Supporting Information

Structure of the pilus assembly protein TadZ from *Eubacterium rectale*: Implications for polar localization

Qingping Xu[1,2], Beat Christen[3], Hsiu-Ju Chiu[1,2], Lukasz Jaroszewski[1,4,5], Heath E. Klock[1,6], Mark W. Knuth[1,6], Mitchell D. Miller[1,2], Marc-André Elsliger[1,7], Ashley M. Deacon[1,2], Adam Godzik[1,4,5], Scott A. Lesley[1,6,7], David H. Figurski[8], Lucy Shapiro[3*], and Ian A. Wilson[1,7**]

[1]Joint Center for Structural Genomics, http://www.jcsg.org
[2]Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
[3]Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305, USA
[4]Center for Research in Biological Systems, University of California, San Diego, La Jolla, CA 92093, USA
[5]Program on Bioinformatics and Systems Biology, Sanford-Burnham Medical Research Institute, La Jolla, CA 92037, USA
[6]Protein Sciences Department, Genomics Institute of the Novartis Research Foundation, San Diego, CA 92121, USA
[7]Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA
[8]Department of Microbiology, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA

Table S1. Primers used for cloning and site-directed mutagenesis

Primer	DNA sequence
ErTadZfw	CTGTACTTCCAGGGCATGAAGATTAAGGTTGCATTGCTTG
ErTadZrv	AATTAAGTCGCGTTATTGCAAGATTTCTTCAAAAAATTCCATC
ErTadZ$_{E162A}$fw	GGAAAAAGGTATTTTATTTAAATATAGCACAGTGTGGCACAACAGATGTTT
ErTadZ$_{E162A}$rv	AAACATCTGTTGTGCCACACTGTGCTATATTTAAATAAAATACCTTTTTCC
pSpeedETfw	GCACCAGATGGGCATTAAACGAGTAT
pSpeedETrv	GATGCCTGGCAGTTCCCTACTCTCG

