Expression and purification of untagged and histidine-tagged folate-dependent tRNA:m5U54 methyltransferase from Bacillus subtilis

Djemel Hamdane, Stéphane Skouloubris, Hannu Myllykallio, Béatrice Golinelli-Pimpaneau

Expression and purification of untagged and histidine-tagged folate dependent tRNA:m5U54 methyltransferase from *Bacillus subtilis*.

Djemel Hamdane[a], Stéphane Skouloubris[b,c], Hannu Myllykallio[d], Béatrice Golinelli-Pimpaneau[a]*

[a]Laboratoire d'Enzymologie et Biochimie Structurales, CNRS, 1 avenue de la Terrasse, 91190 Gif-sur-Yvette, France

[b] Univ Paris Sud, Institut de Génétique et de Microbiologie, UMR 8621, Orsay F-91405, France

[c]CNRS, Orsay F-91405, France

[d] Laboratoire d'Optique et Biosciences, Ecole Polytechnique, CNRS-INSERM, 91128 Palaiseau, France

*: Corresponding author: beatrice.golinelli@lebs.cnrs-gif.fr, tel: 33 1 69824235, fax: 33 1 69823129

.

Abstract

Folate-dependent tRNA m^5U methyltransferase TrmFO is a flavoprotein that catalyzes the C^5-methylation of uridine at position 54 in the TΨC loop of tRNA in several bacteria. Here we report the cloning and optimization of expression in *E. coli* BL21(DE3) of untagged, N-terminus, C-terminus $(His)_6$-tagged TrmFO from *Bacillus subtilis*. Tagged and untagged TrmFO were purified to homogeneity by metal affinity or ion exchange and heparin affinity, respectively, followed by size-exclusion chromatography. The tag did not significantly alter the expression level, flavin content, activity and secondary structure of the protein.

keywords: expression, purification, recombinant TrmFO protein, flavoenzyme, RNA methyltransferase, 5-methyluridine, histidine tag

1. Introduction

All the metabolically stable RNAs are subject to diverse post-transcriptional chemical modifications that are crucial for maintenance of their three-dimensional structure and translation fidelity [1]. One common modified nucleoside is C^5-methylated uridine, invariably

found at position 54 in the TΨC loop of tRNA of almost all organisms. In eukaryotes, archaea and most Gram-negative bacteria, this reaction is catalyzed by enzymes from the TrmA family that use S-adenosyl-L-methionine as the methyl donor [2-5]. An alternative pathway exists in most Gram-positive bacteria and some Gram-negative bacteria [6-8]. In this case, the methylation requires the 5,10-methylenetetrahydrofolate as the methylene donor and reduced flavin adenine dinucleotide $FADH_2$ as a reducing agent (Fig. 1A), *in vitro* [9]. The gene responsible for this class of methyltransferases (TrmFO) has recently been identified [10].

TrmFO proteins from hyperthermophilic bacteria, *Thermus thermophilus (Gram negative)* and *Thermotoga maritima (Gram positive),* have been expressed in *E. coli* in their untagged forms and crystallized [11, 12]. The structure of *T. thermophilus* TrmFO was solved in its free, tetrahydrofolate- (Fig. 1B), and glutathione-bound forms [12]. The structure of *T. thermophilus* TrmFO shares similarity with GidA, another FAD-dependent tRNA modification enzyme, which participates in the incorporation of both 5-aminomethyluridine and 5-carboxymethylaminomethyl to the uridine base at position 34 of some tRNAs [13]. TrmFO consists of two domains, an FAD-binding domain (residues 1-153 and 314-436) and an insertion domain (residues 314-436) (Fig. 1B) [12, 14]. TrmFO holds at its N-terminus end a consensus dinucleotide binding motif, GXGXAGXEA, which is a portion of the Rossmann-fold domain found in a large number of FAD proteins, such as the glutathione reductase family members. [15, 16]. Residues from this motif are in direct interaction with the pyrophosphate moiety of FAD. The characterization of several mutants of TrmFO together with the modeling of the complex with tRNA, gave hints about the catalytic mechanism of the methylation of the target uridine [12]. Yet, many questions still need to be answered, in particular the exact role of the flavin during catalysis. Besides, there is no available structure of TrmFO in complex with an RNA substrate, which would shed light on the tRNA recognition mode.

We have lately been interested in TrmFO from the Gram-positive mesophilic bacterium *Bacillus subtilis*, since the first cloned TrmFO gene came from this organism. Sequence alignment of TrmFO proteins reveals that these proteins are highly conserved. For example, *B. subtilis* TrmFO presents 50.9 % and 49.4% sequence identity with TrmFO from *T. thermophilus* and *T. maritima*, respectively. In an earlier paper, the expression and purification of an N-terminus $(His)_6$-tagged TrmFO from *B. subtilis* have been described [10] but the effect of the tag for protein expresion and/or activity was not studied. The N-terminus is located five residues upstream of the β1 sheet of the flavin domain, hence the tag is in close vicinity to the FAD binding motif (Fig. 1B). It may hinder the proper incorporation of the

FAD coenzyme and/or interfere with the binding of substrates at the nearby active site. Moreover, it is known that histidine tags can alter the solubility of purified proteins [17].

In order to produce large amounts of *B. subtilis* TrmFO required for detailed biophysical characterization as well as crystallization studies, we first decided to optimize the expression and purification protocol previously developed for the N-terminus $(His)_6$-tagged protein (Nt-TrmFO) [10, 18]. We also examined the importance of the $(His)_6$-tag by placing it at the C-terminus extremity of the protein and by engineering an untagged TrmFO protein. Since the C-terminus lies more than 25 Å away from the active site and is solvent accessible (Fig. 1B), positioning the tag at this extremity should not alter the function of the protein. We developed a protocol for purification to homogeneity of the recombinant *B. subtilis* Nt-TrmFO, Ct-TrmFO and untagged TrmFO. These three recombinant forms enable to determine the effect of the presence and position of the histidine tag on the production yield and function of the protein.

Material and methods

Gene cloning

The TrmFO gene was amplified by PCR from *B. subtilis* strain 168 DNA using *Pfu* DNA polymerase (Promega) using the forward and reverse primers indicated in Table S1. In the case of Nt-TrmFO, the pair of primers used was SK20/SK21. The amplified gene was digested with *Bam*HI and *Sma*I and inserted into the expression vector pQE-80L (Qiagen). The resulting plasmid was named pHM408 (Fig. S1A). For Ct- and untagged TrmFO, the pairs of primers used were SK76/SK77 and SK76/SK78, respectively. The amplified genes were digested with *Nco*I and *Xho*I, and inserted into the expression vector pET-28a (Novagen) to generate pHM440 (Ct-TrmFO; Fig. S1B) and pHM441 (untagged TrmFO; Fig. S1C).

Protein expression

The recombinant proteins were overexpressed using pQE-80L and pET-28a derivatives in *E. coli* BL21 or BL21 (DE3) pLysS strains (Invitrogen), respectively. 10 ml of overnight culture was used to inoculate 0.5 L of LB medium containing 100 µg/ml of ampicillin (for Nt-TrmFO) or 50 µg/ml of kanamycin (for Ct-TrmFO and untagged TrmFO) and 100 nM of riboflavin in a 2.8 L Fernback flask. The cells were grown in the dark at 37°C until the OD_{600} was 1 or 2, and then induced with 0.05, 0.5 or 1mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). After overnight incubation at 28°C, the cells were harvested by centrifugation at 3,000 g for 15 min at 4 °C. The pellet was resuspended in 20 ml of 50 mM sodium phosphate

pH 8, 300 mM NaCl, 20 mM imidazole (Buffer A) containing 20 µl of Protein Inhibitor Cocktail (PIC, Sigma) and 5 mM β-mercaptoethanol. The cells were lysed by sonication and the lysate was centrifuged at 3000 g for 40 min.

Purification of Nt-TrmFO and Ct-TrmFO by immobilized metal affinity chromatography

The supernatant was subsequently loaded at 4°C onto a 5 mL Ni^{2+}-NTA agarose affinity column (Qiagen) pre-equilibrated with buffer A. After washing with 5 column volumes of buffer A, TrmFO was eluted with buffer A containing 250 mM of imidazole. The protein was then passed through a PD-10 column (Sephadex-G25 medium; GE Healthcare Inc.) pre-equilibrated with 50 mM Tris-HCl pH 7.8, 150 mM NaCl, 10 % glycerol (v/v). Finally, TrmFO was concentrated to ≈ 500-800 µM using a Vivaspin 20 ml centrifugal concentrator (10,000 MWCO PES membrane; Sartorius) and then filtered through a 0.22 µm filter before storing at -80°C. For crystallization studies, the protein was further purified on a Superdex 200 HR 10/30 column, as described below for the untagged protein.

The flavin content (cofactor / enzyme ratio) of the proteins was obtained by assessing the concentration of free flavin released after heat treatment (3 minutes at 90°C). An extinction coefficient of 11.3 $mM^{-1}.cm^{-1}$ at 450 nm was used to estimate the concentration of free FAD. The concentration of total protein was determined by the Bradford method.

Purification of untagged TrmFO

All the following steps were done using an Äkta purifier system (Amersham Biosciences, GE Healthcare) at 4°C.

