Generation of novel inhibitor-variants for ββα-metal finger nucleases by evolution and rational protein design

Inauguraldissertation

Zur Erlangung des Grades

Doktor der Naturwissenschaften

Dr. rer. nat.

Des Fachbereiches Biologie und Chemie

Der Justus-Liebig-Universität Gießen

Vorgelegt von

Diplom-Biologin

Marika Midon

Gießen, 2010

Die vorliegende Arbeit wurde im Rahmen des Graduiertenkollegs „Enzymes and Multienzyme Complexes Acting on Nucleic Acids“ (GRK 1384) am Institut für Biochemie des Fachbereichs 08 (Biologie und Chemie) der Justus-Liebig-Universität Gießen in der Zeit von November 2006 bis Februar 2010 unter der Leitung von PD. Dr. Gregor Meiß durchgeführt.

Erstgutachter: PD. Dr. Gregor Meiß
Institut für Biochemie, FB08
Justus-Liebig-Universität Gießen
Heinrich-Buff-Ring 58, 35392 Gießen

Zweitgutachter: Prof. Dr. Roland Hartmann
Institut für Pharmazeutische Chemie, FB16
Philipps-Universität Marburg
Marbacher Weg 6-10, 35037 Marburg

Erklärung

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Gießen, den

“Differences of habit and language are nothing at all if our aims are identical and our hearts are open.”

J. K. Rowling: Harry Potter and the Goblet of Fire

„Unterschiede in Lebensweise und Sprache, werden uns nicht im geringsten stören, wenn unsere Ziele die gleichen sind und wir den anderen mit offenen Herzen begegnen.“

J. K. Rowling: Harry Potter und der Feuerkelch

Danksagung

Als erstes möchte ich mich bei Prof. Dr. Alfred Pingoud bedanken, der mich in seinem Institut als Doktorandin aufgenommen und mich stetig unterstützt hat. Seine Zielstrebigkeit, Tatenkraft und sein Scharfsinn sind bewundernswert. Eigenschaften die ihn auch in schweren Zeiten auszeichnen und ich wünsche mir, dass das so bleibt.

Gregor Meiss möchte ich für seine Geduld danken. In all der Zeit hat er mir zuverlässig auf jede meiner Fragen geantwortet und mich zurechtgewiesen, wenn ich mal wieder meinen eigenen Weg gehen wollte. Schon erstaunlich, wie schlau man sein kann.

Danken möchte ich Heike Büngen, die mir als Anfängerin in der Biochemie alle grundlegenden Methoden beigebracht hat. Irgendwie schafft sie es jeden Tag aufs Neue drei Kinder und ein Labor zu organisieren und trotzdem noch für alle ein offenes Ohr zu haben.

Bedanken möchte ich mich bei der DFF-Gruppe, die mich in ihren Reihen aufgenommen haben und die für jeden Spaß zu haben waren. Insbesondere bei Paddy, als Fels in der Brandung, dessen Arbeit ich weiter führen durfte. Danke an Wibke, Steffi und Jana auf die man sich immer verlassen konnte und die stets hilfsbereit waren. Daniel, danke dafür, dass du mir so oft zuhörst und für die gute Laune die du konstant mit ins Labor bringst und das Pfeifen…

George, thank you for a lot of discussions and for being my friend. Peanut!

Danke an Anja, die mit ihrer ruhigen und offenen Art die Strippen im IRTG zieht und der kein Hotel und kein Flug zu teuer ist.

Ein großer Dank geht auch an Claudia, Ina und Karina, die immer alles für mich möglich gemacht haben.

Danke an Peter für viele aufschlussreiche Diskussionen und eine spannende Zeit in Cambridge.

Danke an Ines, der loyalste Mensch der mir je begegnet ist, mit großem Herzen. Danke dafür, dass du mir so oft zu hörst und all meine Launen erträgst. Danke für die vielen gemeinsamen Autobahnkilometer und Staustunden, denn im Zug wäre es mächtig langweilig gewesen und danke das du meine Wäsche wäschst….

Danke an Ilse, Kathi und Silke für ganz viel Spaß und Ablenkung vom Institutsalltag. Fühlt sich gut an, wenn man Freunde hat.

Danke an Dani, Airam und Ramona die mich all die Jahre für mich da waren und mir immer wieder Mut zugesprochen haben.

Ich möchte mich besonders bei meiner Familie bedanken, die immer an mich geglaubt haben. Ohne euch wäre ich nie soweit gekommen. Danke für eure tatkräftige Unterstützung in allen Lebenslagen.

Herr Helm, ich bin nicht eher fertig geworden, aber irgendwann sehen wir uns wieder, denn wenn die Zeit aufhört, beginnt die Ewigkeit…

Index

1 Introduction

1.1 Non-specific nucleases

Non-specific nucleases are ubiquitous enzymes and involved in many essential processes such as DNA repair, recombination, apoptosis, host defense and nutrition. They hydrolyze the phosphodiester backbone of nucleic acids in a sequence/sugar – non-specific manner leading to degradation of DNA/RNA up to the level of nucleotides[1]. This is an essential process for example during apoptotic cell death which in turn contributes to cellular homeostasis and prevents the accumulation of abnormal cells[2]. The two non-specific nucleases EndoG (Endonuclease G) and CAD (Caspase activated DNase) found in higher eukaryotes are responsible for the random degradation of chromosomal DNA during apoptosis. Complete degradation after phagocytosis of the dying cells to nucleotides is mediated by the non-specific nuclease DNase II.

A second type of non-specific nucleases is located in the periplasm of gram-negative bacteria such as Vvn from *Vibrio vulnificus* and EndoI from *E. coli.* These nucleases take part in host defense against the uptake of foreign DNA e.g. during infection by phages. However, the non-specific degradation of DNA leads to reduced transformation rates of theses bacteria as strains lacking Vvn and EndoI can take up DNA more efficently[3; 4]. On the other hand, nucleolytic activity on the surface of some gram-positive bacteria such as *Streptococcus pneumoniae* and *Bacillus subtilis* displayed by the nucleases EndA and NucA is essential for the import of single-stranded DNA fragments in the course of transformation[5; 6]. Other non-specific nucleases are involved in the complete degradation of DNA for the purpose of assimilation of rare nucleotides and phosphate from the environment. These extracellular nucleases are secreted e.g. from *Serratia marcescens* (*Sma*Nuc) and from *Anabaena sp*. (NucA). Whereas under nutrient-limited conditions the extracellular E colicins from *E. coli* such as ColE9 and ColE7 kill competing cells of other *E. coli* strains due to non-specific cleavage of cellular DNA[7; 8].

In addition, non-specific nucleases are also involved in the important mechanisms of DNA repair and recombination. Two well characterized nucleases are ExoI from *E. coli* and the

yeast protein Rad52. They mediate the trimming of broken/cleaved DNA necessary for further repair steps to maintain the necessary stability of the genome[9; 10].

1.1.1 Active site motifs

The non-specific degradation of DNA/RNA is mediated by structurally divergent proteins but only a few active site motifs/structures are involved in the catalytic mechanism for hydrolysis. Each of these motifs utilizes specific mechanisms and exhibits similar features such as conserved amino acid residues and protein folds for nucleolytic cleavage. Interestingly, also sequence/sugar specific nucleases e.g. restriction endonucleases and homing endonucleases share equal active site motifs indicating common ancestors for all types of nucleases and other proteins involved in DNA/RNA hydrolysis such as resolvases and transposases[11].

Homing endonucleases (HEases) are highly specific nucleases with long DNA target sites of 14-40 bp mediating a process termed homing which implies the transfer of the their own coding sequence to cognate alleles lacking the sequence. The process for group I intron and intein encoded nucleases is initiated by a double-strand break at the target site which is required to insert the coding sequence during cell mediated repair. These mobile genetic elements integrate at the target site supported by the DNA repair machinery of the host using the intron-containing allele as a template (homologous recombination)[12]. Based on known crystal structures and sequence comparison five families of HEases have been determined with representative conserved amino acids: LAGLIDADG, H-N-H, His-Cys box, GIY-YIG and PD(D/E)XK (see Table 1). In general, the motifs are located within the active site except for the His-Cys box which coordinates zinc[13]. Members of this family like I-PpoI exhibit an additional H-N-H motif as active site[11]. However, these active site motifs are not restricted to the family of HEases with an exception of the specific LAGLIDADG motif.

The PD(D/E)XK motif for example is the most common motif of Type II restriction endonucleases (REases) (see Figure 1, Table 1)[14]. Type II REases found in bacteria recognize short DNA sequences from 4-8 bp and are part of the restriction modification system (RM). The nucleases are coexpressed in the cell with a specific DNA methyltransferase (MTase). REases are involved in host defense as they degrade incoming phage DNA whereas the genomic host DNA is protected due to methylation of the recognition sites by the MTases[15].

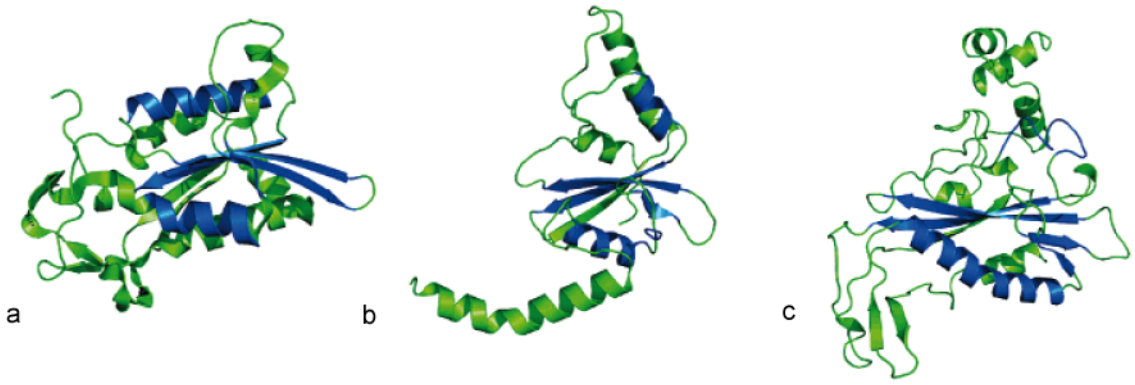

Figure 1: PD(D/E)XK motif containing REases. a) BamHI b) PvuII c) EcoRV. The PD(D/E)XK domain of each nuclease is colored in blue. Only one monomer of each enzyme is shown. Modified after[16].

The second most common fold found in REases is the H-N-H motif which displays the same overall active site architecture like His-Cys box containing nucleases termed ββα-metal finger fold (see Figure 2)[14]. Members of this family exhibit an active site with two antiparallel β strands and an α helix arranged around a central divalent metal ion[17]. Interestingly, also the phage encoded T4 endonuclease VII which cleaves and resolves Holliday junctions, contains an H-N-H motif[18; 19]. Several non-specific nucleases from higher eukaryotes such as the apoptotic nucleases CAD and EndoG and the prokaryotic extracellular nucleases NucA, *Sma*Nuc and the colicins E7 and E9 exhibit the H-N-H motif/ββα-metal finger fold (see Figure 2 and Table 1)[12; 20; 21; 22; 23].

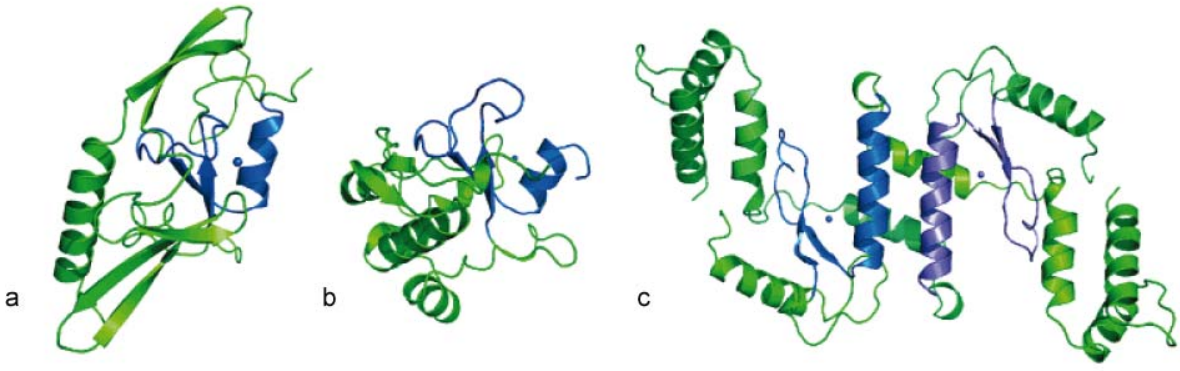

Figure 2: ββα-metal finger motif containing enzymes. a) H-N-H domain of I-PpoI (His-Cys box, HEase). b) H-N-H domain of I-HmuI (H-N-H motif, HEase). c) T4 endonuclease VII (H-N-H motif, resolvase). The ββα-metal finger fold is indicated in blue/purple. Modified after[16].

Table 1: Distribution of active site motifs (incomplete). His-Cys box containing proteins were not considered as the motif does not concern the active site. HEases containing a His-Cys box exhibit an H-N-H motif.

	Type of enzyme	Name
PD(D/E)XK motif	Homing endonuclease	I-Ssp6803I
	Restriction endonuclease	BamHI
	Restriction endonuclease	PvuII
	Restriction endonuclease	EcoRI
	Restriction endonuclease	EcoRV
	Resolvase (Holliday junction)	Hjc
	Transposase	Tn7 transposase TnsA
H-N-H motif	Homing endonuclease	I-HmuI
	Homing endonuclease	I-BasI
	Homing endonuclease	I-TevIII
	Restriction endonuclease	KpnI
	Non-specifc nuclease	Colicin E7
	Non-specifc nuclease	Colicin E9
	Non-specifc nuclease	*Sma*Nuc
	Non-specifc nuclease	NucA
	Non-specifc nuclease	CAD
	Non-specifc nuclease	EndoG
	Non-specific nuclease	Vvn
	Resolvase (Holliday junction)	T4 endonuclease VII
LAGLIDADG motif	Homing endonuclease	I-CreI
	Homing endonuclease	PI-SceI
	Homing endonuclease	I-SceI
	Homing endonuclease	I-DmoI
GIY-YIG motif	Homing endonuclease	I-TevI
	Restriction endonuclease	Eco29kI
	Excinuclease (excision repair)	UvrC

1.1.2 The H-N-H motif

The H-N-H motif characterizes a high number of nucleases or proteins involved in DNA hydrolysis found in all kingdoms of life[24]. The three conserved amino acids histidine-asparagine-histidine could be detected in sequence alignments of enzymes summarized as a Pfam protein family[25]. Further sequence comparison revealed at least eight subsets of the H-N-H family[26].

Recently analyzed nucleases carrying an H-N-H motif exhibit an active site fold as ββα-metal finger structure (see Figure 2) like *Sma*Nuc, NucA, CAD, colicin E7, E9, I-PpoI, I-HmuI and others[11; 17]. This active site motif consists of two anti-parallel β strands and an α helix arranged around a central divalent metal ion (e.g. Mg^{2+}, Mn^{2+}, Ni^{2+}).

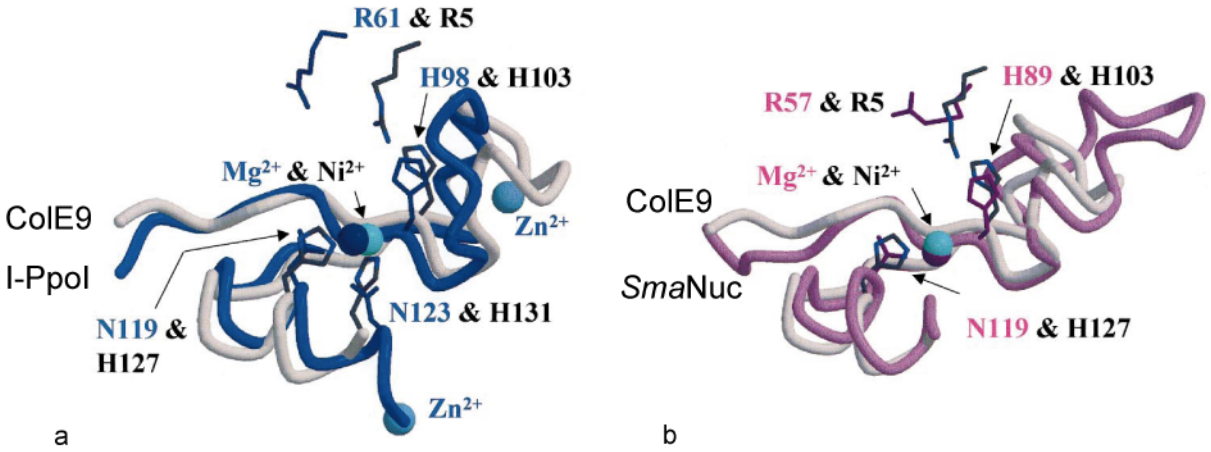

Figure 3: Overlay of the active sites of ColE9/I-PpoI and ColE9/*Sma*Nuc. Side chains of selected amino acid residues involved in catalysis and metal ion binding are highlighted. a) Overlay of ColE9 (gray, labels in black) and I-PpoI (blue, labels in blue). H98 and H103 represent the first histidine of the H-N-H motif acting as the general base for the activation of a water molecule during catalysis. N119, H127 (H-N-H/N) and H131 (indirect binding) are involved in the binding and coordination of the metal ion but not N123. R61 and R5 play a role in the stabilization of the pentacoordinate transition state occurring during catalysis. The two zinc ions of I-PpoI are not involved in catalysis since they support the protein structure. b) Overlay of ColE9 (gray, labels in black) and *Sma*Nuc (in purple). H89 and H103 act as the general base (H-N-H). Only N119 and H127 are involved in metal ion binding (H-N-H/N). R57 and R5 stabilize the pentacoordinate transition state. Modified after[17].

The histidine residue at the first position of the H-N-H motif is involved in catalysis and acts as the general base during nucleic acid cleavage. It activates a water molecule for a nucleophylic in-line attack on the phosphate leaving 5' phosphate and 3' hydroxyl groups. The phosphate is penta-coordinate in the transition state occurring during catalysis to release the P-O bond (see Figure 4).

The nucleophylic attack is supported by the metal ion mediating a polarization of the P-O bond. Additionally, the metal ion is known to be involved in the stabilization of the phosphor anion state and the cleaved product[27].

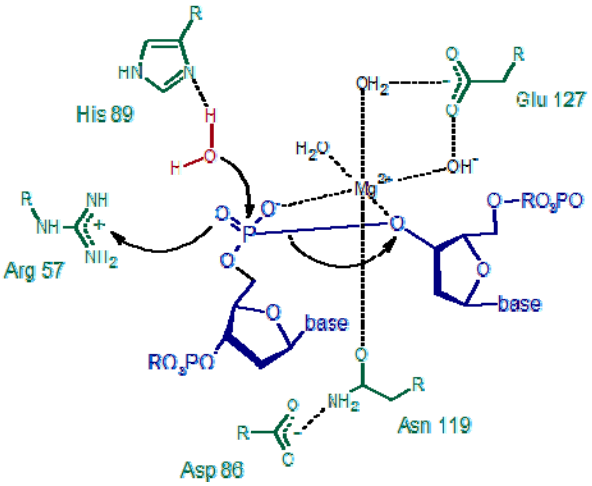

Figure 4: Proposed catalytic mechanism of *Sma*Nuc. Amino acid residues are indicated in cyan, the substrate DNA in deep blue. Histidine 89 acts as the general base to activate a water molecule (red) for a nucleophylic attack on the phosphate leaving 5' phosphate and 3' hydroxyl groups. Arginine 57 stabilizes the penta-coordinative transition state. Asparagine 119 binds to the divalent metal ion magnesium. Modified after[23].

The central asparagine of the H-**N**-H motif is involved in restraining the position of the two β strands hence stabilizing the ββα-metal finger fold shown for the colicins E7 and E9 (see Figure 3)[28; 29]. His-Cys box containing nucleases like I-PpoI bind additionally zinc ions close to the active site to support the protein structure[13; 17]. Metal ion binding to the H-N-**H** motif is mediated by the last histidine residue within the fold which can be replaced by asparagine (H-N-N motif, see Figure 3). According to Maté and Kleanthous H-N-H/N motif containing nucleases fall into two subgroups, those utilizing either a single or two amino acids for metal ion binding (see Table 2)[30]. This second amino acid residue (aspartic acid, histidine, or glutamic acid) is located in front of the general base histidine (D/H/E**H**-N-H) found in CAD, ColE7/E9 and I-HmuI and other recently characterized H-N-H nucleases.

Another subgroup of H-N-H non-specific nucleases additionally exhibits the conserved DRGH motif located on one of the two β strands of the ββα-metal finger motif. However, the last conserved histidine residue within the DRG**H** motif is congruent with the first histidine of the **H**-N-H motif.). It is known from crystallization studies of NucA and *Sma*Nuc that the aspartic acid residue of the **D**RGH motif is involved in proper positioning of the conserved asparagine residue binding to the divalent metal ion cofactor[31; 32]. The arginine (D**R**GH) instead seems to mediate substrate binding to the active site[23]. Other members of this family are the recently structurally characterized nuclease EndoG and EndA from *Streptococcus pneumoniae*.

Table 2: H-N-H/H-N-N motif containing nucleases listed by the number of metal ion contacts (incomplete). Amino acid residues involved in metal ion binding are underlined.

Single metal ion contact	Two metal ion contacts
dEndoG (DRGH/H-N-N motif)	CAD (DH-N-H)
*Sma*Nuc (DRGH/H-N-N motif)	Colicin E7/E9 (HH-N-H)
NucA (DRGH/H-N-N motif)	Vvn (EH-N-N)
I-PpoI (H-N-N motif)	I-HmuI (DH-N-N)

1.1.3 Non-specific nucleases with H-N-H motif

1.1.3.1 Colicins

Colicins are plasmid–encoded antimicrobial proteins produced by *E. coli* from a DNA–damage inducible promoter (SOS promoter) in times of stress to kill competing cells in nutrient-limited environments. They are classified on the basis of cell surface receptors they bind to and abuse pathways for nutrient uptake to enter the cell[8]. Group A colicins use the Tol translocation system whereas members of group B require the Ton system[33].

Colicins consist in general of a central receptor binding domain, an N-terminal translocation domain and a C-terminal cell-killing domain (see Figure 5). After binding of the receptor domain which is structured as long coiled coil, the N-terminal translocation domain interacts with TolB for translocation of the C-terminal killing domain to the inner membrane and into the cell[33]. As the protein needs to span the periplasm the receptor domain which is still bound to the receptor needs to unfurl because the two interactions occur simultaneously[34].

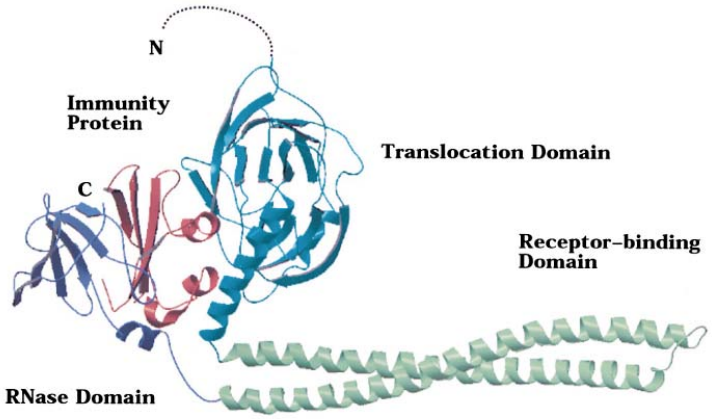

Figure 5: Chrystal structure of the ColE3/Im3 complex. The C-terminal RNase domain indicated in purple is bound to the immunity protein Im3 (red). The N-terminal translocation domain (blue) is in close proximity to the RNase domain. The receptor-binding domain exhibits a long colied-coil (light blue) which needs to unfurl to allow the protein to be connected with the outer membrane and simultaneously translocate its toxic C-terminal domain to the inner membrane and into the cell. Modified after[35].