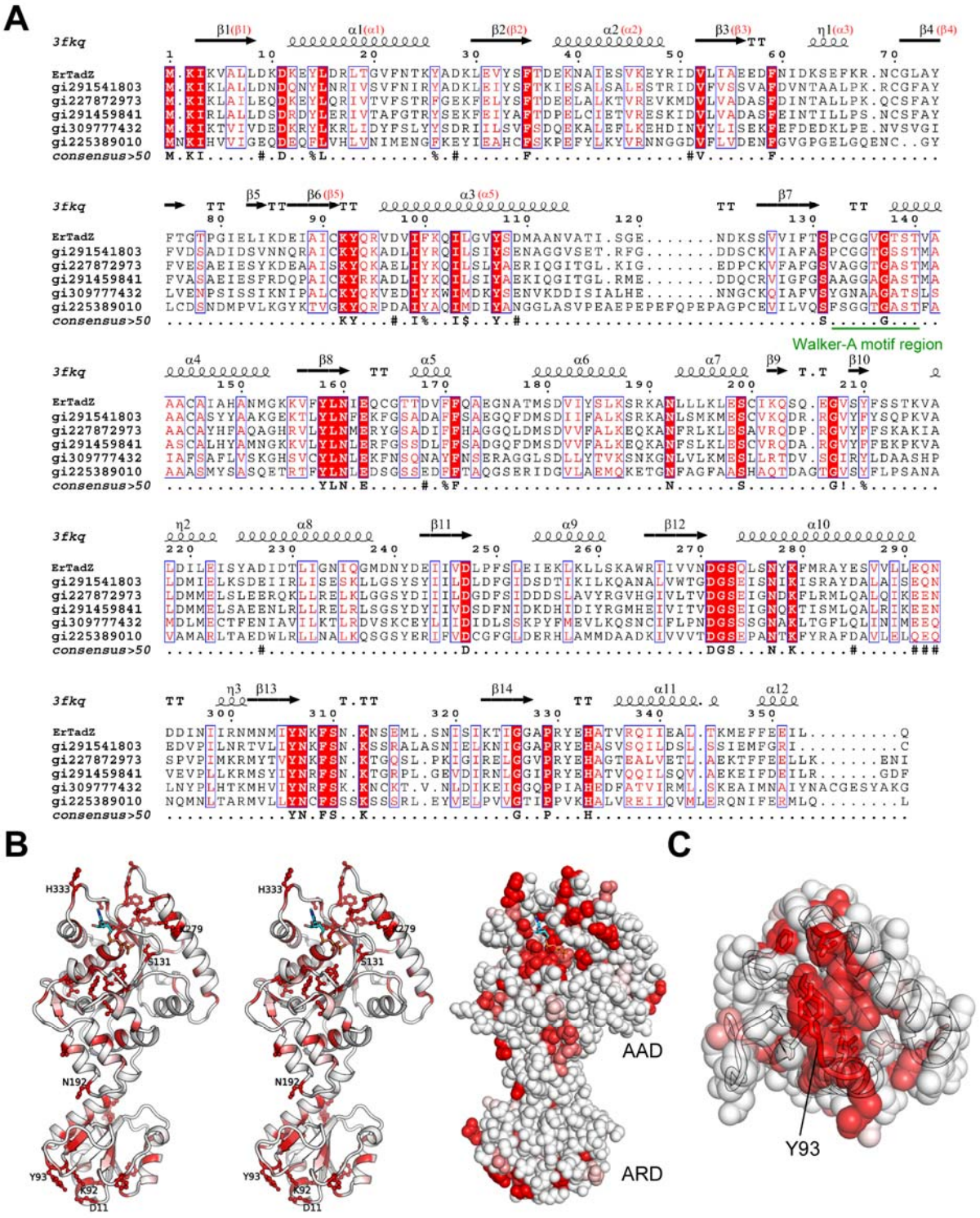

Fig. S1. Sequence conservation of ErTadZ. (**A**) Sequence alignment of ErTadZ (PDB code 3fkq) and its most closely related homologs. The secondary structure elements and sequence numbering of ErTadZ are shown on the top. The corresponding secondary structure elements of a canonical RD are indicated in red in parentheses. Strictly conserved residues are highlighted as white text on a red background. Other highly conserved residues are shown in red text in blue boxes. (**B**) Stereo view of the mapping of sequence conservation onto the structure of ErTadZ. The degree of conservation is shown in a color gradient from red (strictly conserved) to white (not conserved). Strictly conserved residues and ATP (near Ser131) are shown as sticks. (**C**) ErTadZ ARD contains a conserved surface patch near Tyr93. The color gradient is the same as in B.

