RNA transcribed from clustered regularly interspaced short palindromic repeats (CRISPRs) protects many prokaryotes from invasion by foreign DNA such as viruses, conjugative plasmids, and transposable elements. Cas3 (CRISPR-associated protein 3) is essential for this CRISPR protection and is thought to mediate cleavage of the foreign DNA through its N-terminal histidine-aspartate (HD) domain. We report here the 1.8 Å crystal structure of the HD domain of Cas3 from *Thermus thermophilus* HB8. Structural and biochemical studies predict that this enzyme binds two metal ions at its active site. We also demonstrate that the single-stranded DNA endonuclease activity of this *T. thermophilus* domain is activated not by magnesium but by transition metal ions such as manganese and nickel. Structure-guided mutagenesis confirms the importance of the metal-binding residues for the nuclease activity and identifies other active site residues. Overall, these results provide a framework for understanding the role of Cas3 in the CRISPR system.

Clustered regularly interspaced short palindromic repeat (CRISPR) loci are composed of short DNA repeat sequences separated by stretches of variable spacer sequences (1) that are derived from viral and plasmid DNA (2–5). CRISPR loci are located near clusters of CRISPR-associated (*cas*) genes that, together with the RNA transcribed from the CRISPR loci, mediate the resistance pathway (6, 7). Although *cas* gene clusters are extremely diverse in both gene number, sequence, and organization, they can be grouped into three types of CRISPR/Cas system (6). *cas3* is the defining gene of the type I system, which is also the most widespread of the systems, being found in ~95% of all genomes containing *cas* genes (6).

CRISPR/Cas systems operate in three stages: adaptation, expression, and interference. During adaptation, foreign DNA is recognized, processed, and integrated into CRISPR loci as new spacer sequences. In the expression stage, CRISPR loci are transcribed and processed into small RNAs (crRNA). In the final interference stage, crRNAs serve as the guide sequences in the silencing of target nucleic acid.

In the type I system, the cascade (CRISPR-associated complex for antiviral defense) catalyzes the processing of CRISPR transcripts into crRNA (7). Cascade is a 405-kDa complex comprised of five functionally essential Cas proteins (Cse1 (CRISPR subtype *Escherichia coli*) 1, Cse2, Cas5, Cas6, and Cas7) (6–8). Following processing, the mature crRNA remains associated with cascade (7) and targets this complex to bind foreign dsDNA. Upon binding, cascade melts the dsDNA, allowing crRNA to form Watson-Crick base pairs with the complementary strand, whereas the noncomplementary strand is exposed as single-stranded DNA (ssDNA) (8). In the type II (9) and type III (10) systems, the ultimate fate of the targeted nucleic acid is cleavage by Cas nucleases. In the type I system, cascade lacks any detectable DNase activity (8). However, ssDNA endonuclease activity has recently been described for *Streptococcus thermophilus* cas3 (11). In *vivo* studies have shown that *Cas3* is indispensable for the phage-resistant phenotype of *E. coli* (7). Moreover, the nuclease activity of *Cas3* has been linked to the interference stage of the CRISPR/Cas response and is essential for CRISPR-dependent inhibition of biofilm formation in *Pseudoomonas aeruginosa* (12). Thus, a model has been proposed whereby *Cas3* functions by cleaving the ssDNA region exposed by cascade upon target DNA binding (8, 11).

In most cases, *Cas3* proteins are composed of an N-terminal histidine-aspartate (HD) domain, followed by a C-terminal superfamily 2 helicase domain (13, 14). In some genomes, the HD and helicase domains are encoded as separate subunits and are then named Cas3" and Cas3′ separately (6). Throughout the manuscript, we will refer to the HD domain of *Cas3* as *Cas3*<sub>HD</sub> and in accordance with Ref. 6, the HD domain subunit will be referred to as Cas3". The HD domain catalyzes the ssDNA endonuclease activity of *Cas3*, and the helicase domain catalyzes the ATP-dependent unwinding of dsDNA and RNA-DNA duplexes (11). In contrast, *Sulfolobus solfataricus* Cas3" has distinct substrate specificity as it cleaves double-stranded but not single-stranded DNA or RNA (15).

