The Crystal Structure of *Mj*RelBE Protein Complex from *Methanococcus jannaschii*

Dissertation zur Erlangung des akademischen Grades des Doktors der Naturwissenschaften (Dr.rer.nat.)

Vorgelegt dem Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

von Djordje Francuski
aus Belgrad, Serbien

Berlin, Dezember 2008

Die vorliegende Arbeit wurde in der Zeit von Januar 2005 bis Dezember 2007 unter Anleitung von Prof. Dr. Wolfram Saenger am Institut für Chemie-Biochemie/Kristallographie der Freien Universität Berlin im Fachbereich Biologie, Chemie, Pharmazie durchgeführt.

1. Gutachter: Prof. Dr. Wolfram Saenger
2. Gutachter: Prof. Dr. Volker Haucke

Eingereicht am 05. Januar 2009
Tag der mündlichen Prüfung: 20. Mai 2009

Acknowledgements

This thesis would have never been finished without support, friendship and advice from a number of people who I would like to take this opportunity to thank.

First and foremost, I would like to thank Prof. Dr. Wolfram Saenger, for giving me the opportunity to accomplish this thesis in his workgroup and for his continuous faith in me.

To Prof. Dr. Heinz Welfle I would like to express my gratitude for the discussions regarding CD spectroscopy and to Prof. Dr. Volker Haucke I would like to thank for reading this dissertation on a very short notice.

I wish also to thank Dr. Ardeschir Vahedi-Faridi, who has provided friendship and advice through difficult times. There are so many friends and colleagues I would like to thank for their friendship in good and bad times. I am grateful to Wilhelm Weihofen for his critical reading of this dissertation manuscript, Dr. Hai Xu and Dr. Kazutaka Murayama for helping me in my first steps in protein crystallography, Aslan Cicek, Thomas Spräter, Marcus Pech, Jacek Biesiadka, Greg Raszewski, Sebastian Geibel, Maxim Rossmann, Traudy Wandersleben, Daria Slowik, Haydar Bulut and many more for their friendship. Thank you all. I wish also to thank Carsten Jakob, Clemens Langner and Claudia Alings for their help and assistance during this thesis; this goes also to my "Azubis" Jan and Stefanie. To Dr. Werner Schröder I am thanking, not only for N-terminal sequencing, RNA and primer synthesis, but also for inspiring discussions. I would also like to thank Dr. Peter Franke for MALDI-TOF analysis, Prof. Dr. Joachim Behlke for analytical ultracentrifugation analyses, Caterina Farnleitner for the *pET21relBE_MJ* construct and the group of Prof. Dr. Kenn Gerdes for the initial RelBE clones.

I am especially indebted and grateful to my parents and my family that supported me through difficult times. None of this would be possible without their unconditional support. Herby I would like to thank my brother Nebojša and his family, Maja and Maša, for their patience, love and support.

Abbreviations

% (w/v)	x % (w/v), meaning x gram per 100ml of final solution
Å	1 Ångström = 0,1 nm
AA	Amino acids
Alarmones	Small molecules produced by bacteria that are exposed to internal or external stress, such as starvation.
AS	Ammonium sulphate, $(NH_4)_2SO_4$
ATP	Adenosinetriphosphate
AU	Analytical ultracentrifugation
au	arbitrary units
bp	base pairs (DNA or RNA), also kbp (kilo base pairs) and Mbp (mega base pairs)
CD	Circular dichroism
DLS	Dynamic light scattering
DNA	Desoxyribonucleic acid
DNase	Desoxyribonuclease
dsDNA	double stranded DNA
*Ec*RelB	*Escherichia coli* RelB
*Ec*RelBE	*Escherichia coli* RelBE protein complex
*Ec*RelE	*Escherichia coli* RelE
EDTA	Ethylenediaminetetraacetate
HEPES	N-2-Hydroxyethylpiperazine- N'-2-ethane sulfonic acid
His_6-tag	Short amino acid sequence consisting of six histidines
HTH	Helix-turn-helix motive, a fold motive of DNA binding proteins
HPLC	High performance liquid chromatography
IEF	Isoelectric focusing
IPTG	isopropyl-β-D-thiogalactopyranoside
kDa	Kilo Dalton
LB-Medium	Luria Bertani medium
M	mol/l
MAD	Multi wavelength anomalous dispersion
MALDI-TOF-MS	Matrix assisted laser desorption/ionization – time of flight mass spectrometry
MCS	Multiple Cloning Site in a vector
MIR	Multiple isomorphous replacement
MIRAS	Multiple isomorphous replacement with anomalous scattering
*Mj*RelB	*Methanococcus jannaschii* RelB
*Mj*RelBE	*Methanococcus jannaschii* RelBE protein complex
*M*jRelE	*Methanococcus jannaschii* RelE
MPD	2-Methyl-2,4-pentanediol
MR	Molecular replacement
mRNA	Messenger ribonucleic acid

MS	Mass spectrometry
Mw	Molecular weight, g/mol
nt	nucleotides
N-terminal	Amino-terminal
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCD	programmed cell death
PCR	Polymerase chain reaction
PDB	Brookhaven Protein Data Bank
PEG	Polyethyleneglycol
PEI	Polyethyleneimine
*Ph*RelB	*Phyrococcus horikoschii* RelB
*Ph*RelBE	*Phyrococcus horikoschii* RelBE protein complex
*Ph*RelE	*Phyrococcus horikoschii* RelE
PMSF	phenylmethylsulphonyl fluoride
ppGpp	Guanosine tetraphosphate
pppGpp	Guanosine pentaphosphate
PSK	Post Segregation Killing systems
RBS	Ribosome binding site
RelA	The enzyme, also known as ppGpp synthetase I, which synthesizes P4G in response to amino-acid starvation.
rmsd	Root mean square deviation
RNA	Ribonucleic acid
RNAi	RNA Interference
RNase	Ribonuclease
rRNA	Ribosomal ribonucleic acid
SDS	Sodium dodeca sulphate
SIRAS	Single isomorphous replacement with anomalous scattering
SpoT	This enzyme, also known as ppGpp synthetase II, has both P4G synthethase and hydrolase activities.
ssDNA	single stranded DNA
ssRNA	single stranded RNA
TA	toxin-antitoxin systems
Tricine	N-(2-Hydroxy-1,1-bis(hydroxymethyl)ethyl)glycine
Tris	Tris(hydroxymethyl)-aminomethan
tRNA	Transfer ribonucleic acid
UV	Ultraviolet (light)
VIS	Visible (light)
wt	Wild type

INDEX

1 Introduction

Prokaryotes can be classified in two major groups, *eubacteria* and *archaebacteria* which in turn diversify into multiple subgroups. In metabolic pathways and processes of present bacteria rudiments of the processes and pathways of later evolved *eukaryotes* can be found. Although prokaryotes are generally considered “primitive organisms” they feature, depending on their ecologic niche and life style, diversified and specialized metabolic pathways. As they have to adjust to varying food sources, environmental conditions and emerging predators or infestations (like bacteriophage), they evolved mechanisms for rapid adjustment to stress conditions. Prokaryotic life is occupying all habitats, from the troposphere (Christner *et al.*, 2008) to the depths of oceans, in soil and even inside other living cells.

Archaebacteria on the other hand, inhabit the most hostile environments for any livinig organisms, like hot water springs, acid lakes or deep ocean abysses. A widely studied representative of this group is *Methanococcus jannaschii*, which is also the first *archaebacterium* to have its DNA sequenced. It thrives 2000 m deep under the sealevel at temperatures around 85°C and uses carbon dioxide as the primary carbon source.

Eubacteria settle in various environmental niches. They can be found as free-living or as intracellular parasites like members of the *rickettsia* and *chlamydia* genus or some *Mycobacterium* species that can cause tuberculosis and leprosy.

1.1 Short Recapitulation of Cell Biology and the Phylogenic Tree of Life

The living organisms are roughly classified into prokaryotes (in Greek: before nuclei) and eukaryotes (in Greek: true nuclei), with the main difference that eukaryotes have their chromosomes compartmentalized in their nucleus while chromosomes in prokaryotes are not so tightly packed and are not separated from the rest of the cytosol by a membrane. Prokaryotes are further grouped in archea and eubacteria (Figure 1), which all in turn are subdivided in families and species. Viruses occupy a special place in this hierarchy as their origin is still disputed.

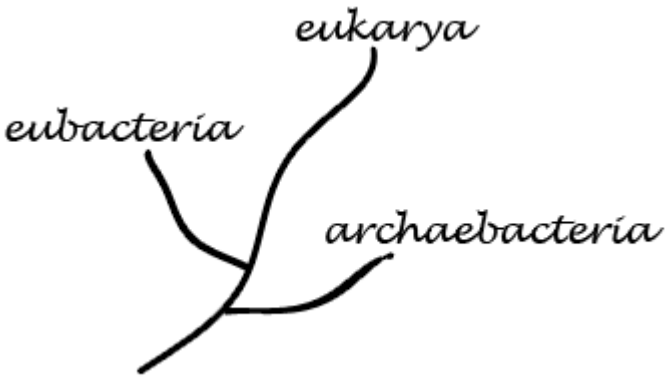

Figure 1 The three main domains of life on earth are depicted in the phylogenic tree. The eubacteria and archaebacteria are often separated from eukaryotes (eukarya) by referring to them as prokaryotes as they lack the distinction of eukaryotes, the cell nuclei.

Prokaryotic cells are structurally simple (Figure 2) but biochemically diverse. They often possess a tough protective coat, called a cell wall, beneath which a plasma membrane encloses a single cytoplasmatic compartment containing DNA, RNA, proteins and small molecules (Alberts *et al.*, 1994). Some bacteria also have a form of cytoskeleton. The cell wall ensures that the cell content is separated from the outer medium and controls transfer of water, nutritients and metabolic products of the cell between the cytosol and the outer medium. Cell wall structure and composition depends on the species and is usually made of more than one layer. The chromosome encodes all the proteins (as enzymes and structural proteins) and RNAs (including the ribosomes and tRNAs) that the cell needs to survive and to multiply. Prokaryotes have more than one possibility to spread their DNA, either through cell division or by conjugation. Some prokaryotes feature extra-chromosomal DNA elements called plasmids which are circular and capable of independent replication. They vary in sizes from 1kbp to over 200 kbp, and their number may vary in a cell from one to 1000 copies. Antibiotic resistance is often coded in plasmids and can be horizontally transferred (between bacterial species) from one cell to another. Plasmids often encode genes for specialized proteins that ensure plasmid maintenance in the host cells.

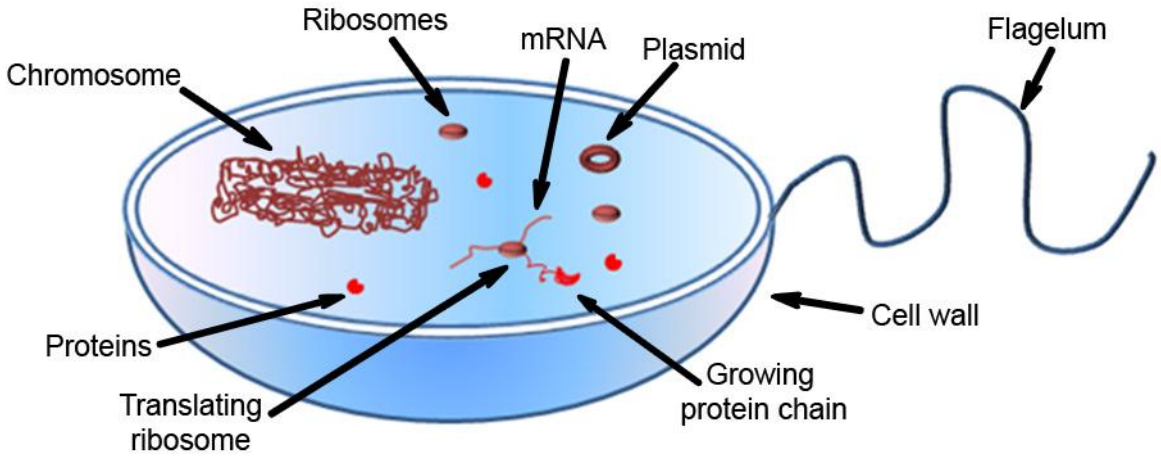

Figure 2 The prokaryotic cell. Some of the components of a prokaryotic cell are depicted. The cell wall physically separates the cell content from the surroundings, the flagellum ensures mobility, the chromosome encodes the proteins and the ribosomes link amino acids in polypeptide chains that form the proteins. The extra-chromosomal elements plasmids are often found in prokaryotic cells.

1.1 Stringent response

The constant environmental changes forced the evolution of global regulatory responses in free-living bacteria that adjust their rate of metabolism. As a response to environmental stress, like amino-acid starvation or carbon-source limitation, these regulatory responses adjust rates of intracellular metabolic processes. This is done by down-regulating synthesis of stable rRNAs and tRNAs. Specifically, during amino-acid starvation the lack of amino acids causes the binding of uncharged tRNAs to the ribosomal A-site. This in turn activates RelA to synthesize the alarmones guanosine tetraphosphate (ppGpp) and pentaphosphate (pppGpp) that in turn bind to and block rRNA promoter regions (Gerdes *et al.*, 2005). In case of carbon starvation an inhibition of SpoT has the same effect and results in an increase of intracellular alarmone concentrations (Figure 3).

Literature reports suggest that some prokaryotes have developed systems comparable to the apoptotic systems of eukaryotes (Engelberg-Kulka and Glaser, 1999; Lewis, 2000; Rice and Bayles, 2003). This is contrary to the general belief of prokaryotes as "selfish" cells while the eukaryotes are "altruistic" in the sense that they would sacrifice themselves in order to save, or promote growth of other cells. This so called programmed cell

death (PCD), has also been brought into context with the plasmid or chromosome coded pairs of toxins and their cognate antitoxins (Lewis, 2000; Rice and Bayles, 2003).

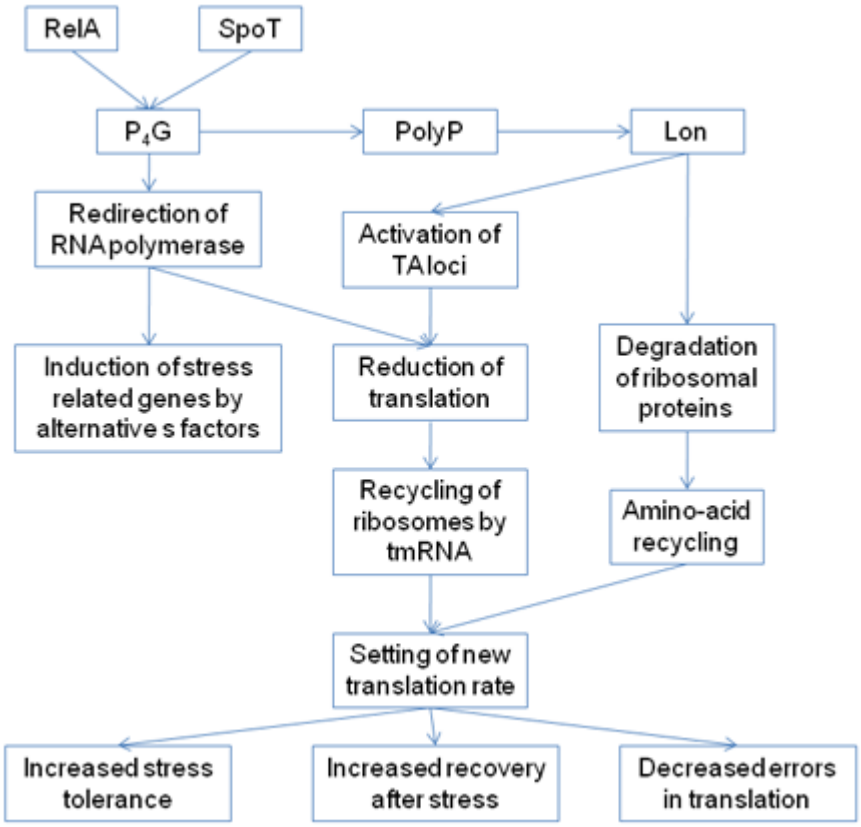

Figure 3 Stringent response scheme (adapted from Gerdes, 2005). Prokaryotic cells under stress conditions increasingly synthesise alarmones (P4G). The increased levels trigger Lon protease to degrade ribosomal proteins and the TA antitoxins, freeing thereby the toxins that inhibit translation. The RNA-polymerase is redirected to transcription of stress related genes. All this leads to reduction of translated proteins. The ribosomes are recycled by tmRNA and a new translation rate is set.

1.2 The Toxin- Antitoxin Systems

In recent years, a series of toxin-antitoxin (TA) systems were discovered. Initially they were identified in plasmids due to their ability to prevent the formation of plasmid free progeny (Gerdes *et al.*, 1986; Ogura and Hiraga, 1983). Subsequently more homologues were discovered in chromosomal DNA (Masuda *et al.*, 1993) of cell lines that have a so called “delayed relaxed” phenotype. This trait was defined by mutations in chromosomal genes encoding the antitoxin RelB of the corresponding toxin RelE (Gotfredsen and Gerdes, 1998). Various names that were used to describe these systems included ‘proteic toxin–

antitoxin loci', 'poison/antitoxin systems', 'plasmid addiction systems' or 'Post Segregation Killing systems' (PSK). The TA systems are also referred to as Type I and Type II, depending on the antitoxin type being antisense RNA or a protein, respectively (Hayes, 2003). In this work the systems regulated with a protein antitoxin will be named 'toxin–antitoxin loci' or TA systems (Gerdes *et al.*, 2005).

Plasmid borne TA systems are part of the plasmid maintenance systems. Plasmid maintenance and inheritance is achieved either through high number of plasmid copies (high-copy plasmids) or by specialized maintenance systems. It has been noted that smaller plasmids tend to use the first method as they minimize the size of their genome throwing out every excess genetic elements, like maintenance systems, but rather increase their copy number. Oppositely to this, large plasmids usually tend to keep their copy number low. Therefore the plasmids account for 2-5% of total DNA in the cell, regardless of plasmid type (low- or high-copy) they carry (Nordström and Austin, 1989). Low-copy plasmids counteract the higher probability of plasmid loss by increasing the number of maintenance cassettes. Each maintenance cassette provides another system that increases the probability of stable plasmid segregation to daughter cells or maintenance in the cell. There are three types of plasmid maintenance systems, (i) resolvases that ensure the separation of oligomerized plasmids, (ii) active plasmid partitioning systems that ensure better than random segregation of plasmids and (iii) the above mentioned TA systems (Nordström and Austin, 1989).

A TA system (Figure 4) is comprised of a toxin and its antitoxin encoded just upstream of the toxin (Alonso *et al.*, 2007). The transcription of this bicistronic operon is inhibited by binding of the antitoxin to the promoter (a multiple palindrome sequence), alone or in complex with the toxin. There are some exceptions, like HigBA (Tian *et al.*, 1996), where the toxin is encoded upstream of the antitoxin, and εζ (Cegłowski *et al.*, 1993), where a third protein, ω, is necessary for transcriptional control. This does not, however, change the overall picture of the TA properties. The antitoxin is susceptible to rapid protease degradation while the toxin is more stable. If the plasmid carrying genes encoding the TA system is lost, the antitoxin in the cell decays while the toxin persists and is free to kill the cells. In this way only those cells that carry the plasmid survive because their antitoxin pool is constantly replenished in excess to the toxin.

In later research, genes of homologous TA systems were discovered in chromosomes (Gotfredsen and Gerdes, 1998) of many bacteria and archaeons. It was evident that these TAs encoded in the chromosome may have another role, although they are in many ways very similar to the plasmid borne TAs. Investigating their biological role, it has also

been proposed that they have a function in bacterial programmed cell death (PCD) (Lewis, 2000; Tony J Greenfield, 2000). According to newer research it is believed that the TA systems encoded in the chromosome rather act as cellular regulators that help the cells adapt their gene expression to environmental stress (Gerdes *et al.*, 2005).

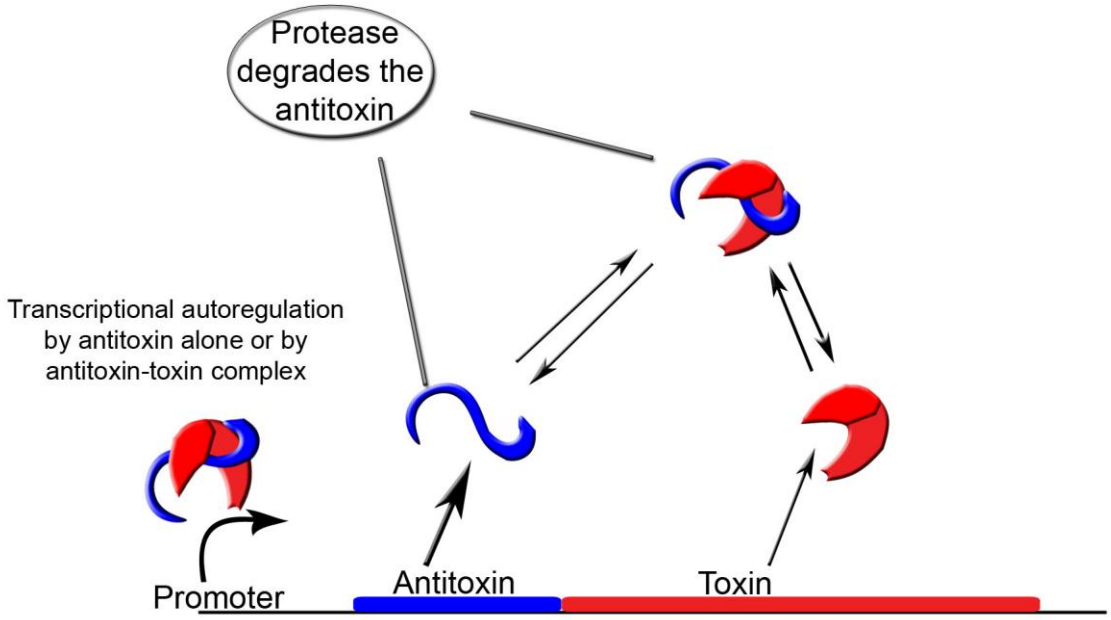

Figure 4 TA system scheme. The antitoxin and toxin are part of the same operon controlled by an autoregulated promoter. The autoregulation is achieved either by the antitoxin or the antitoxin-toxin complex. The antitoxin, being the first in the transcription sequence, is expressed in excess to the toxin. The toxin is a stable protein while the antitoxin is degraded by a protease and if the pool of the antitoxin is not replenished, the toxin is left alone ready to act.