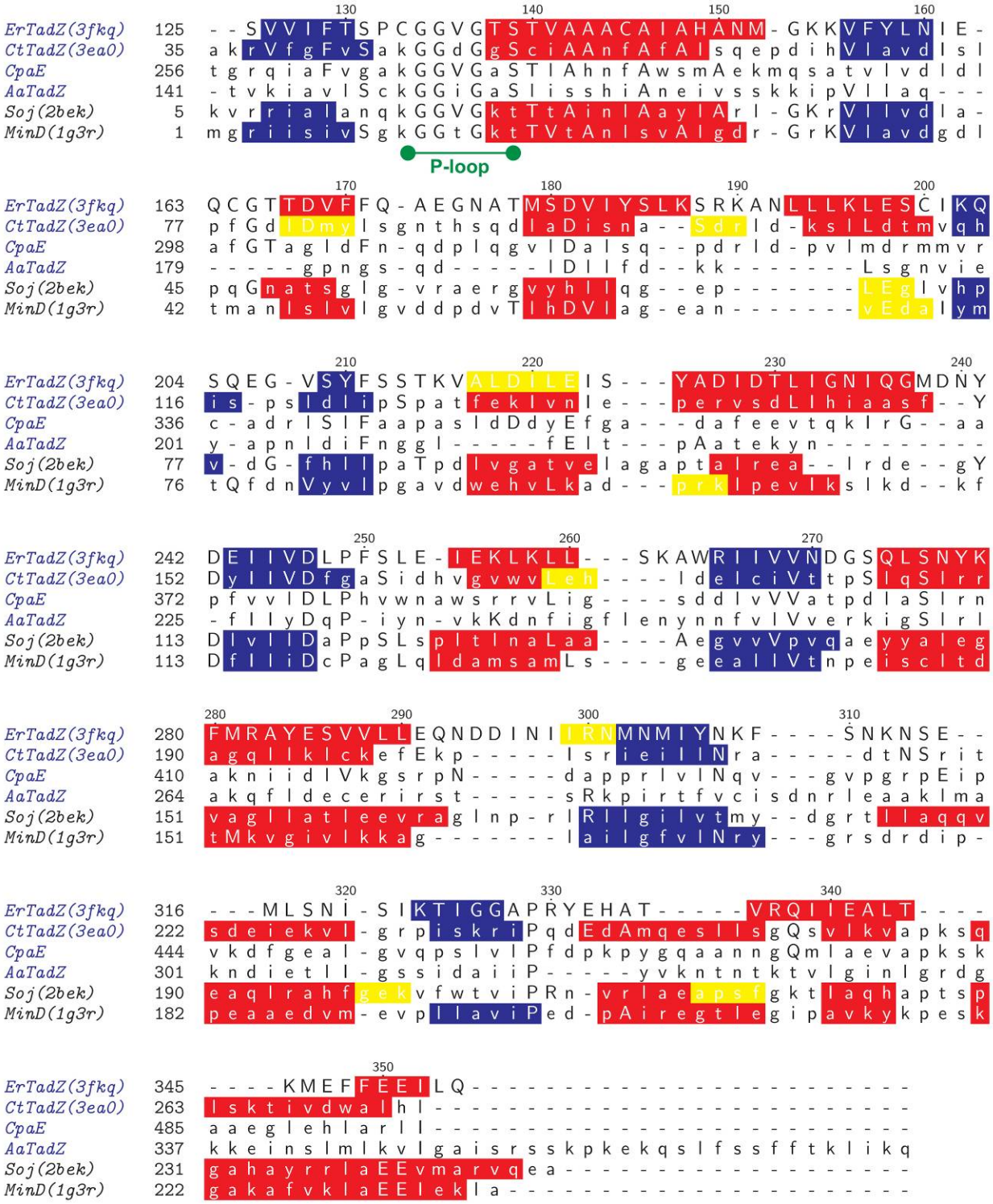

Fig. S2. Multiple sequence alignment of the ATPase domains of ErTadZ (PDB code 3fkq), CtTadZ (PDB code 3ea0), CcTadZ (CpaE), AaTadZ, Soj (PDB code 2bek) and MinD (PDB code 1g3r). The alignment was created by merging a structure-based alignment of the sequences of the four known structures with a sequence-based alignment of CtTadZ, AaTadZ and CcTadZ. The secondary structures elements of proteins with known structures are highlighted (red: α-helix, blue: β-strand, and yellow: 3_{10} helix).

