Acta Biochim Biophys Sin 2026, 58(8): 1719–1735
https://cstr.cn/32197.14.abbs.2026058
https://doi.org/10.3724/abbs.2026058
Original Article

Original Article

Biochemical and structural studies reveal the substrate specificity and catalytic mechanism of MYG1 as a two-metal ion-dependent 3′→5′ exonuclease

Chao Lan[1], Ziyue Chen[1], Guanchao Wang[1], and Jianping Ding[1,2,3,*]

[1]State Key Laboratory of RNA Innovation, Science and Engineering, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China, [2]School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China, and [3]Ruisi Research Center for Life Science, Shanghai 201109, China

*Correspondence address. E-mail: jpding@sibcb.ac.cn

Received 27 February 2026 Accepted 27 March 2026 Published 10 April 2026

Abstract

Nucleases are a class of enzymes that specifically cleave nucleic acids in all living organisms. They play crucial roles in essential biological processes, including the regulation of gene expression, DNA damage repair, and RNA processing and degradation. MYG1 (melanocyte proliferating gene 1) is a highly conserved eukaryotic protein that exhibits 3′→5′ exonuclease activity. This study systematically characterizes the enzymatic properties of MYG1 and determines its structures in complexes with metal ions and various mono- and poly-(deoxy)nucleotides. The functional roles of key residues involved in metal ion binding and substrate binding in the catalytic reaction are examined through site-directed mutagenesis, enzymatic activity assay, and structure determination. Our biochemical and structural data together demonstrate that MYG1 is a Mn^{2+}- or Mg^{2+}-dependent 3′→5′ exonuclease capable of cleaving a variety of nucleic acids with different structures. It exhibits the highest activity for single-stranded RNA and a nucleotide preference for U in single-stranded RNA and dT in single-stranded DNA. Mechanistically, MYG1 functions as a dimer, with the active site formed by the catalytic domain of monomer 1 and the substrate-binding domain of monomer 2, and cleaves nucleic acids through a two-metal ion-mediated catalytic mechanism. These findings establish a molecular basis for further investigations into the biological functions and molecular mechanisms of MYG1 within cells and its potential roles in human diseases.

Key words structure, MYG1, DHH family, substrate specificity, exonuclease, catalytic mechanism

Introduction

Nucleic acids are fundamental macromolecules of life, including DNA and RNA, which together are responsible for the storage and transmission of genetic information. Nucleases are enzymes that specifically recognize and cleave phosphodiester bonds in nucleic acids. Thus, they play a critical role in the metabolism of nucleic acids and are involved in various biological processes, including the regulation of gene expression, DNA damage repair, and RNA processing and degradation [1]. Nucleases can be classified into endonucleases and exonucleases based on their mode of action and deoxyribonucleases and ribonucleases according to their substrate preference [2–4]. They also include specialized categories such as topoisomerases, recombinases, ribozymes, and RNA splicing enzymes. From the perspective of the catalytic mechanism, nucleases can be divided into three major classes: two-metal ion dependent, one-metal ion dependent, and metal ion independent. Nevertheless, within the same mechanistic class, members can exhibit significant differences in three-dimensional structures, metal ion coordination patterns, and biological functions [5]. Nucleases often possess modular domain architectures, achieving efficient recognition and cleavage of specific nucleic acid sequences through the coordinated action of functional domains such as catalytic domains, nucleic acid-binding domains, and recognition domains [5].

DHH family nucleases are a group of evolutionarily highly conserved and functionally diverse nucleases, which are found

© The Author(s) 2026. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/).

widely across bacteria, archaea, and eukaryotes [6,7]. Members of this family share a highly conserved DHH (Asp-His-His) motif within the catalytic domain, which, together with an adjacent substrate-binding domain, forms the active site. The catalytic activity strictly depends on divalent metal ions such as Mn^{2+} or Mg^{2+} and typically employs a two-metal ion-dependent catalytic mechanism. Despite their conserved catalytic core, DHH family nucleases exhibit highly diverse substrate specificity and generally lack strict sequence-specific recognition. As a result, they fulfill a wide range of functional roles and are involved in a variety of biological processes, including DNA repair and recombination, RNA maturation and degradation, stress tolerance, cell signaling, and immune regulation [7–10]. Additionally, dysregulation or dysfunction of DHH family nucleases can result in abnormal RNA processing and genomic instability, leading to their implication in various diseases, including cancers and genetic disorders [11–13].

MYG1 (melanocyte proliferating gene 1) is a newly identified nuclease that possesses a mitochondrial targeting signal and a nuclear localization signal in its N-terminal region, allowing it to localize to both compartments [14,15]. MYG1 homologs are found to exist widely in eukaryotes and are evolutionarily highly conserved [14,15]. Due to the presence of a canonical DHH motif in its N-terminal domain, MYG1 is classified as a member of the DHH family. Biochemical data show that MYG1 possesses 3′→5′ exonuclease activity but lacks endonuclease activity [14]. It exhibits robust activity on single-stranded RNA (ssRNA) and weak activity on single-stranded DNA (ssDNA) but no activity on double-stranded DNA (dsDNA). Additionally, its activity depends on the presence of Mg^{2+}, but no activity is observed in the presence of Mn^{2+}. Functional data demonstrate that in the nucleolus, MYG1 processes pre-ribosomal RNA involved in ribosome assembly and modulates cytoplasmic translation [14]. In the mitochondrial matrix, it processes the 3′-termini of mitochondrial ribosomal and messenger RNAs, thereby regulating the translation of mitochondrial proteins. These findings indicate that MYG1 acts as a coordinator of the nucleo-mitochondrial crosstalk, which is essential for mitochondrial function and cellular respiration. Additionally, transcriptome analyses indicate that MYG1 expression is gradually reduced during spermatogenesis [16] and is significantly higher in embryonic and hematopoietic stem cells than in differentiated hematopoietic cells [17]. Genetic data show that *MYG1*-deficient mice exhibit alterations in stress-induced response [18]. These findings suggest that MYG1 is involved in cell development, differentiation, and the oxidative stress response.

Moreover, dysfunction of MYG1 is closely associated with several human diseases. For example, expression level changes in MYG1 have been implicated in patients with the complex skin disorder vitiligo [19–22]. Additionally, MYG1 is abnormally overexpressed in multiple cancer cell lines and tumor tissues, where it may promote tumorigenesis by enhancing cell proliferation and inhibiting apoptosis [23–25]. Notably, MYG1 expression is upregulated in colorectal cancer (CRC), and highly expressed MYG1 promotes glycolysis and CRC progression independent of its nuclease activity [23].

Despite its evolutionary conservation and functional significance in various biological contexts, the enzymatic properties, structure, substrate specificity, and catalytic mechanism of MYG1 remain largely elusive, and its precise biological function within cells is still unclear. In this study, we systematically characterized the enzymatic properties of MYG1 and determined the crystal structures of MYG1 in complexes with metal ions and various mono- and poly-(deoxy)nucleotides. The functional roles of key residues involved in metal ion binding and substrate binding in the catalytic reaction were examined through site-directed mutagenesis, enzymatic activity assays, and structure determination. Our biochemical and structural data together demonstrate that MYG1 is a Mn^{2+}- or Mg^{2+}-dependent 3′→5′ exonuclease that acts on a wide range of nucleic acids with different structures. It exhibits the highest activity for ssRNA and a nucleotide preference for U in ssRNA and dT in ssDNA. Mechanistically, MYG1 functions as a dimer, with the active site formed by the catalytic domain of monomer 1 and the substrate-binding domain of monomer 2, and cleaves nucleic acids through a two-metal ion-mediated catalytic mechanism. These findings establish a molecular basis for further investigation into its biological functions within cells and its potential roles in human diseases.

Materials and Methods

Cloning, expression, and purification of MYG1

The human *MYG1* gene was amplified via PCR from a cDNA library derived from human cells. A fragment of *MYG1* encoding residues 42–376 was cloned and inserted into the pET28 vector (Novagen, Gibbstown, USA), which includes an N-terminal His_6-SUMO tag. The N-terminal residues 1–41 of MYG1, which contain both mitochondrial and nuclear localization signals, were omitted. The plasmid was transformed into the *Escherichia coli* BL21 (DE3) CodonPlus strain (Weidi Bio, Shanghai, China). Protein overexpression was induced by adding 0.2 mM IPTG at 16°C overnight after the transformed cells grew in LB medium (Oxoid, Basingstoke, UK) supplemented with 0.05 mg/mL kanamycin (Solarbio, Beijing, China) at 37°C until the OD_{600} reached 0.8. Then, the cells were collected by centrifugation and lysed in lysis buffer (30 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM 2-mercaptoethanol, and 1 mM PMSF) using a sonicator. The target protein was purified via affinity chromatography using a Ni-NTA column (Qiagen, Hilden, Germany). The purified protein was incubated overnight with ULP1 protease at 4°C to cleave the His_6-SUMO tag and then purified via size-exclusion chromatography (SEC) using a Superdex G200 10/300 column (GE Healthcare, Waukesha, USA) pre-equilibrated with storage buffer (30 mM Tris-HCl, pH 7.4, and 150 mM NaCl). Considering that the purified MYG1 protein was bound to some metal ions during expression and purification, we treated it with 2 mM EDTA to remove these non-specifically bound metal ions. The EDTA-treated MYG1 protein was further purified using SEC, which was used for structural and biochemical studies unless otherwise specified. MYG1 mutants containing point mutations were generated using 2× Phanta Flash Master Mix (Vazyme Biotech, Nanjing, China) and confirmed by DNA sequencing. Expression and purification of the mutants were the same as for the wild-type protein.

Enzymatic activity assay

All RNA and DNA oligonucleotides were purchased from Sangon (Shanghai, China) and labeled at either the 3′- or 5′-end with 6-carboxfluorescein (6-FAM) (Supplementary Table S1). For a typical nuclease cleavage assay, 100 nM substrate (either ssDNA or ssRNA) was incubated with increasing concentrations (100 to 1000 nM) of freshly prepared (either EDTA-untreated or EDTA-treated) wild-type (WT) or mutant MYG1 in 20 μL of reaction buffer containing

40 mM Tris-HCl (pH 7.8), 100 mM KCl, and 2 mM TCEP (Tris-2-carboxyethyl-phosphine hydrochloride) supplemented with 2 mM Mn^{2+} at 37°C for 60 min (ssRNA) or 120 min (ssDNA) unless otherwise specified. After incubation, 7 μL of 4× loading buffer (50 mM EDTA, 20% formamide, 8 M urea, 0.5% xylene cyanol FF, and 0.5% bromophenol blue) was added to the reaction mixture and then incubated at 70°C for 5 min. The reaction products were resolved on 15% polyacrylamide sequencing gels containing 7 M urea and visualized on an Amersham Typhoon RGB (Cytiva, Marlborough, USA).

To evaluate the impact of various divalent metal ions on the nuclease activity of MYG1, we measured the specific activities of EDTA-untreated MYG1 under the reaction conditions described above, both with and without the addition of 2 mM Mn^{2+} or Mg^{2+}. Additionally, we assessed the specific activities of EDTA-treated MYG1, both with and without the addition of one of six common divalent metal ions (2 mM Mn^{2+}, Mg^{2+}, Ni^{2+}, Ca^{2+}, Zn^{2+}, or Fe^{2+}).

To investigate the cleavage activity of MYG1 toward nucleic acids with different structures, 5′ FAM-labeled nucleic acids (both RNA and DNA) of varying lengths were synthesized and then annealed to generate structured substrates, including double-stranded, 3′-overhang, 3′-flap, and Y-shaped nucleic acids (Supplementary Table S1). The activity of EDTA-treated MYG1 toward these substrates was assessed under the same reaction conditions described above.

To quantitatively analyze the cleavage activity of MYG1 toward ssDNA and ssRNA, we measured the activity of MYG1 using the Qubit™ ssDNA Assay kit (Q10212; Invitrogen, Carlsbad, USA). Briefly, 5 μM substrate (either 25 nt ssDNA-10 or ssRNA-10) (Supplementary Table S1) was incubated with 50 μM EDTA-treated MYG1 in a 30-μL reaction buffer containing 40 mM Tris-HCl (pH 7.8), 100 mM KCl, and 2 mM TCEP supplemented with 2 mM Mn^{2+}. The reaction mixture was incubated at 37°C for varying time periods. After incubation, the sample was processed according to the manufacturer's protocol for the Qubit™ ssDNA Assay kit. Fluorescence values were recorded, and the data were analyzed and plotted using GraphPad Prism (GraphPad Software, San Diego, USA).

Crystallization, data collection, and structure determination

Crystallization of both WT and mutant MYG1 (20 mg/mL) mixed with an equal volume (150 nL) of reservoir solution was performed using the hanging-drop vapor diffusion method at 16°C. The EDTA-treated WT MYG1 was crystallized alone or with the addition of Mn^{2+} (2 mM) or Mn^{2+} (2 mM) and different mononucleotides or nucleotide monophosphates (NMPs) (2 mM AMP, GMP, CMP, UMP, dAMP, dGMP, dCMP, and dTMP). The EDTA-treated H50A, H55A, D57A, E58A, D93A, H107A, and D176A mutants were crystallized in solution with the addition of 2 mM Mn^{2+} and 2 mM AMP. The EDTA-treated WT MYG1 and H50A mutant were crystallized in solution with the addition of Mn^{2+} (2 mM) and a deoxy-trinucleotide (2 mM dACA, ssDNA-11 in Supplementary Table S1). The EDTA-untreated D57A and D93A mutants were crystallized alone without the addition of any nucleic acids or metal ions. The crystallization conditions for all of these crystals are summarized in Supplementary Table S2.