E colicins are group A colicins and occur in complex with their specific inhibitor proteins (immunity protein - Im). They bind with the central receptor-binding domain to BtuB receptors of the outer membrane of gram-negative bacteria usually responsible for Vitamin B_{12} binding and utilize porin OmpF for translocation into the periplasm[34; 36]. The group of E colicins is subdivided into nine types (ColE1-ColE9) due to immunity tests[37; 38]. They have different cytotoxic activities: ColE1 is a membrane-depolarizing agent[39]. ColE2, ColE7, ColE8 and ColE9 are DNases whereas ColE3, ColE5 and ColE6 are RNases[40; 41; 42; 43; 44; 45].

The C-terminally located DNase domains of ColE7 and ColE9 were isolated and could be crystallized in complex with their specific immunity proteins (Im7 and Im9)[28; 46]. Both enzymes revealed a ββα-metal finger structure containing an H-N-H motif as active site. Binding of substrate DNA is mediated by contacting the phosphate backbone of the minor grove bending the DNA towards the major grove. These unspecific contacts result in sugar- and sequence-independent cytotoxic degradation of nucleic acids comparable to the apoptotic nuclease CAD[30; 47].

Cytotoxicity during expression and secretion of E colicins is regulated by the tight binding immunity proteins. The inhibitor Im9 binds to colicine E9 with femtomolar affinity forming a stable complex[48]. However, binding of the specific immunity proteins Im9 and Im7 does not occur at the active site but at an adjacent position (exosite, see Figure 6). Substrate binding to the active site is affected due to steric and electrostatic clashes with the immunity protein[28].

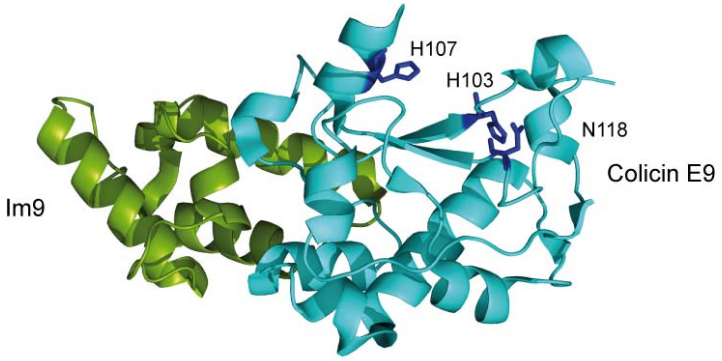

Figure 6: Crystal structure of the DNase domain of colicin E9 and its specific inhibitor Im9[49]. The nuclease (light blue) displays a ββα-metal finger fold as active site with an H-N-H motif (deep blue). The specific immunity protein Im 9 (green) binds at an exosite, blocking binding of the substrate to the active site. (PDB code: 1emv)

1.1.3.2 Caspase activated DNase (CAD)

Homologs of murine Caspase-activated DNase (CAD) are found in all higher eukaryotes and are responsible for the degradation of chromosomal DNA during cell death/apoptosis. CAD and the human ortholog termed DFF40 (DNA fragmentation factor) exist in the cell during non-apoptotic conditions in complex with their specific inhibitor ICAD/DFF45 as heterodimer to avoid cytotoxicity[50]. Besides the inhibitory function of ICAD/DFF45, the expression of the non-specific nuclease alone leads to an inactive protein, indicating a chaperone function for ICAD/DFF45[50]. During apoptosis a caspase-3 dependent activation of the nuclease occurs due to cleavage of the inhibitor ICAD/DFF45 at two positions resulting in the formation of nuclease homodimers[51; 52].

Monomeric CAD/DFF40 exhibits a three domain structure (see Figure 7)[20]. The C3 domain contains the active center which folds into a ββα-metal finger with H-N-H motif[21]. Dimerisation of CAD/DFF40 after activation is mediated by the C2 domain whereas the C1 domain seems to be the major interaction site for the nuclease and its inhibitor[52]. As no crystal structure for the CAD/ICAD complex is available a directed interaction between the inhibitor and the active site of CAD is unknown. Blocking the formation of active homodimeric nuclease by ICAD/DFF45 seems to play the key role in inhibition[20].

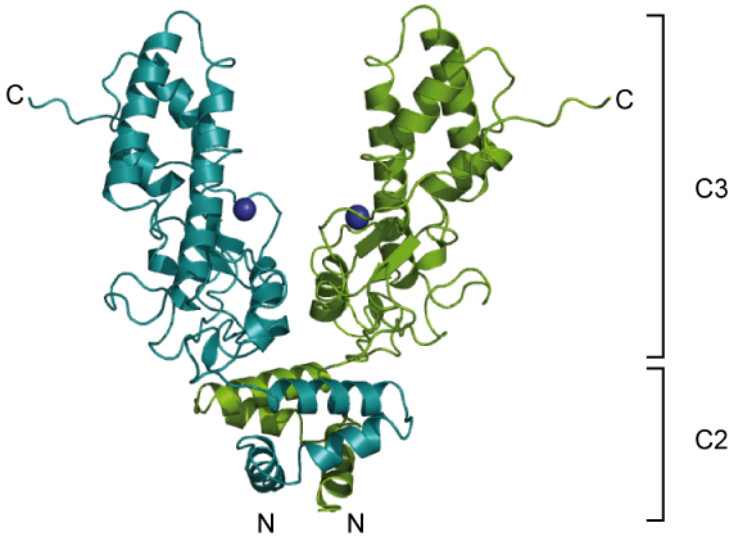

Figure 7: Crystal structure of dimeric CAD[20]. The monomers are indicated in green and blue. CAD exhibits a three domain structure (C1, C2, C3) but only the structure of C2 and C3 could be solved. Two magnesium ions colored in deep blue indicate the active site (ββα-metal finger motif). The structure of the N-terminal domain of the human homolog termed DFF40 was solved in complex with the N-terminal domain of the inhibitor DFF45 (not shown)[52]. (PDB code: 1V0D)

1.1.4 Non-specific nucleases with DRGH/H-N-H motif

1.1.4.1 dEndoG/dEndoGI

The non-specific nuclease EndoG is located in mitochondria and responsible for the degradation of chromosomal DNA in a caspase independent manner during apoptosis. It is also involved in the generation of primers for mitrochondrial DNA replication[53; 54]. The nuclease is a dimer and introduces nicks into double-stranded DNA as the two active sites are located distantly from each other and cleave independently[22]. EndoG gets released from the mitochondria due to induction of apoptosis. However, the destructive activity seems to be prevented by compartmentalization in living cells[22; 55]. Orthologous genes for EndoG displaying high sequence similarity with related functions were detected in several organisms e.g. in the human, bovine and murine system, in *D. melanogaster*, *C. elegans* (cps-6) and yeast (Nuc1p)[22; 53; 56; 57; 58; 59; 60].

Until now only the crystal structure of dEndoG from *D. melanogaster* with its functional inhibitor dEndoGI could be solved[57]. The active site of dEndoG, displays as predicted by sequence comparison, a ββα-metal finger motif and includes the conserved DRGH and H-N-

N motifs which are involved in catalysis and metal ion binding, detectable also in other orthologous genes suggesting the same catalytic mechanism[22].

However, dEndoGI as a functional inhibitor seems to be unique for *D. melanogaster* as sequence comparison did not reveal homologues in other eukaryotes[57]. In addition, dEndoGI shows no structural similarity to NuiA, the tight binding inhibitor of NucA but displays similar inhibition mechanisms[57]. Both inhibitors are involved in metal ion coordination and block access to the active site by mimicking the negative charge of the DNA phosphate backbone. DEndoGI (Asp148 and Asp333) binds to a water molecule of the metal ion hydration shell whereas NuiA (Thr135) forms a direct metal ion bridge by displacing a water molecule of the metal ion hydration shell[31].

DEndoGI displays a two domain structure and binds to homodimeric dEndoG with subpicomolar affinity (see Figure 8). Inhibitory function is mediated by binding of each domain of the inhibitor to the two dEndoG monomers exclusively at the active site as mutation of the conserved DRGH motif and the metal ion coordinative asparagine residue completely abolish binding[57; 61]. The two domains connected by a non-conserved flexible linker show high sequence similarity and are conserved among different *Drosophila* species but display different modes of binding resulting in different affinities of the separated domains for dEndoG[57].

DEndoGI is nuclear encoded and mainly located in the nucleus. The inhibitor presumably acts as “backup system” for viable cells against stray dEndoG produced by failed mitochondrial import or leakage[57]. The fate of dEndoGI during apoptosis remains unknown as the inhibitor needs to be inactive for degradation of chromosomal DNA by dEndoG. However, the inhibitory effect of dEndoGI for other dEndoG homologues has not been investigated so far.

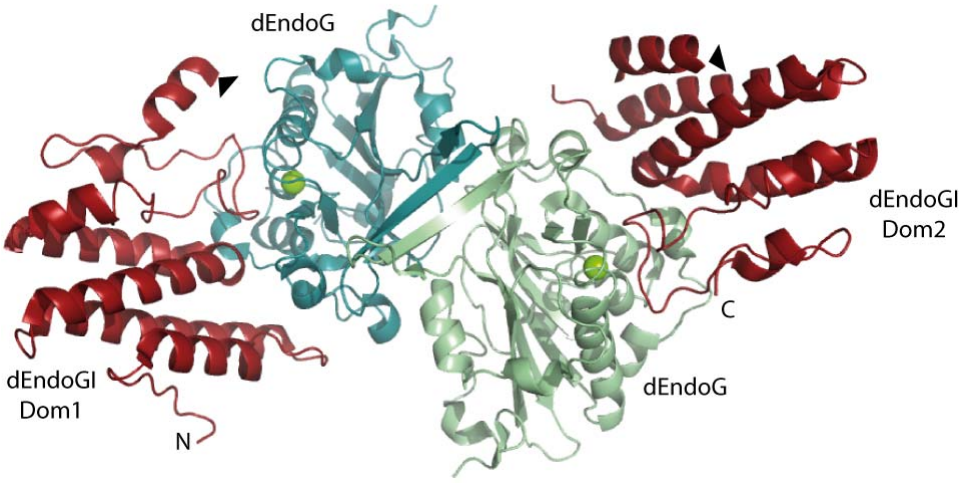

Figure 8: Structure of the dEndoG/dEndoGI complex[57]. The two domains (Dom1 and Dom2) of dEndoGI are indicated in red. A linker (connecting the two helices marked by black triangles) could not be modeled due to flexibility of this region. Homodimeric dEndoG (blue and light blue) is sandwiched by the two dEndoGI domains blocking the active site indicated by the two green magnesium ions. (PDB code: 3ISM)

1.1.4.2 EXOG

A family of paralogous genes probably derived from gene duplication of an ancestral endo/exonuclease gene was found in higher eukaryotes which was termed *EndoG-like-1* (EXOG)[62; 63]. Sequence comparison of human EXOG with human EndoG and homologous genes from yeast (Nuc1p) and *C. elegans* (CPS-6) revealed instead of a DRGH an SRGH motif as part of the active site and in addition a C-terminally located domain predicted to fold as coiled coil (see Figure 9)[63]. Localization studies indicated that EXOG is associated with the inner mitochondrial membrane like yeast and mammalian EndoG as it exhibits an N-terminal located MLS (mitochondrial localization signal) and a predicted helical transmembrane segment[53; 63; 64]. Upon apoptosis EXOG remains attached to the mitochondrial membrane in contrast to EndoG which is released from the intermembrane space[63].

Further investigation of the nucleolytic activity of human EXOG exhibited a 5'-3' exonuclease activity with a preference for single-stranded DNA and nicking activity towards supercoiled DNA[63]. However, mammalian EndoG is an exclusive endonuclease cleaving internal phosphodiester bonds in single and double-stranded DNA and RNA in contrast to the yeast homologue Nuc1p which shows 5'-3' exonuclease and endonuclease activity[56; 60; 65]. Mutation of the unusual SRGH motif of EXOG back to the highly conserved DRGH motif did not cause a significant change of nucleolytic activity or substrate preference indicating that the

single amino acid replacement is not responsible for preferential 5'-3' exonuclease activity. However, the C-terminal domain and a single glycine residue (naturally occurring SNP G277V, presumably a second variant of EXOG, see Figure 9) close to the predicted coiled coil domain also contribute to the nucleolytic activity by an unknown mechanism as deletion and amino acid replacement led to a reduction of DNA degradation[63].

Comparison of EndoG proteins from higher and lower eukaryotes displays a loss of exonuclease activity for mammalian EndoG necessary for recombination and repair of double-strand breaks in comparison to the yeast homolog Nuc1p with dual functionality. Potential complementation of mammalian EndoG can be achieved by EXOG indicating a subdivision of cellular functions between the two proteins. In principle, EXOG has the ability to generate single-stranded gaps in double-stranded DNA required for DNA recombination and repair whereas EndoG displays strong endonuclease activity necessary for the degradation of chromosomal DNA during apoptosis[63].

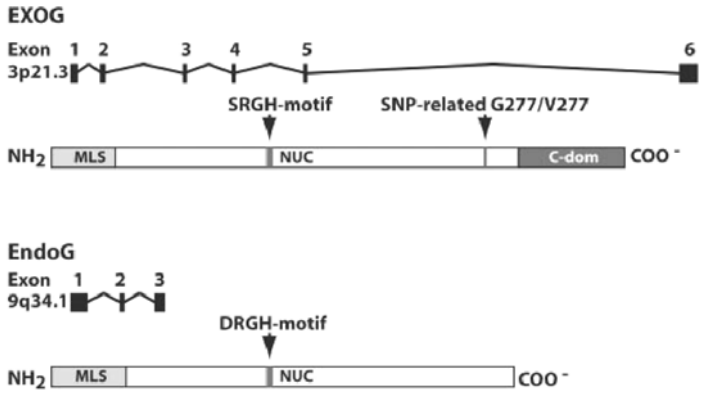

Figure 9: Genomic structure and domain organization of human EXOG and EndoG. The EXOG gene is located on chromosome 3 (3p21.3) and the EndoG gene is located on chromosome 9 (9q34.1). Exons for both genes are indicated as black boxes. Sequence comparison revealed two EXOG variants characterized by a SNP close to the predicted C-terminal coiled coil domain. MLS: mitochondrial localization signal. Copied from[63].

1.1.4.3 NucA

NucA is a non-specific nuclease secreted from the cyanobacterium *Anabaena sp.* PCC 7120 to serve nutritional purposes supporting the assimilation of rare nucleotides and phosphate from the environment[66]. The well characterized nuclease displays a nicking activity and degrades single and double-stranded DNA. It occurs as a monomer in contrast to dimeric

*Sma*Nuc[67]. However, NucA and *Sma*Nuc share 29.5 % sequence similarity and exhibit a similar overall structure[23].

The active site of NucA displays a ββα-metal finger fold with a conserved DRGH/H-N-N motif utilizing divalent metal ions such as magnesium and manganese[68]. The nucleolytic activity of NucA, highly toxic within the cell during expression, gets regulated by its specific inhibitor NuiA (see Table 3)[69]. NucA and NuiA form a 1:1 complex with picomolar affinity that was recently crystallized (see Figure 10)[23; 31]. Although DRGH/H-N-N motif containing nucleases are found in all branches of life, NuiA seems to be unique for cyanobacteria similar to *drosophila*-specific dEndoGI[69].

NuiA binds to the active site of NucA blocking the access for DNA and RNA. Active site interactions include the formation of a salt bridge between inhibitor residue Glu24 and NucA Arg93 protruding from a long loop framing half of the active site. An unusual interaction represents the metal ion bridge between the catalytically active magnesium ion and the C-terminal $Thr135_{NuiA}$. Deletion of this single amino acid leads to a 600-fold increase for the apparent inhibition constant. A second small interaction site close to the active site of NucA is characterized by a salt bridge between $Lys101_{NucA}$ and $Asp75_{NuiA}$. Several hydrogen bonds e.g. between $Arg122_{NucA}$ (D<u>R</u>GH/H-N-H) and $Ser23/Glu24/Thr135_{NuiA}$ contribute to interaction at both sites for a stable complex formation[31]. Additionally, the interaction at the active site gets promoted by the electrostatic potential of the NucA:NuiA interface. The NuiA binding surface exhibits one basic and seven acidic residues thus imitating the negatively charged substrate to support binding to the basic active site surface of NucA (substrate mimicry). However, this mechanism is not unusual as it gets utilized by several inhibitor proteins like Ugi regulating the uracil DNA glycosylase (UDG). In contrast to that, the immunity proteins Im7/9 bind to an exosite to control the nucleases colicin E7/9[70].

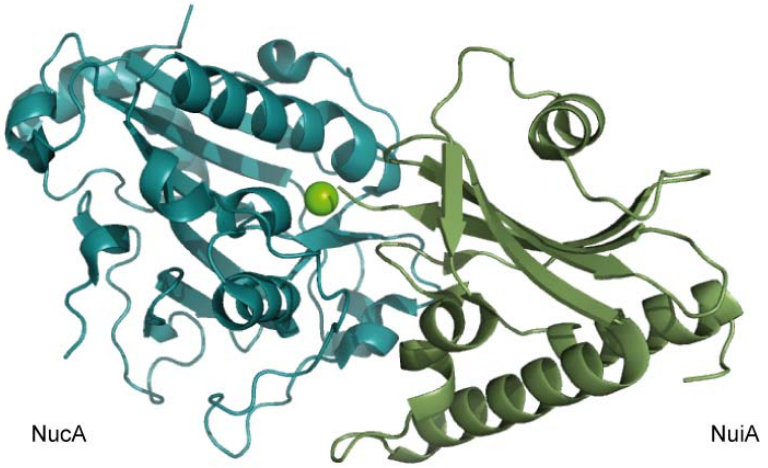

Figure 10: Crystal structure of NucA:NuiA[31]. Structure of the non-specific nuclease NucA (blue) in complex with its specific inhibitor NuiA (green). The magnesium ion (light green) indicates the active site displaying a ββα-metal finger fold. NuiA directly interacts with the active site of NucA, mimicking the nucleic acid substrate. The C-terminus of NuiA coordinates the catalytically important Mg^{2+} ion. (PDB code: 2O3B)

1.1.4.4 *Sma*Nuc

*Sma*Nuc is a non-specific extracellular nuclease secreted by *Serratia marcescens*, a gram-negative anaerobic bacterium[71]. Although *Sma*Nuc shows a high structural and sequence similarity to NucA especially at the active site, regulation within the host cells is presumably not mediated by a specific inhibitor (see Table 3). *Sma*Nuc exists in the cell in an inactive state as it needs the formation of two intramolecular disulfide bonds (see Figure 11) to gain activity during export in an oxidizing environment[72]. Due to the fact that *Sma*Nuc is inactive in a cellular environment, cloning and expression in *E. coli* is rather unproblematic making this nuclease to the best characterized member of the DRGH/H-N-N family. Furthermore comparison of the crystal structure of non-specific *Sma*Nuc with the highly specific homing endonuclease I-PpoI revealed a similar active site for both nucleases (see Figure 3) indicating common ancestors for all types of ββα-metal finger nucleases which became a dogma in the nuclease research field[11].

The nuclease prefers double-stranded DNA in A-form and occurs as a dimer like dEndoG (see Figure 11). Mutational analyses and crystallization studies revealed a small dimer interface including amino acid 180-184 which is located on the opposite site of the active center[73]. A single amino acid replacement (H184R) concerning the dimer interface promotes the

monomeric state but does not affect the catalytic activity indicating the autonomous behavior of the two active sites (nicking activity)[74].

*Sma*Nuc is commercially used (Benzonase™) for the effective removal of nucleic acids for biochemical applications due to its high stability and for a conditional suicide system in yeast. For the second approach *Sma*Nuc expression is regulated by a glucose-repressed ADH2 promoter. Expression of the nucleases upon glucose depletion even in the reducing cytoplasm of yeast leads to effective killing of the host cells. This system is used to prevent an environmental spreading of genetically modified yeast strains[75].

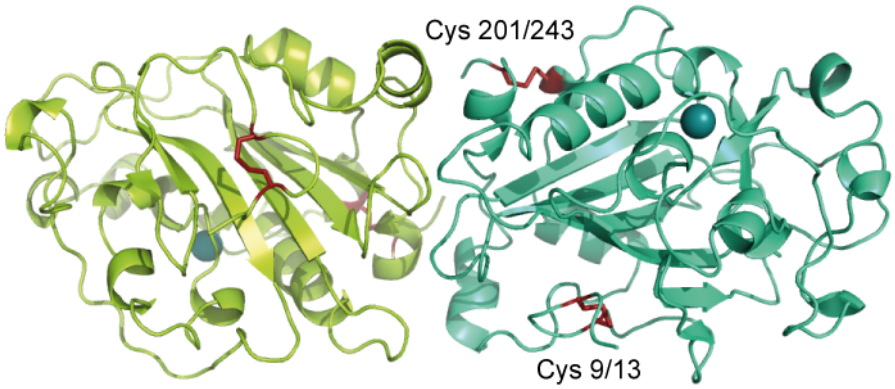

Figure 11: Crystal structure *Sma*Nuc dimer[32]. *Sma*Nuc (green/cyan) exists in the cell in an inactive state as it needs the formation of two intramolecular disulfide bonds (red) between Cys201/243 and Cys9/13 to gain activity. Dimer formation does not concern the active sites (indicated by two magnesium ions, blue) cleaving DNA and RNA independently from each other. (PDB code: 1QAE)

Table 3: Mechanism of inhibition for selected non-specific nucleases.

Nuclease	Inhibitor	Type of inhibition
NucA	NuiA	Tight binding to the active site
*Sma*Nuc	No inhibitor known	Activation due to disulfide bond formation
Vvn	No inhibitor known	Activation due to disulfide bond formation
Colicin E7/E9	Inhibitor Im7/9	Binding to a protein exosite, blocking substrate binding
EndA	No inhibitor known	unknown
Drosophila EndoG	Inhibitor dEndoGI	Tight binding to the active site
CAD	Inhibitor ICAD	Chaperone function, blocking dimer formation

1.1.4.5 EndA

Streptococcus pneumoniae (pneumococci) are pathogenic gram-positive bacteria causing serious infections of the respiratory system of mammalians and lethal meningitis[76]. Other typical streptococcal diseases like STSS (streptococcal toxic shock syndrome), NF (necrotizing fasciitis) and tonsillitis are triggered by group A streptococci (GAS, *Streptococcus pyogenes*)[77]. This group of bacteria exhibit several serotypes expressing different extracellular virulence proteins e.g. streptodornases[78]. Several of these virulence factors are phage encoded supporting recombination and spreading of novel arranged pathogenic genes presumably creating the high diversity of GAS strains[79]. GAS strains express several types of streptodornases which support the infiltration of the innate immune response of mammalians resulting in invasive infections[80; 81]. The reason for this was discovered recently as streptodornases are able to degrade the DNA scaffold of neutrophil extracellular traps (NETs)[82]. NETs are released by activated neutrophils and consist of expelled chromatin, elastase and other antimicrobial proteins and kill captured bacteria (see Figure 12)[83]. This effect was also detected for *Stretococcus peneumoniae* (TIGR4) which express the competence associated nuclease EndA. This surface bound nuclease has a major influence on the degradation of extracellular DNA to allow pneumococci to escape NETs and regain viability[84].