Through extensive data base mining (Pandey and Gerdes, 2005), the number of identified and proposed TA pairs increased dramatically. The currently known toxins can be subdivided into the four TA superfamilies called CcdB, RelE, Doc and ζ, each consisting of distinct gene families that have their representatives in plasmids and chromosomes of a wide range of organisms. There is also a proposed fifth superfamily consisting of an N-terminal domain (PilT-N) that has not yet been studied thoroughly (Anantharaman and Aravind, 2003). This N-terminal PilT-N domain is also found in proteins of the eukaryotic nonsense-mediated RNA decay system (Anantharaman and Aravind, 2003; Shyu *et al.*, 2008).

The **CcdB** toxin superfamily includes the families of CcdB, Kid (PemK), MazF and ChpBK. The **RelE** family encompasses RelE, YoeB (RelE-3), YafQ (RelE-2), HigB, ParE, Txe, YhaV, while **Doc** from Prophage P1 plasmid might only be a trigger in a toxin cascade that ends with MazF (Hazan *et al.*, 2001) and is the only studied representative of its superfamily. The ζ toxin superfamily differs, as mentioned above, from the other TA systems

as the expression of toxin and antitoxin is regulated by a third protein, the repressor ω. The toxin ζ is inactivated by direct binding of the antitoxin ε that blocks the entrance to the proposed ATP site with its N-terminal helix (Meinhart *et al.*, 2003). Site-directed mutagenesis suggested that free ζ may act as a phosphotransferase using ATP to phosphorylate an as-yet unidentified substrate. Recent investigations (Lioy *et al.*, 2006) excluded some possible targets involved in either chromosomal segregation, DNA topology, cell division or translation. An overview of the TA systems is provided in Table 1.

Classification of the corresponding antitoxins is more difficult as they are generally less sequence related than the toxins. Their folds, however, show some surprising connections among the different toxin superfamilies. The antitoxins of the RelE superfamily belong to the group of DNA binding proteins with MetJ/Arc domain fold but some of them (like Axe) belong to the distinct YefM/Phd antitoxin superfamily. These findings connect the TA superfamilies and show the probable path of their evolution. The exchange of antitoxin partners and changes in the toxin target obviously took place during evolution of the TA systems. The most obvious targets in the cell for any toxin are translation, transcription or cell division. Therefore it is not surprising for the toxins from different superfamilies to target the same cellular component (see section 1.3), neither is surprising that the cellular targets inside one superfamily may differ. The changes that the antitoxins suffer by the adaptation of the toxin to its new target during evolution resulted in primary sequence differences and structural diversity among the antitoxins. The task of the antitoxin is to bind firmly to the toxin, block the sites on the toxin surface involved in the toxin-target interaction, and to bind to the promoter region controlling the TA operon.

Table 1 Overview of TA systems. In the column for oligomerization state the experimentally verified oligomerization states are given. In the "spread and gene location" column the phylogenetic distribution of the gene families are given. Table based on (Anantharaman and Aravind, 2003; Gerdes et al., 2005; Sevin and Barloy-Hubler, 2007).

Toxin/ Antitoxin	**Toxin superfamily**					**Antitoxin superfamily**				**Toxin (T) Antitoxin (A) oligomer state**	**Spread and gene location**
	MazF	RelE	Doc	PIN	ζ	AbrB	MetJ/ Arc	YefM/ Phd	HTH motive		
CcdB/ CcdA	+						+			T2A2T2 (Buts *et al.*, 2005)	Plasmid of Gram-negative bacteria
Kid/ Kis	+					+				TA2T concentration dependent (Kamphuis *et al.*, 2007)	Plasmid of Gram-negative and Gram-positive bacteria
MazF/ MazE	+					+				T2A2T2 (Kamada *et al.*, 2003)	Chromosome of Gram-negative and Gram-positive bacteria
ChpBK/ ChpBI	+					+					Chromosome of Gram-negative bacteria
RelE/ RelB		+					+			TAAT (this work and (Takagi *et al.*, 2005)	Chromosome of Gram-negative and Gram-positive bacteria, Archaea
YoeB/ YefM		+						+		TA2T (Kamada and Hanaoka, 2005)	Chromosome of Gram-negative bacteria
YafQ/ DinJ		+					+				Chromosome of Gram-negative bacteria
HigB/ HigA		+							+		Plasmid of Gram-negative and Gram-positive bacteria
ParE/ ParD		+					+				Plasmid of Gram-negative and Gram-positive bacteria
YhaV/ PrlF		+				+				(TAT)2 (Schmidt *et al.*, 2007)	Chromosome of Gram-negative bacteria
Txe/ Axe		+						+			Plasmid of Gram-negative and Gram-positive bacteria
Doc/ Phd			+					+		TA2 (Gazit and Sauer, 1999)	Chromosome and Plasmids of Gram-negative and Gram-positive bacteria, Archaea
VapC/ VapB				+		+					Chromosome of Gram-negative and Gram-positive bacteria, Archaea
ζ/ ε (ω)*					+		+			TA2T (Meinhart *et al.*, 2003)	Chromosome and Plasmids of Gram-positive bacteria

*the promoter regulation is achived via a third protein, see text for details

1.2.1 RNA- Regulated Loci - Post-Segregational Killing (PSK)

Antisense-RNA regulated systems are widespread in nature. In eukaryotes, processes like splicing, editing or rRNA modification, developmental regulation or the more recently discovered RNA interference (RNAi) (Fire *et al.*, 1998) make use of complementary RNAs (antisense-RNAs). However, the first discovered antisense-RNA regulated system was in a prokaryote, in the *Escherichia coli* plasmid ColE1 (Tomizawa *et al.*, 1981). Since then it has been established that antisense RNAs are principal regulators in prokaryotic accessory DNA elements (Weaver, 2007). The antisense-RNAs in plasmids regulate replication, segregational stability or conjugation of plasmids. They have been found also in phages where they fine-tune the decision between lysis and lysogeny or regulate transposition (Hartmann *et al.*, 1995; Wagner and Simons, 1994).

Antisense RNAs are small, diffusible, untranslated, highly structured molecules that bind to their target RNAs (sense RNAs) thereby controlling expression of the target genes (Brantl, 2002). Usually they are encoded in *cis* from the complementary DNA strand but *trans* –encoded antisense RNAs have been also found. The inhibition activities of these antisense-RNAs depend on their sequence and conformation.

Antisense-RNA regulated plasmids include the *inc18* plasmid families, like plasmids from gram positive hosts pIP501 (pAMβ1 and pSM19035 are related to it) and pT181 whose copy number control is achieved over the suppression of rep-mRNA (coding for Rep protein) by its antisense-RNA. Further there is also the *incFII* plasmid family hosted in gram negative *E. coli*. The pR1 plasmid belonging to this family is the best studied example of antisense-RNA regulation. Here the antisense-RNA, CopA, binds to sense-RNA, CopT, which is the leading part of the *repA*-mRNA that codes for the replication control protein RepA. A well studied example is also the first discovered antisense regulated system. It regulates the replication of plasmid ColE1 but in contrast to the above described systems it does not affect the expression of a protein coding gene. The replication initiation on this plasmid is accomplished over a long lived RNA primer that is transcribed from the plasmid by a host RNA polymerase. This primer binds to the plasmid origin region and allows the replication to start. The antisense-RNA binds directly to this RNA primer and inhibits its binding to the plasmid DNA (Tomizawa *et al.*, 1981).

The most studied PSK system, based on antisense-RNA antitoxin, is the *hok/sok* locus (Gerdes *et al.*, 1986; Gerdes *et al.*, 1986) from the plasmid R1. The stable toxin hok, a 52 amino acids (AA) long protein, is encoded by the *hok* gene that is translationally coupled with the overlapping *mok* gene. This long-lived mok mRNA is neither transcribed nor is the target of sok. Mok is slowly processed by RNaseII, by truncation at the 3' end, yielding the hok mRNA. This hok mRNA is capable of transcription but in presence of sok mRNA an RNA duplex is rapidly formed preventing ribosome binding and also presenting a target for RNaseIII dsRNA degradation. This is the way the unstable 64 nucleotides (nt) long antisense-RNA sok is preventing the hok protein being transcribed. Daughter cells that lose the plasmid during cell division still contain the mRNA pool of the original cell. Due to rapid degradation of the sok mRNA there is nothing to stop the transcription of the hok mRNA, and the plasmid deprived cell is doomed to hok caused cell membrane damage and the resulting cell death (Gerdes *et al.*, 1997).

There are other similar systems, like the *par* stability locus of *Enterococcus faecalis* plasmid pAD1 (Brantl, 2002), but there are also chromosomally encoded *hok/sok* loci (Pedersen and Gerdes, 1999) whose function is still unknown. It has been suggested that they ensure stable maintenance of chromosomal regions in which they reside (Gerdes and Wagner, 2007).

1.2.2 The CcdB toxin class

The CcdAB TA system is the first addiction module discovered (Ogura and Hiraga, 1983; Ogura and Hiraga, 1983) followed by KisKid (PemIK) (Bravo *et al.*, 1987; Tsuchimoto *et al.*, 1988) whose chromosomal homologues, MazEF and ChpBIK (Masuda *et al.*, 1993; Metzger *et al.*, 1988), were soon discovered. The structures of the CcdB, Kid and MazF toxins were determined (Hargreaves *et al.*, 2002; Kamada *et al.*, 2003; Loris *et al.*, 1999) and although they share low sequence identity, MazF to Kid at 25%, MazF to CcdB at 16% (Kamada *et al.*, 2003) and Kid to CcdB at 11% (Hargreaves *et al.*, 2002), they share a common structural fold (Table 2).

Table 2 Superimposition quality of CcdB toxin superfamily

Toxin	Sequence identity (%)	Overlapped residues	r.m.s.d. (Å)	Compared to toxin	Reference
MazF	25	89	2.2	Kid	Kamada et al., 2003
MazF	16	98	2.2	CcdB	Kamada et al., 2003
CcdB	11	65	1.3	Kid	Hargreaves et al., 2002

The secondary structures of toxins of this TA class are predominantly β stranded and generally contain seven β strands and three α helices in a conserved pattern. The first three β strands with the last two β strands form a twisted antiparallel β sheet, while the fourth strand together with the fifth strand and the C terminus of the third strand form an additional small antiparallel β–sheet.

From this group the complete structure with the antidote was only determined of the MazEF TA complex, but sequence comparison suggested that both CcdAB and KisKid form the same tertiary structure of a heterohexamer. MazEF forms a heterohexamer where one antidote MazE binds two toxins MazF, $MazF_2$-$MazE_2$-$MazF_2$ (Kamada *et al.*, 2003). The toxins form stable dimers in solution and were also crystallized in this form. This seems to be the active form of the toxins as it was shown in the case of Kid that interacts as a dimer with single stranded RNA even at tenfold excess, as shown by native mass spectrometry (Kamphuis *et al.*, 2006). As described above the MazE and Kis proteins have a similar fold, forming with their N-terminal domains a swapped β barrel combining two antidote molecules that bind with their C terminal domains one toxin dimer each.

Although the NMR structure of CcdA shows a ribbon helix helix (RHH) fold (Madl *et al.*, 2006), rather than a β barrel, it does not change the fact that this forms a dimer with its N-terminal domain and interacts with the toxin dimer with its C-terminal domain. These data suggest heterohexamer formation in solution, but upon DNA binding (as they co-repress their promoter by cooperative binding of the complex to the DNA) the best stoichiometry for binding is 1:1. Known TA complexes of this class are heterohexamers in solution, by a ratio of 2:1 in favor of the antidote. Surprisingly the binding to DNA seems to be performed by a complex closer to 1:1 toxin-antidote ratio. A mixture of 1:1 tends to form

aggregates but it is stabilized in presence of double helical DNA. Therefore it is possible that both MazEF and CcdAB mediate assembly of extended heterocomplex (toxin and antitoxin together) structures on DNA (Kamada *et al.*, 2003).

A well conserved feature of this class is the single area of positive potential on the surface of the toxin homodimer, well-circumscribed by two symmetrical loops on the monomers. These loops change their conformation upon antidote (or target RNA) binding. The well conserved residues that are also crucial for the RNase activity of MazF (D76) and Kid (D75 and H17) (Kamphuis *et al.*, 2006), are placed in this basic region. In contrast, mutagenesis studies on CcdB have suggested that the target interaction surface of CcdB involves the last three residues of its C-terminal helix (Bahassi *et al.*, 1995), that is, on the opposite side of the homodimer than the proposed active site in Kid or MazF. Since CcdB is a gyrase poison, it can be thought of as an adoption of the fold to a new target. It is very likely that these systems evolved from a common ancestor and later specialized on different targets. To counteract the toxins, their cognate antitoxins had to evolve accordingly, explaining the differences among the antitoxins and in their binding mode to their toxins.

1.2.3 The RelE toxin class

The toxin RelE belongs to a larger structural family of endoribonucleases. It is also the name carrier of its own TA system superfamily. The known structural representatives of this toxin superfamily (Table 1), feature a similar ribonuclease fold are RelE from *Pyrococcus horikoshii* (Takagi *et al.*, 2005; Wilson and Nierhaus, 2005) (denoted as *Ph*RelE), YoeB from *E. coli* (Kamada and Hanaoka, 2005) and RelE from *M. jannaschii* (denoted as *Mj*RelE) described herein. Structurally related proteins that are not part of TA systems but share the endoribonuclease fold are Colicins D and E5 (both tRNases (Graille *et al.*, 2004; Lin *et al.*, 2005)), RNase Sa from *Streptomyces aureofaciens* (and its family members) and RegB endoribonuclease that is involved in the control of the bacteriophage T4 multiplication cycle (Uzan *et al.*, 1988) .

The main representative of this TA superfamily, the RelBE system, was first discovered in the *E. coli* chromosome (Gotfredsen and Gerdes, 1998). Further representatives were found in various bacteria and archaea, some of them with almost identical sequences while others were first identified as novel TA systems. Connections between these proteins

with low sequence homology (Anantharaman and Aravind, 2003; Gerdes *et al.*, 2005) weren't revealed until later. The toxins of the RelE superfamily are characterized by a five-stranded β-sheet with four of the strands antiparallel to each other while the first and last strands are parallel. The β-sheet is flanked on one side by two α-helices α1, α2 and on the other side by α-helix α3 that, together with the loop connecting strands β2 , β3 forms a deep cleft that encompasses the RNase active site located on the five-stranded β-sheet as later described in more detail. A comparable architecture is also found in genuine RNases that have very low sequence homology to RelE.

The *relBE* promoter is strongly activated during nutritional stress and the protease Lon is required (Christensen *et al.*, 2001) for degradation of antitoxin RelB. There is competition between the translation release factor RF1 and RelE for ribosome binding, and it was suggested that RelE enters the ribosomal A-site to cleave mRNA (Pedersen *et al.*, 2003) at the second position of the A-site codon, preferentially at stop codons (UAG, UAA, UGA), but at sense codons as well (UCG, CAG) (Christensen and Gerdes, 2003; Pedersen *et al.*, 2003), leading to global inhibition of protein synthesis and eventually to cell growth arrest. This toxic effect may be reversed by expression of tmRNA (which is also cleaved by RelE) (Christensen and Gerdes, 2003) that recycles ribosomes by entering the A-site of ribosomes stalled on truncated mRNAs.

The toxic action of RelE is counteracted by antitoxin RelB which prevents the binding to mRNA by wrapping around the toxin. It was suggested by Takagi *et al.* (Takagi *et al.*, 2005) that the size of RelE is increased by antitoxin binding so that it is sterically not able to enter the A-site, and that the antitoxin is competing with RNA for binding to the RelE toxin.

1.3 Applications

There have been different approaches to TA systems applications. They have already been used as aids in cloning in plasmids but much effort has been invested in development of novel antibiotics that would target these TA systems. A new development is

the use of bacterial toxins in targeted treatment of cancer cells. In Table 3 the known cellular targets of TA systems are given.

Table 3 Known targets of TA toxins and proteases targeting the cognate antitoxins.

TA family (antitoxin / toxin)	Target of toxin	Protease degrading the antitoxin
CcdAB	Inhibition of replication through DNA gyrase	Lon
ParDE	Inhibition of replication through DNA gyrase	Unknown
RelBE	Translation through mRNA cleavage	Lon
MezEF	Translation through mRNA cleavage	ClpXP / Lon
Kis Kid	Translation through mRNA cleavage	Lon
PhD Doc	Translation through Ribosome inhibition	ClpAP

1.3.1 TA system toxins as cloning aids

The first TA system discovered soon found its application. The CcdB toxin is used as a control mechanism and enhancement for insertion of genes in plasmids. This is used in the Gateway® technology (Invitrogene™).

The system is comprised of the entry clone, destination vector and the necessary enzymes and *E. coli* strains. As described in the Gateway® cloning kit, the first step to Gateway cloning is the insertion the gene of interest into an entry clone (the Gateway). The entry clone contains the system vector and the gene of interest, but can also be cloned just as a DNA fragment that contains the gene of interest with the required recombination sites L1 and L2 flanking the gene. The two vectors, entry vector (or just the constructed PCR fragment) and the destination vector are mixed together with the required enzymes. The genetic elements are then exchanged among them via a recombination reaction (Figure 5). The required enzymes are Int (integrase), IHF (integration host factor) and Xis (excisionase). Int and Xis are enzymes derived from the λ bacteriophage genome while IHF is an *E. coli* protein (all contained in the LR Clonase Mix).

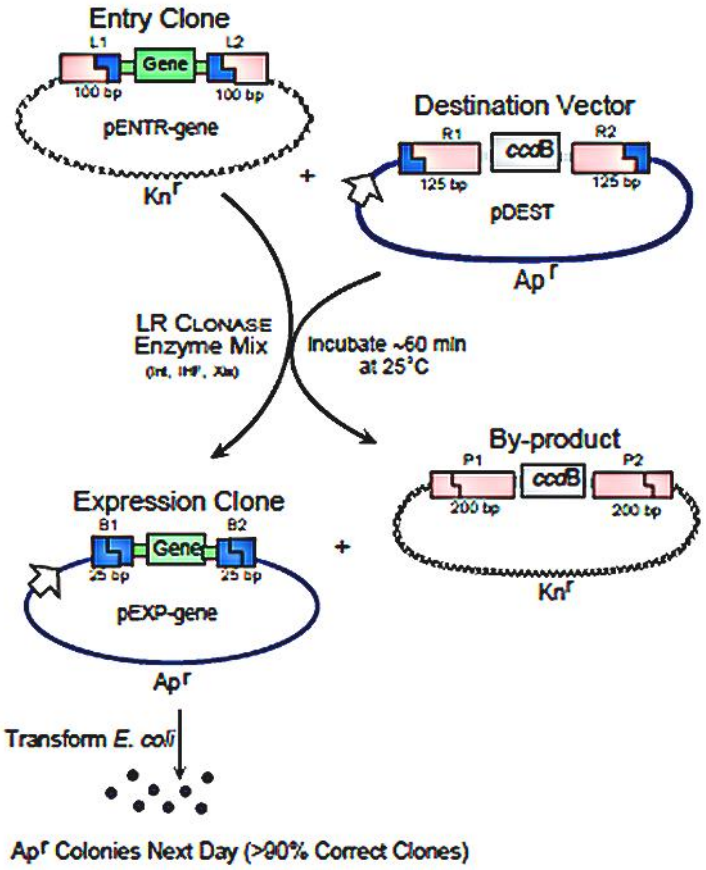

Figure 5 The Gateway® cloning kit containing the *ccdB* gene (grey). The recombination sites L1, L2 (red) and R1, R2 (blue) recombine with each other, L1 with R1 sites and L2 with R2 sites. The gene of interest (green) is transferred from the entry clone to the destination vector and the *ccdB* vice versa. The resulting expression clone is viable while the by-product kills its host cells. (Source: Gateway® cloning technology manual from Invitrogen™)

The *ccdB* gene is located in the destination vector in-between the recognition sites and *attR* recombination sites. When the recombination reaction is performed the two genetic elements, one from the entry clone containing the gene of interest and one from the destination plasmid containing the ccdB gene, switch their vectors. After the competent cells are transformed by adding the resulting plasmid mixture, the cells containing the desired gene will survive, while those with the old vector or the plasmid that was not successfully recombined are killed by the containing *ccdB* gene.

An *E. coli* strain carrying the antitoxin *ccdA* gene, counteracting the ccdB toxicity, is supplied with the kit in order to amplify and maintain the destination vectors.

1.3.2 TA systems as antibiotic targets (Alonso *et al.*, 2007; Hayes, 2003; Kedzierska *et al.*, 2007)

Pathogenic bacteria are subjected to a selective pressure by indiscriminate overuse of antibiotics. Bacteria develop resistance to one or to multiple antibiotics in order to cope with this. These genetic traits, adding antibiotic resistance, can be transferred among different bacterial species by means of DNA exchange processes like transformation, conjugation, transposition, or transduction. Some of the multiple antibiotic-resistance genes have been shown to be clustered into mobile DNA elements, called *integrons*, which are responsible for the recruiting of multiple smaller mobile gene cassettes that encode antibiotic-resistance genes. Thus, infectious diseases are in constant flux due to mechanisms of horizontal gene spread.

A number of possible solutions for tackling these problems are studied with synthesis of new antibiotics still representing the major path of research. The first step in this quest for new antibiotics is to identify adequate targets in the bacterial metabolism. As the number of determined macromolecular structures increases steadily, so is the amount of targets increased for possible docking experiments. Some other strategies include targeting of virulence factors, bacteriophage therapy, and use of genomics approaches.

Once the genomes of several pathogens had been sequenced, it was possible to search for possible targets for antimicrobials. This means to search for systems whose balance is important for the welfare of bacterial populations and to find a way to influence this balance. This is the case with TA systems that are thought to be involved in stress regulation. As these systems are wide spread in the eubacterial and achaebacterial world, the number of pathogens that could be targeted is vast. Many of the TA systems are even coded multiple times in some chromosomes (like RelBE in *E. coli*). Although, according to the present studies, triggering of the toxin does not kill the cells, it does lead to a reduction of the colony-forming units of a cell culture by several orders of magnitude. Thus the TA systems may lead research to a totally new avenue of natural compounds that could act as drugs counteracting the pathogenicity of bacterial strains by activating their “built in” autodestructs.

