doi:10.1016/j.jmb.2010.04.052 J. Mol. Biol. (2010) 400, 71–81

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Structural Basis of TBX5–DNA Recognition: The T-Box Domain in Its DNA-Bound and -Unbound Form

Christian U. Stirnimann[1]†, Denis Ptchelkine[1,2]†, Clemens Grimm[2] and Christoph W. Müller[1]*

[1]*European Molecular Biology Laboratory, Structural and Computational Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany*

[2]*European Molecular Biology Laboratory, Grenoble Outstation, 38042 Grenoble, France*

Received 22 February 2010; received in revised form 21 April 2010; accepted 26 April 2010
Available online 5 May 2010

Edited by K. Morikawa

TBX5, a member of the T-box transcription factor family, plays an important role in heart and limb development. More than 60 single point or deletion mutations of human TBX5 are associated with Holt–Oram syndrome that manifests itself as heart and limb malformations in 1 out of 100,000 live births. The majority of these mutations are located in the TBX5 T-box domain.

We solved the crystal structures of the human TBX5 T-box domain in its DNA-unbound form and in complex with a natural DNA target site allowing for the first time the comparison between unbound and DNA-bound forms. Our analysis identifies a 3_{10}-helix at the C-terminus of the T-box domain as an inducible recognition element, critically required for the interaction with DNA, as it only forms upon DNA binding and is unstructured in the DNA-unbound form.

Using circular dichroism, we characterized the thermal stability of six TBX5 mutants containing single point mutations in the T-box domain (M74V, G80R, W121G, G169R, T223M, and R237W) and compared them with wild-type protein. Mutants G80R and W121G show drastically reduced thermal stability, while the other mutants only show a marginal stability decrease. For all TBX5 mutants, binding affinities to specific and nonspecific DNA sequences were determined using isothermal titration calorimetry. All TBX5 mutants show reduced binding affinities to a specific DNA target site, although to various degrees. Interestingly, all tested TBX5 mutants differ in their ability to bind unspecific DNA, indicating that both sequence-specific and unspecific binding might contribute to the misregulation of target gene expression.

Keywords: Holt–Oram syndrome; TBX5; protein–DNA interaction; crystal structure; DNA binding affinities

Introduction

Holt–Oram syndrome was first described in 1960 and named after its originators Mary Holt and Samuel Oram.[1] It manifests as inborn imperfections, where forelimb and cardiac congenital defects such as atrial and ventricular septal defect are observed. Holt–Oram syndrome is the most common heart–limb syndrome and affects 1 out of 100,000 live births. It is an autosomal-dominant disease that is caused by mutations in the T-box transcription factor *TBX5* gene.[2–4] More than 60 mutations including nonsense, frameshift, splice site, deletion, and missense mutations in the *TBX5* gene have been identified. Most of the mutations are found within the T-box domain ("The human *TBX5* gene mutation database"‡[5]), a DNA binding domain of about 180 amino acid residues that is highly conserved between the members of T-box family transcription

*Corresponding author. E-mail address: Christoph.Mueller@embl.de.

† C.U.S. and D.P. contributed equally to this work.

Present addresses: D. Ptchelkine, Commissariat à l'Energie Atomique, CEA Saclay, Bât.152 SIMOPRO, France; C. Grimm, Institut für Biochemie, Biozentrum der Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany.

Abbreviations used: TBE, T-box binding element; NES, nuclear export signal; ITC, isothermal titration calorimetry; PDB, Protein Data Bank.

‡ www.uni-leipzig.de/~genetik/TBX5

factors. The T-box domain of Brachyury, the founder member of the T-box family, recognizes *in vitro* a 20-bp almost perfect palindromic consensus site [T(G/C)ACACCT/AGGTGTGAAATT, half-site underlined] with high affinity, as identified by SELEX experiments.[6] So far, only half-sites have been identified in natural promoters[7] of T-box target genes and the *in vivo* relevance of the palindromic binding site has therefore been questioned. For TBX5, the consensus sequence was defined as the following 8-bp-long DNA half-site (A/G)GGTGT (C/G/T)(A/G) based on binding studies with synthetic oligonucleotides.[8,9] TBX5 can also bind to the *in vitro* selected palindrome, although binding to the palindrome is not preferred compared to the half-site binding,[9] making it unlikely that TBX5 dimerization is required for its biological activity. In addition to the DNA binding domain, TBX5 possesses a poorly conserved N-terminal extension (residues 1–50). The C-terminal region following the T-box domain harbors a transactivation domain including a nuclear localization signal.[10]