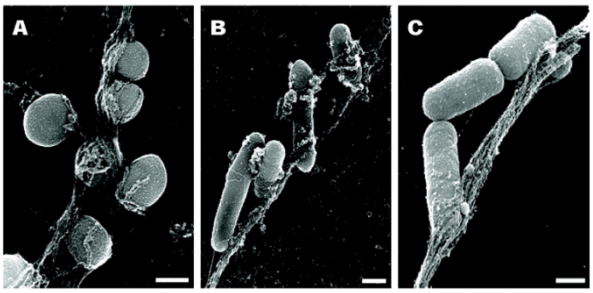

Figure 12: Bacteria associated with neutrophil fibers. A) SEM of *S. aureus*, B) *S. typhimurium* and C) *S. flexneri* trapped by NETs. Bar, 500 nm. Modified after[83].

All streptodornases are closely related and exhibit a high sequence similarity to EndA[85]. Both types of nucleases revealed an H-N-N motif but the conserved DRGH sequence could only be detected for EndA, whereas most streptodornases contain a DRSH motif. Further analysis of

the N-terminal sequence of streptodornases predicted a leader peptide for extracellular secretion. This signal peptide is cleaved off during secretion at least shown for streptodornase Sda1[85]. In contrast to that, EndA needs no further processing of the N-terminal membrane localization signal for surface exposure[5].

As shown by mutational analyses EndA is the major nuclease involved in degradation of incoming DNA to single-stranded fragments during pneumococci transformation[5; 86]. So far, no specific inhibitor or regulative mechanism for EndA was discovered. Investigation of one downstream and two upstream adjacent genes (epuA, epuB and epdA) revealed that none of the gene products are essential for viability or are involved in the transformation mechanism[5].

Interestingly a detailed analysis of streptodornases and their cleavage properties showed that bacterial tRNA has the ability to inhibit nucleolytic activity[87]. This might explain the fact that cloning and expression of this non-specific nucleases in *E. coli* is unproblematic although no inhibitory proteins are known for these nucleases.

1.2 Intention of the work

The two non-specific nucleases NucA and *Sma*Nuc display a high structural and sequence similarity. Nevertheless *Sma*Nuc is inactive in the cell as it needs the formation of disulfide bonds and is presumably not regulated by a protein inhibitor whereas NucA and NuiA form a tight binding complex. Due to the homology of the two proteins especially within the active sites it should be possible to engineer a functional protein inhibitor for *Sma*Nuc on the basis of NuiA via random mutagenesis and selection. One approach for such a selection system is presented in this work, which might be useful to engineer inhibitor proteins for other DRGH motif containing nucleases such as EndA. This competence associated nuclease displayed on the surface of *Streptococcus pneumoniae* is known to be a pathogenic factor and a potential inhibitor might serve as starting point for future drug design[31]. This work presents the first molecular/biochemical characterization of recombinant EndA and provides evidence that the non-specific nuclease bears an H-N-H/N active site fold.

2 Materials and methods

2.1 Materials

2.1.1 Chemicals and biochemicals

All chemicals used were of high purity and are listed in Table 4.

Table 4: Chemicals.

Chemical/ biochemical	Company
Acetic acid	Roth
40 % Acrylamide:bisacrylamide solution (29:1)	AppliChem
40 % Acrylamide:bisacrylamide solution (19:1)	AppliChem
Agar	AppliChem
Agarose Ultra Pure™	Invitrogen
Aluminium sulfate	AppliChem
Ampicillin	AppliChem
APS (Ammoniumpersulfate)	Merck
Boric acid	Merck
Bromophenol blue	Merck
BSA (Bovine serum albumin)	NEB
Calcium chloride	Merck
Chloramphenicol	AppliChem
Chloric acid, 37 %	Merck
Coomassie® Brilliant blue G250	AppliChem
DEPC (Diethylpyrocarbonate)	Fluka
DTT (1,4-dithiothreitol)	AppliChem
dNTPs (deoxyribonucleotides)	Fermentas
EDTA (Ethylene diamine tetraacetate)	AppliChem
Ethanol	Merck
Ethidium bromide	Roth
Formamide	Merck
Glycerol	AppliChem
Glycin	Merck
High molecular weight herring sperm DNA	Sigma Aldrich
Imidazole	Merck

IPTG (Isopropyl β-D-1-thiogalactopyranoside)	AppliChem
Lubrol	Sigma Aldrich
Kanamycin	AppliChem
Magnesium chloride	Merck
MES monohydrate [2-(N-Morpholino)ethanesulfonic acid]	AppliChem
Methanol	VWR
2-Mercaptoethanol	Merck
MOPS (3-(N-morpholino)propanesulfonic acid)	Biomol
O-Phosphoric acid, 85 %	Roth
PEG 4000 (polyethylene glycol) 50 %	Fermentas
Potassium chloride	Merck
Potassium dihydrogen phosphate	Merck
di-Potassium hydrogen phosphate	Merck
Nonfat dried milk powder	AppliChem
Sucrose	AppliChem
Sodium acetate, anhydrous	Applichem
Sodium chloride	Merck
Sodium hydroxide	Merck
SDS (sodium dodecyl sulfate)	AppliChem
TEMED (Tetramethylethylenediamine)	Merck
Tetracycline	AppliChem
Tris	Merck
Triton X-100	AppliChem
Tryptone	AppliChem
Tween 20	Merck
Urea	AppliChem
Yeast extract	AppliChem
Xylene cyanol	Merck

2.1.2 *E. coli* Strains

Table 5: *E. coli* Strains.

Strain	**Company**	**Genotype**
XL1-Blue MRF'	Stratagene	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB lacIqZΔM15* Tn*10* (Tet^R)].
BL21-Gold(DE3)	Stratagene	*E. coli* B F^- *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte
Origami	Novagen	*Δ(ara-leu)7697 ΔlacX74 ΔphoA Pvu*II *phoR araD139 ahpC galE galK rpsL* F'[*lac*$^+$ *lacI*^{q} *pro] gor522*::Tn*10 trxB* (Kan^R,Str^R, Tet^R)

2.1.3 Electrophoresis ladders

All ladders were prepared and used according to the supplied manual.

Table 6: Ladders.

Name	Type	Company
PageRuler™ Unstained Protein Ladder	Protein ladder	Fermentas
PageRuler™ Prestained Protein Ladder	Protein ladder, Western blot analysis	Fermentas
GeneRuler™ 1 kb DNA Ladder	DNA ladder	Fermentas
RiboRuler™ RNA Ladder, Low Range	RNA ladder	Fermentas

2.1.4 Kits

Table 7: Kits and Systems.

Name	Application	Company
PURExpress™ *In Vitro* Protein Synthesis Kit	Cell-free transcription/translation	New England Biolabs
PureYield™Plasmid Miniprep System	Small-scale plasmid preparation	Promega
PureYield™Plasmid Midiprep System	Medium-scale plasmid preparation	Promega
PureYield™Plasmid Maxiprep System	Large-scale plasmid preparation	Promega
RNeasy Mini Kit	RNA purification and concentration	Qiagen
Wizard® SV Gel and PCR Clean up system	Purification of DNA fragments from agarose gels or after PCR.	Promega
T7 Transcription Kit	*In Vitro* transcription of RNA	Fermentas

2.1.5 Nucleases

2.1.5.1 Restriction endonucleases

All restriction endonucleases were supplied either by New England Biolabs (NEB) or Fermentas and used in the recommended activity buffer according to the manual.

Table 8: Restriction endonucleases.

Resctriction endonuclease	**Cleavage site (5'-3')**
BamHI	G|GATCC
BglII	A|GATCT
EcoRI	G|AATTC
HindIII	A|AGCTT
KasI	G|GCGCC
NcoI	C|CATGG
SalI	G|TCGAC
SphI	G|CATGC
XhoI	C|TCGAG

2.1.5.2 Non-specific nucleases

Table 9: Non-specific nucleases.

Non-specific nuclease	**Source**
DNase I	Fermentas
NucA	Heike Büngen

2.1.6 Nucleic Acids

2.1.6.1 Plasmids

Table 10: Plasmids for protein expression.

Plasmid	Source
pETM-30	EMBL Protein Expression and Purification Facility
pQE-30	Qiagen
pBBnuiA/nucA(I)	Christian Korn
pHisnuc	Wolfgang Wende
pHisnuc K55A, pHisnuc H89N	Peter Friedhoff
pET25b(+) EndA2a2 H160A	Patrick Schäfer
pQE-30 NuiA	Heike Büngen
pQE30-DsRem	Heike Büngen

2.1.6.2 Primer

Table 11: Primer. Introduced restriction endonuclease cleavage sites used for cloning procedures are indicated in bold letters and introduced mutations in red. Corresponding sequences for epitope tagging are underlined.

Name	Sequence (5'-3')	Comment
74-76 random	TTAGCAGAGCCACTACTCCCNNN NNNNNNTATGAGGATGAAGAAAATGC	Mutagenesis NuiA
74-76 us	GGGAGTAGTGGCTCTGCTAA	Mutagenesis NuiA
78-80	CTACTCCCCAAGACTGGTATGAA AATGCTGTAGTTGCTAA	Mutagenesis NuiA
78-80 us	ATACCAGTCTTGGGGAGTAG	Mutagenesis NuiA
109-111 random	CGCAGGTGTATCGACTGGGTNNNN NNNNNCTTGATGTTTATGTTATTGG	Mutagenesis NuiA
109-111 us	ACCCAGTCGATACACCTGCG	Mutagenesis NuiA
3'EndA XhoI	CTC**CTCGAG**TCACTGA GTTACAGTTACTTCTC	XhoI
3'nucA	CCGGC**GTCGAC**TAAT TATCAACTTTACTCTC	SalI

3'NucS SalI	GTC**GTCGAC**TCAGTTTT TGCAGCCCATGAG	SalI
3'nuiA BglII	AGA**AGATCT**TCAAGTT TCCACAACTTTAG	BglII
5'EndA NcoI	CCA**CCATGGG**CGCACC TAATAGTCCCAAAAC	NcoI
5'NucA flag SphI	GCA**GCATGC**GGCG**CCATGG**ATT ACAAGGATGACGACGATAAGCA AGTGCCACCATTAACTGAA	SphI, NcoI, FLAG-tag
5'NucS	GTC**GTCGAC**GACACG CTCGAATCCATCGAC	SalI
5'NucS flag SphI	GCA**GCATGC**ACTAGTATGGATTA CAAGGATGACGACGATAAGGACA CGCTCGAATCCATCGAC	SphI, FLAG-tag
5'nuiA(I)	GGCATGCAT**GGATCC**ACC AAAACCAACTCAGAAATT	BamHI
5'nuiA EcoRI	GAA**GAATTC**TTAAAGAGGAGAAAT TAACTATGACCAAAACCAACTCAGA	EcoRI
D157Afor(C)	CCCATGCAGTCGCAAGAGGTCATTT	EndA Mutation to A
D157Arev(B)	AAATGACCTCTTGCGACTGCATGGG	EndA Mutation to A
DsReMrmcs	GTC**GTCGAC**TCGCGAGGCGCCGA TATCCTGGGAGCCGGAGTGGCGGGC	SalI
EndAFLAGrev(D)	TATTCATTACTTATCGTCGTCATCCTT GTAATCGCCTCCCTGAGTTACAGTTAC	FLAG-tag
fQE	CCCGAAAAGTGCCACCTG	Sequencing
N179Afor(C)	CCTCAACAAGCGCACCTAAAAACAT	EndA Mutation to A
N179Arev(B)	ATGTTTTTAGGTGCGCTTGTTGAGG	EndA Mutation to A
N182Afor(C)	GCAATCCTAAAGCAATTGCTGTTCA	EndA Mutation to A
N182Arev(B)	TGAACAGCAATTGCTTTAGGATTGC	EndA Mutation to A
N191Afor(C)	CAGCCTGGGCAGCACAGGCACAAGC	EndA Mutation to A
N191Arev(B)	GCTTGTGCCTGTGCTGCCCAGGCTG	EndA Mutation to A
NucA KasI fw	GGC**GGCGCC**GCGGCCGCCAT GCAAGTGCCACCATTAACTG	KasI
NucA SalI rw	GTC**GTCGAC**GAGCTCCT AATTATCAACTTTACTCTC	SalI
R69 random	ACATTGACAGCTTTTTTAGCNN NGCCACTACTCCCCAAGACTG	Mutagenesis NuiA
R69 us	GCTAAAAAAGCTGTCAATGT	Mutagenesis NuiA
R158Afor(C)	ATGCAGTCGATGCAGGTCATTTGTT	EndA Mutation to A

R158Arev(B)	AACAAATGACCTGCATCGACTGCAT	EndA Mutation to A
RBSEndAfor(A)	TAAG[AAGGAGA]TATACCAAT GGCACCTAATAGTCCCAAA	[RBS]
RevMutA160H	ACCACCGATTAATGCATAGCC TAACAAATGACCTCGATCGAC	Back mutation EndA
RMutA160HNsiI	GTCGATCGAGGTCATTTGTTAG GCTATGCATTAATCGGTGGT	Back mutation EndA
rQE	GTTCTGAGGTCATTACTGG	Sequencing
SphIDRMf flag	GCA**GCATGC**ATGATTACAAG GATGACGACGATAAGATGGA CAACACCGAGGACGTC	SphI, FLAG-tag
uniT7RBSfor(A´)	GAAAT{TAATACGACTCACTAT A}GGGAGACCACAACGGTTTCCCT CTAGAAATAATTTTGTTTAACTTT AAG[AAGGAGA]TATACC	{T7 promoter sequence}, [RBS]

2.1.6.3 Substrates

Table 12: Oligo-DNA substrates.

Name	Sequence (5'-3')	Modification 5'	Modification 3'
FS2B	GAAAAAACCCCCCTTTTTTC (molecular beacon)	6-Fam	BHQ-1

Table 13: Circular DNA substrates.

Name	Feature	Size	Company
pBluescript SK(+)	Double-stranded plasmid	2958 bp	Stratagene
ΦX174 RF I DNA	Double-stranded, circular form of ΦX174 DNA	5386 bp	New England Biolabs
ΦX174 RF II DNA	Double-stranded, nicked, circular from of ΦX174 DNA	5386 bp	New England Biolabs
ΦX174 Virion DNA	Single-stranded, viral DNA	5386 b	New England Biolabs

2.1.7 Other enzymes/proteins

Table 14: Other enzymes/proteins.

Enzyme/protein	Company
T4 DNA Ligase (5 U/µl)	Fermentas
RNase Inhibitor (40,000 units/ml)	New England Biolabs
TEV Protease (His-tagged, 2 U/µl)	MoBiTec

2.1.8 Protein purification

Table 15: Materials protein purification.

Device	Feature	Company
Ni-NTA Agarose	Nickel-charged resin	Qiagen
CENTRIPLUS® Centrifugal Filter Devices	MW cut-off: 10,000 Da	MILLIPORE
Dialysis membrane	MW cut-off: 14.000 Da	Roth
Protino® Ni-IDA 2000 packed columns	Gravity-flow column chromatography	Macherey-Nagel
Superdex™75, size exclusion column	Separation 3-70 kDa proteins	GE Healthcare

2.1.9 Polymerases

Table 16: Polymerases.

Polymerases	Application	Source
Taq DNA Polymerase	Routine PCR, error-prone PCR	Heike Büngen, Ina Dern
Pfu DNA Polymerase	High fidelity PCR	Heike Büngen, Ina Dern

2.1.10 Materials Western blotting

2.1.10.1 Antibodies

Table 17: Antibodies.

Antibody	Company
ANTI-FLAG M2-Peroxidase antibody	SIGMA
Anti-GST Antibody	Pharmacia Biotech
Anti-Sheep/Goat IgG-POD, Fab fragments	Boehringer Mannheim

2.1.10.2 Other materials for Western blotting

Table 18: Other materials for Western blotting.

	Name	Company	Feature
Membrane	Amersham Hybond™-P	GE Healthcare	Hydrophobic polyvinylidene difluoride (PVDF) membrane
Filter paper	Blotting paper MN 218 B	Macherey-Nagel	For blotting procedures
Film	Amersham Hyperfilm™ ECL	GE Healthcare	For chemiluminescent-based detection
Detection reagents	Amersham ECL Plus™ Western Blotting Detection Reagents	GE Healthcare	Detection of Horseradish Peroxidase labeled antibodies

2.2 Methods

2.2.1 Microbiological methods

2.2.1.1 Media

LB Medium and LB agar plates

10 g/l tryptone
5 g/l yeast extract
5 g/l NaCl

The pH was adjusted to 7.5 with NaOH and the medium autoclaved. For LB agar plates up to 1.5 % (v/w) agar was added before autoclaving. After autoclaving and cooling (40 °C) the medium antibiotics in appropriate concentrations (see Table 19) could be added and plates were poured in Petri dishes.

Table 19: Concentration of used antibiotics.

Name	Stock concentration	Final concentration
Ampicillin	25 mg/ml (in H_2O)	100 µg/ml
Kanamycin	25 mg/ml (in H_2O)	25 µg/ml
Tetracycline	12 mg/ml (in 50 % Ethanol)	10 µg/ml

2.2.1.2 Preparation of electrocompetent *E. coil* cells

A 25 ml LB preculture with appropriate antibiotic was inoculated with the needed *E. coli* strain and grown over night at 37 °C. The preculture (5 ml) was transferred to a 500 ml LB culture and grown to an OD_{600} of 0.5-0.8. The culture was incubated on ice for 30 min and the cells harvested by centrifugation (15 min, 4000 × *g*, 4 °C). The pellet was resuspended with 250 ml 10 % sterile, cold glycerol and centrifuged again. This washing step was repeated twice with 150 ml and 40 ml 10 % glycerol. In a final step the cells were carefully resuspened

in 2 ml 10 % glycerol and aliquots of 80 µl prepared. The aliquots were quick-frozen in liquid nitrogen and stored at -80 °C. All steps were performed with sterile equipment.

2.2.1.3 Transformation of *E. coil* by electroporation

Needed number of aliquots were thawed on ice and either 50-150 ng of plasmid DNA or 4-5 µl ligation reaction were added to the cells and the solution gently mixed. The DNA/cell suspension was transferred between the electrodes of prechilled electroporation cuvettes. An electrical impulse was applied to the cuvettes (U=1350 V). Transformed cells were rinsed with 500 µl LB medium, transferred to an Eppendorf tube and incubated 1 h at 37 °C. The cell/medium mixture was centrifuged for 3 min at 4000 × *g*, 250 µl media removed and the pellet resuspended in the residual medium. A tenth part was plated on an agar plate with required antibiotics and the remaining mixture on a second plate. The plates were incubated over night at 37 °C.

2.2.1.4 Plasmid preparations

All plasmid preparations were performed with the PureYield™Plasmid Mini/Midi/Maxiprep System according to the manual.

2.2.2 Molecular biology methods

2.2.2.1 Electrophoresis

2.2.2.1.1 Agarose gel electrophoresis

TBE buffer	5× Loading dye
100 mM Tris	250 mM EDTA
100 mM Boric acid	25 % Sucrose
2.5 mM EDTA	1 % SDS
	0.01 % Bromphenol blue
	0.01 % Xylene cyanol
The pH was set automatically at 8.3. The buffer was filtered.	The pH was adjusted to 8.0 with NaOH.

For all analytical and preparative agarose gels 0.8 % agarose was melted in TBE buffer and appropriate gels were prepared. Samples were diluted with 5× loading dye, transferred to the wells and an electrical field applied to the gel chamber with certain voltage according to the size of the gel. A DNA ladder was used as reference (see Table 6). The agarose gels were stained with ethidium bromide (stock solution 1 %, for staining dilution 1:1000 in water) and documented. DNA fragments were cut from preparative gels and the DNA isolated with the Wizard® SV Gel and PCR Clean up system according to the manual.

2.2.2.1.2 SDS gel electrophoresis

Stacking gel 6 %	Separation gel 15 %
1 % SDS	1 % SDS
130 mM Tris/HCl pH 6.8	420 mM Tris/HCl pH 8.0
6 % acrylamide (29:1)	15 % acrylamide (29:1)
0.08 % APS	
0.2 % TEMED	

The solutions were prepared separately and APS and TEMED added to the separation gel, mixed and poured into the assembled gel caster for polymerization. After polymerization of the separation gel the stacking gel was poured and the comb assembled.

Laemmli buffer	5× Laemmli loading buffer
25 mM Tris	160 mM Tris/HCl pH 6.8
0.1 % SDS	2 % SDS
800 mM Glycin	2-Mercaptoethanol
	40 % Glycerin
The pH was set automatically at 8.3.	0.1 % Bromphenol blue

Laemmli loading buffer was added to the protein samples and loaded on the gel together with a protein ladder (unstained, see Table 6). After separating the proteins (35 mA, 60 min) the gel chamber was disassembled and the gel washed three times for 10 min with water. The bands were visualized by Coomassie staining and afterwards washed with water to reduce unspecific staining. For protein expression analysis samples from *E. coli* cultures (300 µl – 800 µl according to the OD_{600}) were resuspended after centrifugation in 30 µl Laemmli loading buffer and heated up to 95 °C for 3 min, chilled on ice and 8 µl loaded on an SDS-PAGE.

Coomassie staining solution

0.1 % Coomassie® Blue G 250
10 % Ethanol
5 % Aluminum sulfate
2 % Phosphoric acid

The solution was mixed over night and filtered the next day. The staining solution was used up to three times.

2.2.2.1.3 Denaturing electrophoresis

5 % acrylamide/urea gel

7 M urea
5 % acrylamide (19:1)
0.08 % APS
0.2 % TEMED
In 10 ml TBE (see 2.2.2.1.1 Agarose gel electrophoresis)

All components were mixed and poured in an assembled gel caster and a comb adjusted.

2× RNA loading dye

10 % glycerol
0.025 % SDS
0.025 % bromophenol blue
0.025 % xylene cyanol
0.5 mM EDTA
Filled up with formamide

The samples were mixed with loading dye, heated for 10 min at 65 °C and chilled on ice.

The gel was prepared by an electrophoresis in TBE buffer at 8 Watt for 30 min without samples. After cleaning the wells, the samples were loaded and separated at 3 Watt for 1 h. The gel was stained with ethidium bromide (see 2.2.2.1.1 Agarose gel electrophoresis) and documented.

2.2.2.2 Western blot analysis

An SDS-PAGE was performed according to chapter 2.2.2.1.2 and a prestained protein ladder used (see Table 6).

Transfer buffer	Tris buffered Saline-Tween 20 (TTBS)
50 mM Tris	100 mM Tris/HCl pH 7.5
40 mM Glycin	150 mM NaCl
0.05 % SDS	0.1 % Tween 20
20 % Methanol	

The membrane was activated for 10 sec in methanol and washed with water. Membrane, filter paper, and SDS gel were incubated in transfer buffer for 15 min and assembled between cathode and anode of a semi dry blotter (BlueFlash, Serva) according to the manual. The transfer was performed at 45 mA for 2 h. After disassembling the membrane was incubated in 50 ml blocking solution (4 % nonfat dried milk powder in TTBS) over night at 4 °C and washed the next day with TTBS. For each antibody 10 ml fresh blocking solution was prepared, transferred to the membrane and the antibody added (dilution: see Table 20). The membrane was incubated (light protected) for 1 h at RT and washed with TTBS 3 times for 10 minutes. If required a secondary antibody was added to the membrane. After washing the membrane the binding was visualized by Amersham ECL Plus™ Western Blotting Detection Reagents according to the manual.

