7312–7320 *Nucleic Acids Research, 2009, Vol. 37, No. 21* *Published online 25 September 2009*
doi:10.1093/nar/gkp770

Crystal structure of the EndoG/EndoGI complex: mechanism of EndoG inhibition

Bernhard Loll[1], Maike Gebhardt[1], Elmar Wahle[2] and Anton Meinhart[1,*]

[1]Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Jahnstrasse 29, 69120 Heidelberg and [2]Institute of Biochemistry and Biotechnology, Martin-Luther-University Halle-Wittenberg, Kurt-Mothes-Strasse 3, 06120 Halle, Germany

Received August 7, 2009; Revised August 30, 2009; Accepted September 1, 2009

ABSTRACT

EndoG is a ubiquitous nuclease that is translocated into the nucleus during apoptosis to participate in DNA degradation. The enzyme cleaves double- and single-stranded DNA and RNA. Related nucleases are found in eukaryotes and prokaryotes, which have evolved sophisticated mechanisms for genome protection against self-antagonizing nuclease activity. Common mechanisms of inhibition are secretion, sequestration into a separate cellular compartment or by binding to protein inhibitors. Although EndoG is silenced by compartmentalization into the mitochondrial intermembrane space, a nucleus-localized protein inhibitor protects cellular polynucleotides from degradation by stray EndoG under non-apoptotic conditions in *Drosophila*. Here, we report the first three-dimensional structure of EndoG in complex with its inhibitor EndoGI. Although the mechanism of inhibition is reminiscent of bacterial protein inhibitors, EndoGI has evolved independently from a generic protein-protein interaction module. EndoGI is a two-domain protein that binds the active sites of two monomers of EndoG, with EndoG being sandwiched between EndoGI. Since the amino acid sequences of eukaryotic EndoG homologues are highly conserved, this model is valid for eukaryotic dimeric EndoG in general. The structure indicates that the two active sites of EndoG occupy the most remote spatial position possible at the molecular surface and a concerted substrate processing is unlikely.

INTRODUCTION

Endonuclease G (EndoG) is a nuclear encoded, sugar-non-specific nuclease highly conserved among eukaryotes (1–3). Bacterial homologues of EndoG have been described as well, for example in *Anabaena* and *Serratia marcescens* (4,5). The bacterial homologues of EndoG are secreted from the cell and probably serve a nutritional function by providing precursors for nucleic acid biosynthesis or act as bactericides to inhibit growth of competing organisms. Related enzymes also serve as virulence factors during lung infection (6,7). In eukaryotes, EndoG serves a different function. The enzyme was originally proposed to be involved in mitochondrial DNA replication (8) and may be involved in the initiation of genomic inversions in Herpes Simplex type-1 virus (9). EndoG is targeted to the mitochondrial intermembrane space, and is now thought to be released from mitochondria upon induction of apoptosis together with cytochrome *c* and other proapoptotic proteins (1,3,10–12). Upon release, EndoG is translocated into the nucleus and helps to degrade chromosomal DNA (3,12). The role of EndoG in apoptosis is still ambiguous since an EndoG knock-out in mice did not show impaired apoptotic DNA degradation (13,14). However, the lack of apoptotic DNA degradation might also be explained by redundancy with additional nucleases (15). The major enzyme responsible for apoptotic DNA degradation is the caspase-activated DNase (CAD), which is regulated through association with its inhibitor ICAD. ICAD serves a dual function: it has a chaperone-like role, since no functional CAD is produced without ICAD. Moreover, ICAD functions as an inhibitor for CAD; it is cleaved by caspases upon induction of apoptosis and thereby releases activated CAD.

Bacteria have evolved sophisticated mechanisms of protection against self-antagonizing activity of EndoG-like nucleases. The activity of the *Serratia* nuclease depends on the formation of cystine bonds, which occurs only during secretion (16). *Anabaena* possesses a protein inhibitor (NuiA) that forms a metal ion bridge to the nuclease NucA (17). Similarly, cytoplasmatic protein inhibitors for other secreted nucleases such as pancreatic RNase (18), barnase (19) or colicin E3 (20) have been described,

*To whom correspondence should be addressed. Tel: +49 6221 486 505; Fax: +49 6221 486 585; Email: anton.meinhart@mpimf-heidelberg.mpg.de
The atomic coordinates and structure factor amplitudes have been deposited in the Protein Data Bank (3ISM).

suggesting that bacteria as well as eukaryotes have backup systems to protect cellular nucleic acids from sorting failures.

An inhibitor of EndoG (EndoGI) is encoded in the genomes of different *Drosophila* species (21). Most EndoGI is located in the nucleus, separated from the mitochondrial EndoG, and has been suggested to act as life insurance for the cell against stray EndoG produced by failure of mitochondrial import or leakage from the organelle. EndoGI binds to EndoG with subpicomolar affinity, and binding is dependent on the active site architecture of the nuclease. EndoGI consists of two highly similar domains linked within one polypeptide chain. Each domain is able to inhibit EndoG on its own, and, accordingly, analytical ultracentrifugation showed that one molecule of EndoGI binds two EndoG molecules (21). EndoG from *Drosophila* as well as other EndoG variants from different eukaryotes exist as homodimers (2,21,22), and each inhibitor domain of EndoGI most probably binds to the active site of one EndoG monomer (21).

The active site of EndoG displays a ββα-metal finger topology, harboring conserved residues involved in metal ion binding and nuclease activity (2,23). Based on sequence similarity with related bacterial nucleases, homology models have been proposed for EndoG. However, contradicting results concerning the active site architecture were obtained, requiring structural investigations of EndoG.

We therefore crystallized EndoG from *Drosophila melanogaster* in the presence of the inhibitor EndoGI. For the first time, we report the three-dimensional structure of eukaryotic EndoG. We demonstate that EndoGI inhibits EndoG by blocking the active site and the oligonucleotide-binding groove. EndoGI is structurally unrelated to the previously described nuclease inhibitors and most probably has evolved independently of any ancestral bacterial inhibitor.