DEAE-Sepharose anion exchange chromatography

After lysis, the cells supernatant was loaded at 2.5 mL/min onto a HiTrap DEAE Sepharose column (5ml ; GE Healthcare Inc.) preequilibrated in 25 mM Tris-HCl pH 8, 10 mM NaCl (buffer B). The enzyme was eluted with 75 ml of a linear gradient of 0.01-1M NaCl in buffer B. The fractions containing the protein were pooled and concentrated to 2 ml before loading onto a PD-10 column equilibrated with 10 mM potassium phosphate pH 8 (buffer C).

Heparin-Sepharose chromatography

The sample was then loaded at 2.5 mL/min onto a HiTrap Heparin Sepharose column (5ml, GE healthcare Inc.) equilibrated with buffer. The protein was eluted with 75 ml of a linear gradient of 0-2M NaCl in buffer C.

Size-exclusion chromatography

As final step of purification, the tagged or untagged TrmFO proteins were loaded at 0.5 ml/min onto a Superdex 200 HR 10/30 column (GE healthcare Inc.) equilibrated in 50 mM sodium phosphate pH 8, 150 mM NaCl. Native molecular mass determination was performed by calibrating the column using a molecular weight standard mixture (Biorad) containing thyroglobulin, γ-globulin, ovalbumin, myoglobin and vitamin B12.

Circular dichroism Spectroscopy

The secondary protein structure was analyzed by Circular dichroism (CD) spectroscopy. Far UV CD spectra were recorded on a JASCO J810 spectropolarimeter (Tokyo, Japan) over the 190 to 260 nm wavelength range in 10 mM potassium phosphate pH 8 at 20°C. Spectra were recorded in duplicate in a 0.4 cm pathlength quartz cuvette using a scan speed of 50 $nm \cdot min^{-1}$, a band width of 2 nm and a response time of 2 s. The percentage content of the α-helical and β-sheet structures in the proteins was estimated according to Scholtz et al [19].

Methylation Activity

The tRNA m^5U54 methyltransferase activity of TrmFO was determined, as described previously [20], using an *E. coli* [α-^{32}P]UTP-labeled $tRNA^{Ala1}$ transcript. Initially, both NADH and NADPH were used in the enzymatic assays, since the electron donor of FAD was not known (Fig. 1A) [18]. After the observation that only NADH is oxidized by TrmFO (Fig. S2), NADPH was not used any longer and only NADH was added in the activity assay. 50–100 fmol of [α-^{32}P]-labeled $tRNA^{Ala1}$ was incubated at 37°C in a 50 µl reaction mixture containing 50 mM N-[2-hydroxyethyl] piperazine-N-[2-ethanesulfonic acid]-Na buffer (HEPES-Na, Sigma), pH 7.5, 100 mM ammonium sulfate, 0.1 mM EDTA, 0.5 mM NADH (reduced nicotinamide adenine dinucleotide, Sigma), 0.5 mM (6R,S)-N5,N10-CH2H4PteGlu-Na2 (methylenetetrahydrofolate), 25 mM mercaptoethanol (Promega) and 20% glycerol. The reaction was started by the addition of 0.2 µM of purified TrmFO. After 30 minutes incubation, the tRNA product was phenol extracted and digested with nuclease P1 (Roche), which generates 5'-nucleoside monophosphates. The hydrolysate was analyzed by 2D thin-layer chromatography on cellulose plates (Machery Nagel)(Fig. 5A, B). The relative amount of m^5U formed per tRNA molecule was determined by measuring the radioactivity in the TLC spots using a PhosphorImager screen and quantification with the ImageQuant software.

NAD(P)H oxidation by TrmFO under aerobic condition

The ability of TrmFO to oxidize NADH and/or NADPH was determined under steady state conditions in 100 mM HEPES pH 7.5, 100 mM ammonium acetate, 20 % glycerol. Assays were performed using ~ 1 μM TrmFO and variable concentrations of NAD(P)H (0 - 0.5 mM). The amount of NAD(P)H oxidized was monitored by the absorbance decrease at 340 nm (ε_{340} = 6.21 $mM^{-1} \cdot cm^{-1}$) using a Cary 50 (Varian, Inc) spectrophotometer. The rate versus NAD(P)H concentration was analyzed according to Michaelis Menten.

Results and discussion

Protein expression

The expression of the Nt-TrmFO was performed using pQE-80L as the expression vector, whereas pET-28a was used for the expression of the Ct-TrtmFO and untagged TrmFO (Fig. S1). *E. coli* BL21 and *E. coli* BL21 (DE3) pLysS were used as the expression host strains with the pQE-80L and pET-28a vectors, respectively. The low copy *cis*-repressed pQE-80L vector allows tightly regulated N-terminus $(His)_6$-tagged protein expression in any *E. coli* host strain. The vector contains, 5' to the cloning site, the strong inducible T5 promoter, which is recognized by *E. coli* RNA polymerase, under control of two *lac* operator sequences and, 3' to the cloning site, the *lacI* gene coding for the *lac* repressor. Overexpression of the *lac* repressor strongly represses protein expression prior to induction by IPTG, which displaces the repressor from the *lac* operator. Protein expression in the *E. coli* BL21 (DE3) pLysS/pET host-vector couple allows tight regulation and high level of protein expression. The host strain genome contains the gene for bacteriophage T7 RNA polymerase, under the control of the *lac* promoter and the *lac* operator, and the gene coding for the *lac* repressor protein. The pLysS plasmid in the host strain encodes T7 lysozyme, which lowers the background expression level of target genes under the control of the T7 promoter but does not interfere with the level of expression after induction by IPTG. The pET28a vector contains, 5' to the cloning site, the T7 promoter specific for T7 RNA polymerase and the *lac* operator and, 3' to the cloning site, a C-terminal histidine tag sequence. Transcriptions of both the T7 RNA polymerase and the cloned genes are activated upon addition of IPTG. To construct the expression vector for untagged TrmFO, a stop codon was included in the reverse oligonucleotide used in the gene amplification step of the cloning so that the $(His)_6$-tag sequence of the pET28a vector was not fused to the C-terminus of the protein (Table S1; Fig. S1C).

The optimization of protein expression was achieved by varying the concentration of IPTG (0.05, 0.5 and 1 mM) along with the incubation time before induction (see Fig. S3 in the case of Ct-TrmFO). The highest amount of soluble TrmFO proteins is obtained when

protein expression is induced at OD_{600}=1 with 50 μM IPTG. After the Nickel affinity column, 34 mg of soluble Nt-TrmFO was obtained per 500 ml of culture (Table 1A), which is tenfold higher than the previously reported yield [18]. The amount of soluble Ct-TrmFO obtained is 65% that of Nt-TrmFO (Table 1A). This difference could be due to the expression vector rather than to the position of the tag since pET-28a was used for Ct-TrmFO instead of pQE-80L for Nt-TrmFO. After purification, the amount of TrmFO obtained is about the same in the case of Ct- or untagged TrmFO, which were expressed using the same vector. Therefore, the presence of the histidine tag at the C-terminus has no significant effect on the production yield of the protein, in contrast to what is observed for other proteins [17].

Isolation of a spontaneous mutation in TrmFO

During the course of this study, we noticed that some of our clones carrying *B. subtilis* TrmFO protein bore unwanted mutations. For instance, a R312G mutant isolated after cloning was found to be defective in tRNA methylation, with an activity corresponding only to 40% that of the wild type enzyme (Fig 5). This result may partially explain the low *in vitro* activity of *B. subtilis* Nt-TrmFO reported by [10]. Indeed, in the crystal structure of *T. thermophilus* TrmFO, Arg319 (equivalent to Arg312 in *B. subtilis*) lies near the folate-binding site of the protein (Fig. 1C). Its main chain carbonyl oxygen is hydrogen bonded to the amino group of the folate pteridin ring and is located 4.7 Å away from the C8 methyl group of the FAD cofactor. Therefore, the R312G mutation in *B. subtilis* Nt-TrmFO may actually diminish the affinity of the protein for the folate cofactor and thereby decrease the methylation activity. This suggests that overexpression of a functionally active *B. subtilis* TrmFO in *E. coli* cells may inhibit cell growth, thus resulting in isolation of inactivating mutations. As earlier studies pointed out a likely high affinity of TrmFO enzymes for folate, our data suggest strain adaptability to the sequestration of the folate pool present in the cell by the overexpressed enzyme. This observation explains why we carefully optimized the induction protocol for the purification of *B. subtilis* TrmFO and prompted us to regularly control the *trmFO* gene by DNA sequencing.

Purification method

Since the TrmFO protein contains a flavin prosthetic group, a visible chromophore, elution of the protein during the different chromatographic steps was recorded by following the absorbance at 280 nm (protein) and at 450 nm (FAD cofactor).

The Qiagen Ni-NTA protein purification system is based on the remarkable selectivity of Ni^{2+}-NTA (nickel-nitrilotriacetic acid) resin for proteins containing a surface-exposed affinity tag of six consecutive histidine residues. After one step of chromatography on a (Ni^{2+}-NTA)

agarose affinity column, the Nt- and Ct-TrmFO recombinant proteins were purified to near homogeneity (Fig. 2A & B, lanes 4). This degree of purity is sufficient for the biochemical and biophysical characterization of the protein. A further purification step on a size-exclusion chromatography column was achieved for the crystallization trials.