1.3.3 Toxins of TA systems targeting eukaryotic cells

There are many toxins produced by prokaryotic cells, some of them are toxic to eukaryotic cells as well and a few of these have found practical applications (Fitzgerald, 1996). These include Shigatoxin (STX) and its relatives from *S. dysenteriae* and certain strains of *E. coli* whose toxins are known to inhibit both prokaryotic and eukaryotic translation machinery (the so called ribosome-inactivating proteins or RIPs) (Yamamoto *et al.*, 2002). The effects of these prokaryotic toxins on eukaryotic cells are exhibited as morphological signs of apoptosis, which is described as membrane budding, chromatin condensation, fragmentation and reduction of cell volume and a typical ladder pattern of DNA degradation.

Two TA systems have been examined concerning their killing efficiency for eukaryotic cells, one representative from the RelE and one from MazF toxin groups. They are both organized as bicistronic operons, *kis/kid* – the representative of MazF toxin group, and *relB/relE* - the representative of the RelE toxin group. This bicistronic genetic organization is seldom found in eukaryotes but widely spread in prokaryotes (Alonso *et al.*, 2007).

TAs have the advantage to other prokaryotic toxins used to treat eukaryotic cells, that they have antagonists (antitoxins). Thereby a control of which cells are going to be targeted and which will be protected from the action of the toxin, can be achieved. This is, for example in human cancer cells, achieved by transcriptional regulators known to be inactivated, such as p53. The same can be applied to perform regulated knockouts with tissue-specific promoters (de la Cueva-Méndez *et al.*, 2003).

1.4 *Escherichia coli*

E. coli are well known gram-negative, rod-shaped, facultative anaerobic bacteria usually found in the lower intestine of warm-blooded animals but is also capable of survival for brief periods of time outside the body. The optimum environmental conditions for *E. coli* growth are temperatures from 10-45°C with an optimum at 37°C. *E. coli* cells propel themselves with flagella. The species was first discovered by Theodore Escherich in 1885 and became the model organism for studying prokaryotic life, especially for gram negative eubacteria. There are a number of *E. coli* strains but only a small number of them are dangerous to human health. Most of these disease causing strains account only for mild

diarrhoea but some, like strain O157:H7, can cause serious illness or death in the elderly, the very young or the immunocompromised. Since O157:H7 is abundant in human and animal faeces and not usually found in other niches, it was used since 1892 as indicator of faecal contamination. Today, although some strains do cause diseases in humans, *E. coli* has an important role in modern biotechnology allowing for the mass production of proteins in industrial fermentation processes. This is due to its long history of laboratory culture, simplicity of manipulation, and because it is a very versatile host for the production of heterologous proteins. However it cannot be used to produce the larger, complex proteins which contain multiple disulfide bonds or proteins that also require post-translational modification for activity.

1.5 *Methanococcus jannaschii*

Methanococcus jannaschii is an autotrophic hyperthermophilic organism shaped in irregular cocci, lacks a cell wall and belongs to the kingdom of Archaea. Extreme environments such as hypothermal vents at the bottom of the oceans in which water reaches boiling temperature or pressure is extremely high, is a preferred living space for *M. jannaschii* (Bult *et al.*, 1996). It was first isolated in 1982 in sediment samples taken from a "white smoker" chimney located 2600 m deep under the ocean in the East Pacific Rise, near the western coast of Mexico. This extreme habitat with temperatures from 48 -94°C with an optimum at 85°C, with pressure at 200atm, a pH of 5.2 -7, and salinity of 1.0 -5.0% NaCl makes this a usual member of the unusual kingdom of *archebacteria*. For comparison the surface temperatures of oceans is from 0 to 30°C, with pH between 7.5 -8.4, and salinity from 3.5 -3.8% NaCl. The cellular structure of *M. jannaschii* reveals that they possess a cell envelope that is made up of cytoplasmic membrane and a protein surface layer (S-layer) (Sleytr *et al.*, 2007) and their motility is achieved by utilizing polar bundles of flagella. *M. jannaschii* is strictly anaerobic and uses only carbon dioxide as its sole carbon source. Its main pathway for energy production is through methanogenesis, a process during which hydrogen is used as an energy source to reduce carbon dioxide to methane.

1.6 The Aim of This Work

In the beginning of this work (January 2005) there were no structures of RelE and RelB available and based only on the primary sequence information the superfamily categorization of the toxins was not reliable. It was known from other studies that RelE cleaves RNA but the three-dimensional structure of the protein was necessary in order to elucidate the way the RNA is bound and cleaved.

The structures of a RelE homologue from *Pyrococcus horikoschii* and the member of the RelE superfamily YoeB were determined in the course of 2005 (Kamada and Hanaoka, 2005; Takagi *et al.*, 2005) but the connection among them was still not fully realized.

The aim of this work was determination and function elucidation of the *Methanococcus jannaschi*i RelB and RelE proteins (*Mj*RelB and *Mj*RelE) and their complex (*Mj*RelBE). The work on the structure of *Escherichia coli* RelB and RelE (*Ec*RelB and *Ec*RelE) was pursued in parallel.

2 Materials and Methods

All chemicals used in this work were of analytical grade, if not stated otherwise. Kits and equipment were used according to the producer's manuals or with modified protocols as described in this section.

2.1 Chemicals, Enzymes and Kits

Chemicals used

Agarose, low tempetature melting (LMP)	Invitrogen
Ammonium sulphate $(NH_4)_2SO_4$	Merck
Bromphenol blue	Serva
Dinatrium-Malonate	Merck
DNA ladder	New England Biolabs (NEB)
Glutaraldehyde	Sigma-Aldrich, Merck
IPTG (isopropyl-β-D-thiogalactopyranoside)	Roche
Polyethylenglykol (M_W = 200 – 20.000 Da)	Sigma-Aldrich, Merck
Polyethyleneimine	Serva GmbH
PMSF	Serva
Protease-Inhibitor-Cocktail Complete	Roche
Protein ladder Broad Range 2-212 kDa	NEB
Sterile filter (0,22 µm and 0,45 µm)	Millipore
6-Amino-Capronic acid	Sigma-Aldrich, Merck
2-Methyl-2,4-pentanediol	Merck

Enzymes and proteins

EcoRI	NEB
NdeI	NEB
BamHI	NEB
XhoI	NEB
NcoI	NEB
Vent-DNA-Polymerase	NEB
Pfu-DNA-Polymerase	Stratagene
T4-Ligase	NEB
DNaseI	Merck
Thrombin	Sigma-Aldrich
Bovin-sreum-albumin (BSA)	Sigma-Aldrich

Kits

Qiaprep Spin Plasmid Miniprep kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen
Quick Change	Qiagen

Expression vektors and competent *E. coli* cells

E. coli BL21(DE3)	Novagen

E. coli Rosetta(DE3)	Novagen
E. coli ER2566	NEB
E. coli DH5*a*	NEB
pET28a	Novagen
pET21a	Novagen
pTrc99a	Stratagen

Cell medium

LB-Medium	Caseinpepton	10 g/l
	Yeast extrakt	5 g/l
	NaCl	5 g/l
	pH 7,5 (RT, 1 M NaOH)	

LB-Medium with antibiotics

Ampicillin	100 µg/ml
Chloramphenicol	20 µg/ml
Kanamycin	40 µg/ml

LB-Medium with Agar (LB-Agar)

Agar-Agar	15 g/l

Chromatography media and instrumentation

Chromatogrphy system FPLC	GE Healthcare (previously Amersham Pharmacia)
Chromatography system ÄKTA	GE Healthcare (previously Amersham Pharmacia)
Chromatography system BIOCAD	Perceptive Biosystems
Fraktion collector SF2120	Applied Biosystems

Columns

HisTrap 1ml	GE Healthcare (previously Amersham Pharmacia)
Talon beads	Clontech, TakaraBio Inc.
HE20 Heparin	POROS, Applied Biosystems
HS20	POROS, Applied Biosystems
HQ20	POROS, Applied Biosystems
PL-SAX (-$N(CH_3)_3^+$)	Polymer Laboratories
PL-SCX (-SO_3^-)	Polymer Laboratories
HiLoad Superdex75 16/60	GE Healthcare (previously Amersham Pharmacia)

Desalting

HiPrepDesalting	GE Healthcare
ZipTip® C_{18} desalting for MALDI-TOF	Millipore

Concentrators

Amicon Ultra (cutoff: 3, 10, 30 kDa); 5 and 15 mL Volume	Millipore, Schwalbach

Protein purifications

Purification of *M. jannaschii* proteins:

RelBE complex

Liquid Chromatography, Ion exchange: (HE20 Heparin, affinity and cation exchange column)

Running Buffer (A): 50 mM NaCl, 50 mM Tris pH 8.0 at 4°C

Elution Buffer (B): 2 M NaCl, 50 mM Tris pH 8.0 at 4°C

Gelfiltration (HiLoad Superdex75 16/60)

Running Buffer (C): 500 mM NaCl, 20 mM Tris pH 8.0 at 4°C

Running Buffer (C2): 500 mM NaCl, 20 mM HEPES pH 7.5 at 4°C

RelE(R62S)his
Ni-Affinity-Chromatography (HisTrap – Amersham Biosciences)
Running Buffer (D): 20 mM Imidazol, 500 mM NaCl, 50 mM Tris pH 8.0 at 4°C
Elution Buffer (E): 500 mM Imidazol, 500 mM NaCl, 50 mM Tris pH 8.0 at 4°C

Gelfiltration (HiLoad Superdex75 16/60)
Running Buffer (C): 500 mM NaCl, 20 mM Tris pH 8.0 at 4°C

Purification of *E. coli* proteins:

Ammonium sulphate precipitation
Precipitation buffer stock 4 M $(NH_4)_2SO_4$, 50 mM NaCl, 50 mM Tris pH 8.0 at 4°C

Dialysis
Dialysis buffer 50 mM NaCl, 50 mM Tris pH 8.0 at 4°C

Liquid Chromatography, Ion exchange: (HS20 cation exchange column)
Running Buffer (F): 50 mM NaCl, 50 mM Tris pH 8.0 at 4°C
Elution Buffer (G): 2 M NaCl, 50 mM Tris pH 8.0 at 4°C

Gelfiltration (HiLoad Superdex75 16/60)
Running Buffer (H): 100 mM NaCl, 20 mM Tris pH 8.0 at 4°C

Denaturating conditions
Denaturating Buffer (I): 8 M Urea, 100 mM NaCl, 20 mM Tris-HCl pH 8.0 at 4°C
Elution buffer (J): 100 mM Imidazole, 100 mM NaCl, 20 mM Tris-HCl pH 8.0 at 4°C

Sample Buffers and staining solution for Electrophoresis:
DNA-Samplebuffer (6x):

0.25	%	Bromphenol blue (w/v)
0.25	%	Xylen Cyanol (w/v)
40	%	D(+)-Saccharose (w/v)

SDS-Samplebuffer (5x):

50	m*M*	Tris-HCl pH 7.0
10	%	Laurylsulfate (w/v)
0.1	%	Bromphenol blue (w/v)
1	%	2-Mercaptoethanol (v/v)
14	%	1,4-Dithiothreitol (w/v)
10	%	Glycerol (v/v)

Stainingsolution for SDS-PAGE gels:

0.5	%	Coomassie Brilliant Blue G250 (w/v)
6.25	%	Ethanol, techn. (v/v)
4.25	%	Perchloracid (v/v)

Crystalisation

12 well plates, Modell Groningen	Nelipak bv (Niederlande)
96 well plates	Greiner BioOne
Plastic film	3M (Hampton Research)
Mounted Cryo loops	Hampton Research
Cover plates 22 × 22 mm, Thickness 2	Laborhandel
Glass capillaries, ∅ = 0,3; 0,5; 0,7; 1,0; 1,5 mm	Glas Müller (Berlin)

Quartz capillaries, ∅ = 0,3; 0,5; 0,7; 1,0; 1,5 mm	Glas Müller (Berlin)
Crystal-Screen I	Hampton Research
Crystal-Screen II	Hampton Research
Index-Screen	Hampton Research
Natrix-Screen	Hampton Research
Salt RX	Hampton Research
PEG/Ion-Screen	Hampton Research
Quick-Screen	Hampton Research
MPD-Grid-Screen	Hampton Research
Ammonium Sulfat-Grid-Screen	Hampton Research
PEG/LiCl-Grid-Screen	Hampton Research
NaCl-Grid-Screen	Hampton Research
Malonat-Grid-Screen	Hampton Research
PEG6000-Grid-Screen	Hampton Research
Additive-Screens I-III	Hampton Research
Detergent-Screen I	Hampton Research

Instruments

Agarose-Gelelektrophorese chamber QS710	Kodak Biomax
Thermocycler Tpersonal, Trio Thermoblock	Biometra
Elektroporator Easyject Prima	EquiBio
Binokularmicroscope SZ60	Olympus
PAGE-Elektrophoresis	Biometra
PhastGel Electrophoresis system	Pharmacia (GE Healthcare)
French® Pressure Cell Press	Sim Aminco
Cell homogenizer EmulsiFlex-C5	Avestin
Mar345 Image Plate	mar-Research
Röntgengenerator FR-571 with rotating Anode	Enraf-Nonius
Cryosytem	Oxford
DLS Laser Spectroscatter 201	RiNA GmbH
CD Spektrometer JASCO-J600	Jasco Labor und Datentechnik
UV-VIS-Spektrometer	Shimadzu
Small volume RNA/DNA and Protein UV-VIS photometer, GeneQuant	GE Healthcare (previously Amersham Pharmacia)

2.2 Molecular Biology Methods

2.2.1 Transformation and Stock cultures

The incorporation of DNA into competent cells (competent for acquisition of plasmid DNA) is performed through heat shock (for chemically competent cells) or through electroporation (Potter, 1988) (for electrocompetent cells). The necessary cells are prepared as described in the protocols supplied by the manufacturer (EquiBio) and stored at -80°C till needed. The transformation of the cells is performed as recommended by the manufacturer (EquiBio).

2.2.2 Plasmid Preparation and Sequencing

In order to check the cloned genes for the correct sequence and subsequently transform new cells with the engineered plasmids, plasmid DNA had to be isolated and sequenced. Plasmid DNA was extracted from Bacteria using QIAquick® Plasmid Miniprep-DNA kits from Qiagen, as described in the manufacturer protocols. Plasmid sequencing was carried out by GATC company (Konstanz).

The method of plasmid isolation employs modified alkaline method of cell lysis and a silica-based fiber matrix that binds DNA in the presence of a chaotropic salt. The sample is first incubated with a ribonuclease and cell lysis solution, and then in a DNA-binding solution. The plasmid DNA binds to the fiber matrix while the contaminants are washed from the matrix with a wash buffer. The purified plasmid DNA is eluted from the fiber matrix with a low-ionic-strength buffer and captured in a microcentrifuge tube. The result is purified plasmid DNA that is ready for restriction, ligation or sequencing reactions.

2.2.3 Primer design for cloning and mutagenesis

Primers are short single stranded DNA fragments that bind to their complementary sequence on the target DNA strands, either partially or in full length. If the insertion of a point mutation is necessary, a primer can be designed with one or more nucleotide mismatches in the middle of their target sequence. This is done in a way that the

primer still binds with good specificity and strength, but during the amplification of the DNA sequence the engineered mutation is inserted. Another type of designs with partial binding to the target DNA is when it is intended to add new digestion sites. For this purpose the primer binds only with one part (3' or 5') to the DNA while the mismatching part, that has the additional DNA restriction enzyme recognition sequence, "hangs" lose. On the other hand, when it is intended to specify the region of DNA that is to be amplified, the primers would be designed to bind in full length to the target DNA. When designing primers, there are a few rules to be followed. They are usually designed to be specific for the sequence so to avoid unspecific priming they should be longer than 16 nucleotides (mathematically this makes coincidental overlaps less probable as for this length there are $4^{16}= 4*10^9$ combinations of nucleotide arrangement). The melting temperature of the primers should be between 50 and 70°C, and they should not have complementary sequences that could bind to each other forming hairpins. There are programs available to simplify the primer design task by calculating the melting temperature of each primer and search for regions that could form hairpins. The program "Oligonucleotide Properties Calculator" (http://www.basic.northwestern.edu/biotools/oligocalc.html) (Kibbe, 2007) was used for this purpouse.

2.2.4 Digestion of DNA with Restriction Endonucleases

For the purpose of trimming DNA fragments or cutting them out of longer DNA (such as plasmids or partial constructs), commercially available endonucleases from New England Biolabs (NEB) were used. Each endonuclease has its specific target sequence for DNA digestion and a buffer in which it is most effective. The manufacturer (NEB) supplies the specific buffers and provides the tables with descriptions of the activity of the enzymes in each of these buffers. As most of the supplied restriction endonucleases are sufficiently active in more than one buffer, a double digest (digesting both sides of a DNA fragment with different restriction endonuclease in one reaction) were also used.

2.2.5 Extraction of DNA from Agarose Gels and Reaction Mixtures

In order to separate amplified regions of DNA from primers or unspecific bands, agarose gels with ethidium bromide were cast and DNA samples loaded. The DNA fragments

in the gel separate by size when current is applied and the bands, containing the DNA of same sizes, are visualized under UV-light. Ethidium bromide intercalates in the DNA and upon transfer of the ethidium bromide in the hydrophobic environment of the base pairs, its fluorescence is no longer quenched by water and the absorbed UV light is re-emitted (by fluorescence) in the visible part of the spectrum. The bands of the expected sizes can then be cut out of the gel and DNA extracted with QIAquick® Gel-Extraction-Kit from Qiagen or with corresponding kits from Promega or Stratagene, as described in the standard protocols supplied with the corresponding kit. The base principal of these kits is to dissolve the agarose, buffer the solution to the necessary pH for binding DNA onto a provided disposable ion exchange column, wash the bound DNA and elute it in the appropriate buffer (usually 10 mM Tris pH 8.0). The same kits have been used, with minor adjustments, also for direct purification of DNA from digestion reactions.

2.2.6 Cloning of the *relB* and *relE* genes

Professor Gerdes kindly supplied the starting plasmids containing the *Methanococcus jannaschii* (strain DSM2661) and *Escherichia coli* (strain K12) chromosomally encoded *relB* and *relE* genes, *pSC2526MJ#2* and *pSC2524HE* respectively. *E. coli* strains DH5*a* and TOP10 were used for subcloning and ER2566, BL21 and Rosetta™ for protein expression.

As the initial constructs did not yield the expected quantity of the products and some variations in the design where necessary, new constructs have been prepared for both the *Escherichia coli* and *Methanococcus jannaschii* genes.

The *pET21relBE_MJ* [*] was constructed by adding a NdeI restriction site in front of the *relB* gene and HindIII at the end of the relE gene. The fragment was cloned in *pET21b* without any tags but with persistent spontaneous mutations at the position RelE Arg62, always to Ser or Lys. To prove that this spontaneous mutation inactivates the toxin a separate cloning of the toxin was performed by adding NdeI restriction site at the 5' end of the mutated *relE* gene. The cloning was performed by PCR, the primer generating the NdeI restriction site at the 5' end of *relE* and the standard T7rev (Standard vector primer) were used for the amplification of the required DNA fragment. After the PCR products were separated on 1%

[*] The *pET21relBE_MJ* construct was made by Caterina Farnleitner

agarose gel, extracted and digested by NdeI and HindIII restriction enzymes, they were ligated in the *pET28a* vector that was previously treated with the same restriction enzymes. The so obtained construct was named *pET28relE(R62S)*. The transcription of the cloned genes in the expression vectors (like the pET vectors) is dependent on the T7 RNA polymerase. As the T7 RNA polymerase is not coded in the vector, it has to be supplied by the host cell. Transformation was generally first performed in *E. coli* cloning strains, like DH5α, as these strains do not have the T7 RNA polymerase integrated in their chromosomes so it is easier to establish a plasmid in them. After successful cloning and verification of the gene sequence by DNA sequencing, the plasmid was isolated and transformed to a new host, containing a chromosomal copy of the T7 RNA polymerase gene, for expression. In case of *M. jannaschii relB* and *relE* genes the *E. coli* Rosetta™ strain (carrying the pRARE helper plasmid) was the strain of choice as it can synthesize tRNAs for codons rarely used in *E. coli*, but are frequently found in genes originated from archeons and eukaryotes. In this way the non-toxicity of the *relE* gene carrying the R62S mutation was verified.

As the R62S mutation was present even in the original pSC2526MJ#2 gene, an effort was made to "cure" the mutation. Two methods were used, the Quick Change® kit from Qiagen and the sequential primer extension PCR. In the first method the manufacturer protocol was followed with one mutation primer *MJrelE(S62R)_demut_for*. In the second method two primers were constructed, *MJrelE(S62R)_demut_for* and *MJrelE(S62R)_demut_rev*, which were used in two separate PCRs with the *pET21relBE MJ* construct as template. The purified products of the first two PCRs were used in the subsequent PCR together with two flanking standard vector primers. The so generated DNA fragment containing the antidote *relB* and demutated toxin *relE*, was cut by the NdeI and HindIII restriction enzymes and inserted in the pET21 vector again. Both of these restriction sites were already present in the original plasmid sequence of *pET21relBE_MJ* and it was not necessary to engineer them in the sequence.

The construct for *E. coli relBE* genes expression without the His-tag, *pTrc99relBE_EC*, was constructed by cloning the *E. coli relB* and *relE* genes from *pSC2524HE* to the *pTrc99* vector. With the used primers an NcoI restriction site was added to the 5' end of *relB* and HindIII to 3' end of *relE*. The obtained construct was first transformed in *E. coli* DH5α host strain from which they were later isolated, sequenced and, upon sequence verification by DNA sequencing, transformed in ER2566 *E. coli* host strain for

expression. In the same manner a construct of relB alone was constructed by insertion of the gene with NcoI and XhoI restriction sites at its ends in the pET21 vector.

In the following Table 4 the used primers are listed:

Table 4 List of primers. Mutaded bases (red) and inserted restriction sites (blue) are colour labelled.