MATERIALS AND METHODS

Protein expression and purification

Full length EndoGI (dEndoGI) and N-terminally truncated EndoG (lacking the mitochondrial localization sequence and a short unstructured polypeptide stretch; further referred to as dEndoG), both from *D. melanogaster,* were co-expressed from a bicistronic expression construct (21). *Escherichia coli* C43(DE3) cells (Lucigen) transformed with the bicistronic expression construct were grown at 310 K in LB-medium containing 100 mg/l ampicillin to an OD_{600} of 0.6, cooled to 291 K and protein expression was induced with 0.5 mM IPTG for 16 h. Harvested cells were resuspended in buffer A (50 mM Tris–HCl pH 8.0, 250 mM NaCl, 2 mM $MgSO_4$ and 10 mM 2-mercaptoethanol). Cell walls were broken by sonication and the cleared supernatant was loaded onto a 1 ml HisTrap HP column (GE Healthcare) equilibrated with buffer A. Proteins were eluted as a complex in a gradient to buffer A containing additional 250 mM imidazole. The dEndoG/dEndoGI complex was further purified by size exclusion chromatography on a Superdex^{TM}75 column (GE Healthcare) equilibrated with 50 mM Hepes–NaOH pH 7.5, 200 mM NaCl, 2 mM $MgSO_4$ and 2 mM dithioerythritol. For crystallization, the protein complex was concentrated to 17 mg/ml as determined by the BioRad Protein Assay.

For nuclease activity assays, the dEndoG/dEndoGI complex was incubated with 10 mM EDTA and bound to MonoQ column (GE Healthcare) equilibrated with buffer B (50 mM Tris–HCl pH 8.0 and 2 mM dithioerythritol) with additional 10 mM EDTA. While dEndoGI bound to the column, dEndoG was found in the flow-through, which was loaded onto a 1 ml HiTrap Heparin (GE Healthcare) equilibrated with buffer B subsequently. Pure dEndoG was eluted in a gradient to buffer B containing additional 1 M NaCl.

Nuclease activity assay

In limited digestion experiments, 7.5 fmol of dEndoG were incubated with 80 ng plasmid DNA in 30 µl reaction buffer (10 mM sodium phosphate buffer pH 6.0, 2 mM $MgSO_4$ and 2 mM dithioerythritol). Aliquots were taken at different time intervals and the reaction was stopped by addition of 10 mM EDTA. To determine the pH-dependence of dEndoG activity in reaction buffer at different pH, the reaction was stopped after 1 min incubation. Samples were electrophoretically separated over a 0.5% (w/v) agarose gel and stained with SYBR Gold nucleic acid stain (Invitrogen).

Crystallization, crystal harvesting and data collection

Crystals grew at 277 K in a vapor diffusion setup applying a reservoir solution containing 20% (w/v) PEG 2000 MME. Crystals were harvested into a cryo protectant solution containing 100 mM Tris–HCl pH 8.5 and 32% (v/v) PEG 400 and subsequently flash cooled in liquid nitrogen. Diffraction data were collected at the beamline X10SA at the Swiss Light Source (Switzerland) at 100 K and processed with XDS (24).

Phasing and structure refinement

Initial phases were determined by molecular replacement (25) using the structure of NucA (26). Model building was performed in cycles of manual building with COOT (27) and automatic refinement using the simulated annealing protocol from CNS (28) and REFMAC (29). At final stages of refinement, phases were improved with REFMAC using TLS refinement. Model quality was evaluated with PROCHECK (30) and MolProbity (31). Solvent-accessible surface areas were calculated with AREAIMOL (29). Figures were prepared using PYMOL (32), and electrostatic surface potentials were calculated with APBS (33). Sequence alignments were performed with CLUSTALW (34) and illustrated with ALSCRIPT (35).

Table 1. Data collection and refinement statistics

Data collection	Native
Space group	$P2_12_12_1$
Unit cell [a, b, c (Å)]	79.8, 109.2, 121.1
Wavelength (Å)	1.0075
Resolution (Å)[a]	43.2–2.2 (2.3–2.2)
No. reflections[a]	46 364 (5385)
Completeness (%)[a]	95.9 (90.5)
$<I/\sigma(I)>$[a]	14.8 (4.2)
R_{meas}[a,b]	8.7 (40.3)
Redundancy	4.1 (4.2)
Refinement	
R_{work}[a,c]	18.4 (20.1)
R_{free}[a,d]	22.3 (23.1)
Overall B factor (Å^2)	31.8
Chain A dEndoG (1998 atoms) (Å^2)	30.9
Chain B dEndoG (2051 atoms) (Å^2)	31.7
Chain C dEndoGI Dom1 (1280 atoms) (Å^2)	30.7
Chain C dEndoGI Dom2 (1089 atoms) (Å^2)	34.9
Magnesium (two atoms) (Å^2)	22.0
Buffer (one molecule) (Å^2)	44.2
Water (328 molecules) (Å^2)	31.3
R.m.s.d.[e] from ideal geometry, bond length (Å)	0.010
Bond angles (°)	1.150

[a]Values in parentheses refer to the highest resolution shell.
[b]$R_{meas} = \Sigma_h\ [n/(n-1)]^{1/2}\ \Sigma_i\ |I_h - I_{h,i}| / \Sigma_h \Sigma_i\ I_{h,I}$, where I_h is the mean intensity of symmetry-equivalent reflections and n is the redundancy.
[c]$R_{work} = \Sigma_h\ |F_o - F_c| / \Sigma\ F_o$ (working set, no σ cut-off applied).
[d]R_{free} is the same as R_{cryst}, but calculated on 5% of the data excluded from refinement.
[e]R.m.s.d. from target geometries.

RESULTS AND DISCUSSION

Overall structure of the EndoG/EndoGI complex

A model of the dEndoG/dEndoGI complex was refined against diffraction data to 2.2 Å resolution (data and refinement statistics are given in Table 1). Although some parts of the model were missing (see Supplementary Table S1), the model could be refined to a free R-factor of 22.3%. Overall the model has excellent stereochemistry with 91.5% of the residues lying in the most favored regions of the Ramachandran plot and 8% of the residues in additional allowed regions.

In agreement with previous studies (21), two dEndoG molecules are bound by one dEndoGI molecule in the crystal structure (Figure 1). The two dEndoG molecules form a compact dimer and are sandwiched between two domains of dEndoGI (further referred to as Dom1 and Dom2). The linker connecting the two inhibitor domains, however, could not be modeled into the residual ambiguous electron density. Thus, the physical connection of the two domains remains elusive (Supplementary Table S1). The termini of the two domains that need to be connected are separated by 65 Å, a distance that could easily be spanned by the missing 44 residues. Consequently, one dEndoGI molecule can inhibit both nuclease domains, without steric interference of the homodimer arrangement of dEndoG. Finally, MALDI-TOF experiments performed on dissolved protein crystals revealed that the EndoGI polypeptide chain was intact. Proteolytic cleavage of dEndoGI into two separate domains can therefore be excluded.