Table 20: Antibody dilution.

Antibody	**Dilution**
Anti-FLAG	1:500
Anti-GST (primary antibody)	1:2000
Anti-Sheep/Goat (secondary antibody)	1:2500

2.2.2.3 Polymerase chain reaction (PCR)

2.2.2.3.1 *Taq* and *Pfu* PCR

Table 21: Overview PCR cycle. For a PCR reaction step 2 – 3 were repeated up to 30 times.

	***Taq* PCR**	***Pfu* PCR**
1. Initial denaturation step	3 min, 95 °C	
2. Denaturation step	1 min, 95 °C	
3. Primer annealing step	1 min, 55 °C	
4. Extending step	A 1 kbp fragment corresponds to 1 min extension time, 72 °C	A 1 kbp fragment corresponds to 1.5 min extension time, 68 °C
5. Final extending step	3 min, 72 °C	3 min, 68 °C

Taq PCR	*Pfu* PCR
Primer forward (500 nM)	Primer forward (200 nM)
Primer reverse (500 nM)	Primer reverse (200 nM)
1× *Taq* activity buffer	1× *Pfu* activity buffer
dNTPs (200 µM each)	dNTPs (200 µM each)
Taq DNA polymerase (5U)	*Pfu* DNA polymerase (5U)
Template (resuspended colonies, or plasmid DNA 10-50 ng)	Template (plasmid DNA 10-50 ng)

The final volume for each PCR reaction was 50 µl. The temperature cycling was performed according to Table 21. *Taq* PCR was mainly used for routine screening of *E. coli* colonies (colony PCR). *Pfu* PCR was performed to amplify DNA fragments for cloning procedures due to the proofreading function (3'-5' exonuclease activity) of the polymerase.

10× *Taq* activity buffer	10× *Pfu* activity buffer
100 mM Tris/HCl pH 8.8	200 mM Tris/HCl pH 9.0
500 mM KCl	100 mM $(NH_4)SO_4$
1 % Triton X-100	100 mM KCl
15 mM $MgCl_2$	1 % Triton X-100
	25 mM $MgSO_4$

2.2.2.3.2 Overlap extension PCR (OEP)

Overlap extension PCR was used for site specific mutagenesis[88]. Two separate PCR reactions were set up to amplify two overlapping gene segments of the full length sequence (see Figure 13). The two internal primer b and c were complementary and mediate the amino acid replacement. Each internal primer was combined with the corresponding flanking primer a and d and the reactions performed as *Pfu* PCR (see 2.2.2.3.1 *Taq* and *Pfu* PCR). The PCR products AB and CD were separated on an agarose gel, sliced from the gel and purified (Wizard® SV Gel and PCR Clean up system). In a second PCR the two fragments with complementary ends containing the mutated sequence were used as template (50 ng each) to amplify the full length gene with the flanking primer a and d. The fragment containing the full length CDS was separated on an agarose gel, sliced from the gel and purified.

The obtained DNA fragments coding for *endA* wt (template: pET25b(+) EndA2a2 H160A) and variants (corresponding primers see Table 22) were used as template for an *in vitro* protein synthesis reaction of the proteins (see 2.2.2.6.1 *In vitro* protein synthesis of EndA). OEP was performed for mutation of selected amino acids of NuiA (template: pQE-30 NuiA, see Table 23) and the obtained mutated genes coexpressed with *Sma*Nuc in a bicistronic arrangement in *E. coli* for the selection of functional inhibitor variants (see 3.3.3 Selection of functional inhibitor variants).

Table 22: OEP primer for *endA* wt and variant templates. The two PCR products (template: pET25b(+) EndA2a2 H160A) for *endA* wt were used as a template to amplify full length wt *endA* with primer a´ (uniT7RBSfor(A´)) and d (see Figure 13). *EndA* wt sequence served as a template to generate variants D157A, R158A, N179A, N182A and N191A.

	Primer for product AB	**Primer for product CD**
endA wt	a: RBSEndAfor(A) b: RevMutA160H	c: RMutA160HNsiI d: EndAFLAGrev(D)
endA D157A	a: RBSEndAfor(A) b: D157Arev(B)	c: D157Afor(C) d: EndAFLAGrev(D)
endA R158A	a: RBSEndAfor(A) b: R158Arev(B)	c: R158Afor(C) d: EndAFLAGrev(D)
endA N179A	a: RBSEndAfor(A) b: N179Arev(B)	c: N179Afor(C) d: EndAFLAGrev(D)
endA N182A	a: RBSEndAfor(A) b: N182Arev(B)	c: N182Afor(C) d: EndAFLAGrev(D)
endA N191A	a: RBSEndAfor(A) b: N191Arev(B)	c: N191Afor(C) d: EndAFLAGrev(D)

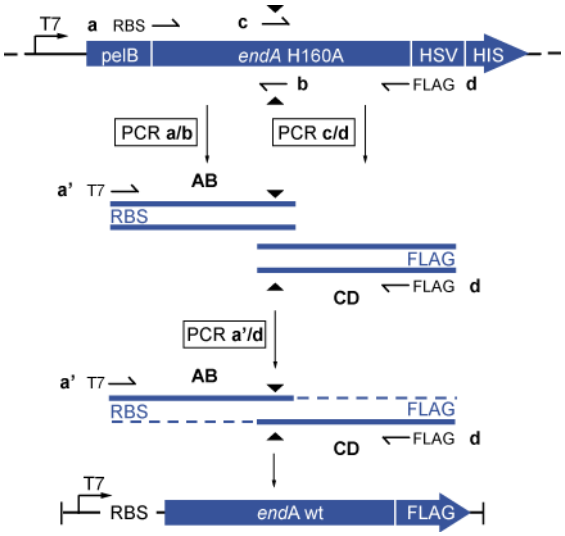

Figure 13: Overlap extension PCR for *endA* wt and variants. *EndA* wt sequence served as a template to generate variants D157A, R158A, N179A, N182A, and N191A. The two PCR products for each variant were used to amplify full length mutated *endA* with primer a´ and d.

Table 23: OEP primer for mutation of selected amino acid residues of NuiA. The two PCR products (template: pQE-30 NuiA) for each variant were used as a template to amplify full length mutated *nuiA* with primer a and d.

	Primer for product AB	**Primer for product CD**
nuiA R69	a: 5'nuiA EcoRI b: R69 random	c: R69 us d: 3'nuiA
nuiA 74-76	a: 5'nuiA EcoRI b: 74-76 random	c: 74-76 us d: 3'nuiA
nuiA 78-80	a: 5'nuiA EcoRI b: 78-80	c: 78-80 us d: 3'nuiA
nuiA 109-111	a: 5'nuiA EcoRI b: 109-111 random	c: 109-111 us d: 3'nuiA

2.2.2.3.3 Error-prone PCR NuiA

10× Mut dNTPs	PCR reaction
2 mM dGTP	Primer forward 5'nuiA EcoRI (300 nM)
2 mM dATP	Primer reverse 3'nuiA BglII (300 nM)
10 mM dCTP	1× Mut dNTPs
10 mM dTTP	1× *Taq* activity buffer
	Taq DNA polymerase (10 U)
	Template 50 ng (pBBnuiA/nucA(I))

The final volume for each PCR reaction was 50 µl. The temperature cycling was performed according to Table 21 with an extension time of 1 min. Biased dNTPs, exceeding amount of *Taq* DNA polymerase and increased extension time led to higher mutation rates during PCR[89]. Obtained mutated genes were coexpressed with *smaNuc* in a bicistronic arrangement in *E. coli* for selection of functional inhibitor variants (see 3.3.3 Selection of functional inhibitor variants).

2.2.2.4 Ligation reaction for established plasmids

PCR products of genes were ligated into expression vectors to express recombinant proteins. The sequence of interest was amplified via *Pfu* PCR with appropriate primers containing a restriction endonuclease cleavage site. The PCR product was separated on an agarose gel, sliced out from the gel and purified. Two separate cleavage reactions with corresponding restriction endonucleases were performed according to the manual supplied with the restriction enzyme to digest the insert and the vector for complementary ends and the DNA purified (Wizard® SV Gel and PCR Clean up system).

Ligation reaction

25 ng vector DNA
Insert 5:1 molar ratio over vector
2 µl 10× reaction buffer (supplied by Fermentas)
2 µl 50 % PEG 4000
1 µl T4 DNA Ligase
Water up to 20 µl

The reaction was incubated on RT for 3 h and heat inactivated for 10 min at 65 °C before transformation of strain XL1-Blue MRF' with the ligated DNA.

2.2.2.4.1 Bicistronic plasmids

Table 24: Cloning procedure for bicistronic plasmids. Overview of required templates and primers to establish the indicated plasmids shown in Figure 14. All reactions were performed as *Pfu* PCR (see 2.2.2.3.1 *Taq* and *Pfu* PCR).

	Template	**Primer**	**Obtained Plasmid**
1) PCR Bicistron	pBBnuiA/nucA (I)	5'nuiA(I) 3'nucA	pQE-BCSV
2) PCR NucA	pBBnuiA/nucA (I)	5'NucA flag SphI NucA SalI rw	pQE-NFNA
3) PCR *Sma*Nuc	pHisNuc	5'NucS flag SphI 3'NucS SalI	pQE-NFNS
4) PCR *Sma*Nuc K55A	pHisNuc K55A	5'NucS flag SphI 3'NucS SalI	pQE-NFNS K55A
5) PCR *Sma*Nuc H89N	pHisNuc H89N	5'NucS flag SphI 3'NucS SalI	pQE-NFNS H89N
6) PCR DsRedMM	pQE30-DsRem	SphIDRMf flag DsReMrmcs	pBCSV
7) PCR NucA	pBBnuiA/nucA (I)	NucA KasI fw NucA SalI rw	pBCNA
8) PCR *Sma*Nuc	pHisNuc	5'NucS 3'NucS	pBCNS

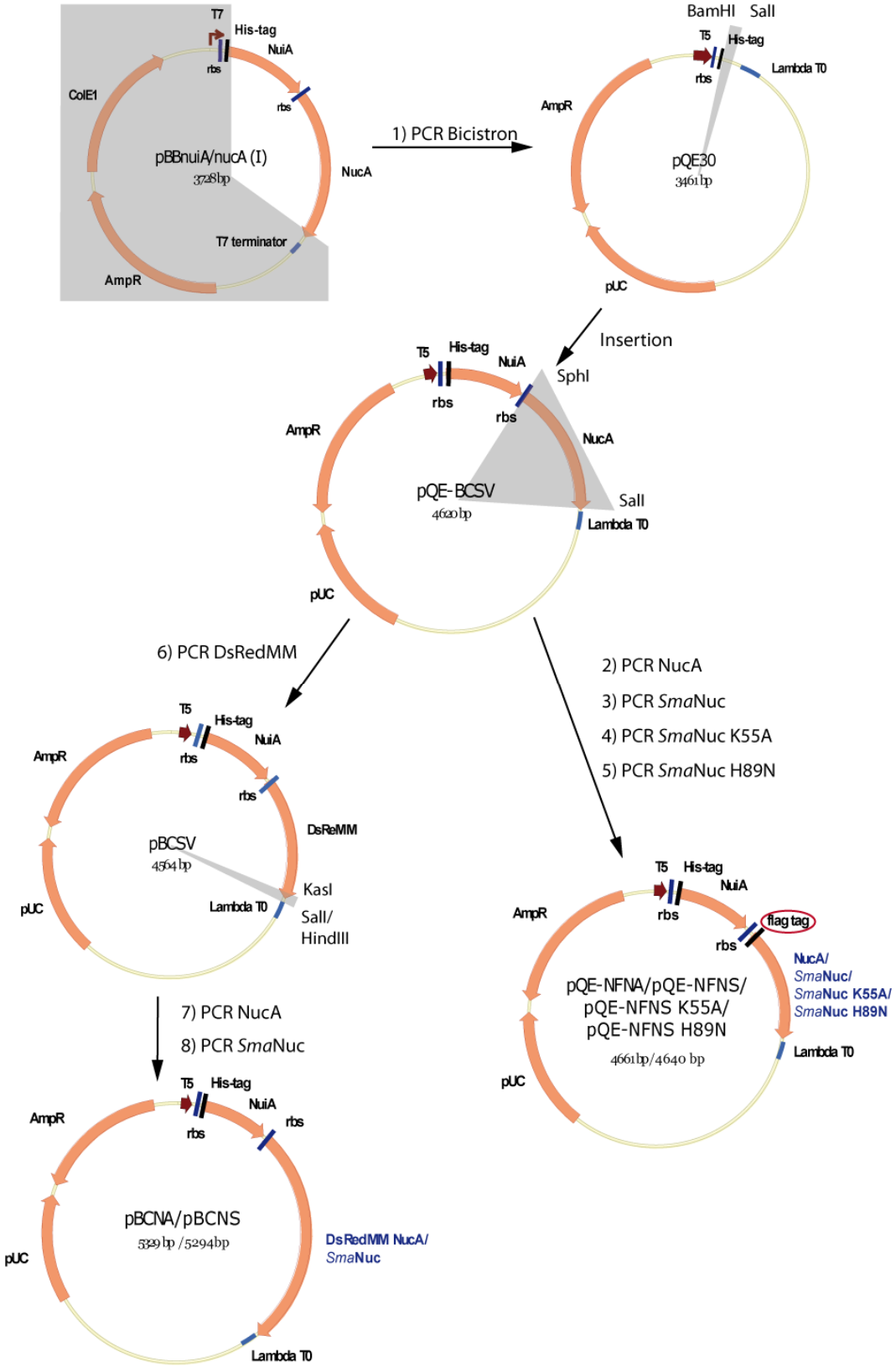

Figure 14: Cloning procedure bicistronic plasmids: Overview of the cloning procedure for the bicistronic plasmids: pBCNA/pBCNS and pQE-NFNA and pQE-NFNS, pQE-NFNS K55A, pQE-NFNS H89N. Cleaved-off or non amplified plasmid fragments are marked in gray.

2.2.2.4.2 pETM-30 EndA H160A

Table 25: Cloning procedure for pETM-30 EndA H160A. Overview of required template and primers for pETM-30 EndA H160A. A *Pfu* PCR was performed to amplify *endA* (see 2.2.2.3.1 *Taq* and *Pfu* PCR).

Template	Primer PCR EndA H160A	Plasmid backbone	Obtained Plasmid
pET25b(+) EndA2a2 H160A	3'EndA XhoI 5'EndA NcoI	pETM-30 (XhoI, NcoI)	pETM-30 EndA H160A

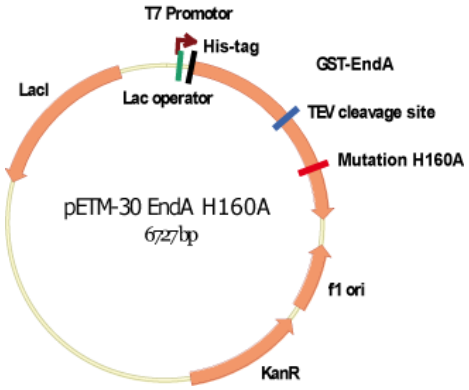

Figure 15: pETM-30 EndA H160A. The plasmid mediates expression of EndA H160A N-terminal fused to His-GST. The linker contains a TEV Protease cleavage site for clipping off the purification tags.

2.2.2.5 *In vitro* transcription

For an *in vitro* transcription of DNA the T7 Transcription Kit was used according to the manual. As the transcript served as a non-specific RNA substrate the supplied control template was used which results in a 401 b RNA transcript. To 50 µl of the transcription reaction 5 µl DNaseI were added and incubated for 25 min at 37 °C. The RNA was cleaned with the RNeasy Mini Kit according to the manual to remove degraded DNA.

2.2.2.6 Protein expression and purification

2.2.2.6.1 *In vitro* protein synthesis of EndA

In vitro protein synthesis reaction for EndA wt and variants (D157A, R158A, N179A, N182A, N191A) was performed with the PURExpress™ *In Vitro* Protein Synthesis Kit according to the manual. For each 25 µl reaction 250 ng of template DNA and 0.5 µl RNase inhibitor were added. The template DNA was generated by OEP (see 2.2.2.3.2 Overlap extension PCR (OEP)) containing a T7 promoter sequence and a RBS introduced by the two flanking primer a and a´ (see Figure 13). EndA and variants were C-terminal epitope tagged (FLAG-tag) for Western blot analysis due to amplification of the template with flanking primer d.

2.2.2.6.2 Expression and purification of EndA H160A

Transformation	Competent *E. coli* expression strain BL21-Gold(DE3) was transformed with plasmid pETM-30 EndA H160A and plated on LB agar plates containing kanamycin.
Preculture	A 25 ml LB culture containing kanamycin was inoculated with a single colony and incubated over night at 37 °C.
Main culture and expression	5 ml preculture were transferred to a 500 ml LB culture (kanamycin) and incubated at 37 °C until the culture reached an OD_{600} of 0.4. The protein expression was induced by 1 mM IPTG for 4 h at 37 °C.
Harvesting and cell disruption	The culture was spun down (4000 × *g* at 4 °C for 10 min), the supernatant decanted and the cell pellet washed with 20 ml STE. The cells were resuspended in 10 ml LEW buffer, transferred to a glass beaker and disrupted by sonication on ice (15 × 15 sec bursts and 15 sec cooling periods). The crude lysate was centrifuged (6.000 × *g* for 20 min at 4 °C), the

	supernatant transferred to a fresh tube and the pellet discarded.
Purification of GST-EndA H160A	The column (Protino® Ni-TED 2000 packed column) was equilibrated with 4 ml LEW buffer, the cleared lysate applied and the column washed twice with 4 ml HS-wash buffer. The protein was eluted 3× with 2 ml elution buffer.
Dialysis of GST-EndA H160A and TEV Protease mediated cleavage	The collected fractions were transferred to dialysis tubes and dialyzed over night at 4 °C against TEV activity buffer. The two fractions with highest concentration were mixed and concentrated with CENTRIPLUS® (Centrifugal Filter Devices) according to the manual to a concentration of 50 µM. 1 ml of His-GST tagged EndA H160A was incubated with 50 U TEV Protease over night at 4 °C.
Purification of EndA H160A (Ni-NTA Agarose)	After TEV Protease cleavage the samples were transferred to dialysis tubes and dialyzed over night against the binding buffer. 1 ml Ni-NTA slurry per ml protein solution was spun down (5 min, 200 × *g*) and the supernatant removed by pipetting. The protein solution was added to the Ni-NTA Agarose, resuspended and incubated 4 h at 4 °C. The mixture was spun down (5 min, 200 × *g* at 4 °C) and the supernatant transferred to fresh tubes. TEV Protease, GST and uncleaved protein should be bound to the matrix. The samples were again concentrated to 30 µM (see above).
Size Exclusion	A Superdex™75 column was equilibrated with size exclusion buffer and 1.5 ml of the sample loaded for each run. 1 ml fractions were collected for the appearing peak at 12.5 ml. The samples were pooled and concentrated to 40 µM (see above). BSA (1 g), Ovalbumin (0.5 g), α Chymotrypsin (0.5 g), Cytochrome C (0.2 g) were loaded on the column as protein standards and the elution peaks defined to calculate the apparent molecular mass of EndA.

Table 26: Buffer compositions for EndA H160A purification.

Buffer	
STE	10 mM Tris/HCl pH 8.0, 100 mM NaCl, 0.1 mM EDTA
LEW buffer (Protino® Purification System)	50 mM NaH_2PO_4/NaOH pH 8.0; 300 mM NaCl
HS-wash buffer	50 mM NaH_2PO_4/NaOH pH 8.0; 750 mM NaCl
Elution buffer (Protino® Purification System)	50 mM NaH_2PO_4/NaOH pH 8.0; 300 mM NaCl; 250 mM imidazole
TEV activity buffer	50 mM Tris/HCl pH 8.0, 1 mM DTT, 0.5 mM EDTA
Binding buffer	50 mM NaH_2PO_4/NaOH pH 8.0, 300 mM NaCl, 20 mM imidazole
Size exclusion buffer	50 mM NaH_2PO_4/NaOH pH 8.0, 300 mM NaCl, 0.01 % Triton X-100

2.2.3 Activity assays

2.2.3.1 Activity assays EndA

2.2.3.1.1 Sequence alignment of DRGH/H-N-H motif containing nucleases

Sequences of the aligned DRGH-motif containing nucleases were retrieved using the basic local alignment search tool BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) via the NCBI (National Center for Biotechnology Information) web server with the amino acid sequences of *Anabaena* nuclease (gi: 39041) or EndA (gi: 47374) as the query. Alignment of the primary structure of selected proteins was performed using ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html) via the web server of the European Bioinformatics Institute (EBI). According to the results obtained from our mutational and biochemical analyses the alignments were manually corrected using GeneDoc (http://www.nrbsc.org/gfx/genedoc/) or BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) software packages to match all amino acid residues with identical functions.

2.2.3.1.2 In-gel activity assay (zymography)

Unpurified protein synthesis reaction for wt and variants (3 µl; see 2.2.2.6.1 *In vitro* protein synthesis of EndA) was separated on a 15 % SDS gel containing 30 µg/ml high molecular weight herring sperm DNA. After washing 3× for 10 min with water to remove SDS the gel was incubated at RT in 1× activity buffer (40 mM Tris/HCl pH 7.7, 2 mM $MgCl_2$) to renature the nuclease and stained with ethidium bromide.

2.2.3.1.3 Single radiation diffusion (SRED) assay

Petri dishes were prepared with 20 ml of 1 % agarose solution in 40 mM Tris/HCl, pH 7.7, 2 mM $MgCl_2$ containing 30 µg/ml high molecular weight herring sperm DNA. The agarose gel was pre-stained with ethidium bromide. Crude protein synthesis reaction mixture (1 µl) and as control 1 µl DNaseI (Fermentas, 1 U/µl) with 1 µl of 40 mM Tris/HCl, pH 7.7, 2 mM $MgCl_2$ were applied to wells with a diameter of 2 mm. Plates were incubated over night at RT. The diameter [mm²] of the halos was measured[90]. For calibration, 0.062 µl, 0.25 µl, 1 µl, 1,5 µl and 4 µl of the wild type EndA protein synthesis reaction mixture were used. In the calibration curve, the diameter of the halos were plotted against the amounts of the wild type EndA protein synthesis reaction mixture. This curve was used to calculate the relative activity of the variants, whose concentration in the respective crude protein synthesis reaction mixture was estimated by a Western blot analysis (see Figure 21).

2.2.3.1.4 Nucleolytic activity on DNA

10× activity buffer	Reaction mix
400 mM Tris/HCl pH 7.5	10 µl 10× activity buffer
20 mM $MgCl_2$	10 µl unpurified protein synthesis reaction 1:10 diluted in 1× activity buffer (see 2.2.2.6.1)
	10 µl substrate DNA (150 ng/µl): ΦX174 RF I, ΦX174 RF II, ΦX174 Virion
	Filled up with water to 100 µl

The cleavage reaction was performed with unpurified protein synthesis reaction for EndA wt at RT. After 30 sec, 1, 2, 4, 8 and 16 min 10 µl aliquots were transferred to 4 µl loading dye to stop the reaction and separated on a 0.8 % agarose gel.

