THE PROTEIN SOCIETY

The three-dimensional structure of TrmB, a transcriptional regulator of dual function in the hyperthermophilic archaeon *Pyrococcus furiosus* in complex with sucrose

Michael Krug, Sung-Jae Lee, Winfried Boos, Kay Diederichs, and Wolfram Welte*

Department of Biology, University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany

Received 7 December 2012; Revised 6 March 2013; Accepted 6 March 2013
DOI: 10.1002/pro.2263
Published online 10 April 2013 proteinscience.org

Abstract: TrmB is a repressor that binds maltose, maltotriose, and sucrose, as well as other α-glucosides. It recognizes two different operator sequences controlling the TM (Trehalose/Maltose) and the MD (Maltodextrin) operon encoding the respective ABC transporters and sugar-degrading enzymes. Binding of maltose to TrmB abrogates repression of the TM operon but maintains the repression of the MD operon. On the other hand, binding of sucrose abrogates repression of the MD operon but maintains repression of the TM operon. The three-dimensional structure of TrmB in complex with sucrose was solved and refined to a resolution of 3.0 Å. The structure shows the N-terminal DNA binding domain containing a winged-helix-turn-helix (wHTH) domain followed by an amphipathic helix with a coiled-coil motif. The latter promotes dimerization and places the symmetry mates of the putative recognition helix in the wHTH motif about 30 Å apart suggesting a canonical binding to two successive major grooves of duplex palindromic DNA. This suggests that the structure resembles the conformation of TrmB recognizing the pseudopalindromic TM promoter but not the conformation recognizing the nonpalindromic MD promoter.

Keywords: archaea; winged helix-turn-helix; crystal structure; transcriptional repressor; coiled-coil

Introduction

TrmB was first recognized as a sugar-specific repressor controlling the expression of the TM (Trehalose/Maltose) operon encoding the subunits of a maltose/trehalose-specific ABC transporter, as well as a trehalose metabolizing enzyme in the hyperthermophilic Archaeon *Thermococcus litoralis*. *trmB*,[1] the gene encoding the repressor is part of this gene cluster. Presumably, as a result of lateral gene transfer, the same (identical) gene cluster was also found in another hyperthermophilic Archaeon, *Pyrococcus furiosus*.[2] TrmB binds maltose, maltotriose, sucrose, trehalose and glucose.[3] Common to these sugars is the identical structure of the non reducing glucosyl moiety. In vitro approaches

Abbrevations: ABC, ATP-binding cassette; BRE, transcription factor B recognition element; CC, coiled coil; DBD, DNA binding domain; EBD, effector binding domain; EMSA, electrophoretic mobility shift assay; MD, maltodextrin system; TGM, thermococcales glycolytic motif; TM, trehalose-maltose system; Tgr, thermococcales glycolytic regulator; wHTH, winged helix-turn-helix MOTIF; TrmB, transcriptional regulator of the maltose system.

Grant sponsor: DFG; Grant number: WE 962/12-1.
Michael Krug's current address is Soft Matter and Functional Materials, Macromolecular Crystallography, Elektronenspeicherring BESSY II, Albert-Einstein-Straße 15, 12489 Berlin, Germany.
Sung-Jae Lee's current address is Department of Biology, Kyung Hee University, 26 Kyunghee-daero, Dongdaemun-gu, Seoul 130-701, Korea.

*Correspondence to: Wolfram Welte, Department of Biology, University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany. E-mail: wolfram.welte@uni-konstanz.de

[electrophoretic mobility shift assay (EMSA), run off transcription, footprint analysis, mutational analysis] showed that TrmB recognizes the pseudopalindromic sequence ATACTTTTAGTAT overlapping the BRE-TATA box of the TM promoter upstream of the transcription start site. Of all the sugars that are recognized by TrmB only maltose and trehalose release the repressor from the TM promoter and prevent inhibition of transcription of TM operon genes in vitro.[4] However, sucrose and maltotriose maintain binding of the repressor at the TM promoter.[3]

P. furiosus harbors a second gene cluster encoding an ABC transporter specific for maltotriose (maltodextrins) and a maltodextrin metabolizing enzyme, the MD operon.[2,5] This gene cluster lacks an obvious candidate for specific regulation of expression. We found that TrmB is also able to regulate the MD operon.[3] In this case, TrmB is bound downstream of the BRE-TATA box overlapping the transcriptional start site of the operon. The TrmB binding sequence only contains one half (ATACT) of the pseudopalindromic operator sequence seen in the TM promoter. Thus, the mode of transcriptional repression by TrmB is different for the TM and the MD promoter. The affinity for TrmB is higher in the TM than in the MD promoter. In addition, maltose and trehalose, the two sugars that prevent repression by TrmB of the TM promoter do not release repression of the MD promoter. Here, sucrose and maltotriose prevent repression by TrmB of the MD operon and act as inducers. In addition, the inducing effect of maltotriose on the MD operon is prevented in the presence of maltose.[6]

The surprising property of TrmB to recognize two different operators and to be released from represssion by different inducers is corroborated by mutational analysis of TrmB. The exchange of Tyr 50 by Asn abolishes the ability of TrmB to repress the TM operon[6] but not the MD operon. In addition the E87A mutation abolishes repression of the MD operon but has no effect on the repression of the TM operon (S-J Lee, unpublished observation). Thus, the conformations of TrmB in complex with the TM and the MD promoters must be different.