The HD domain superfamily of metal-dependent phosphohydrolases catalyzes both phosphomonoesterase and phosphodiesterase reactions for a broad range of substrates (16). All HD domains contain three highly conserved motifs (motifs I, II, and V) (16). Motif II contains the signature...
histidine-aspartate sequence, and motifs I and V contain conserved histidine and aspartate residues, respectively. Crystal structures of several HD domains reveal a canonical divalent metal ion-binding site formed by these residues (17–19). Two additional motifs, containing conserved histidine residues (motifs III and IV), are found in some superfamily members, including Cas3 and Cas3° (16). The function of these additional motifs is unknown but likely involves metal ion binding.

As part of our investigation into the molecular basis of the CRISPR system, we undertook crystallographic and biochemical studies of the HD domain of Thermus thermophilus Cas3. We report here the crystal structure of this domain alone and with bound Ni2+. We show that this domain is a metal-dependent ssDNA endonuclease, active in the presence of transition metals but not Mg2+ or Ca2+, and that two metal ions are likely bound at the active site. We also characterize potential active site residues through structure-guided mutagenesis experiments.

**EXPERIMENTAL PROCEDURES**

**Cloning and Mutagenesis**—The 780-bp sequence encoding Cas3HDdom was amplified from the *Thermus thermophilus* HB8 cas3 gene (TTHB187) and cloned into the pHAT2 expression vector (20) between its Ncol and EcoRI restriction sites. The resulting plasmid encodes the Cas3HDdom polypeptide fused to an N-terminal His6 tag. Alanine mutations were introduced into the cas3HDdom gene by the QuikChange site-directed mutagenesis method (Stratagene). All mutations were verified by DNA sequencing.

**Expression and Purification**—Wild-type and mutant pHAT2-cas3HDdom plasmids were transformed into the T7 EXPRESS strain of *E. coli* (New England Biolabs). The cells were grown at 37 °C in Luria-Bertani medium to an A600 of ~0.4. Expression was induced by the addition of 0.2 mM isopropyl 1-thio-β-D-galactopyranoside. Following overnight incubation at 20 °C, the cells were harvested by centrifugation. Cell pellets were lysed in lysis buffer (20 mM Tris-HCl, pH 8.0, 1 mM NaCl, and 10% glycerol) and then clarified by centrifugation at 18,000 rpm for 30 min. The lysates were loaded on a 5-ml immobilized S200 column (GE Healthcare) pre-equilibrated with gel-filtration buffer lacking EDTA, and concentrated to 8 mg/ml using an Ultracel 10 K centrifugal filter unit (Millipore). Purified proteins were >95% pure as judged by SDS-PAGE and Coomassie staining.

**Crystallization**—Crystals of Cas3HDdom were obtained using the sitting-drop vapor-diffusion method by mixing 2 μl of Cas3HDdom at ~8 mg/ml with 1 μl of solution containing 10 mM phosphate–citrate buffer at pH 4.2 and 10% PEG 300. Crystals were stabilized and cryo-protected in a solution containing 10 mM sodium acetate, pH 4.2, and 30% PEG 300, and then flash-frozen in liquid nitrogen. Platinum and osmium derivative crystals were obtained by soaking crystals for 90 min in a stabilization solution containing either 1 mM of K2PtCl4 or (NH4)2OsCl6.

**Data Collection and Structure Determination**—X-ray diffraction data were collected at beamline 9.2 at the Stanford Synchrotron Radiation Light Source and processed with either XDS (21) or HKL2000 (22). SHELX (23) was used to find the positions of osmium sites in the (NH4)2OsCl6 derivative crystal first. The phases derived from these osmium sites were then used to calculate a difference Fourier map to find the heavy atom positions in the platinum derivative. All phases were calculated using SOLVE and improved by solvent-flattening in RESOLVE (24). Model building was carried out in Coot (25), and the model was refined with PHENIX (26). Coordinates of the metal-free and metal-bound structures have been deposited with the Protein Data Bank (PDB) codes 3SK9 and 3SKD, respectively.