The untagged TrmFO protein was purified in three steps consisting of anion exchange, heparin affinity and size-exclusion chromatography. The theoretical isoelectric point of untagged TrmFO is 5.83. Diethylaminoethyl- (DEAE) is a weak anion exchanger. At pH 8, untagged TrmFO is negatively charged and well retained on the column since it eluted at around 0.5 M NaCl (Fig. 3A). Several contaminants co-eluted with TrmFO from the DEAE Sepharose column (Fig. 2C, lane 3), necessitating other purification steps.

Heparin chromatography is an adsorption chromatography, in which a wide range of biomolecules, including DNA or RNA binding enzymes [21-23], can be specifically and reversibly adsorbed by heparins immobilized on an insoluble support. Heparins are negatively charged polydispersed linear polysaccharides so that heparin chromatography is not only an affinity but also an ion exchange chromatography. In the case of TrmFO, the heparin chromatography step was very effective as shown in Fig. 3B & 2C, lane 4. TrmFO eluted at around 0.4 M NaCl and was well separated from contaminant proteins that eluted much earlier.

Gel filtration was used as a final step of purification for each TrmFO protein. As shown in Fig. 4, this technique is well suited for the separation of soluble TrmFO from aggregates. Interestingly, we observed that the untagged TrmFO had a higher tendency to form aggregates than the tagged proteins. After separation of untagged TrmFO from aggregates, the main peak, when reinjected on the gel filtration column, was found to be homogeneous. This result suggests that the soluble TrmFO is not in equilibrium with the aggregate state. The protein yield is 7% (Table 1) and high quantities of purified TrmFO were obtained: 25.5, 12.5 and 13.2 mg per 500 ml of culture for the Nt-, Ct- and untagged TrmFO, respectively. The flavin content of non-covalently bound FAD in the three purified proteins was 40%. Thus, the tag did not significantly prevent the incorporation of the flavin cofactor during protein synthesis in *E. coli*.

To improve the yield of the holoenzyme, we attempted to add free FAD to a freshly purified sample of TrmFO. Basically, the protein was incubated with 5 to 10 equivalent of FAD for 1 hour at 4°C at different pH (from 8 to 11)[24]. Unfortunately, only a limited range of pH could be investigated since rapid and irreversible aggregation of the enzyme occurred at low pHs. The excess of free flavin was separated from the holoprotein on a PD-10 column and the

flavin content was measured. This classical flavin reconstituting procedure did not increase the flavin content. Several other attempts to reconstitute the holoenzyme led to a decrease of the tRNA methylation activity.

Secondary structure analysis by CD spectroscopy

CD is a widely used technique to estimate the secondary structure of proteins and to determine whether a protein is correctly folded or not. The far-ultraviolet CD spectrum (180–250 nm) of TrmFO displayed negative bands at 207 and 222 nm, which is typical of a protein containing a high fraction of alpha helices (Fig. S4). The spectra of the three forms of TrmFO are highly similar (data not shown). Analysis of the CD data indicates that the secondary structure of TrmFO is composed of 41% α-helices and 13% of β-sheets. This content agrees with that calculated for the *T. thermophilus* structure [12], which indicates that all the TrmFO proteins isolated during this work are correctly folded.

Determination of the molecular mass of recombinant TrmFO by gel filtration

Native molecular mass of the different TrmFO proteins was determined by gel filtration on a Superdex 200 HR 10/30 column (Table 1B, Fig. S5). The apparent molecular mass of the His-tagged proteins (47.9 and 47.7 kDa) corresponds well to the monomeric form of the protein (theoretical mass of 49.5 and 49.2, including His_6-tag). Yet, the apparent molecular mass of the untagged protein (33.6 kDa) is 14.8 kDa lower than expected for a monomeric protein (theoretical mass=48.1 kDa). The untagged TrmFO proteins from *T. thermophilus* and *T. maritima* elute from a Superdex 200 column as monomers, with a molecular weight of 29 kDa (theoretical mass=48.9 kDa) and 42 kDa (theoretical mass=49.7 kDa), respectively [11, 12]. The apparent molecular mass of untagged TrmFO from *T. thermophilus* is similar to that of untagged TrmFO from *B. subtilis*. In size-exclusion chromatography, proteins are separated according to their molecular size in solution, or more precisely to their hydrodynamic volume. Thus, with the same molecular mass, a non-globular protein will elute earlier than a globular protein. The shape of TrmFO from *T. thermophilus* protein appears globular (Fig. 1B) and the protein shares high sequence similarity with *B. subtilis* TrmFO. The strange behavior of the untagged TrmFO proteins on Superdex 200 gel filtration column could result from the existence of an extended conformation due to flexibility of the two domains relative to each other (Fig. 1A). In contrast, the histidine tag appears to stabilize the globular three-dimension structure of TrmFO in solution since the tagged proteins behave normally on gel filtration.

Enzymatic activity

<u>Oxidative half reaction</u>

Like many other flavoenzymes, TrmFO can function as an NADH oxidase, consuming molecular oxygen and producing NAD^+ and H_2O_2. As shown in Fig. S2, *B. subtilis* TrmFO oxidizes NADH rather than NADPH, the latter being used by the *T. thermophilus* enzyme [12]. The K_m for NADH is 5 μM and the k_{cat} is 0.22 min^{-1}. Thus, it is likely that NADH gives a hydride to the FAD of *B. subtilis* TrmFO to form $FADH_2$. The preferential binding of NADH to TrmFO was also confirmed by isothermal titration calorimetry (data not shown). Interestingly, NADH is also used as the electron source to reduce FAD in the structurally close tRNA modifying GidA enzymes [13].

Methylation activity

All expressed enzymes display similar tRNA methylation activities (Fig. 5). In our activity test, in which an excess of enzyme over tRNA substrate is used, ~ 1 mol of m^5U per tRNA was formed in 30 min. Previous enzyme activity measurements in the same conditions yielded only 0.4 mol of m^5U per tRNA [10]. This result emphasizes that the location of the tag does not impair the function of TrmFO.

Conclusions

We have successfully cloned, optimized the expression and purified three constructs of TrmFO. More than 10 mg of pure protein was produced per 0.5 L of culture for each TrmFO form. The presence and the location of the histidine tag did not alter the yield of soluble protein produced, the overall folding, and the tRNA methylation activity of the protein. However, in the absence of the tag, an extended conformation of TrmFO appears to be stabilized in solution.

Acknowledgements

We thank Delphine Housset for help in producing the proteins and Dr R. Moser, Merck-Eprova, AG, Switzerland, for providing the folate derivatives. This research was funded by THYMET grant (PCV07_189094) from the Agence Nationale de la recherche (ANR-PCV) and by the CNRS.

References

[1] P.F. Agris, Decoding the genome: a modified view. Nucleic Acids Res 32 (2004) 223-238.

[2] J.T. Kealey, X. Gu, D.V. Santi, Enzymatic mechanism of tRNA (m5U54)methyltransferase. Biochimie 76 (1994) 1133-1142.

[3] A. Alian, T.T. Lee, S.L. Griner, R.M. Stroud, J. Finer-Moore, Structure of a TrmA-RNA complex: A consensus RNA fold contributes to substrate selectivity and catalysis in m5U methyltransferases. Proc Natl Acad Sci U S A 105 (2008) 6876-6881.

[4] J. Urbonavicius, S. Auxilien, H. Walbott, K. Trachana, B. Golinelli-Pimpaneau, C. Brochier-Armanet, H. Grosjean, Acquisition of a bacterial RumA-type tRNA(uracil-54, C5)-

