Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure analysis of the conserved methyltransferase domain of human trimethylguanosine synthase TGS1

Thomas Monecke, Achim Dickmanns, Anja Strasser and Ralf Ficner*

Department of Molecular Structural Biology, Institute of Microbiology and Genetics, Georg-August-University Göttingen, Justus-von-Liebig-Weg 11, D-37077 Göttingen, Germany

Correspondence e-mail: rficner@uni-goettingen.de

Received 11 December 2008
Accepted 24 January 2009

PDB Reference: methyltransferase domain of TGS1, 3egi, r3egisf.

Methyltransferases play an important role in the post-transcriptional maturation of most ribonucleic acids. The modification of spliceosomal UsnRNAs includes N2-dimethylation of the m^7G cap catalyzed by trimethylguanosine synthase 1 (TGS1). This 5′-cap hypermethylation occurs during the biogenesis of UsnRNPs as it initiates the m_3G cap-dependent nuclear import of UsnRNPs. The conserved methyltransferase domain of human TGS1 has been purified, crystallized and the crystal structure of this domain with bound substrate m^7GpppA was solved by means of multiple-wavelength anomalous dispersion. Crystal structure analysis revealed that m^7GpppA binds *via* its adenosine moiety to the structurally conserved adenosylmethionine-binding pocket, while the m^7 guanosine remains unbound. This unexpected binding only occurs in the absence of AdoMet and suggests an incomplete binding pocket for the m^7G cap which is caused by the N-terminal truncation of the protein. These structural data are consistent with the finding that the crystallized fragment of human TGS1 is catalytically inactive, while a fragment that is 17 amino acids longer exhibits activity.

1. Introduction

S-Adenosyl-L-methionine (AdoMet) dependent methyltransferases (MTases) are involved in many different cellular processes including the post-transcriptional modification of RNAs. Some 74 different methylated RNA nucleosides have been identified in the three kingdoms of life. The methylation of guanosine concerns atoms N1, N2, N7 and 2′O and various combinations of these methylations have been found (m^1G, m^2G, m^7G, Gm, $m^{2,2}G$, $m^{2,7}G$, $m^{2,2}Gm$, $m^{2,2,7}G$, m^1Gm, $m^{2,7}Gm$; Limbach *et al.*, 1994). All RNA methylations are introduced post-transcriptionally by AdoMet-dependent MTases, most of which belong to the class I MTases, which are characterized by a Rossmann-fold-like $\alpha\beta$ structure (Schubert *et al.*, 2003).

One member of this family is the trimethylguanosine synthase 1 (TGS1), which catalyzes the N2-dimethylation of the m^7G cap of spliceosomal uridyl-rich small nuclear RNAs (UsnRNAs) and of some small nucleolar RNAs (snoRNAs) (Hausmann & Shuman, 2005*a*; Maxwell & Fournier, 1995; Mouaikel, Bujnicki *et al.*, 2003; Mouaikel *et al.*, 2002). TGS1 enzymes from *Saccharomyces cerevisiae* (yTGS1), *Giardia lamblia*, *Schizosaccharomyces pombe*, *Drosophila melano-*

gaster, *Trypanosoma brucei* and human cells (hTGS1) have been characterized with respect to their biochemical properties as well as their interaction with UsnRNPs or snoRNPs (small nuclear/nucleolar ribonucleoprotein particles) (Colau *et al.*, 2004; Enunlu *et al.*, 2003; Girard *et al.*, 2008; Gunzl *et al.*, 2000; Hausmann *et al.*, 2007, 2008; Hausmann & Shuman, 2005*a*,*b*; Komonyi *et al.*, 2005; Misra *et al.*, 2002; Mouaikel, Bujnicki *et al.*, 2003; Mouaikel, Narayanan *et al.*, 2003; Mouaikel *et al.*, 2002; Plessel *et al.*, 1994; Ruan *et al.*, 2007; Watkins *et al.*, 2004; Zhu *et al.*, 2001). There is a significant difference between organisms regarding the size of TGS1, which varies from 239 residues in *S. pombe* to 853 residues in *Homo sapiens*, as well as its cellular localization. Yeast TGS1 was shown to act in the nucleus exclusively (Mouaikel *et al.*, 2002), whereas hTGS1 methylates spliceosomal UsnRNAs in the cytoplasm and snoRNAs in the nucleus (Colau *et al.*, 2004; Mouaikel, Narayanan *et al.*, 2003; Verheggen *et al.*, 2002).