**Analysis of Methanococcus jannaschii Cas3**—To access the structure of *M. jannaschii* Cas3°, we calculated difference electron density maps (2.3 Å resolution) from the deposited coordinates and structure factors (PDB code 3M5F) using the phenix.model_vs_data script (27). The Rwork/Rfree from this calculation is 22.9/26.3%, in good agreement with the published Rwork/Rfree of 23.1/26.0%.

**Nuclease Assay**—Cas3HDdom nuclease assays were performed as described previously (11). Magnesium chloride (Mg2+), manganese chloride (Mn2+), nickel sulfate (Ni2+), copper chloride (Cu2+), cobalt chloride (Co2+), calcium chloride (Ca2+), and zinc chloride (Zn2+) were used in metal ion substitution reactions at the indicated concentrations. All reactions were terminated with 20 mM EDTA. The products of reactions were separated by electrophoresis through 1% agarose gels and visualized by ethidium bromide staining.

**Thermofluor Assay**—The apparent melting temperature values of wild-type and mutant cas3HDdom were determined as described previously (28). Experiments were performed in a buffer containing 20 mM Tris-HCl, pH 8.0, and 200 mM NaCl. The final concentration of protein was 10 μM (ε = 55460 M⁻¹ cm⁻¹). Reactions were heated from 20 to 80 °C, and the fluorescence intensity was recorded at 0.2 °C intervals. Fluorescence intensities were plotted as a function of temperature, and the midpoint of the unfolding transition taken as an estimation of the melting temperature.

**RESULTS**

**Crystal Structure of Cas3HDdom**—Sequence analysis predicts that the Cas3HDdom spans residues 5–260 of *T. thermophilus* HB8 cas3. The DNA sequence encoding this region was cloned into the pHAT2 expression vector (20), expressed in *E. coli*, and purified to homogeneity by affinity and size-exclusion chromatography. Cas3HDdom was crystallized by vapor diffusion with PEG 300 as the precipitant. Divalent metal ions inhibited crystallization and therefore were not included in crystallization experiments. Crystals belong to the space group P412121 (a = b = 48.5 Å and c = 205.9 Å) and contain one Cas3HDdom molecule per asymmetric unit. The structure was determined by multiple isomorphous replacement and refined at 1.8 Å resolution to an Rwork of 16.8% and an Rfree of 19.5%. A representative section of...
Cas3 HD Domain Structure

![Diagram of Cas3 HD Domain Structure]

FIGURE 1. Crystal structure of *T. thermophilus* Cas3<sup>HD</sup>dom. A, unbiased $F_o - F_e$ electron density map contoured at 3σ. The residues, which are represented as sticks, were omitted from the map calculation. B, ribbon trace of the Cas3<sup>HD</sup>dom structure. The HD domain motifs are colored as follows: motif I (red), motif II (green), motif III (orange), motif IV (blue) and motif V (pink). The black sphere represents the cognate metal binding site. The N and C termini are labeled. C, Amino acid sequence conservation scores mapped onto the surface of two orthogonal views of the Cas3<sup>HD</sup>dom using CONSURF (41). The structure to the left is in the same orientation as shown in B. The conservation scale is drawn below the two views of the structure. The black sphere represents the cognate metal-binding site. D, Cas3<sup>HD</sup>dom metal ion-binding site. Residues colored by element are in the metal ion-bound configuration. Residues colored yellow are in the protein alone configuration. The Ni<sup>2+</sup> (blue) and two water molecules (red) are represented as spheres. Anomalous difference electron density map contoured at 5σ (white mesh) reveals the site of the Ni<sup>2+</sup> ion.