The *Thermococcales T. litoralis*, *P. furiosus*, *P. horikoshii*, *P. abyssi,* and *T. kodakaraensis* encode at least five different TrmB-like proteins. Whereas TrmB itself was only found in *P. furiosus* (where it is identical to the TrmB in *T. litoralis*) and in *P. horikoshii*,[4] in particular, TrmBL1 of *P. furiosus*, the homolog of Tgr (for Thermococcales glycolytic regulator) of *T. kodakaraensis* is noteworthy. Recognizing a certain palindromic motif (TGM),[7] it functions as a repressor for operons encoding glycolytic enzymes and as an activator for operons encoding gluconeogenic enzymes.[8] The putative helix-turn-helix (HTH) motif of these TrmB paralogs shows high sequence conservation.[9] In particular, Tyr at position 50 in TrmB is conserved in all members of the TrmB like proteins which, as we show in this publication, is part of the recognition helix of the wHTH motif that interacts with the TM operator.

The crystal structure of an N-terminally truncated version of TrmB lacking the first 109 amino acids ($TrmB_{\Delta 2\text{-}109}$) was previously determined in complex with maltose indicating its function as a sugar binding domain.[10] Here, we present the structure of full length TrmB with bound sucrose. Because sucrose binds to TrmB[10] and maintains binding of the repressor at the TM promoter[3]the structure of the complex should be similar to that of the structure of TrmB in complex with the pseudopalindromic TM promoter, ATACTTTTAGTAT.

Results

The properties of the protein that was crystallized as full length TrmB

Of the different N- and C-terminal His-tagged versions of TrmB that we tried to purify and to concentrate for crystallization only one was sufficiently soluble (to ~6 mg mL^{-1}) to be crystallized (Materials and Methods section). The protein construct also encoded an N-terminal His tag (MRGSHHHHH HTDP) and a C-terminal extension consisting of VDLQPSLIS. In addition, sequencing revealed that during cloning the mutation Val161Ala had occurred. According to EMSA analysis, the mutant protein exhibited wild type behavior in repressing the TM operon.[6] The addition of sucrose (1 m*M* final concentration) was essential to obtain crystals. No crystals could be obtained in presence of maltose, glucose or in absence of sugar and combinations of these conditions with various DNA sequences containing the identified pseudopalindromic TM operator binding sequence.

Crystal structure of TrmB

TrmB crystallized in space group $P3_221$, with unit cell axes $a = b = 158.5$ Å, $c = 79.2$ Å. The crystals contain one molecule of TrmB in the asymmetric unit corresponding to a solvent content of more than 82% v/v. A high resolution cutoff of 3.0 Å was chosen; $CC_{1/2}$[11] in the highest shell of 0.80 indicates that this cutoff is rather conservative, even though I/σ is only slightly higher than 1. The three-dimensional structure of TrmB was refined giving R_{work} and R_{free} values of 23.1 and 26.3%, respectively (Table I). The high solvent content correlates well with large overall B-factors. There is no electron density for the N-terminal His Tag and the first six amino acids of TrmB. The model extends from Ile 7 to the C-terminal Ser 342 and includes four subsequent residues VDLQ of the artificial C-terminal

Table I. *Crystallographic Statistics*

Spacegroup	$P3_221$
Unit cell dimensions (Å)	158.5, 158.5, 79.2
Total/unique refl.	167,456/23,245
Resolution (Å)	40-2.99 (3.17–2.99)
R_{meas}	0.072 (1.344)
Completeness (%)	99.3 (97.1)
$CC_{1/2}$ in highest shell (reflection pairs)	0.80 (3847)
$\langle I/\sigma_I \rangle$	18.3 (1.5)
Model R/R_{free}-values (%)	23.1/26.3 (46.3/46.7)
R.m.s. deviation from ideal geometry	
Bonds (Å)	0.010
Angles (°)	1.39
Coordinate error (Å)	0.4
B-factors	
Protein (mean value)	128.0
Sucrose (mean value)	111.2
H_2O (four molecules)	95.1
Ramachandran statistics (%)	
Most favored/additional-/generously allowed-/disallowed regions	85.8/11.9/1.3/1.0

The values given in parentheses are for the highest resolution shell. The free *R*-value was calculated from 5% of the data, which were removed at random before the structure was refined. The Ramachandran statistics was calculated with PROCHECK.[31] $CC_{1/2}$ statistics were calculated with phenix.cc_star, which is part of the Phenix package.

extension. The structure (Fig. 1) contains the previously determined C-terminal sugar-binding domain (referred to as EBD henceforth),[10] an N-terminal DNA-binding domain (referred to as DBD henceforth) consisting of alpha helices α1-α4, two β-strands β1-β2 resembling a winged-helix-turn-helix motif (referred to as wHTH henceforth) and helix α5 with a high propensity for coiled-coil formation (referred to as CC henceforth).[13]

The dimeric structure of TrmB

In solution TrmB occurs in a dimeric form.[3] Among the three crystallographic dimers with intermolecular contacts [Fig. 2(a and d)] only the one shown in Fig. 2(a) stands out as a candidate for a dimer in solution. Its formation buries 5971 Å^{2}from solution[16,17] due to apposition of two symmetry mates of amphipathic CC helices [Fig. 2(a–c)]. Hydrophobic amino acid pairs from CC and CC', Phe81/Ile91′, Phe84/Leu 88′, and their symmetry mates thus form the basis of the interaction between two monomers. The structure and the pairwise helix-crossing angle[18] indicate that it is not a typical left-handed 7 helices over 2 turns coiled-coil but rather a slightly right-handed 11 residues over 3 turns one.

The resulting TrmB dimer exhibits a conspicuous overall similarity with known structures of MerR family members[19] despite significant differences in the effector- and the DNA binding domains which preclude a structural superposition.