UsnRNA 5′-cap hypermethylation plays an important role during the biogenesis of UsnRNPs (Dickmanns & Ficner, 2005). In higher eukaryotes, the maturation of UsnRNPs comprises a nucleocytoplasmic transport cycle. Newly transcribed snRNAs U1, U2, U4 and U5 are exported to the cytoplasm in an m^7G cap-dependent manner, where assembly with seven Sm proteins occurs (Hamm *et al.*, 1990; Mattaj, 1986). This assembly process is mediated by the survival of motor neuron complex (SMN complex), a large multiprotein complex (Neuenkirchen *et al.*, 2008). The m^7G cap is subsequently hypermethylated by TGS1 and the resulting m_3G cap is recognized by the nuclear import adaptor snurportin1 (Huber *et al.*, 1998; Strasser *et al.*, 2005), which binds to the general nuclear import receptor importinβ. Hence, the m_3G cap serves as nuclear import signal that indicates the completed assembly of the core UsnRNP particle. Therefore, the interaction of TGS1 with UsnRNP proteins SmB/B′ and D1 as well as with the SMN complex appears to correlate with the ordered process of RNP assembly and subsequent cap hypermethylation (Mouaikel, Narayanan *et al.*, 2003; Mouaikel *et al.*, 2002).

Biochemical studies have revealed that TGS1 is specific for m^7G-capped RNA and m^7GTP, which represents the minimal substrate, while nonmethylated 5′-cap RNA or GTP are not N2-dimethylated (Hausmann & Shuman, 2005*a*,*b*; Hausmann *et al.*, 2008). TGS1 catalyzes two successive methyl-transfer reactions using AdoMet as a methyl-group donor, which includes the formation of the intermediate product N^2,N^7-dimethylguanosine (Hausmann & Shuman, 2005*a*; Hausmann *et al.*, 2008). Furthermore, a three-dimensional structure model of yTGS1 was generated by means of homology modelling, based on which the m^7G cap-binding pocket was predicted (Mouaikel, Bujnicki *et al.*, 2003).

In order to verify the proposed structural model and the mode of m^7G cap recognition, we have crystallized the predicted MTase domain of hTGS1 in the presence of the substrate dinucleotide m^7GpppA. Our crystal structure analysis and additional biochemical studies demonstrate that the predicted MTase domain is catalytically inactive owing to a lack of m^7G cap binding. We show that additional N-terminal residues enlarging the canonical MTase domain are required for enzymatic activity.

2. Materials and methods

2.1. Protein expression and purification

Human TGS1 fragments (amino acids 636–853 and 653–853) were subcloned from pGEX-6P-1 full-length TGS1 (accession No. Q96RS0) into the *Bam*HI/*Xho*I sites of pGEX-6P-3 (GE Healthcare, Germany) and verified by sequencing. The following primers were used for subcloning: MT636_forward, 5′-CGC**GGATCC**CCTGAAATAGCTGCTGTTCCTGAGC-3′ (*Bam*HI site in bold), MT653_forward, 5′-CGC**GGATCC**AGGCTCTTCTCCCGTTTTGATG-3′ (*Bam*HI site in bold), and MT853_reverse, 5′-CCG**CTCGAG**TTAGGTTTCAGAGGCTGGTCTTCG-3′ (*Xho*I site in bold).

The native GST-fusion constructs for the activity tests were expressed in *Escherichia coli* BL21 (DE3) (Invitrogen, USA) at 289 K in ampicillin-containing 2YT medium, which was supplemented with 2%(*w*/*v*) glucose. Expression of constructs was induced at OD_{600} = 0.8, adding IPTG to a final concentration of 500 µ*M*. Immediately after induction, 2%(*v*/*v*) ethanol and 50 m*M* K_2HPO_4 were added to the growth culture. The cells were harvested after 18 h by centrifugation (5000*g*, 20 min, 277 K) and resuspended in lysis buffer containing 50 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl, 2 m*M* EDTA and 2 m*M* DTT. All subsequent steps were carried out at 277 K unless stated otherwise. Cells were disrupted using a 110S microfluidizer (Microfluidics, USA). The clarified lysate (30 000*g*, 30 min, 277 K) was subsequently loaded onto a GSH-Sepharose column (GE Healthcare, Germany), which was equilibrated with lysis buffer. Unbound proteins were removed by washing with two column volumes (CV) of lysis buffer. In order to eliminate RNA contaminants, the loaded column was washed with one CV of a high-salt buffer containing 50 m*M* Tris–HCl pH 7.5, 1 *M* NaCl and 2 m*M* DTT. After re-equilibration in lysis buffer, the bound fusion protein was eluted with lysis buffer additionally containing 25 m*M* reduced glutathione. GST-$hTGS1_{653–853}$ was incubated with PreScission protease (GE Healthcare, Germany) at 277 K overnight in order to cleave the fusion protein into GST and $hTGS1_{653–853}$ containing a multiple cloning site remainder of Gly-Pro-Leu-Gly-Ser at the N-terminus. $hTGS1_{653–853}$ was further purified using a Superdex S75 (26/60) gel-filtration column (GE Healthcare, Germany) equilibrated in a buffer containing 20 m*M* Tris–HCl pH 7.5, 200 m*M* NaCl and 2 m*M* DTT. The elution volume of $hTGS1_{636–853}$ and $hTGS1_{653–853}$ on the gel-filtration column corresponded to a monomeric state of the protein. The resulting pure protein was concentrated to 6 mg ml^{-1} using a Vivaspin concentrator with MWCO 10 000 Da (Sartorius, Germany) and stored in aliquots at 193 K.