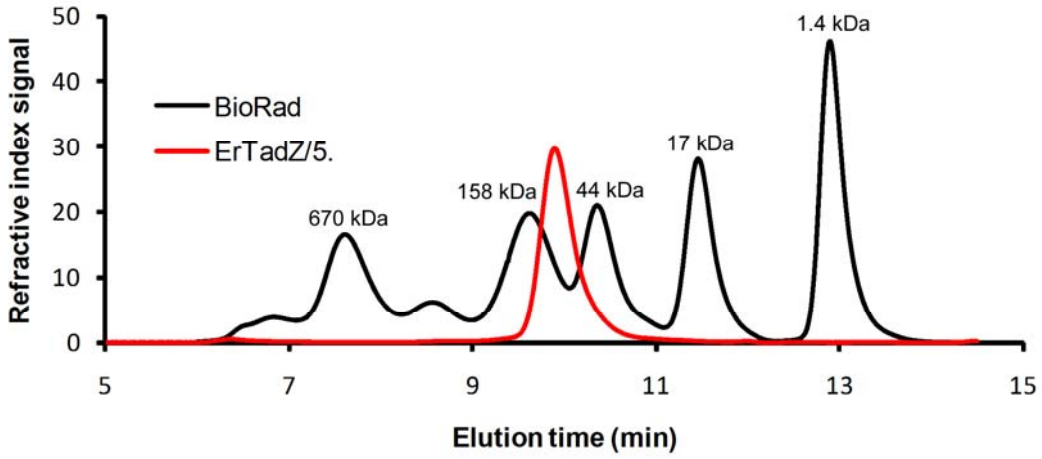

Fig. S3. Analytical size exclusion chromatography of ErTadZ (red line). The molecular weight standards (black line) that were used to estimate the native molecular weight of ErTadZ included thyroglobulin (670 kDa), bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), myoglobin (14 kDa) and vitamin B12 (1.35 kDa). For convenience of display, the refractive index signal for ErTadZ was scaled by a factor of 5.

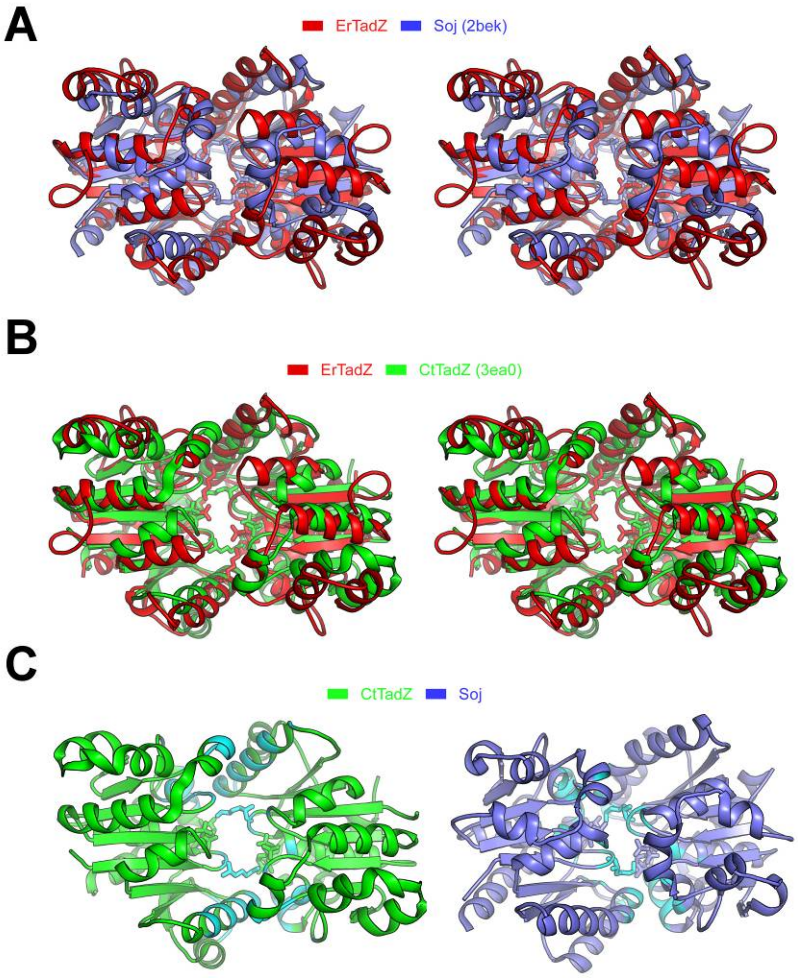

Fig. S4. Comparison of TadZ dimers. (**A**) Stereo view of the superimposition of ErTadZ (red) and Soj (blue), and (**B**) Stereo view of the superimposition of of ErTadZ (red) and CtTadZ (green). (**C**) CtTadZ dimer (green) and Soj dimer (blue) with residues near the domain interfaces highlighted in cyan.