methyltransferase by Archaea through an ancient horizontal gene transfer. Mol Microbiol 67 (2008) 323-335.
[5] H. Walbott, N. Leulliot, H. Grosjean, B. Golinelli-Pimpaneau, The crystal structure of *Pyrococcus abyssi* tRNA (uracil-54, C5)-methyltransferase provides insights into its tRNA specificity. Nucleic Acids Res 36 (2008) 4929-4940.
[6] J.M. Romeo, A.S. Delk, J.C. Rabinowitz, The occurrence of a transmethylation reaction not involving S-adenosylmethionine in the formation of ribothymidine in *Bacillus subtilis* transfer-RNA. Biochem Biophys Res Commun 61 (1974) 1256-1261.
[7] A.S. Delk, J.C. Rabinowitz, Biosynthesis of ribosylthymine in the transfer RNA of *Streptococcus faecalis:* a folate-dependent methylation not involving S-adenosylmethionine. Proc Natl Acad Sci U S A 72 (1975) 528-530.
[8] W. Schmidt, H.H. Arnold, H. Kersten, Tetrahydrofolate-dependent biosynthesis of ribothymidine in transfer ribonucleic acids of Gram-positive bacteria. J Bacteriol 129 (1977) 15-21.
[9] A.S. Delk, D.P. Nagle, Jr., J.C. Rabinowitz, Methylenetetrahydrofolate-dependent biosynthesis of ribothymidine in transfer RNA of *Streptococcus faecalis*. Evidence for reduction of the 1-carbon unit by FADH2. J Biol Chem 255 (1980) 4387-4390.
[10] J. Urbonavicius, S. Skouloubris, H. Myllykallio, H. Grosjean, Identification of a novel gene encoding a flavin-dependent tRNA:m5U methyltransferase in bacteria--evolutionary implications. Nucleic Acids Res 33 (2005) 3955-3964.
[11] N. Cicmil, Crystallization and preliminary X-ray crystallographic characterization of TrmFO, a folate-dependent tRNA methyltransferase from *Thermotoga maritima*. Acta Crystallogr Sect F Struct Biol Cryst Commun 64 (2008) 193-195.
[12] H. Nishimasu, R. Ishitani, K. Yamashita, C. Iwashita, A. Hirata, H. Hori, O. Nureki, Atomic structure of a folate/FAD-dependent tRNA T54 methyltransferase. Proc Natl Acad Sci U S A 106 (2009) 8180-8185.
[13] S. Meyer, A. Scrima, W. Versees, A. Wittinghofer, Crystal structures of the conserved tRNA-modifying enzyme GidA: implications for its interaction with MnmE and substrate. J Mol Biol 380 (2008) 532-547.
[14] I. Moukadiri, S. Prado, J. Piera, A. Velazquez-Campoy, G.R. Bjork, M.E. Armengod, Evolutionarily conserved proteins MnmE and GidA catalyze the formation of two methyluridine derivatives at tRNA wobble positions. Nucleic Acids Res 37 (2009) 7177-7193.
[15] O. Vallon, New sequence motifs in flavoproteins: evidence for common ancestry and tools to predict structure. Proteins 38 (2000) 95-114.
[16] O. Dym, D. Eisenberg, Sequence-structure analysis of FAD-containing proteins. Protein Sci 10 (2001) 1712-1728.
[17] C.G. Xu, X.J. Fan, Y.J. Fu, A.H. Liang, Effect of location of the His-tag on the production of soluble and functional *Buthus martensii* Karsch insect toxin. Protein Expr Purif 59 (2008) 103-109.
[18] J. Urbonavicius, C. Brochier-Armanet, S. Skouloubris, H. Myllykallio, H. Grosjean, In vitro detection of the enzymatic activity of folate-dependent tRNA (Uracil-54,-C5)-methyltransferase: evolutionary implications. Methods Enzymol 425 (2007) 103-119.
[19] J.M. Scholtz, H. Qian, E.J. York, J.M. Stewart, R.L. Baldwin, Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water. Biopolymers 31 (1991) 1463-1470.
[20] H. Grosjean, L. Droogmans, M. Roovers, G. Keith, Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. Methods Enzymol 425 (2007) 55-101.
[21] M.A. O'Connell, W. Keller, Purification and properties of double-stranded RNA-specific adenosine deaminase from calf thymus. Proc Natl Acad Sci U S A 91 (1994) 10596-10600.

[22] S.C. Nam, C. Kang, Expression of cloned cDNA for the human mitochondrial RNA polymerase in Escherichia coli and purification. Protein Expr Purif 21 (2001) 485-491.
[23] E. Moncke-Buchner, P. Mackeldanz, D.H. Kruger, M. Reuter, Overexpression and affinity chromatography purification of the Type III restriction endonuclease EcoP15I for use in transcriptome analysis. J Biotechnol 114 (2004) 99-106.
[24] M. Husain, V. Massey, Reversible resolution of flavoproteins into apoproteins and fee flavins. Methods Enzymol 53 (1978) 429-437.

Table 1A. Purification of TrmFO from 500 mL of *E. coli* culture.

Protein	Purification step	Total protein (mg)[1]	TrmFO[2] (mg)
Nt-TrmFO	Ni-NTA	61	34
	Superdex S200	51	25.5
Ct-TrmFO	Ni-NTA	46	22
	Superdex S200	27	12.5
Untagged-TrmFO	DEAE Sepharose	131.8	94
	Heparin Sepharose	49.5	32
	Superdex S200	33	13.2

[1] determined by the Bradford assay.
[2] based on the flavin absorbance (does not take into account the apoprotein).

Table 1B. Hydrodynamic parameters of tagged and untagged TrmFO determined by size-exclusion chromatography on a Superdex 200 HR 10/30 column.

	Untagged TrmFO	Nt-TrmFO	Ct-TrmFO
Elution volume (mL)	15.66	15.04	15.01
K_{AV}	0.472	0.427	0.425
$M_{theoretical}$ (kDa)	48.12	49.46	49.18
$M_{experimental}$ (kDa)	33.57	47.86	47.72

Figure legends

Fig. 1 A Enzymatic reaction catalyzed by TrmFO (MTHF: 5,10-methylenetetrahydrofolate; THF: tetrahydrofolate; FAD: flavin adenine dinucleotide). B Stereoview of the structure of TrmFO from *T. thermophilus* (PDB code 3G5R) showing the two domains architecture of the protein. The flavin domain is shown in purple and the insertion domain in red. The FAD cofactor is in yellow, tetrahydrofolate in green, the N and C-terminal regions in green and cyan, respectively.
C Stereoview showing details of the folate-binding site and the location of Arg319 in pink.

Fig. 2 Purification of TrmFO analyzed by SDS PAGE (10%). (A) Nt-TrmFO and (B) Ct-TrmFO. Lane 1: total protein before cell disruption. Lane 2: crude cell extract. Lane 3: unbound protein fraction after Ni-NTA column. Lane 4: eluted protein fraction after Ni-NTA column. (C) untagged TrmFO. Lane 1: total protein before cell disruption. Lane 2: crude cell extract. Lane 3: eluted protein fraction after the DEAE Sepharose column. Lane 4: eluted protein fraction after the heparin Sepharose column. Lane 5: eluted protein fraction after the gel filtration column.

Fig. 3
Purification of untagged TrmFO.
A HiTrap DEAE Sepharose chromatography. The four peaks at the beginning of the chromatogram represent different injections of the sample. The SDS PAGE done on the fractions, which were not retained on the column, indicates that they do not contain TrmFO.
B HiTrap heparin Sepharose chromatography
Elution of TrmFO was recorded by following the absorbance at 280 nm (—) and 450 nm (- - -).

Fig. 4
Purification of recombinant TrmFO by size-exclusion chromatography on a Superdex 200 HR 10/30 column. (A) and (B) are the elution profiles after affinity Ni-NTA chromatography of the Nt- and Ct-TrmFO proteins, respectively. (C) Elution profile after the heparin Sepharose column of untagged TrmFO. The absorbance was followed at 280 nm (—) and 450 nm (- - -, 3-fold magnification). The insets are Coomassie blue stained SDS-PAGE (10 %) of the collected fractions for each TrmFO form.

Fig. 5
tRNA m^5U MTase activity. *E. coli* [α-^{32}P]UTP-labeled $tRNA^{Ala1}$ transcript was incubated with methylenetetrahydrofolate, NADH in the absence (A) or presence (B) of TrmFO, as described in Material and Methods. After incubation, the tRNA transcript was digested by nuclease P1, and the resulting nucleotides were analyzed by two-dimensional TLC on cellulose plates and autoradiography. (C) The activity of the R312G mutant, tagged and untagged proteins was quantified by measuring the relative amount of radioactivity in the spots corresponding to pm^5U and pU, taking into account the total number of uridines in the tRNA molecule. The error bar corresponds to the average of at least three distinct experiments.