The selenomethionine-containing human TGS1 fragment encompassing amino acids 653–853 was expressed according to the protocol described by Reuter & Ficner (1999). The purification of SeMet-$TGS1_{653–853}$ was performed as described

Table 1
Statistics of data sets for selenomethionine (SeMet) $TGS1_{653–853}$.

Values in parentheses are for the highest resolution shell.

Data set	Peak	Inflection
Data collection		
Space group	$P4_32_12$	
Unit-cell parameters (Å, °)	$a = b = 213.9$, $c = 62.4$, $\alpha = \beta = \gamma = 90$	
Wavelength (Å)	0.9799	0.9801
Resolution range (Å)	50.00–2.90 (3.00–2.90)	50.00–2.20 (2.28–2.20)
No. of reflections	32845	71603
Completeness (%)	100.0 (99.9)	97.0 (76.2)
R_{merge}† (%)	5.6 (16.2)	4.8 (31.1)
Average $I/\sigma(I)$	31.8 (11.5)	34.7 (3.3)
Redundancy	9.4 (7.4)	8.7 (4.4)
Mosaicity (°)	0.35	0.34
No. of Se sites per ASU	12	
Refinement		
Resolution (Å)	50.0–2.2	
Molecules per ASU	4	
No. of atoms		
Protein	6030	
Ligand	108	
Waters	631	
R_{work}‡ (%)	21.0	
R_{free}§ (%)	25.2	
Figure of merit	0.82	
Average B factors (Å^2)		
Protein	38.0	
Ligand	37.7	
Waters	46.4	
R.m.s. deviations		
Bond lengths (Å)	0.010	
Bond angles (°)	1.291	
Ramachandran statistics (%)		
Most favoured	92.0	
Allowed	8.0	
Generous	0.0	
Disallowed	0.0	

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where the sum i is over all separate measurements of the unique reflection hkl. ‡ $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$. § R_{free} as R_{work} but summed over a 5% test set of reflections.

for the native proteins with the exception that the DTT concentration was elevated to 5 m*M* in all buffers in order to prevent oxidation of the selenium.

2.2. HPLC-based activity assay

In order to analyze the activity of the purified human TGS1 fragments, an HPLC-based activity assay was developed. 25 µ*M* purified protein was mixed with 0.5 m*M* cap analogue m^{7}GpppA (KEDAR, Poland) and 2 m*M* AdoMet (Sigma–Aldrich, Germany) in 1× PBS. The reaction mixture, with a total volume of 10 µl, was incubated at 310 K and the reaction was stopped by addition of 1 µl 1 *M* $HClO_4$ and incubation on ice for 1 min. The solution was neutralized by adding 20 µl 2 *M* sodium acetate. Precipitated protein was pelleted by centrifugation (16 000*g*, 10 min, 293 K) and the supernatant was loaded onto a reversed-phase HPLC column (Prontosil C18-AQ, Bischoff Chromatography, Germany), which was equilibrated in phosphate buffer *A* containing 100 m*M* K_2HPO_4/KH_2PO_4 pH 6.5. The substrates and products of the reaction were eluted from the column by applying a linear gradient from 0 to 60% buffer *B*, which consisted of buffer *A* with an additional 50%(*v*/*v*) acetonitrile. Commercially available m^{7}GpppA, AdoMet, AdoHcy (Sigma–Aldrich, Germany) and m^{2,2,7}GpppA (KEDAR, Poland) served as references for column calibration.