Architecture of homodimeric EndoG

A Dali search (36) revealed that the mixed αβ fold of dEndoG closely resembles NucA from *Anabaena* (26) and *Serratia* nuclease (37). Like the related bacterial endonucleases, dEndoG contains a central twisted antiparallel β-sheet consisting of six β-strands flanked by helices and a second short β-sheet formed by strands βD and βE. In contrast to the bacterial nucleases, strands βG and βH of dEndoG form a third, small two-stranded antiparallel β-sheet (wing) which protrudes from the globular domain (Figures 1 and 2A). Interestingly, the wings of both monomers dimerize and form an intermolecular β-sheet. Additional residues in adjacent loop regions contribute to dimer formation (Figure 2A and Supplementary Table S2), and the total buried surface area involved in dimer formation is 1600 Å^2, 7.5% of the total solvent accessible surface. The domain interface is mainly hydrophilic, and two cavities filled with water molecules are buried in the domain interface. A superposition of the two dEndoG molecules revealed no significant differences, and root-mean-square deviation (r.m.s.d.) of 0.39 Å for all main chain atoms was calculated when the first six amino-terminal residues were excluded.

Residues forming the dEndoG dimer interface are conserved from yeast to human (Figure 2A and B). The dimer observed for dEndoG is thus a valid model for eukaryotic EndoG in general. This agrees with previous reports showing that EndoG from different organisms exists as homodimers *in vitro* (2,21,22). Based on the sequence homology to the *Serratia* nuclease, a different mode of dimer formation has been proposed (2). However, the β-wing forming the intermolecular β-sheet is not present in related bacterial endonucleases, and thus bacterial *Serratia* nuclease dimer formation is different. The wing was probably inserted later during evolution; it is conserved throughout all eukaryotic EndoG variants (Figure 2A).

The active site of EndoG is highly conserved

An additional conserved patch is located at the surface of dEndoG, opposite the dimer interface, where a hydrated magnesium ion is bound (Figure 2B and C). Mutational studies suggested that the highly conserved DRGH motif located in this conserved patch (spanning Asp152 to His155 in dEndoG) is involved in both metal ion and oligonucleotide binding and thus is the key motif for catalysis (2,23). A histidine residue was proposed to activate water breakage of the phosphodiester bond. Indeed, in the crystal structure, the Nε of His155 forms a hydrogen bond to water molecule 2 belonging to the first hydration shell of the magnesium ion (Figure 2C). Four additional water molecules are involved in the coordination of the catalytic metal ion: water molecule 3 forms a hydrogen bond to the side chain carbonyl function of Gln181 and both water molecules 4 and 5 are held in place by the carboxylate group of Glu195. Only water molecule 6 is not contacted by the protein moiety and is solvent exposed. A single bond between the protein and the metal ion is formed

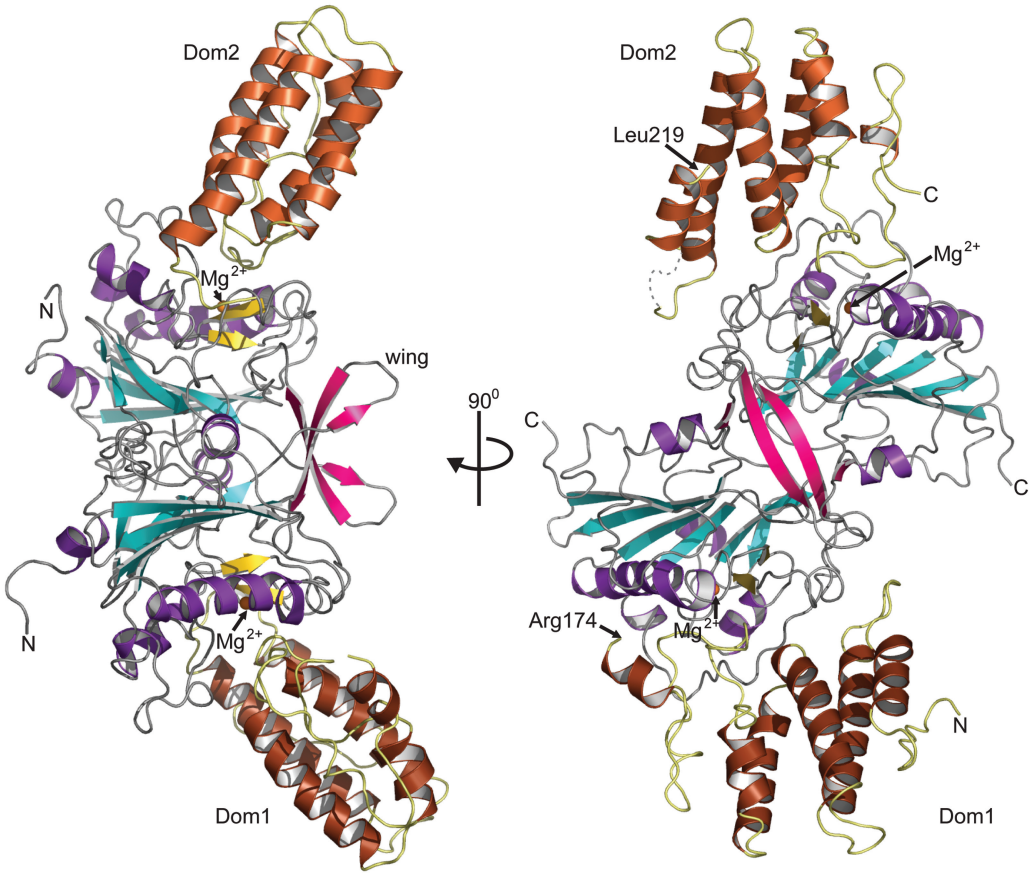

Figure 1. Overall structure of the dEndoG/dEndoGI complex. Helices of dEndoG are illustrated as purple ribbons and strands of the central β-sheet in cyan. The small two-stranded β-sheet (β-strands D and E) involved in metal ion binding is highlighted in yellow, and the bound metal ion is shown as an orange sphere. The two wings forming an intermolecular β-sheet are shown in red (β-strands G and H). Dom1 and Dom2 of Endo GI, which sandwich the dEndoG dimer, are shown as brown ribbons. The missing linker between the two domains would start at Arg174 in Dom1 and connect the polypeptide chain with Leu219 in Dom2.

by the side chain carbonyl function of Asn187 (Figure 2C).

In addition to residues involved in metal ion binding and activation of a water molecule, Arg57 of the *Serratia* nuclease was proposed to be required for catalysis by neutralizing the extra negative charge on the phosphate group in the transition state (38). However, a cysteine residue is located at the equivalent position in the sequences of dEndoG and its mammalian homologues. Mutation of this cysteine residue in bovine EndoG did not result in nuclease inactivation (2), suggesting that this residue is not involved in the transition state. The loop region harboring these residues is substantially different in the dEndoG and *Serratia* nuclease structures. A conserved arginine residue (Arg124 in dEndoG) is located further downstream in the sequences of different EndoG homologues and is placed in the active site of dEndoG at a similar position as observed for Arg57 the *Serratia* nuclease (Figure 2C and Supplementary Figure S1). Mutation of this arginine residue in human EndoG impairs nuclease activity (23), and thus Arg124 is probably involved in formation of the transition state in dEndoG.