Unbiased electron density is shown in Fig. 1A. The final model displays good geometry with no Ramachandran outliers (Table 1) and contains all of the Cas3<sup>HD</sup>dom sequence except residues 81–101 and 183–188, for which electron density was not interpretable. No divalent metal ions were apparent in the electron density map. The tertiary structure of Cas3<sup>HD</sup>dom, composed of 10 α-helices and two β-strands, is illustrated in Fig. 1B. Overall, Cas3<sup>HD</sup>dom adopts a globular structure with a concave surface formed by the five conserved motifs of the HD superfamily (Fig. 1C) (16).

To investigate the molecular basis of divalent metal ion binding by Cas3<sup>HD</sup>dom, we measured diffraction data from crystals soaked in a stabilization solution containing 100 μM nickel sulfate. Soaks were performed at the crystallization pH of 4.2 because attempts to increase the pH severely reduced the quality of the x-ray diffraction data. Anomalous difference electron density maps, calculated from soaked data, contained a single strong peak (∼25σ) positioned at the canonical metal ion-binding site. The Ni<sup>2+</sup> ion is coordinated, with octahedral geometry, by two water molecules and four conserved residues, within motifs I (His-24), II (His-69 and Asp-70), and V (Asp-205) (Fig. 1D). Superposition of the structures of Cas3<sup>HD</sup>dom with and without bound metal ion results in a root mean square deviation of 0.28 Å over 235 Ca atoms, demonstrating that the structure of Cas3<sup>HD</sup>dom remains largely unchanged upon metal ion binding. The only significant difference is seen at the metal ion-binding site where the side chain of His-69 rotates to coordinate the metal ion (Fig. 1D).

Comparison with Other HD Domain Proteins—As expected, a search of the structural database by DALI (29) shows that Cas3<sup>HD</sup>dom is related to other HD domains, including many unpublished structures deposited by structural genomics initiatives. The most closely related is the unpublished structure of a Cas3<sup>β</sup> protein (Mj0384) from *Methanocaldococcus jannaschii* (PDB code 3M5F). These two proteins share only 14% amino acid identity, yet their structures align with a root mean square deviation of 3.2 Å over 142 Ca atoms (Z score of 7.9) (Fig. 2A). Although overall similar, there are some notable differences between the structures. First, two regions differ in topology. A 13-amino acid loop (residues 176–191) connects the seventh and eighth helices of Cas3<sup>HD</sup>dom, whereas in Cas3<sup>β</sup>, an insertion results in an additional α-helix connecting these helices (residues 168–202) via an alternate path (Fig. 2B). The C terminus of Cas3<sup>HD</sup>dom also ends with a helix-strand-strand arrangement that is absent in the Cas3<sup>β</sup> structure (Fig. 2A). However, this difference may be the result of disorder in the electron density map of Cas3<sup>β</sup>. The last 30 residues of Cas3<sup>β</sup> are not modeled, and the electron density describing the last helix (which contains motif V) lacks clear side chain features. The coordinates of many of the residues in this helix have been truncated in the model. The second difference between Cas3<sup>HD</sup>dom and Cas3<sup>β</sup> is in the configuration of the metal ion binding sites. Two Ca<sup>2+</sup> ions are modeled at the active site of *M. jannaschii* Cas3<sup>β</sup> (Fig. 3A). Strikingly, neither of these metal ions is found at the canonical metal binding site of HD domains (site A in Fig. 3), perhaps due to the disorder observed in the electron density in the region of motif V. Instead, four conserved residues within motifs II, III, and IV coordinate one of the Ca<sup>2+</sup> ions (site B in Fig. 3A). The significance of the position of the second Ca<sup>2+</sup> ion (site C in Fig. 3A) is unclear as it is located 4.0 Å from the nearest protein atom or water molecule.