Figure 1

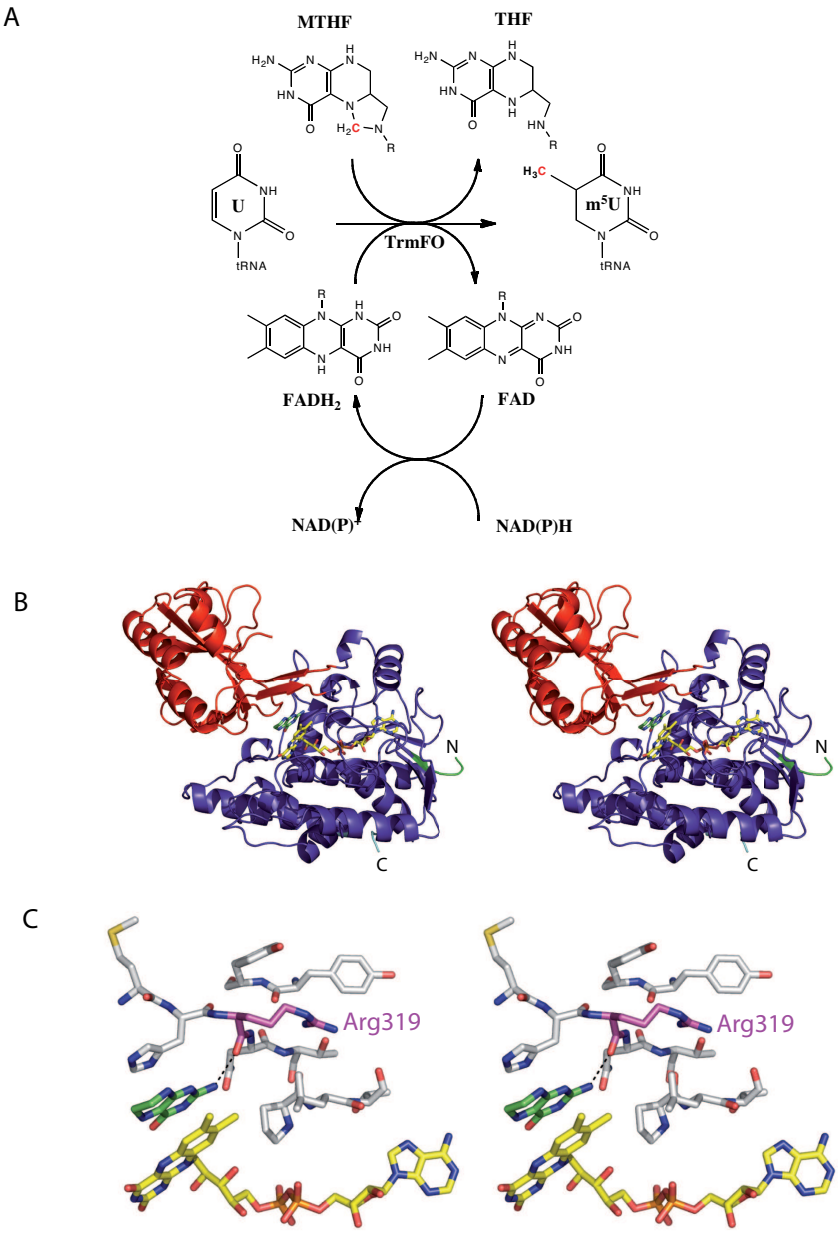

A
MTHF
THF
U
tRNA
TrmFO
m5U
tRNA
FADH2
FAD
NAD(P)+
NAD(P)H
B
N
C
N
C
C
Arg319
Arg319

Figure 2

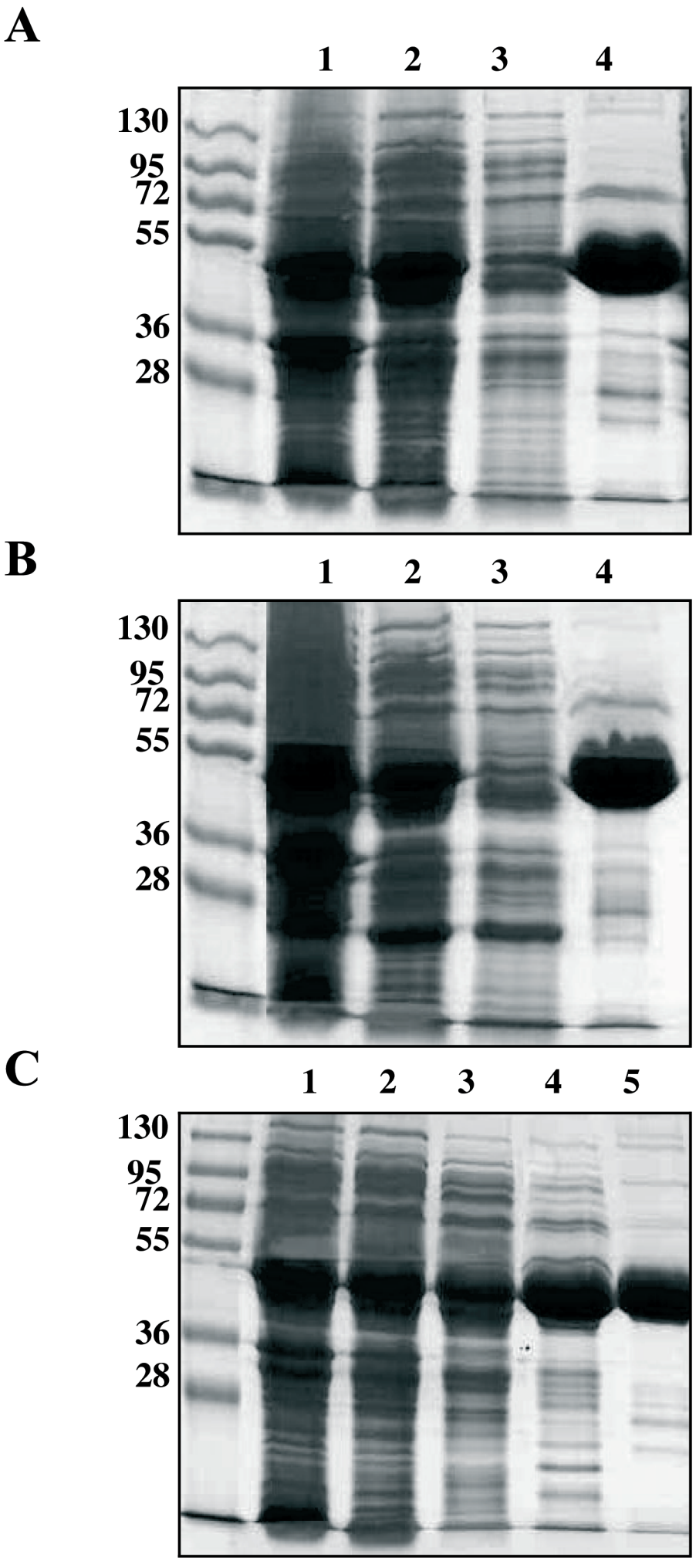
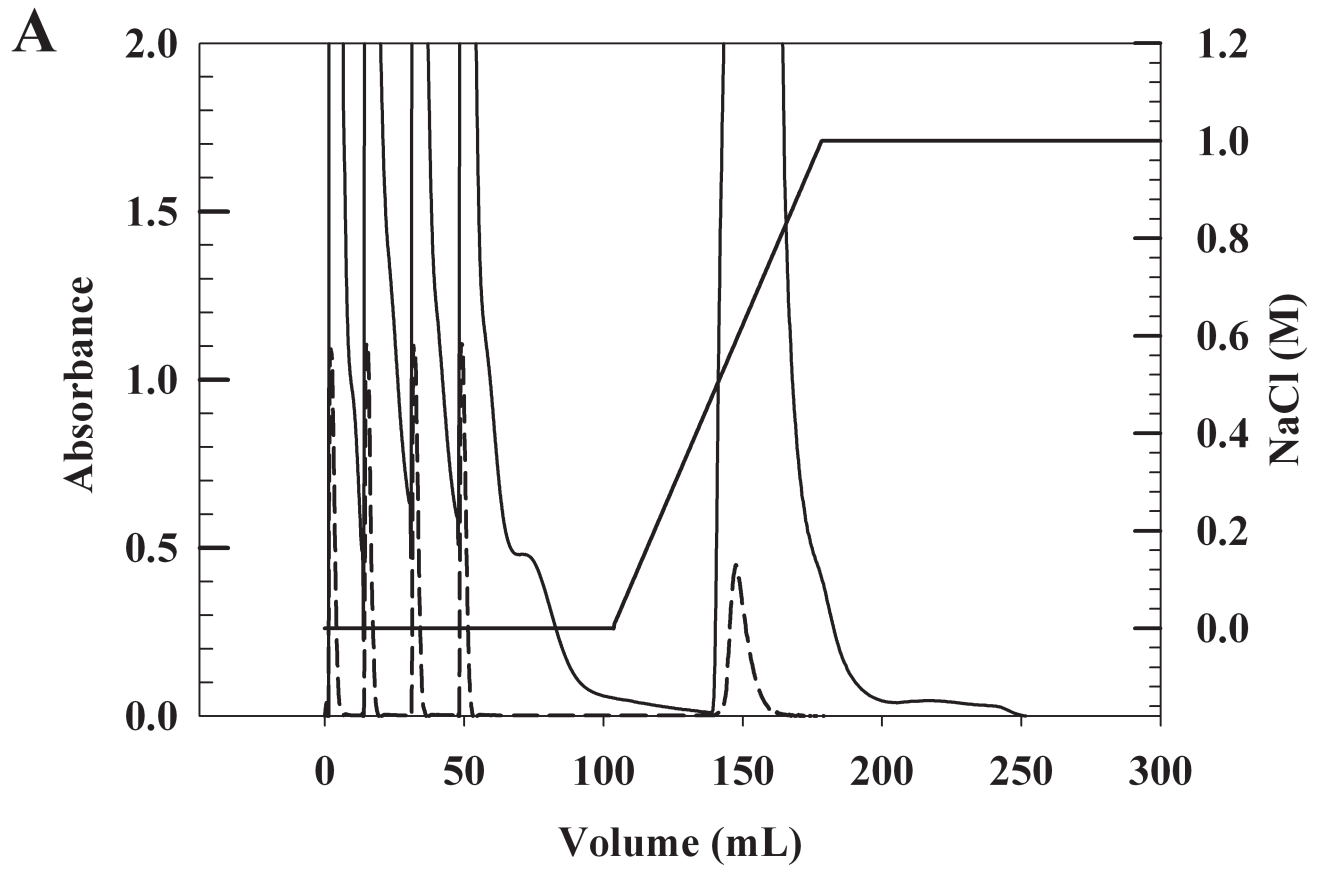

A
1 2 3 4
130
95
72
55
36
28
B
1 2 3 4
130
95
72
55
36
28
C
1 2 3 4 5
130
95
72
55
36
28

Figure 3


A
Absorbance
NaCl (M)
Volume (mL)
0.0
0.5
1.0
1.5
2.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
50
100
150
200
250
300