Intriguingly, the active site of dEndoG contains several additional water molecules (Figure 2C) that are also found in the *Serratia* nuclease and NucA at equivalent positions (Supplementary Figure S1). It is tempting to speculate that these water molecules occupy positions where the phosphate backbone of the substrate is bound and that the eukaryotic and bacterial homologues share not only a highly conserved catalytic mechanism but also a conserved mode of oligonucleotide binding.

Furthermore, residues involved in the formation of the overall architecture of the active site are conserved in eukaryotic and bacterial endonucleases. Most prominent is Asp152 in dEndoG, located in the DRGH motif. One oxygen atom of the carboxylate group is the acceptor for a hydrogen bond from the amide of the side chain of Asn187, which binds to the metal ion (Figure 2C). Mutation of this aspartate residue to alanine in bovine EndoG impaired nuclease activity (2). The human EXOG homologue contains a serine residue at the equivalent position in the amino acid sequence, and a replacement by asparate apparently does not influence the activity (39). *In silico* mutation of Asp152 to serine in dEndoG revealed that this mutation can establish a

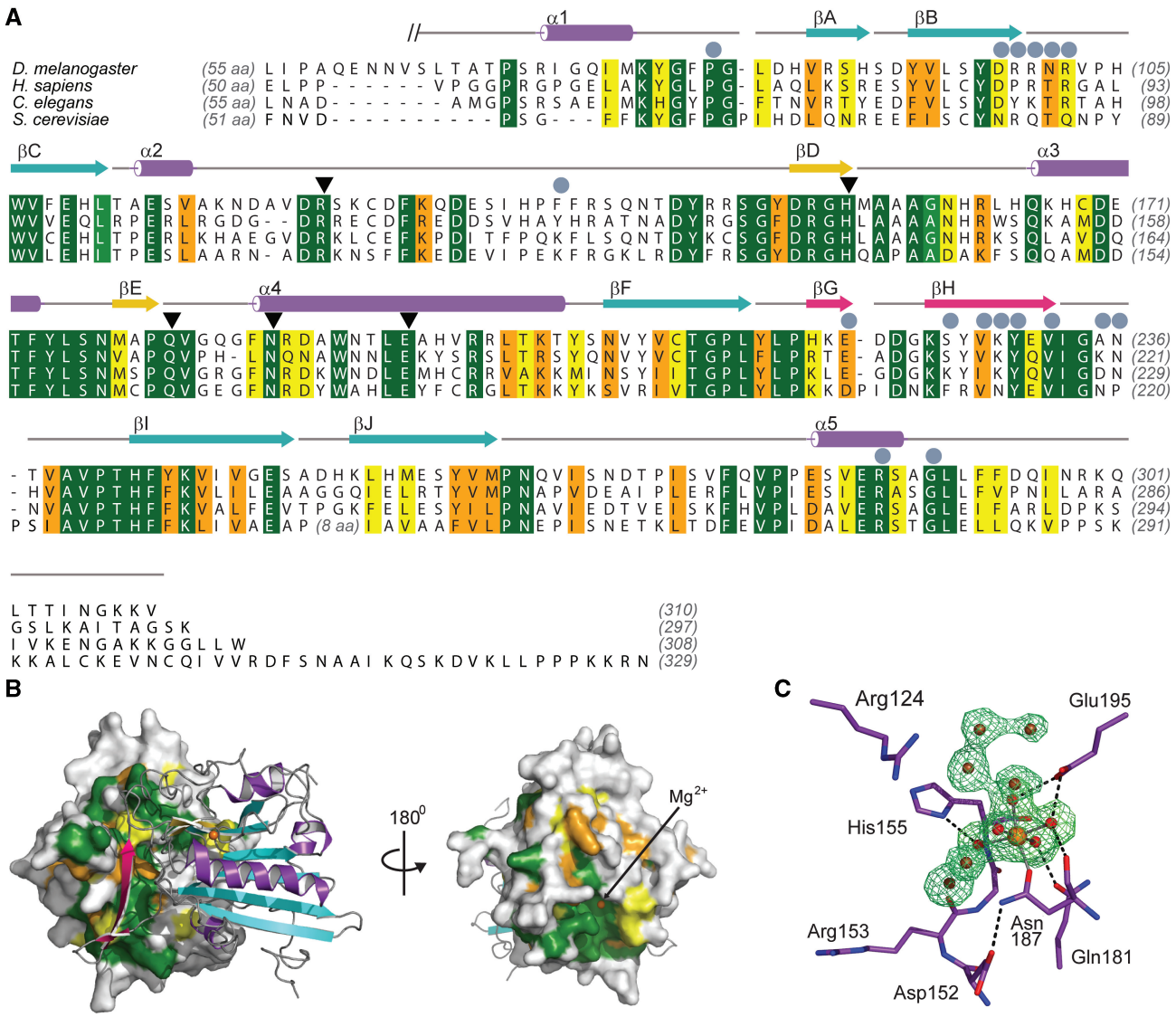

Figure 2. Sequence conservation of EndoG. (**A**) Sequence alignment of EndoG homologues for *D. melanogaster* (NP_610737), *Homo sapiens* (NP_004426), *Caenorhabditis elegans* cps–6 (NP_491371) and *S. cerevisiae* Nuc1p (NP_012327). Identical residues are colored dark green and according to the decreasing similarity from light green through orange to yellow. Secondary structure elements are colored according to Figure 1. Helices are drawn as cylinders and strands as arrows. Residues which are involved in formation of the dEndoG homodimer interface are labeled with gray circles and residues directly involved in catalysis are marked with black triangles. (**B**) Conserved patches at the molecular surface of dEndoG. Left: A surface representation of one dEndoG monomer is colored according the amino acid sequence alignment in (A), and the second molecule is drawn as a ribbon model similarly to Figure 1. Right: the dEndoG homodimer turned by 180°, showing the conserved patch around the active site of dEndoG. (**C**) Active site of dEndoG. Residues involved in metal ion coordination and/or catalytic function are shown as stick models, the central Mg^{2+} ion as an orange sphere and water molecules from the metal ion hydration shell and solvent water molecules are shown as red spheres. An electron density map contoured at 3.0 σ for the omitted metal ion and water molecules is shown as a green mesh.

similar hydrogen bond pattern, and thus the active site configuration is probably maintained.

