 20*, 570−578.
17. Blanquet, S., Iwatsubo, M., and Waller, J. P. (1973) *Eur. J. Biochem. 36*, 213−226.
18. Fourmy, D., Meinnel, T., Mechulam, Y., and Blanquet, S. (1993) *J. Mol. Biol. 231*, 1068−1077.
19. Fourmy, D., Mechulam, Y., and Blanquet, S. (1995) *Biochemistry 34*, 15681−15688.
20. Hountondji, C., Schmitter, J. M., Beauvallet, C., and Blanquet, S. (1990) *Biochemistry 29*, 8190−8198.
21. Kim, S., Jo, Y. J., Lee, S. H., Motegi, H., Shiba, K., Sassanfar, M., and Martinis, S. A. (1998) *FEBS Lett. 427*, 259−262.
22. Senger, B., Despons, L., Walter, P., Jakubowski, H., and Fasiolo, F. (2001) *J. Mol. Biol. 311*, 205−216.
23. Leslie, A. (1990) *Crystallographic computing V*, Oxford University Press, Oxford, U.K.
24. Collaborative Computational Project No. 4. (1994) *Acta Crystallogr. D50*, 760−763.
25. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
26. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
27. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
28. Cassio, D., Lemoine, F., Waller, J. P., Sandrin, E., and Boissonnas, R. A. (1967) *Biochemistry 6*, 827−836.
29. Blanquet, S., Fayat, G., Waller, J. P., and Iwatsubo, M. (1972) *Eur. J. Biochem. 24*, 461−469.
30. Mayaux, J. F., and Blanquet, S. (1981) *Biochemistry 20*, 4647−4654.
31. Lawrence, F., Blanquet, S., Poiret, M., Robert-Gero, M., and Waller, J. P. (1973) *Eur. J. Biochem. 36*, 234−243.
32. Guillon, J. M., Meinnel, T., Mechulam, Y., Lazennec, C., Blanquet, S., and Fayat, S. (1992) *J. Mol. Biol. 224*, 359−367.
33. Dardel, F., and Bensoussan, P. (1988) *Comput. Appl. Biosci. 4*, 483−486.
34. Serre, L., Verdon, G., Choinowski, T., Hervouet, N., Risler, J. L., and Zelwer, C. (2001) *J. Mol. Biol. 306*, 863−876.
35. Crepin, T., Schmitt, E., Mechulam, Y., Sampson, P. B., Vaughan, M. D., Honek, J. F., and Blanquet, S. (2003) *J. Mol. Biol. 332*, 59−72.
36. Starzyk, R. M., Webster, T. A., and Schimmel, P. (1987) *Science 237*, 1614−1618.
37. Ghosh, G., Pelka, H., and Schulman, L. H. (1990) *Biochemistry 29*, 2220−2225.
38. Ghosh, G., Kim, H. Y., Demaret, J. P., Brunie, S., and Schulman, L. H. (1991) *Biochemistry 30*, 11767−11774.
39. Meinnel, T., Mechulam, Y., Le Corre, D., Panvert, M., Blanquet, S., and Fayat, G. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 291−295.
40. Schmitt, E., Meinnel, T., Panvert, M., Mechulam, Y., and Blanquet, S. (1993) *J. Mol. Biol. 233*, 615−628.
41. Nomanbhoy, T., Morales, A. J., Abraham, A. T., Vortler, C. S., Giege, R., and Schimmel, P. (2001) *Nat. Struct. Biol. 8*, 344−348.
42. Dessen, P., Fayat, G., Zaccai, G., and Blanquet, S. (1982) *J. Mol. Biol. 154*, 603−613.
43. Dessen, P., Blanquet, S., Zaccai, G., and Jacrot, B. (1978) *J. Mol. Biol. 126*, 293−313.
44. Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.
45. Barton, G. J. (1993) *Protein Eng. 6*, 37−40.

BI0356247