2.3. Crystallization and structure determination

The human TGS1 fragment was crystallized by the vapour-diffusion method in sitting-drop 24-well Cryschem plates (Hampton Research, USA). SeMet-containing TGS1 comprising amino acids 653–853 was crystallized in the presence of a sevenfold molar excess of the cap dinucleotide m^{7}GpppA. 1 µl reservoir solution containing 1.5 *M* sodium formate and 0.1 *M* MES pH 6.1 was mixed with 1 µl of the prepared protein-substrate solution (6 mg ml^{-1}). Single crystals with dimensions of 70 × 70 × 300 µm grew within 3 d at 293 K and belonged to space group $P4_32_12$, with unit-cell parameters $a = b = 213.9$, $c = 62.4$ Å. The crystals were soaked in cryosolution containing an additional 20%(*v*/*v*) glycerol for 5 s and subsequently mounted on a goniometer head in a 100 K cryostream. Peak, inflection-point and high-energy remote data sets were collected from an SeMet crystal on beamline BW7A at EMBL/DESY in Hamburg; the appropriate wavelengths were determined using a fluorescence scan. The remote data set was not used for phasing and refinement as it showed an increased R_{merge} compared with the peak and inflection-point data owing to radiation damage. The crystal was rotated in steps of 0.3° for the peak data set over a total range of 120° and in steps of 0.2° over the same range for the inflection-point data, while the distance was changed in between. Since SeMet-containing crystals diffracted to higher resolution than native crystals, phasing as well as refinement was performed using the SeMet-derivative crystals only. Data were integrated, scaled and reduced with the *HKL*-2000 suite (HKL Research, USA) and phases were obtained using *HKL2MAP* (Pape & Schneider, 2004). The resulting experimental electron-density map was used in *ARP*/*wARP* (Morris *et al.*, 2003) to build an initial model. The model was refined against the high-resolution inflection-point data set by iterative cycles of *REFMAC*5 (Murshudov *et al.*, 1997) and manual model building in *Coot* (Emsley & Cowtan, 2004). Waters were built using *Coot* and validated by hand. The structure was refined to good stereochemistry at a resolution of 2.2 Å to a final R_{work} of 21.0% and an R_{free} value of 25.2%. Owing to structural differences in the individual monomers, noncrystallographic symmetries (NCSs) were excluded from the refinement process. The Ramachandran plot of the refined structure model of human $TGS1_{653–853}$ generated with *SFCHECK* (Vaguine *et al.*, 1999) shows that 92% of the refined residues are located within the most favoured regions and 8% in additionally allowed regions; no residues lie in the generously allowed or disallowed regions.

The four monomers in the asymmetric unit show *D*2 symmetry with three twofold axes perpendicular to each other and consist of the following residues. Monomer 1 is defined by residues 649–848; residues 767–771 were not built owing to missing electron density. Molecules 2 and 3 are represented by

residues 649–847, but residues 767–773 are missing from the model. Monomer 4 consists of amino acids 649–847; there was no electron density for residues 662–665 and 768–773 and thus they were not built. Figures were generated using *PyMOL* (DeLano, 2002). The statistics of the X-ray diffraction data sets and structure refinement are summarized in Table 1.

3. Results and discussion

The conserved methyltransferase (MTase) domain of human TGS1 corresponds to the C-terminal 200 residues of the protein, while the function of the N-terminal 652 residues in snRNP biogenesis is yet unclear. We generated a truncated human TGS1 containing only the minimal MTase domain according to the homology model predicted for yeast TGS1 (Mouaikel, Bujnicki *et al.*, 2003), which comprises residues 653–853 ($TGS1_{653–853}$). This truncated hTGS1 was expressed in *E. coli*, purified and crystallized as described in §2. Crystals were only obtained when at least one of the reaction partners AdoMet, AdoHcy or m^7GpppA was present in the crystallization buffer; all attempts to crystallize the apoenzyme failed.

Cocrystallization with m^7GpppA yielded crystals that belonged to space group $P4_32_12$, whereas cocrystallization with AdoMet led to trigonal crystals which turned out to be almost perfectly twinned. Since the crystallographic phase problem could not be solved by means of molecular replacement, a selenomethionine (SeMet) derivative of $TGS1_{653–853}$ was produced and crystallized in the presence of the cap analogue m^7GpppA. A two-wavelength MAD experiment provided an interpretable electron-density map and the resulting crystal structure was refined at 2.2 Å resolution (Table 1). The overall structure of $TGS1_{653–853}$ closely resembles the canonical fold of class I methyltransferases, which is characterized by a central seven-stranded β-sheet flanked by several α-helices on both sites. However, $TGS1_{653–853}$ contains three additional N-terminal β-strands, of which the first extends the central β-sheet to eight strands with topology $\beta1\uparrow\beta9\uparrow\beta10\downarrow\beta8\uparrow\beta7\uparrow\beta4\uparrow\beta5\uparrow\beta6\uparrow$ (Fig. 1). This N-terminal β-strand 1 is located next to $\beta9$ and mediates important crystal-packing contacts. Strand $\beta1$ of each monomer in the asymmetric unit packs against strand $\beta1$ of a monomer of the adjacent asymmetric unit in an antiparallel fashion, thus forming an extended β-sheet.

The two additional short β-strands $\beta2$ and $\beta3$ connecting $\beta1$ and helix $\alpha1$ are only present in two of the four monomers in the asymmetric unit of the crystal (Fig. 2). These structural deviations are caused by different crystal-packing contacts, suggesting conformational flexibility of this region.