The sugar-binding site

TrmB was crystallized with bound sucrose. As mentioned in the introduction, transcriptional regulation by TrmB requires that the structure of TrmB in complex with bound sucrose differs from the complex with bound maltose. Only the maltose bound form of the truncated version of TrmB containing the EBD, is known.[10] It was therefore of interest to ascertain how the EBD would change when sucrose is bound to TrmB. The superposition of the two structures binding maltose and sucrose is shown in Fig. 3. Surprisingly, only very small differences within the binding pocket can be seen. The observed conformation in the crystals thus seems to be stable in full-length TrmB with bound sucrose and in the isolated EBD in complex with maltose but further considerations dicussed below lead us to conclude that full-length TrmB with bound maltose must have a different stable conformation.

In the EBD two subdomains can be recognized. The N-terminal subdomain consists of an eight-stranded sheet flanked by two large helices on one side and one large helix Eα3 on the other side (Fig. 5), termed "sugar recognition helix" henceforth. The C-terminal subdomain connected by a short hinge forms a strand, a helix, and an irregular, flattened, seven stranded ß-barrel with its axis parallel to the strands of the N-terminal subdomain. The sugar

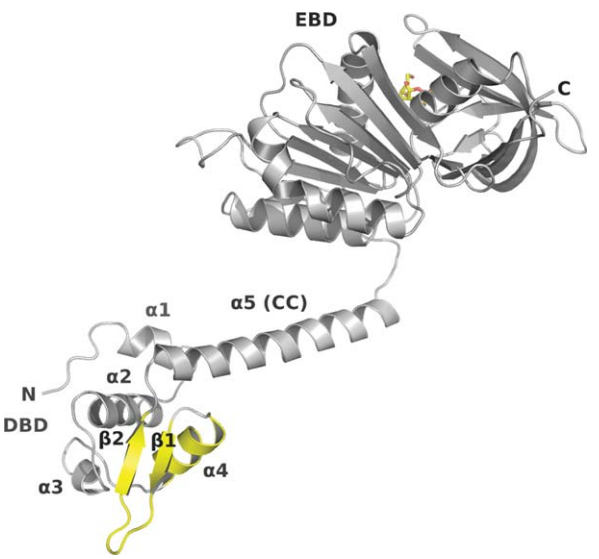

Figure 1. Structure of TrmB in ribbon representation with bound sucrose in yellow wireframe. The N-terminal DBD domain consists of a winged Helix-Turn-Helix domain with helix 4 as the recognition helix. The wing and the recognition helix are colored yellow. Helices (α) and strands (β) are numbered. For comparison with the sequence see Fig. 5. The electron density map for α1 and α3 is weak, indicating high flexibility. The DBD is connected via helix α5 (CC) and a short linker with the sugar-binding EBD domain harboring sucrose (yellow). An interactive view is available in the electronic version of the article.