Prior to data collection, the crystals were cryoprotected using reservoir solution supplemented with 20% glycerol and then flash-cooled in liquid nitrogen. Diffraction data were collected at 100 K on beamlines of the Shanghai Synchrotron Radiation Facility (SSRF) and National Facility for Protein Science in Shanghai (NFPSS) and processed with autoPROC [26], HKL2000 [27], and XDS [28]. The structure of WT MYG1 bound with Na^+ (MYG1-Na) was solved by molecular replacement (MR) implemented in Phenix [29], using the Alphafold2-predicted MYG1 structure as the search model. All other structures of both WT and mutant MYG1 were determined by MR using the MYG1-Na structure as the search model. Manual model building was performed using Coot [30], and structural refinement was carried out using Phenix. The stereochemistry quality of the final structure models was validated by the Validation and Deposition Services at the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB). Molecular graphics figures were prepared using PyMOL (https://pymol.org/). Statistics for diffraction data, structure refinement, and quality of the final structure models are summarized in Supplementary Table S3.

Isothermal titration calorimetry analysis

The binding affinities of WT MYG1 with different NMPs were measured by isothermal titration calorimetry (ITC) at 25°C using a MicroCal PEAQ-ITC instrument (Malvern Instruments, Worcestershire, UK). The protein sample was prepared in buffer containing 30 mM Tris-HCl (pH 7.4) and 150 mM NaCl. For each experiment, a total of 20 injections of 2 μL NMP or dNMP (300 μM; with the first injection being 0.5 μL) were titrated into the sample cell containing 30 μM MYG1, with a 120-s interval between injections. The same method was used to measure the binding affinities of the D57A mutant with polynucleotides (6 nt ssDNA-13 and ssRNA-12 in Supplementary Table S1). For each experiment, a total of 20 injections of 2 μL ssDNA or ssRNA (100 μM; with the first injection being 0.5 μL) were titrated into the sample cell containing 10 μM D57A mutant, with a 120-s interval between injections. The ITC data were analyzed using MicroCal PEAQ-ITC Analysis Software (Malvern Instruments) with a one-site binding model. Each experiment was independently repeated twice.

Results

Expression and purification of MYG1

MYG1 is a dually localized nuclease targeting both the mitochondria and the nucleus. Previous studies have shown that its N-terminal residues 1–39 encompass mitochondrial and nuclear localization signals [15]. The Alphafold2-predicted MYG1 structure suggests that residues 1–41 form an unstructured region. To facilitate protein expression, purification, and crystallization, we constructed a truncated MYG1 variant (residues 42–376) by omitting the N-terminal residues 1–41. This construct was successfully expressed as a recombinant protein using an *E. coli* expression system. The target protein was purified via a two-step purification process involving Ni-NTA affinity chromatography and SEC. The purified MYG1 protein was found to bind to unknown metal ion(s) and exhibited partial activity toward ssDNA and ssRNA (see results later). To simplify the biochemical and structural studies, we treated the purified protein with 2 mM EDTA to remove non-specifically bound metal ions. The EDTA-treated MYG1 protein was further purified by SEC, which exhibited high purity and homogeneity as analyzed by SEC and SDS-PAGE (Figure 1A). Size-exclusion chromatography-multiangle light scattering (SEC-MALS) analysis revealed that the EDTA-treated MYG1 protein exists as a dimer in solution (Figure 1B). Both EDTA-untreated and EDTA-

treated MYG1 proteins were used for structural and biochemical studies.

MYG1 is a Mn^{2+}/Mg^{2+}-dependent 3′→5′ exonuclease

Previous studies have shown that MYG1 possesses 3′→5′ exonuclease activity toward both ssDNA and ssRNA [14]. To assess the nuclease activity of MYG1, we performed cleavage activity assays using 5′-FAM-labeled ssDNA and ssRNA as substrates. The results showed that EDTA-untreated MYG1 exhibited nuclease activity toward both ssDNA and ssRNA in a dose-dependent manner without the addition of any metal ions but notably higher activity when supplemented with Mn^{2+} or Mg^{2+} (Figure 1C,D), indicating that EDTA-untreated MYG1 was bound to some metal ions. In contrast, EDTA-treated MYG1 by itself exhibited little activity toward both ssDNA and ssRNA; however, it displayed significantly higher activity in the presence of Mn^{2+} or Mg^{2+}, with a particularly pronounced effect on ssRNA (Figure 1E).

To characterize the enzymatic properties of MYG1, we first evaluated the impact of various divalent metal ions on its nuclease activity. EDTA-treated MYG1 alone exhibited negligible cleavage activity toward both ssDNA and ssRNA (Figure 1E). Upon the addition of six common divalent metal ions individually to the reaction solution, MYG1 displayed moderate activity toward ssDNA in the presence of Mn^{2+} or Mg^{2+}, with Mg^{2+} being slightly more favorable than Mn^{2+} (Figure 1E). Additionally, it displayed very weak activity in the presence of Ni^{2+} or Ca^{2+}. Similarly, MYG1 exhibited significantly higher activity toward ssRNA in the presence of Mn^{2+} or Mg^{2+}, with Mn^{2+} being slightly more favorable than Mg^{2+} (Figure 1E). Additionally, it displayed moderate activity in the presence of Ni^{2+} and weak activity in the presence of Ca^{2+} or Zn^{2+}. These results align with previous studies showing that MYG1 is a member of the DHH family [14] and that the catalytic activity of DHH family nucleases relies on divalent metal ions with a preference for Mn^{2+} and Mg^{2+} [7]. Consequently, to simplify biochemical studies, we used EDTA-treated MYG1 supplemented with 2 mM Mn^{2+} in all subsequent activity assays unless otherwise specified.

It is noteworthy that a previous study indicated that the activity of MYG1 is dependent on the presence of Mg^{2+}, but no activity is observed in the presence of Mn^{2+} [14]. In contrast, our cleavage activity assays demonstrated that EDTA-treated MYG1 exhibited activity toward both ssRNA and ssDNA in the presence of both Mg^{2+} and Mn^{2+}. This discrepancy may be attributed to differences in protein preparation procedures. Specifically, in our study, we utilized EDTA-treated MYG1 to investigate the effects of various divalent metal ions on its nuclease activity, whereas the previous study employed purified MYG1 without EDTA treatment for their activity assays. Our biochemical data indicated that purified, EDTA-untreated MYG1 was bound to some divalent metal ions, which may interfere with the assessment of metal ion dependence in the activity assays. Additionally, variations in the activity assay conditions, including different concentrations of divalent metal ions and different types of substrates, may also contribute to this discrepancy.

Next, we synthesized ssDNA and ssRNA substrates of varying lengths, each labeled with a FAM at either the 3′- or 5′-end (Supplementary Table S1), and determined the cleavage directionality of MYG1 based on the pattern of the cleavage products by electrophoresis. The results indicated that MYG1 cleaved nucleic acids sequentially from the 3′-end, confirming that it is a 3′→5′ exonuclease (Figure 1F).

Moreover, we examined the ability of MYG1 to cleave nucleic acids (RNA and DNA) with different structures, including double-stranded, 3′-overhang, 3′-flap, and Y-shaped forms (Supplementary Table S1). The activity assays revealed that MYG1 could cleave these substrates with varying activities, exhibiting a preference for RNA, particularly ssRNA (Figure 1G).

In summary, our biochemical data demonstrate that MYG1 is a Mn^{2+}/Mg^{2+}-dependent 3′→5′ exonuclease capable of cleaving a broad range of nucleic acids with different structures, with the highest activity observed toward ssRNA.

Crystal structure of MYG1

To elucidate the molecular basis for substrate binding and the catalytic mechanism of MYG1, we determined the crystal structures of WT MYG1 bound to Na^{+}, Mn^{2+}, or Mn^{2+} and various NMPs (AMP, GMP, CMP, UMP, dAMP, dCMP, dGMP, and dTMP). Crystallization of the EDTA-treated WT MYG1 alone without the addition of any divalent metal ions yielded crystals, leading to the determination of the structure of MYG1 bound to Na^{+} (MYG1-Na) (Supplementary Table S3). In this structure, clear electron density was observed for two metal ions at the active site. Given the high concentration of NaCl in the crystallization solution, we assigned the bound metal ions as Na^{+}. Structure refinement showed that the Na^{+} ions fit the electron density well (Supplementary Figure S1A), exhibited reasonable B-factors comparable to the protein (Supplementary Table S3), and possessed six coordinating ligands with appropriate geometry and bond lengths (see results later).

Crystallization of the EDTA-treated WT MYG1 in the presence of Mn^{2+} (2 mM) yielded crystals, leading to the determination of the structure of MYG1 bound to Mn^{2+} (MYG1-Mn) (Supplementary Table S3). In this structure, two metal ions are identified at the active site. The presence of 2 mM Mn^{2+} in the crystallization solution led us to assign the bound metal ions as Mn^{2+}, which fit the electron density well (Supplementary Figure S1A), exhibited reasonable B-factors comparable to the protein (Supplementary Table S3), and had six coordinating ligands with appropriate geometry and bond lengths (see results later).

Crystallization of EDTA-treated WT MYG1 in the presence of Mn^{2+} and eight different NMPs (AMP, GMP, CMP, UMP, dAMP, dCMP, dGMP, and dTMP) resulted in crystals, allowing us to determine the structures of MYG1 in complexes with Mn^{2+} and NMPs (MYG1-Mn-NMPs) (Supplementary Table S3). In these structures, two Mn^{2+} ions are bound at the active site (Supplementary Figure S1A). Additionally, clear electron density was observed for one NMP (NMP1) at the active site, designated the N1 site. Moreover, in the structures of MYG1 complexed with AMP, dCMP, and dTMP, clear electron density was observed for another NMP (NMP2) bound adjacently to the N1 site (designated the N2 site), although this NMP is not linked to the nucleotide at the N1 site by a phosphodiester bond (Supplementary Figure S1A). In contrast, in the other five structures, only sparse electron density was observed at the N2 site, which was modeled as water molecules. This variation is most likely due to differences in crystallization conditions, local binding environments, and the chemical properties of the nucleic acids, resulting in varying ligand occupancy.

In all of these structures, most residues of MYG1 are clearly defined in electron density maps, except a few residues located on

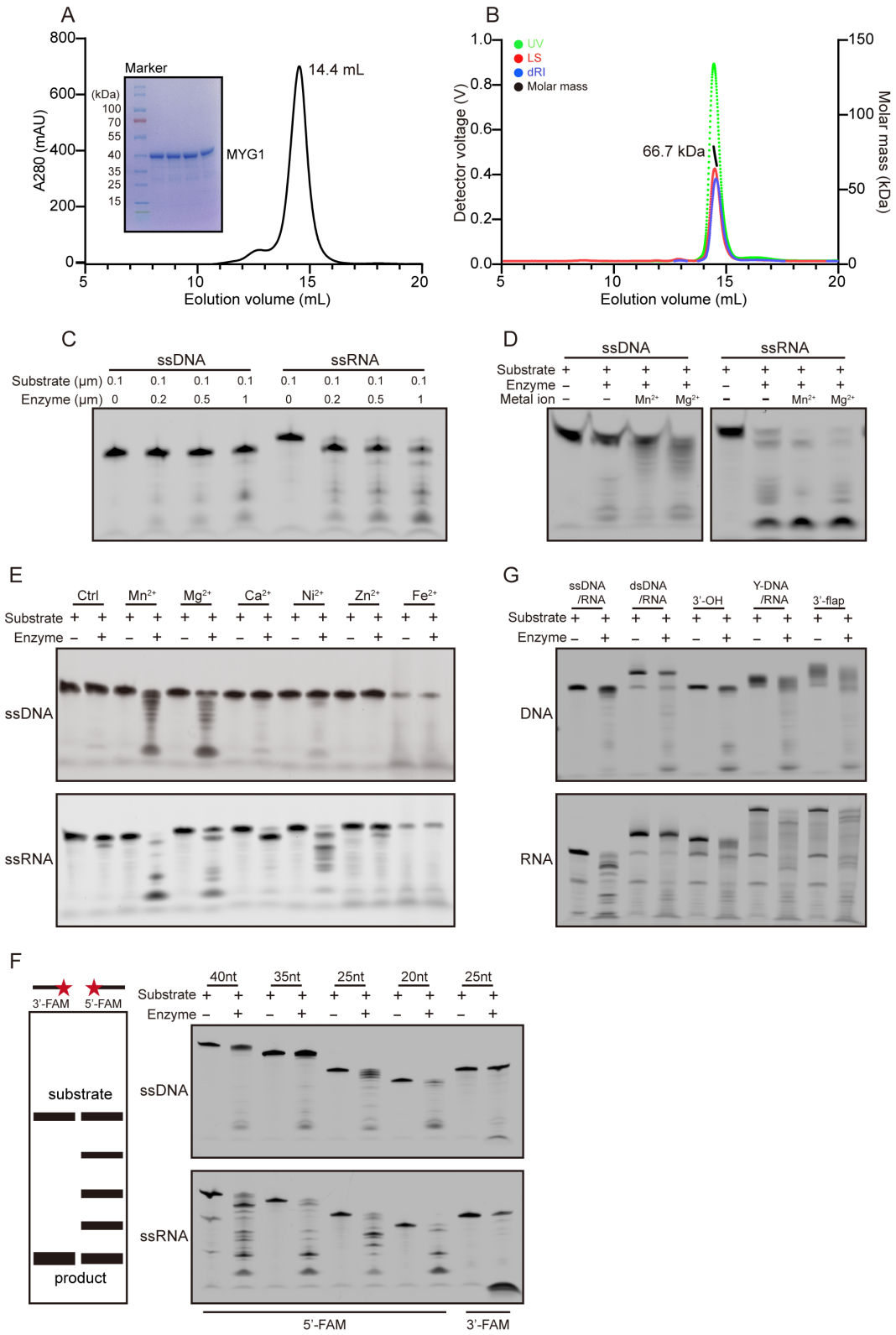


Figure 1. Purification, biochemical characterization, and enzymatic properties of MYG1 (A) SEC and SDS-PAGE analyses of the EDTA-untreated MYG1. (B) SEC-MALS analysis of the EDTA-treated MYG1. The black line indicates the apparent molecular mass (~66.7 kDa) of MYG1, corresponding to approximately twice the calculated molecular mass of the MYG1 monomer (37.8 kDa). (C) Cleavage activity assay of the EDTA-untreated MYG1. (D) Cleavage activity assay of the EDTA-untreated MYG1 in the presence of Mg^{2+} or Mn^{2+}. (E) Cleavage activity assay of the EDTA-treated MYG1 in the presence of different divalent metal ions. (F) Cleavage directionality of the EDTA-treated MYG1 and its activity on substrates of varying lengths. (G) Cleavage activity assay of the EDTA-treated MYG1 on nucleic acids with different structures.

