Figure S1 Phylogeny of A43/B16/C25 protein family. Archaeal E' sequences are used as the outgroup. Support values from 500 boostrap replicates are indicated next to branches present in greater than 50% of the topologies. Branch lengths for the consensus tree were estimated using PAML. Members of each subunit protein family as defined by this phylogeny are bounded by colored boxes: A43, pink; B16, yellow; C25, blue. The giardial A43 homologue is highlighted in bold lettering. The full organism name and the respective GenBank identifier for each sequence are listed in Table S2. The scale bar represents 10 amino acid replacements per 100 positions.

Figure S2 Phylogeny of A12.2/B12.6/C11/TFIIS protein family with giardial Zrd proteins. Archaeal TFS/RpoM sequences are used as the outgroup. Support values for the class V consensus tree (see Methods) are indicated next to each branch present in greater than 50% of the significant topologies. Branch lengths of the consensus tree were estimated using PAML. The scale bare represents 10 amino acid changes per 100 positions. The three giardial Zrd proteins are highlighted in bold lettering. Abbreviated sequence identifiers are: *Arabidopsis thaliana*, *Drosophila melanogaster*, *Giardia lamblia*, *Homo sapiens*, *Methanococcus jannaschii*, *Oryza sativa*, *Plasmodium falciparum*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Sulfolobus solfataricus*.

Figure S3 The TATA-binding protein (TBP) sequences from selected Archaea and Eucarya are aligned to show the similarities of the first and second copies of the direct repeat sequence. Consensus amino acids are shown for each alignment column with at most 3 exceptions to a single amino acid, or at most 2 exceptions to two amino acids. The exceptions to the consensus are highlighted in cyan. The conserved phenylalanine residues that intercalate into the DNA helix are also marked with asterisks on the outside of the alignment. Complete names of the taxa included are: *Methanococcus jannaschii*, *Archaeoglobus fulgidus*, *Methanobacterium thermoautotrophicum*, *Pyrococcus woesei*, *Sulfolobus solfataricus*, *Aeropyrum pernix*, *Arabidopsis thaliana*, *Schizosaccharomyces pombe*, *Homo sapiens*, *Drosophila melanogaster*, *Tetrahymena thermophila*, *Dictyostelium discoideum*, *Nosema locustae*, *Entamoeba histolytica*, *Guillardia theta* nucleomorph, *Acanthamoeba castellanii*, *Plasmodium falciparum*, and *Giardia lamblia*. For some sequences, amino-terminal and/or carboxy-terminal residues have been omitted (indicated by ellipsis).