The Cas3<sup>β</sup> residues interacting with the site B Ca<sup>2+</sup> ion are conserved in Cas3<sup>HD</sup>dom (Fig. 3B), suggesting that Cas3<sup>HD</sup>dom could bind two metal ions. Yet, in the electron density maps generated from our Ni<sup>2+</sup>-soaking experiments, we fail to observe a metal ion at site B. The motif IV histidine residues (His-137 and His-138), which coordinate the site B metal ion in
Cas3 HD Domain Structure

FIGURE 2. Comparison of Cas3 \( \text{HD} \) \( \text{dom} \) with other HD domains. A, structure of \( \text{T. thermophilus} \) Cas3 \( \text{HD} \) \( \text{dom} \) (yellow and red) superimposed on the structure of \( \text{M. jannaschii} \) Cas3 (white). The additional helix-strand-strand element in the structure of Cas3 \( \text{HD} \) \( \text{dom} \) is colored red. The molecules are oriented as shown in B. B, side-by-side view of the structures of \( \text{T. thermophilus} \) Cas3 \( \text{HD} \) \( \text{dom} \) (left) and \( \text{M. jannaschii} \) Cas3 (right). The helical insertion found in \( \text{M. jannaschii} \) Cas3 and its equivalent loop in \( \text{T. thermophilus} \) Cas3 are colored blue. The helices either side of this region are colored green. C, ribbon trace of the minimal core five \( \alpha \)-helices of the HD domain taken from the structure of Cas3 \( \text{HD} \) \( \text{dom} \). Residues from the five HD domain motifs are colored as shown in B. A surface representation of Cas3 \( \text{HD} \) \( \text{dom} \) is shown in the background.

FIGURE 3. Metal ion-binding sites. The residues that form the metal ion-binding sites of \( \text{M. jannaschii} \) Cas3 (A), \( \text{T. thermophilus} \) Cas3 \( \text{HD} \) \( \text{dom} \) (B), and a protein of unknown function (PDB 2PQ7) from a Thermotogales species (C). The text color of residue labels indicates the motif that the residue belongs to as shown in Fig. 1B. Labeled metal ions are shown as green spheres. His-138 of \( \text{T. thermophilus} \) Cas3 \( \text{HD} \) \( \text{dom} \) is modeled in two alternative conformers; for clarity, only one of these conformers is shown.

Cas3, are oriented away from the binding site in the Cas3 \( \text{HD} \) \( \text{dom} \) structure (Fig. 3B). However, these residues are in a loop that is constrained by crystal packing contacts and, as a result, may be unavailable for metal ion binding. We were unable to crystallize Cas3 \( \text{HD} \) \( \text{dom} \) in the presence of divalent metal ions. In solution, Cas3 \( \text{HD} \) \( \text{dom} \) may bind two metal ions, one at site A and the other at site B. At least five structures of HD domains have been determined (PDB codes 2PQ7, 3HCl, 3CCG, 2OGL, and 2O08) that have a metal ion bound at each of these sites (an example of which is shown in Fig. 3C). These five HD domains are of unknown function, but they all contain a conserved histidine residue in motif III and either one or two conserved histidine residues in motif IV.

The availability of the crystal structures of many HD superfamily members permits us to define the minimal fold of the HD domain. An inspection of the overlay of HD domain structures reveals a common core of five \( \alpha \)-helices that, along with their connecting loops, house the five motifs that define the HD superfamily (Fig. 2C) (16). Beyond this core structure, different members of the HD superfamily...
have unique structural elements that presumably help specify the individual functions of each HD domain family.

Nuclease Activity of Cas3HDdom—S. thermophilus Cas3 has been shown to cleave ssDNA but not dsDNA in a Mg\(^{2+}\)-dependent manner. To investigate the activity of the T. thermophilus Cas3HDdom, we incubated various concentrations of the protein in the presence of Mg\(^{2+}\) and ssDNA (M13mp18). The reactions were then analyzed by electrophoresis through agarose gels, and the DNA species was visualized by ethidium bromide staining. Activity of Cas3HDdom on ssDNA in the presence of Mg\(^{2+}\) was detected. We also assayed cleavage of dsDNA (PvuII-linearized pUC19) and again observed no cleavage (Fig. 4).