solvent-exposed surface regions (Supplementary Table S3). MYG1 consists of an N-terminal domain and a C-terminal domain, which are spatially separated without direct interaction but are connected by a long α-helix (α7) (Figure 2A). The N-terminal domain is composed of a central parallel four-stranded β-sheet (β1–β4) flanked by one α-helix (α2) on one side and five α-helices (α1 and α3–α6) on the other, and the C-terminal domain consists of a central mixed β-sheet (β5–β9) surrounded by three α-helices (α8–α10) (Figure 2A). Based on structural analysis of MYG1 (see results later) and its structural comparison to other DHH nucleases, we defined the N-terminal domain of MYG1 as the catalytic domain and the C-terminal domain as the substrate-binding domain.

Our biochemical data indicate that MYG1 exists as a dimer in solution (Figure 1B). In all of these structures, there is one MYG1 molecule in the asymmetric unit (ASU) (Supplementary Table S3). However, structural analysis revealed that two crystallographic 2-fold axis-related MYG1 molecules form a compact dimer in a "head-to-tail" arrangement (Figure 2B). Specifically, the catalytic domain of monomer 1 interacts with the substrate-binding domain of monomer 2, with the residues of the DHH motif (D106-H107-H108) in the catalytic domain of monomer 1 oriented toward the substrate-binding domain of monomer 2 to form the active site, and vice versa. The dimer interface is characterized by extensive intermolecular hydrophilic and hydrophobic interactions, which buries 5109 Å^2 or 27% of the solvent-accessible surface of each monomer, indicating a very tight dimer interface. These results indicate that the dimer is the basic functional unit that contains two active sites. Due to the extensive interactions at the dimer interface, attempts to obtain an MYG1 monomer through mutations of residues at the interface for functional validation were unsuccessful.

Structure of the metal ion-binding site

The active site of MYG1 is located at the interface between the catalytic domain of monomer 1 and the substrate-binding domain of monomer 2 (Figure 2B). The active site comprises a metal ion-binding site and a substrate-binding site, which are positioned adjacent to each other. A characteristic feature of DHH family proteins is the binding of two metal ions in a highly negatively charged pocket of the active site [31,32]. Consistent with this, the metal ion-binding site of MYG1 is primarily formed by several strictly conserved residues from the catalytic domain of one MYG1 monomer, including H50, H55, D57, E58, D93, H107, and D176,

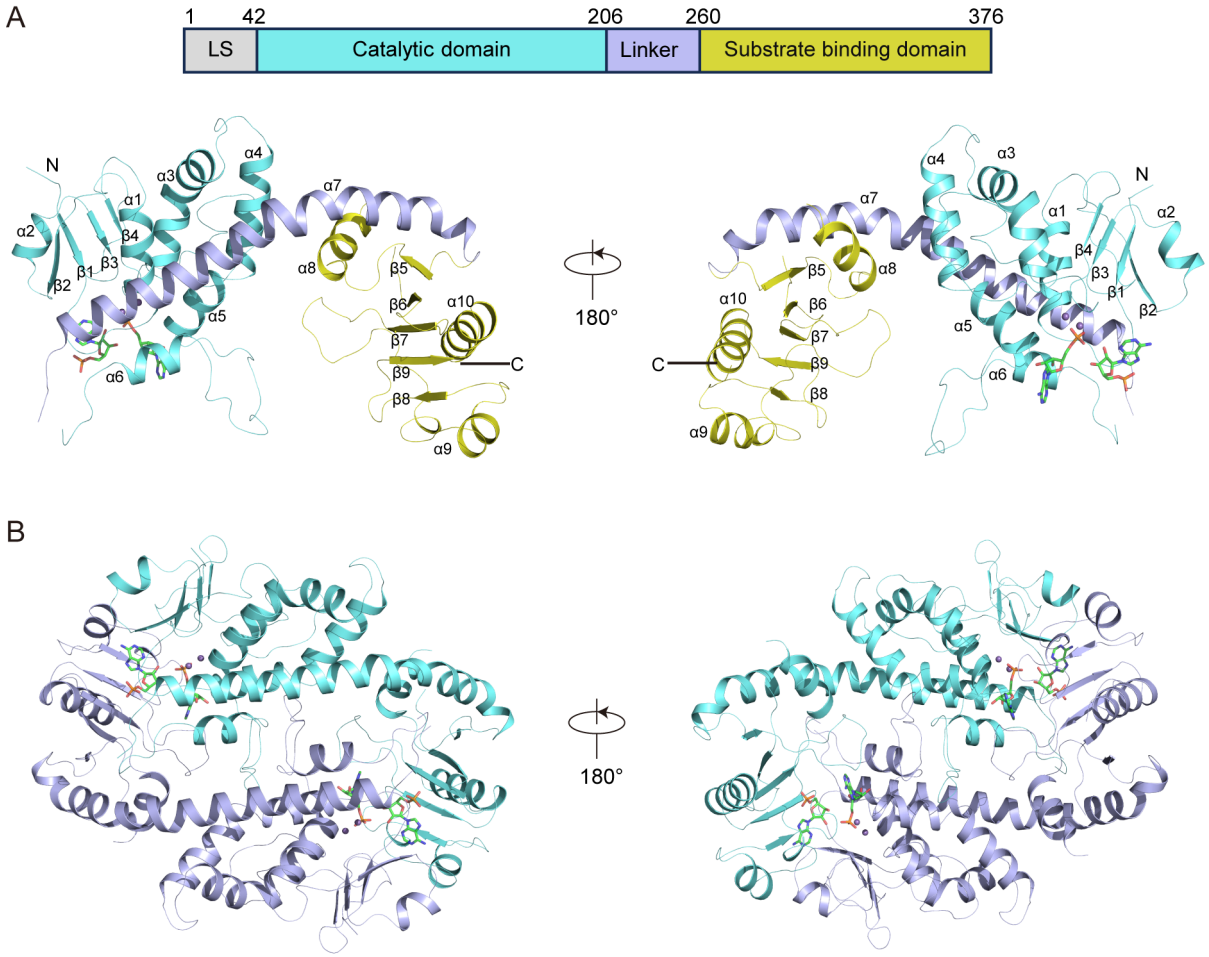


Figure 2. Crystal structure of MYG1 (A) Overall structure of MYG1-Mn-AMP in two different orientations. A schematic diagram showing the domain architecture of MYG1 is placed above the overall structure. The catalytic domain is colored cyan, the substrate-binding domain is colored yellow, and the linker region is colored purple. The secondary structure elements are labeled. The bound Mn^{2+} and AMP are shown in sphere and stick models. (B) Overall structure of the MYG1-Mn-AMP dimer. The catalytic domain of monomer 1 (cyan) interacts with the substrate-binding domain of monomer 2 (purple) to form the active site. The active site is indicated with the bound Mn^{2+} and AMP.

which create an electronegative pocket to coordinate the two metal ions (Figure 3A).

In the MYG1-Na structure, two Na^+ ions (designated Na1 and Na2) are located at the metal-binding site. Na1 is coordinated by the sidechains of D57, D93, H107, and D176 and two water molecules (Wat1 and Wat2), while Na2 is coordinated by the sidechains of H50, H55, E58, and D93 and two water molecules (Wat1 and Wat3) (Figure 3B). In the MYG1-Mn structure, two Mn^{2+} ions (designated Mn1 and Mn2) occupy the same positions and are coordinated by the same ligands as the two Na^+ ions in the MYG1-Na structure (Figure 3C). In the MYG1-Mn-NMP structures, two Mn^{2+} ions also occupy the same positions and are coordinated by the same ligands as in the MYG1-Mn structure, except that two phosphate oxygens of NMP1 replace Wat2 and Wat3, contributing one coordinating ligand each to Mn1 and Mn2 (Figure 3D).

Similar to other DHH family nucleases, the residues coordinating the metal ions in MYG1 are primarily aspartates and histidines, including the first histidine (H107) of the DHH motif, reflecting the structural conservation within this family. To investigate the functional role of the key residues involved in metal ion binding, we performed site-directed mutagenesis and activity assays. The results indicated that the D57A, D93A, and D176A mutants

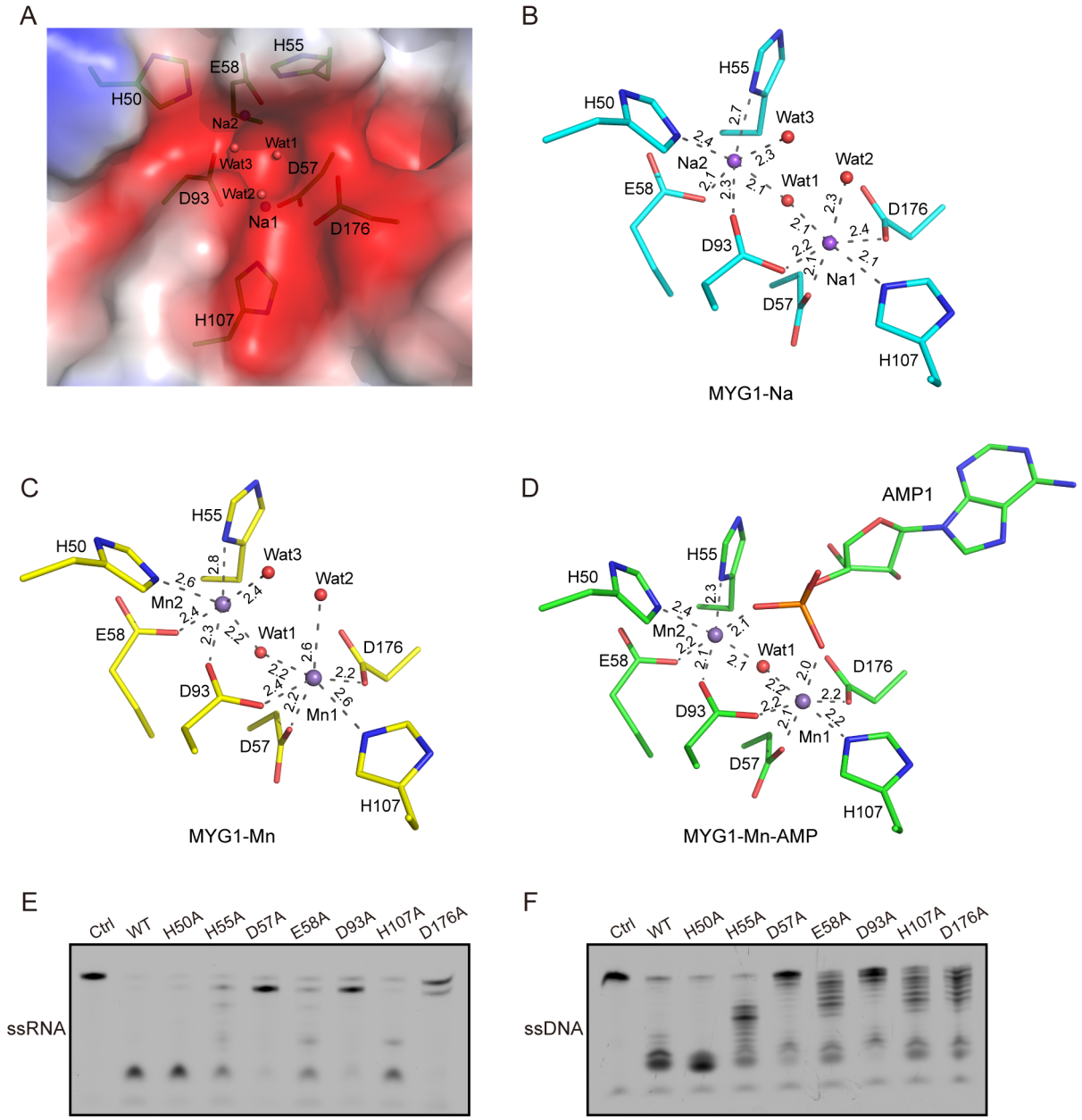


Figure 3. Structure of the metal ion-binding site of MYG1 (A) Electrostatic surface representation of the metal ion-binding site in the MYG1-Na structure. The Na^+ ions and water molecules at the active site are shown in sphere and the key residues forming the metal ion-binding site are shown in stick models. (B) Structure of the metal ion-binding site in the MYG1-Na structure. The coordination bonds are indicated with dashed lines and shown with bond lengths. (C) Structure of the metal ion-binding site in the MYG1-Mn structure. (D) Structure of the metal ion-binding site in the MYG1-Mn-AMP structure. (E) Cleavage activity assay of WT MYG1 and the mutants involved in metal ion binding on ssRNA. (F) Cleavage activity assay of WT MYG1 and the mutants involved in metal ion binding on ssDNA.

exhibited little or very weak activity, while the H55A, E58A, and H107A mutants displayed weak activity (Figure 3E,F). Surprisingly, the H50A mutant exhibited comparable activity for ssRNA and slightly higher activity for ssDNA compared to WT MYG1 (Figure 3E,F). These results suggest that mutations of the residues coordinating the metal ions likely affect metal ion binding and subsequently the catalytic reaction, leading to impaired activity of MYG1. Notably, mutations affecting Mn1 binding had more severe effects on activity than those impacting Mn2 binding.

