

residues 274–322. The rest of the protein wraps around the central helix at both the N-terminal and C-terminal ends. The resulting structure has a distinctive question-mark-like appearance (Fig. 1c). The structure closely resembles a lower-resolution Tex structure that was recently submitted to the Protein Data Bank (PDB) by the New York Structural GenomiX Research Consortium (PDB accession code: 2OCE).

Although the Tex structure is fairly compact, it can be largely described as a series of distinct domains. The first 84 residues adopt an HtH structure (Fig. 2), although primary sequence analysis has not previously identified this motif in Tex. The next 189 residues (85–273) wrap around the bottom of the central helix (H15; residues 274–322), which extends nearly the full length of the structure, with a 30° kink occurring at lysine 288. Extending from the C-terminal end of this helix, the rest of the structure forms three domains that had been previously predicted from analysis of the amino acid sequence:⁷ YqgF (residues 329–455), helix–hairpin–helix (HhH) (residues 501–557), and S1 (residues 654–730) domains. The S1 domain is tethered to the rest of the structure by a stretch of 25 residues (629–653) that traverses the top of the molecule and contains little defined secondary structure but is clearly ordered in the electron density.

HtH

Despite very low (<15%) sequence identity, the N-terminal HtH domain overlaps with other HtH structures (e.g., DNA helicase hel308, PDB accession code: 2P6U) with a root-mean-square deviation

(RMSD) of 2.2 Å over 64 C α atoms. HtH motifs typically bind double-stranded DNA (dsDNA) by inserting a helix into the major groove of the DNA duplex.¹⁶ The observed conformation of the HtH region in Tex does not appear to be competent for canonical dsDNA binding because the DNA would significantly clash with the surrounding Tex structure. However, Tex may be capable of binding a single-stranded nucleic acid substrate through this domain. The strongest “hit” from the DALI search (z score = 5.6, RMSD = 2.2 Å) is to the ratchet domain of an archeal DNA helicase, Hel308.¹⁷ Hel308 unwinds dsDNA and uses the third helix of the HtH motif to bind a single strand of the DNA. Superposition of the DNA-bound Hel308 complex (PDB accession code: 2P6R) onto the Tex structure places the single-stranded DNA (ssDNA) through a narrow, elongated cavity ($\sim 6 \times 20$ Å) that passes through the center of the Tex structure (Fig. 3). Although we currently have no direct evidence that Tex binds nucleic acid substrates in this region, there is nothing obvious from the Tex structure that would preclude single-stranded nucleic acid binding at this site in some contexts.

YqgF homologous domain

The central YqgF domain of Tex (residues 329–455) belongs to the YqgFc domain family (SMART SM00732²⁰), a domain described as RNase H-like and typified by the *E. coli* protein YqgF.^{2,7} Little functional or biochemical data are available for YqgF domain-containing proteins, even though these proteins are highly conserved and occur across a wide variety of bacterial genomes.²¹ YqgF domain family

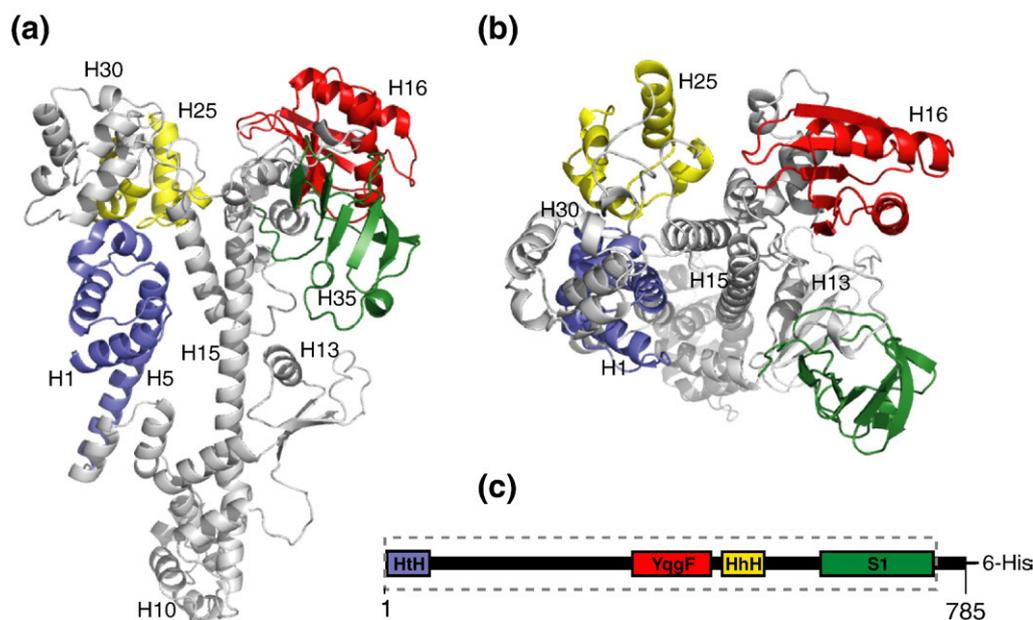


Fig. 2. Tex domain arrangement. (a and b) Orthogonal views of ribbon diagram of the Tex structure. Structural motifs identified from primary sequence and structural analyses are colored. Blue, HtH motif. Red, YqgF homologous domain. Yellow, tandem HhH motif. Green, S1 domain. (c) Tex domain structure, colored as in (a) and (b). The segment boxed with broken lines, comprising residues 1–730, indicates the region of Tex sequence observed in the crystal structure.