We also assayed cleavage of circular single-stranded M13mp18, and the indicated amounts of Cas3HDdom were incubated for 2 h at 37 °C. Reaction mixtures with EDTA or no added metal served as controls. All reactions were quenched with 20 mM EDTA, and the products were then resolved by a 1% agarose gel and visualized by ethidium bromide staining. Activity of Cas3HDdom on dsDNA in the presence of Mg\(^{2+}\) was abolished (K73A), suppressed (W102A and S209A), or had little effect of that magnitude (Table 2). As HD domains have unique structural elements that presumably help specify the individual functions of each HD domain family, these results confirm that the activity we observed is not the result of a contaminating protein. The other alanine mutations (K73A, H105A, and S209A) have a predicted to bind metal ions abolished this nuclease activity. This assay can assess ligand binding because ligands that bind more tightly to the folded form of the protein than to the unfolded form are likely to increase the apparent Tm of that protein.

Metal Ion Binding Increases Thermal Stability of Cas3HDdom—To further characterize the interaction between Cas3HDdom and divalent cations, we performed a Thermofluor assay. This assay measures the change in the fluorescence signal of SYPRO orange dye as it interacts with a protein undergoing thermal unfolding. The fluorescence signal of the dye is quenched in an aqueous environment but becomes unquenched when exposed to the hydrophobic core of the protein upon unfolding. The midpoint of the unfolding transition is taken as an approximation of the melting temperature (Tm).

Mutational Analysis of Cas3HDdom—A series of point mutant proteins were generated in which putative active site residues (His-24, His-69, Asp-70, Lys-73, His-105, His-138, His-139, Ser-202, Ser-209, and Asp-205) or the surface-exposed aromatic residues that surround this site (Trp-102 and Phe-253) were replaced with alanine. The location of these mutations in Cas3HDdom is highlighted in Fig. 7A. Alanine mutants were expressed and purified in the same manner as the wild-type protein (Fig. 7B). To ensure that any potential defects observed in nuclease activity could not be attributed to global misfolding and to assess divalent metal ion binding, the apparent Tm of each alanine mutant was determined in the absence or presence of Ni\(^{2+}\) (Table 2). In the absence of Ni\(^{2+}\), nine of the 12 mutant proteins have Tm values similar to or greater than that of wild-type protein, indicating that these mutations did not destabilize the fold of Cas3HDdom. Three of the mutant proteins (K73A, H105A, and S209A) have a Tm lower than wild-type, suggesting that some destabilization did occur. For wild-type Cas3HDdom, the addition of Ni\(^{2+}\) results in an increase in apparent Tm (ΔTm) of 15 °C. The mutation of residues not implicated in metal ion binding (Lys-73, Trp-102, Ser-202, Ser-209, and Phe-253) results in similar or larger ΔTm values, supporting the evidence that these residues do not participate in metal ion binding. With the exception of His-105, the mutation of residues predicted to bind metal ions (His-24, His-69, Asp-70, His-138, His-139, and Asp-210) results in a decreased ΔTm (Table 2), suggesting that, in solution, these residues are involved in metal ion binding.

We next tested each of the alanine mutants for ssDNA endonuclease activity using the M13mp18 phage cleavage assay in the presence of Ni\(^{2+}\) (Fig. 7C). Under these conditions, wild-type Cas3HDdom generated cleavage products that migrated as a tight smear on an agarose gel, whereas the mutation of residues predicted to bind metal ions abolished this nuclease activity. These results confirm that the activity we observed is not the result of a contaminating protein. The other alanine mutations abolished (K73A), suppressed (W102A and S209A), or had little
Cas3 HD Domain Structure

or no affect (S202A and F252A) on the nuclease activity of Cas3HDdom.