Structure of the product-binding site

In all eight MYG1-Mn-NMP structures, one NMP (NMP1) binds to the N1 site of the substrate-binding site (Figure 4A). Additionally, in the structures of MYG1 complexed with AMP, dCMP, and dTMP, another NMP (NMP2) binds to the N2 site (Figure 4B). Although the two NMPs are not linked by a phosphodiester bond, the ribose moiety of NMP2 is positioned near the phosphate group of NMP1 (Supplementary Figure S2). The bound NMPs engage in both hydrophilic and hydrophobic interactions with several residues of MYG1 (Figure 4). Specifically, in the structures of MYG1 complexed with AMP, GMP, CMP, and UMP, the base of NMP1 is sandwiched between the phenyl rings of F313 and F348, forming strong π-π stacking interaction (Figure 4A). The 2′-OH of the ribose moiety forms a hydrogen bond with the mainchain carbonyl of V175. The three oxygen atoms of the phosphate group coordinate with the two Mn^{2+} ions and form a hydrogen bond with the sidechain of H108. Additionally, the N6 amine of the purine ring in AMP1 forms two hydrogen bonds with the mainchain carbonyl of F313 and the sidechain hydroxyl of S346. The N1 atom of the purine ring in GMP1 forms a hydrogen bond with the mainchain carbonyl of F313, and the O6 atom forms a hydrogen bond with the sidechain hydroxyl of S346. The N4 amine of the pyrimidine ring in CMP1 forms two hydrogen bonds with the mainchain carbonyl of F313 and the sidechain hydroxyl of S346. In contrast, only the O4 atom of the pyrimidine ring in UMP1 forms a hydrogen bond with the sidechain hydroxyl of S346. Consistent with the structural data, the ITC binding assays indicated that MYG1 exhibited comparable binding affinity for AMP, GMP, and CMP (K_D of 4.3–4.9 μM) but very weak binding for UMP (Supplementary Figure S3).

In the MYG1-Mn-AMP structure, the base of AMP2 is sandwiched between the phenyl ring of F342 and the sidechain of R80, forming strong π-π stacking and cation-π interactions (Figure 4B). The N1 atom of the purine ring forms a hydrogen bond with the sidechain of R327. The 2′-OH of the ribose moiety forms a hydrogen bond with the sidechain of R80, while the phosphate group forms two hydrogen bonds with the sidechain of R302.

In the structures of MYG1 complexed with dAMP, dCMP, dGMP, and dTMP, dNMP1 makes similar interactions with MYG1 as

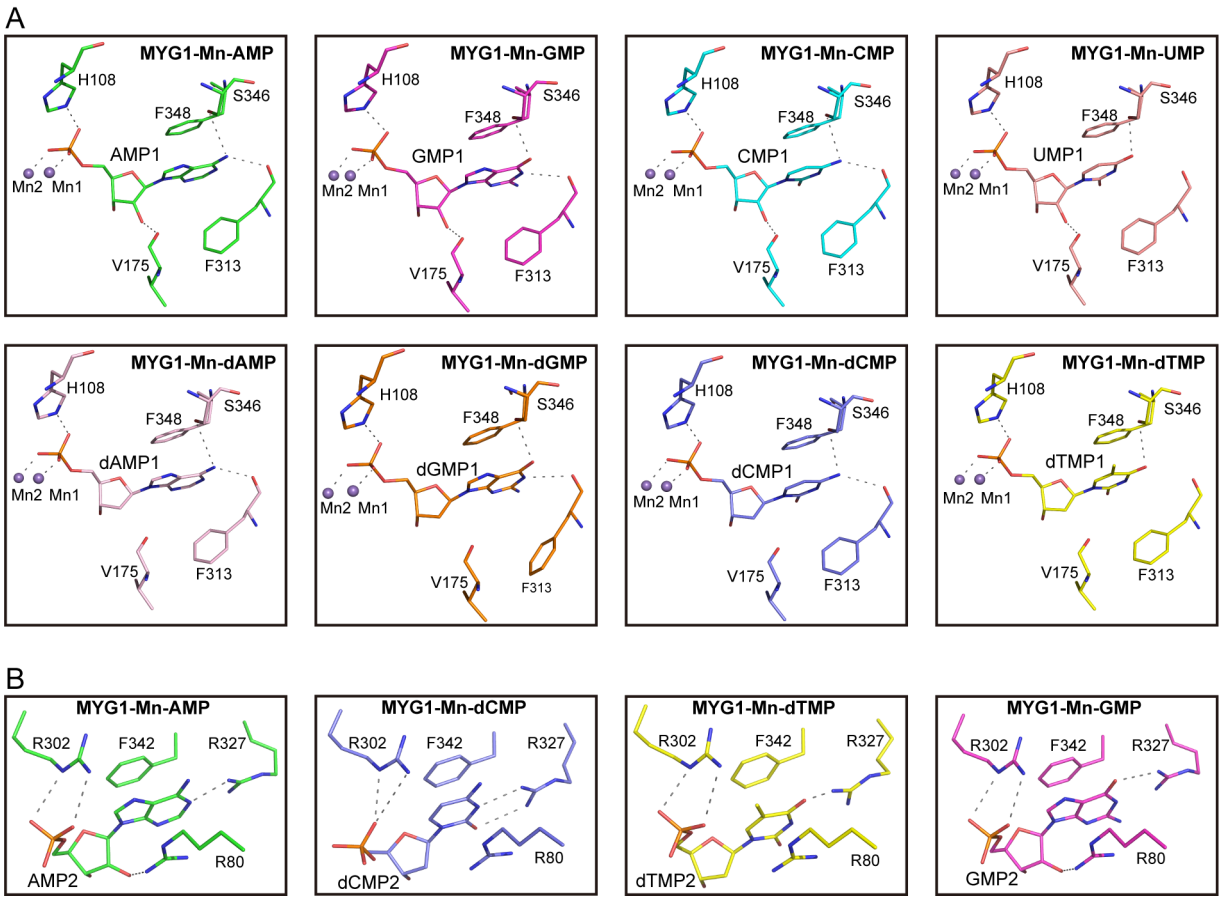


Figure 4. Structure of the product-binding site of MYG1 (A) Interactions of NMP at the N1 site (NMP1) with the surrounding residues and the metal ions in different MYG1-Mn-NMP structures. The hydrogen bonds and coordination bonds are indicated with dashed lines. (B) Interactions of NMP at the N2 site (NMP2) with the surrounding residues in different MYG1-Mn-NMP structures. The MYG1-Mn-GMP structure does not contain a GMP at the N2 site, and GMP2 was docked into the N2 site based on the position of AMP2.

observed in the corresponding NMP-bound MYG1 structures (Figure 4A). In the MYG1-Mn-dCMP/dTMP structures, dNMP2 also engages in similar interactions with MYG1 as observed in the MYG1-Mn-AMP structure, except that the O2 and N3 atoms of the pyrimidine ring in dCMP form two hydrogen bonds with the sidechain of R327, while the O4 atom of the pyrimidine ring in dTMP forms a hydrogen bond with the side chain of R327 (Figure 4B). Due to the absence of the 2′-OH in the ribose moiety of dNMP, the ribose does not interact with the protein in these structures. Consistently, ITC binding assays demonstrated that MYG1 exhibited comparable binding affinity for dAMP, dGMP, and dCMP (K_D of 7.3–10.9 μM) but no detectable binding for dTMP (Supplementary Figure S3). Additionally, MYG1 exhibited slightly higher binding affinities (approximately 2-fold) for NMPs (AMP, GMP, and CMP) than for dNMPs (dAMP, dGMP, and dCMP).

Collectively, the structural analysis of WT MYG1 in complexes with eight different NMPs revealed a conserved nucleotide-binding mode at the N1 and N2 sites across different nucleotides. Notably, F313, S346, and F348 play critical roles in recognizing and binding the nucleotide base at the N1 site, while R80, F342, and R327 are involved in recognizing and binding to the nucleotide base at the N2 site. Additionally, the phosphate group of NMP1 coordinates with both metal ions, and the ribose moiety of NMP2 is positioned near the phosphate group of NMP1. Based on the spatial relationship between the two nucleotides and their proximity to the metal ions, we designated the N1 site as the product-binding site. Therefore, these NMP-bound MYG1 structures represent product-bound states.

Structure of the substrate-binding groove

To define the substrate-binding groove for longer nucleotide substrates, we carried out co-crystallization experiments of MYG1 with various long nucleotides. Initial co-crystallization of EDTA-treated MYG1 in the absence of Mn^{2+}, but the presence of polynucleotides of different lengths (5 nt ssDNA-12 and 10 nt ssRNA-11, as listed in Supplementary Table S1), failed to yield crystals. However, co-crystallization of EDTA-treated MYG1 in the presence of Mn^{2+} and various polynucleotides successfully produced crystals. However, structure determination revealed that in these structures, only one NMP binds to the N1 site, similar to the MYG1-Mn-NMP structures. It appears that the polynucleotides were cleaved during crystallization, as the EDTA-treated MYG1 is catalytically active in the presence of Mn^{2+} (Figure 1E).

On the other hand, during crystallization of the EDTA-untreated mutants with mutations in key residues coordinating the metal ions without supplement with Mn^{2+} and polynucleotides, we obtained crystals for the D57A and D93A mutants. Unexpectedly, structure determination revealed that both a metal ion and a polynucleotide are bound at the active site (Supplementary Figure S1B and Supplementary Table S3). In both structures, clear electron density was observed for one metal ion at the Mn2 site, while the Mn1 site is empty due to the loss of coordination by D57 or D93. As our biochemical data indicated that EDTA-untreated MYG1 was bound with certain divalent metal ions and exhibited nuclease activity, we assigned the bound metal ion as Mn^{2+} based on the electron density fitting and reasonable B-factor comparable to the protein, which maintains similar coordination geometry and bond lengths as those in the MYG1-Mn-AMP structure (Supplementary Figure S4A and Supplementary Table S3). In the D57A mutant structure, clear electron density was observed for four nucleotides at the active site, whereas in the D93A mutant structure, there was relatively weak electron density for one nucleotide at the N1 site and clear electron density for one nucleotide at the N2 site (Supplementary Figure S1B). The phosphate and ribose moieties of the bound nucleotides were well defined in the electron density maps, and the phosphate and ribose moieties of adjacent nucleotides were connected. However, the specific types of bases could not be distinctly identified; therefore, we modeled the nucleotides as polyU. It appears that the bound Mn^{2+} and polynucleotides co-purified with the protein due to the inactivity of these mutants, suggesting that these structures represent the natural substrate-bound states.

Additionally, co-crystallization of EDTA-treated WT MYG1 and the H50A mutant in the presence of Mn^{2+} and a deoxynucleotide (dACA) yielded crystals, allowing us to determine the structures of WT MYG1 and the H50A mutant in complexes with Mn^{2+} and two nucleotides, respectively (Supplementary Table S3). In the WT MYG1 structure, two Mn^{2+} ions are bound at the same positions and coordinated by the same ligands as those in the MYG1-Mn-AMP structure (Supplementary Figure S4A). In the H50A mutant structure, only one Mn^{2+} is bound at the Mn1 site, which is coordinated by the same ligands as that in the MYG1-Mn-AMP structure (Supplementary Figure S4A). Additionally, in both structures, two nucleotides bind to the N1 and N2 sites, which could be clearly modeled as dA and dC connected by a phosphodiester bond (Supplementary Figure S1B). These results indicated that the bound deoxynucleotide was not cleaved during crystallization, likely due to the weak activity of WT MYG1 and the H50A mutant toward ssDNA in the presence of Mn^{2+} (Figures 1E and 3F).

Structural comparisons revealed that the overall structure and the active-site configuration in these polynucleotide-bound structures are essentially the same as those in the MYG1-Mn and MYG1-Mn-NMP structures (Supplementary Figure S5). This suggests that the binding of either a mononucleotide (product) or a polynucleotide (substrate) does not induce significant conformational changes in the overall structure or at the active site. Since the D57A-Mn-UUUU structure contains more nucleotides, we used this structure to analyze the substrate-binding groove and the interactions of the substrate with MYG1.