Name	***Sequence***	***Description***
relE(S62R) demut for.	gg att aga gta gga aat tat agt ata ggt att gag gtt aat gga g	5' primer for demut of relE(R62S) mutant to wild type (45bp G/C33.3% Tm=74.7, Ta=69.7)
relE(S62R) demut rev.	c tcc att aac ctc aat acc tat act ata att tcc tac tct aat cc	3' primer for demut of relE(R62S) mutant to wild type (45bp G/C33.3% Tm=74.7, Ta=69.7)
relE NdeI for.	cga caa ttc cat atg aaa gtg tta ttt gc	5' primer for inserting NdeI restriction site at 5' end of MJ relE (29bp G/C34% Tm=56, Ta=?)
ECrB NcoI for.	ttt ttt tcc atg ggt agc att aac ctg cgt att g	5' primer for inserting NcoI restriction site at 5' end of EC relB (34bp G/C38% Ta=61)
ECrB XhoI rev.	aaa aaa ctc gag ttc atc cag cgt cac acg	3' primer for inserting XhoI restriction site at 3' end of EC relB, no stop (30bp G/C67% Ta=61)
ECrB NcoI for.old	cgc gga tcc atg ggt agc att aac	5' primer for inserting NcoI restriction site at 5' end of EC relB
ECrE HindIII rev.	ccc aag ctt tca gag aat gcg ttt	3' primer for inserting HindIII restriction site at 3' end of EC relE

2.3 Protein preparation

2.3.1 Protein expression in *E. coli*

The genes inserted in the vectors are under the control of the *T7* promoter that regulates the transcription of the downstream genes by the T7 polymerase. In the pET vector system there is an additional *lac* operator where the lacI repressor binds. The lacI repressor is therefore inhibiting both the transcription of the T7 RNA polymerase on the host chromosome and the transcription of the genes inserted in the vector. Addition of IPTG to a growing culture of the expression clones induces T7 RNA polymerase, which in turn can transcribe the

target DNA in the plasmid. In the same time (in the pET vector system) the lacI repressor releases the *lac* operator in front of the insert DNA in the plasmid.

The used constructs require for growth of cell cultures and for protein expression almost the same conditions. The only differences were due to different selection antibiotic requirements. The vectors *pTrc99a* and *pET21a* carry ampicillin resistance while the *pET28a* carries kanamycine resistance. In case when the Rosetta™ cells were used an additional selection marker, chloramphenicol, had to be added to ensure the maintenance of the *pRARE* helper plasmid that has to be maintained in the Rosetta™ cells.

Overnight cultures of 100 ml LB were inoculated from plates with single cell colonies, supplemented with selection antibiotic and grown at 37°C.

The bacterial cultures were grown from 1:100 diluted overnight culture in a shaker at 37 °C in LB medium supplemented with the required selection antibiotic (depending on the construct). At the optical density of 0.6 at 600 nm gene expression was induced by 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The culture was centrifuged (5,000 g, 4°C, 10 min) 3.5 h after induction; and cells were resuspended in the lysis buffer containing 50 mM NaCl, 50 mM Tris-HCl, pH 8.0 and lysed or stored for later at -20°C.

2.3.2 Cell Disruption

The pelleted cells were diluted with buffer A in three times the volume of the approximate cell weight and disrupted with a French®Press at 12,000 psi at 4°C or with the cell homogenizer. For the purification of the *E.coli* RelB, RelE or RelBE, a Protease-Inhibitor-Cocktail COMPLETE tablet (per 20 g of cell pellet) was added in the resuspended cells before cell disruption. The cell debris were pelleted by centrifugation in a cooled ultracentrifuge at 90 000 g (30000 rpm in an UC TI60 rotor) and the clear lysate was carefully removed. As the cell lysate of the *M. jannaschii* RelBE contained a lot of DNA, which posed a problem for sterile filtering, a spatula tip of dry DNaseI was added and incubated for 20 min at 4°C. The obtained cell lysates were sterile filtered through 0.22 µm filters.

2.3.3 Purification of Proteins

2.3.3.1 Purification Protocol for *Methanococcus jannaschii* RelBE

The filtered lysate was loaded on a heparin column, and using the ÄKTA chromatography system a gradient was run from the starting buffer A to 45% buffer B and then with a step gradient of 55% buffer B (1.1 M NaCl) the protein was eluted as a concentrated peak. The so obtained protein is already more that 90% pure so only a final polishing step was necessary to reach the required purity of >95%. This was done, after the protein was concentrated to 1ml total volume, by Gel Filtration on a S75 column in buffer C. For the final purifications, after the crystallization protocol was refined, the Gel Filtration buffer was changed to 500 mM NaCl, 20 mM HEPES pH 7.5 (buffer C2).

The RelE(R62S)his N-terminal His_6-taged construct was purified with a HisTrap column and starting buffer D and elution buffer E using a step gradient. The protein eluted at 30% buffer E (300 mM imidazole) and the His_6-tag was removed overnight by addition of 0.2U of thrombin per milligram of protein. After a concentration step to 1ml total volume the protein was applied to the S75 gel filtration column in buffer C.

2.3.3.2 Purification Protocol for *Echerichia coli* RelBE

The first step in the purification of the *E.coli* RelBE construct without the His_6-tag (construct *pTrc99aRelBE_EC*) is an amonium sulphate (AS) precipitation. In the cleared lysate the Precipitation buffer (4 M AS, 50 mM Tris pH 8.0) was slowly added so that the final protein solution had an AS concentration of 1.8M. The precipitated proteins were pelleted in a HB4 Rotor at 10.000 g for 30 min at 4°C. The soluble fraction was dialyzed overnight in dialysis buffer A (50 mM NaCl, 50 mM Tris pH 8.0) and sterile filtered through 0.22 µm filters. The dialyzed sample was further purified by ion exchange over a HS20 column with starting buffer F and elution buffer G. The eluted protein was concentrated and loaded on a S75 gel filtration column with buffer H.

The *E.coli* RelBE construct *pSC2524HE*, having a His_6-tag, was purified over a HisTrap and S75 columns when intended for crystallization. For activity essays the purification was changed. Filtered soluble cell extract was incubated with Talon for 12 h at 4°C on a laboratory rocker. A column was washed with 10 column volumes of denaturating

buffer I (8 M Urea, 100 mM NaCl, 20 mM Tris-HCl pH 8.0 at 4°C). Denaturant washing removes most of RelB protein leaving histidine tagged RelE bound to Talon. The RelE protein was eluted from the column with elution buffer J (100 mM imidazole, 100 mM NaCl, 20 mM Tris-HCl pH 8.0 at 4°C). RelE was dialyzed against buffer C (100 mM NaCl, 20 mM Tris-HCl pH 8.0 at 4 °C) and concentrated to 1 mg/ml.

2.3.4 Concentrating the Proteins

The proteins were concentrated with ultrafiltration concentrators with a cut- off at 5 kDa (for RelE) or 15 kDa (for the RelBE complex).

2.4 Protein Properties Measurements

2.4.1 Concentration Measurements

The protein concentration was measured during the purification with Bradford method (Bradford, 1976), and the final concentrations of the pure proteins were determined from the UV absorption at 280 nm using the molar absorbance of 2560, 7680, 10240, 15930 $(cmM)^{-1}$ for *Mj*RelB, *Mj*RelE, *Mj*RelBE and *Ec*RelBE, respectively. Molar extinction coefficients of proteins considering their amino acid compositions were calculated using "ProtParam" (http://www.expasy.ch/tools/protparam.html).

2.4.2 Polyacrylamide Gel Elektrophorese (PAGE) and Isoelectric Focusing (IEF)

To assess the contents and purity of a protein sample SDS-PAGE (Laemmli, 1970) was used according to standard protocols (Sambrook *et al.*, 1989) and stained with silver or coomassie blue, depending on the required sensitivity. As the studied proteins are small (between 5 and 12 kDa) usually the 17% acrylamide gels were the right percentage to

choose. With the *M. jannaschii* RelBE complex the problem of RelB and RelE bands smearing and overlapping occurred, as the resolution of the standard SDS-PAGE is poor for small sized proteins. Therefore a modified protocol was chosen that uses tricine instead of tris and has an additional separation layer (see Appendix 7.1 for details). The SDS-PAGE gels were stained by standard coomassie blue or silver staining using the Price lab protocol (http://bcaws.biochem.uiowa.edu/price/pricelab/silver_staining.htm). To further characterize the proteins and also check their purity, as the SDS-PAGE only separates the proteins according to their size, IEF was used. The IEF separates proteins on basis of their isoelectric point, and consequently unmasks a possible impurity of the same size. Further, IEF gives the correct isoelectric point of the proteins in native form, which may differ from the theoretical value. For IEF experiments the PhastGel-System and corresponding gels, all from Amersham Biosciences (now GE Healthcare) were used.

2.4.3 MALDI-TOF- Mass Spectrometry

The MALDI (**M**atrix-**A**ssisted **L**aser **D**esorption/**I**onization)- TOF (**T**ime **O**f **F**light)- MS (**M**ass **S**pectrometry) is a precise instrumentation for mass determination of molecules. A matrix (picolinic acid in this case) is added to the sample and co- crystallized with it on the target plate of the instrument. The matrix serves to encourage ionization of the analyte and to rapidly and efficiently absorb the laser irradiation so that the analyte (protein, DNA/RNA…) is ionized and evaporated into the electric acceleration field. The MALDI-TOF instrument uses pulses of laser light to vaporize the protein/matrix in a process known as "desorption" and during this process some molecules become ionized through protonation and can be accelerated in the electric field towards the detector. The laser pulse can also fragment the molecule into a variety of charged and neutral particles (which can lead to spurious bands or artefacts in the resulting spectrum). As the molecules are accelerated to common kinetic energies in the mass analyzer, the heavier molecules travel slower than the light ones and the double charged ones travel twice as fast as the single charged. The resulting spectrum is a function of the mass through charge ratio, and the higher ionized species are recognized by their apparent smaller size in the spectrum (with an apparent size difference being a whole number compared to the single charged molecule).

For MALDI-TOF-MS analysis it is essential to desalt the sample as salt interferes with the measurement. The desalting is done with ZipTipps.

With this method it is possible to determine the actual size of the protein that is prepared for crystallization as some amino acids may be missing at the C or N-termini. Further it can be used to verify whether some modification on the protein was successful (as selenium methionine exchange).

The MALDI-TOF-MS analysis was kindly done by Dr. Peter Franke (Institut für Biochemie, FU Berlin).

2.4.4 Dynamic Light Scattering (DLS)

DLS measures the hydrodynamic radii of small particles and molecules. The measurement is based on the time a particle needs to travel through a monitored volume. On the basis of the temperature and the agility of the particle the mass and size of the particle can be elucidated. The detection of the particle is performed by a photo diode capturing the laser light reflected by the particles in a normal angle towards the source.

For crystallographic applications it is the relative spread of particle sizes that is of interest. It has been shown that single sized (monomodal) samples tend to crystallize more readily than polydisperse samples. Further it is possible to follow the formation of oligomers as the buffer composition is changed.

The samples are prepared by sterile filtration or centrifugation, and 20 µl of concentrated protein is used for each measurement. The used concentration is usually over 1 mg/ml and up to 10 mg/ml, depending on the count rates this concentration produces on the instrument detector.

2.4.5 Circular Dichroismus – Spectroscopy (CD)

The CD-spectroscopy is a method of choice for assesing the secondary structure contents of polypeptides as well for following the changes in protein fold with temperature, buffer composition, pH or cofactor addition.

The CD spectra were recorded with JASCO-J600 CD Spektrometer (JASCO Labor- und Datentechnik) in the range of 190 nm – 240 nm. The buffer for *M. jannaschii*

RelBE protein complex was 500 mM NaCl, 20 mM sodium phosphate pH 7.5 and for *E. coli* RelBE protein complex 100 mM NaCl, 20 mM sodium phosphate pH 7.5. The protein concentration was adjusted so that the measurement was performed in the optimal sensitivity range for the instrument. All measurements were performed in 0.02 cm thermostated cuvettes connected to a thermostat allowing measurements from 15°C to 88°C. The presented mean residue ellipticities ([Θ] in deg cm^2 $dmol^{-1}$) were obtained from raw data (ellipticities in millidegrees) by subtracting the corresponding spectra of the buffer solution and taking into account the mass concentration (*c* in mg/ml), molecular mass (*M*w in mg/dmol), number of amino acid residues (*N*), and optical path length (*d* in cm) through the relation $[\Theta]_{mwr}=\Theta$ *M*w/(*N c d*). The used molecular weights and residue numbers are 16580 Da, 138 AA for *Mj*RelBE and 18194 Da, 157 AA for *Ec*RelBE.

The secondary structure content at 20°C was calculated by the program K2D (Andrade *et al.*, 1993) and changes in protein folding on various temperatures was monitored.

2.4.6 Analytical Ultracentrifugation

This method is used for the determination of concentration dependent association of biomolecules in solution. As the RelBE complex is forming a heterotetramer it was interesting to determine the dissociation constants of the heterodimer – heterodimer or the RelE – RelB binding.

The sedimentation equilibrium experiments and data analysis was kindly performed by Prof. J. Behlke (MDC, Berlin - Buch).

2.4.7 Endoribonuclease Activity Tests

To test in vitro endoribonuclease activity of the RelE proteins two short RNA fragments were synthesized, the *RNAcut*-UACAUAGUGAGGCGCUAAGCAGAGAUUUAAGCAA containing the preferable cutting sites (UAG, UAA, UGA) (Christensen and Gerdes, 2003; Pedersen *et al.*, 2003) and *RNAn*-UACAUACUCAGGCUCCAAGCAGAGAUUUCAGCAA with less specific sites (UCG, CAG) (Christensen and Gerdes, 2003; Pedersen *et al.*, 2003). Stock solutions of the sample

RNA were aliquoted and frozen at -20°C and none of the aliquots were used twice as traces of contaminant RNases would slowly degrade the samples. The buffers were prepared with diethyl dicarbonate (DEPC) and sterilized water in order to avoid contaminant RNases. All pipetting was done with special pipette tips with built in micro-filters to prevent contamination through the pipette. Contamination by other RNases would show as false positives and therefore the ssRNA stock was aliquoted upon arrival and these aliquots were used as source of samples only for one experiment each.

Samples were taken in different time intervals and loaded on an acrylamide gel. The gels were stained with ethidium bromide to visualize the RNA fragments.

2.4.8 Glutaraldehyde Cross- linking

The protein sample was incubated in buffer D with 0.01% and 0.02% glutaraldehyde and samples of the reaction mixture were taken at different time intervals from 0 to 30 minutes. The reaction is terminated by addition of 1 M Tris-HCl, pH 8.0. The samples were analyzed with SDS-PAGE.

2.5 Crystallization

The vapour diffusion technique was the main method of crystallization. For preliminary crystallization trials the Greiner 96 well plates for sitting drop vapour diffusion in combination with Hampton crystal screens were used. The reservoir was filled with 100 µl of the screening solution while the crystallization drop was mixed from 1.1 µl protein and 1.1 µl reservoir solution. There are three indentations in each of the 96 compartments of the plate preset for the crystallization drop so up to three protein concentrations could be screened simultaneously if enough material was purified. Generally two indentations, separated by an empty one, were used to avoid unwanted mixing of the drops. The plate was sealed with plastic film (3M Hampton Research) and stored at 4 or 18°C. For fine screening of crystallization conditions the 12 well plates for hanging drop vapour diffusion were used. The reservoir was filed with 600-800 µl of screening solution (for screens that were meant to be brought to the synchrotron as little as 300 µl) and the crystallization drop was mixed from 1-3

µl of protein and 1-3 µl of reservoir solution. The crystallization drop was placed on a siliconised glass plate and placed over the reservoir with the drop facing downwards. The glass plate can accommodate up to 4 smaller drops so varying the protein concentration over one reservoir is also possible in this configuration. The plates were prepared and stored at 4 or at 18°C. All plates were inspected immediately after preparation. This was done in order to identify dust particles, before they might be confused for real crystals, and to see how the protein reacts upon the addition of precipitants. These inspections were performed regularly, every day for the first week and after that weekly or monthly. All conditions and changes in the crystallization drops were logged in special forms.

2.5.1 Seeding

In order to obtain better diffracting crystals a variation of the micro seeding method was utilized. The micro seeding method separates the processes of nucleation and crystal growth as the best conditions for each of these processes may differ. In this method formed fresh crystals are crushed with a small metal tool and then the seeds (microscopic debris of the crystal) transferred to a pre- equilibrated solution with precipitant concentration just below supersaturation.

The variation of the method consisted in preparing a stock of seeds by diluting the whole drop with the crushed crystal in 300 µl of a stabilizing solution (with more precipitant than in the original drop) from which the pre- equilibrated drops were seeded. The seeding was performed with a horse tail hair through multiple drops (a so called *streak seeding*). The hair would be dipped in the solution with the seeds and then, without dipping it again in the seed solution, streaked through a few identical drops in order to obtain a dilution series of seeds in the subsequent drops. The new crystals would form mainly around the line where the hair passed through the drop, and in each of the subsequently seeded drop in the row the number of the crystals would decrease but their size, to a smaller degree, increased.

2.5.2 Crystallization of *M. jannaschii Mj*RelBE Protein Complex

Initial crystals were grown in hanging drop method by adding 2.5 µl of protein sample in buffer C and 2.5 µl of reservoir solution of the following composition: 10% 2-

methyl-2,4-pentanediol (MPD), 0.1 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5. The drop was equilibrated over 500 µl of reservoir solution, and after two days single, irregularly shaped crystals grew to a size of 0.4-0.9 mm. For seeding experiments, a single crystal was ground with a spatula in the crystallization drop. The debris were diluted with 200 µl of stabilizing solution consisting of 22% MPD, 250 mM NaCl, 0.1 M HEPES pH 7.5, vortexed and further used as seed stock. The seed stock was stable for more than two months and could be used for later seeding. Final crystals were prepared by streak seeding in a hanging drop over a 400 µl reservoir consisting of 20% MPD, 0.1 M Bis-Tris pH 5.5. The crystallization drop was prepared with 2.5 µl protein solution and 2.5 µl reservoir solution. Streak-seeding was performed by dipping a horse tail hair in the seed stock and streaking with it in a row through three prepared drops in order to obtain a dilution series of the seeding. The best crystals were obtained after two days in the last seeded drop. The obtained crystals showed a regular rhombohedral habitus and were directly flash frozen without further cryoprotectant.

2.5.3 Heavy Atom Derivatives

Heavy atom derivatisation was done by co-crystallization, or by soaking with 5 mM heavy atom reagent for two to 24 hours. Heavy atom stock solutions were prepared by dissolving the solid compounds in water (100 mM stock) while the dilutions (10, 25 and 50 mM) were prepared by diluting the stock solution in the protein buffer (buffer C). The co-crystallization was performed by pipetting on a siliconized glass plate 1 µl of the heavy atom compound (from 10, 25 or 50 mM dilution), 2 µl of protein solution and 2 µl of the well solution, yielding a 5 µl crystallization drop. Heavy atom soaking was performed as follows. The crystals were transferred from their crystallization drop into a fresh drop, on a fresh siliconized glass plate, containing the same reservoir solution and the added heavy atom compound. The glass plate was then turned over a sealed well with the same composition of the reservoir solution as in the original crystallization condition of the referring crystal.

2.5.4 Crystal Mounting

The obtained crystals were mounted either in a quartz capillary (for room temperature X-ray data acquisition) or frozen in a nylon loop (for cryo X-ray data

acquisition). For capillary mounting the crystals are carefully sucked in the capillary together with their mother liquid and positioned at 2/3 of the capillary. The excess liquid is carefully removed with thin filter paper to prevent the crystal from swirling around in the liquid during data collection. On one side of the capillary a drop of the mother liquid is sucked in and positioned preferably near the crystal. Both ends of the capillary are then sealed with liquid wax. The drop from the mother liquid sealed inside the capillary ensures the right vapour pressure that protects the crystal from drying out and cracking.

For data collection in synchrotron X-ray beams, frozen crystals are necessary. This is done to minimize the effects of radiation damage at high brilliance tightly focused X-ray beams available at synchrotron beamlines. Even for home generators this method is useful as data collection takes days till the complete dataset is recorded. The high energy X-ray radiation causes breakage of chemical bonds releasing free radicals that diffuse through the crystal and attack other bonds destroying in this way the order in the crystal. Freezing the crystal to low temperatures, at around 100K, slows diffusion down and prolongs the life of the crystal in the X-ray beam. As the main component of the mother liquid is usually water, the problem of ice formation has to be overcome. Water ice poses problems in protein crystallography as it can damage the crystal and it produces so-called ice-rings in the diffraction image. The ice-rings form because ice is usually a mixture of many small water crystals so there are always a number of lattices in position for diffraction. This produces rings in the diffraction images, characteristic for powder samples. Flash cooling together with cryoprotectans (such as MPD, glycerol, PEG, glucose…) protect against ice formation and facilitate the transition from liquid to amorphous solid (vitrification). As for crystallisation of the *Mj*RelBE protein complex alcohols were used, the first choice was to hold on to the same alcohol and only raise its fraction in the motherliquor. The transition was performed by transferring the crystal through drops with stepwise increased alcohol concentration. This procedure was also useful for backsoaking the derivatized crystals and removing the excess heavy atom reagent. For ethanol grown crystals 30% ethanol was sufficient, 30% iso-propanol for the crystals grown in this alcohol and 20-22% of MPD for crystals grown from MPD conditions.

2.6 Crystallographic Methods

2.6.1 Measuring of X-Ray Diffraction Data

The diffraction data were measured in the rotation method by oscillating the crystal for a small angle (from 0.1° up to 2°) during one exposure and then proceeding to the next part of the reciprocal sphere. This continues till all the required data are acquired, which depends on the symmetry of the crystal in question. For triclinic cell it would require at least 180° rotation and for cubic (of the 432 point group) only at least 35°. More data usually means more redundancy but due to other factors, as radiation damage, it is not always recommended to continue collecting data after a complete dataset is acquired.