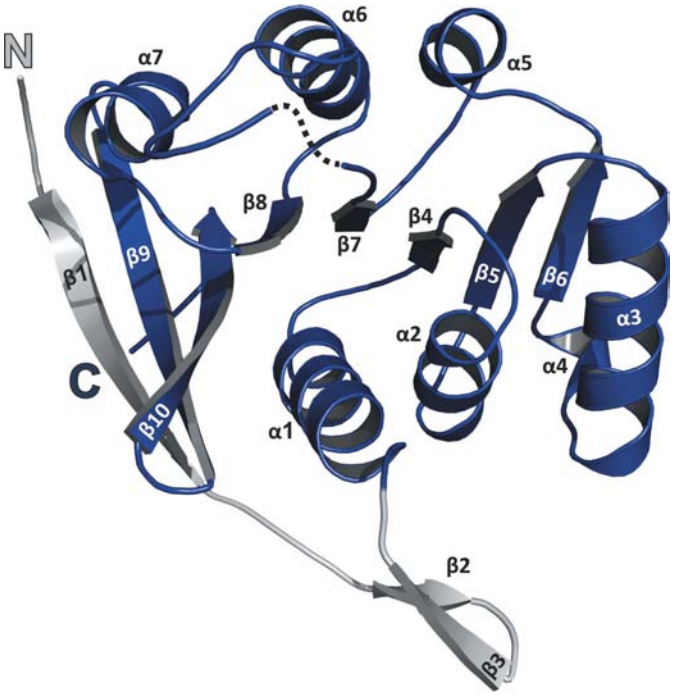

Figure 1
Overall structure of human TGS1 (amino acids 653–853) in cartoon representation. The canonical class I methyltransferase domain fold ($\alpha\beta\alpha$-sandwich) is coloured blue, while the additional N-terminal extension (β-strands 1–3) is depicted in grey. The secondary-structure elements as well as the N- and C-termini are labelled. The missing connection between residues Trp766 and Ala774, which is not defined in the electron-density map, is shown as a dotted line.

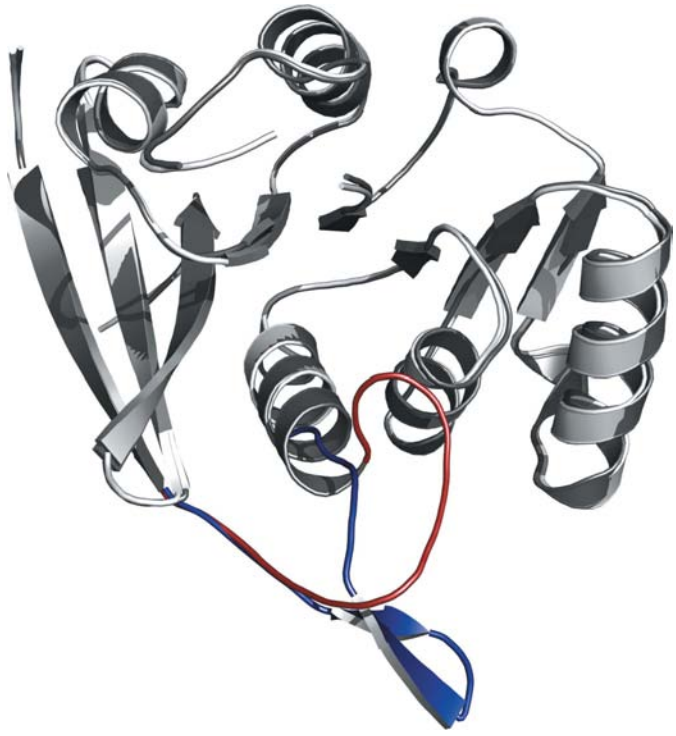

Figure 2
Superposition of the two different conformations present in the asymmetric unit. The four molecules in the asymmetric unit are evenly split into two conformational states. For each state only one molecule is depicted in the overlay, since the respective second molecule superposes almost perfectly. Within both conformations the major structural differences are restricted to the region connecting β-strand 1 to α-helix 1, as it harbours two β-strands in one state and no secondary-structure elements in the other. Therefore, this region is highlighted in blue and red for molecules 1 and 2, respectively, whereas the major parts of both molecules, coloured light and dark grey, respectively, are nearly identical.

Interestingly, the substrate m^{7}GpppA is bound *via* its adenosine diphosphate moiety in the structurally conserved AdoMet-binding cleft, while m^{7}G is fully disordered (Fig. 3). The adenine base is sandwiched between the hydrophobic side chains of Phe698 and Ile720 and its N6 atom forms a hydrogen bond to the carboxylate of Asp747. Both ribose hydroxyls are hydrogen bonded by the side chain of Asp719 and the β-phosphate is bound by the side chain of Lys724. This binding mode of the adenosine moiety of m^{7}GpppA closely resembles that of bound AdoMet or *S*-adenosyl-L-homocysteine (AdoHcy) in other class I methyltransferases (not shown).

The results of these crystallographic studies suggest that the crystallized TGS1$_{653–853}$ is not capable of binding the m^{7}G cap in the correct way. The observed binding of m^{7}GpppA to the AdoMet pocket occurs owing to a defective m^{7}G cap-binding pocket and only in absence of AdoMet in the crystallization buffer.

In order to confirm this interpretation of the structural data, the enzymatic activity of TGS1$_{653–853}$ was measured using a newly established HPLC-based assay. The purified protein was incubated with both substrates and the reaction was stopped by precipitation of the protein. After removal of the precipitated protein by centrifugation, the reaction substrates and products were separated by reversed-phase HPLC and quantified. The reversed-phase HPLC column was calibrated using commercially available standards of all substrates and products (Fig. 4*a*).