DISCUSSION

Cas3 is functionally essential (7, 12) and is the signature gene of the type I CRISPR/Cas system (6). The HD nuclease domain of Cas3 is proposed to cleave the ssDNA revealed upon cascade binding to target DNA (11). Consistent with this hypothesis, we show that T. thermophilus Cas3HDdom cleaves ssDNA but not dsDNA. This result also establishes that the helicase domain of Cas3 does not alter the substrate specificity of its HD domain. Mg2+ activated the endonuclease activities of S. thermophilus Cas3 (11) and S. sulfataricus Cas3 (15). In contrast, the transition metal ions Mn2+, Co2+, Ni2+, and Zn2+ activate the endonuclease activity of T. thermophilus Cas3HDdom but not Mg2+ or Ca2+ (Fig. 7C). It is also noteworthy that T. thermophilus Cas3HDdom appears much more active in the presence of transition metal ions, particularly Ni2+, than S. thermophilus Cas3 is in the presence of Mg2+ (11). More quantitative data will be needed to establish whether this is significant. Which metal ion, or ions, is used in vivo by T. thermophilus Cas3 remains to be determined. We cannot rule out the possibility that the in vitro requirement for transition metal ions could be a sign of a missing cofactor found in the cell. However, in T. thermophilus, the total intracellular concentration of Mn2+, Ni2+, and Zn2+ are ~160, 150, and 550 μM, respectively, whereas the concentration of Co2+ is undetectable (37). Considering that only 20 μM of Mn2+ or Ni2+ is needed to activate Cas3HDdom in vitro, these two ions are the most likely in vivo candidates of the ions studied.

The crystal structure of T. thermophilus Cas3HDdom provides the first view of an HD domain with nuclease activity. Cas3HDdom adopts a globular structure with a concave surface that contains the active site and presumably binds substrate DNA. Comparison of the structure of the Cas3HDdom with bound Ni2+ and other HD domains with bound divalent metal ions suggests that Cas3HDdom binds two metal ions at its active site. In line with this, the mutation of residues predicted to form the metal ion binding sites results in proteins with smaller ΔTm values upon addition of Ni2+ compared with the wild-type protein. The D70A mutant has the smallest ΔTm value, consistent with the observation that among two metal-ion dependent enzymes, the most critical residue for metal binding is often an aspartate (38). His-105 is the only residue predicted to bind metal ion that, when mutated, has a ΔTm value comparable with that of wild-type protein (Table 2). However, this residue most likely is a metal ion ligand as it is highly conserved and coordinates metal binding in structures of other HD domains (PDB codes 2PQ7, 3HC1, 3CCG, 2OGI, and 2OOG).

The metal-binding data and analysis presented here also imply that all HD domain proteins with histidine residues in motifs III and IV will have two metal ions at their active sites. Thus, unlike the HD domains characterized to date (17–19), proteins in this subset of HD domains, which include Cas3 and Cas3’, are likely to utilize a two metal-ion mechanism for catalysis (39, 40). The fact that Cas3HDdom appears to bind two metal ions in the absence of substrate is also somewhat distinct, as it is generally observed that metal ion binding by two-metal-ion dependent enzymes requires the presence of cognate substrate (38).

We used structure-guided mutagenesis to confirm both the importance of metal ion-binding residues to the activity of Cas3HDdom and to examine the role of other residues close to metal ion-binding sites. The mutation of the residues predicted to bind metal ions, including His-105 (Fig. 3B), completely ablates the nuclease activity of Cas3HDdom under the conditions tested. These results, coupled with both our structural analysis and ΔTm data (Table 2), are consistent with two
Cas3 HD Domain Structure

FIGURE 7. Mutational analysis of Cas3$^{\text{HDdom}}$. A, ribbon trace of Cas3$^{\text{HDdom}}$ (white). Residues selected for mutation are represented as sticks (red), and the Ni$^{2+}$ ion as a sphere (blue). B, an SDS-PAGE of the purified mutants, stained with Coomassie Blue. C, ssDNA endonuclease activity of the mutants. Reaction mixtures containing 10 mM Tris-Cl, pH 7.5, 60 mM KCl, 10% glycerol, 100 nM Cas3$^{\text{HDdom}}$, and 60 nM NiSO$_4$ and 4 nM circular single-stranded M13mp18 DNA were incubated for 50 min at 37 °C. Reactions with EDTA or no Cas3$^{\text{HDdom}}$ (−) served as controls. The reaction products were resolved by a 1% agarose gel and visualized by ethidium bromide staining.