Contradicting proposals for residues that bind metal ions have been made previously, and the residue equivalent to Glu286 in dEndoG was proposed to be involved in coordination of a second metal binding site in the human homologue (23). However, Glu286 forms a salt bridge within the homodimer interface, and no additional metal ion could be identified although 2 mM Mg^{2+} was present in the crystallization setup. Thus Glu286 is more likely a structural residue and not directly involved in metal ion cofactor binding and catalysis.

The oligonucleotide-binding site of EndoG

Conserved surface exposed residues cluster and line a positively charged groove around the metal ion binding site (Figures 2B and 3). A superposition of dEndoG with the related ββα-metal finger nuclease Vvn bound to DNA (40) reveals that residues lining this positively charged groove are probably involved in oligonucleotide binding. The majority of them are arginine residues (Arg124, Arg153 of the DRGH motif, Arg163 and Arg199) with their guanidinium groups lining the central part of the positively charged groove. Additional positively charged residues more distal from the active site extend this

putative oligonucleotide-binding site to 30 Å in its longest dimension.

The two oligonucleotide-binding grooves of EndoG are at opposing sides of the homodimer and do not form an extended intermolecular oligonucleotide-binding site (Figure 1). This corresponds to the active site arrangement in the *Serratia* nuclease even though the mechanism of dimerization differs. Given the distance and orientation of the two active sites, they cannot bind to and process double-stranded oligonucleotides simultaneously without dramatic bending of the double strand. It is more likely that the two active sites process substrates independently of each other. Such an independent mechanism is further supported by our observation that in the crystal structure the two monomers are identical and a cooperative action of the individual active sites seems rather unlikely. Finally, limited digestion of supercoiled (sc) plasmid DNA by dEndoG primarily produced open closed (oc) intermediates (Figure 4). Denaturation of these samples revealed that both strands had accumulated a large number of cuts; unit length DNA strands were not visible. This is inconsistent with the two active sites making two concerted, closely spaced cuts in both strands, as this would result in linearization of the DNA without oc intermediates. Thus, the enzyme seems to cut stochastically within the two strands, and linearization does not result until independent cuts in the complementary strands are sufficiently close together that base pairing can no longer maintain the double-stranded structure. Similar to bovine EndoG (2), dEndoG has optimal activity around pH 6.0 (Figure 4C). Notably, at pH 7.0 or higher the amount of short intermediates is reduced and both the oc and linearized form accumulate. Thus, dEndoG seems to function differently than the related bacterial *Serratia* nuclease, which does not produce an oc form of plasmid DNA (41). As the final oligonucleotide products of dEndoG only accumulate upon longer incubation, they are also derived from random cutting, i.e. the two active sites in one homodimer do not cut simultaneously within one stand and do not act as molecular rulers.

Overall structure of dEndoGI

Each domain of dEndoGI is composed of α-helix bundles forming the hydrophobic core, which is sealed by an N-terminal lid helix and a C-terminal extension almost totally lacking secondary structure elements (Figure 1 and Supplementary Figure S2). The two domains are homologous to each other and are also highly conserved among different *Drosophila* species (Supplementary Figure S2). In contrast, the flexible linker connecting the two domains is less well conserved. The solenoid fold of each domain is reminiscent of that of HEAT repeats. Dom1 and Dom2 are structurally very similar: a superposition revealed an r.m.s.d. of only 1.0 Å for all main chain atoms, neglecting a loop region distal from the nuclease binding site (residues 78–81 in Dom1 and 264–266 in Dom2). This loop region differs in length and geometry of the polypeptide chain.

A DALI search of Dom1 and Dom2 (36) revealed that the structurally most closely related domain is the C-terminal domain of the ATP-dependent helicase Ku80 (42). Based on its solenoid fold, this domain has been proposed to be involved in protein–protein interactions. Although there is no obvious reason to suspect a functional connection of dEndoGI and Ku80, we note that Ku80 was also identified in an amino acid sequence search when the sequence of Dom1 was used as query. Further structurally related proteins are the subunit VPS35 of the cargo-recognition complex (43), a yeast importin β (44).

The inhibitor of the bacterial EndoG homologue NucA, NuiA, has a distinctly different fold, containing a central β-sheet flanked by helices (17,45). Thus, dEndoGI has probably evolved from a protein-protein interaction module independently of NuiA. Although not structurally related to dEndoGI, the pancreatic RNase inhibitor is also composed of tandem repeats. In the case of the RNase inhibitor, leucine-rich repeats form a horseshoe like architecture (18). Apparently, dEndoG and pancreatic RNase inhibitor have evolved from different protein–protein interaction modules that have adapted to become tight binders and, eventually, inhibitors.

Inhibition of dEndoG by dEndoGI

One of the ridges formed by loop regions connecting the α-helix repeats of dEndoGI is rich in acidic residues which form a highly negatively charged patch at the molecular surface of Dom1 and Dom2 (Figure 3). These negatively charged surface patches seal the putative oligonucleotide-binding groove of each dEndoG monomer by electrostatic interactions. Each domain covers roughly 6% of the entire dEndoG dimer surface. A plethora of salt bridges are formed between carboxylate groups from the ridge of dEndoGI to guanidinium groups of dEndoG (Supplementary Table S3). Thus, dEndoGI not only forms a simple antipole that binds the positively charged oligonucleotide-binding grooves; rather, the negatively charged side chains are likely to mimic the phosphate

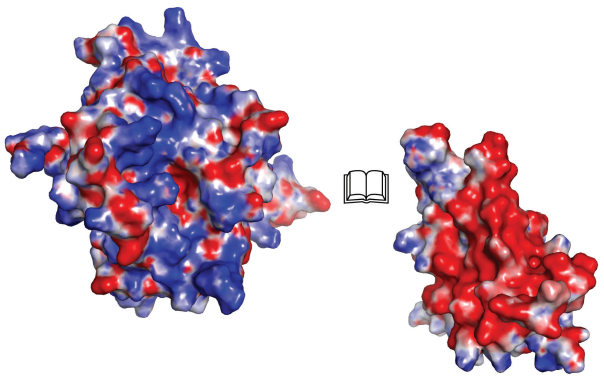

Figure 3. Bookview of the EndoG/EndoGI complex. For simplicity, only one dEndoG monomer (left) and the corresponding Dom1 of dEndoGI (right) are shown as surface representations. Coloring is according to the electrostatic surface potential over the range of ±9 kT/e (blue/red).