2.2.3.1.5 Nucleolytic activity on RNA

10× activity buffer	Reaction mix
400 mM MOPS pH 7.5	4 µl 10× activity buffer
20 mM $MgCl_2$	4 µl unpurified protein synthesis reaction (see 2.2.2.6.1)
In DEPC-treated water	250 ng RNA transcript (see 2.2.2.5)
	Filled up with DEPC-treated water to 40 µl

The cleavage reaction was performed with unpurified protein synthesis reaction for EndA wt and the supplied positive control (see 2.2.2.6.1) at RT. After 1, 2 and 15 min 10 µl aliquots were transferred to 10 µl RNA loading dye to stop the reaction and separated on a 5 % denaturing urea gel.

2.2.3.1.6 Chemical rescue of EndA H160A

Triple buffer system	10× activity buffer
500 mM Na-acetate	1:5 Dilution of triple buffer pH 5.5-9.5:
500 mM MES	100 mM Na-acetate
1 M Tris	100 mM MES
	200 mM Tris
	50 mM MgCl

Five triple buffers and 500 mM imidazole were prepared with different pH values: 5.5, 6.5, 7.5, 8.5, 9.5 adjusted with HCl or NaOH[91]. Five corresponding 10× buffers were prepared. EndA H160A was diluted to 200 nM in 1× activity buffer with pH 5.5-9.5.

	pH 5.5	pH 6.5	pH 7.5	pH 8.5	pH 9.5	Control
10× activity buffer			5 µl			1 µl
500 nM imidazole			20 µl			4 µl
pBSK (150 ng/µl)			5 µl			1 µl
200 nM EndA H160A			5 µl			-
H_2O			15 µl			4 µl

Five cleavage reactions with pH 5.5 – 9.5 were performed at RT to test the nucleolytic activity of EndA H160A rescued by imidazole. After 2, 5, 10 and 20 min 10 µl aliquots were transferred to 4 µl loading dye to stop the reaction. Five control reactions without EndA and pH 5.5-9.5 were incubated at RT for 20 min, 4 µl loading dye added and all samples separated on a 0.8 % agarose gel. The intensity of the bands was measured by TINA2.0 and the ratio of open-circular to supercoiled plasmid calculated and the velocity of the reaction (slope) plotted.

2.2.3.2 Inhibition of NucA by NuiA

The inhibition effect of NuiA (apparent inhibition constant K_i^{app}) on its specific nuclease NucA was monitored by altered degradation of a fluorescent substrate (molecular beacon). The oligonucleotide labeled with a fluorophore (5': 6 FAM) and a quencher (3': BHQ-1) forms a hairpin-like structure mediated by complement bases at both ends with a central loop. The fluorescence signal increases due to cleavage of the molecular beacon as close proximity of fluorophore and quencher is necessary for the quenching effect.

The fluorescence signal for several NucA/NuiA ratios (see Table 27) for three independent measurements was recorded (NucA concentration of 3 nM). However, previous unpublished results indicate that only 2.5 % of the molecules in that nuclease preparation are active since NucA tends to form inactive aggregates. As the intrinsic active fraction of NucA was calculated to be 75 pM, NuiA was added in concentrations corresponding to this amount. The preincubation for stable complex formation was performed on ice for 20 min (50 µl reaction in 1× reaction buffer).

Table 27: NucA/NuiA ratios and corresponding concentrations.

Ratio NucA/NuiA	**10× concentration**	**Final concentration**	**Intrinsic concentration**
0:1	30 nM:0	3 nM:0	75 pM:0
0:0.1	30 nM:75 pM	3 nM:7.5 pM	75 pM:7.5 pM
0:1	30 nM:750 pM	3 nM:75 pM	75 pM:75 pM
0:2	30 nM:1.5 nM	3 nM:150 pM	75 pM:150 pM
0:5	30 nM:3.75 nM	3 nM:375 pM	75 pM:375 pM
0:8	30 nM:6 nM	3 nM:600 pM	75 pM:600 pM
0:10	30 nM:7.5 nM	3 nM:750 pM	75 pM:750 pM

10× reaction buffer	Cleavage reaction (cuvette)
500 mM Tris/HCl pH 7	1.25 µl FS2B (20 µM)
50 mM $MnCl_2$	10 µl 10× reaction buffer
500 mM NaCl	78.75 µl H_2O
0.01 % Triton X-100	10 µl NucA/NuiA 10× preincubated complex

The samples were measured with a Fluorescence Spectrophotometer F-4500 (HITACHI) in an Ultra-Micro Fluorescence Cell (Hellma, light path: 3 × 3 mm, 45 µl). The cuvette was prepared with substrate and buffer/water. The reaction was started by adding the NucA/NuiA 10× preincubated complex after a stable fluorescence signal was detected (excitation wave length: 494 nm, detection wave length: 520 nm) and measured for 3 min.

3 Results

3.1 Purification and chemical rescue of EndA H160A

EndA on the surface of *Streptococcus pneumoniae* was earlier identified as entry nuclease involved in the specific degradation of incoming DNA during transformation[5]. Sequence alignments with other well characterized non-specific nucleases revealed several conserved amino acid residues presumably located within the active site[68]. For the purpose of a biochemical characterization and verification of the active site it was necessary to clone wt endA in *E.coli*. The recombinant expression of EndA should result in sufficient protein for a further analysis. However, this turned out to be highly problematic since the obvious nucleolytic activity of EndA prevented successful cloning of the wild type gene. Nevertheless, the nuclease and variants containing single amino acid replacements concerning the active site could be produced in a cell-free expression system. Detailed tests for nucleolytic activity of the synthesized proteins revealed a complete loss of function for variant EndA H160A. This variant appeared to be the preferred candidate for successful cloning and a high scale expression in *E.coli*.

3.1.1 Purification of recombinant EndA H160A

Variant EndA H160A could be successfully cloned and expressed in *E. coli* probably due to a complete loss of nucleolytic activity (see 3.2) caused by this single mutation. EndA H160A was expressed as a His-GST double tagged version in *E. coli* with a molecular weight of 55.6 kDa and the fusion protein purified by Ni-affinity chromatography (see Figure 16). The double tag was clipped off by TEV Protease at a specific recognition site between GST and EndA H160A (see Figure 17 a) introduced by pETM-30.

His-tagged GST and TEV Protease were separated from EndA H160A using Ni-NTA Agarose. The untagged nuclease could be recovered from the supernatant as verified by a loss of signal in a Western blot analysis employing an α-GST antibody (see Figure 17 b). The

theoretical molecular weight of His-tagged GST was calculated with 29 kDa but the protein band runs significantly lower in an SDS-PAGE below that of EndA H160A (27 kDa).

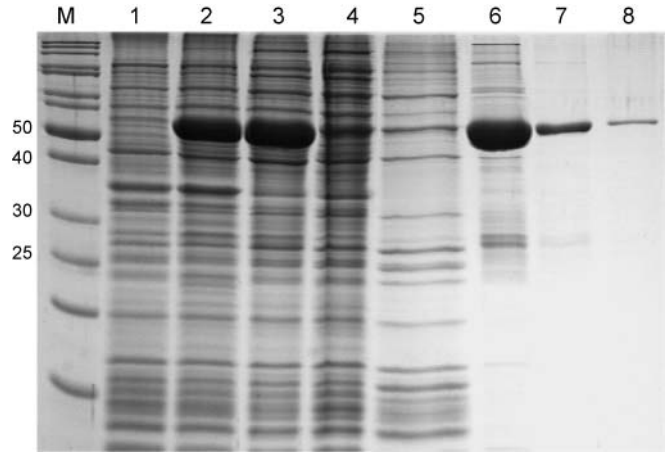

Figure 16: Purification of His-tagged GST-EndA H160A by Ni-affinity chromatography. (1) *E. coli* cells before induction, (2) after induction, (3) supernatant after cell disruption and centrifugation, (4) column flow through, (5) column wash, (6-8) elution 1-3. Molecular weight His-GST tagged EndA H160A: 55.6 kDa.15 % SDS-PAGE.

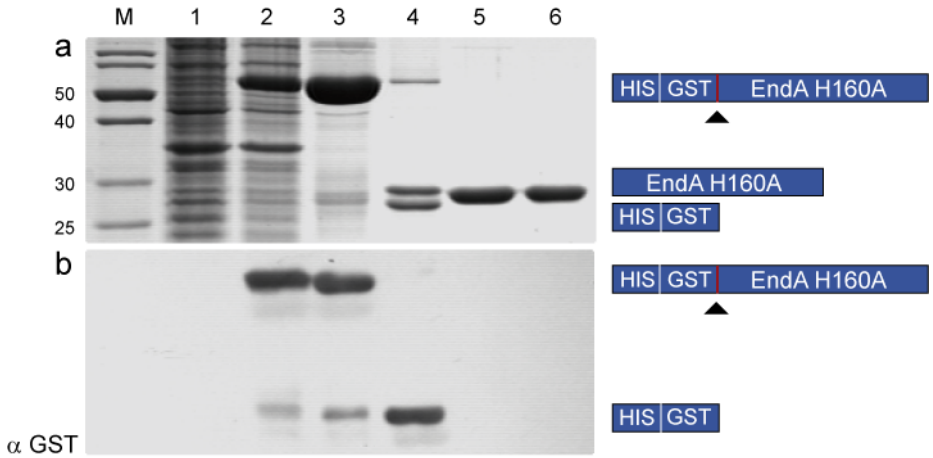

Figure 17: TEV Protease mediated cleavage of His-tagged GST-EndA H160A. (a) (1) *E. coli* cells before induction, (2) after induction, (3) column elution 1, (4) after incubation with TEV Protease: full length protein was cleaved resulting in His-GST (lower band) and EndA H160A (upper band), (5) supernatant after incubation with Ni-NTA Agarose: His-tagged proteins bind to the matrix, untagged EndA H160A remains in the supernatant, (6) after size exclusion. 15 % SDS-PAGE (b) Western blot analysis (α-GST). Loaded samples correspond to lanes in panel a. Loss of signal indicates the successful separation of EndA H160A from GST.

In a final purification step residual contamination was removed by size exclusion chromatography (see Figure 18) represented by several small peaks (II) whereas EndA H160A eluted as a single major peak at 12.5 ml (I). The retention coefficient was calculated

to be 0.69. Protein standards in a size exclusion chromatography were used to calculate the apparent molecular weight of EndA H160A with 30.6 kDa based on a curve fitted to the determined retention coefficients of the standards. The calculated apparent molecular weight of EndA H160A is in good agreement with the theoretical molecular weight of 26.8 kDa indicating that EndA is a monomer under these conditions.

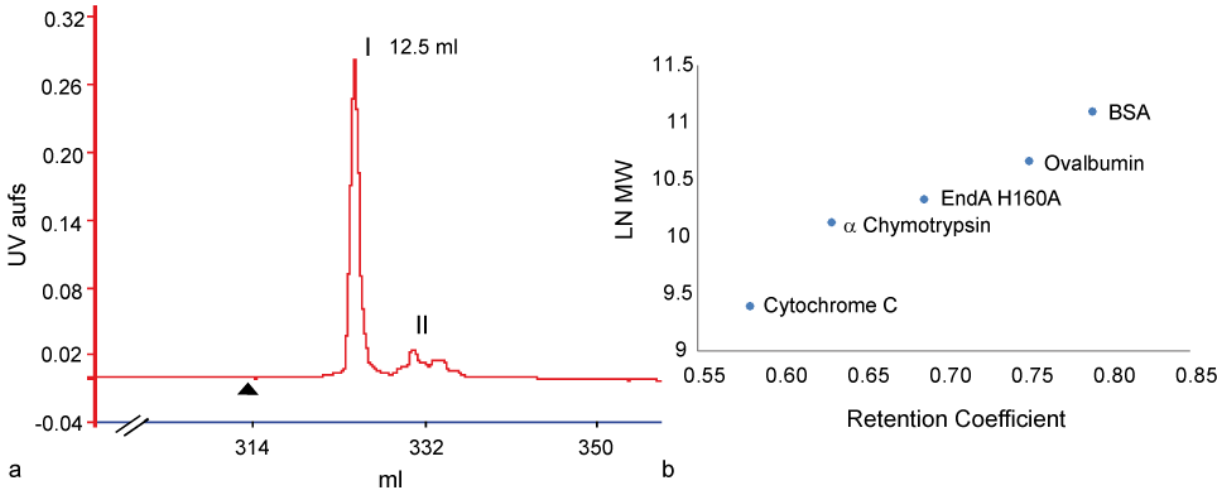

Figure 18: Purification of EndA H160A by size exclusion chromatography. (a) Purification of EndA H160A by size exclusion chromatography: After clipping off the His-GST double tag and incubation with Ni-NTA Agarose the collected flow through containing only untagged EndA H160A was applied to a Superdex™75 column and eluted at a flow rate of 1 ml/min. The black triangle indicates the time point of sample loading. Elution of proteins was monitored at 280 nm. Peak I represents EndA separated from a residual contaminant (II). (b) The retention coefficients (Superdex™75 column) of EndA and protein standards were determined and plotted against the natural logarithm of the molecular mass. The molecular weight of 30.6 kDa obtained for EndA, based on the fitted curve, corresponds reasonably well with the calculated molecular weight of 26.8 kDa indicating that EndA is a monomer under these conditions.

3.1.2 Chemical rescue of EndA H160A

Since no nucleolytic activity for EndA H160A was detectable in any activity assay that was applied (see 3.2.1) catalysis was tried to be restored by supplementing the reaction buffer with excess imidazole. This molecule in principal should complement the deleted side chain that had been removed by substitution of the histidine residue by alanine (Figure 2)[92; 93]. The ratio of open circular to supercoiled DNA in a plasmid-activity assay was calculated based on the intensity of the DNA-bands obtained at selected time points at different pH values. Indeed, EndA H160A showed nucleolytic activity at pH values ranging from pH 7.5 to 8.5 at a concentration of 200 mM imidazole as detected by an increase with time of open circular plasmid DNA (see Figure 19). As a control, the plasmid DNA was incubated with 200 mM

imidazole without enzyme for 20 min at the same pH values as used for the rescue experiments without any effect on the plasmid DNA.

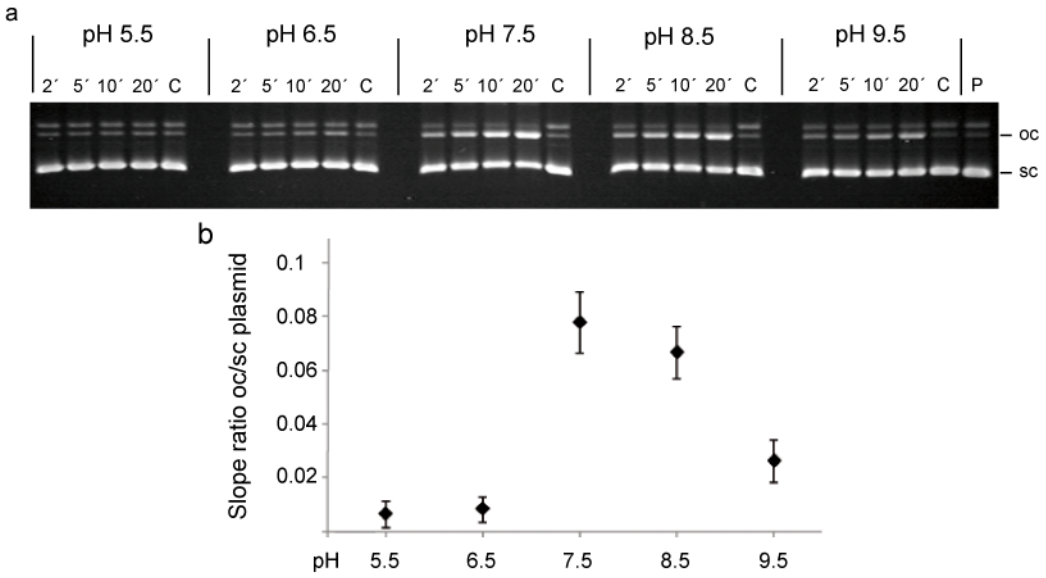

Figure 19: Chemical rescue of the nuclease activity of catalyticly inactive EndA H160A. (a) 200 mM imidazole was incubated with 20 nM EndA H160A and plasmid DNA at pH 5.5, 6.5, 7.5, 8.5 and 9.5. Plasmid DNA (P) incubated for 20 minutes in the absence of EndA H160A and plasmid DNA (C) were loaded as controls. 0.8 % agarose gel. (b) The experiment was repeated and the ratio of open circular (oc) to supercoiled (sc) plasmid DNA plotted against time (n=3). EndA H160A showed nucleolytic activity at slightly basic pH values ranging from pH 7.5 to 8.5. Time point 0 corresponds to the control (P).

3.2 Activity assays for EndA wt and variants

All efforts to clone and express *endA* wt in *E. coli* were not successful (data not shown) probably due to the intrinsic toxicity of the nuclease for the host already during cloning procedures. A cell-free expression system for the synthesis of EndA wt and several variants was used instead. Guided by a sequence alignment of DRGH-motif containing proteins (see Figure 20) certain amino acid residues were selected for the mutational analysis of the presumptive active site of EndA. Asp157, Arg158, and His160 from the conserved DRGH-motif, all localized in one of the two β strands of the ββα-metal finger structure, were exchanged for alanine[31; 32; 57]. His160 within the DRG**H**-motif corresponds to the first histidine residue in the **H-**N-H-motif and plays an essential role in catalysis acting as the general base[16; 23]. Furthermore, two asparagine residues were exchanged (Asn179 and

Asn182) for alanine. As candidates for the central asparagine residue in the H-**N**-H-motif these residues might be involved in hydrogen bond network formation and stabilization of the ββα-metal finger structure[29; 30]. According to the sequence alignment another asparagine residue (Asn191) might assume the function of the metal ion ligand corresponding to the histidine residue at the last position of the H-N-**H** motif. In fact, DRGH-motif nucleases generally contain an H-N-N motif[31; 32; 57].

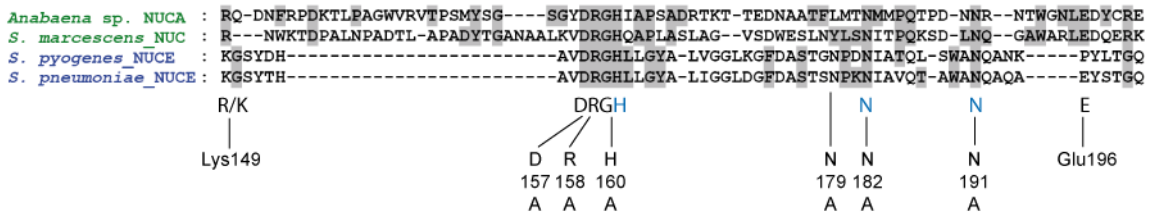

Figure 20: Sequence alignment for DRGH-motif containing nucleases (detail): Indicated are amino acid residues selected for a substitution by alanine. Mutations at positions 157, 158 and 160 concern the conserved DRGH-motif. Histidine at position 160 also represents the first amino acid within the **H**-N-H/N-motif and plays a central role in catalysis. One of two asparagines at positions 179 and 182 could represent the first asparagine residue of the H-**N**-H/**N**-motif, with asparagine 191 being the second. Lysine 149 and Glutamine 169 were selected for remodeling the active centre of wild type EndA based on the alignment and the proposed mechanism for the *Serratia* nuclease (see Figure 30)[23].

3.2.1 In-gel activity assay

Even though synthesized protein for the supplied control reaction of the cell-free transcription/translation system was detectable by Coomassie staining wild type EndA and EndA variants could not be detected by Coomassie staining after separation on SDS-PAGE gels. However, a strong nuclease activity was observed for all variants, except for EndA H160A and the control template supplied with the *in vitro* expression cocktail, using an in-gel activity assay after separation of synthesized proteins on SDS-PAGE gels containing high molecular weight herring sperm DNA (see Figure 21 a). All EndA variants could be detected by Western blot analysis employing an α-FLAG antibody (see Figure 21 b, c). Intriguingly, the amounts of protein produced in the cell-free expression system apparently were inversely correlated to the activity of a given variant (for activity see Figure 22). The variants were produced in different amounts although the template concentration used for the protein synthesis reaction was equal in all reactions. It is very likely that the synthesized nucleases in the coupled transcription/translation system are able to degrade the template DNA according

to their specific activity. Thus, the amount of protein produced for a given variant increases in a feed-back loop with the loss of activity of a given nuclease variant caused by the amino acid exchange. This is most obvious when comparing wild type EndA with the inactive variant H160A, which was produced to the highest protein concentration measured via Western blot analysis. Interestingly, the variant EndA N182A is an exception to this rule. Though it only shows a moderate cleavage activity, this variant apparently was instable producing only a weak signal in the Western blot when analyzed immediately after the synthesis reaction and a decreased signal when analyzed after a storage time of two weeks (data not shown).

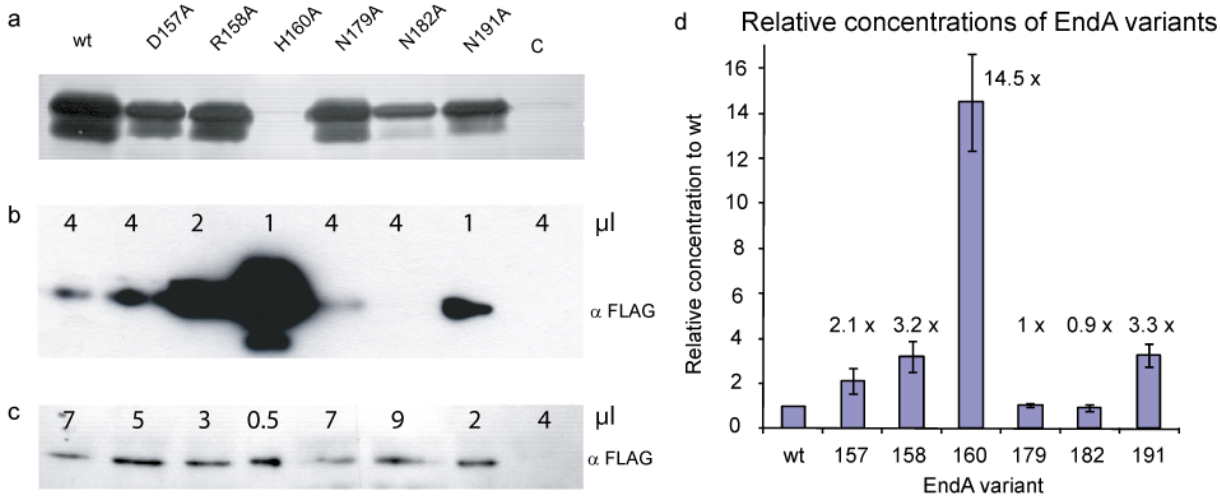

Figure 21: In-gel activity assay and Western blot analysis for EndA wt and variants. (a) Nuclease activity of unpurified protein synthesis reaction of EndA wt, D157A, R158A, H160A, N179A, N182A and N191A separated on a 15 % SDS gel containing high molecular herring sperm DNA. Negative control (C): synthesized protein of supplied control reaction from the *In vitro* Protein Synthesis Kit. (b) (c) Western blot analysis of FLAG-tagged EndA wt and variants (α-FLAG). Amount loaded of the crude synthesis reaction as indicated. Different exposure times for Western blot (b) and (c). (d) Relative concentrations for EndA variants to EndA wt calculated based on the Western blot analysis.