The crystals have been measured at the Enraf-Nonius FR-571 home generator, at synchrotron BESSY II, Berlin (PSF beamlines BL 14-1 and BL 14-2) and at ESRF, Grenoble (beamline ID 14-2). The crystals were prepared for X-ray measurement by shock freezing them in liquid nitrogen or mounting them in a glass capillary. The first method was always used when collecting data at the intense synchrotron beams to minimize radiation damage. The second method was occasionally used at the "in house" rotating anode X-ray generator equipped with an Image-plate detector (mar345, MAR research, Norderstedt). This second method was used to test the crystals prior the testing of cryoprotectants to assure that the observed diffraction quality (or the lack of any diffraction) of the crystals is not an artefact of the cryoprotectant.

2.6.2 Indexing and Integration of X-Ray Data

The collected X-ray data was preliminary processed with one program just to assess the quality of the crystal while data were still recorded. If the data was promising or showed problems during processing other programs were tested on the dataset. The programs for indexing and integrating used are DENZO (or HKL2000) (Otwinowski and Minor, 1997), MOSFLM (Leslie, 1992) and XDS (Kabsch, 1993). Merging of symmetry equivalent reflections and scaling of the diffraction images was performed with SCALEPACK (Otwinowski and Minor, 1997) or SCALA (Collaborative Computational Project, 1994). TRUNCATE from the CCP4-Program package was used for transforming the intensities in the SCALEPACK output to structure factor amplitudes that are used in further programs for data processing.

The best data (native and heavy atom derivatives) that was finally used for the structure determination were processed with DENZO / SCALEPACK.

2.6.3 Structure Determination

The main problem in crystallography is the so called "Phase problem". The diffraction data contain the structure factors magnitudes but the information of their phases is lost in a diffraction experiment. As we need both structure factors and phases in order to reconstruct the electron density distribution of the unit cell, it is first necessary to obtain the phases. There are a number of methods, and their varieties, for phase determination, direct methods, Multiple Isomorphous Replacement (MIR), Multiple Anomalous Dispersion (MAD) and Molecular Replacement (MR).

The easiest solution to the phase problem is found if a structure of a homologues protein has been previously determined. There have been several structures of RelE toxin family determined, one of them from an archaebacterium, *Pyrococcus horikoschii*. However for a successful MR the protein used for a model shares usually at least 30% sequence identity and has less than 1 Å root mean square deviation (r.m.s.d) of the main chain to the target structure. The highest sequence identity is with the *P. horikoschii*, 24% and failed to give good structural solution.

Attempts were made to obtain the phases with the MAD method from heavy atom soaked crystals but this approach failed as well due to low anomalous signal. Finally after obtaining good datasets with three heavy atom compounds the phases were determined using the MIRAS method.

2.6.4 Structure Determination of *Mj*RelBE Complex with the Multiple Isomorphous Replacement with Anomalous Signal (MIRAS) Method

The MIRAS method is a combination of MAD and MIR. The required data are the native dataset and the MAD datasets (peak, inflection and remote) of each of the derivatives. The best two of the available three heavy atom derivatives were chosen, ethylen mercury phosphate ($C_2H_7HgO_4P$) and cis-dichlorodiammine-platinum(III) (cis-$Pt(NH_3)_2Cl_2$), having 2.6 and 2.7 Å resolution respectively. Isomorphism of the derivatives with the native dataset was low, as the cell constants among the crystals differed over 1%, which made it

more difficult for a reliable solution of the phase problem. However with these datasets the solution of the phase problem was possible. Heavy atom search, substructure refinement and density modification were done in SHARP (Vonrhein *et al.*, 2007), which identified 4 Mercury and 11 Platinum sites. Density modification was done in the autoSHARP procedure by the SOLOMON program (Abrahams and Leslie, 1996).

2.6.5 Model Building and Refinement

The aim of refinement is to obtain a model that simultaneously best fits both the geometry requirements and the X-ray diffraction data. The used programs include the maximal- likelihood method based REFMAC5 (Murshudov *et al.*, 1997) with TLS refinement and simulated annealing and maximal likelihood programs from the CNS program package.

Prior to refinement a portion, usually 5% of the reflections were selected in thin shells evenly spread over the entire resolution range, and set aside. These reflections were used for cross validation of the refinement process by using the R-factor Rfree. Reflections that contribute to Rfree were not used throughout the entire refinement process. Initial model building was done with RESOLVE (Terwilliger, 2003) that was able to build 20% of the Cα positions correctly, and the complete model was built manually using COOT (Emsley and Cowtan, 2004). The model was refined with REFMAC5 (Murshudov *et al.*, 1997) from the CCP4 package (Collaborative Computational Project, 1994) and using simulated annealing followed by energy minimization and B value refinement from the CNS (Brünger *et al.*, 1998) package. Model bias was removed by calculating an annealed omit map and further rebuilding. Water molecules were placed in the electron density map at the end of refinement. The last step of refining in REFMAC5 included a TLS (Winn *et al.*, 2001) refinement step that improved the R and Rfree factors by another 3%. For the TLS refinement four TLS groups were defined that corresponded to the four polypeptide chains in the asymmetric unit. All refinements of the B-factors were isotropic.

2.6.6 Structure Validation and Figure Preparation

The validation of the stereo chemical properties of the polypeptide chains was performed with the programs PROCHECK and WHATCHECK (Hooft *et al.*, 1996). For superimpositions of polypeptide chains the programs LSQKAB (Kabsch, 1976) and TOP (Lu, 2000) were used. For the topographical representation of the polypeptide folding the TopDraw (Bond, 2003) program was used and the rest of protein figures were prepared with the program PYMOL (DeLano Scientific LLC, http://pymol.sourceforge.net/) or the Molscript program in combination with Raster3D. Secondary structure annotation was based on the DSSP program (Kabsch and Sander, 1983) output.

2.7 Predicting the Protein Fold

2.7.1 Sequence Comparison, Folding Recognition, Homology Modeling and Protein Interfaces Calculation

For database mining the BLAST (Altschul *et al.*, 1997) program was used. Multiple sequence comparisons were performed using the ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/) or the M-Coffee (Moretti *et al.*, 2007) server. For the presentation of the compared sequences the ESPript (Gouet *et al.*, 1999) server was used. Secondary structure prediction was performed with JUFO (Meiler and Baker, 2003). Fold recognition was performed with 3D-PSSM (Kelley *et al.*, 2000) and homology model building with PHYRE (Bennett-Lovsey *et al.*, 2008). Protein interfaces and molecular assemblies were calculated with the PISA service (Krissinel and Henrick, 2007).

3 RESULTS

The objects of this work are the members of the RelE toxin family and their cognate antitoxins from chromosomes of two organisms, the *Ec*RelBE TA protein complex from *Escherichia coli* (strain K12) and *Mj*RelBE TA protein complex from *Methanococcus jannaschii* (strain DSM2661).

The antitoxin *Mj*RelB of *M. jannaschii* is a 52 amino acid long protein with a theoretical pI of 5.16 while the corresponding toxin *Mj*RelE is an 88 amino acid long protein with a very basic pI of 10. The *E. coli* homologues are slightly larger, with the antitoxin *Ec*RelB being 79 amino acids long with a theoretical pI of 4.18 while the corresponding toxin *Ec*RelE is 95 amino acids long protein with a pI of 9.67.

3.1 Characterization of the *Ec*RelBE and *Mj*RelBE Protein Complexes

3.1.1 Cloning, Expression and Purification

The starting points of this work were the corresponding genes of *Ec*RelBE and *Mj*RelBE provided in plasmids *pSC2524HE* and *pSC2526MJ#2* respectively, both kindly provided by the group of Kenn Gerdes, Odense/Denmark.

3.1.1.1 Cloning, Expression and Purification of EcRelBE

The *Ec*RelBE carrying *pSC2524HE* construct with an N-terminal His_6-tag when expressed in ER2566 cells produced up to 1.5 mg per liter medium of *Ec*RelBE. However in the later stages of the work, due to problems in obtaining crystals, another construct was prepared without the His_6-tag (*pTrc99aRelBE_EC)*. This complicated the purification of the protein, and additional purification steps were needed in order to achieve the desired purity of the protein (>95%). However, this construct did not yield the desired crystals either.

For the purification of the His_6-tagged proteins Ni sepharose was used to retrieve His_6-tagged proteins from *E. coli* lysates and protein was further purified by gel filtration. The protein complex *Ec*RelBE that did not feature a His_6-tag (*pTrc99aRelBE_EC*) required AS precipitation and furthermore a cation exchange step prior to gel-filtration, as described in "Materials and Methods" (section 2.3.3.2).

3.1.1.2 Cloning, Expression and Purification of *Mj*RelBE

As the original construct harboring the *Mj*RelBE gene showed poor expression it had to be cloned in a new plasmid (see section 2.2.6). The so obtained construct *pET21relBE_Mj* was transformed into *E. coli* BL21 Rosetta™ cells (see section 2.2.6) that enable the expression of eukaryotic and archebacterial genes containing rare codons. This combination yielded expression levels of 6.5 mg protein from 1 L *E. coli* culture and allowed the purifcation of protein amounts sufficient for subsequent crystallization attempts. The DNA sequencing of the cloned *relE_Mj* gene showed that a mutation, leading to a R62S change on the amino acid level, was picked up during the cloning process. From all the other colonies obtained after component *E. coli* cells transformation, only one other clone, containing the R62K mutation, was identified. Other colonies contained frame shifted *relE_Mj* genes resulting in stop codons at the beginning of the sequence. To verify if the R62S and R62K mutation affects the RelE activity, the mutated *relE_Mj* genes were cloned into a vector without the antitoxin *relB_Mj* and tested for expression. The result was a good expression of the mutated protein *Mj*RelE that indicated the importance of Arg62. As the importance of this residue was established, the effort was made to reverse the mutation by mutating the *relE* gene back (S62R) to obtain the wild type gene. Attempts to clone the so obtained wild-type toxin gene alongside the antitoxin gene *relB_Mj* failed since the obtained construct proved to be too toxic for the *E. coli* strain used for cloning and the only transformants that grew on the selection plates featured frame shifts (producing stop codons) or empty vectors.

Heparin coated resin binds many DNA binding and basic proteins with high affinity. Since the theorectical pI of the *Mj*RelBE protein complex is 9.5, chromatography on heparin resin with buffers A and B was used as a purification step. To selectively purify the *Mj*RelBE protein complex, advantage was taken of the fact that the complex retains its positive charge up to a pH value of at least 8.0 at which many host proteins already pass their

pI. Most *E. coli* proteins do not bind at pH 8.0 to the heparin resin but the *Mj*RelBE complex binds very strongly and elutes only after raising the ionic strength of the buffer to 1M NaCl. In this way and at this stage the protein complex purity exceeded > 95% and the subsequent gel filtration step served mainly to remove aggregated MjRelBE complexes.

During the development of an efficient purification protocol it was noted that the *Mj*RelE protein and the *Mj*RelBE protein complex tend to aggregate. However, this problem could be avoided by increasing the NaCl concentration to 500 mM. As *Methanococcus jannaschii* thrives also in higher salt conditions (up to 0.85 M NaCl) this result could be understood as an adaptation to such environmental conditions.

Another puzzling effect was the behaviour of the *Mj*RelBE protein complex when loaded on a standard SDS-PAGE. The bands of the complex components could not be distinguished as they run in a continuous "smear". It was only possible in Tricine gels (Figure 6) to separate the *Mj*RelE and *Mj*RelB protein bands from each other. The tricine gel allowes better separation of smaller sized proteins.

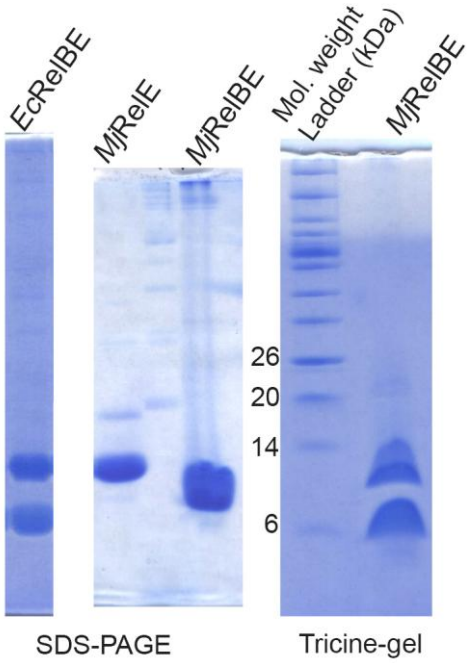

Figure 6 The Tricine gel effect. In SDS-PAGE the *Ec*RelBE runs at expected sizes as does the *Mj*RelE. The *Mj*RelBE complex when loaded on a SDS-PAGE is smeared regardless of the protein concentration. Only in Tricine gels the *Mj*RelBE complex is resolved into its components that run in predicted sizes according to their molecular weight. Gels stained with coomassie.

3.1.2 DLS and CD Measurements

The monodispersity of a sample is one of few parameters that is regarded as a prerequisite for successful crystallization trials. To determine the appropriate NaCl

concentration that result in a solution of monodispers *Mj*RelBE complex, a series of Dynamic light scattering (DLS) experiments with increasing salt concentrations were undertaken.

The readout of DLS experiments are apparent particle sizes. The determined particle sizes for the *Mj*RelBE complex incresed dramatically when the NaCl concentration of the corresponding buffer was below 400 mM (Figure 7). Based on this knowledge Gel filtration buffers were prepared with 500 mM NaCl. High salt concentration is one of few parameters that are usually avoided in the protein buffer used for crystallization trials but as there were no other acceptable options the crystallization trials were performed with the protein in 20 mM Tris pH 8.0 with 500 mM NaCl, and later changed (based on the results discussed in chapter 3.2.1) to 20 mM HEPES pH 7.5 with 500 mM NaCl.

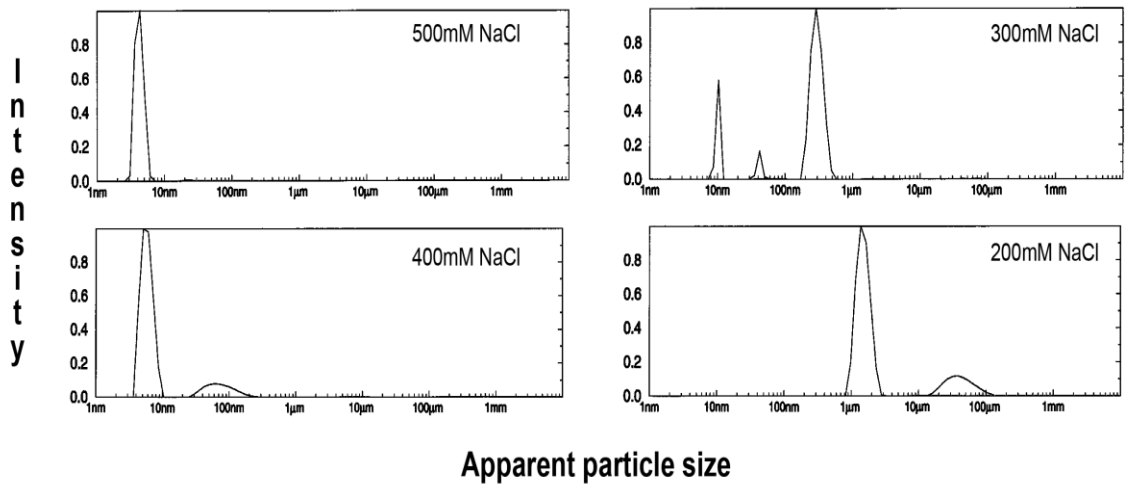

Figure 7 DLS measurements of 20mg/ml *Mj*RelBE in 20 mM HEPES pH 7.5 at various NaCl concentrations. The apparent particle size is plotted on the x scale giving the distribution of the particle sizes in solution. The aggregation behaviour upon decrease of NaCl concentration is evident.

The *Ec*RelBE protein complex showed a single uniform size when the NaCl concentration in the buffer exceeded 50 mM NaCl. Therefore, *Ec*RelBE was stored in 100 mM NaCl, 20 mM Tris pH 8.0 buffer for crystallization trials.

In order to assess the thermal stability of the proteins and to determine the α/β secondary structure contents, CD spectra of the purified RelBE complexes were measured at a range of temperatures

The CD spectra (Figure 8a) of *Ec*RelBE showed rapid unfolding when temperature reached 52°C, and this process was irreversible.

As expected, the proteins from the thermophilic *M. jannaschii* proved to be very heat stable up to the limits of the instrumental capabilities (88°C). Cooling the sample back down resulted in indentical CD spectra (Figure 8b) before the *Mj*RelBE was heated up. The calculated α/β content varied very little with temperature and was calculated to be 25.5% α helical, 20.17% β strands and 54% random. The actual α/β content of *Mj*RelBE in the finally determined structure is 39% α helices, 24% β strands and 37% random coil. This shows a mediocre agreement with the values determined by CD spectroscopy and underlines the limitations of this method.

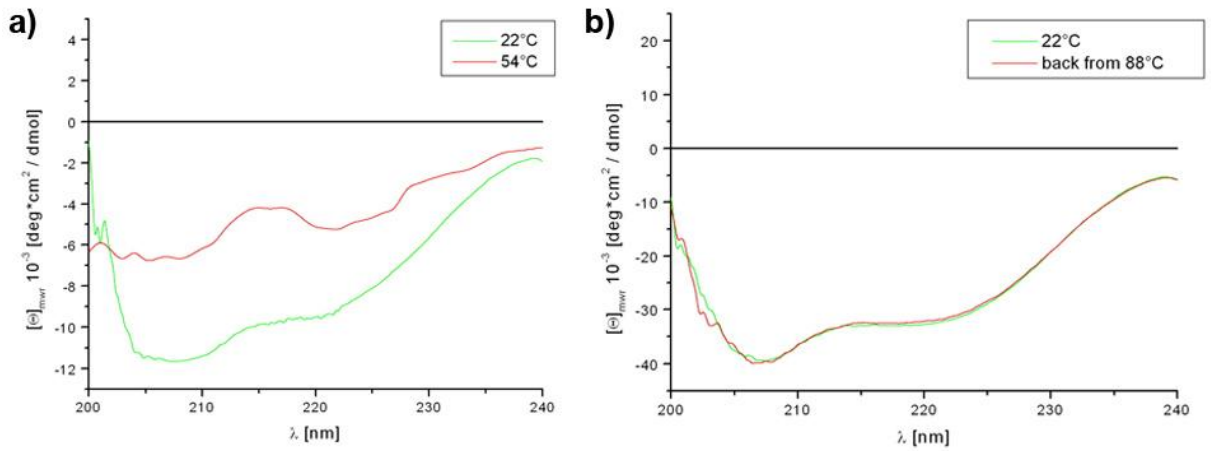

Figure 8 CD spectra: **a)** of 0.7 mg/ml *Ec*RelBE in folded and unfolded state; **b)** of 1.4 mg/ml *Mj*RelBE before and after heating to 88°C (no unfolding occurs during this process)

3.1.3 Size-exclusion, Analytical Ultracentifugation, Glutaraldehyde Cross-linking, N-terminal Sequencing and MALDI-TOF

In order to determine the behaviour and oligomerization state of the RelBE protein complexes in solution a set of experiments was carried out.

For *Ec*RelBE, size exclusion chromatography showed peaks corresponding to an approximate size of 50 kDa. This would correspond to a heterotetramer in solution. This result was supported by (Bhardwaj, 2005) reporting a concentration dependent equilibrium between heterotetramer and heterodimer observed in analytical ultracentrifugation (UC) experiments. MALDI-TOF mass spectrocopy (Figure 9) and N-terminal sequencing revealed that *Ec*RelE misses the first two residues (methionine and alanine) and *Ec*RelB the first 15

amino acids. N-terminal clipping of these residues may have been caused by *E. coli* host proteases.

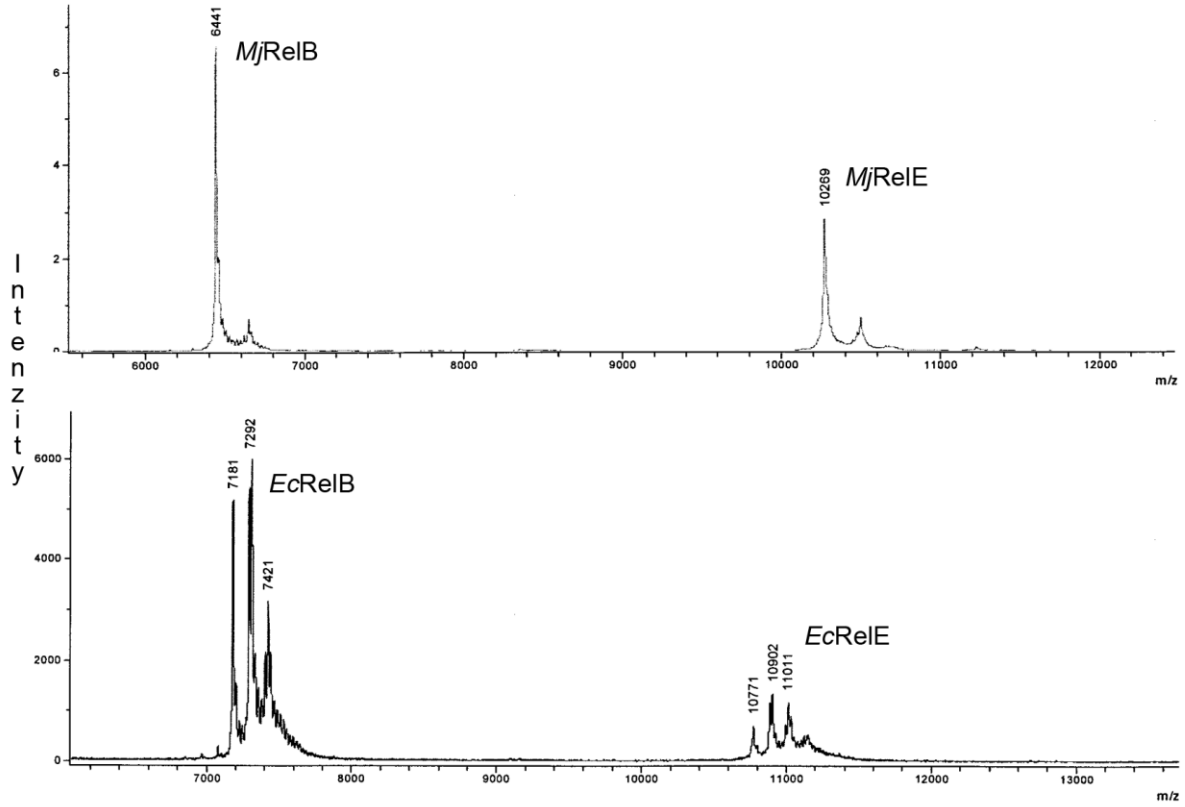

Figure 9 MALDI-TOF of the *Mj*RelBE and *Ec*RelBE protein complexes. The molecular masses of the expressed proteins are 6443 Da for *Mj*RelB (adduct of 6311 Da *Mj*RelB and 134 Da pinacolic acid from the matrix) and 10271 Da *Mj*RelE. The molecular masses of the expressed *E. coli* proteins are 7292 Da for *Ec*RelB and 10902 Da for *Ec*RelE

Size exclusion chromatography and UC experiments (Figure 10) of the *Mj*RelBE protein complex indicated a concentration dependent oligomerization behaviour as well. Higher concentrated *Mj*RelBE near 1 mg/ml, especially if the salt concentration was below 500 mM NaCl, tended to form higher oligomerization states or to aggregate. It was typical for the complex (concentration 10-22 mg/ml) to be eluted from the gel chromatography column as a heterooctamer, with an apparent molecular size of 64 kDa (see Appendix 7.6). In less concentrated solutions (below 0.5 mg/ml) the formation of an equilibrium between heterodimer and heterotetramer was observed, with the equilibrium being in favor of the latter (Bhardwaj, 2005). N-terminal sequencing and MALDI-TOF experiments showed that *Mj*ReB is present in full length whereas for *Mj*RelE the first two residues, Met1 and Lys2, were found missing (Figure 9).