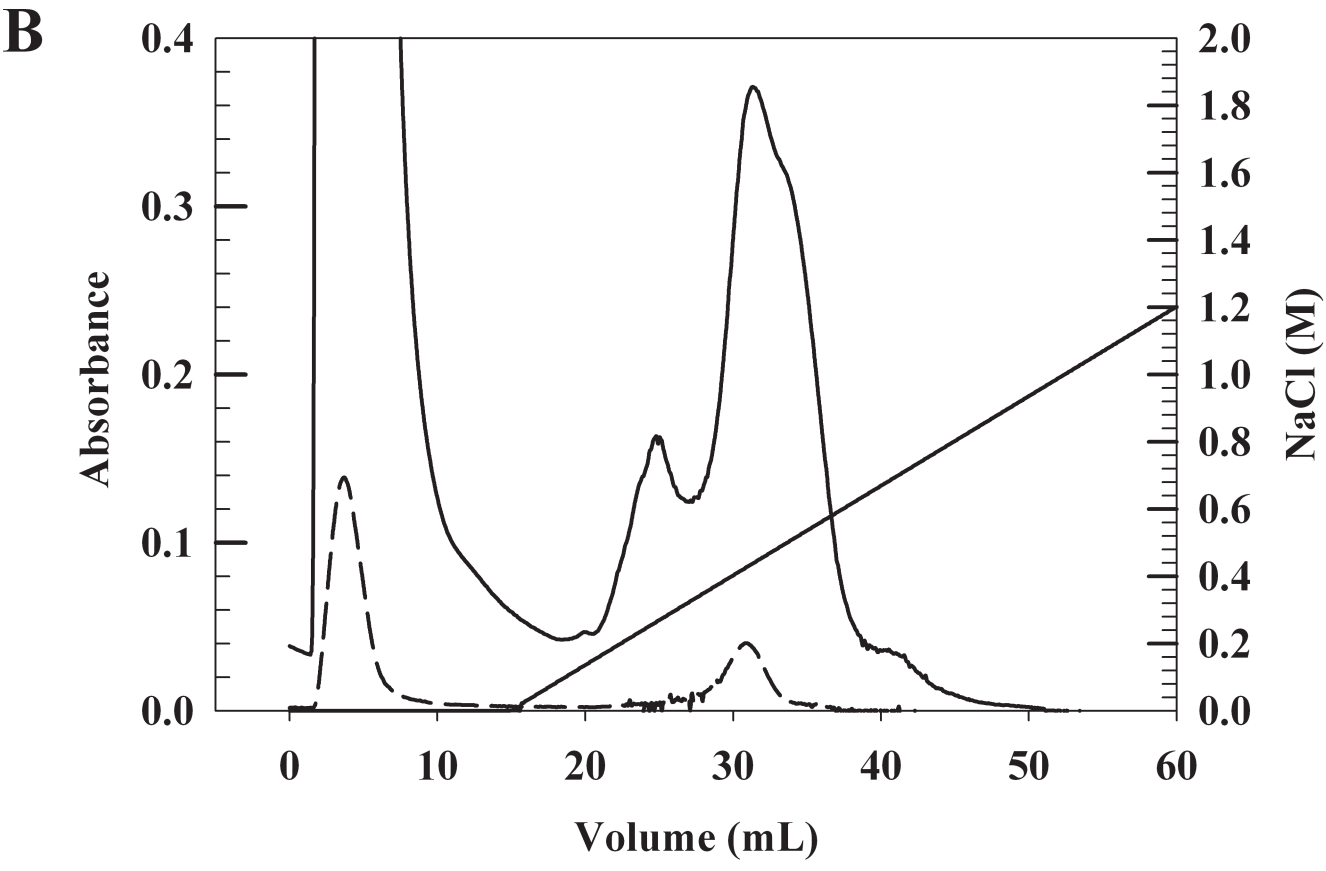

B
Absorbance
NaCl (M)
Volume (mL)
0.0
0.1
0.2
0.3
0.4
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
0
10
20
30
40
50
60