As expected, on testing the crystallized TGS1$_{653–853}$ neither formation of m^{2,2,7}GpppA nor of AdoHcy could be observed (Fig. 4*b*), confirming the data derived from the crystal structure. In contrast, a larger TGS1 fragment comprising residues 636–853 (TGS1$_{636–853}$) and thus containing 17 additional N-terminal residues was capable of converting m^{7}GpppA to m^{2,2,7}GpppA accompanied by the conversion of AdoMet to AdoHcy (Fig. 4*c*).

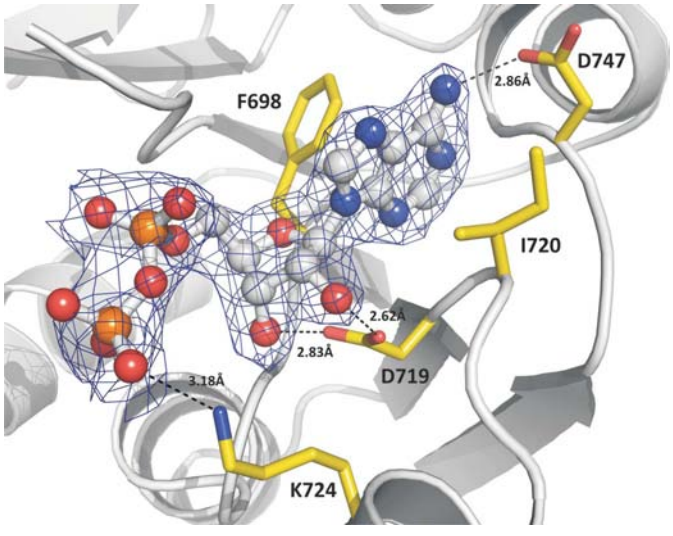

Figure 3
Detailed view of the bound adenosine moiety (ppA) from the co-crystallized m^{7}GpppA cap dinucleotide. TGS1$_{653–853}$ is shown in cartoon representation (light grey) and the residues involved in binding are labelled and highlighted as yellow sticks. The bound adenosine diphosphate (ADP) is surrounded by an $|F_o| - |F_c|$ OMIT map contoured at a level of 3σ (ADP was omitted).