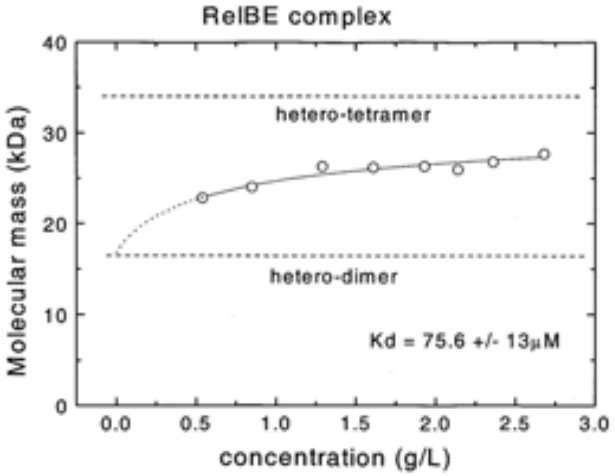

Figure 10 Ultracentrifugation results for *Mj*RelBE protein complex. Two regions are marking the concentration dependent oligomerization state: the hetero-dimer and the hetero-tetramer state. A curve is fitted to the measurements and the dotted extrapolation of the curve indicates the strong binding in the hetero-tetramer complex so the vast majority of the protein is in this oligomerization state in solution.

When the first datasets of *Mj*RelBE crystals were recorded, the question arose how the toxin and antitoxin molecules would spatially arrange to form a heterodimer or -tetramer. The answer to this question was of importance since structure determination by molecular replacement failed when the model of *Pyrococcus horikoschii Ph*RelBE, which was already available at that time, was used as a search model. Although the sequence identities of the RelE molecules are low (24.4% using BLOSUM62 similarity matrix) it was expected that the three-dimensional structures would be almost identical because they belong to the same toxin family. However, extensive trials to determine the *Mj*RelBE structure by molecular replacement failed. The question of the organization of the toxin-antitoxin complex was partially answered by the gel filtration retention time which indicated heterotetrameric and heterooctameric organization (Bhardwai thesis and Apendix 7.6). A simple cross-link experiment was set up in order to investigate this question further (Figure 11). The results, without further complementary experiments, were inconclusive but indicated the possibility of a RelB-RelB interaction, contrary to the published structure of *Ph*RelBE. They apparently indicated a formation of *Mj*RelB homodimers (12 kDa) and *Mj*RelBE heterodimers (16kDa). Simmilar experiments with *Ec*RelBE did not yield any interpretable results.

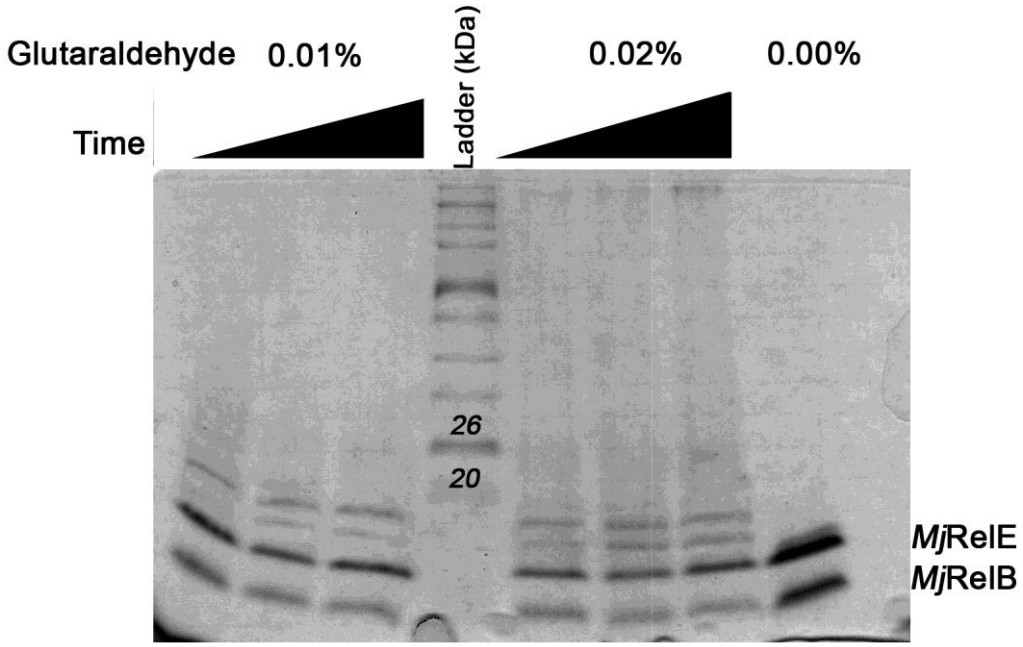

Figure 11 Glutaraldehyde cross-linking experiment. The samples were incubated 5, 10 and 15 minutes each with 0.01% and 0.02% glutaraldehyde and loaded on a tricine gel.

3.1.4 Endoribonuclease Tests

To verify the hypothesis that RelEs can cleave RNAs only in the presence of the ribosome (see Introduction section 1.2.3), two substrate single stranded RNAs were designed and synthesized (see Materials section 2.4.7). According to (Pedersen *et al.*, 2003) RelEs are supposed to cut the *RNAcut-* seqences but only minor cleavage of the *RNAn-* control sequences should be observed. As *Mj*RelE could only be expressed with a mutation in the active site rendering the enzyme most likely inactive, the RNase activity test was solely performed with the *Ec*RelE protein.

The *Ec*RelE protein used in the RNA cleavage assays was purified as described in Methods and Materials (section 2.3.3.2 and Results section 3.1.1.1), the completion of the refolding process was verified by CD spectrometry. Although at the beginning there were indications of *Ec*RelE ribonuclease activity (the sample denoted with "old" in Figure 12), in later experiments the results could not be confirmed (Figure 12). The lane with the positive test containing RNase T1 and ssRNA shows how RNA can be rapidly digested even with

minute amounts of the enzyme. If the ssRNA stock went through more than two freeze-thaw cycles and samples from it used for experiments, a pattern of degradation appeared in the controls.

Degradation of the RNA chain may originate from a contamination by RNases although every experiment was performed in sterile conditions with sterile and RNA free equipment (as described in Material section 2.4.7).

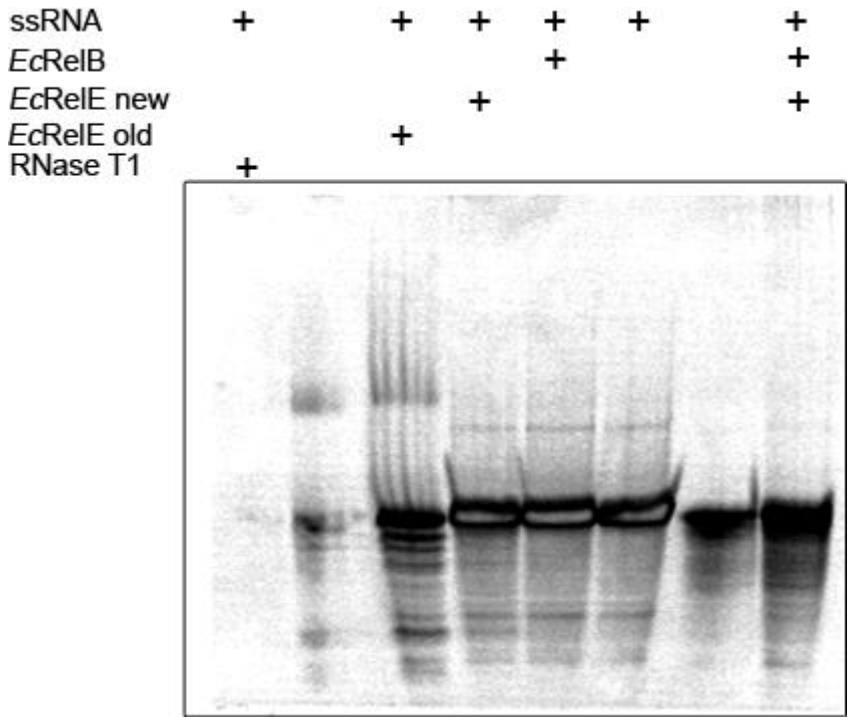

Figure 12 Ribonuclease test. The sample ssRNA was incubated with the proteins, loaded on a denaturating polyacrylamide gel and stained with ethidium bromide. The ssRNA sample containing RNaseT1 is thoroughly digested (utmost left lane) while the ssRNA in the RelE containing samples is undigested.

This result (Figure 12) was expected for *Ec*RelE but unfortunately the experiment could not be carried out on *M*jRelE, which was the centre of interest, because the protein *Mj*RelE could not be obtained in the active form.

3.2 Crystal Structure of the Archaebacterial *Mj*RelBE Protein Complex

3.2.1 Crystallization of the *Mj*RelBE Protein Complex

Crystallization is one of the narrowest bottlenecks in modern macromolecular crystallography. It was not different in this work as the search for the initial and refined crystallization conditions took hard and long efforts.

Despite all effort no crystals of *Ec*RelBE were obtained.

In the crystallization drops of *Mj*RelBE the standard observation was that of phase separation or heavy precipitation. For the *Mj*RelBE protein complex the first indications of crystal formation were obtained in Hampton Crystal screen HT in a 20% ethanol condition (Figure 13a). These nucleation clumps were still too small to be mounted in the X-ray beam but offered a good starting point for further refinement of crystallization conditions. More crystals could be reproducibly grown if the protein concentration was raised to 15 mg/ml (Figure 13b). The best crystals with ethanol as precipitant were obtained by addition of 100 mM Tris pH 7.5 and protein concentration of 10 mg/ml (Figure 13c). However crystals grown in this condition are intertwined and the diffraction pattern was of low resolution and not interpretable.

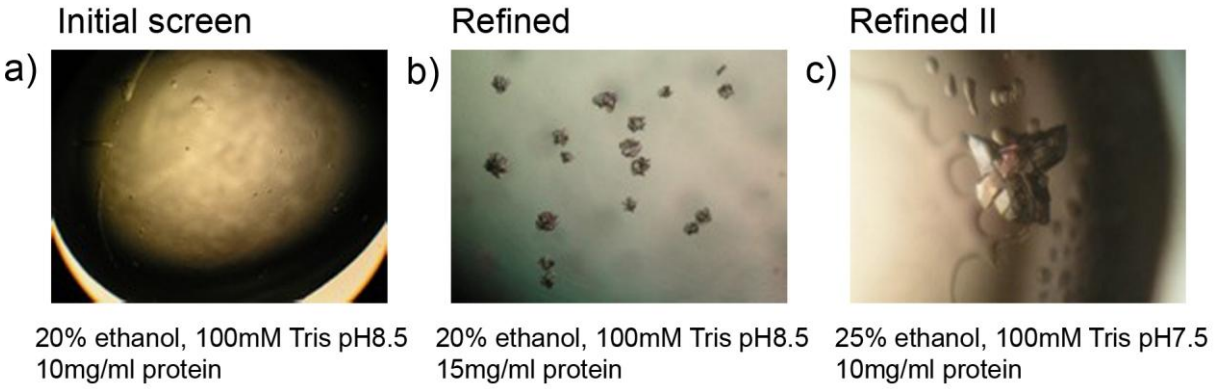

Figure 13 First crystals of the *Mj*RelBE protein complex. These starting crystals were not suitable for X-ray data collection. **a)** The original nucleation clumps from the Hampton screen, **b-c)** refined crystallisation conditions.

To refine crystallization condition further, various alcohols replacing ethanol as precipitaing agent were tested. Large and irregularly shaped crystals with largest dimensions of 0.9 mm were obtained from conditions containing 20-30% iso-propanol buffered with 100 mM Tris pH 7.0 (Figure 14a). X-ray diffraction data collected form these crystals showed reflections corresponding to a resolution of 4-8 Å, however, in conjunction with an estimated

mosaicity of 3° and high anisotropy of reflections depending on the crystal and its orientation, the obtained data was still insufficient for structure determination. Attempts to index these images gave conflicting results concerning the determination of the Laue group. Finally, irregularly shaped crystals of various sizes from 0.1 to 0.7 mm were obtained form a crystallization condition composed of 4-22% MPD and 100 mM HEPES pH 7.5 (Figure 14c). These crystals diffracted anisotropically to a maximum resolution of 2.6 Å. The space group was determined to $P2_1$ and first complete native and derivative datasets were collected.

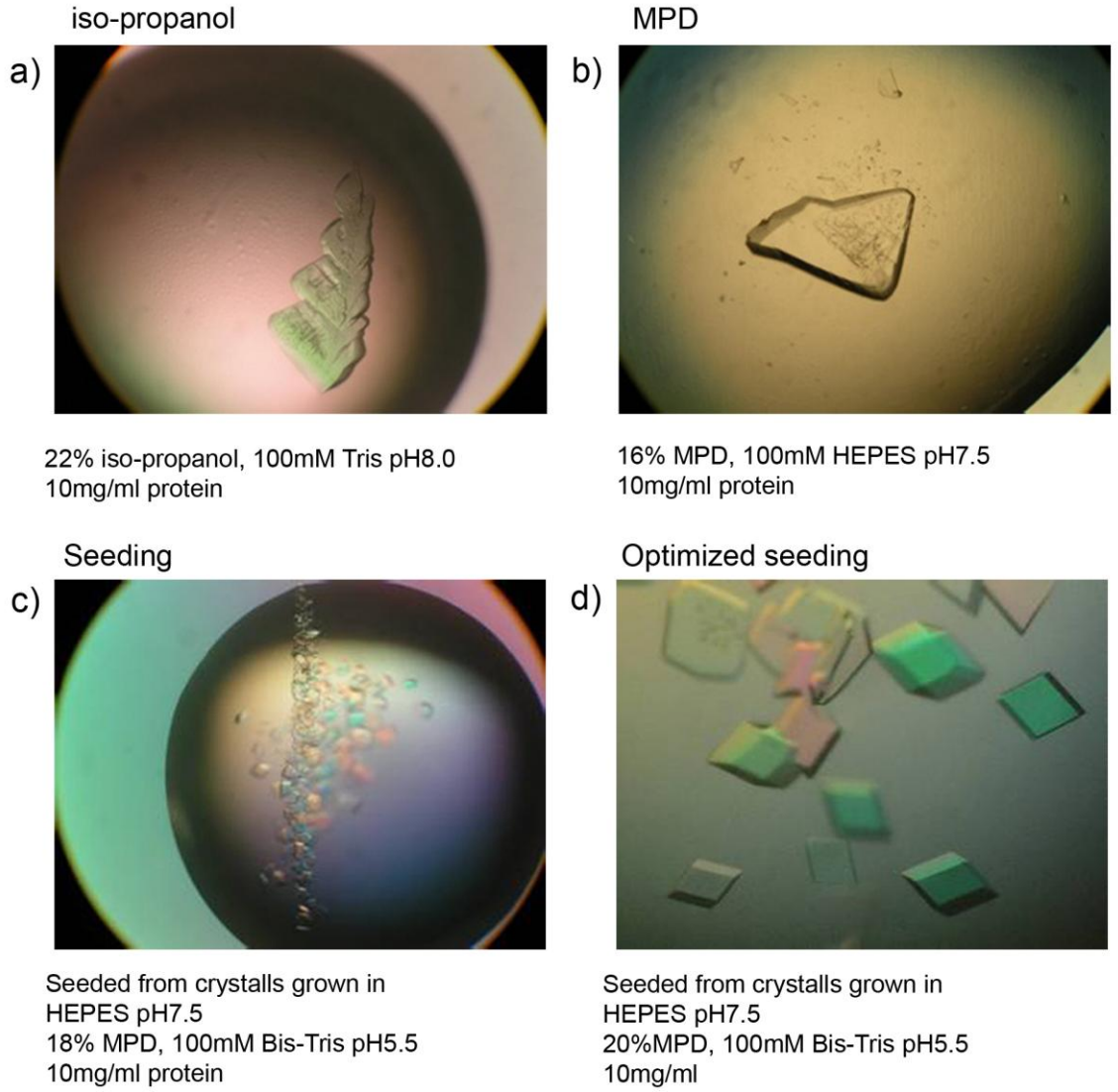

Figure 14 Refining crystallization conditions. These crystals diffracted X-rays well and data collection was possible. **a)** The oddly shaped crystals grown in iso-propanol are a collection of crystals grown in different directions. **b)** Crystals grown in MPD conditions are composed of better ordered layers of crystal lattices from which derivative and native data was collected. **c)** Rigorous grounding of the MPD grown crystals and the use of the so obtained nucleation seeds resulted in **d)** clearly shaped crystals for which higher resolution data was collected (up to 2.1 Å).

In order to obtain higher resolution native datasets crystallization conditions were further refined to yield better crystals. An optimized protocol for seeding (as described in "Materials and Methods") proved to be particularly useful and rhombohedrally shaped crystals of uniform sizes between 0.3 and 0.5 mm were obtained that diffracted x-rays up to 2.1 Å resolution.

For phase determination X-ray diffraction datasets of crystals containing heavy atom compounds were collected. There are in total three methionins in the *Mj*RelB and *Mj*RelE sequences but only 2 present in the expressed protein (see 3.1.3), and out of those 2 one is placed at the N-terminus of the antitoxin with a high chance of being disordered in the crystal. This would leave one SeMet per 138 residues, which is far from the recommended one SeMet per 40 AA residues. Therefore experiments to soak crystals or to co-crystallise proteins with heavy atom compounds were pursued rather than selenomethionine (SeMet) substitution. More than 40 heavy atom compounds were tested out of which 12 where non-destructive to the crystals (determined by visual inspection). Co-crystallizing *Mj*RelBE with ethylene mercury phosphate ($C_2H_7HgO_4P$) and trimethylleadacetate ($C_5H_{11}O_2Pb$), respectively, yielded crystlas from which multiple anomalous dispersion (MAD) datasets were collected resolutions ranging from 2.6 Å to 2.8 Å were used to solve the phase problem. Another data set was collected for cis-dichlorodiammine-platinum(III) (cis-$Pt(NH_3)_2Cl_2$) with 2.7 Å resolution. Crystals soaked with any other of the 12 non-destructive compounds diffracted poorly or showed no X-ray fluorescence when subjected to synchrotron X-ray light of the appropriate wavelength (indicating that these heavy atoms are not present in the crystal).

3.2.2 Data Collection, Phase Problem Solution and Structure Refinement

The space group of crystals grown in MPD conditions (Figure 14b-d) was determined to be $P2_1$. Unit cell parameters of native and derivative crystals are given in Table 5. Notably, the refined three cell dimensions are near 58 Å and the beta angle is near 90 degrees, suggesting the presence of a higher symmetric (cubic) space group. However, this possibility was ruled out since the statistics after scaling favoured the $P2_1$ space group and ultimately structure determination confirmed that the space group is indeed $P2_1$ (see below).

The structure of a previously published *Mj*RelBE complex homolog from *P. horikoshii* (*Ph*RelBE, 24% sequence identity of all residues (Takagi *et al.*, 2005) was used as a search model in molecular replacement with EPMR (Kissinger *et al.*, 1999), Amore (Navaza, 1994) and Phaser (McCoy *et al.*, 2007). Different search models were produced by using the complex as published or using individual *Ph*RelE or *Ph*RelBE subunits and/or deleting non conserved protein regions. However, all molecular replacemnt trials failed (McCoy *et al.*, 2007) and other means for phase determination had to be utilized. For this purpose heavy atom labelled protein crystals were prepared by heavy atom soaking.

As the anomalous signal from these derivative crystals was weak, the phase problem was solved with multiple isomorphous replacement (MIRAS) where the difference in heavy atom contribution to the overall diffraction intensities is used for phase angle determination. As the phase angle cannot be determined correctly with only one derivative, the acquisition of diffraction data from another isomorphous heavy atom derivatized crystal is necessary. The phase angle ambiguity can then be removed and correct phases calculated.

Table 5 Maximal unit cell parameter deviation. The coresponding unit cell parameters of the native and derivative crystals (used in MIRAS) differ more than 1% and therefore at the edge of what is commonly refered to as isomorphous crystals.

Crystal	Cell lengths (Å)			Cell angles (°)		
	a	b	c	α	β	γ
Native	**52.794**	57.992	**58.747**	90.000	**92.297**	90.000
ethylen mercury phosphate (derivative I)	**52.041**	58.371	**59.754**	90.000	**91.716**	90.000
cis-dichlorodiammine-platinum (derivative II)	**52.695**	57.918	**57.659**	90.000	**92.741**	90.000
Mean unit cell parameters	52.510	58.094	58.720	90.000	92.251	90.000
Maximal cell parameter deviation (%)	**1.426**	0.776	**3.506**	0.000	**1.118**	0.000

For the multiple isomorphous replacement method, the collection of one native and two derivative datasets from isomorphous crystals is required (with unit cell parameters varying not more then 1%). If the cell parameters vary more (non-isomorphism), the phase solution becomes more difficult to obtain.