3.2.2 Single radiation enzyme diffusion (SRED) assay of EndA wt and variants

The nucleolytic activity of the crude protein synthesis reaction mixture was quantified by measuring the size of the halo in a single radiation enzyme diffusion (SRED) assay (see Figure 22 a). For each enzyme variant 1 µl of the synthesis reaction mixture was spotted onto an agar plate containing high molecular weight DNA. Whereas all EndA variants showed nuclease activity, the control template supplied for the *in vitro* expression cocktail showed

none. 1 U of bovine pancreatic DNase I was used as a positive control. The relative concentrations of each variant were determined by Western blotting (see Figure 21). To calibrate the system different amounts of the wild type EndA synthesis reaction mixture were used in a separate SRED assay and the size of the halos measured to obtain a calibration curve (see Figure 22 b). Since the concentrations of the variants relative to that of wild type EndA was known from the Western blot analysis, the relative activities of the variants could be determined (see Figure 22 c).

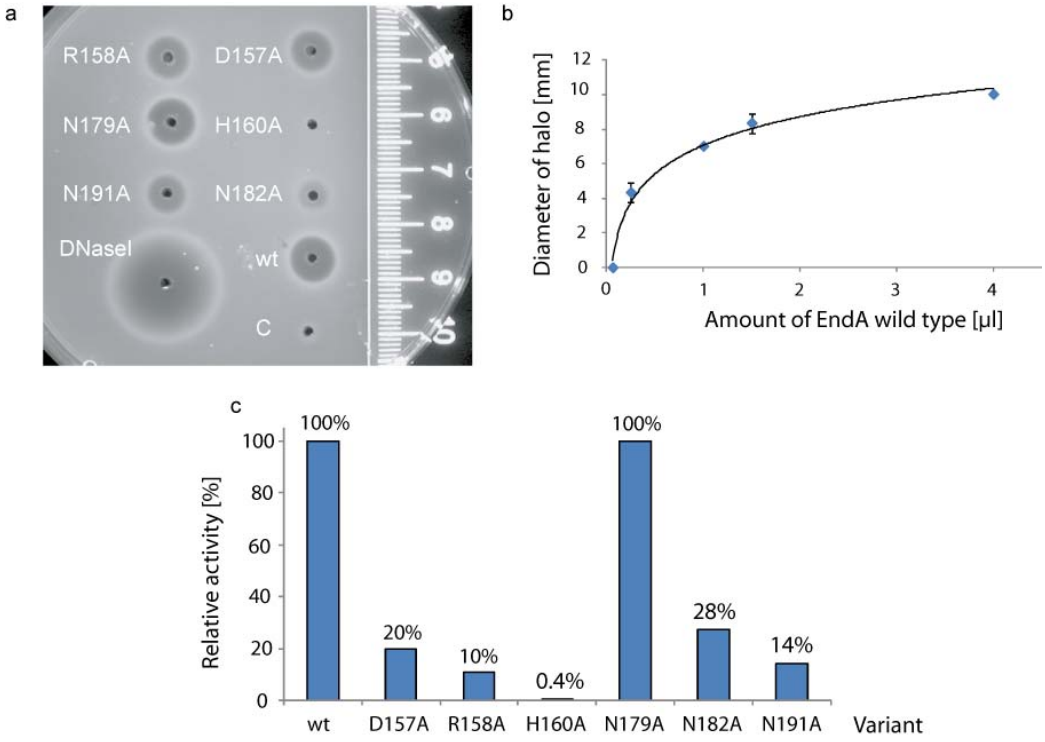

Figure 22: Single radiation enzyme diffusion (SRED) assay: (a) Relative nuclease activity of the *in vitro* protein synthesis reaction mixture containing wild type EndA, D157A, R158A, H160A, N179A, N182A, N191A tested with high molecular herring sperm DNA embedded in agarose (prestained with ethidium bromide). The reaction mixture with the control template supplied with the *In vitro* Protein Synthesis Kit was used as a negative control (C), DNaseI as positive control. The ruler indicates the diameter of the halos. (b) Calibration curve for the estimation of relative nucleolytic activity from the size of the halos. The size of the halos (diameter in mm) was plotted against the amount (µl) of the protein synthesis reaction mixture containing wild type EndA. (c) The relative nucleolytic activity of each EndA variant contained in the crude reaction mixture was determined from the diameter of the halos using the calibration curve in (b) and the concentration of the EndA variants in the crude reaction mixture determined by Western blot analysis. The activity of wild type EndA is set to 100 %.

3.2.3 Nucleolytic activity of EndA wt on circular DNA substrates

In a second type of assay wild type EndA was tested on single-stranded, double-stranded supercoiled and nicked phage DNAs, respectively, with a dilution of 1:100 of the crude synthesis reaction mixture (see Figure 23). No pronounced preference of wild type EndA could be detected for single or double-stranded DNA as all substrates were cleaved to the same extent. The reaction mixtures obtained with the control template supplied for the *in vitro* expression cocktail and the EndA H160A template did not show any nucleolytic activity in this assay (data not shown).

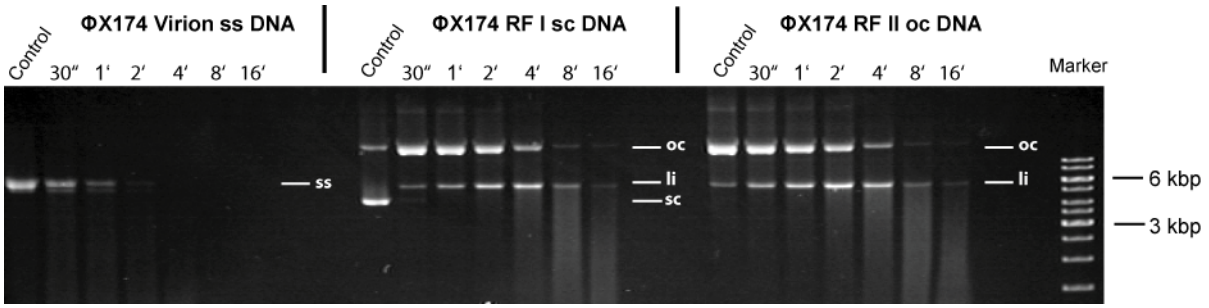

Figure 23: Substrate preference of EndA wt. Nucleolytic activity of EndA wt (*in vitro* synthesis reaction) was tested on ΦX174 virion DNA (single-stranded, circular), RF I DNA (double-stranded, circular) and RF II DNA (double-stranded, nicked). Complete degradation occurs. ss – single-stranded, sc – supercoiled, oc – open circle (nicked), li – linear. 0.8 % agarose gel.

3.2.4 Nucleolytic activity of EndA wt on RNA

For further investigation of the nucleolytic activity of EndA wt, RNA degradation was monitored using an RNA transcript as substrate. Incubation of unpurified synthesis reaction for EndA wt with the substrate RNA led to degradation of the transcript. However, no degradation occurs during incubation with the supplied positive control reaction (see Figure 24). As differently sized DNA and RNA molecules are intermediates of the synthesis reaction, hence visualized on a gel, the unpurified synthesis reaction of EndA wt and positive control supplied with the *In Vitro* Protein Synthesis Kit was loaded as reference.

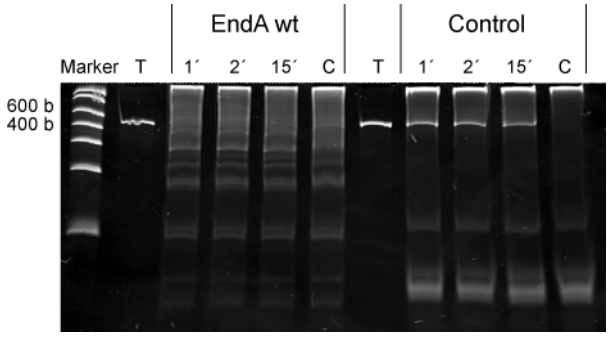

Figure 24: Nucleolytic activity of EndA wt on RNA. 401 b RNA transcript (T) was incubated either with EndA wt protein synthesis reaction (left) or with the supplied positive control (right). Since the synthesis reaction was not purified differently sized DNA and RNA as intermediates of the reaction were visualized additionally to the added transcript. Unpurified synthesis reaction of EndA wt and supplied positive control was loaded as a reference (C). Complete degradation occurs. 5 % urea PAGE.

3.3 Bicistronic selection system

In the course of this work a bicistronic plasmid-based selection system was established for screening and selection of functional inhibitors for non-specific nucleases in particular for *Sma*Nuc. This assay should provide a basis for the establishment of new screening methods to engineer inhibitor variants for virulent surface bound EndA as a promising starting point for a future drug design.

Since cloning and expression of non-specific nucleases is known to be problematic due to intrinsic toxicity, the coexpression of a specific functional inhibitor is utilized to prevent activity of the nuclease during expression, e.g. for colicin E7/E9 and their immunity proteins[46; 94]. Expression of both proteins can be either mediated by two individual promoter sequences controlled independently on one/two plasmids or by one promoter sequence and two coding sequences/ribosomal binding sites as bicistronic arrangement. As an example, the bicistronic expression of NuiA/NucA in *E. coli* is unproblematic as the two proteins occur as a complex in insoluble inclusion bodies with excess of NuiA over NucA[95].

In the developed assay functional inhibitors were selected upon expression of nuclease and inhibitor variants from a bicistron in *E. coli.* The toxic effect of the non-specific nuclease could be down regulated which is visualized by viable hosts/appearing colonies after transformation. The *nuiA/nucA* bicistron served as control plasmid expressing a functional inhibitor for the toxic nuclease whereas the *nuiA*/*smaNuc* construct should cause toxicity during cloning procedures and expression in *E. coli*. The second construct was used to establish a *nuiA* variant library by error-prone and randomized primer PCR with *smaNuc* at second position of the bicistron to screen for surviving colonies which express functional inhibitors for *Sma*Nuc after transformation of *E. coli*. Since the expressed protein at the first position of the bicistron is expressed in excess, in principal even inhibitor variants with weaker binding affinity should be selected.

3.3.1 Verification of basal expression

Expression of the genes arranged as bicistron was controlled by a phage-derived T5 promoter sequence which gets recognized by the *E. coli* RNA polymerase[96]. This promoter as part of the pQE-30 plasmid is inducible by adding IPTG to the medium but also shows basal (leaky) expression in the uninduced state visualized by expressing the fluorescent protein DsRedMonomer (see Figure 25) from control pasmids. Freshly transformed *E. coli* (XL1-Blue MRF') colonies harboring the indicated plasmids were grown as 3 ml cultures without addition of IPTG and spun down. DsRedMonomer was expressed from the second position of the biscistron either alone (pBCSV) or as N-terminal fusion to NucA (pBCNA) and *Sma*Nuc (pBCNS) without specific induction of the protein expression. Therefore, all further selection experiments were done in the absence of IPTG since basal expression of the bicistron results in a sufficiently high concentration of the highly active nucleases.

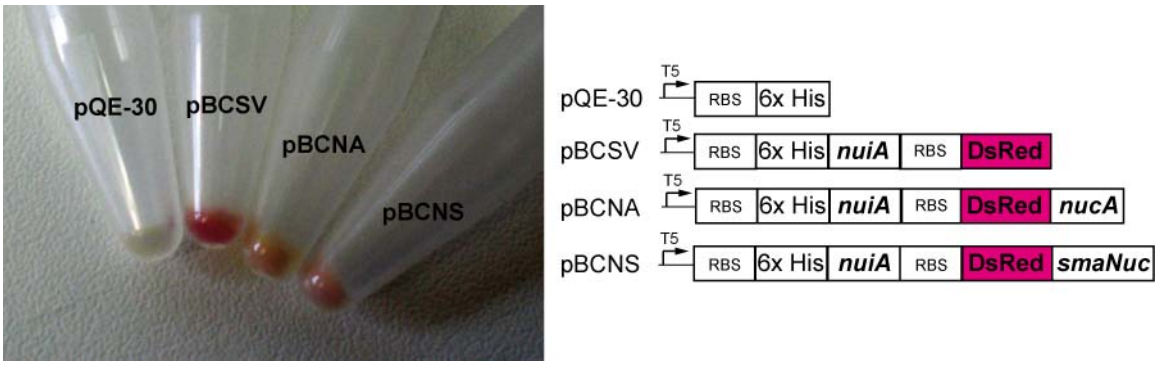

Figure 25: Verification of basal/leaky expression in *E. coli* by DsRedMonomer. Freshly transformed *E. coli* (XL1-Blue MRF') colonies harboring indicated plasmids (right panel) were grown as 3 ml cultures without induction of protein expression by IPTG and spun down (left panel). Pink color represents DsRedMonomer (DsRed) either alone or as an N-terminal fusion to NucA/*Sma*Nuc indicating an expression from the second position of the bicistron without induction (basal/leaky expression). Colonies containing plasmid pQE-30 were used as reference. T5: T5 promoter, 6×His: 6× His-tag, RBS: ribosomal binding site.

For the bicistronic selection system DsRedMonomer was replaced by a FLAG-tag to detect the nuclease expression by a Western blot analysis (upper panel, Figure 26). The fusion of DsRedMonomer to the nucleases presumably decreases their activity due to the molecular weight of this protein (25 kDa).

Freshyl transformed *E. coli* colonies containing the indicated bicistronic plasmids were grown as 3 ml cultures and pelletized aliquots (same amount loaded according to measured OD_{600}) after cell lysis separated by SDS-PAGE. Expression of NuiA (16.5 kDa) from the first position of the bicistron could be detected for all constructs but results in a lower concentration of the protein in cells harboring plasmid pQE-NFNA (lower panel, Figure 26). Cells containing pQE-30 were loaded as negative control. Expression of NuiA from plasmid pQE-30 represents the positive control. NucA and *Sma*Nuc could not be detected in a Coomassie stained SDS gel but the proteins were visualized by Western blot analysis employing an α-FLAG antibody. The theoretical molecular weight of both nucleases is around 28 kDa. NucA (black arrow, upper panel, Figure 26) occurring in a lower concentration within the cell also runs significantly lower in an SDS gel compared to *Sma*Nuc. Two other control plasmids were established expressing *Sma*Nuc variants K55A (residual *in vitro* activity 30 %) and H89N (almost no activity *in vitro* detectable)[23; 97].

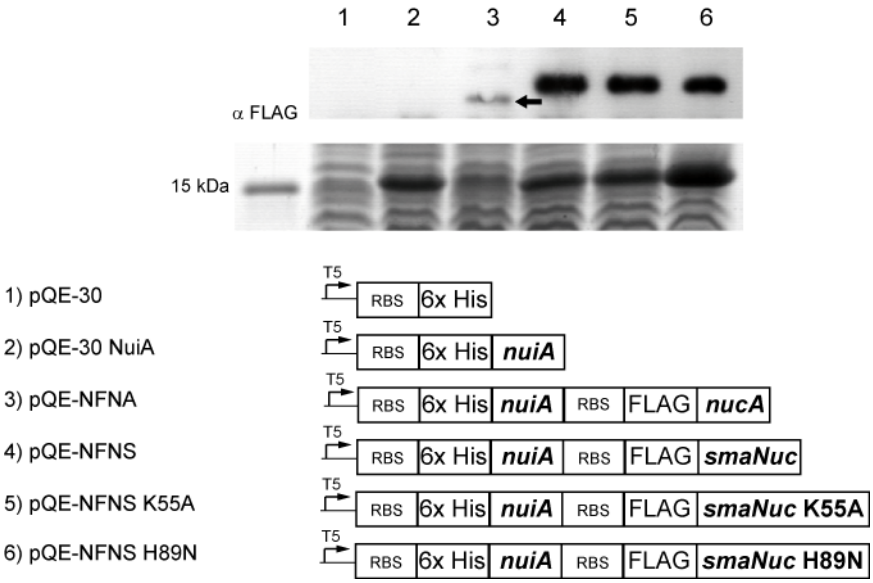

Figure 26: Expression of *nuiA*- and nucleases-genes in the bicistronic plasmid selection system: Freshly transformed *E. coli* (XL1-Blue MRF') colonies harboring indicated plasmids (right panel) were grown as 3 ml cultures without induction of protein expression by IPTG, spun down and separated after cell lysis on a 15 % SDS gel (lower left panel). The same amount of lyzed cells was loaded according to the measured OD_{600}. Expression of 16.5 kDa NuiA could be detected. However, the amount of expressed nucleases was too low for visualization by Coomasssie staining but could be detected in a Western blot analysis (α-FLAG, upper panel).

3.3.2 Establishment of a bicistronic selection system

The highly active nuclease NucA is regulated by the tight binding inhibitor NuiA within the cell whereas no inhibitor is known for *Sma*Nuc[69]. However, *Sma*Nuc exists in an inactivated form due to reducing conditions within the cell, which prevents the formation of disulfide bonds. *Sma*Nuc gains nucleolytic activity in an oxidizing environment upon secretion[72]. On the one hand, this fact renders *Sma*Nuc an easy candidate for cloning and expression as coexpression of an inhibitor is not required. On the other hand, *Sma*Nuc needs to be activated for inhibitor screening as the selection system is based on preventing toxicity caused by nuclease activity. Activation could be achieved by transformation of a particular *E. coli* strain (Origami) with *Sma*Nuc expressing plasmids (see Figure 27). This strain enhances the formation of disulfide bonds in the cytoplasm due to mutations in the thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes[98].

Either XL1-Blue MRF' or Origami competent cells were transformed with 10 ng of the indicated plasmids, plated after appropriate dilution and the colonies (cfu) counted for three independent experiments. The mean value of the number of obtained colonies was calculated (see Figure 27). Plasmid pQE-30, pQE-30 NuiA and pQE-NFNA served as controls. The number of colonies containing plasmid pQE-30 indicates the transformation competence for both *E. coli* strains (XL1-Blue MRF' and Origami) as no protein was expressed. Expression of NuiA from plasmid pQE-30 NuiA and coexpression of NuiA/NucA from plasmid pQE-NFNA were unproblematic and led to equal amounts of colonies after transformation for XL1-Blue MRF' and Origami compared to the colony numbers obtained after transformation with plasmid pQE-30. Although *Sma*Nuc wt should be inactive during expression in XL1-Blue MRF' cells due to the reducing conditions the decreased number of colonies reflects toxicity of the nuclease even under reducing conditions. The effect could be increased by transforming Origami with plasmid pQE-NFNS expressing *Sma*Nuc wt due to enhanced intracellular disulfide bond formation. However, expression in XL1-Blue MRF' of *Sma*Nuc K55A with a residual activity of 30 % did not exhibit toxicity but led to almost no growth after transformation of Origami with pQE-NFNS K55A. For *Sma*Nuc variant H89N almost no residual activity could be detected *in vitro* but a decreased number of colonies indicates

toxicity of this variant *in vivo* during basal expression in Origami cells. Transformation of XL1-Blue MRF' with the plasmid pQE-NFNS H89N did not exhibit reduced colony numbers.

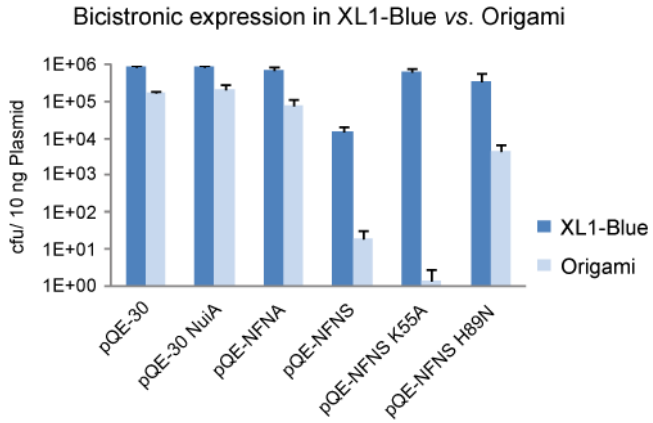

Figure 27: Viability (cfu) of *E. coli* harboring indicated plasmids. *E. coli* strains XL1-Blue MRF' and Origami were transformed with indicated plasmids and plated with appropriate dilutions. The number of colonies for three different experiments was counted and the mean calculated. Transformation of Origami cells with *Sma*Nuc expressing plasmids displayed a decreased number of colonies due to enhanced activity/toxicity of the nuclease compared to expression in XL1-Blue MRF'. *Sma*Nuc needs to be activated by the formation of disulfide bonds which is enhanced due to mutations in the thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes in Origami cells. Although *Sma*Nuc H89N did not exhibit activity *in vitro*, toxicity could be observed *in vivo* in Origami cells. cfu: colony forming units.

3.3.3 Selection of functional inhibitor variants

The basal expression of *Sma*Nuc from plasmid pQE-NFNS and pQE-NFNS K55A in Origami revealed a strong toxic phenotype reflected by a low number of colonies after transformation. This distinct effect seemed to be useful for the selection of functional inhibitor proteins for *Sma*Nuc, as coexpression of an inhibitor reduces toxicity. For the selection process a genetic library of inhibitor variants based on the sequence of *nuiA* was created by error-prone PCR and PCR using randomized primer.

Several residues of NuiA are known to have specific interactions with NucA. Based on the published crystal structure these residues were selected (i) for site directed mutagenesis targeted by randomized primers or (ii) for deletion (AA 78-80) in order to change the binding interface of NuiA to create suitable inhibitors for *Sma*Nuc (see Figure 28)[31]. OEP for each primer pair was performed to randomize either amino acid residues 68, 74-76 or 109-111. The obtained fragments for each reaction and error-prone PCR fragments of *nuiA* were used to replace *nuiA* wt in plasmid pQE-NFNS and pQE-NFNS K55A. After transformation of

Origami cells with the products of each ligation reaction viable colonies were picked and the containing plasmids sequenced to find mutations introduced into the *nuiA* gene. Unfortunately, no viable colonies appeared after transformation of Origami cells with the libraries created by site directed mutagenesis neither established in plasmid pQE-NFNS nor pQE-NFNS K55A. Yet, viable colonies could be picked after transformation of Origami cells with the library obtained by error-prone PCR and isolated plasmids were send for sequencing. Unfortunately, the sequencing revealed mutations within the promoter sequence of the plasmids and nearly no mutations within the coding region of the *nuiA* gene. As a result, no functional inhibitor for *Sma*Nuc based on NuiA could be isolated.