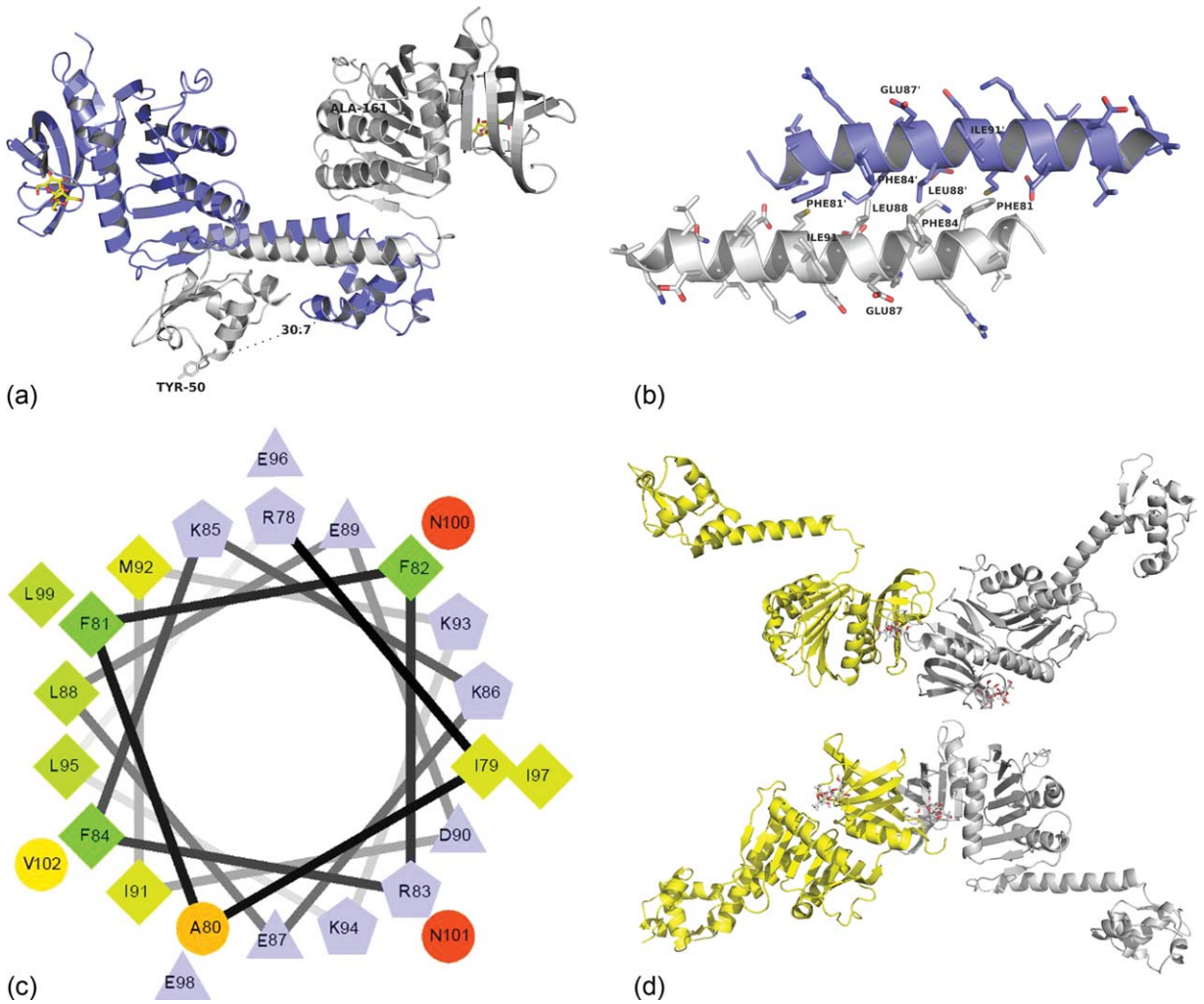

Figure 2. a: Structure of the TrmB dimer in ribbon representation and bound sucrose in yellow wireframe. The structure represents the dimer created by the -X, Y-X, 2/3-Z crystallographic symmetry operation. One monomer is colored grey, the other mauve. The protein presumably dimerizes by forming a coiled coil of the CC helices of the two monomers. The dimer can be considered as a result of domain swapping of the DBDs between two copies of an ancestral protein consisting of the EBD and the DBD with the CC helices as a hinge loop.[14]The distances between the two recognition helices (α4) are indicated. Tyr50 is essential for TM promoter binding but not for MD promoter binding. b: The coiled-coil formed by two crystallographic symmetry mates of CC in ribbon representations with side chains in stick representations. The zipper-like arrangement of hydrophobic residues Phe81/Ile91′, Phe84/Leu 88′, Leu88/Phe 84′, and Ile91/Phe81′ can be seen. c: Helical wheel projection of CC. The diagram was made using the tool by Don Armstrong:[15] http://www.ncbi.nlm.nih.gov/pubmed/12646391. Two packing induced, physiologically irrelevant dimers.

binding site is on the surface of the cleft between the two subdomains (Fig. 1). In the case of bound maltose, six of the seven amino acid residues that are in contact with the sugar ligand are located in the C-terminal subdomain and are holding the nonreducing glucosyl moiety[10] (Fig. 3). The seventh (Ser 229) is in contact with the reducing glucosyl moiety.[10] It is the only residue from the N-terminal subdomain[10] (Fig. 3) and it is located on the sugar recognition helix. In the case of bound sucrose in full length TrmB the nonreducing glucosyl moiety common to both, sucrose and maltose is bound in a different orientation but interacts with the same six amino acids from the C-terminal subdomain (Asn 305, Gly 320, Met 321, Val 324, Ile 325, and Glu 326). The glucosyl ring of sucrose is flipped by about 180° in comparison to the nonreducing glucosyl ring of maltose. In addition, the fructosyl ring of sucrose is tilted about 90° with respect to the reducing glucosyl residue of maltose, as typically seen in crystal structures (Figs. 3 and 4). The density map suggests that this results in a loss of contact of sucrose with the sugar recognition helix from the N-terminal subdomain, but the limited resolution does not allow to draw unequivocal conclusions. The sugar recognition helix is involved in a crystal contact.

DNA recognition of TrmB

The crystal structure of TrmB in complex with sucrose should be similar to the conformation existing

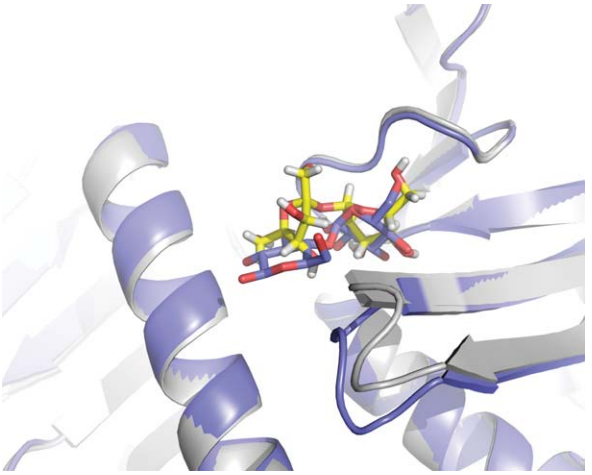

Figure 3. Superposition of the sugar-binding EBD in complex with maltose and sucrose. The sugar-binding EBD domains, $TrmB_{\Delta 2\text{-}109}$ (colored mauve) with bound maltose (carbon atoms in blue, oxygens in red) and full length protein (colored light grey) with bound sucrose (carbon atoms in yellow, oxygens in red) are shown. The non reducing glucosyl moieties of maltose and sucrose are bound by the same amino acid residues. Whereas the reducing glucosyl moiety of maltose interacts with the sugar recognition helix by hydrogen bonding (see Results), this is less clear for the fructosyl moiety of sucrose. The superposition of the structures was done with the software THESEUS[20] using a maximum likelihood approach.

in complex with the pseudopalindromic operator sequence ATACTTTTAGTAT of the TM promoter. Indeed, the 30 Å distance between the two symmetry mates of α4 in the proposed TrmB dimer [See Fig. 2(a)] would allow binding to adjacent major grooves of B-DNA as found for the pair of recognition helices in other dimeric HTH proteins bound to DNA. The orientation of the α4 helices, however, precludes the usual binding of both helices on the same side of the DNA hinting to a conformational change on binding to the TM operator. In accordance with the role of α4 as a recognition helix, the mutation Tyr50Asn within α4 (see Fig. 5) abolishes the ability of TrmB to repress the TM promoter.[6] The ability of TrmB to repress the MD promoter is not affected by this mutation (Lee *et al*., unpublished observation). Thus, α4 cannot interact in the same conformation with the half palindrome DNA sequence of MD as with the TM pseudopalindrome. The conformation of TrmB that recognizes the MD promoter must be different.