The phase problem was solved by using the datasets) from ethylene mercury phosphate (derivative I) (Table 6), cis-dichlorodiammine-platinum (derivative II) and a native

dataset from the crystals grown by the optimized seeding procedure. The MIRAS method was successful despite the large difference among the native and derivative crystals in cell lengths and angles.

Table 6 Data collection statistics of the native and derivative datasets, phase solution and for the refinement of the structure from the native *Mj*RelBE dataset.

Data collection

	Native	Ethylen mercury phosphate	Dichlorodiammine-platinum
Space group	$P2_1$	$P2_1$	$P2_1$
Cell parameters			
a, b, c (Å);	52.79 57.99 58.75	52.04 58.37 59.75	52.69 57.92 57.66
α, β, γ (°)	90.00 92.30 90.00	90.00 91.72 90.00	90.00 92.74 90.00
Wavelength (Å)	0.9490	1.0074	1.0722
Resolution limits (Å)	50.0-2.13 (2.13-2.18)	50.0-2.6 (2.60-2.70)	50.0-2.7 (2.70-2.80)
Total reflections	232747	382504	193774
Unique reflections	20062	11235	9759
R sym[a]	0.050 (0.42)	0.050 (0.34)	0.064 (0.35)
I / σ(I)	27.03 (3.13)	28.00 (6.50)	18.88 (4.14)
Completeness (%)	99.6 (98.9)	99.9 (100.0)	99.7 (99.6)
Redundancy	4.2 (4.1)	7.6 (7.7)	5.7 (5.5)
Mosaicity (°)	0.39	0.99	0.97

a. *R*sym = Σhkl Σj|Ij-<*I*>|/ΣhklΣjIj, where <*I*> is the mean intensity of reflection *hkl*.

The high mosaicity in the derivates, problems with intensity distribution among diffraction spots (reflections) in all datasets and high anisotropy in all datasets, were causing problems during data processing. Structure refinement statistics are summarized in (Table 7)

Table 7 Structure refinement statistics on the native dataset.

Refinement

Resolution (Å)	20.0-2.2
Unique reflections	17159
R_{work} (%)[a] /R_{free} (%)[b]	0.238 / 0.274
No. of atoms	
Protein	2191
Water	56
B-factors (average)(Å^2)	50.04
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.058
Geometry (%)	
Most favored	93.8
Additionally allowed	5.7
Generously allowed	0.4

a. *R*factor = Σhkl||*F*obs|-|*F*calc||/ Σhkl|*F*obs|; where Fobs and Fcalc are respectively, the observed and calculated structure factor amplitude for reflections hkl included in the refinement.
b. *R*free is the same as *R*factor but calculated over a randomly selected fraction (5%) of reflection data not included in the refinement

3.2.3 Crystal Packing

The toxin *Mj*RelE and its antitoxin *Mj*RelB form a tight heterodimer and, as shown above (see 3.1.3), tend to form higher oligomerization states in solution. A stable form of the *Mj*RelBE protein complex seemed to be a heterotetramer. In the crystal the asymmetric unit consists of a heterotetramer where the two *Mj*RelBE heterodimers are related by a twofold non-crystallographic symmetry (NCS) axis (the molecule depicted with the NCS axis Figure 15). There are two molecules in the unit cell, related by a two-fold screw axis along the b axis.

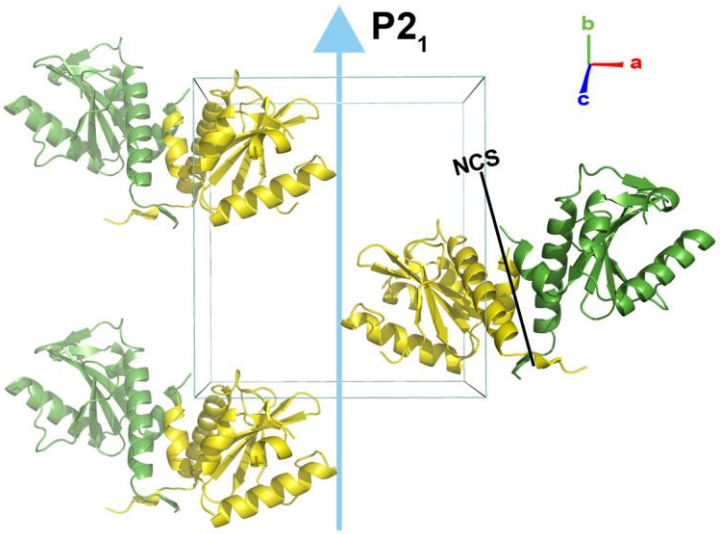

Figure 15 The $P2_1$ crystal packing viwed from front with the $P2_1$ screw axis almost parallel to the b unit cell axis. The twofold NCS axis of the heterotetramer is depicted for only one of the symmetry related molecules.

Figure 16a depicts the crystal packing viewed with the b crystal cell axis parallel to the plane and Figure 16b viewed from the top (down the b axis). The symmetry related molecules in this representation are all related by translation.

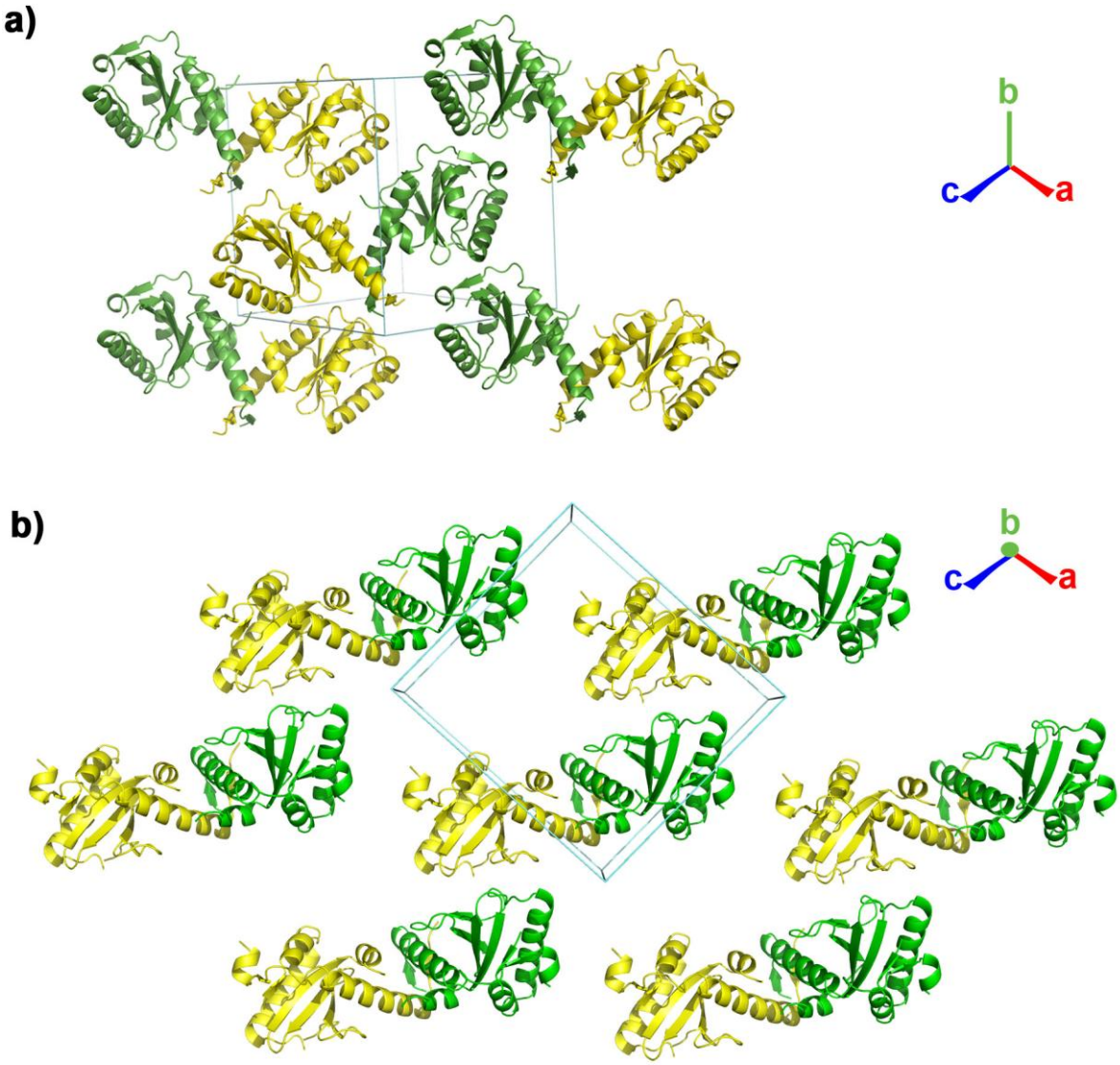

Figure 16 Crystal packing of the *Mj*RelBE crystals. **(a)**, with the b axis parallel to the plane, and from the top **(b)**, viewed down the b axis. For clarity only the symmetry related chains in one plane are given per view. The unit cell is presented in blue, heterodimers I (containing $MjRelB_I$ and $MjRelE_I$) in yellow and heterodimer II (containing $MjRelB_{II}$ and $MjRelE_{II}$) in green.

3.3 Structure Description

*Mj*RelBE forms a heterotetramer consisting of two heterodimers that are related by a non-crystallographic symmetry axis (NCS) (Figure 17). Heterodimer I ($MjRelBE_I$) consists of $MjRelB_I$, $MjRelE_I$, and heterodimer II ($MjRelBE_{II}$) of $MjRelB_{II}$, $MjRelE_{II}$ (in the PDB file they are denoted as chain A, B, C and D, respectfully). In the heterotetramer the N-

terminal β-strands β1 and α-helices α1 of the antitoxins *Mj*RelB$_{I}$ and *Mj*RelB$_{II}$ are connected by hydrogen bonds and van der Waals interactions, while helix α1 of RelB is tightly bound to a cleft (se below) of RelE. The remaining two loops, strand β2 and helix α2 of the antitoxins *Mj*RelB$_{I,II}$ are tightly wrapped around the toxins *Mj*RelE$_{I}$ and *Mj*RelE$_{II}$ in each of the *Mj*RelBE heterodimers.

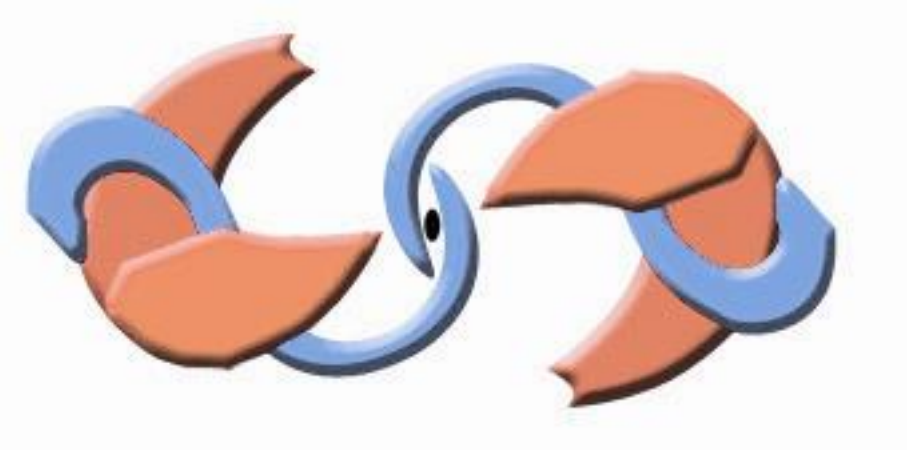

Figure 17 Schematic representation of the *Mj*RelBE heterotetramer. The *Mj*RelE toxins (red) is griping the *Mj*RelB antitoxins (blue) that form a β-strand with the second *Mj*RelB, and in this way connect two *Mj*RelBE heterodimers to form a heterotetramer. The NCS (black ellipse) is positioned in the middle of the complex passing in between the central β-sheet.

In the schematic representation (Figure 17) the toxins (red) are represented as pincers grabbing the antitoxins (blue). The two heterodimers are related by a twofold NCS that runs through their shared β-sheet formed by the β1 strands of *Mj*RelB$_{I}$ and *Mj*RelB$_{II}$. A rotation by 180° would superpose the two heterodimers, with some differences discussed below in section 3.3.5.

N-terminal sequencing and MALDI-TOF spectra showed that full length antitoxin *Mj*RelB was present before and after crystallization while the first two residues of the toxin *Mj*RelE were missing. The electron density was not interpretable for residues 1-7 and 49-52 of *Mj*RelB$_{I}$ and for residues 1-8 and 48-52 of *Mj*RelB$_{II}$ while for toxin *Mj*RelE$_{I}$ electron density was not found for Asp39.

Since the topographies of the heterodimers *Mj*RelBE$_{I}$ and *Mj*RelBE$_{II}$ are similar, only the former, if not explicitly stated otherwise, is described here. The core of the toxin *Mj*RelE (Figure 18, red) is formed by an antiparallel five-stranded β-sheet except for strand β1 that is parallel to the adjoining strand β5. The β-sheet is flanked by three helices, helices α1 and α2 being on one side forming a U-shaped base (Figure 19, red) while the C-terminal helix α3 is on the other side. Helix α3 and the loop connecting strands β2 and β3 form a deep

cleft accommodating the active site (see below) that is blocked by the long helix α1 of antitoxin *Mj*RelB (blue).

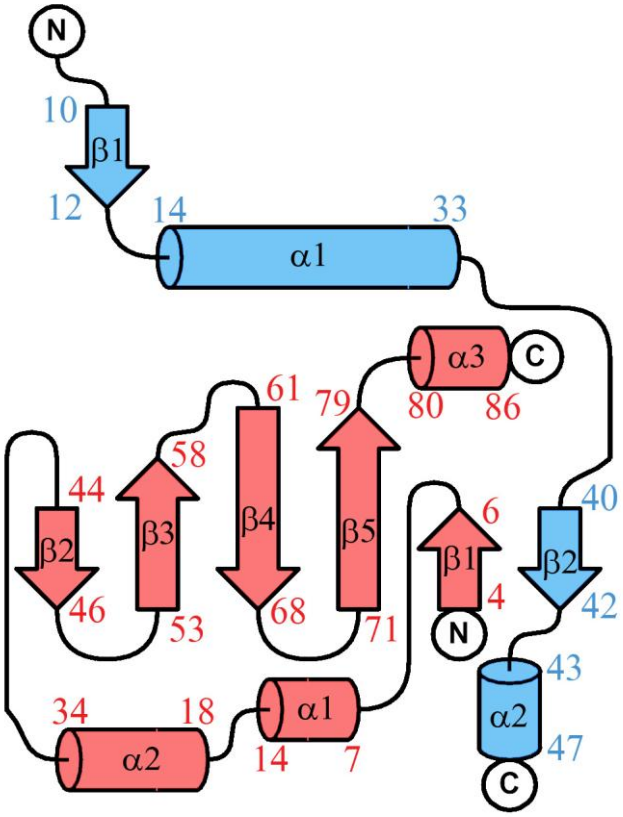

Figure 18 Topographical representation of *Mj*RelE (red) and *Mj*RelB (blue). The starting and ending residue numbers of secondary structure elements are written next to them. Strand β2 from antitoxin *Mj*RelB forms an antiparallel β-sheet with strand β1 of toxin *Mj*RelE, and helix α1 of *Mj*RelB is positioned above the five stranded central β-sheet of *Mj*RelE and gripped from both sides by helix α3 and the β2-β3 loop of *Mj*RelE. Two of these heterodimers form a heterotetramer *Mj*Rel(BE)$_2$ by the interaction of the antitoxins *Mj*RelB that form an antiparallel β-sheet with their β1 strands, and by binding of α3 of *Mj*RelE$_I$ to α1 of *Mj*RelB$_I$.

The antitoxin *Mj*RelB is wrapped around toxin *Mj*RelE, with strand β2 forming an antiparallel sheet with β1 of the toxin *Mj*RelE, while strands β1 and β1' (' indicating the NCS related partner) of the two antitoxins form another short antiparalel β-sheet that contributes to the stabilization of the heterotetramer.

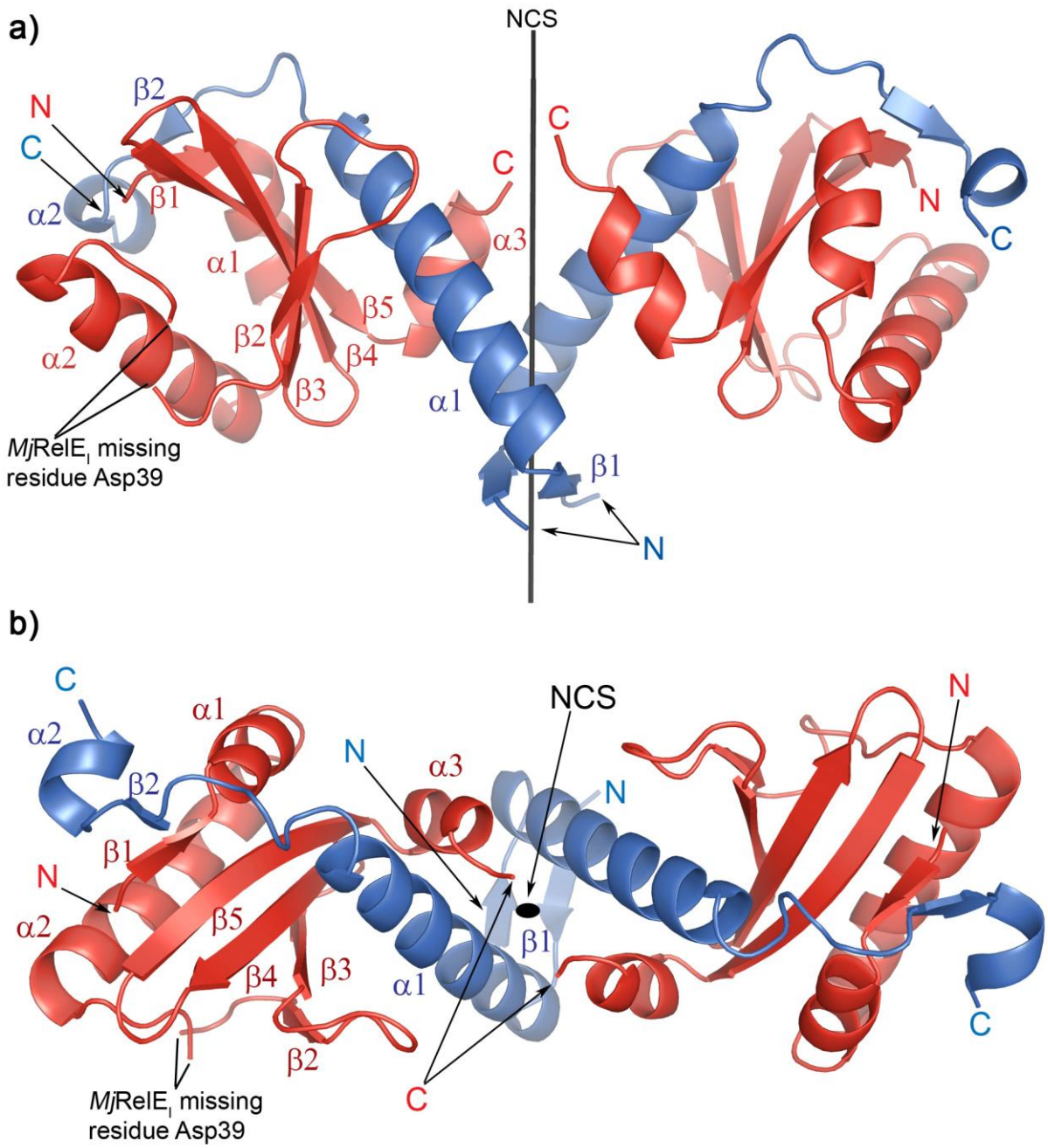

Figure 19 The *Mj*RelBE heterotetramer. The antitoxins *Mj*RelB are colored blue and the *Mj*RelE toxins red with N and C-termini marked with letters in the same color. The secondary structure elements of the *Mj*RelBE$_I$ heterodimer are marked and the twofold NCS axis is marked with a black line (figure a) and an ellipse (figure b). **a)** Side view of the heterotetramer; **b)** top view, down the NCS axis, of the heterotetramer.

3.3.1 *Mj*RelE – *Mj*RelB Complex Formation

*Mj*RelE features a large cleft that accommodates the long helix α1 of *Mj*RelB. Besides the large contact surface between the cleft formed by *Mj*RelE and the long helix α1 of *Mj*RelB, further interactions involve the long loop (residues 34-39) connecting α1 and β2 of *Mj*RelB, its short strand β1 and the short C-terminal helix α2. In the heterodimer the interactions between the toxin and antitoxin bury 2740 $Å^2$ of the solvent-accessible area. The contribution of the buried surface to the dissociation free energy is around 17 kcal/mol (the exact value varies slightly between the two heterodimers due to slight structural differences), as calculated by the PISA service (Krissinel and Henrick, 2007).

The cleft formed by *Mj*RelE (Figure 20 and Figure 21) is positively charged with Lys7, Lys45, Lys48, Arg55, Arg57, Arg62 (this position is replaced by Ser62 in the here described inactive variant *Mj*RelEArg62Ser), Arg75, Arg76, and Arg80. Additionally there are polar residues Tyr50, Tyr53, Tyr86, and hydrophobic residues Phe87, Pro88.