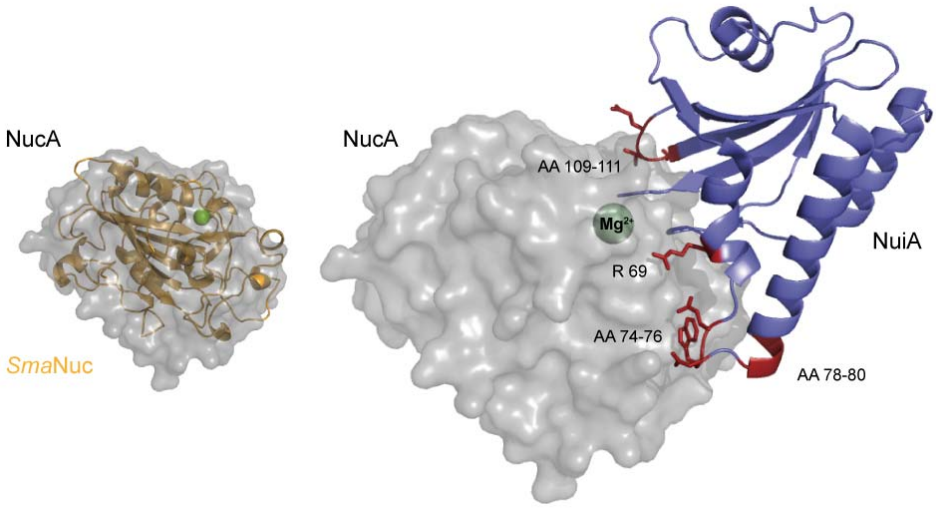

Figure 28: Crystal structure of the NucA/NuiA complex. Left: NucA displayed as gray surface structurally aligned to *Sma*Nuc (orange). Mg^{2+} is indicated as a green sphere. Right: NucA displayed as gray surface, Mg^{2+} indicated in green as part of the active site and NuiA visualized as blue cartoon with amino acids (red) selected for mutagenesis by randomized primers. Amino acids 78-80 were chosen for deletion. The selected amino acids form specific contacts to NucA but do not interact with the active site. (PDB code NucA/NuiA: 203B, *Sma*Nuc:1SMN)

3.4 Inhibition NucA by NuiA

The inhibition effect of the tight binding inhibitor NuiA for the nuclease NucA was monitored by degradation of a fluorescent substrate (multiple turn over reaction) determining the slope (reaction velocity) of the increasing fluorescent signal. As the substrate forms a molecular beacon the fluorophore at the 5' end is in close proximity to the quencher molecule at 3' end of the DNA hence absorbing the emitted signal. The fluorophore gets released after degradation of the hairpin shaped substrate by NucA which leads to an increasing fluorescence signal. The obtained linear increase could be fitted to a linear regression curve and the slope which

corresponds to the relative velocity of the reaction was determined. In order to determine an inhibition constant for the reaction, the relative velocity for each NucA/NuiA ratio measured was plotted against the NuiA concentration. Based on Equation 1 a curve was fitted (non-linear least-squares fit) to the obtained data (see Figure 29) keeping the enzyme concentration [E], inhibitor concentration [I] and the starting velocity v_0 as constant parameters to calculate K_i^{app} [99]. The apparent inhibition constant (mean value) derived from three independent experiments was calculated to be 0.6 ± 0.15 pM.

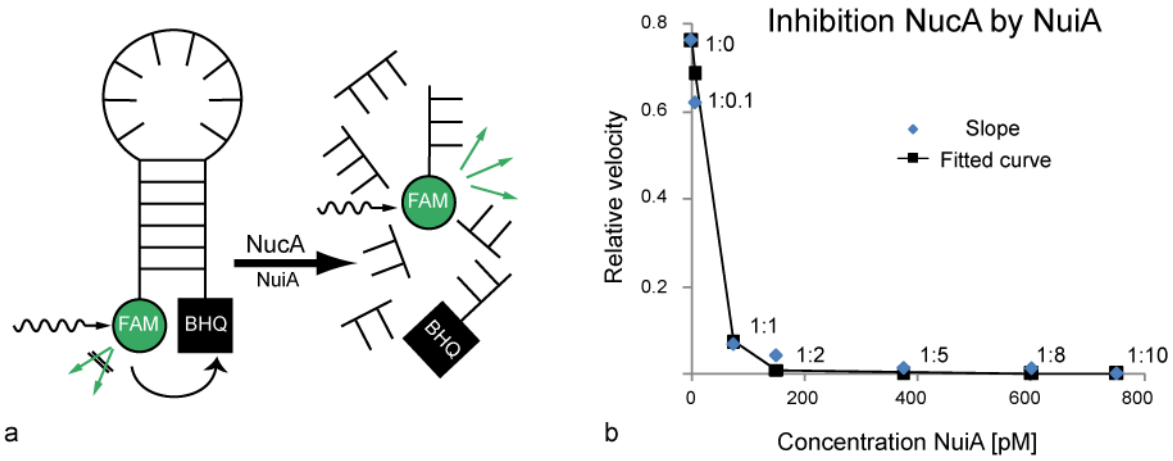

Figure 29: Fluorescence activity assay for NucA at different inhibitor (NuiA) concentrations. (a) Illustration of the fluorescence assay. Close proximity of the fluorophore (6 FAM) and quencher (BHQ – Black Hole Quencher) leads to absorption of the transferred energy. After degradation of the DNA substrate by NucA the fluorophore gets released leading to an increasing fluorescence signal. (b) The relative reaction velocity for degradation of the fluorescent substrate (molecular beacon) by NucA (intrinsic concentration 75 pM) at different NuiA concentrations (1:0, 1:0.1, 1:1, 1:2, 1:5, 1:8, 1:10) was determined. A curve (black) for the measured slopes (blue) was fitted by Equation 1 to calculate the apparent inhibition constant K_i^{app}. The apparent inhibition constant (mean value) derived from three independent experiments was calculated with 0.6 ± 0.15 pM.

$$v = v_0 \frac{[E]-[I]-K+\sqrt{([E]-[I]-K)^2+4[E]K}}{2[E]}$$

Equation 1: Reaction velocity for reversible tight binding inhibition. v is the initial reaction velocity observed at inhibitor concentration [I], v_0 is the initial reaction velocity observed in the absence of inhibitor, [E] is the active enzyme concentration and K the apparent inhibition constant K_i^{app} [99; 100].

4 Discussion

4.1 EndA

Non-specific nucleases play an important role in nucleic acid metabolism in all kingdoms of life[1; 7]. Due to the intrinsic toxicity of these enzymes, cloning of genes coding for nucleases and their recombinant expression in *E. coli* or other host organisms can be highly problematic unless, for example, nuclease inhibitors are available for intracellular coexpression or the enzyme is produced as an inactive pro-enzyme that becomes activated by secretion into the extracellular space[28; 72; 94; 95]. EndA from *S. pneumoniae* is a secreted surface exposed endonuclease for which up to date no inhibitory protein is known[5].

All cloning attempts for EndA wt led to mutations in the region coding for the active site and/or insertion of stop codons in the open reading frame preventing expression of active full length protein. EndA could only be cloned and expressed as an inactive variant in which the putative active site histidine residue His160 was exchanged for alanine (H160A). This variant could be produced in high amounts intracellularly as a His-GST double tagged protein (see Figure 16). Surprisingly, EndA H160A protein fused to pelB turned out to be highly soluble upon expression in *E. coli* in contrast to other non-specific nucleases of the DRGH-motif containing type such as *Anabaena* nuclease or *Serratia* nuclease (see Figure 16)[95; 101].

In order to perform a more detailed mutational and biochemical analysis of EndA an OEP was used to splice DNA templates coding for wild type EndA and desired variants (see Figure 13). The proteins were synthesized using a cell-free expression system. The subsequent analyses of *in vitro* translated EndA and EndA variants confirmed expectations concerning amino acid residues of the active site of this DRGH-motif containing nuclease. According to the data presented here and in agreement with the results obtained for sugar non-specific nucleases such as *Serratia* nuclease, *Anabaena* nuclease or Endonuclease G, His160 from the DRG**H**-motif of EndA is crucial for catalysis, most likely by acting as a general base (see Figure 30)[22; 23; 57]. This assignment is supported by chemical rescue of the activity of the H160A variant with excess imidazole in the reaction buffer (see Figure 19). In the H160A variant imidazole apparently occupies the space that is normally filled by the side chain of the active site histidine residue in the wild type enzyme but absent in this variant. EndA H160A therefore is able to catalyze the cleavage of phosphodiester bonds in the presence of imidazole at a slightly basic pH.

Replacement of amino acid residues from the conserved **DR**GH-motif of EndA at the first and second position by alanine (variants D157A and R158A) also diminishes nuclease activity, albeit to a lower extent (see Figure 22). It is known from the above mentioned studies concerning *Serratia* nuclease, *Anabaena* nuclease and EndoG, that these residues are involved in proper positioning of the conserved asparagine residue, which is binding the divalent metal ion cofactor in the enzyme's active site (see Figure 30)[31; 32].

A so called H-N-H-motif is characteristic for a large number of proteins covering a broad variety of functions[16]. Sequence analyses show a conserved H-N-N-motif in EndA (see Figure 20). As two asparagines (N179, N182) were likely candidates for the second position within the H-**N**-N motif, both of them were selected for an alanine replacement. However, only variant N182A shows a marked drop in activity indicating that N182 very likely is part of the active site (see Figure 22) and probably involved in constraining the loop structure of the ββα-metal finger as shown for ColE7 at the corresponding position, thus contributing to the stability of the protein[29]. This explains the finding that although variant N182A is less active than wild type EndA, the concentration of the synthesized variant N182A is lower compared to the variants with mutations affecting catalytic residues in the active site (see Figure 21 and Figure 22). The amount of produced protein for all other variants is in general accompanied by a decrease in activity due to degradation of the template used in the cell-free protein synthesis reaction.

The conserved aspargine residue, which corresponds to the last histidine residue in the H-N-**H**-motif is known to be the direct ligand responsible for metal ion binding as in the case of *Serratia* nuclease, *Anabaena* nuclease, and mitochondrial EndoG[30; 31; 57]. Replacement of Asp191 by alanine in EndA leads to decrease in activity suggesting that this residue is involved in catalysis, presumably by binding the divalent metal ion cofactor (see Figure 30).

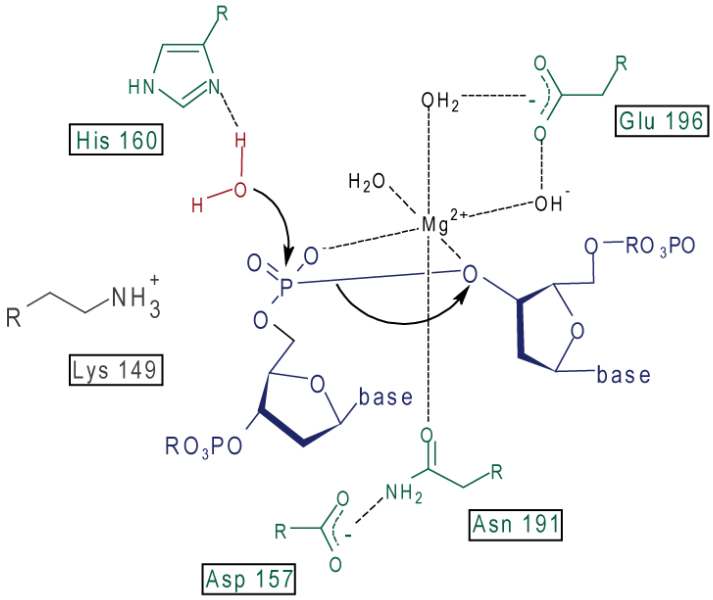

Figure 30: Catalytic mechanism of wild type EndA: The proposed catalytic mechanism of EndA is based on the established catalytic mechanism of the *Serratia* nuclease[23]. His160 acts as the general base for the activation of a water molecule. The magnesium ion is coordinated by Asn191 which is held in place by Asp157. Glu196 is assumed to be indirectly involved in magnesium ion binding. The transition state is stabilized by Lys149.

Streptococci take up single-stranded linear DNA segments during transformation[86; 102]. Using different phage DNA and an RNA transcript as substrate no pronounced preference could be detected for the isolated recombinant enzyme (see Figure 19). EndA exerts a strong nicking activity towards supercoiled DNA. This cleavage mechanism is characteristic for monomeric as well as dimeric non-specific nucleases with independently acting subunits and supported by the fact that EndA behaves as a monomer in size exclusion chromatography (Figure 18). As single-stranded DNA serves as substrate for EndA, the question of regulation of the enzyme activity on the membrane surface of streptococci arises. For the 17 kDa entry nuclease from *Bacillus subtilis* an inhibition of nuclease activity was observed *in vitro* by an 18 kDa protein (Nin)[103]. Several Com proteins on the membrane of *B. subtilis* are involved in binding and transport of DNA into the cell similar as suggested for *S. pneumoniae*[6]. It is tempting to speculate that one of the Com proteins is interacting with EndA, thereby regulating its function and perhaps being responsible for an inhibitory activity within the cell.

.

4.2 Inhibitor selection system

In vivo selection systems in yeast and *E. coli* are interesting tools for screening of large libraries used to engineer and select proteins with specific features e.g. homing endonucleases with new DNA target sites[104; 105]. Benefiting from this basic concept a screening system for the selection of functional inhibitor proteins for DRGH/H-N-H motif containing non-specific nucleases was established. The developed assay utilizes the ability of inhibitor proteins to diminish the toxic/lethal activity of the nuclease in the cell due to a coexpression. The regulating effect is directly visualized by viable hosts/appearing colonies after transformation of *E. coli* with plasmids containing the genes arranged in a bicistron. The expression of nuclease and inhibitor is mediated by an inducible phage-derived promoter. However, in case of the applied system gene expression occurs without a specific induction of the promoter as visualized by the fluorescent protein DsRedMonomer directly fused to the nuclease (NucA, *Sma*Nuc) at the second position of the bicistron (see Figure 25). This basal or leaky expression leads to a lower protein concentration within the cell, which presumably supports the solubility of the nucleases, as they are known to form inclusion bodies upon overexpression in *E. coli*. Aggregation in the cell leads to a reduced nucleolytic activity, which in turn reduces the toxicity and disturbs the screening process. Solubility is also promoted by the fact that the nuclease at the second position of the bicistron has in general a lower expression level in comparison to the inhibitor protein right behind the promoter sequence (see Figure 26)[95].

Nevertheless the biased expression from the bicistronic arrangement is more than sufficient to inhibit the nucleolytic activity of NucA by NuiA as no toxic phenotype could be observed (see Figure 27). Coexpression of *Sma*Nuc wt and NuiA in XL1-Blue MRF' cells should be as unproblematic as for NucA and NuiA since *Sma*Nuc needs the formation of two disulfide bonds to become active. However, a weak toxic phenotype could be observed indicating a low nucleolytic activity without the effective formation of the intramolecular bonds. An impressive increase of activation could be observed after transformation of Origami cells with the last-mentioned plasmid. Although *Sma*Nuc variant K55A only has a residual activity of 30 % nearly no colonies were observed after transformation of Origami cells and no toxic phenotype was observed for XL1-Blue MRF' cells in comparison to *Sma*Nuc wt. This striking

high level of activation for the variant K55A was unexpected since the variant is impaired in substrate binding.

A reason for the obtained unexpected result might be the low toxic phenotype during basal expression of *Sma*Nuc wt in XL1-Blue MRF' which already supports a selection process. The selection probably occurs during necessary high scale plasmid preparations since XL1-Blue MRF' cells grow better when harboring plasmids with mutated promoter sequences. These replication errors are promoted because they lead to a lower or almost no nuclease expression. In turn, Origami cells harboring the preselected plasmids survive without coexpression of a functional inhibitor. The appearing “false-positive” colonies make an efficient screening for inhibitory proteins rather difficult.

Nevertheless, even screening supported by plasmid pQE-NFNS K55A did not result in the selection of functional inhibitor proteins. All selected and sequenced plasmids revealed mutated promoter sequences. This proves the fact that the most effective way to relieve the selection pressure is a mutated promoter sequence blocking expression and not the coexpression of a functional inhibitor. Additionally, the cells are able to diminish the expression of toxic genes by an unknown way as an external trigger for the expression e.g. by IPTG is missing. NuiA and NucA have obviously a lower concentration in the cell than *Sma*Nuc as shown by a Western blot analysis (see Table 26).

Taken together the presented screening method is not suitable for an effective screening and selection of functional inhibitors for *Sma*Nuc. Furthermore, the screening approach is not transferable to EndA since this extremely toxic nuclease up to date could not be cloned and expressed in active form in *E. coli*. A possible screening for functional inhibitors might be achieved by utilizing the yeast two-hybrid system. This screening method is used to detect protein interactions e.g. for a nuclease and its inhibitor. A possible procedure includes cloning and expression of the inactive EndA variant H160A and coexpression of a NuiA library, which might be applicable also for the *Sma*Nuc variant H89N. However, this variant still shows a weak activity displayed by the toxic effect in Origami cells, although the mutation concerns the general base histidine residue, which corresponds to histidine at position 160 of EndA. A weak toxic effect during expression in a yeast two-hybrid screen might support the selection efficiency. The toxic effect results in a double selection since yeast cells without a functional inhibitor will not survive as no binding/inhibition occurs.

The inhibitory effect of selected candidates could be easily monitored by the presented fluorescence assay. The calculated apparent inhibition constant of 0.6 ± 0.15 pM corresponds

well with the published value of 3.2 ± 1.9 pM calculated based on a plasmid activity assay[31]. This method therefore could be an effective tool to monitor the inhibitory effect directly. Plasmid assays only allow an indirect calculation after gel electrophoresis and densitometric analysis of DNA bands after staining, whereas for the fluorescence assay the inhibition constant can directly be calculated based on the obtained progress curves.

5 Summary

EndA is a membrane-attached surface exposed DNA-entry nuclease previously known to be required for genetic transformation of *Streptococcus pneumoniae*. More recent studies have shown that the enzyme also plays an important role during establishment of invasive infections by degrading neutrophil extracellular traps (NETs), enabling streptococci to overcome the innate immune system in mammals. This work presents the first mutational and biochemical analysis of recombinant forms of EndA produced either in a cell-free expression system or in *E. coli*. Histidine at position 160 and Asn191 were identified to be essential for catalysis and Asn182 (corresponding to the central asparagine residue in the well known H-**N**-H-motif found in many nucleases) to be required for the stability of EndA. The role of His160 (**H**-N-H) as the putative general base in the catalytic mechanism is supported by chemical rescue of the H160A variant of EndA with imidazole added in excess.

Although EndA appeared to be highly toxic for the cell, no functional inhibitor for EndA could be isolated up to now. Such a protein inhibitor might be an interesting starting point for a possible drug design. A detailed comparison of EndA with other non-specific nucleases revealed a high sequence homology especially at the active site with NucA from *Anabaena sp*. However, NucA occurs in a complex with its specific inhibitor NuiA. As NuiA blocks the active site of NucA it should be feasible to select functional inhibitors from a NuiA based library for EndA and other homologous nucleases such as the *Serratia* nuclease. Unfortunately, the established plasmid based assay appeared to be too instable and no candidate inhibitor could be isolated. Furthermore, the highly toxic nuclease EndA could not be cloned and expressed in an active form in *E. coli* so far. A yeast two-hybrid assay might be used to screen a NuiA library for a protein-protein interaction with the inactive/non-toxic EndA H160A variant.

6 Zusammenfassung

EndA ist eine unspezifische Nuklease, die auf der Oberfläche des pathogenen Bakteriums *Streptococcus pneumoniae* lokalisiert ist. Dort ist sie an der Fragmentierung doppelsträngiger DNA zu einzelsträngigen Segmenten beteiligt, die während kompetenter Phasen des Bakteriums aufgenommen werden. Des Weiteren konnte nachgewiesen werden, dass EndA effektiv das DNA Gerüst so genannter NETs (*Neutrophil Extracellular Traps*) degradiert. Mit diesen netzartigen Strukturen sind bakterizide Proteine wie Histone und Elastasen assoziiert, die zum Absterben gebundener Bakterien führen können. Mikroben, die durch den spezifischen DNA Abbau somit der Immunabwehr von Säugern entkommen, können schwere invasive Krankheiten auslösen.

In dieser Arbeit wurden gezielt EndA Varianten mit Punktmutationen erzeugt, die das konservierte H-N-H/N Motiv betreffen. Die entsprechenden Varianten konnten aufgrund ihrer Toxizität nur mit Hilfe eines zellfreien Expressionssystems hergestellt werden. Allerdings konnte die inaktive Variante H160A in *E. coli* exprimiert und die nukleolytische Aktivität durch einen Überschuss an Imidazol im Reaktionspuffer wieder hergestellt werden (*chemical rescue*). Der Verlust der katalytischen Aktivität entstand durch den Austausch der generellen Base Histidin zu Alanin innerhalb des **H**-N-H/N Motivs. Des Weiteren konnte gezeigt werden, das Asparagin 182 (H-**N**-H/N) an der Stabilisierung des Proteins beteiligt ist, während Asparagine 192 das benötigte zweiwertige Metallion im aktiven Zentrum bindet (H-N-H/**N**).

Bis heute wurde für die toxische Nuklease EndA kein Inhibitor entdeckt, der ein möglicher Ansatz für eine entsprechende Therapie für durch Streptokokken verursachte Krankheiten sein könnte. Allerdings sollte es möglich sein, basierend auf dem bekannten Inhibitor NuiA für die homologe Nuklease NucA einen funktionellen Inhibitor für EndA und andere homologe Nukleasen (*Serratia* Nuklease) herzustellen. Der in dieser Arbeit entwickelte Selektionsassay erwies sich allerdings als zu instabil, sodass kein Inhibitor isoliert werden konnte. Eine mögliche Alternative könnte ein *yeast two-hybrid* Assay sein, mit dessen Hilfe entsprechende NuiA Varianten aufgrund ihrer Interaktion mit EndA H160A isoliert und im Weiteren auf ihre inhibitorische Funktion überprüft werden könnten.

7 References

1. Rangarajan, E. S. & Shankar, V. (2001). Sugar non-specific endonucleases. *FEMS Microbiol Rev* **25**, 583-613.
2. Nagata, S., Nagase, H., Kawane, K., Mukae, N. & Fukuyama, H. (2003). Degradation of chromosomal DNA during apoptosis. *Cell Death Differ* **10**, 108-16.
3. Wu, S. I., Lo, S. K., Shao, C. P., Tsai, H. W. & Hor, L. I. (2001). Cloning and characterization of a periplasmic nuclease of Vibrio vulnificus and its role in preventing uptake of foreign DNA. *Appl Environ Microbiol* **67**, 82-8.
4. Durwald, H. & Hoffmann-Berling, H. (1968). Endonuclease-I-deficient and ribonuclease I-deficient Escherichia coli mutants. *J Mol Biol* **34**, 331-46.
5. Puyet, A., Greenberg, B. & Lacks, S. A. (1990). Genetic and structural characterization of endA. A membrane-bound nuclease required for transformation of Streptococcus pneumoniae. *J Mol Biol* **213**, 727-38.
6. Provvedi, R., Chen, I. & Dubnau, D. (2001). NucA is required for DNA cleavage during transformation of Bacillus subtilis. *Mol Microbiol* **40**, 634-44.
7. Hsia, K. C., Li, C. L. & Yuan, H. S. (2005). Structural and functional insight into sugar-nonspecific nucleases in host defense. *Curr Opin Struct Biol* **15**, 126-34.
8. Braun, V., Pilsl, H. & Gross, P. (1994). Colicins: structures, modes of action, transfer through membranes, and evolution. *Arch Microbiol* **161**, 199-206.
9. Lee, S. D. & Alani, E. (2006). Analysis of interactions between mismatch repair initiation factors and the replication processivity factor PCNA. *J Mol Biol* **355**, 175-84.
10. Mortensen, U. H., Lisby, M. & Rothstein, R. (2009). Rad52. *Curr Biol* **19**, R676-7.
11. Friedhoff, P., Franke, I., Meiss, G., Wende, W., Krause, K. L. & Pingoud, A. (1999). A similar active site for non-specific and specific endonucleases. *Nat Struct Biol* **6**, 112-3.
12. Stoddard, B. L. (2005). Homing endonuclease structure and function. *Q Rev Biophys* **38**, 49-95.
13. Flick, K. E., Jurica, M. S., Monnat, R. J., Jr. & Stoddard, B. L. (1998). DNA binding and cleavage by the nuclear intron-encoded homing endonuclease I-PpoI. *Nature* **394**, 96-101.
14. Orlowski, J. & Bujnicki, J. M. (2008). Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses. *Nucleic Acids Res* **36**, 3552-69.
15. Pingoud, A., Fuxreiter, M., Pingoud, V. & Wende, W. (2005). Type II restriction endonucleases: structure and mechanism. *Cell Mol Life Sci* **62**, 685-707.
16. Eastberg, J. H., Eklund, J., Monnat, R., Jr. & Stoddard, B. L. (2007). Mutability of an HNH nuclease imidazole general base and exchange of a deprotonation mechanism. *Biochemistry* **46**, 7215-25.
17. Kuhlmann, U. C., Moore, G. R., James, R., Kleanthous, C. & Hemmings, A. M. (1999). Structural parsimony in endonuclease active sites: should the number of homing endonuclease families be redefined? *FEBS Lett* **463**, 1-2.