Discussion

The exceptional property of TrmB is its ability to bind to two different operator sequences, TM and MD and the differential dependence of its affinity on bound sugars. Sucrose as well as maltotriose, which act as inducers at the MD promoter, nevertheless maintain the binding of TrmB to the TM promoter. In contrast, maltose prevents maltotriose to act as inducer at the MD promoter. Maltose that releases TrmB from the TM promoter maintains binding to the MD promoter. Therefore, binding of maltose and sucrose to TrmB must result in two different arrangements of the two DBDs. The two structures of the EBD, the complex of truncated TrmBΔ2-109 with maltose published earlier by us and the EBD of full length TrmB in complex with sucrose reported here, do not offer an immediate clue for the possible mechanisms promoting these functions.

Effector binding domain EBD

The two structures of the EBD in complex with maltose and sucrose are surprisingly similar (Fig.

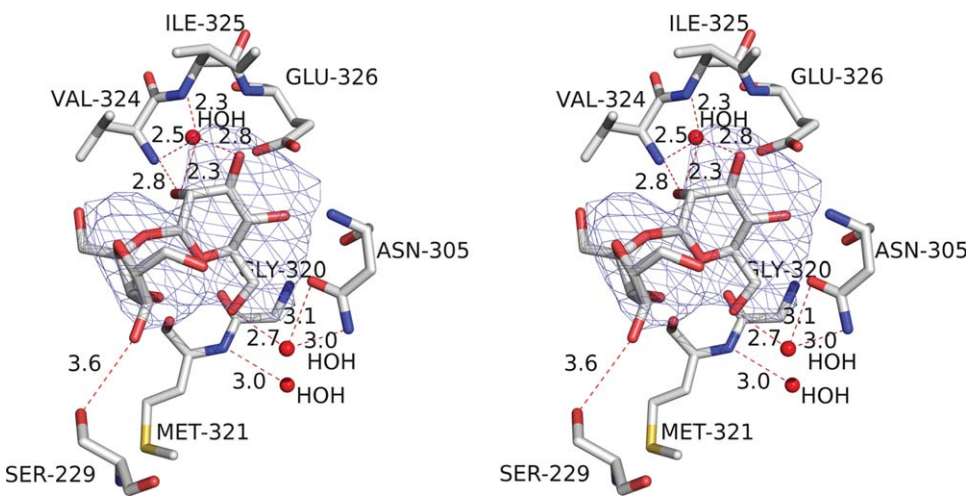

Figure 4. The sugar binding EBD in complex with sucrose in stereo. The TrmB residues forming interactions with sucrose are indicated. Distances of potential hydrogen bonds are given in Å units. The omit electron density map of sucrose is shown at the 5 σ level. For comparison, the interactions of maltose with the EBD of the truncated version of TrmB are shown in Fig. 4 of Krug *et al*.[10]

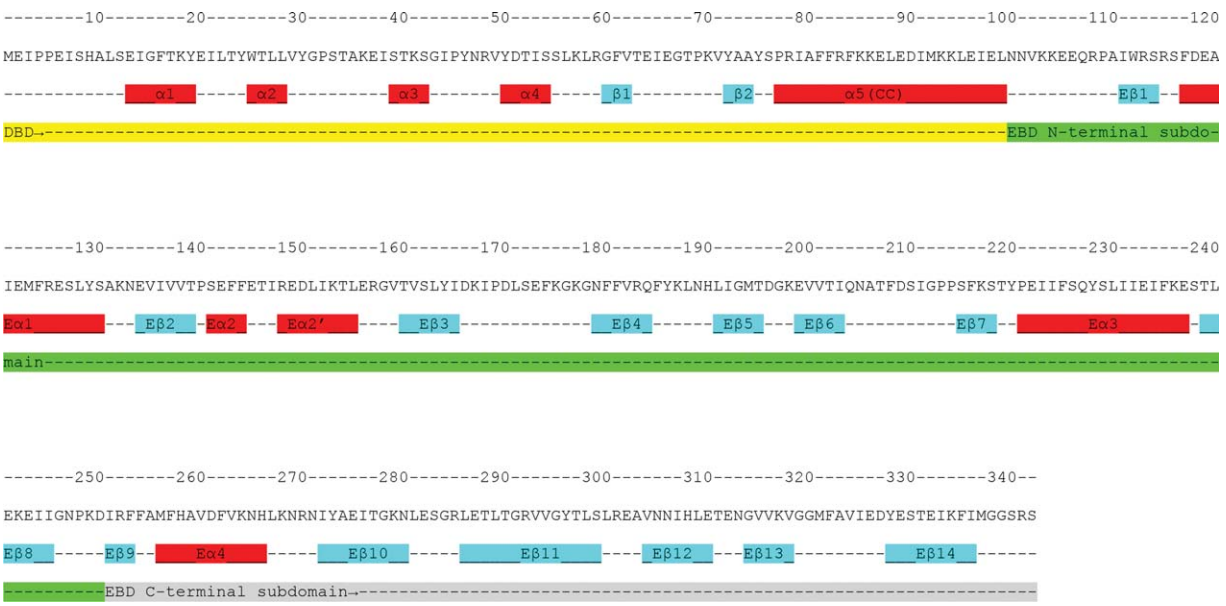

Figure 5. TrmB sequence and secondary structure elements. The α-helices and β-strands are indicated as red and blue bars, respectively. A preceding letter E marks EBD secondary structure elements. The DBD, the EBD N-terminal subdomain and the EBD C-terminal subdomain are underlined by yellow, green, and grey bars, respectively. The Eα3 is the sugar recognition helix.