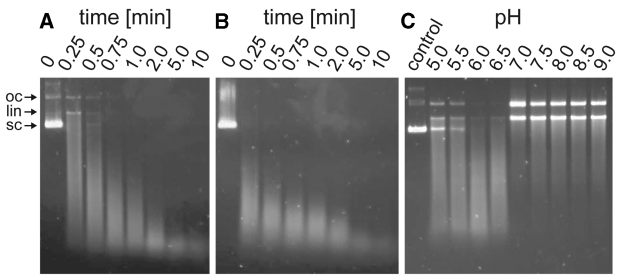

Figure 4. Nuclease activity assays of dEndoG. (**A**) Limited digestion of plasmid DNA. (**B**) Same sample as in (A) after melting. (**C**) pH-profile of dEndoG nuclease activity. The reaction was stopped after 1 min of incubation.

groups of the oligonucleotide substrate. Indeed, a superposition of dEndoG with the Vvn nuclease bound to DNA (40) showed that the position of active site proximal carboxylate groups of the two domains of dEndoGI superimpose with the phosphate groups of the DNA in the Vvn structure (Supplementary Figure S3). This observation not only could serve as a guide for docking an oligonucleotide substrate to EndoG, but also the structure of the dEndoG/dEndoGI complex could serve as a starting point for the *de novo* design of inhibitors for this type of nucleases, similarly as recently suggested for the *Anabaena* NucA/NuiA complex (17).

Another feature in common to the inhibitory actions of dEndoGI and NuiA is that both inhibitors are involved in metal ion coordination. Whereas dEndoGI blocks access to the active site by binding to a water molecule of the metal ion hydration shell (Figure 5), the C-terminal residue of NuiA displaces a water molecule from the hydration shell of the metal ion and establishes a direct metal ion bridge (17). Binding to the two active sites is likely the key determinant for the affinity of dEndoGI to dEndoG, since chelating the metal ion cofactor leads to disassembly of the otherwise stable dEndoGI/dEndoG complex (see 'Materials and Methods' section). Further support for this comes from previous mutational studies showing that mutation of the metal ion coordinating Asn187 in dEndoG (Figure 5) impaired binding of dEndoGI (21). Finally, the latter report described that mutation of the DRGH motif to DAGA totally abolished binding of dEndoG to dEndoGI, supporting that additional salt bridges formed between carboxylate groups of dEndoGI (Asp145 in Dom1 and Asp330 in Dom2) and Arg153 of each dEndoG (Figure 5) significantly contribute to affinity. Further salt bridges are formed between the carboxylate group of Asp151 in Dom1 and Asp336 in Dom2 to the guanidinium group of Arg124 of dEndoG, which is presumably involved in the formation of the transition state.

When Dom1 and Dom2 are compared, striking differences in binding to the hydrated metal ion (Figure 5) are observed. In Dom1, a solvent water molecule is sandwiched between the carboxylate group of Asp148 and the solvent exposed water molecule from the metal ion hydration shell. In a bifurcated manner, this water molecule binds a second, additional solvent water

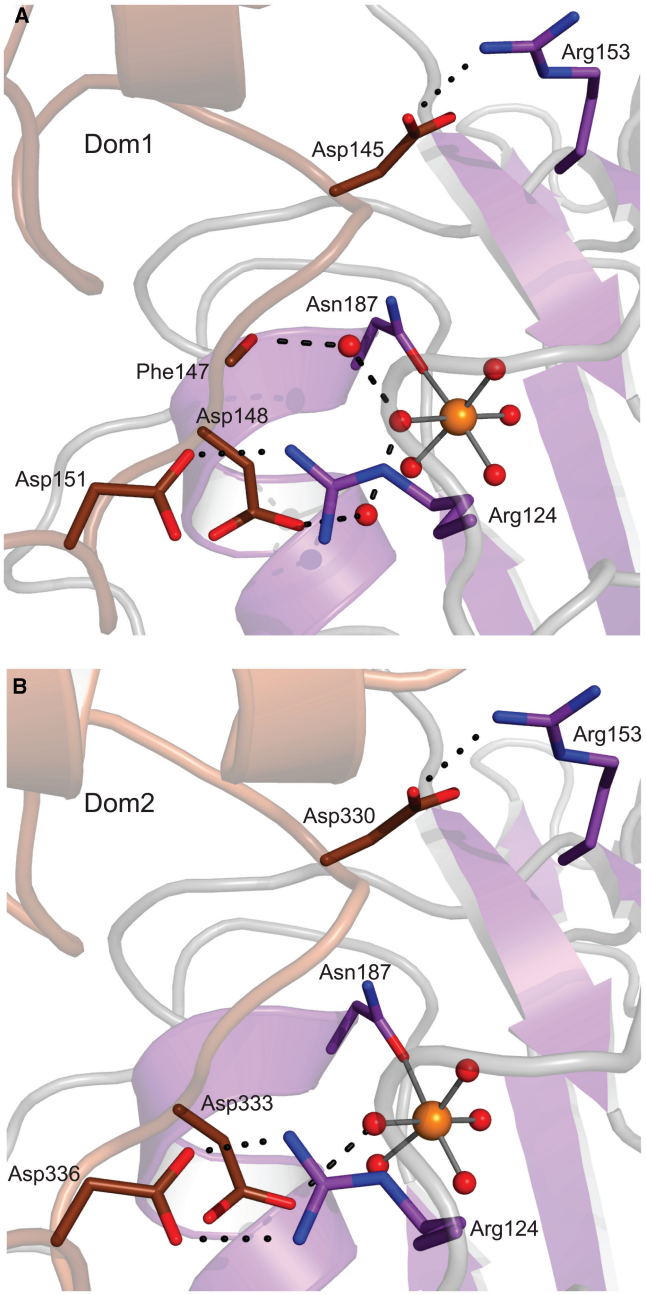

Figure 5. Inhibition of the active site of dEndoG by dEndoGI. The interface between Dom1 of dEndoGI (brown) and the active site of one dEndoG monomer (violet) is shown in (**A**) and the equivalent interface between Dom2 and the second dEndoG molecule in (**B**). Residues involved in binding are shown as sticks and for simplicity only salt bridges of important arginine residues (dotted black lines) and hydrogen bonds to water molecules (dashed black lines) are presented.

molecule which in turn forms a hydrogen bond to the carbonyl group of Phe147 (Figure 5A). In contrast, in Dom2 the side chain of Asp333 forms a direct hydrogen bond to the respective water molecule coordinating the metal ion (Figure 5B). These differences may contribute to the observed differences in the reported affinity of the separated domains for dEndoG where Dom2 binds with higher affinity to dEndoG than Dom1 does (21).