Structural analysis of the D57A-Mn-UUUU complex revealed that the substrate-binding groove is formed by residues from two MYG1 monomers of the dimer, including T53 and H55 from the β1-α1 loop, R80 from the β2-α2 loop, H108 from the β4-α3 loop, V175 from the α5-helix of the catalytic domain in monomer 1, G268 and E274 from the α8-helix, R327 from the α9-helix, and R302, Q304, F313, F342, S346, and F348 from the β6-β9 sheet of the substrate-binding domain in monomer 2 (Figure 5A,B). The substrate-binding groove is largely positively charged, which recognizes and binds to the substrate through coordination bonds, hydrogen bonds, and π-π and cation-π stacking interactions (Figure 5A,B). The first and second nucleotides bind to the N1 and N2 sites, respectively, and interact with surrounding residues in a manner similar to that observed in the MYG1-Mn-NMP structures, except that the phosphate of the second nucleotide forms an additional hydrogen bond with the sidechain of H55 (Figure 5B). The binding site for the third nucleotide is designated the N3 site, where the O2 atom of the nucleotide base forms a hydrogen bond with the sidechain of T53, the 2′-OH of the ribose moiety forms a hydrogen bond with the

sidechain of R302, and the phosphate group forms a hydrogen bond with the sidechain of Q304 (Figure 5B). For the fourth nucleotide, while the phosphate and ribose moieties were clearly defined in the electron density map, the base had poor electron density and thus was not modeled (Supplementary Figure S1B). The binding site for the fourth nucleotide is designated the N4 site, where the 2′-OH of the ribose moiety forms a hydrogen bond with the mainchain carbonyl of G268, and the phosphate group forms a hydrogen bond with the mainchain amide of E274 (Figure 5B).

In the D93A-Mn-UU structure, the two bound nucleotides interact with the protein in a manner similar to the first two nucleotides in the D57A-Mn-UUUU structure (Supplementary Figure S6). In the MYG1-Mn-dAC and H50A-Mn-dAC structures, the two bound nucleotides also interact with surrounding residues in a manner

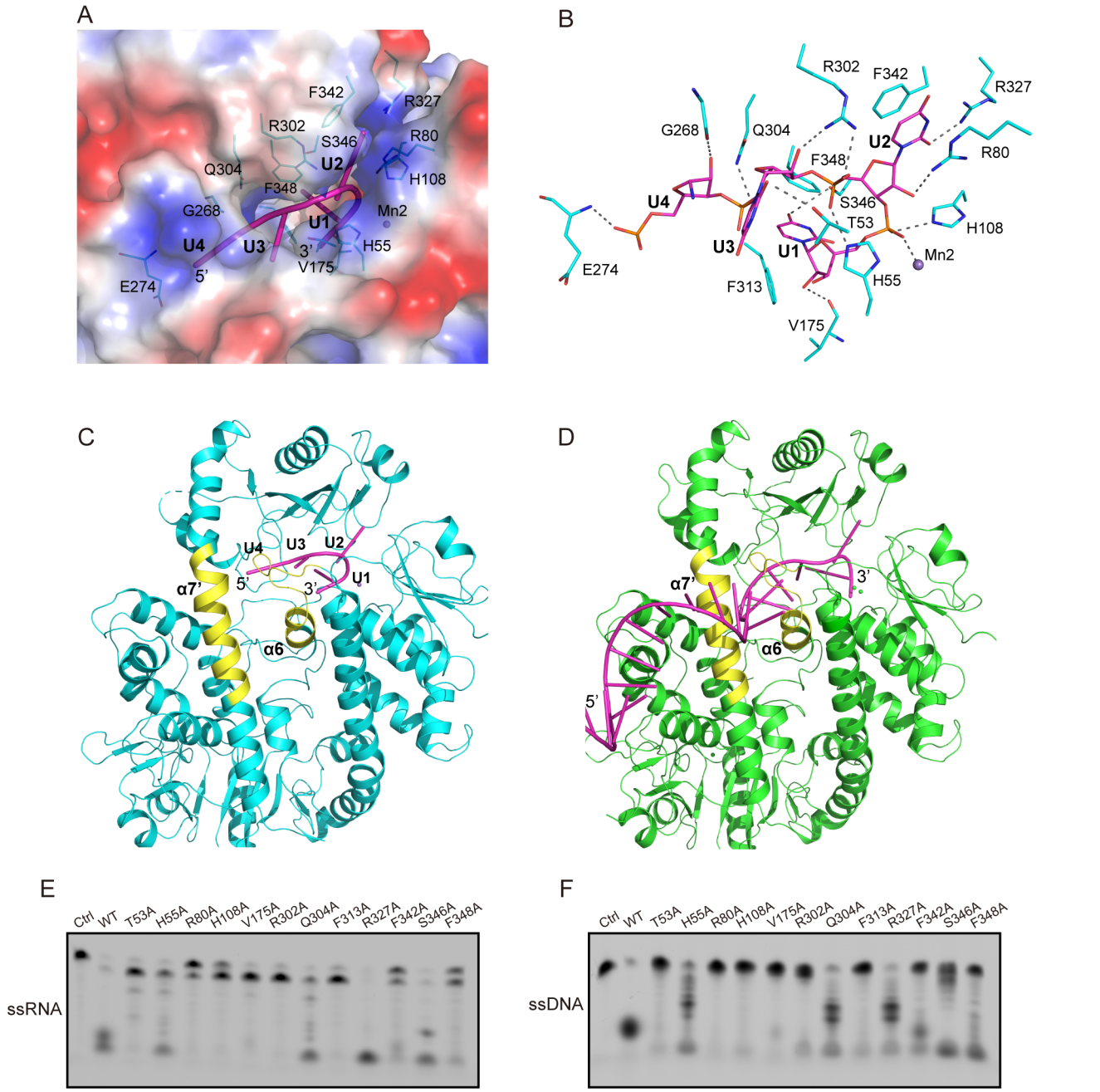


Figure 5. Structure of the substrate-binding groove of MYG1 (A) Electrostatic surface representation of the substrate-binding groove in the D57A-Mn-UUUU structure. The bound polyU substrate is shown in ribbon and the key residues forming the substrate-binding groove are shown in stick models. (B) Interactions of the polyU substrate with the surrounding residues and the metal ion in the D57A-Mn-UUUU structure. The hydrogen bonds and coordination bonds are indicated with dashed lines. (C) Dimer representation of the D57A-Mn-UUUU structure showing the entrance and directionality of the bound polyU substrate. For clarity, only one polyU substrate is shown. The α6-helix of monomer 1 and the α7′-helix of monomer 2 are highlighted in yellow. (D) AlphaFold3-predicted structure model of the MYG1 dimer in complex with a 20-nt ssRNA. The bound ssRNA is shown in ribbon. For clarity, only one polyU substrate is shown. (E) Cleavage activity assay of WT MYG1 and the mutants involved in substrate binding on ssRNA. (F) Cleavage activity assay of WT MYG1 and the mutants involved in substrate binding on ssDNA.

similar to the first two nucleotides in the D57A-Mn-UUUU structure, except that the ribose does not interact with the protein due to the absence of the 2′-OH in the ribose moiety (Supplementary Figure S6).

Based on the trajectory of the bound polynucleotides and their spatial relationships to the metal ions, we propose that a long polynucleotide enters the substrate-binding groove through the opening formed by the α6-helix of monomer 1 and the α7′-helix of monomer 2 in the 5′→3′ direction and then extends into the active site for the catalytic reaction (Figure 5C). This suggestion is supported by the AlphaFold3-predicted structure of MYG1 in complex with a long polynucleotide, showing that the long polynucleotide enters the substrate-binding groove through this opening in the 5′→3′ direction and extends toward the active site (Figure 5D).

To verify the functional role of the key residues involved in substrate binding (T53, H55, R80, H108, V175, R302, Q304, F313, R327, F342, S346, and F348), we performed site-directed mutagenesis and examined the enzymatic activity of these mutants. The activity assays indicated that most of these mutants, including T53A, R80A, H108A, V175A, R302A, F313A, F342A, and F348A, exhibited significantly decreased activity toward both ssRNA and ssDNA, while the H55A and Q304A mutants displayed moderately reduced activity compared to WT MYG1 (Figure 5E,F). Intriguingly, the R327A and S346A mutants exhibited comparable activity for ssRNA or slightly decreased activity for ssDNA compared to WT MYG1 (Figure 5E,F). Since S346 and R327 are involved in recognizing and binding the nucleotide bases at the N1 and N2 sites, respectively, it is possible that these two residues may be associated with the nucleotide preference of substrates (see discussion later).

Collectively, our structural and biochemical data indicate that the substrate-binding groove is formed by residues from two MYG1 monomers in the functional dimer, including T53, H55, R80, H108, and V175 from the catalytic domain of monomer 1 and G268, E274, R302, Q304, F313, R327, F342, S346, and F348 from the substrate-binding domain of monomer 2. Mutations of most of the residues involved in substrate binding affect the precise binding of the substrate, leading to impaired activity of MYG1.

Mn1 is essential, but Mn2 is not essential for the catalytic reaction of MYG1

Our biochemical data indicate that mutations in residues coordinating the metal ions impair the enzymatic activity of MYG1 (Figure 3E, F). Notably, mutations of the residues involved in Mn1 binding have more severe impacts on the activity than those involved in Mn2 binding. To investigate the molecular basis for the functional roles of these residues, we crystallized the EDTA-treated H50A, H55A, D57A, E58A, D93A, H107A, and D176A mutants in the presence of Mn^{2+} and AMP and determined the structures of these mutants in complexes with Mn^{2+} and AMP (Supplementary Table S3). In these structures, clear electron density was observed for one or two Mn^{2+} ions and one or two AMPs at the active site (Supplementary Figure S1C). These structures together with those of the D57A and D93A mutants in complexes with Mn^{2+} and polyU (D57A-Mn-UUUU and D93A-Mn-UU) clarify the functional roles of the residues coordinating the metal ions in the catalytic reaction.

In the D57A-Mn-AMP structure, both Mn1 and Mn2 are present at the active site, maintaining similar coordination geometry and ligands as observed in the MYG1-Mn-AMP structure, except that the coordination of D57 to Mn1 is replaced by a water molecule (Supplementary Figure S4B). Conversely, in the D57A-Mn-UUUU structure, only Mn2 is observed, while Mn1 is absent, with Mn2 maintaining similar coordination geometry and ligands as observed in the D57A-Mn-AMP structure (Supplementary Figure S4A). In the D93A-Mn-AMP and D93A-Mn-UU structures, only Mn2 is observed, exhibiting a distorted coordination geometry with five ligands due to the loss of coordination from D93 (Supplementary Figure S4A,B). Similarly, in the D176A-Mn-AMP structure, only Mn2 is bound at the active site, maintaining similar coordination geometry and ligands as observed in the MYG1-Mn-AMP structure, except that Wat1 is replaced by another phosphate oxygen of the nucleotide (Supplementary Figure S4B). In contrast, in the H107A-Mn-AMP structure, both Mn1 and Mn2 are present at the active site: while Mn2 retains similar coordination geometry and ligands as in the MYG1-Mn-AMP structure, Mn1 exhibits a distorted coordination geometry with five ligands due to the loss of coordination from H107 (Supplementary Figure S4B). These results indicate that among the residues coordinating Mn1, the D57A, D93A, and D176A mutations significantly weaken or abolish Mn1 binding, explaining the markedly reduced activity of these mutants, whereas the H107A mutation moderately weakens Mn1 binding, correlating with the moderately reduced activity of the H107A mutant (Figure 3E,F).

In the H55A-Mn-AMP and E58A-Mn-AMP structures, both Mn1 and Mn2 are observed, with coordination geometry and ligands similar to those in the MYG1-Mn-AMP structure, except that the coordination of Mn2 by H55 or E58 is replaced by a water molecule (Supplementary Figure S4B). These results indicate that among the residues coordinating Mn2, the H55A and E58A mutations only weaken Mn2 binding, accounting for the moderately reduced activity of the H55A and E58A mutants (Figure 3E,F).

Interestingly, although H50 is also involved in coordinating Mn2, in the H50A-Mn-AMP structure, only Mn1 is observed, while Mn2 is absent, with Mn1 maintaining similar coordination geometry and ligands as that in the MYG1-Mn-AMP structure (Supplementary Figure S4B). Additionally, two AMPs were identified at the active site: AMP1 had relatively poor electron density, while AMP2 was well defined (Supplementary Figure S1C). Moreover, in the H50A-Mn-dAC structure, only Mn1 is observed, which also maintains similar coordination geometry and ligands as that observed in the MYG1-Mn-AMP structure (Supplementary Figure S4A). However, two linked dAC nucleotides are bound to the N1 and N2 sites in a manner similar to the first two nucleotides in the D57A-Mn-UUUU structure (Supplementary Figure S6). Since the H50A-Mn-dAC and H50A-Mn-AMP structures represent the substrate-bound and product-bound states before and after the catalytic reaction, respectively, these results indicate that the H50A mutation abolishes Mn2 binding and weakens nucleotide binding at the N1 site, which would facilitate product release after the catalytic reaction. This may explain the slightly enhanced enzymatic activity of the H50A mutant (Figure 3E,F).

In summary, our biochemical and structural data demonstrate that metal ions are crucial in the catalytic reaction of MYG1. Specifically, Mn1 is essential for the cleavage activity, whereas Mn2 is not critical for the catalytic reaction.