3). Sucrose which in contrast to maltose is maintaining repression at the TM operon must nevertheless interact differently with the EBD. The low resolution of the structure reported here indicates that the nonreducing glucosyl ring in both structures interacts with the same set of residues of the C-terminal subdomain. The kinked conformation of sucrose arranging the fructose ring perpendicular to the glucose ring will therefore lead to different interactions with the sugar recognition helix as the elongated conformation of the two glucosyl rings of maltose.The fact that all sugars that are bound by TrmB harbor the same nonreducing glucosyl residue suggests that they all are held by the same six amino acid residues of the C-terminal subdomain and that they mediate their differential signals via different interactions with the N-terminal subdomain. Furthermore, binding of maltose to the full length protein in absence of DNA shows cooperativity with a Hill coefficient of two,[3] whereas binding of maltose to the truncated version of TrmB, comprised of only the EBD, shows noncooperative maltose binding. In contrast to the binding of maltose, half maximal binding of sucrose to full length TrmB is noncooperative and occurs at the same concentration as maltose,[3] whereas the binding affinity of maltose to the truncated TrmB is higher.[10]Therefore, the maltose bound structure of full length TrmB and of the truncated protein cannot be identical. This suggests that cooperative binding of maltose to TrmB is linked to a specific interaction with the sugar recognition helix nonexistent in the sucrose bound form and that part of the binding energy is consumed for an allosteric transition affecting the mutual arrangement of the EBD and the DBD. This transition that results in an EBD conformation similar to that seen in the crystal structures must then convert the high affinity for binding to the pseudopalindromic TM operator sequence to a high affinity toward binding the nonpalindromic MD operator sequence.

TrmB possesses a wHTH motif in its DBD

The assignment of α4 as a recognition helix for the pseudopalindromic operator sequence of TM is supported by the observation that the Tyr50Asn mutant lacks transcriptional repression of TM[6] and abolishes the TM operator mass shift seen with EMSA in the wild-type protein.[6] Furthermore, the two short beta strands after the recognition helix suggest that TrmB is a member of the wHTH family[21] with β1 and β2 corresponding to S2 and S3 of the canonical wHTH fold and the loop in between corresponding to wing1 (Fig. 1). The first strand to be expected before α3 would then be missing in TrmB and the expected second wing, W2, would either be very short or absent.

TrmB shares dimerization by a coiled-coil with other DNA binding proteins

The coiled-coil dimerization motif and a DBD with a wHTH motif are found in structures of other proteins: Sto12a from *Sulfolobus tokodaii*, 109 residues,[22] Sso10a from *Sulfolobus solfataricus*, 95 residues,[23] a putative, uncharacterized DNA binding protein from *Archaeoglobus fulgidus* (protein accession code 1SFX), 109 residues, and the MerR family

proteins BmrR[24] and MtaN.[25] Similarly, as in TrmB, the latter two also possess a C-terminal drug binding domain comprising ~150 residues. However, these proteins perform significantly different regulatory functions. TrmB acts as a transcriptional repressor in the DNA bound state and dissociates on binding inducer molecules, whereas the MerR family proteins function as activators for RNA polymerase in complex with coactivator molecules.

Apart from general structural features, TrmB shares no significant sequence similarity and no structural details with the aforementioned proteins. The sequence of secondary structure elements in the DBD for TrmB is α1-α2- α3-α4-β1-β2 with α4 as the recognition helix and wing1 as loop between β1 and β2, whereas for BmrR it is β1-α1-α2-β2-β3-α3-α4. In BmrR α2 is the recognition helix and wing1 is the loop between β2 and β3. A second wing is replaced in BmrR by α3-α4.

TrmB binds to pseudopalindromic (TM) and nonpalindromic (MD) operator DNA. This is conceivable from knowledge of other wHTH containing proteins

In vitro experiments showed that TrmB can repress transcription of both, pseudopalindromic TM and nonpalindromic MD operator sequence.[3,6] In the absence of maltose, the affinity of TrmB for MD is, however, lower than for TM. Using EMSA, the concentration of TrmB to shift the MD promoter maximally (about 2 μ*M*) was found to be four times higher than for the TM promoter,[3] suggesting a K_d for TM binding below 1 μ*M*.

Scanning mutagenesis of the pseudopalindromic TM operator DNA strongly suggests that TrmB binds there as a symmetric dimer with α4 as the recognition helix contacting adjacent major grooves.[3] The conformation of TrmB complexed with sucrose [Fig. 2(a)] should be close to this dimer structure, which may also be similar to the dimer of uncomplexed TrmB observed in solution at room temperature.[3]

The structure reported by us gives no clues as to the quaternary structure and conformation of TrmB in complex with the MD operator. Binding of monomeric as well as dimeric species to palindromic or nonpalindromic DNA have been reported for other HTH containing proteins: SpoIIID[26] is proposed to bind DNA as a monomer. In addition to the HTH, a C-terminal basic region is proposed to contribute contacts with DNA.

OmpR[27] possesses an atypical wHTH motif. One monomer binds to DNA with high affinity and a second monomer can bind subsequently with lower affinity so that a symmetric or asymmetric protein-DNA complex results.

E. coli LexA repressor can bind to half-sequences of its palindromic operator binding sequence, but with 1000-fold lower affinity.[28,29]LexA has the same sequence of secondary structure elements in its wHTH domain, as TrmB.