18. Mizuuchi, K., Kemper, B., Hays, J. & Weisberg, R. A. (1982). T4 endonuclease VII cleaves holliday structures. *Cell* **29**, 357-65.

19. Biertumpfel, C., Yang, W. & Suck, D. (2007). Crystal structure of T4 endonuclease VII resolving a Holliday junction. *Nature* **449**, 616-20.

20. Woo, E. J., Kim, Y. G., Kim, M. S., Han, W. D., Shin, S., Robinson, H., Park, S. Y. & Oh, B. H. (2004). Structural mechanism for inactivation and activation of CAD/DFF40 in the apoptotic pathway. *Mol Cell* **14**, 531-9.

21. Scholz, S. R., Korn, C., Bujnicki, J. M., Gimadutdinow, O., Pingoud, A. & Meiss, G. (2003). Experimental evidence for a beta beta alpha-Me-finger nuclease motif to represent the active site of the caspase-activated DNase. *Biochemistry* **42**, 9288-94.

22. Schafer, P., Scholz, S. R., Gimadutdinow, O., Cymerman, I. A., Bujnicki, J. M., Ruiz-Carrillo, A., Pingoud, A. & 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**, 217-28.

23. Meiss, G., Gimadutdinow, O., Haberland, B. & Pingoud, A. (2000). Mechanism of DNA cleavage by the DNA/RNA-non-specific Anabaena sp. PCC 7120 endonuclease NucA and its inhibition by NuiA. *J Mol Biol* **297**, 521-34.

24. Gorbalenya, A. E. (1994). Self-splicing group I and group II introns encode homologous (putative) DNA endonucleases of a new family. *Protein Sci* **3**, 1117-20.

25. Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R. & Bateman, A. The Pfam protein families database. *Nucleic Acids Res* **38**, D211-22.

26. Mehta, P., Katta, K. & Krishnaswamy, S. (2004). HNH family subclassification leads to identification of commonality in the His-Me endonuclease superfamily. *Protein Sci* **13**, 295-300.

27. Doudeva, L. G., Huang, H., Hsia, K. C., Shi, Z., Li, C. L., Shen, Y., Cheng, Y. S. & Yuan, H. S. (2006). Crystal structural analysis and metal-dependent stability and activity studies of the ColE7 endonuclease domain in complex with DNA/Zn2+ or inhibitor/Ni2+. *Protein Sci* **15**, 269-80.

28. Kleanthous, C., Kuhlmann, U. C., Pommer, A. J., Ferguson, N., Radford, S. E., Moore, G. R., James, R. & Hemmings, A. M. (1999). Structural and mechanistic basis of immunity toward endonuclease colicins. *Nat Struct Biol* **6**, 243-52.

29. Huang, H. & Yuan, H. S. (2007). The conserved asparagine in the HNH motif serves an important structural role in metal finger endonucleases. *J Mol Biol* **368**, 812-21.

30. Mate, M. J. & Kleanthous, C. (2004). Structure-based analysis of the metal-dependent mechanism of H-N-H endonucleases. *J Biol Chem* **279**, 34763-9.

31. Ghosh, M., Meiss, G., Pingoud, A. M., London, R. E. & Pedersen, L. C. (2007). The nuclease a-inhibitor complex is characterized by a novel metal ion bridge. *J Biol Chem* **282**, 5682-90.

32. Miller, M. D., Cai, J. & Krause, K. L. (1999). The active site of Serratia endonuclease contains a conserved magnesium-water cluster. *J Mol Biol* **288**, 975-87.

33. Lazdunski, C. J., Bouveret, E., Rigal, A., Journet, L., Lloubes, R. & Benedetti, H. (1998). Colicin import into Escherichia coli cells. *J Bacteriol* **180**, 4993-5002.

34. Housden, N. G., Loftus, S. R., Moore, G. R., James, R. & Kleanthous, C. (2005). Cell entry mechanism of enzymatic bacterial colicins: porin recruitment and the thermodynamics of receptor binding. *Proc Natl Acad Sci U S A* **102**, 13849-54.

35. Soelaiman, S., Jakes, K., Wu, N., Li, C. & Shoham, M. (2001). Crystal structure of colicin E3: implications for cell entry and ribosome inactivation. *Mol Cell* **8**, 1053-62.

36. Di Masi, D. R., White, J. C., Schnaitman, C. A. & Bradbeer, C. (1973). Transport of vitamin B12 in Escherichia coli: common receptor sites for vitamin B12 and the E colicins on the outer membrane of the cell envelope. *J Bacteriol* **115**, 506-13.

37. Watson, R., Rowsome, W., Tsao, J. & Visentin, L. P. (1981). Identification and characterization of Col plasmids from classical colicin E-producing strains. *J Bacteriol* **147**, 569-77.

38. Cooper, P. C. & James, R. (1984). Two new E colicins, E8 and E9, produced by a strain of Escherichia coli. *J Gen Microbiol* **130**, 209-15.

39. Cramer, W. A., Dankert, J. R. & Uratani, Y. (1983). The membrane channel-forming bacteriocidal protein, colicin El. *Biochim Biophys Acta* **737**, 173-93.

40. Schaller, K. & Nomura, M. (1976). Colicin E2 is DNA endonuclease. *Proc Natl Acad Sci U S A* **73**, 3989-93.

41. Toba, M., Masaki, H. & Ohta, T. (1988). Colicin E8, a DNase which indicates an evolutionary relationship between colicins E2 and E3. *J Bacteriol* **170**, 3237-42.

42. Eaton, T. & James, R. (1989). Complete nucleotide sequence of the colicin E9 (cei) gene. *Nucleic Acids Res* **17**, 1761.

43. Chak, K. F., Kuo, W. S., Lu, F. M. & James, R. (1991). Cloning and characterization of the ColE7 plasmid. *J Gen Microbiol* **137**, 91-100.

44. Boon, T. (1971). Inactivation of ribosomes in vitro by colicin E 3. *Proc Natl Acad Sci U S A* **68**, 2421-5.

45. Akutsu, A., Masaki, H. & Ohta, T. (1989). Molecular structure and immunity specificity of colicin E6, an evolutionary intermediate between E-group colicins and cloacin DF13. *J Bacteriol* **171**, 6430-6.

46. Ko, T. P., Liao, C. C., Ku, W. Y., Chak, K. F. & Yuan, H. S. (1999). The crystal structure of the DNase domain of colicin E7 in complex with its inhibitor Im7 protein. *Structure* **7**, 91-102.

47. Hsia, K. C., Chak, K. F., Liang, P. H., Cheng, Y. S., Ku, W. Y. & Yuan, H. S. (2004). DNA binding and degradation by the HNH protein ColE7. *Structure* **12**, 205-14.

48. Wallis, R., Moore, G. R., James, R. & Kleanthous, C. (1995). Protein-protein interactions in colicin E9 DNase-immunity protein complexes. 1. Diffusion-controlled association and femtomolar binding for the cognate complex. *Biochemistry* **34**, 13743-50.

49. Kuhlmann, U. C., Pommer, A. J., Moore, G. R., James, R. & Kleanthous, C. (2000). Specificity in protein-protein interactions: the structural basis for dual recognition in endonuclease colicin-immunity protein complexes. *J Mol Biol* **301**, 1163-78.

50. Enari, M., Sakahira, H., Yokoyama, H., Okawa, K., Iwamatsu, A. & Nagata, S. (1998). A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD. *Nature* **391**, 43-50.

51. Liu, X., Zou, H., Slaughter, C. & Wang, X. (1997). DFF, a heterodimeric protein that functions downstream of caspase-3 to trigger DNA fragmentation during apoptosis. *Cell* **89**, 175-84.

52. Zhou, P., Lugovskoy, A. A., McCarty, J. S., Li, P. & Wagner, G. (2001). Solution structure of DFF40 and DFF45 N-terminal domain complex and mutual chaperone activity of DFF40 and DFF45. *Proc Natl Acad Sci U S A* **98**, 6051-5.

53. Li, L. Y., Luo, X. & Wang, X. (2001). Endonuclease G is an apoptotic DNase when released from mitochondria. *Nature* **412**, 95-9.

54. Cote, J. & Ruiz-Carrillo, A. (1993). Primers for mitochondrial DNA replication generated by endonuclease G. *Science* **261**, 765-9.

55. Ohsato, T., Ishihara, N., Muta, T., Umeda, S., Ikeda, S., Mihara, K., Hamasaki, N. & Kang, D. (2002). Mammalian mitochondrial endonuclease G. Digestion of R-loops and localization in intermembrane space. *Eur J Biochem* **269**, 5765-70.

56. Dake, E., Hofmann, T. J., McIntire, S., Hudson, A. & Zassenhaus, H. P. (1988). Purification and properties of the major nuclease from mitochondria of Saccharomyces cerevisiae. *J Biol Chem* **263**, 7691-702.

57. Loll, B., Gebhardt, M., Wahle, E. & Meinhart, A. (2009). Crystal structure of the EndoG/EndoGI complex: mechanism of EndoG inhibition. *Nucleic Acids Res*.

58. Vincent, R. D., Hofmann, T. J. & Zassenhaus, H. P. (1988). Sequence and expression of NUC1, the gene encoding the mitochondrial nuclease in Saccharomyces cerevisiae. *Nucleic Acids Res* **16**, 3297-312.

59. Parrish, J., Li, L., Klotz, K., Ledwich, D., Wang, X. & Xue, D. (2001). Mitochondrial endonuclease G is important for apoptosis in C. elegans. *Nature* **412**, 90-4.

60. Widlak, P., Li, L. Y., Wang, X. & Garrard, W. T. (2001). Action of recombinant human apoptotic endonuclease G on naked DNA and chromatin substrates: cooperation with exonuclease and DNase I. *J Biol Chem* **276**, 48404-9.

61. Temme, C., Weissbach, R., Lilie, H., Wilson, C., Meinhart, A., Meyer, S., Golbik, R., Schierhorn, A. & Wahle, E. (2009). The Drosophila melanogaster gene CG4930 encodes a high affinity inhibitor for endonuclease G. *J Biol Chem*.

62. Moritani, M., Nomura, K., Tanahashi, T., Osabe, D., Fujita, Y., Shinohara, S., Yamaguchi, Y., Keshavarz, P., Kudo, E., Nakamura, N., Yoshikawa, T., Ichiishi, E., Takata, Y., Yasui, N., Shiota, H., Kunika, K., Inoue, H. & Itakura, M. (2007). Genetic association of single nucleotide polymorphisms in endonuclease G-like 1 gene with type 2 diabetes in a Japanese population. *Diabetologia* **50**, 1218-27.

63. Cymerman, I. A., Chung, I., Beckmann, B. M., Bujnicki, J. M. & Meiss, G. (2008). EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Res* **36**, 1369-79.

64. Buttner, S., Eisenberg, T., Carmona-Gutierrez, D., Ruli, D., Knauer, H., Ruckenstuhl, C., Sigrist, C., Wissing, S., Kollroser, M., Frohlich, K. U., Sigrist, S. & Madeo, F. (2007). Endonuclease G regulates budding yeast life and death. *Mol Cell* **25**, 233-46.

65. Ruiz-Carrillo, A. & Renaud, J. (1987). Endonuclease G: a (dG)n X (dC)n-specific DNase from higher eukaryotes. *EMBO J* **6**, 401-7.

66. Muro-Pastor, A. M., Flores, E., Herrero, A. & 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**, 3021-30.

67. Meiss, G., Franke, I., Gimadutdinow, O., Urbanke, C. & Pingoud, A. (1998). Biochemical characterization of Anabaena sp. strain PCC 7120 non-specific nuclease NucA and its inhibitor NuiA. *Eur J Biochem* **251**, 924-34.

68. Ghosh, M., Meiss, G., Pingoud, A., London, R. E. & Pedersen, L. C. (2005). Structural insights into the mechanism of nuclease A, a betabeta alpha metal nuclease from Anabaena. *J Biol Chem* **280**, 27990-7.

69. Muro-Pastor, A. M., Herrero, A. & Flores, E. (1997). The nuiA gene from Anabaena sp. encoding an inhibitor of the NucA sugar-non-specific nuclease. *J Mol Biol* **268**, 589-98.

70. Mol, C. D., Arvai, A. S., Sanderson, R. J., Slupphaug, G., Kavli, B., Krokan, H. E., Mosbaugh, D. W. & Tainer, J. A. (1995). Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA. *Cell* **82**, 701-8.

71. Miller, M. D., Benedik, M. J., Sullivan, M. C., Shipley, N. S. & Krause, K. L. (1991). Crystallization and preliminary crystallographic analysis of a novel nuclease from Serratia marcescens. *J Mol Biol* **222**, 27-30.

72. Ball, T. K., Suh, Y. & Benedik, M. J. (1992). Disulfide bonds are required for Serratia marcescens nuclease activity. *Nucleic Acids Res* **20**, 4971-4.

73. Miller, M. D. & Krause, K. L. (1996). Identification of the Serratia endonuclease dimer: structural basis and implications for catalysis. *Protein Sci* **5**, 24-33.

74. Franke, I., Meiss, G., Blecher, D., Gimadutdinow, O., Urbanke, C. & Pingoud, A. (1998). Genetic engineering, production and characterisation of monomeric variants of the dimeric Serratia marcescens endonuclease. *FEBS Lett* **425**, 517-22.

75. Balan, A. & Schenberg, A. C. (2005). A conditional suicide system for Saccharomyces cerevisiae relying on the intracellular production of the Serratia marcescens nuclease. *Yeast* **22**, 203-12.

76. Kadioglu, A., Weiser, J. N., Paton, J. C. & Andrew, P. W. (2008). The role of Streptococcus pneumoniae virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol* **6**, 288-301.

77. Bronze, M. S. & Dale, J. B. (1996). The reemergence of serious group A streptococcal infections and acute rheumatic fever. *Am J Med Sci* **311**, 41-54.

78. Ferreira, B. T., Benchetrit, L. C., De Castro, A. C., Batista, T. G. & Barrucand, L. (1992). Extracellular deoxyribonucleases of streptococci: a comparison of their occurrence and levels of production among beta-hemolytic strains of various serological groups. *Zentralbl Bakteriol* **277**, 493-503.

79. Beres, S. B., Sylva, G. L., Barbian, K. D., Lei, B., Hoff, J. S., Mammarella, N. D., Liu, M. Y., Smoot, J. C., Porcella, S. F., Parkins, L. D., Campbell, D. S., Smith, T. M., McCormick, J. K., Leung, D. Y., Schlievert, P. M. & Musser, J. M. (2002). Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proc Natl Acad Sci U S A* **99**, 10078-83.

80. Podbielski, A., Zarges, I., Flosdorff, A. & Weber-Heynemann, J. (1996). Molecular characterization of a major serotype M49 group A streptococcal DNase gene (sdaD). *Infect Immun* **64**, 5349-56.

81. Sumby, P., Barbian, K. D., Gardner, D. J., Whitney, A. R., Welty, D. M., Long, R. D., Bailey, J. R., Parnell, M. J., Hoe, N. P., Adams, G. G., Deleo, F. R. & Musser, J. M. (2005). Extracellular deoxyribonuclease made by group A Streptococcus assists pathogenesis by enhancing evasion of the innate immune response. *Proc Natl Acad Sci U S A* **102**, 1679-84.

82. Buchanan, J. T., Simpson, A. J., Aziz, R. K., Liu, G. Y., Kristian, S. A., Kotb, M., Feramisco, J. & Nizet, V. (2006). DNase expression allows the pathogen group A Streptococcus to escape killing in neutrophil extracellular traps. *Curr Biol* **16**, 396-400.

83. Brinkmann, V., Reichard, U., Goosmann, C., Fauler, B., Uhlemann, Y., Weiss, D. S., Weinrauch, Y. & Zychlinsky, A. (2004). Neutrophil extracellular traps kill bacteria. *Science* **303**, 1532-5.

84. Beiter, K., Wartha, F., Albiger, B., Normark, S., Zychlinsky, A. & Henriques-Normark, B. (2006). An endonuclease allows Streptococcus pneumoniae to escape from neutrophil extracellular traps. *Curr Biol* **16**, 401-7.

85. Aziz, R. K., Ismail, S. A., Park, H. W. & Kotb, M. (2004). Post-proteomic identification of a novel phage-encoded streptodornase, Sda1, in invasive M1T1 Streptococcus pyogenes. *Mol Microbiol* **54**, 184-97.

86. Lacks, S. (1962). Molecular fate of DNA in genetic transformation of Pneumococcus. *J Mol Biol* **5**, 119-31.

87. Wannamaker, L. W., Hayes, B. & Yasmineh, W. (1967). Streptococcal nucleases. II. Characterization of DNAse D. *J Exp Med* **126**, 497-508.

88. Heckman, K. L. & Pease, L. R. (2007). Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc* **2**, 924-32.

89. Cadwell, R. C. & Joyce, G. F. (1994). Mutagenic PCR. *PCR Methods Appl* **3**, S136-40.

90. Nadano, D., Yasuda, T. & Kishi, K. (1993). Measurement of deoxyribonuclease I activity in human tissues and body fluids by a single radial enzyme-diffusion method. *Clin Chem* **39**, 448-52.

91. Ellis, K. J. & Morrison, J. F. (1982). Buffers of constant ionic strength for studying pH-dependent processes. *Methods Enzymol* **87**, 405-26.

92. Toney, M. D. & Kirsch, J. F. (1989). Direct Bronsted analysis of the restoration of activity to a mutant enzyme by exogenous amines. *Science* **243**, 1485-8.

93. Carter, P. & Wells, J. A. (1987). Engineering enzyme specificity by "substrate-assisted catalysis". *Science* **237**, 394-9.

94. Wallis, R., Reilly, A., Barnes, K., Abell, C., Campbell, D. G., Moore, G. R., James, R. & Kleanthous, C. (1994). Tandem overproduction and characterisation of the nuclease domain of colicin E9 and its cognate inhibitor protein Im9. *Eur J Biochem* **220**, 447-54.

95. Korn, C., Meiss, G., Gast, F., Gimadutdinow, O., Urbanke, C. & Pingoud, A. (2000). Genetic engineering of Escherichia coli to produce a 1:1 complex of the anabaena sp. PCC 7120 nuclease NucA and its inhibitor NuiA. *Gene* **253**, 221-9.

96. Gentz, R. & Bujard, H. (1985). Promoters recognized by Escherichia coli RNA polymerase selected by function: highly efficient promoters from bacteriophage T5. *J Bacteriol* **164**, 70-7.

97. Friedhoff, P., Kolmes, B., Gimadutdinow, O., Wende, W., Krause, K. L. & Pingoud, A. (1996). Analysis of the mechanism of the Serratia nuclease using site-directed mutagenesis. *Nucleic Acids Res* **24**, 2632-9.

98. Prinz, W. A., Aslund, F., Holmgren, A. & Beckwith, J. (1997). The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the Escherichia coli cytoplasm. *J Biol Chem* **272**, 15661-7.

99. Kuzmic, P., Sideris, S., Cregar, L. M., Elrod, K. C., Rice, K. D. & Janc, J. W. (2000). High-throughput screening of enzyme inhibitors: automatic determination of tight-binding inhibition constants. *Anal Biochem* **281**, 62-7.

100. Williams, J. W. & Morrison, J. F. (1979). The kinetics of reversible tight-binding inhibition. *Methods Enzymol* **63**, 437-67.

101. Friedhoff, P., Gimadutdinow, O., Ruter, T., Wende, W., Urbanke, C., Thole, H. & Pingoud, A. (1994). A procedure for renaturation and purification of the extracellular Serratia marcescens nuclease from genetically engineered Escherichia coli. *Protein Expr Purif* **5**, 37-43.

102. Gabor, M. & Hotchkiss, R. D. (1966). Manifestation of Linear Organization in Molecules of Pneumococcal Transforming DNA. *Proc Natl Acad Sci U S A* **56**, 1441-1448.

103. Vosman, B., Kuiken, G., Kooistra, J. & Venema, G. (1988). Transformation in Bacillus subtilis: involvement of the 17-kilodalton DNA-entry nuclease and the competence-specific 18-kilodalton protein. *J Bacteriol* **170**, 3703-10.

104. Eklund, J. L., Ulge, U. Y., Eastberg, J. & Monnat, R. J., Jr. (2007). Altered target site specificity variants of the I-PpoI His-Cys box homing endonuclease. *Nucleic Acids Res* **35**, 5839-50.

105. Gruen, M., Chang, K., Serbanescu, I. & Liu, D. R. (2002). An in vivo selection system for homing endonuclease activity. *Nucleic Acids Res* **30**, e29.

8 Supplementary information

8.1 Abbreviations

A	Ampere
ATP	Adenosine triphosphate
AmpR	Ampicillin resistance
bp	Base pair(s)
BCSV	Bicistronic selection vector
BCNA	Bicistronic selection vector NucA
BCNS	Bicistronic selection vector *Sma*NucA
CAD	Caspase activated DNase
CDS	Coding sequence
ColE1	ColE1 origin of replication
cfu	Colony forming units
Da	Dalton
dNTPs	deoxyribonucleotide
DNA	Deoxyribonucleic acid
ds	Double-stranded
e.g.	Exempli gratia; for example
f1 ori	f1 origin of replication
g	Gram
EndoG	Endonuclease G
GST	Glutathione S-transferase
HEases	Homing endonucleases
k	Kilo
KanR	Kanamycin resistance
l	Liter
Lambda T0	Lambda T0 terminator
LacI	Lactose repressor
LB	Luria-Bertani
li	Linear
m	Milli/Meter

M	Molar
min	Minute
MW	Molecular weight
n	Nano
NFNA	Non fluorescent selection vector NucA
NFNS	Non fluorescent selection vector *Sma*Nuc
NucA	Nuclease of *Anabaena spec.*
oc	Open circle
OD	Optical density
OEP	Overlap extension PCR
p	Pico
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pUC	pUC origin of replication
RBS	Ribosomal binding site
REases	Restriction endonucleases
rpm	Revolutions per minute
RNA	Ribonucleic acid
RT	Room temperature
sc	Supercoiled
SDS	Sodium dodecyl sulfate
sec	Second
*Sma*Nuc	*Serratia* nuclease
ss	Single-stranded
T5	T5 promoter
T7	T7 promoter
UV	Ultraviolet
vs.	Versus
w/v	Weight/volume
µ	Micro
× *g*	Gravity