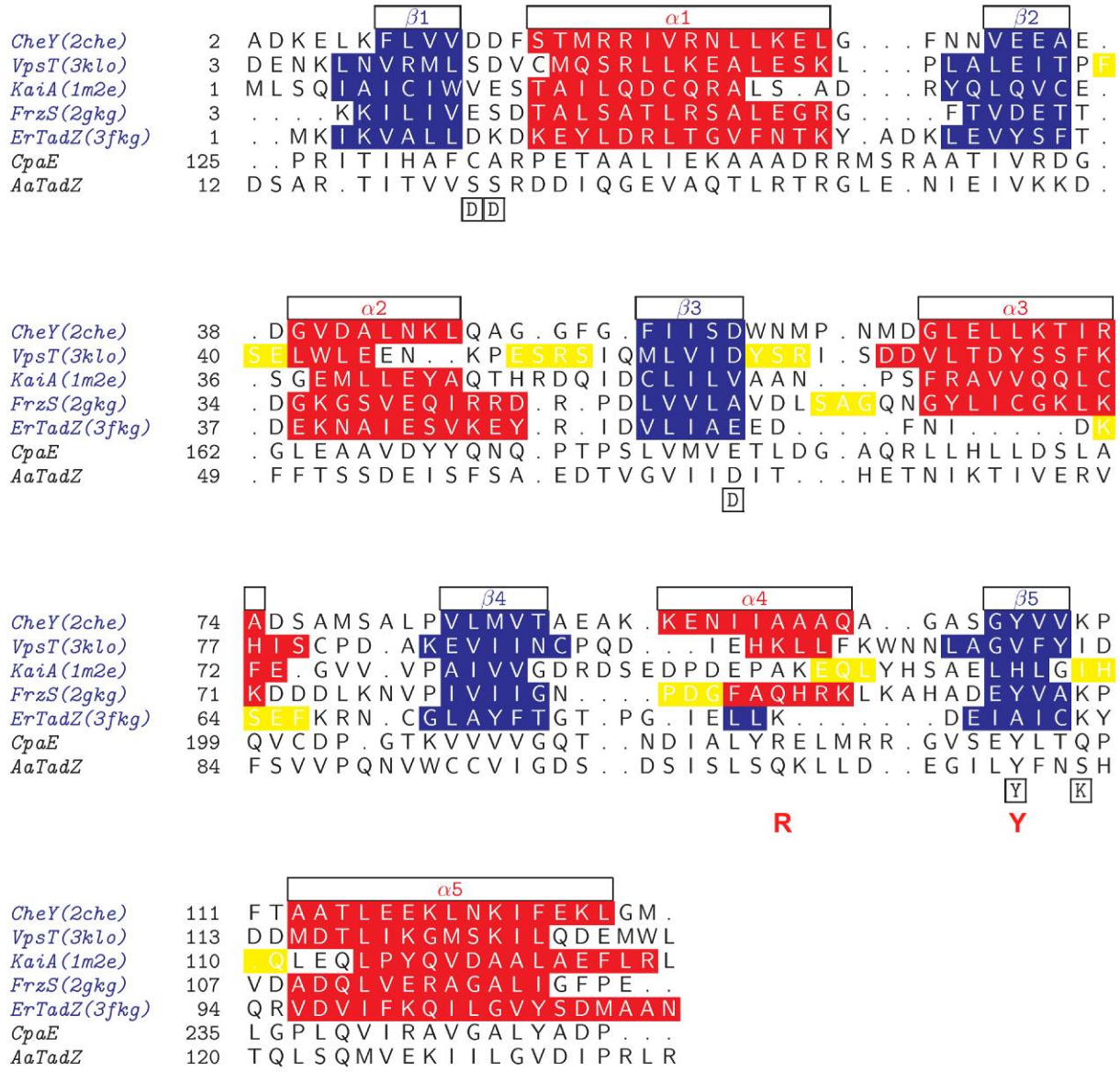

Fig. S5. Multiple sequence alignment of CheY (PDB code 2che) and ARDs of VpsT (PDB code 3klo), KaiA (PDB code 1m2e), FrzS (PDB code 2gkg), ErTadZ (PDB code 3fkg), CcTadZ (CpaE) and AaTadZ. The secondary structure profile of CheY is shown on the top. The secondary structures elements of proteins with known structures are highlighted (red: α-helix, blue: β-strand, and yellow: 3_{10} helix). Functionally critical residues of CheY (letters in boxes) and the two highly conserved residues of CcTadZ (R/Y in bold red) are shown at the bottom. The alignment was created by merging a structure-based alignment of the five sequences with known structures with a sequence-based alignment of ErTadZ, AaTadZ and CcTadZ.