An additional difference between Dom1 and Dom2 binding dEndoG is the hydrophobic interaction between Leu146 of Dom1, respectively Tyr331 in Dom2, and Phe129 of dEndoG. Whereas the aromatic ring system in Dom2 (Tyr331) makes a hydrophobic stacking interaction with the side chain of Phe129, the side chain of Leu129 in Dom1 shows alternative conformations and apparently is flexible. This seems to be a general difference in EndoGI from different species, since Leu129 is strictly conserved in Dom1 and at the position equivalent to Tyr331 the aromatic side chain character is maintained in Dom2 (Supplementary Figure S1). Additional minor differences in binding of Dom1 or Dom2 of dEndoGI to dEndoG were observed and are listed in Supplementary Table S3.

CONCLUSIONS

Eukaryotic apoptotic endonuclease G and the related bacterial endonucleases share a conserved active site and oligonucleotide-binding groove. Similarly to the related bacterial *Serratia* nuclease, EndoG is a homodimeric protein, although the mode of dimerization differs. The *Drosophila* inhibitor EndoGI is composed of two highly conserved domains that are linked by a flexible and non-conserved polypeptide stretch (Supplementary Figure S1). The crystal structure revealed that each domain blocks the active site of one subunit in the dEndoG dimer and seals the oligonucleotide-binding groove. However, structural differences are found between the two domains, and Dom1 binds less tightly to one dEndoG monomer. Intriguingly, this domain shows a higher degree of amino acid sequence conservation when compared with Dom2 (Supplementary Figure S1). It seems plausible that Dom1 of EndoGI might serve additional functions which lead to the observed sequence conservation. Another possible explanation could be that the amino acid sequence variation of Dom2 between different *Drosophila* homologues serves a species-specific adaption and represents a modulation of tight binding, since Dom2 binds to dEndoG with higher affinity than Dom1 (21).

Each domain of dEndoGI is composed of HEAT repeats; dEndoGI apparently originated from a protein-protein interaction module that eventually was transformed into an inhibitor. Many trafficking proteins are composed of HEAT repeats, with some of the most closely related structural homologues of the dEndoGI domains involved in these processes. Notably, importin β which was shown to be involved in EndoG translocation in *Saccharomyces cerevisiae* (3), is structurally related to dEndoGI. So far, EndoGI homologues have only been identified in various *Drosophila* species (21). It remains to be seen whether inhibitors of EndoG also exist in other eukaryotes. Common bioinformatic tools may fail to identify EndoGI homologues solely based on amino acid sequence alignments, since HEAT repeat proteins show weak sequence conservation, and most conserved residues lead to a common hydrophobic core (46). Thus, the structure presented serves as a starting point for a structure based search for EndoGI homologues in other eukaryotic organisms.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank M. Cryle, J. Reinstein, R. Shoeman for helpful discussions and R. Shoeman for performing the MALDI-TOF experiments. The authors are grateful to I. Schlichting for continuous encouragement and support. The authors thank E. Hofmann and the scientific staff at the beamline X10SA, Paul Scherrer Institute (Villigen, Switzerland) for data collection and I. Vetter for support of the crystallographic software.

FUNDING

Grants to A.M. and E.W. from the Deutsche Forschungsgemeinschaft. A.M. is member of CellNetworks—Cluster of Excellence (EXC81). Funding for open access charge: Max Planck Society

Conflict of interest statement. None declared.

REFERENCES

1. Parrish,J., Li,L., Klotz,K., Ledwich,D., Wang,X. and Xue,D. (2001) Mitochondrial endonuclease G is important for apoptosis in *C. elegans*. *Nature*, **412**, 90–94.
2. Schafer,P., Scholz,S.R., Gimadutdinow,O., Cymerman,I.A., Bujnicki,J.M., Ruiz-Carrillo,A., Pingoud,A. and Meiss,G. (2004) Structural and functional characterization of mitochondrial EndoG, a sugar non-specific nuclease which plays an important role during apoptosis. *J. Mol. Biol.*, **338**, 217–228.
3. Buttner,S., Eisenberg,T., Carmona-Gutierrez,D., Ruli,D., Knauer,H., Ruckenstuhl,C., Sigrist,C., Wissing,S., Kollroser,M., Frohlich,K.U. *et al.* (2007) Endonuclease G regulates budding yeast life and death. *Mol. Cell*, **25**, 233–246.
4. Muro-Pastor,A.M., Flores,E., Herrero,A. and Wolk,C.P. (1992) Identification, genetic analysis and characterization of a sugar-non-specific nuclease from the cyanobacterium *Anabaena* sp. PCC 7120. *Mol. Microbiol.*, **6**, 3021–3030.
5. Ball,T.K., Saurugger,P.N. and Benedik,M.J. (1987) The extracellular nuclease gene of *Serratia marcescens* and its secretion from *Escherichia coli*. *Gene*, **57**, 183–192.
6. Beiter,K., Wartha,F., Albiger,B., Normark,S., Zychlinsky,A. and Henriques-Normark,B. (2006) An endonuclease allows *Streptococcus pneumoniae* to escape from neutrophil extracellular traps. *Curr. Biol.*, **16**, 401–407.
7. Buchanan,J.T., Simpson,A.J., Aziz,R.K., Liu,G.Y., Kristian,S.A., Kotb,M., Feramisco,J. and Nizet,V. (2006) DNase expression allows the pathogen group A *Streptococcus* to escape killing in neutrophil extracellular traps. *Curr. Biol.*, **16**, 396–400.
8. Cote,J. and Ruiz-Carrillo,A. (1993) Primers for mitochondrial DNA replication generated by endonuclease G. *Science*, **261**, 765–769.
9. Huang,K.J., Zemelman,B.V. and Lehman,I.R. (2002) Endonuclease G, a candidate human enzyme for the initiation of genomic inversion in herpes simplex type 1 virus. *J. Biol. Chem.*, **277**, 21071–21079.
10. Li,L.Y., Luo,X. and Wang,X. (2001) Endonuclease G is an apoptotic DNase when released from mitochondria. *Nature*, **412**, 95–99.