MYG1 exhibits varying cleavage activity toward different nucleotides

DHH family nucleases exhibit highly diverse substrate specificity

and generally lack strict sequence-specific recognition [7]. Intriguingly, in our cleavage activity assays, certain intermediate product bands appeared to be stronger than others (Figure 1C–G), suggesting that MYG1 exhibits varying cleavage activity toward different nucleotides. To further investigate this notion, we synthesized various ssRNAs and ssDNAs with polyA, polyC, polyU/T, G-rich, and mixed sequences as substrates for cleavage activity assays (Supplementary Table S1). The results showed that among the ssRNA substrates, MYG1 exhibited high activity toward the polyU sequence, moderate activity toward the polyA sequence, mixed and G-rich sequences, and weak activity toward the polyC sequence (Figure 6A). Among the ssDNA substrates, MYG1 displayed moderate activity toward the polyT sequence, weak activity toward the mixed and polyC sequences, and very weak activity toward the polyA and G-rich sequences (Figure 6A). To verify these findings, we synthesized three polyU ssRNAs containing three consecutive A, C, or G bases in the middle (designated polyU-A/C/G) (Supplementary Table S1). The activity assays showed that for the polyU-C ssRNA, cleavage exhibited a clear pause in the middle of product formation, while the polyU-A/G ssRNAs displayed no obvious pause, confirming that MYG1 had weak activity toward the polyC sequence and moderate activity toward the polyA and G-rich sequences (Figure 6B). These results indicate that MYG1 possesses a certain degree of nucleotide preference, exhibiting higher cleavage activity toward U in ssRNA and T in ssDNA compared to other nucleotides.

Typically, differences in cleavage activity toward various nucleotides may arise from varying binding abilities with substrate and/or product, among other factors, leading to different rates of substrate translocation and/or product release after the catalytic reaction, thus affecting processivity [33]. To understand the molecular basis for the nucleotide preference of MYG1, we performed detailed structural analyses of both WT MYG1 and various mutants in complexes with either mononucleotides or polynucleotides.

In all WT and mutant MYG1 structures containing a nucleotide at the N1 site, the nucleotide base is sandwiched between the phenyl rings of F313 and F348 via π-π stacking interactions. Additionally, different bases engage in varying hydrogen-bonding interactions with surrounding residues (Figure 4A). Specifically, the N6 amine of the purine ring in AMP or dAMP forms two hydrogen bonds with the sidechain hydroxyl of S346 and the mainchain carbonyl of F313. The N1 atom of the purine ring in GMP or dGMP forms a hydrogen bond with the mainchain carbonyl of F313, while the O6 atom forms a hydrogen bond with the sidechain hydroxyl of S346. The N4 amine of the pyrimidine ring in CMP or dCMP forms two hydrogen bonds with the sidechain hydroxyl of S346 and the mainchain carbonyl of F313. In contrast, only the O4 atom of the pyrimidine

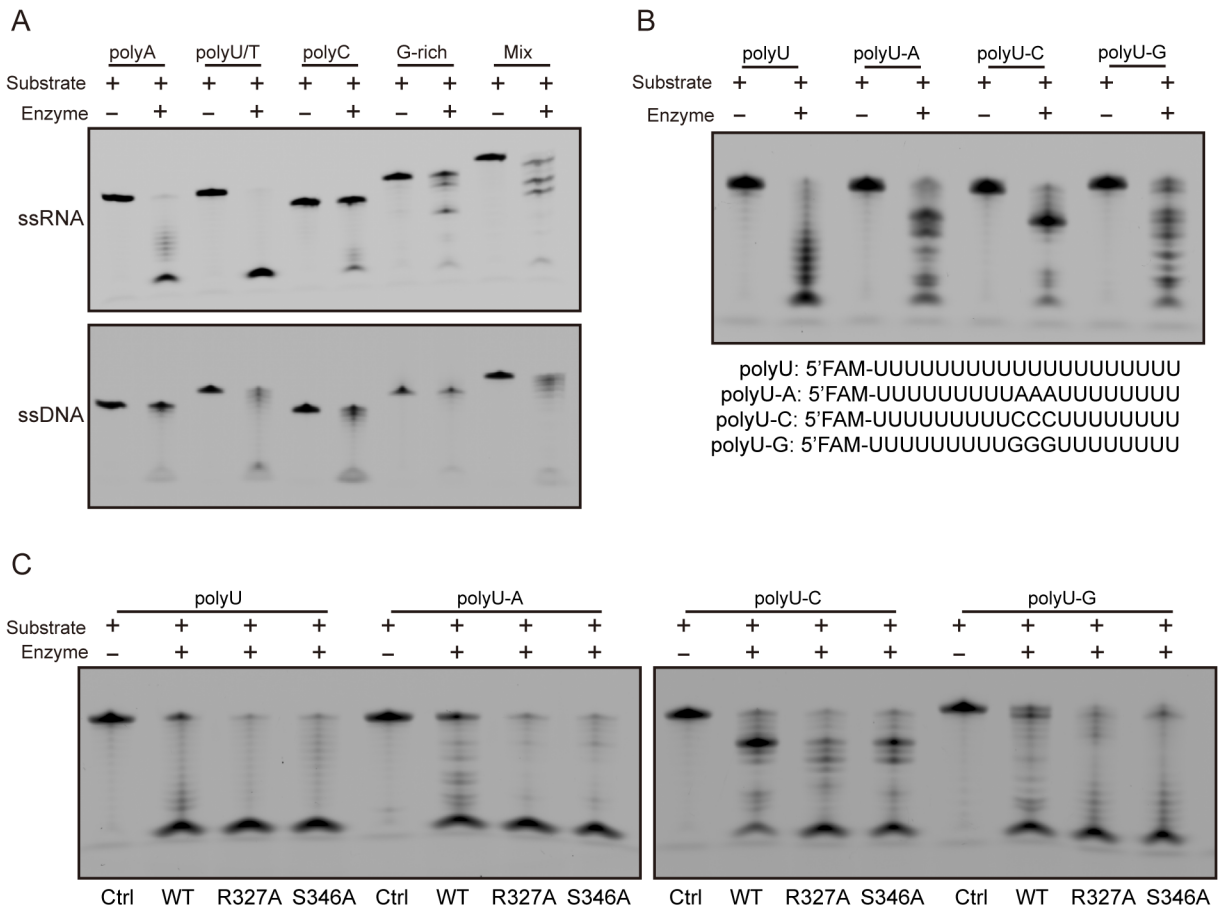


Figure 6. Nucleotide preference of MYG1 (A) Cleavage activity assay of the EDTA-treated MYG1 on nucleic acids with different sequences. (B) Cleavage activity assay of the EDTA-treated MYG1 on polyU ssRNAs, both with or without insertion of three consecutive A, G, or C in the middle. (C) Cleavage activity assay of the EDTA-treated WT MYG1 and the R327A and S346A mutants on polyU ssRNAs, both with or without insertion of three consecutive A, G, or C in the middle.

ring in UMP or dTMP forms one hydrogen bond with the sidechain hydroxyl of S346. These results indicate that the sidechains of F313 and F348 anchor the base of the nucleotide at the N1 site, while the sidechains of F313 and Ser346 are responsible for base discrimination. Notably, the base of UMP or dTMP forms one fewer hydrogen bond with the protein compared to other nucleotides, which may facilitate its release from the product-binding site after the catalytic reaction, thereby increasing the cleavage activity.

In all WT and mutant MYG1 structures containing a nucleotide at the N2 site, the nucleotide base is sandwiched between the phenyl ring of F342 and the guanidinium group of R80 via strong π-π and cation-π stacking interactions (Figure 4B). Additionally, different bases also engage in varying hydrogen bonds with surrounding residues. Specifically, the N1 atom of the purine ring in AMP or dAMP forms a hydrogen bond with the sidechain of R327. The O2 and N3 atoms of the pyrimidine ring in CMP or dCMP form two hydrogen bonds with the sidechain of R327. The O4 atom of the pyrimidine ring in UMP or dTMP forms a hydrogen bond with the side chain of R327. Since no structure containing GMP or dGMP at the N2 site was solved, we docked GMP or dGMP at the N2 site based on the position of AMP, revealing that the O6 atom of the purine ring in GMP or dGMP could form a hydrogen bond with the sidechain of R327 (Figure 4B). These results indicate that the sidechains of F342 and R80 anchor the base of the nucleotide at the N2 site, while the sidechain of R327 is responsible for base discrimination. Additionally, the base of CMP or dCMP forms one more hydrogen bond with the protein compared to other nucleotides, which may slow down substrate translocation after the catalytic reaction, thereby decreasing the cleavage activity.

Our structural analyses indicate that S346 and R327 specifically recognize and interact distinctly with the nucleotide bases at the N1 and N2 sites, respectively. The activity assays revealed that the R327A mutant exhibited moderately increased activity toward both ssRNA and ssDNA with mixed sequences, while the S346A mutant displayed comparable activity for ssRNA and slightly decreased activity for ssDNA compared to WT MYG1 (Figure 5E,F). Additionally, the cleavage bands corresponding to intermediate and final products of the R327A and S346A mutants were significantly different from those of WT MYG1. These results suggest that the varying cleavage activity of MYG1 toward different nucleotides may be related to the specific recognition and binding of nucleotide bases by R327 and S346. To further test this notion, we examined the activity of the R327A and S346A mutants toward polyU ssRNAs, both with or without the insertion of three consecutive A, G, or C bases in the middle (Supplementary Table S1). The activity assays showed that both the R327A and S346A mutants exhibited slightly higher activity than WT MYG1 toward the polyU sequence (Figure 6C). Additionally, the two mutants displayed higher activity than WT MYG1 toward the polyU-A/G/C sequences, although their activities were comparable (for polyU-A) or slightly lower (for polyU-G and polyU-C) compared to those with the polyU sequence. Notably, the activity enhancement was significant for the polyU-C sequence. These results can be explained by our structural data showing that the R327A mutant lost its interaction with the nucleotide base at the N2 site, thereby increasing the substrate translocation rate after the catalytic reaction. In contrast, the S346A mutant lost its interaction with the nucleotide base at the N1 site, thereby facilitating the product release rate after the catalytic reaction. Collectively, our structural and biochemical data demonstrate that the varying cleavage activity of MYG1 toward different nucleotides, with a preference for U in ssRNA and T in ssDNA, is attributed to the specific recognition and binding of different nucleotide bases by R327 and S346.

Molecular basis for the substrate preference of MYG1 for ssRNA over ssDNA

Previous as well as our biochemical data indicate that MYG1 exhibits much higher activity for ssRNA than for ssDNA (Figure 1C–G). This finding is further supported by a time-dependent nuclease digestion assay (Figure 7A). Additionally, we quantitatively measured the cleavage activity of MYG1 toward ssRNA and ssDNA. The results revealed that the cleavage efficiency of MYG1 for ssRNA was approximately 10-fold higher than that for ssDNA (Figure 7B,C).

To understand the molecular basis for the differing activity of MYG1 toward ssRNA and ssDNA, we compared the structures of WT and mutant MYG1 in complexes with both nucleotides and deoxynucleotides. The only structural difference between RNA and DNA is the presence of the 2′-OH group on the ribose in RNA. In the D57A-Mn-UUUU and D93A-Mn-UU structures, the 2′-OH of the ribose moiety in all nucleotides forms a hydrogen bond with surrounding residues of MYG1 (Figure 5B and Supplementary Figure S6). In contrast, in the MYG1-Mn-dAC and H50A-Mn-dAC structures, the ribose moiety does not interact with the protein due to the absence of the 2′-OH group (Supplementary Figure S6). These results indicate that the nucleotides of ssRNA form additional hydrogen-bonding interactions with MYG1 compared to those of ssDNA. Consistently, ITC binding assays showed that MYG1 exhibited slightly higher binding affinity (approximately 2-fold) for NMPs than for dNMPs, and the D57A mutant exhibited nearly 15-fold higher binding affinity for ssRNA than for ssDNA (Figure 7D,E and Supplementary Figure S3). Collectively, our biochemical and structural data demonstrate that MYG1 binds to ssRNA more tightly than to DNA, which may explain its higher activity toward ssRNA compared to ssDNA.