Figure 4

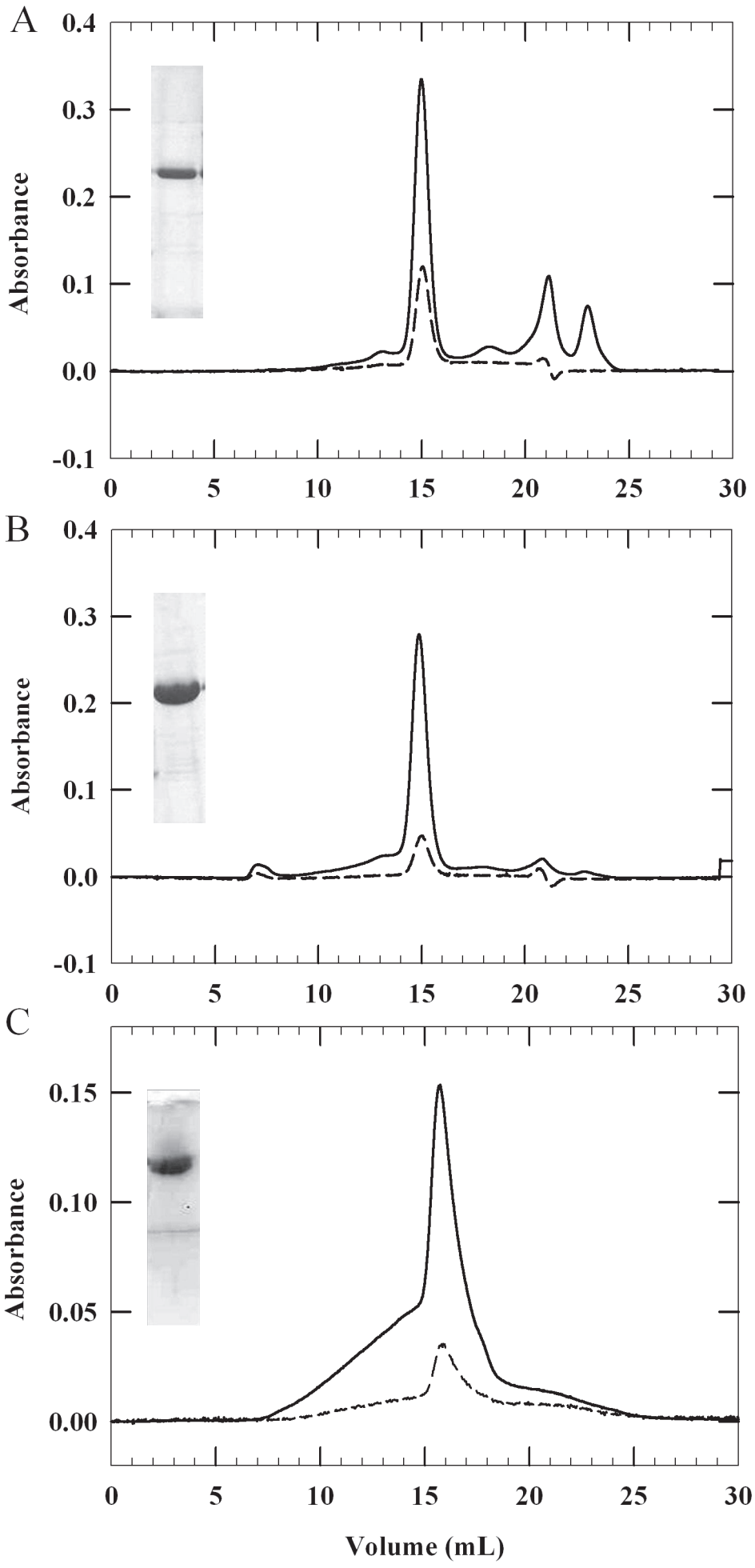
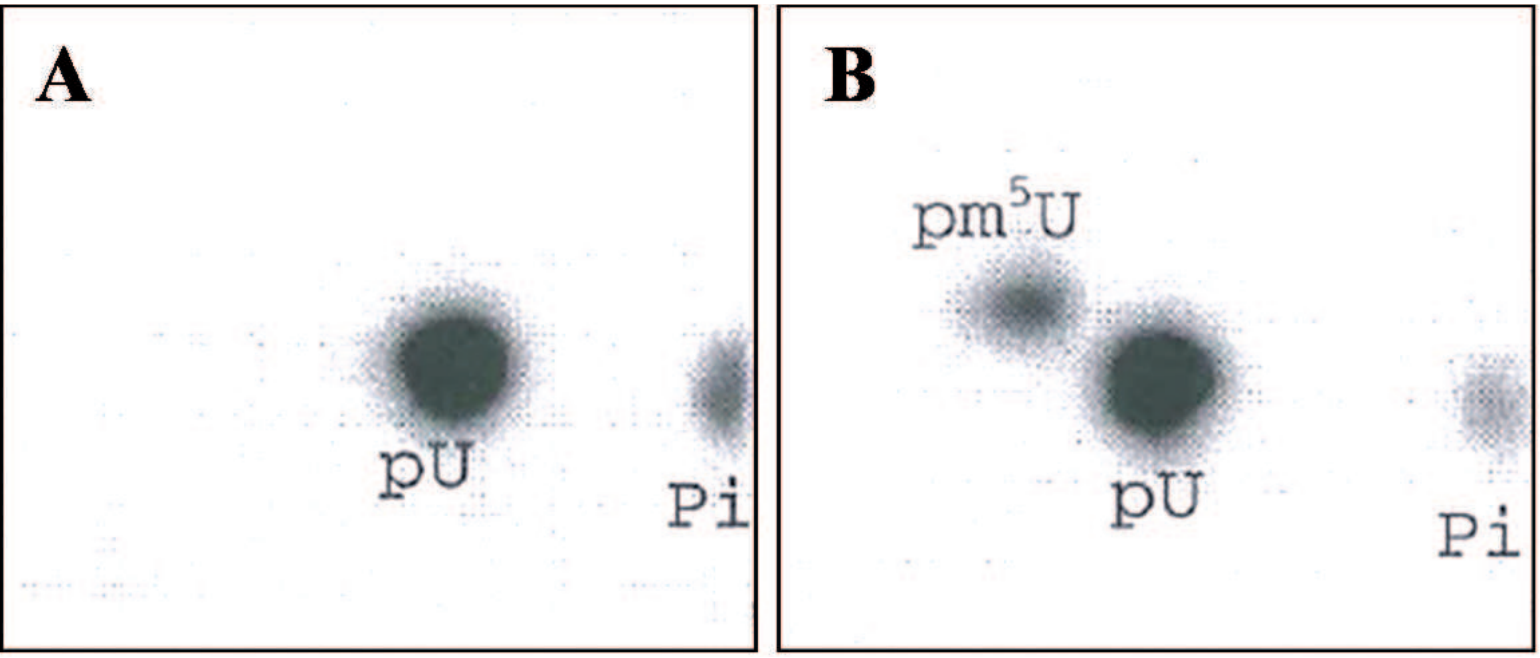
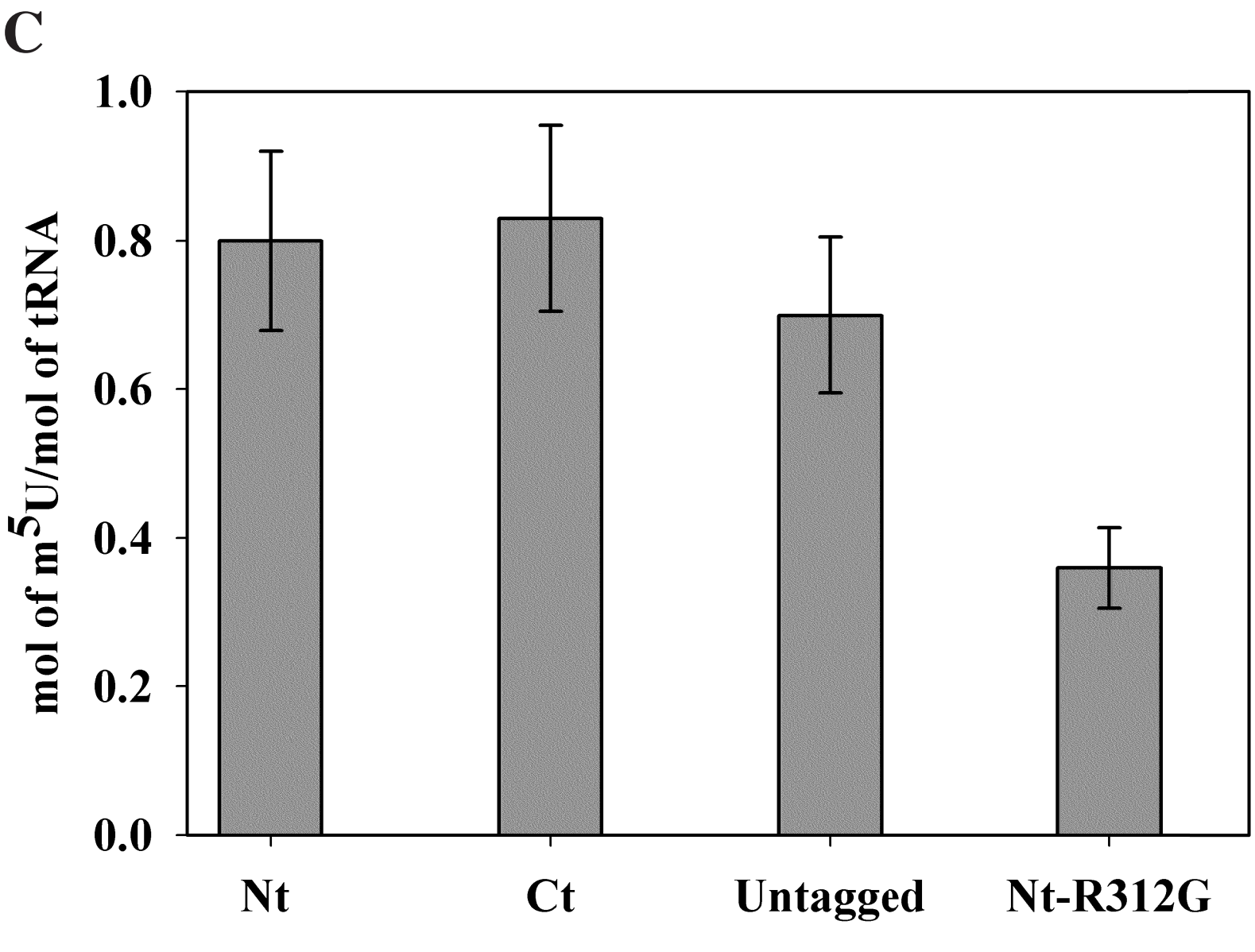

A
B
C
Absorbance
Absorbance
Absorbance
0.4
0.3
0.2
0.1
0.0
-0.1
0.15
0.10
0.05
0.00
0
5
10
15
20
25
30
Volume (mL)

Figure 5


A
pU
Pi
B
pm5U
pU
Pi


C
1.0
0.8
0.6
0.4
0.2
0.0
mol of m5U/mol of tRNA
Nt
Ct
Untagged
Nt-R312G

Supplementary data

Supplementary Table

Table S1: Oligonucleotides used for cloning of the TrmFO gene

Name	Sequence (5' to 3')
SK20	CGGGATCCatgaaccaacaaacagtgaatgta
SK21	TCCCCCGGGctatattgttttcgaaattgtttg
SK76	CATGCCATGGGCaaccaacaaacagtgaat
SK77	CCGCTCGAGtattgttttcgaaattgt
SK78	CCGCTCGAGctatattgttttcgaaattgt

The sequence in small characters corresponds to the genomic sequence

Supplementary Figures

Fig. S1: Structure of the expression plasmids. A pHM408 (Nt-TrmFO). B pHM440 (Ct-TrmFO). C pHM441 (untagged TrmFO).

Fig. S2: Expression of soluble Ct-TrmFO at different IPTG concentrations (0.05, 0.5 and 1 mM) and different incubation length before induction (OD_{600}=1, column 1; OD_{600}=2, column 2). Analysis was done on a 10% SDS PAGE gel stained by Coomassie blue. The arrow indicates the mass corresponding to the TrmFO protein. Top: Total protein after cell disruption. Bottom: Total protein from soluble fraction after cell disruption.

Fig. S3: Far-UV CD spectrum of Nt-TrmFO. The spectra of the Ct-TrmFO and untagged TrmFO are similar.

Fig. S4: Determination of the molecular weight of the various TrmFO proteins
The Superdex 200 HR 10/30 column was calibrated with a standard mixture of proteins. From the elution volume of each protein V_e, the elution volume parameter $K_{AV} = (V_e-V_0)/(V_t-V_0)$ was determined, with V_t, the total column volume and V_0, the void volume. K_{AV} is a linear function of the logarithm of the molecular weight (MM).

Fig. S5: Steady-state NAD(P)H oxidase activity of TrmFO. The reactions, carried out aerobically in the presence of the indicated concentrations of NAD(P)H, were initiated by the addition of 1 μM TrmFO.