11. van Loo,G., Schotte,P., van Gurp,M., Demol,H., Hoorelbeke,B., Gevaert,K., Rodriguez,I., Ruiz-Carrillo,A., Vandekerckhove,J., Declercq,W. *et al.* (2001) Endonuclease G: a mitochondrial protein released in apoptosis and involved in caspase-independent DNA degradation. *Cell Death Differ.*, **8**, 1136–1142.
12. Zhao,T., Zhang,H., Guo,Y. and Fan,Z. (2007) Granzyme K directly processes bid to release cytochrome *c* and endonuclease G leading to mitochondria-dependent cell death. *J. Biol. Chem.*, **282**, 12104–12111.
13. Irvine,R.A., Adachi,N., Shibata,D.K., Cassell,G.D., Yu,K., Karanjawala,Z.E., Hsieh,C.L. and Lieber,M.R. (2005) Generation and characterization of endonuclease G null mice. *Mol. Cell Biol.*, **25**, 294–302.
14. David,K.K., Sasaki,M., Yu,S.W., Dawson,T.M. and Dawson,V.L. (2006) EndoG is dispensable in embryogenesis and apoptosis. *Cell Death Differ.*, **13**, 1147–1155.
15. Parrish,J.Z. and Xue,D. (2006) Cuts can kill: the roles of apoptotic nucleases in cell death and animal development. *Chromosoma*, **115**, 89–97.
16. Ball,T.K., Suh,Y. and Benedik,M.J. (1992) Disulfide bonds are required for *Serratia* marcescens nuclease activity. *Nucleic Acids Res.*, **20**, 4971–4974.
17. Ghosh,M., Meiss,G., Pingoud,A.M., London,R.E. and Pedersen,L.C. (2007) The nuclease a-inhibitor complex is characterized by a novel metal ion bridge. *J. Biol. Chem.*, **282**, 5682–5690.
18. Kobe,B. and Deisenhofer,J. (1993) Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature*, **366**, 751–756.
19. Hartley,R.W. (1989) Barnase and barstar: two small proteins to fold and fit together. *Trends Biochem. Sci.*, **14**, 450–454.
20. Carr,S., Walker,D., James,R., Kleanthous,C. and Hemmings,A.M. (2000) Inhibition of a ribosome-inactivating ribonuclease: the crystal structure of the cytotoxic domain of colicin E3 in complex with its immunity protein. *Structure*, **8**, 949–960.
21. Temme,C., Weissbach,R., Lilie,H., Wilson,C., Meinhart,A., Meyer,S., Golbik,R., Schierhorn,A. and Wahle,E. (2009) The *Drosophila melanogaster* Gene *cg4930* encodes a high affinity inhibitor for endonuclease G. *J. Biol. Chem.*, **284**, 8337–8348.
22. Cote,J., Renaud,J. and Ruiz-Carrillo,A. (1989) Recognition of (dG)n.(dC)n sequences by endonuclease G. Characterization of the calf thymus nuclease. *J. Biol. Chem.*, **264**, 3301–3310.
23. Wu,S.L., Li,C.C., Chen,J.C., Chen,Y.L., Lin,C.T., Ho,T.Y. and Hsiang,C.Y. (2009) Mutagenesis identifies the critical amino acid residues of human endonuclease G involved in catalysis, magnesium coordination, and substrate specificity. *J. Biomed. Sci.*, **16**, 6.
24. Kabsch,W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.*, **26**, 795–800.
25. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Cryst.*, **40**, 458–674.
26. Ghosh,M., Meiss,G., Pingoud,A., London,R.E. and Pedersen,L.C. (2005) Structural insights into the mechanism of nuclease A, a $\beta\beta\alpha$ metal nuclease from *Anabaena*. *J. Biol. Chem.*, **280**, 27990–27997.
27. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Cryst.*, **D60**, 2126–2132.
28. Brunger,A.T. (2007) Version 1.2 of the Crystallography and NMR system. *Nat Protoc.*, **2**, 2728–2733.
29. CCP4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **50**, 760–763.
30. Laskowski,R.A., Mac Arthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, **26**, 283–291.
31. Lovell,S.C., Davis,I.W., Arendall,W.B. 3rd, de Bakker,P.I., Word,J.M., Prisant,M.G., Richardson,J.S. and Richardson,D.C. (2003) Structure validation by Cα geometry: φ, ψ and Cβ deviation. *Proteins*, **50**, 437–450.
32. DeLano,W.L. (2002) The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA, USA.
33. Baker,N.A., Sept,D., Joseph,S., Holst,M.J. and McCammon,J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
34. Chenna,R., Sugawara,H., Koike,T., Lopez,R., Gibson,T.J., Higgins,D.G. and Thompson,J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.*, **31**, 3497–3500.
35. Barton,G.J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.*, **6**, 37–40.
36. Holm,L., Kaariainen,S., Rosenstrom,P. and Schenkel,A. (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics*, **24**, 2780–2781.
37. Miller,M.D. and Krause,K.L. (1996) Identification of the *Serratia* endonuclease dimer: structural basis and implications for catalysis. *Protein Sci.*, **5**, 24–33.
38. Friedhoff,P., Kolmes,B., Gimadutdinow,O., Wende,W., Krause,K.L. and Pingoud,A. (1996) Analysis of the mechanism of the *Serratia* nuclease using site-directed mutagenesis. *Nucleic Acids Res.*, **24**, 2632–2639.
39. Cymerman,I.A., Chung,I., Beckmann,B.M., Bujnicki,J.M. and Meiss,G. (2008) EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Res.*, **36**, 1369–1379.
40. Li,C.L., Hor,L.I., Chang,Z.F., Tsai,L.C., Yang,W.Z. and Yuan,H.S. (2003) DNA binding and cleavage by the periplasmic nuclease Vvn: a novel structure with a known active site. *Embo J.*, **22**, 4014–4025.
41. Franke,I., Meiss,G. and Pingoud,A. (1999) On the advantage of being a dimer, a case study using the dimeric *Serratia* nuclease and the monomeric nuclease from *Anabaena* sp. strain PCC 7120. *J. Biol. Chem.*, **274**, 825–832.
42. Zhang,Z., Hu,W., Cano,L., Lee,T.D., Chen,D.J. and Chen,Y. (2004) Solution structure of the C-terminal domain of Ku80 suggests important sites for protein-protein interactions. *Structure*, **12**, 495–502.
43. Hierro,A., Rojas,A.L., Rojas,R., Murthy,N., Effantin,G., Kajava,A.V., Steven,A.C., Bonifacino,J.S. and Hurley,J.H. (2007) Functional architecture of the retromer cargo-recognition complex. *Nature*, **449**, 1063–1067.
44. Liu,S.M. and Stewart,M. (2005) Structural basis for the high-affinity binding of nucleoporin Nup1p to the *Saccharomyces cerevisiae* importin-beta homologue, Kap95p. *J. Mol. Biol.*, **349**, 515–525.
45. Kirby,T.W., Mueller,G.A., DeRose,E.F., Lebetkin,M.S., Meiss,G., Pingoud,A. and London,R.E. (2002) The nuclease A inhibitor represents a new variation of the rare PR-1 fold. *J. Mol. Biol.*, **320**, 771–782.
46. Andrade,M.A., Petosa,C., O'Donoghue,S.I., Muller,C.W. and Bork,P. (2001) Comparison of ARM and HEAT protein repeats. *J. Mol. Biol.*, **309**, 1–18.