Catalytic mechanism of MYG1

The cleavage activity of DHH family nucleases strictly depends on divalent metal ions, such as Mn^{2+} or Mg^{2+}, and typically employs a two-metal ion-dependent catalytic mechanism [7]. In the catalytic reaction, the two metal ions are proposed to play distinct roles. Metal ion 1 (Me1) is essential for catalysis, as it activates a water molecule to serve as the nucleophile. Metal ion 2 (Me2) primarily stabilizes the transition state, thereby lowering the energy barrier of the reaction, and works cooperatively with Me1 to facilitate efficient hydrolysis of the phosphodiester bond. Indeed, our biochemical data indicate that the cleavage activity of MYG1 requires divalent metal ions, with a preference for Mn^{2+} and Mg^{2+} (Figure 1E). Our structural data demonstrate that two Mn^{2+} ions bind to the active site of MYG1: Mn1 is coordinated by D57, D93, H107, D176, a water molecule, and the phosphate group of NMP1, while Mn2 is coordinated by H50, H55, E58, D93, a water molecule, and the phosphate group of NMP1 (Figure 3D). Our mutagenesis, activity assays, and structure analyses further reveal that mutations in residues coordinating Mn1 (D57A, D93A, and D176A) significantly weaken or abolish Mn1 binding, leading to markedly reduced activity (Figure 3E,F and Supplementary Figure S4). Although H107 is also involved in Mn1 coordination, the H107A mutation only moderately weakens Mn1 binding, resulting in moderately reduced

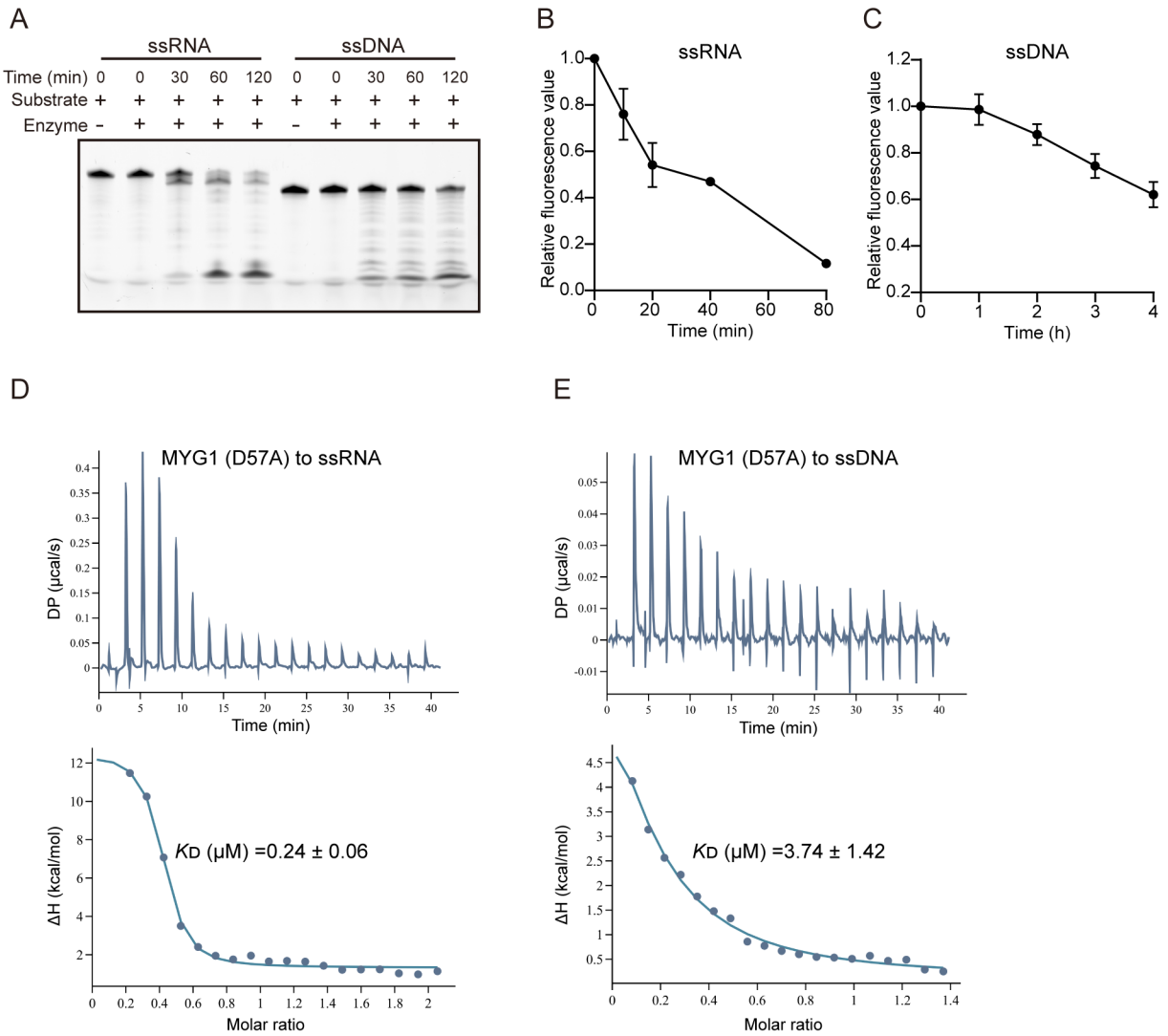


Figure 7. Substrate preference of MYG1 for ssRNA over ssDNA (A) Time-dependent cleavage activity assay of the EDTA-treated MYG1 for ssRNA and ssDNA. (B,C) Semi-quantitative analysis of the cleavage activity of the EDTA-treated MYG1 for ssRNA and ssDNA. (D,E) ITC analysis of the binding affinity of the D57A mutant for ssRNA and ssDNA.

activity. On the other hand, mutations in residues coordinating Mn2 (D55A and E58A) moderately weaken Mn2 binding, leading to moderately reduced activity (Figure 3E,F and Supplementary Figure S4). However, the H50A mutation, which also participates in Mn2 coordination, abolishes Mn2 binding and weakens nucleotide binding at the N1 site, resulting in slightly enhanced activity.

Although the two-metal ion-dependent DHH family nucleases are diverse and the key residues involved in the catalytic reaction vary, the hydrolysis of the phosphodiester bond can generally be summarized in three basic steps: nucleophilic attack, transition state formation, and bond cleavage. Our structural and biochemical data for MYG1 collectively support its nucleic acid cleavage via a two-metal ion-dependent mechanism. Based on the functional roles of the key residues involved in metal ion binding and substrate binding, we propose the catalytic mechanism of MYG1 as follows (Figure 8).

After a polynucleotide substrate enters the catalytic active site, residues F313, F348, and S346 recognize and bind to the first 3′-terminal nucleotide of the substrate, and residues R80, F342, and R327 recognize and bind to the second nucleotide. The phosphodiester bond linking the first and second nucleotides coordinates with both metal ions. Me1 activates a water molecule to generate OH^-, which acts as the nucleophile to attack the phosphorus atom, forming a pentacoordinate trigonal bipyramidal transition state. Meanwhile, Me1 and Me2 stabilize this transition state, and H108 acts as a general acid to donate a proton to the 3′-O^- group of the ribose moiety in the second nucleotide after P-O bond cleavage. Finally, the cleaved product dissociates from the active site, allowing the polynucleotide substrate to translocate and proceed to the next cleavage reaction. In this mechanism, Me1 plays a key catalytic role, while Me2 serves to stabilize the transition state and the leaving product. Notably, our biochemical and structural data indicate that Me1 is essential for the catalytic reaction, whereas Me2 is not. In conclusion, our structural and biochemical data demonstrate that MYG1 is a two-metal ion-dependent nuclease but can also cleave substrates using a single metal ion.

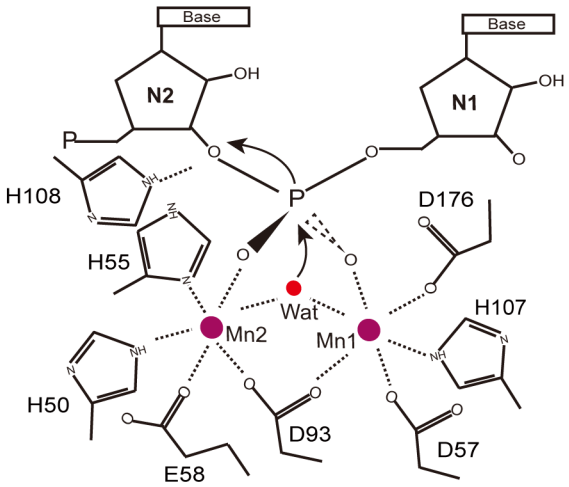


Figure 8. Catalytic mechanism of MYG1 The metal ions are shown as magenta circles, and the water molecule that generates the nucleophilic hydroxide is shown as a red circle. Arrows indicate the direction of nucleophilic attack and electron transfer.

Discussion

MYG1 is a newly identified member of the DHH family of nucleases. Previous studies have shown that it possesses 3′→5′ exonuclease activity, exhibiting high activity on ssRNA and weak activity on ssDNA but no activity on dsDNA [14]. However, the enzymatic properties, structure, substrate specificity, and catalytic mechanism of MYG1 remain largely elusive. In this study, we systematically characterized the enzymatic properties of MYG1 and successfully determined the crystal structures of MYG1 in complexes with metal ions and various mono- and poly-(deoxy)nucleotides. The structural and biochemical data indicate that MYG1 is a Mn^{2+}- or Mg^{2+}-dependent 3′→5′ exonuclease that can cleave a wide range of nucleic acid substrates with different structures, exhibiting the highest activity for ssRNA. Additionally, MYG1 exhibits varying cleavage activity toward different nucleotides, showing a preference for U in ssRNA and T in ssDNA. Mechanistically, MYG1 employs a two-metal ion-mediated catalytic mechanism but can also cleave substrates using a single metal ion. Furthermore, the DHH motif (D106-H107-H108) of MYG1 plays dual functional roles in the catalytic reaction. Although D106 is not involved in the catalytic reaction, H107 coordinates to Me1, and H108 acts as a general acid to protonate the 3′-O^- group of the ribose moiety in the second nucleotide after P-O bond cleavage.

MYG1 consists of an N-terminal catalytic domain and a C-terminal substrate-binding domain connected by a long α-helix. Intriguingly, a structural homolog search for MYG1 using the Foldseek server [34] revealed no significant matches in the Protein Data Bank (PDB). When the individual domains were examined, the N-terminal domain of MYG1 also showed no structurally similar proteins, while the C-terminal domain exhibited similarity to the C-terminal domain of the exonuclease RecJ, a member of the DHH family [35]. Further structural comparison of MYG1 with representative DHH family nucleases, including the exonuclease RecJ [35], NanoRNase NmA [36], Hef-associated nuclease (HAN) [10], and inorganic pyrophosphatase (PPase) [31], revealed that they all contain an N-terminal catalytic domain and a C-terminal substrate-binding domain connected by a long straight or broken α-helix; however, the overall topology is significantly different (Supplementary Figure S7). Although other DHH nucleases exist either as monomer or dimer in solution, they all function as monomers, and the active site is located at the interface of the catalytic domain and the substrate-binding domain within the same molecule (Supplementary Figure S7). Specifically, the DHH motif of the catalytic domain faces the substrate-binding domain of the same molecule to form the active site. The substrate-binding groove (or cleft) adopts a large and open conformation in the substrate-free state but transitions to a narrow and closed conformation upon substrate binding [31,36]. In contrast, MYG1 exists and functions as a dimer in both solution and structure with a very tight dimerization interface. Specifically, the DHH motif of the catalytic domain is oriented away from the substrate-binding domain of the same monomer and instead faces toward the substrate-binding domain of an adjacent monomer to form the active site. Consequently, the substrate-binding groove is narrow and partially covered by structural elements of the substrate-binding domain.

The structural differences in the substrate-binding groove may explain the unique substrate preference of each nuclease in the DHH family. For instance, the substrate-binding groove of Hef-associated nuclease (HAN) is relatively large, making it suitable for binding long oligonucleotides [10]. This explains its high activity on long single-stranded substrates but low activity on short- or double-stranded substrates. Additionally, HAN can degrade both ssDNA and ssRNA, indicating that it lacks specific recognition residues to discriminate the ribose 2′-OH group [10]. RecJ can only hydrolyze long ssDNA but not ssRNA because a conserved tyrosine residue (Tyr114) is positioned near the binding site of the deoxyribose of the first nucleotide, which would cause steric clash with the 2-OH group in ribose, explaining its specificity for ssDNA [8]. In contrast, bacterial NrnA can hydrolyze short ssRNA, pAp, and c-di-GMP, but not ssDNA [36]. In the NrnA structure, the catalytic domain and the substrate-binding domain are linked by a broken α-helix but are positioned separately, forming a widely open substrate-binding groove. Upon substrate binding, the substrate-binding groove undergoes conformational changes to accommodate different types of substrates. This structural feature accounts for its preference for short nanoRNA, c-di-GMP, and pAp [36,37]. The substrate-binding groove in MYG1 is narrow and partially covered by structural elements of the substrate-binding domain, allowing only single-stranded nucleic acids to enter and pass through the channel. This may explain its preference for ssRNA and ssDNA. Additionally, the specific recognition of the 2′-OH group in the ribose by residues V175, R80, and R302 mediates the differential activity of MYG1 toward ssRNA versus ssDNA. Nevertheless, our biochemical data show that in addition to ssRNA and ssDNA, MYG1 also exhibits weak activity toward double-stranded nucleic acids with different structures, suggesting that the substrate-binding groove of MYG1 may undergo conformational changes to adopt an open conformation for binding double-stranded nucleic acids. Further structural and biochemical studies are required to examine this possibility.

Our biochemical data also indicate that MYG1 possesses a certain degree of nucleotide preference, with varying cleavage activity toward different nucleotides, exhibiting a preference for U in ssRNA and dT in ssDNA. Typically, exonucleases do not exhibit sequence specificity; however, some nucleases display certain base preferences. For example, Dis3l2 shows a preference for polyU sequences through uracil-specific interactions spanning the first 12 nucleotides of a polyU tail [38]. The specific recognition of nucleotides by

nucleases often relies on hydrogen-bonding interaction and/or shape complementarity due to the varying shapes and functional groups of different nucleotide bases [39]. Our structural data indicate that MYG1 recognizes and binds to different nucleotides at the N1 and N2 sites via distinct hydrogen-bonding interactions, which may influence the product dissociation rate and/or substrate translocation rate, resulting in varying cleavage activity toward different nucleotides.

Although DHH family nucleases are generally considered to utilize a two-metal ion-dependent catalytic mechanism, some members (*e.g.*, *Pyrococcus furiosus* RecJ and the NrnA homolog MSMEG_2630) have been observed with only one metal ion at the active site in their crystal structures [40,41]. It was suggested that the binding of the second metal ion may require the presence of a substrate or specific catalytic conditions [41]. In this study, two Mn^{2+} ions were observed to bind to the active site of MYG1. The structural and biochemical data indicate that MYG1 utilizes a two-metal ion-dependent catalytic mechanism. Specifically, Mn1 is responsible for activating a water molecule to generate the nucleophile, and Mn2 stabilizes the transition state and the leaving product. Interestingly, mutation of H50, a residue involved in coordinating Mn2, disrupts Mn2 binding and weakens product binding, resulting in slightly enhanced enzymatic activity. These results indicate that Mn1 is essential for the catalytic activity of MYG1, but Mn2 is not. In other words, MYG1 can exert catalytic activity using a single metal ion-dependent mechanism. This finding could be exploited in the design of engineered DHH nucleases with novel substrate specificity and altered catalytic activity.