TBX5 is known to interact with promoter regions of several genes and among them the cardiac-specific *natriuretic peptide precursor type A* (*ANF* or *Nppa*) and the *connexin* (*cx40*) genes are best described.[7,11,12] It has been shown that TBX5 activates its target genes synergistically with the homeodomain transcription factor Nkx2-5.[7,11] Interaction studies by electrophoretic mobility shift assay experiments showed that the N-terminal extension and the T-box of TBX5 are essential for interaction with Nkx2-5, whereas the C-terminal domain is not.[11] In addition, TBX5 interacts with GATA4, a Zn-finger transcription factor, which together with Nkx2-5 co-activates *ANF* expression.[13]

The recognition of the T-box DNA binding site by T-box proteins is essential for their function. So far, crystal structures of two complexes between the T-box domains of Brachyury from *Xenopus laevis* (Xbra)[14] and human TBX3[15] and T-box binding sites have been solved and characterized. In both crystal structures, the T-box domain was bound to a palindromic T-box binding element (TBE) derived from the *in vitro* selected consensus sequence.[6] Here, we present the crystal structure of human TBX5 bound to a natural T-box binding site taken from the *ANF* promoter and, for the first time, the DNA-unbound structure of a T-box domain. These high-resolution structures allow us to analyze the conformational changes associated with TBX5 binding to DNA. Our structural analysis is supported by the biophysical characterization of six TBX5 proteins containing single point mutations and the wild-type protein. For wild-type and mutant proteins, we determined the dissociation constants (K_d) for a natural T-box binding site taken from the *ANF* promoter and compared them with the K_d values measured for an unspecific DNA duplex. The characterized mutations significantly influence not only specific but also unspecific binding to DNA. In summary, our structural analyses together with biophysical studies provide molecular insights into the interaction mechanisms of TBX5 with DNA.

Results and Discussion

Overall structure of TBX5–DNA complex

The TBX5–DNA complex (Fig. 1a) solved at 2.2 Å resolution in space group $P6_122$ is the first structure of a T-box protein bound to a natural DNA target site comprising a single half-site. In both previously determined T-box/DNA complex structures, the two T-box domains bound to the *in vitro* selected palindromic binding site (Fig. 1b) interacted through small, non-conserved interfaces and also possess distinctly different quaternary structures. At the interface, the two TBX3 domains are rotated by 10° in opposite directions compared to the Xbra T-box domains.[14,15] The structural consequences of the presence of a half-palindromic site in the TBX5–DNA complex compared to the other two T-box/DNA complexes are discussed below.

The T-box domain of TBX5 is composed of a seven-stranded β-barrel domain that is closed by a smaller β-pleated sheet (Fig. 1a). As expected, the overall structure is very similar to TBX3[15] and the two T-box domains (62% identical residues) can be superimposed with an overall C^α rmsd of 2.1 Å over 174 atoms. While the main structural elements in the TBX5, TBX3, and Xbra scaffolds are similar, one major conformational difference between them is found in the region encompassed by β-strands F and G, where the two TBX3 monomers contact each other (residues 239–245). In TBX3, this region adopts a 3_{10}-helical conformation,[15] while in Xbra, the corresponding region forms an extended loop (residues 173–177) that also contacts a neighboring protein molecule (Fig. 2a).[14] In TBX5, the corresponding region (residues 190–195) is mobile since no electron density is observed for these residues. Hence, this part of the protein is variable in the T-box family (Fig. 2b). The relative orientation of the two monomers in the Xbra and TBX3 structures is constrained by their positions on the DNA bearing the palindromic site, while the TBX5 T-box domain is bound to a natural half-palindromic site taken from the ANF promoter (Fig. 1b). The conformational restrains in the region between strands F and G imposed by monomer–monomer contacts in TBX3 and through the contact with a crystallographically related molecule in the Xbra–DNA complex are absent in the TBX5 structure, therefore allowing greater conformational flexibility of this region.

TBX5 and TBX5–DNA complex are monomeric in solution

The TBX5–DNA complex crystals (space group $P6_122$) contain a second DNA-free TBX5 molecule in the asymmetric unit. Both TBX5 molecules are connected through an intermolecular β-sheet between strand G (amino acids 199–207) of molecule A and the corresponding strand of molecule B related by a non-crystallographic dyad. Both strands are additionally linked *via* an intramolecular disulfide