The P2 C repressor structure was determined in absence of DNA.[30] It is a dimer containing two HTH domains. Model calculations indicate that a 30 bp DNA containing the known nonpalindromic operator sequence repeat could form a complex with two protein dimers so that in each dimer one HTH domain binds the operator sequence and the other binds an asymmetrical DNA sequence.

The observation that the mutant Tyr50Asn (in α4, Fig. 4) abolished transcriptional regulation of TM but not of MD supports the hypothesis that TrmB binds to its two operators in different conformations.

Conclusions

In conclusion, we propose a working model for TrmB-controlled transcription of the TM and the MD operon. In absence of inducer molecules TrmB binds to both operator sequences, because the protein is flexible enough to adopt either of its two DNA binding conformations. Effector molecules like sucrose lock TrmB in its symmetric dimer conformation with high affinity for the pseudopalindromic TM operator sequence. On the other hand, effector molecules like maltose interact with the sugar binding helix in a way leading to allosteric conformational changes in both EBDs. They bind cooperatively and cause a conformation which lacks affinity to TM but promotes affinity to the nonpalindromic MD operator sequence.

Materials and Methods

Expression and Purification of TrmB

PCR was performed using chromosomal DNA of *P. furiosus* as template. The amplification accidentally introduced a Val161Ala mutation, which was not noticed during the initial work. Primers carried at their 5′ ends *Nco*I and *BamH*I sites. The fragment was ligated into pQE30 (Qiagen) containing an N-terminal 6 x His tag. This introduced a N-terminal MRGSHHHHHHTDP and a C-terminal VDLQPSLV extension. The plasmid confers ampicillin resistance, contains the *lacI*^{q} gene and expressed *trmB* under an IPTG-inducible promoter. *E. coli* strain SF120 (defective in multiple proteases) was used as expression host. Cells were grown in NZA medium [10 g NZ-amine (Sheffield Product) 5 g yeast extract, 7.5 g NaCl per liter] at 28°C. At OD 0.8 (600 nm) 0.2 m*M* IPTG was added. Cells were further grown at 37°C for 5 h and harvested by centrifugation. The pellet was resuspended with 10 m*M* 2-(Cyclohexylamino)ethanesulfonic acid (CHES), pH 9.0, 200 m*M* NaCl, 50 m*M* imidazole, 3% 1,4 dioxan (referred to as buffer henceforth). The suspension was passed three

times through a French pressure cell at 11,000 p.s.i. After centrifugation at 50,000*g* for 60 min, the supernatant was heated to 80°C for 20 min and centrifuged at 100,000*g* for 60 min. The supernatant was loaded onto a Histrap column from GE-Healthcare equilibrated with buffer. The column was washed with 20 column volumes of buffer and elution was achieved in three steps with buffer complemented with 100 m*M*, 200 m*M*, and 500 m*M* imidazole, respectively. The protein eluted in the last step. A total of 15 mg pure TrmB was obtained from 8 L of culture. The protein was stored at 4°C.

Crystallization

For crystallization the protein was concentrated using Vivaspin concentrators (Vivascience) equipped with a 30 kDa cut-off membrane. The final protein concentration was between 5.0 and 7.0 mg mL^{-1}. For cocrystallization with sucrose concentrated TrmB, solution was adjusted to a final sucrose concentration of 1 m*M*, incubated at 80°C for 60 min and incubated at 4°C over night.

Screening was done using various commercially available kits of the Nextal Screening Suite (Qiagen) using the sitting drop vapor diffusion method.

Fine screening with hanging drop vapor phase equilibration yielded the best crystals with a reservoir of 0.1*M*, *N*,*N*-Bis(2-hydroxyethyl)glycine (BICINE), pH 9.0, 0.65*M* $(NH_4)_2SO_4$ at 18°C and an initial protein drop of 6 μL obtained by mixing equal volumes of concentrated protein solution and reservoir solution.

Data collection and data analysis

The crystals were flash frozen in liquid propane. A buffer corresponding to the equilibrated crystallization drops was used. About 30% (v/v) glycerol was used as cryoprotectant. Data acquisition was done at the Swiss Light Source (SLS) (beamline X06SA) equipped with the PILATUS 6M detector and were processed using XDS and subsequently merged using XSCALE.[31] The temperature during data collection was set to 100 K.

Crystal structure determination and refinement

Crystallographic phases were determined by molecular replacement using the structure of the sugar-binding domain of TrmB[10] as search model by using the software Phaser,[32] which uses algorithms based on the maximum likelihood method to compute the rotational and translational searches. Model building was done using the graphical model building program COOT[33] and refinement was done using the software PHENIX.[34]

PDB Accession code

Atomic coordinates of TrmB in complex with sucrose are available from the Protein Data Bank under accession code 3QPH.

Acknowledgments

The authors thank Dr. Jutta Nesper for help with molecular biology work. The fruitful discussions with Dr. Michael Thomm and his group from the University of Regensburg are gratefully acknowledged. The authors owe gratitude to Dr. Andrei Lupas from the MPI in Tübingen for help in classifying the coiled-coil and also thank the staff of the SLS beamlines for support.