Figure S1

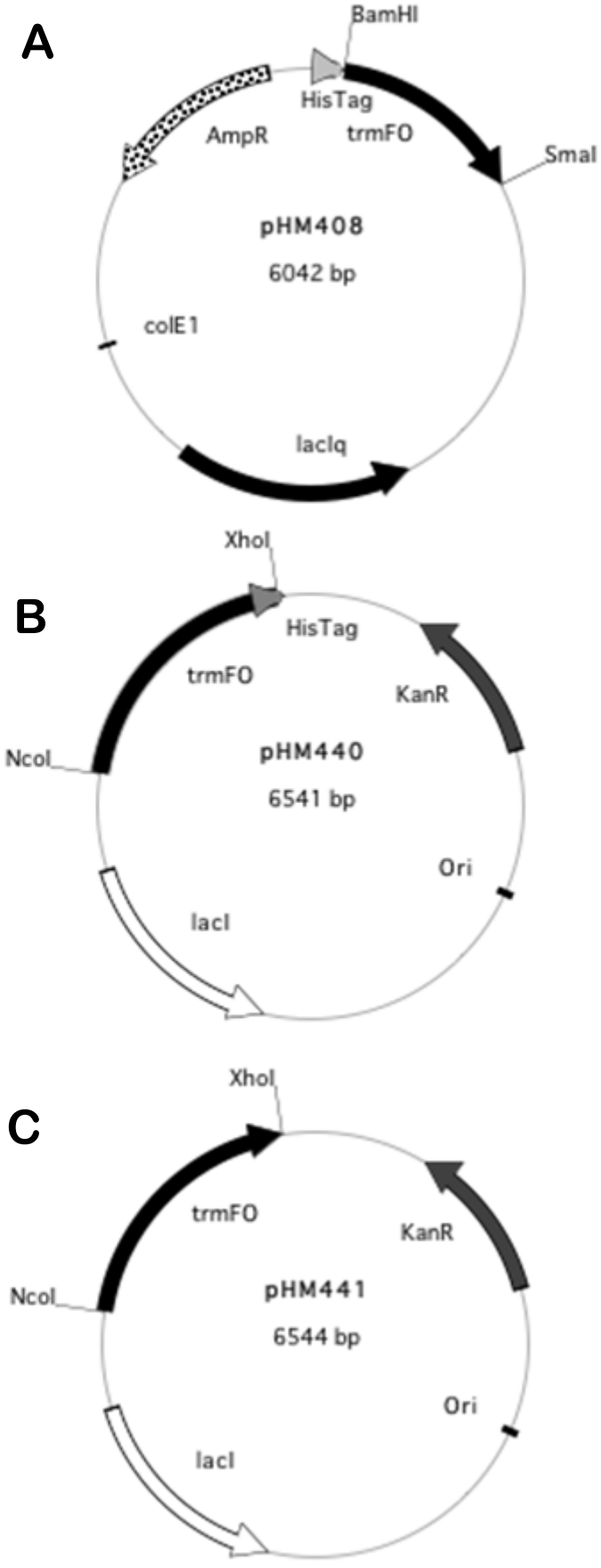

A
BamHI
HisTag
AmpR
trmFO
SmaI
pHM408
6042 bp
colE1
lacIq
B
XhoI
HisTag
trmFO
KanR
NcoI
pHM440
6541 bp
Ori
lacI
C
XhoI
trmFO
KanR
NcoI
pHM441
6544 bp
Ori
lacI

Figure S2

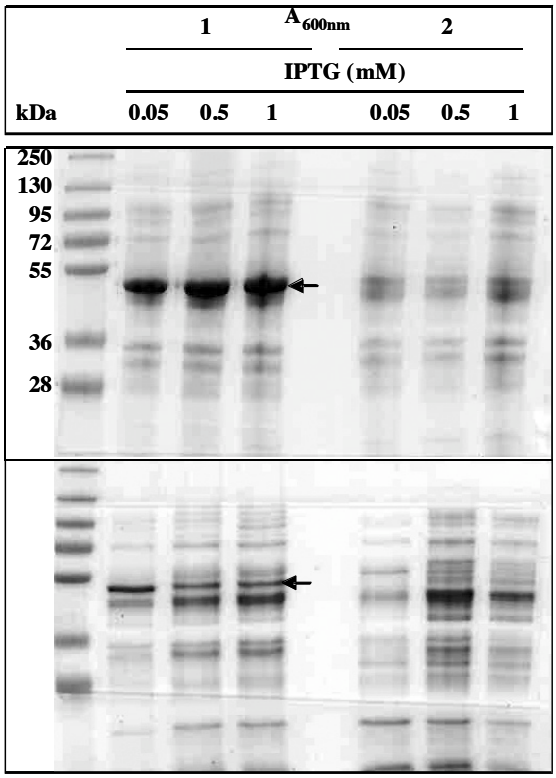


Figure S3

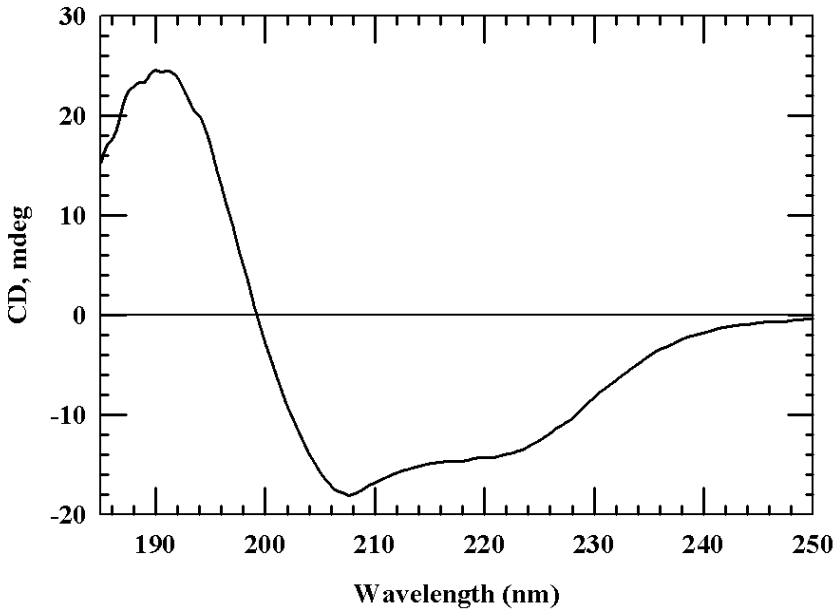

30
20
10
0
-10
-20
CD, mdeg
190
200
210
220
230
240
250
Wavelength (nm)

Figure S4

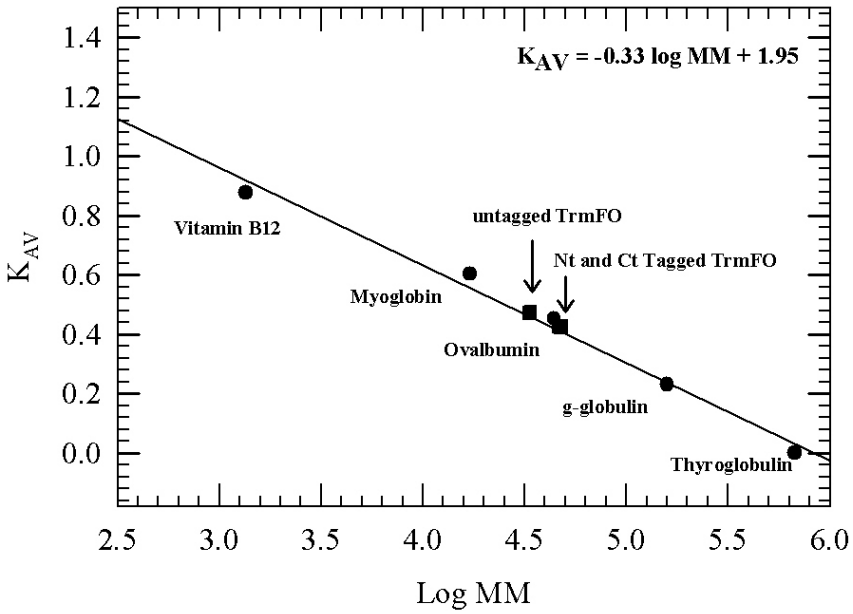
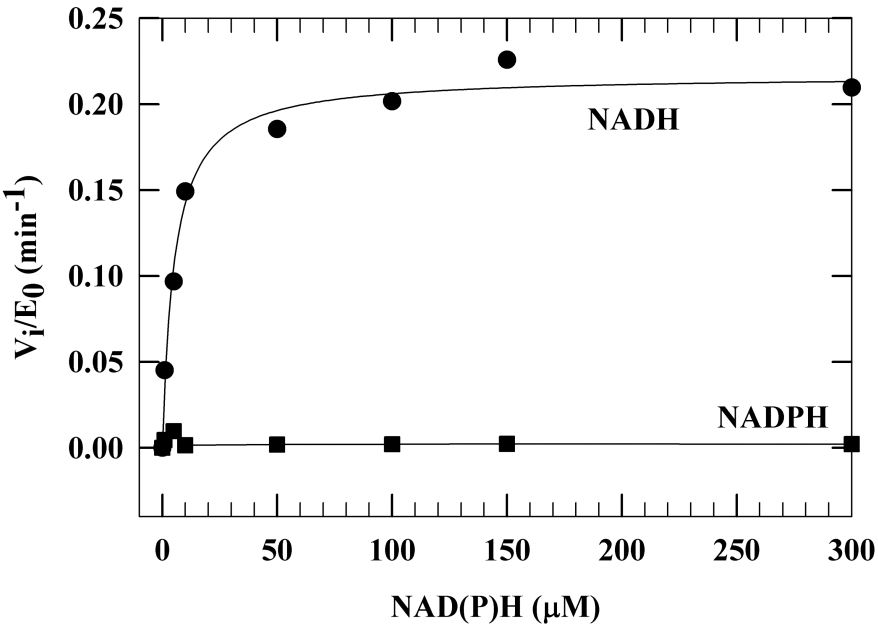

K_{AV} = -0.33 log MM + 1.95
Vitamin B12
Myoglobin
untagged TrmFO
Nt and Ct Tagged TrmFO
Ovalbumin
g-globulin
Thyroglobulin
K_{AV}
Log MM
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0

Figure S5


0.25
0.20
0.15
0.10
0.05
0.00
V_i/E_0 (min^{-1})
NADH
NADPH
0
50
100
150
200
250
300
NAD(P)H (μM)