Our biochemical and structural data identify key residues involved in metal ion binding and substrate binding. Sequence alignment of MYG1 homologs spanning eukaryotes to prokaryotes (including *Homo sapiens*, *Mus musculus*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, and *Agrobacterium fabrum*) shows a moderate level of sequence conservation (identity: 13.2%, similarity: 67.3%). Notably, the key residues involved in metal ion binding, including H50, H55, D57, E58, D93, H107, and D176, as well as those involved in substrate binding, including H108, F313, F342, and F348, are also highly conserved across these species (Supplementary Figure S8). These findings suggest that MYG1 employs a conserved substrate binding and catalytic mechanism across evolutionarily distant species. However, the alignment also reveals that residues implicated in ribose and base recognition, specifically T53, R80, V175, R302, R327, and S346, exhibit lower conservation. This divergence implies species-specific variations in substrate preference, potentially enabling MYG1 homologs to participate in distinct biological pathways. Collectively, sequence analyses indicate that while the catalytic mechanism of MYG1 is likely conserved among its orthologs, the protein may have acquired divergent biological functions in different species through adaptive evolution of substrate-selective residues.

DHH family proteins play crucial roles in important cellular processes, such as DNA replication, recombination, repair, and nucleic acid metabolism, in both prokaryotes and eukaryotes. However, current knowledge about the physiological function of MYG1 in eukaryotic cells is very limited. Previous studies have shown that MYG1 participates in ribosomal RNA processing and influences mitochondrial RNA metabolic homeostasis [14] Additionally, it was reported that knockdown of *MYG1* inhibits proliferation in tumor cells [23]. Nevertheless, the specific signaling pathways in which MYG1 is involved and its regulatory mechanisms remain to be elucidated. The structural and biochemical features revealed in this study provide an important molecular basis for further exploration of its biological functions.

Supplementary Data

Supplementary data are available at *Acta Biochimica et Biophysica Sinica* online.

Data Availability

The crystal structures of the MYG1-Na, MYG1-Mn, MYG1-Mn-AMP, MYG1-Mn-GMP, MYG1-Mn-CMP, MYG1-Mn-UMP, MYG1-Mn-dAMP, MYG1-Mn-dGMP, MYG1-Mn-dCMP, MYG1-Mn-dTMP, H50A-Mn-AMP, H55A-Mn-AMP, D57A-Mn-AMP, E58A-Mn-AMP, D93A-Mn-AMP, H107A-Mn-AMP, D176A-Mn-AMP, MYG1-Mn-d(AC), H50A-Mn-d(AC), D57A-Mn-UUUU, and D93A-Mn-UU complexes have been deposited in the Protein Data Bank with accession codes 24AH, 24AJ, 24AN, 24AO, 24AQ, 24AP, 24BS, 24BU, 24AW, 24AY, 24BX, 24BZ, 24BD, 24BL, 24CA, 24BM, 24BV, 24BP, 24BQ, 24BR, 24BW, respectively. Source data are provided with this paper. All other data can be accessed upon reasonable request from the corresponding authors.

Author Contributions

C.L. conducted the biochemical and structural studies and data analysis. Z.C. contributed to the structure determination and data analysis. G.W. participated in data analysis and discussion. J.D. conceived the study, contributed to the experimental design and data analysis, and wrote the manuscript.

Acknowledgement

We would like to thank the staff at beamlines BL02U1 and BL10U2 of the Shanghai Synchrotron Radiation Facility (SSRF) and BL18U1 and BL19U1 of the National Facility for Protein Science in Shanghai (NFPSS) for their technical assistance in diffraction data collection. We also appreciate the technical support from the Core Facility for Molecular Biology at the Center for Excellence in Molecular Cell Science, as well as valuable discussion with other members of our group.

Funding

This work was supported by the grants from the National Key Research and Development Program of China (Nos. 2020YFA0509000 and 2020YFA0803200) and the National Natural Science Foundation of China (No. 32071190).

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Nelson DL, Cox MM. *Lehninger principles of biochemistry*. Eighth edition. ed. 2021, Austin: Macmillan Learning
2. Cochrane JC, Strobel SA. Catalytic strategies of self-cleaving ribozymes. *Acc Chem Res* 2008, 41: 1027–1035
3. Déclais AC, Lilley DM. New insight into the recognition of branched DNA structure by junction-resolving enzymes. *Curr Opin Struct Biol* 2008, 18: 86–95
4. Schoeffler AJ, Berger JM. DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. *Quart Rev Biophys* 2008, 41: 41–101

5. Yang W. Nucleases: diversity of structure, function and mechanism. *Quart Rev Biophys* 2011, 44: 1–93
6. Aravind L, Koonin EV. A novel family of predicted phosphoesterases includes *Drosophila* prune protein and bacterial recJ exonuclease. *Trends Biochem Sci* 1998, 23: 17–19
7. Srivastav R, Sharma R, Tandon S, Tandon C. Role of DHH superfamily proteins in nucleic acids metabolism and stress tolerance in prokaryotes and eukaryotes. *Int J Biol Macromol* 2019, 127: 66–75
8. Cheng K, Xu H, Chen X, Wang L, Tian B, Zhao Y, Hua Y. Structural basis for DNA 5'-end resection by RecJ. *eLife* 2016, 5: e14294
9. Fang M, Zeisberg WM, Condon C, Ogryzko V, Danchin A, Mechold U. Degradation of nanoRNA is performed by multiple redundant RNases in Bacillus subtilis. *Nucleic Acids Res* 2009, 37: 5114–5125
10. Feng L, Chang CC, Song D, Jiang C, Song Y, Wang CF, Deng W, ***et al.*** The trimeric Hef-associated nuclease HAN is a 3′→5′ exonuclease and is probably involved in DNA repair. *Nucleic Acids Res* 2018, 46: 9027–9043
11. D'Angelo A, Garzia L, André A, Carotenuto P, Aglio V, Guardiola O, Arrigoni G, ***et al.*** Prune cAMP phosphodiesterase binds nm23-H1 and promotes cancer metastasis. *Cancer Cell* 2004, 5: 137–149
12. Fenwick AL, Kliszczak M, Cooper F, Murray J, Sanchez-Pulido L, Twigg SRF, Goriely A, ***et al.*** Mutations in CDC45 encoding an essential component of the pre-initiation complex, cause Meier-Gorlin syndrome and Craniosynostosis. *Am J Hum Genet* 2016, 99: 125–138
13. Zollo M, Ahmed M, Ferrucci V, Salpietro V, Asadzadeh F, Carotenuto M, Maroofian R, ***et al.*** PRUNE is crucial for normal brain development and mutated in microcephaly with neurodevelopmental impairment. *Brain* 2017, 140: 940–952
14. Grover R, Burse SA, Shankrit S, Aggarwal A, Kirty K, Narta K, Srivastav R, ***et al.*** Myg1 exonuclease couples the nuclear and mitochondrial translational programs through RNA processing. *Nucleic Acids Res* 2019, 47: 5852–5866
15. Philips M-, Vikeså J, Luuk H, Jønson L, Lilleväli K, Rehfeld JF, Vasar E, ***et al.*** Characterization of MYG1 gene and protein: subcellular distribution and function. *Biol Cell* 2009, 101: 361–377
16. Chan WY, Lee TL, Wu SM, Ruszczyk L, Alba D, Baxendale V, Rennert OM. Transcriptome analyses of male germ cells with serial analysis of gene expression (SAGE). *Mol Cell Endocrinol* 2006, 250: 8–19
17. Golan-Mashiach M, Dazard JE, Gerecht-Nir S, Amariglio N, Fisher T, Jacob-Hirsch J, Bielorai B, ***et al.*** Design principle of gene expression used by human stem cells: Implication for pluripotency. *FASEBJ* 2005, 19: 147–149
18. Philips MA, Abramov U, Lilleväli K, Luuk H, Kurrikoff K, Raud S, Plaas M, ***et al.*** Myg1-deficient mice display alterations in stress-induced responses and reduction of sex-dependent behavioural differences. *Behav Brain Res* 2010, 207: 182–195
19. Dwivedi M, Laddha NC, Begum R. Correlation of increased MYG1 expression and its promoter polymorphism with disease progression and higher susceptibility in vitiligo patients. *J Dermatological Sci* 2013, 71: 195–202
20. Kingo K, Philips MA, Aunin E, Luuk H, Karelson M, Rätsep R, Silm H, ***et al.*** MYG1, novel melanocyte related gene, has elevated expression in vitiligo. *J Dermatological Sci* 2006, 44: 119–122
21. Philips MA, Kingo K, Karelson M, Rätsep R, Aunin E, Reimann E, Reemann P, ***et al.*** Promoter polymorphism -119C/G in MYG1 (C12orf10) gene is related to vitiligo susceptibility and Arg4Gln affects mitochondrial entrance of Myg1. *BMC Med Genet* 2010, 11: 56
22. Traks T, Keermann M, Karelson M, Rätsep R, Reimann E, Silm H, Vasar E, ***et al.*** Polymorphisms in melanocortin system and *MYG 1* genes are associated with vitiligo. *Acad Dermatol Venereol* 2019, 33: e65
23. Chen J, Duan S, Wang Y, Ling Y, Hou X, Zhang S, Liu X, ***et al.*** MYG1 drives glycolysis and colorectal cancer development through nuclear-mitochondrial collaboration. *Nat Commun* 2024, 15: 4969
24. Han X, Li A, Wang W, Du L, Wang C, Huang G. MYG1 promotes proliferation and inhibits autophagy in lung adenocarcinoma cells via the AMPK/mTOR complex 1 signaling pathway. *Oncol Lett* 2021, 21: 334
25. Liu X, Zhu Y, Huang W, Chen J, Lan J, Long X, Zhou J. MYG1 interacts with HSP90 to promote breast cancer progression through Wnt/β-catenin and Notch signaling pathways. *Exp Cell Res* 2025, 446: 114448
26. Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T, ***et al.*** Data processing and analysis with the *autoPROC* toolbox. *Acta Crystlogr D Biol Crystlogr* 2011, 67: 293–302
27. Otwinowski, Z.W. Minor Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 1997, 276: 307–26
28. Kabsch W. *XDS*. *Acta Crystlogr D Biol Crystlogr* 2010, 66: 125–132
29. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, ***et al.*** *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystlogr D Biol Crystlogr* 2010, 66: 213–221
30. Emsley P, Cowtan K. *Coot*: model-building tools for molecular graphics. *Acta Crystlogr D Biol Crystlogr* 2004, 60: 2126–2132
31. Ahn S, Milner AJ, Fütterer K, Konopka M, Ilias M, Young TW, White SA. The "open" and "closed" structures of the type-C inorganic pyrophosphatases from *Bacillus subtilis* and *Streptococcus gordonii*. *J Mol Biol* 2001, 313: 797–811
32. Wakamatsu T, Kitamura Y, Kotera Y, Nakagawa N, Kuramitsu S, Masui R. Structure of RecJ exonuclease defines its specificity for single-stranded DNA. *J Biol Chem* 2010, 285: 9762–9769
33. Adámková K, Trundová M, Kovaľ T, Husťáková B, Kolenko P, Dušková J, Skálová T, ***et al.*** Substrate preference, RNA binding and active site versatility of *Stenotrophomonas maltophilia* nuclease SmNuc1, explained by a structural study. *FEBS J* 2025, 292: 129–152
34. van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, Söding J, ***et al.*** Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* 2024, 42: 243–246
35. Yamagata A, Kakuta Y, Masui R, Fukuyama K. The crystal structure of exonuclease RecJ bound to Mn^{2+} ion suggests how its characteristic motifs are involved in exonuclease activity. *Proc Natl Acad Sci USA* 2002, 99: 5908–5912
36. Schmier BJ, Nelersa CM, Malhotra A. Structural basis for the bidirectional activity of Bacillus nanoRNase NrnA. *Sci Rep* 2017, 7: 11085
37. He Q, Wang F, Liu S, Zhu D, Cong H, Gao F, Li B, ***et al.*** Structural and biochemical insight into the mechanism of Rv2837c from *Mycobacterium tuberculosis* as a c-di-NMP phosphodiesterase. *J Biol Chem* 2016, 291: 14386–14387
38. Faehnle CR, Walleshauser J, Joshua-Tor L. Mechanism of Dis3l2 substrate recognition in the Lin28–let-7 pathway. *Nature* 2014, 514: 252–256
39. Coulocheri SA, Pigis DG, Papavassiliou KA, Papavassiliou AG. Hydrogen bonds in protein-DNA complexes: where geometry meets plasticity. *Biochimie* 2007, 89: 1291–1303
40. Li MJ, Yi GS, Yu F, Zhou H, Chen JN, Xu CY, Wang FP, ***et al.*** The crystal structure of *Pyrococcus furiosus* RecJ implicates it as an ancestor of eukaryotic Cdc45. *Nucleic Acids Res* 2017, 45: 12551–12564
41. Srivastav R, Kumar D, Grover A, Singh A, Manjasetty BA, Sharma R, Taneja B. Unique subunit packing in mycobacterial nanoRNase leads to alternate substrate recognitions in DHH phosphodiesterases. *Nucleic Acids Res* 2014, 42: 7894–7910