References

1. Lee S-J, Engelmann A, Horlacher R, Qu Q, Vierke G, Hebbeln C, Thomm M, Boos W (2003) TrmB, a sugar-specific transcriptional regulator of the trehalose/maltose ABC transporter from the hyperthermophilic archaeon *Thermococcus litoralis*. J Biol Chem 278:983–990.
2. Diruggiero J, Dunn D, Maeder DL, Holley-Shanks R, Chatard J, Horlacher R, Robb FT, Boos W, Weiss RB (2000) Evidence of recent lateral gene transfer among hyperthermophilic archaea. Mol Microbiol 38:684–693.
3. Lee S-J, Moulakakis C, Koning SM, Hausner W, Thomm M, Boos W (2005) TrmB, a sugar sensing regulator of ABC transporter genes in *Pyrococcus furiosus* exhibits dual promoter specificity and is controlled by different inducers. Mol Microbiol 57:1797–1807.
4. Lee S-J, Surma M, Hausner W, Thomm M, Boos W (2008) The role of TrmB and TrmB-like transcriptional regulators for sugar transport and metabolism in the hyperthermophilic archaeon *Pyrococcus furiosus*. Arch Microbiol 190:247–256.
5. Koning SM, Konings W, Driessen AJM (2002) Biochemical evidence for the presence of two alpha429 glucoside ABC-transport systems in the hyperthermophilic archaeon *Pyrococcus furiosus*. Archaea 1:19–25.
6. Lee S-J, Surma M, Seitz S, Hausner W, Thomm M, Boos W (2007) Differential signal transduction via TrmB, a sugar sensing transcriptional repressor of *Pyrococcus furiosus*. Mol Microbiol 64:1499–1505.
7. Van de Werken HJ, Verhees CH, Akerboom J, de Vos WM, van der Oost J (2006) Identification of a glycolytic regulon in the archaea *Pyrococcus* and *Thermococcus*. FEMS Microbiol Lett 260:69–76.
8. Kanai T, Akerboom J, Takedomi S, van de Werken HJ, Blombach F, van der Oost J, Murakami T, Atomi H, Imanaka T (2007) A global transcriptional regulator in *Thermococcus kodakaraensis* controls the expression levels of both glycolytic and gluconeogenic enzyme-encoding genes. J Biol Chem 282:33659–33670.
9. Schmid AK, Reiss DJ, Pan M, Koide T, Baliga NS (2009). A single transcription factor regulates evolutionarily diverse but functionally linked metabolic pathways in response to nutrient availability. Mol Syst Biol 5:282.
10. Krug M, Lee S-J, Diederichs K, Boos W, Welte W (2006) Crystal structure of the sugar binding domain of the archaeal transcriptional regulator TrmB. J Biol Chem 281:10976–10982.
11. Karplus PA, Diederichs, K (2012) Linking crystallographic model and data quality. Science 336:1030–1033.
12. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291.

13. Lupas A, Van Dyke M, Stock J (1991) Predicting coiled coils from protein sequences. Science 252:1162–1164.
14. Gronenborn AM (2009) Protein acrobatics in pairs–dimerization via domain swapping. Curr Opin Struct Biol 19:39–49.
15. Zidovetzki R, Rost B, Armstrong DL, Pecht I (2003) Transmembrane domains in the functions of Fc receptors. Biophys Chem 100:555–575.
16. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the CCP4 suite and current developments. Acta Cryst D67: 235–242.
17. Lee B, Richards FM (1971) The interpretation of protein structures: estimation of static accessibility. J Mol Biol 55:379–400.
18. Lupas A, Gruber M (2005) The structure of α-helical coiled coils. Adv Prot Chem 70:37–78.
19. Brown NL, Stoyanov JV, Kidd SP, Hobman JL (2003) The MerR family of transcriptional regulators. FEMS Microbiol Rev 27:145–163.
20. Theobald DL, Wuttke DS (2006) THESEUS: maximum likelihood superpositioning and analysis of macromolecular structures. Bioinformatics 22:2171–2172.
21. Gajiwala KS, Burley SK (2000). Winged helix proteins. Curr Op Struct Bio 10: 110–116.
22. Shinkai A, Sekine S, Urishibata A, Terada T, Shirouzu M, Yokoyama S (2007). The putative DNA-binding protein Sto12a from the thermoacidophilic Archaeon *Sulfolobus tokodaii* contains intrachain and interchain disulfide bonds. J Mol Biol 372: 1293–1304.
23. Chen LC, Chen L-R, Zhou XE, Wang Y, Kahsai MA, Clark AT, Edmondson SP, Liu ZJ, Rose JP, Wang B-C, Meehan EJ, Shriver JW (2004) The hyperthermophile protein Sso10a is a dimer of winged helix DNA-binding domains linked by an antiparallel coiled coil rod. J Mol Biol 341:73–91.
24. Heldwein EE, Brennan RG (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. Nature 409:378–382.
25. Newberry KJ, Brennan RG (2004) The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N terminus. J Biol Chem 279:20356–20362.
26. Himes P, McBryant SJ, Kroos L (2010) Two regions of *Bacillus subtilis* transcription factor SpoIIID allow a monomer to bind DNA. J Bact 192:1596–1606.
27. Rhee JE, Sheng W, Morgan LK, Nolet R, Liao X, Kenney LJ (2008) Amino acids important for DNA recognition by the response regulator OmpR. J Biol Chem 283:8664–8677.
28. Butala M, Žgur-Bertok DW, Busby SJ (2009) The bacterial LexA transcriptional repressor. Cell Mol Life Sci 66:82–93.
29. Zhang APP, Pigli YZ, Rice PA (2010) Structure of the LexA–DNA complex and implications for SOS box measurement. Nature 466:883–886.
30. Massad T, Skaar K, Nilsson H, Damberg B, Henriksson-Peltola P, Haggård-Ljungquist E, Högbom M, Stenmark P (2010) Crystal structure of the P2 C-repressor: a binder of non-palindromic direct DNA repeats. Nucl Acid Res 38:7778–7790.
31. Kabsch W (2010) XDS. Acta Cryst D 66:125–132.
32. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. Acta Cryst D63:32–41.
33. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Cryst D60:2126–2132.
34. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Cryst D58:1948–1954.