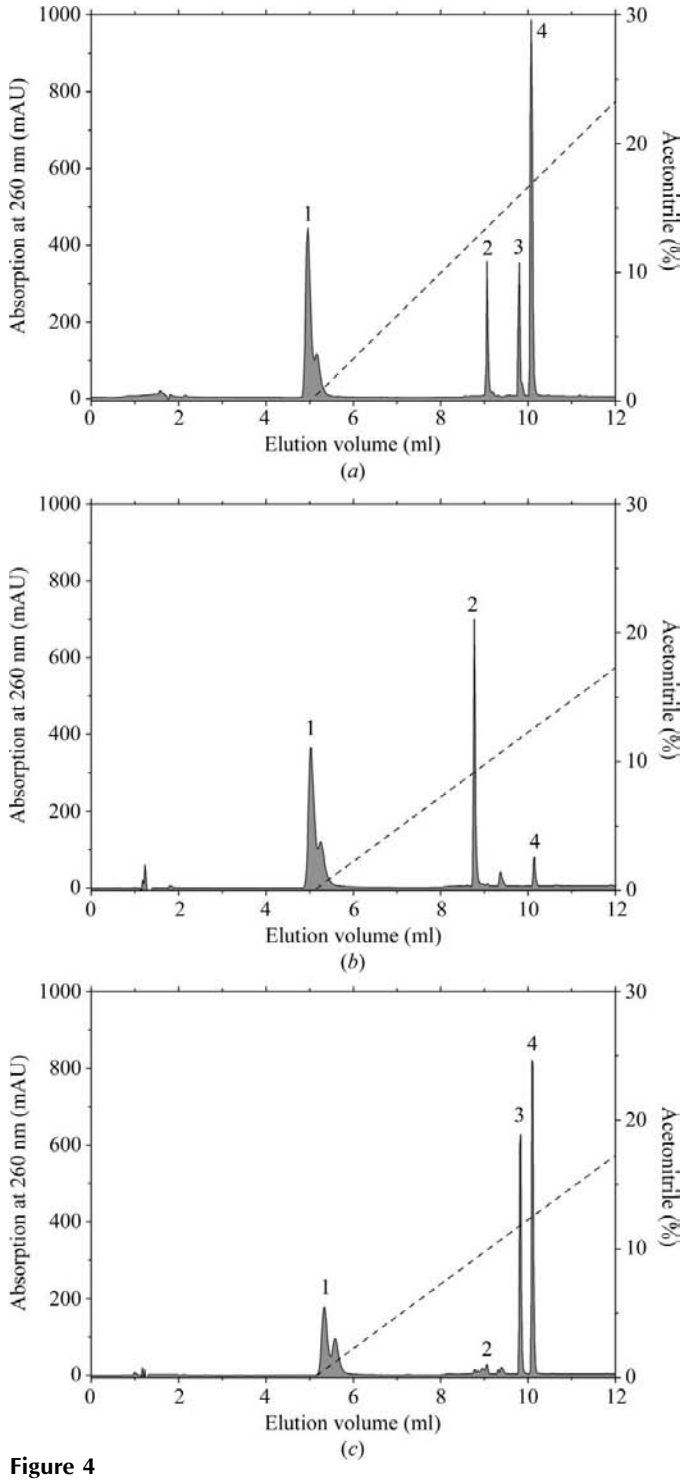

Figure 4
HPLC-based activity test of hTGS1 fragments. (*a*) Calibration run using a mixture of the substrates *S*-adenosyl-L-methionine (AdoMet; 1) and m^{7}GpppA (cap analogue; 2) as well as both reaction products m^{2,2,7}GpppA (3) and *S*-adenosyl-L-homocysteine (AdoHcy; 4) as references. (*b*) Activity test for the crystallized TGS1 comprising residues 653–853. 25 μM purified protein was mixed and incubated with 0.5 mM cap analogue m^{7}GpppA and 2 mM AdoMet for 120 min. (*c*) The 17-amino-acid longer fragment (amino acids 636–853) shows catalytic activity as judged by the utilized substrates after 120 min incubation.

A structural model of yeast TGS1 (yTGS1) has previously been generated using a methyltransferase from *Methanococcus jannaschii* (MJ0882; PDB code 1dus; Huang *et al.*, 2002) as a template structure. This homology model contains yTGS1 residues Met70–Glu267, which correspond to amino acids Leu664–Ala850 in hTGS1 (Mouaikel, Bujnicki *et al.*, 2003). From this model it was predicted that residues Phe60–Cys262 form the minimal globular MTase domain, which in turn correspond to residues Phe655–Leu845 in hTGS1. The m^7G cap was thought to be sandwiched between residues Trp178 (Trp766 in hTGS1) and Trp75 (Trp669 in hTGS1) (Bujnicki & Rychlewski, 2002; Mouaikel, Bujnicki *et al.*, 2003); residues Asp103 and Asp126 (Asp696 and Asp719 in hTGS1, respectively) participate in formation of the AdoMet-binding pocket. Surprisingly, all these residues are present in the crystallized but inactive $TGS1_{653–853}$ fragment, raising questions regarding the molecular basis for the lack of its activity. In particular, the question arises whether the additional 17 residues are sufficient to fulfil the following two functions: to span the distance to the active site and to participate in the formation of a functional substrate-binding pocket.

Bridging the distance of approximately 30 Å from the N-terminus to the catalytic site would require approximately eight amino acids in an extended conformation, assuming a path leading directly through the protein. It is more plausible that the main chain would have to circumvent the protein with its globular shape, thus requiring even more residues. The consideration mentioned above together with the fact that an eight-stranded β-sheet has not been observed to date in any structure of a methyltransferase domain implies that a structural rearrangement of the N-terminal β-strands is more likely to occur. This hypothesis is further supported by secondary-structure predictions, which reveal an α-helical conformation of the region containing β-strands 1–3 (data not shown). An alternative explanation for the difference in catalytic activity between $hTGS1_{636–853}$ and $hTGS1_{653–853}$ is that the additional 17 residues may contribute to or facilitate protein dimerization or oligomerization and therefore lead to activation. However, this possibility can be excluded since both tested hTGS1 constructs exist as monomers in solution as judged by gel-filtration chromatography analyses.

In summary, the structure of the human TGS1 methyltransferase domain presented here corresponds to an inactive truncated domain with a maximum of 17 amino-acid residues missing that would be required to gain catalytic activity. Whether this is achieved by a large structural rearrangement of the N-terminus or simply by reaching the active site within these 17 residues remains to be clarified.

4. Conclusions

The crystallized conserved methyltransferase domain of hTGS1 is catalytically inactive, even though it contains all the residues that have been predicted to be involved in substrate binding and catalysis by means of structural homology modelling. The presence of an additional 17 residues extending the N-terminus leads to an active methyltransferase. Crystal structure analysis reveals that the AdoMet-binding pocket of $hTGS1_{653–853}$ is functional, while no complex with bound m^7G could be obtained. It remains unclear how the additional N-terminal residues complete the active site; it is most likely that they contribute to the binding of the m^7G cap. Hence, further structural and mutational studies are required in order to understand the substrate specificity and catalytic mechanism of hTGS1.

We are very grateful to Rémy Bordonné and Utz Fischer for providing the full-length clone of hTGS1. We thank the staff of EMBL beamline BW7A at DESY (Hamburg) for excellent support during data collection, as well as Stephanie Schell for assistance in protein preparation and crystallization. Furthermore, we are very thankful to Markus Rudolph for help with the activity assay. This work was supported by the DFG (SFB523).