TABLE 2

| Melting temperatures of wild-type and mutant Cas3$^{\text{HDdom}}$ |
|-------------------------------------------------------------|
|                  | 0 μM Ni$^{2+}$ | 100 μM Ni$^{2+}$ |
| WT               | 49.6 ± 0.2     | 64.8 ± 0.6       |
| H24A             | 53.6 ± 0.4     | 62.8 ± 0.4       |
| H69A             | 49.8 ± 1.8     | 62.2 ± 0.8       |
| D70A             | 52.6 ± 0.2     | 55.2 ± 0.4       |
| K73A             | 45.0 ± 0.4     | 65.0 ± 0.6       |
| W102A            | 49.0 ± 0.4     | 67.2 ± 0.4       |
| H105A            | 45.8 ± 0.4     | 61.2 ± 0.4       |
| H137A            | 50.8 ± 0.4     | 62.0 ± 0.2       |
| H138A            | 49.2 ± 0.2     | 61.8 ± 0.4       |
| S202A            | 51.2 ± 0.8     | 66.2 ± 0.4       |
| D205A            | 50.6 ± 0.6     | 61.0 ± 0.4       |
| S209A            | 47.2 ± 0.2     | 62.4 ± 0.4       |
| F233A            | 49.6 ± 0.4     | 65.4 ± 0.6       |

metal ions bound at the Cas3$^{\text{HDdom}}$ active site and establish the importance of these ions for nuclease activity. Mutation of the invariant Lys-73 produced an inert enzyme (Fig. 7C). Because of the proximity of this residue to the metal ion binding sites (Fig. 7A) and its positive charge, it is likely that it helps correctly position a phosphate group of the substrate for catalysis. Mutation of Trp-102 or Ser-209 also resulted in a protein with a reduced activity (Fig. 7C). The position of these residues within the substrate-binding cleft (Fig. 7A) suggests that they may play a role in substrate recognition.

Studies of S. solfataricus Cas3 have shown that this enzyme has distinct substrate specificity compared with S. thermophilus and T. thermophilus Cas3, as it cleaves dsDNA but not ssDNA (15). Additionally, comparison of mutational studies between T. thermophilus cas3 (presented here) and S. solfataricus Cas3$^3$ (15) suggests that the active site geometry of S. solfa-
taricus Cas3$^3$ may also be distinct. First, mutation of either His-69 or His-105 ablates Cas3$^{\text{HDdom}}$ nuclease activity (Fig. 7C). However, mutation of the corresponding residues in S. solfataricus Cas3$^3$ results in a protein with near wild-type activity (15). Secondly, mutation of Glu-92 in S. solfataricus Cas3 has a different functional role or mechanism of action within Cas3$^3$ but not Cas3, inactivated nuclease activity. Inspection of the structure of M. jannaschii Cas3$^3$ suggests that this glutamate may replace the second histidine residue of motif IV (His-143 in T. thermophilus and His-124 in M. jannaschii). This histidine forms part of metal ion site B (Fig. 3) and is essential for the nuclease activity of T. thermophilus Cas3$^{\text{HDdom}}$ (Fig. 7C). In the structure of M. jannaschii Cas3$^3$, Glu-92 is in close proximity to site B, but its side chain is oriented away from the metal ion (Fig. 3A).

S. solfataricus Cas3$^3$ lacks the second histidine of motif IV. Thus, Glu-92 could substitute for this histidine in coordinating the metal ion, potentially explaining the importance of this residue for the nuclease activity of S. solfataricus Cas3$^3$. The significance of this mutational data and the difference in substrate specificity awaits further studies of the Cas3 and Cas3$^3$ proteins. However, these results may indicate that S. solfataricus Cas3$^3$ has a different functional role or mechanism of action within the CRISPR response.

Acknowledgments—We thank Gabriel Brandt and Jennifer Kavran for critical reading of the manuscript and Jurgen Bosch for use of a real-time PCR machine.

REFERENCES
1. Nakata, A., Amemura, M., and Makino, K. (1989) J. Bacteriol. 171, 3553–3556