References

Bujnicki, J. M. & Rychlewski, L. (2002). *BMC Bioinformatics*, **3**, 10.
Colau, G., Thiry, M., Leduc, V., Bordonne, R. & Lafontaine, D. L. (2004). *Mol. Cell. Biol.* **24**, 7976–7986.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, California, USA.
Dickmanns, A. & Ficner, R. (2005). *Top. Curr. Genet.* **12**, 179–204.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Enunlu, I., Papai, G., Cserpan, I., Udvardy, A., Jeang, K. T. & Boros, I. (2003). *Biochem. Biophys. Res. Commun.* **309**, 44–51.
Girard, C., Verheggen, C., Neel, H., Cammas, A., Vagner, S., Soret, J., Bertrand, E. & Bordonne, R. (2008). *J. Biol. Chem.* **283**, 2060–2069.
Gunzl, A., Bindereif, A., Ullu, E. & Tschudi, C. (2000). *Nucleic Acids Res.* **28**, 3702–3709.
Hamm, J., Darzynkiewicz, E., Tahara, S. M. & Mattaj, I. W. (1990). *Cell*, **62**, 569–577.
Hausmann, S., Ramirez, A., Schneider, S., Schwer, B. & Shuman, S. (2007). *Nucleic Acids Res.* **35**, 1411–1420.
Hausmann, S. & Shuman, S. (2005*a*). *J. Biol. Chem.* **280**, 4021–4024.
Hausmann, S. & Shuman, S. (2005*b*). *J. Biol. Chem.* **280**, 32101–32106.
Hausmann, S., Zheng, S., Costanzo, M., Brost, R. L., Garcin, D., Boone, C., Shuman, S. & Schwer, B. (2008). *J. Biol. Chem.* **283**, 31706–31718.
Huang, L., Hung, L., Odell, M., Yokota, H., Kim, R. & Kim, S.-H. (2002). *J. Struct. Funct. Genomics*, **2**, 121–127.
Huber, J., Cronshagen, U., Kadokura, M., Marshallsay, C., Wada, T., Sekine, M. & Luhrmann, R. (1998). *EMBO J.* **17**, 4114–4126.
Komonyi, O., Papai, G., Enunlu, I., Muratoglu, S., Pankotai, T., Kopitova, D., Maroy, P., Udvardy, A. & Boros, I. (2005). *J. Biol. Chem.* **280**, 12397–12404.
Limbach, P. A., Crain, P. F. & McCloskey, J. A. (1994). *Nucleic Acids Res.* **22**, 2183–2196.
Mattaj, I. W. (1986). *Cell*, **46**, 905–911.
Maxwell, E. S. & Fournier, M. J. (1995). *Annu. Rev. Biochem.* **64**, 897–934.
Misra, P., Qi, C., Yu, S., Shah, S. H., Cao, W. Q., Rao, M. S., Thimmapaya, B., Zhu, Y. & Reddy, J. K. (2002). *J. Biol. Chem.* **277**, 20011–20019.
Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.
Mouaikel, J., Bujnicki, J. M., Tazi, J. & Bordonne, R. (2003). *Nucleic Acids Res.* **31**, 4899–4909.
Mouaikel, J., Narayanan, U., Verheggen, C., Matera, A. G., Bertrand, E., Tazi, J. & Bordonne, R. (2003). *EMBO Rep.* **4**, 616–622.

Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J. & Bordonne, R. (2002). *Mol. Cell*, **9**, 891–901.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Neuenkirchen, N., Chari, A. & Fischer, U. (2008). *FEBS Lett.* **582**, 1997–2003.
Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
Plessel, G., Fischer, U. & Luhrmann, R. (1994). *Mol. Cell. Biol.* **14**, 4160–4172.
Reuter, K. & Ficner, R. (1999). *Acta Cryst.* D**55**, 888–890.
Ruan, J. P., Ullu, E. & Tschudi, C. (2007). *Mol. Biochem. Parasitol.* **155**, 66–69.
Schubert, H. L., Blumenthal, R. M. & Cheng, X. (2003). *Trends Biochem. Sci.* **28**, 329–335.
Strasser, A., Dickmanns, A., Luhrmann, R. & Ficner, R. (2005). *EMBO J.* **24**, 2235–2243.
Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.
Verheggen, C., Lafontaine, D. L., Samarsky, D., Mouaikel, J., Blanchard, J. M., Bordonne, R. & Bertrand, E. (2002). *EMBO J.* **21**, 2736–2745.
Watkins, N. J., Lemm, I., Ingelfinger, D., Schneider, C., Hossbach, M., Urlaub, H. & Luhrmann, R. (2004). *Mol. Cell*, **16**, 789–798.
Zhu, Y., Qi, C., Cao, W. Q., Yeldandi, A. V., Rao, M. S. & Reddy, J. K. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10380–10385.