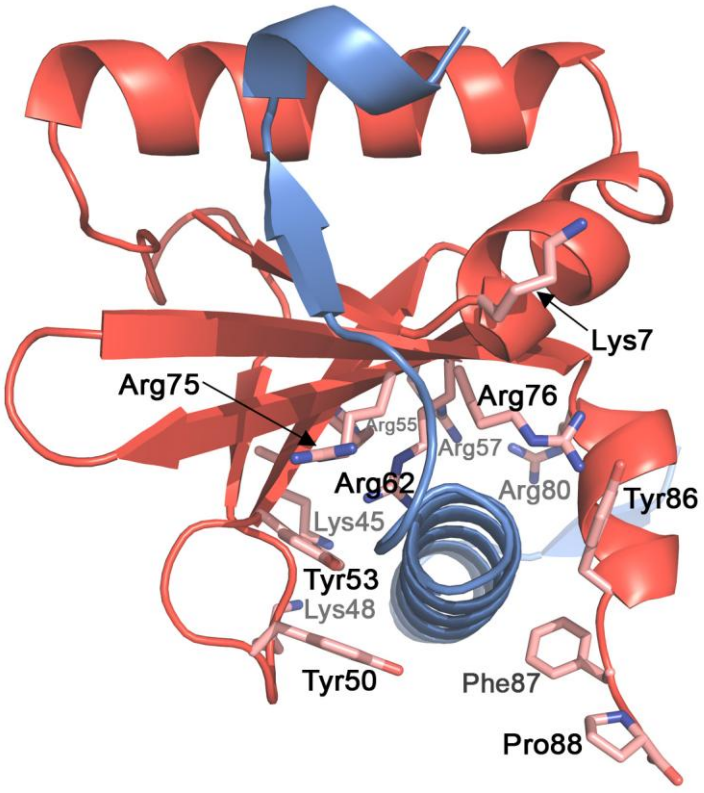

Figure 20 The charged residues located in the *Mj*RelE binding cleft. Positively charged residues Lys7, Lys45, Lys48, Arg55, Arg57, Arg62 (modelled), Arg75, Arg76, and Arg80 are mainly located inside the while the uncharged Tyr50, Tyr53, Tyr86 and hydrophobic Phe87 and Pro88 are placed on the outer walls of the cleft. *Mj*RelB is coloured blue. Residues located further from the observer are labelled in light grey.

The interactions with *Mj*RelB are mainly electrostatic and involve, on $MjRelB_I$ Glu21, Asp25, Lys30 and on $MjRelB_{II}$ Glu21, Glu22, Asp25, Glu34 (see contact tables in the Appendix 7.5). The loop between helix α1 and strand β2 of *Mj*RelB interacts with *Mj*RelE

through hydrogen bonds between *Mj*RelB Asp38OD2, Glu39N and *Mj*RelE Lys7O, Thr8N, Thr8Oγ. Further hydrogen bonds are found in the β-sheet formed by strands β2 of *Mj*RelB (Glu39O, Leu41N, Leu41O) and β1 of *Mj*RelE (Phe5O, Phe5N, Lys7N).

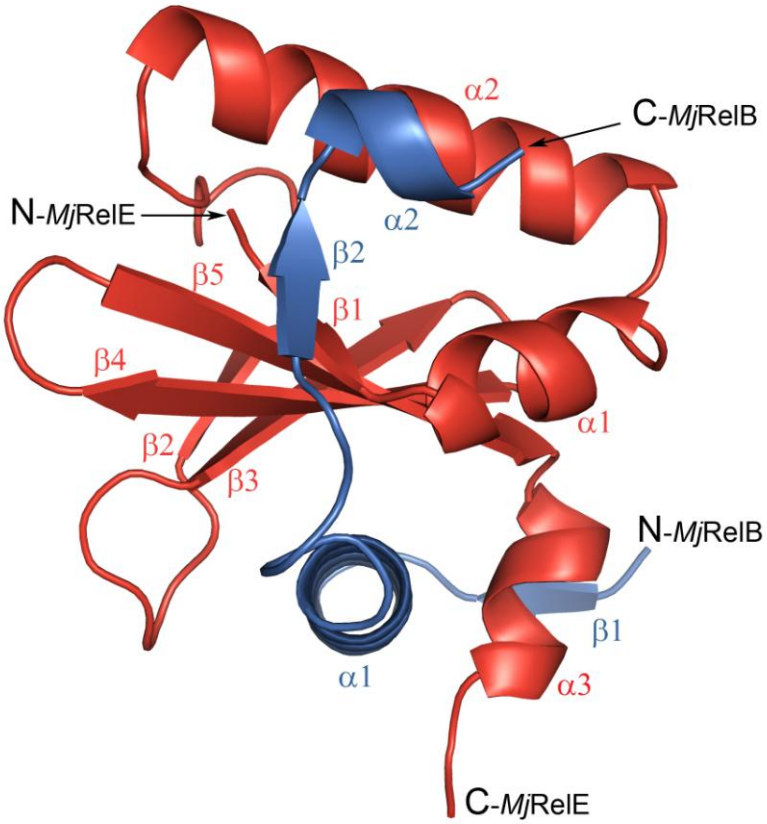

Figure 21 *Mj*RelBE heterodimer. *Mj*RelB (blue) helix α1 is gripped on both sides by *Mj*RelE (red) helix α3 and the β2-β3 loop and nested on top of the central β-sheet of *Mj*RelE. The β1 strand of *Mj*RelE is forming with the *Mj*RelB strand β2 a short β-sheet while *Mj*RelB α2 forms only some hydrophobic contacts with α2 of *Mj*RelE. β1 of *Mj*RelB is too far to form any interactions with the toxin *Mj*RelE but forms a short antiparallel β-sheet with strand β1' of the NCS-related *Mj*RelB'.

3.3.2 *Mj*RelBE Heterotetramer Formation

Analytical ultracentrifugation confirmed the formation of a heterotetramer from the heterodimers in a wide variety of protein concentrations with a dissociation constant of 75.6 µM. It was expected to find a heterotetramer of the kind as reported for the *Ph*RelBE homolog which, however, proved to be wrong as shown by the models that differ grossly (as will be discussed in section 3.4.1).

In the heterotetramer the antitoxins *Mj*RelB$_I$ and *Mj*RelB$_{II}$ form an antiparallel β-sheet (Figure 22a) with strands β1 and β1' (' indicating the NCS related partner). Further interactions involve hydrophobic contacts between helices α1 and α1' through residues Ile20, Ile23, and Ile20', Ile23' and interactions between residues Phe10, Tyr17, and Phe10', Tyr17'. The Phe10 residue is part of the β1 strand whilst the Tyr17 is part of the α1 strand following the first β strand of the antitoxin. These two residues (Phe10 and Tyr17), one from heterodimer I and the other from heterodimer II and wise versa, are participating in π-stacking where the two aromatic rings are placed at about 4 Å distance on top of each other (Figure 22b). The hydrogen bonds between the β1 and β1' strands form the characteristic pattern for an antiparallel β-sheet. Beside the hydrogen bonds participating in the β-sheet formation, there is another hydrogen bond between Ser13' hydroxyl group of *Mj*RelB$_{II}$ and Phe10N of *Mj*RelB$_I$ but this interaction is not present in the vice versa direction (heterodimer I Ser13 and heterodimer II Phe10). This is further discussed in section 3.3.5. Further down the antitoxin chains, towards the C-termini, the antitoxins interact with each other by hydrophobic interactions. The α1 helices of the antitoxins are turned with their hydrophobic residues to each other, thereby increasing the bonding surface that connects the two heterodimers.

Interactions between the *Mj*RelBE heterodimers include also cross interactions of toxin-antitoxin from the opposite dimers. These interactions include hydrogen bonds between residues in the C-terminal helix α3 of *Mj*RelE$_I$ (residues Arg80, Lys81 and Tyr84) and residues in helix α1' of *Mj*RelB$_{II}$ (residues Glu16 and Lys19).

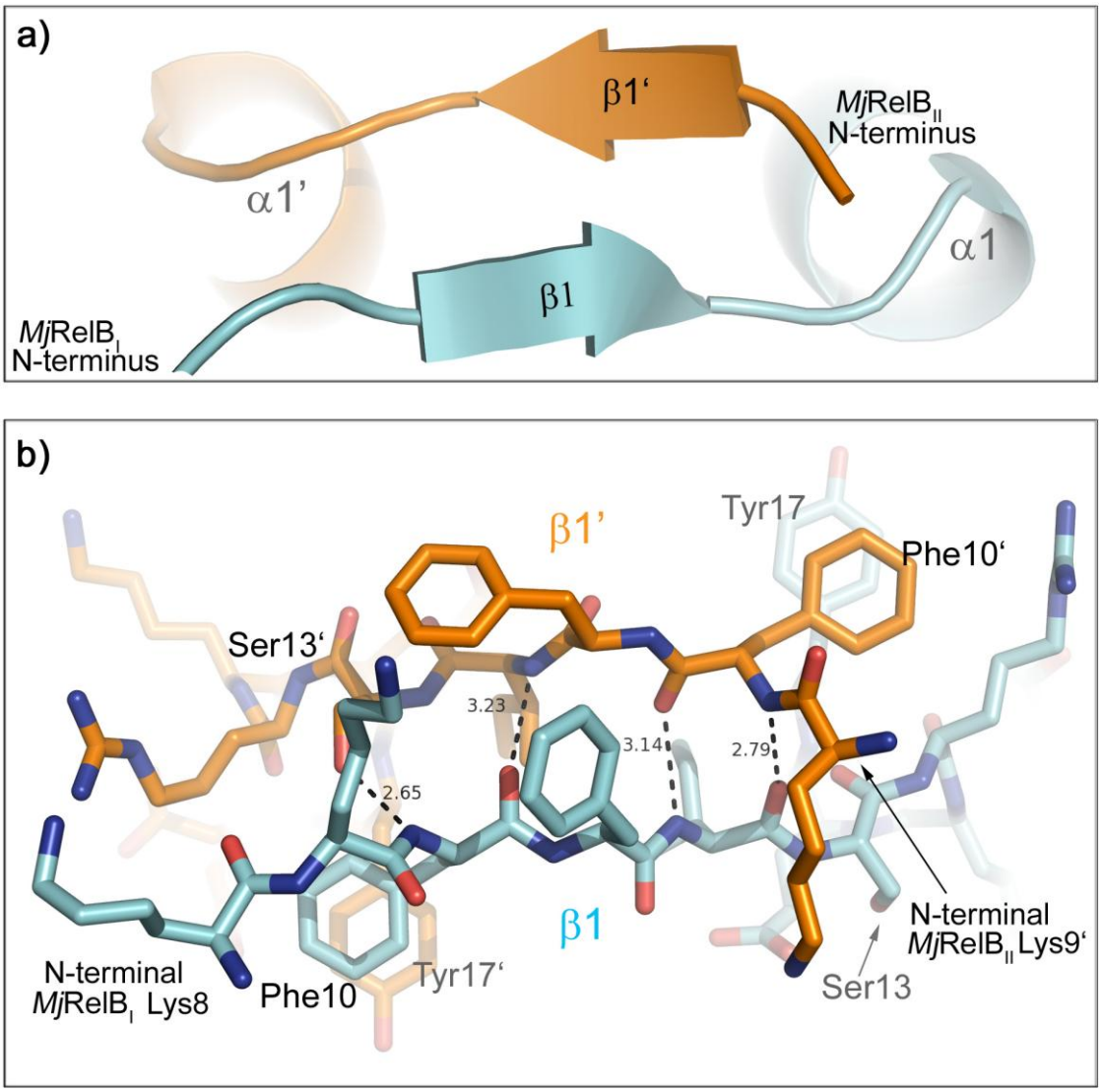

Figure 22 The *Mj*RelB$_{I}$-*Mj*RelB$_{II}$ β-sheet. Heterodimer I antitoxin (*Mj*RelB$_{I}$) is coloured blue (residues marked with ‘) and heterodimer II antitoxin (*Mj*RelB$_{II}$) orange. Hydrogen bonds and their lengths (in Angstroms) are presented with dashed lines. **a)** Cartoon representation of the *Mj*RelB$_{I}$-*Mj*RelB$_{II}$ β-sheet with the N-termini of each chain marked. **b)** Stick representation of the same β-sheet, with the π-stacking Tyr17 and Phe10.

However, the interactions between the heterodimers are not symmetric as it would be expected. Since the chains do not overlap entirely when the NCS operator is applied, so are some of the interactions between the chains slightly different or even unique in this particular structure. The differences in interactions are discussed in section 3.3.5.

3.3.3 Temperature Factors (B-factors)

Although the data range is up to 2.2 Å resolution, the definition of the electron density is not in all parts of the model of the same quality. In all crystals there are small imperfections and the corresponding amino acids are not placed at exactly the same position inside its unit cell as compared to residues from other unit cells. This causes the electron density to be "smeared" in space as it, the electron density, represents the average position of each atom in the unit cell through all the unit cells in the crystal. It is the imperfections that cause the main distortions and poorly defined regions of the electron density. The cause for these imperfections is the flexibility of certain parts of the chain that can assume different conformations in each of the unit cells of the crystal. What we observe in the X-ray experiment is the average of all the conformations which may result in a poorly defined electron density or even total absence of the electron density in certain fragments of the polypeptide chain. Often there are alternative conformations for amino acid residues found in proteins. In the crystal the residues adopt one or the other conformation which are energetically almost equal. In these cases we describe the atomic positions and the movement of the residue not by the temperature factor but with another factor, the occupancy. These two parameters are closely related as they both describe the disordering of atomic positions. In higher resolution structures (better than 1.9 Å) the temperature factor is expressed as "anisotropic B-factor" because the high resolution of data allows calculating the different intensities of the B-factor in different directions (ellipsoids). If the occupancy would not be taken in to account, the anisotropic B-factor would show high anisotropy pointing towards the alternative position of the atom.

The high overall B-factors in the structure of *Mj*RelBE, 50.04 $Å^2$, show that this structure has problems with local atom positional disorder and with highly flexible parts (Figure 23).

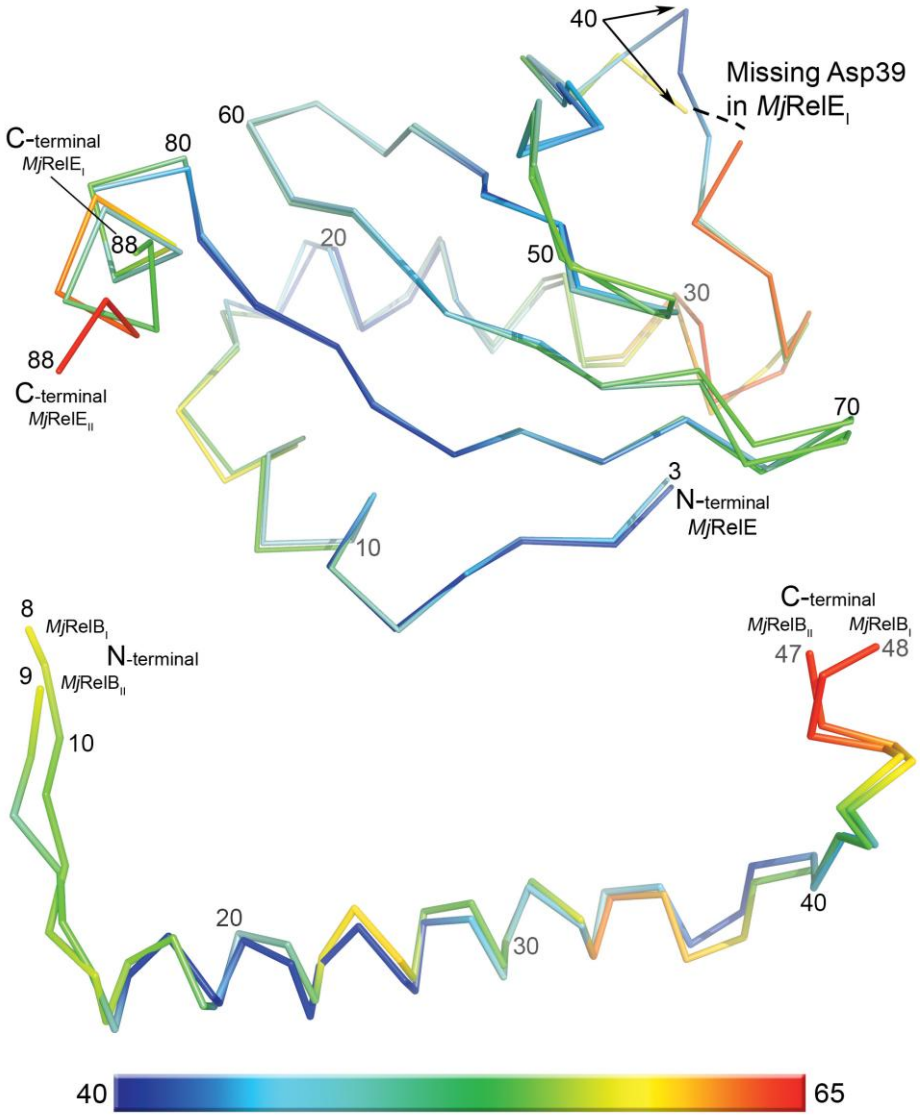

Figure 23 Ribbon representation of *Mj*RelE and *Mj*RelB coloured by B-factors of the main chain atoms. The termini and every tenth residue are numbered. Chains from heterodimers I and II are overlaid on each other for comparison. The bar at the bottom of the figure represents the colour-scale of the B-factors ranging from 40-65 Å^2. The C- and N-termini are denoted and the chain breake in *Mj*RelE$_I$ that is due to missing electron density. The highest B-factors, and therefore highest flexibility, of the chains are at the C-termini and in the flexible loop of *Mj*RelE$_I$ (residues 35-43).

Areas with the highest B-factors are the C-termini of all chains where the B-factors reach 62 Å^2, and the flexible loop of *Mj*RelE$_I$ where the temperature factors testify of a higher mobility of these residues.

The electron density in these areas is less well defined and one residue of the toxin in heterodimer I, Asp39, could not be fitted to it (Figure 24). This residue is at a place where parts of the chain with very different temperature factors meet.

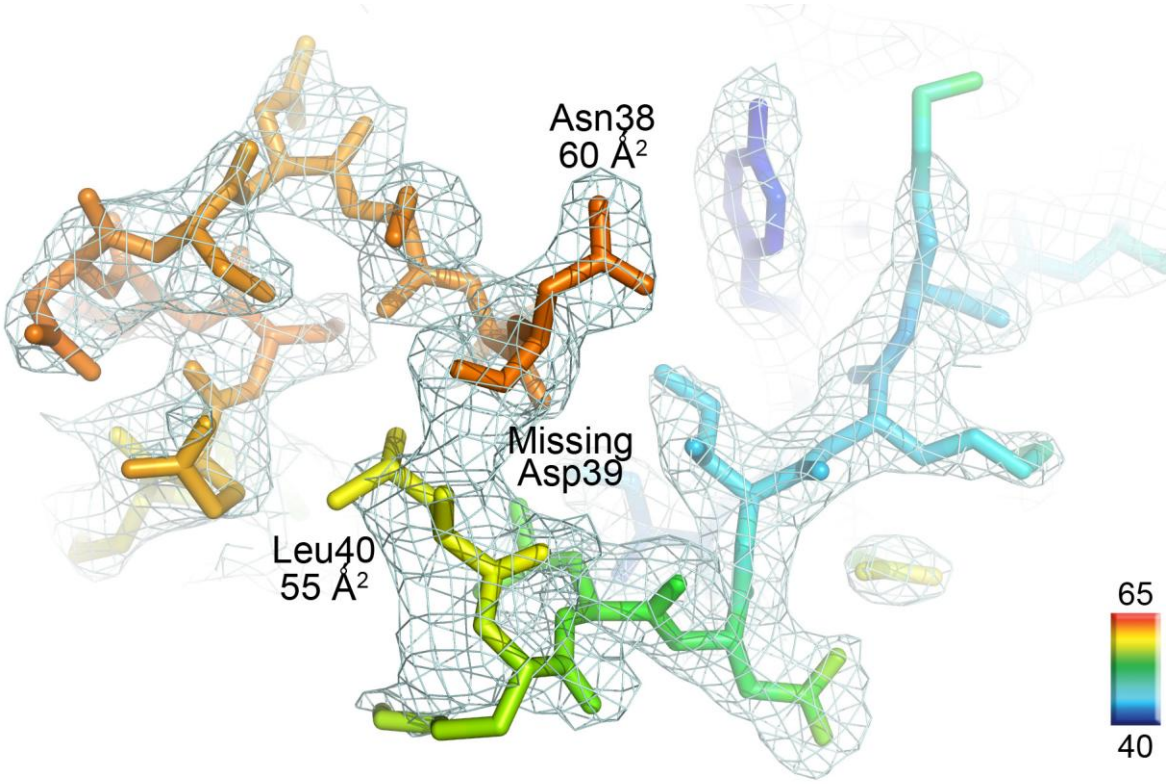

Figure 24 The *Mj*RelE$_I$ flexible loop with the missing residue Asp39 indicated. Atoms are coloured according to their B-factors. The neighbouring residues B-factors for the Cα atoms are higher than the overall B-factor. The flexible loop starts at residue Asn35 (with B-factor 58.6 Å^2) and ends at residue Lys45 (with B-factor 47.1 Å^2). Electron density carved at 1 sigma.

The loop obviously adopts different conformations and the modelling of it was exceptionally difficult. In *Mj*RelE$_I$ it was possible to model the loop in different conformations but in none of them all of the residues could be built. Despite calculations of an omit map, the loop seemed correctly built in different positions. Nevertheless, the best fit, slightly better than the other possibilities, was achieved with the conformation deposited in the PDB file.

The interesting fact is that the α2-β2 loop in the second heterodimer *Mj*RelE$_{II}$ has lover B-values, even lower than the neighbouring residues. In this chain, *Mj*RelE$_{II}$, the residue missing in heterodimer I is present and the whole loop is well defined.

Another obvious difference in B-factors between the chains is within the C-termini of *Mj*RelE where the last few residues of *Mj*RelE$_I$ are better defined than the ones in the second toxin chain.

The C-termini of *Mj*RelB from both heterodimers are extruding out of the bulk of the protein into the solvent. Although the antitoxin protein was present in full length prior to crystallization, the last few residues could not be modelled due to poor definition of the

electron density in these regions. The B-factors of the placed residues in the C-terminal parts of the antitoxins show how these parts become disordered as further the chain separates from the rest of the protein complex. The opposite effect can be seen in the α1 helix of the antitoxins that are gripped by the toxin. This helix is constrained from movement and the accompanying B-factors are therefore lower than in the rest of the chain. For the toxin chains *Mj*RelE, the best defined regions, and simultaneously the regions with the lowest B-factors, are the β-sheets. The β-sheet is positioned in the centre of the *Mj*RelBE protein complex and therefore the best defined region.

The causes for differences in B-factors between the corresponding regions of the two NCS related heterodimers are the different crystal contacts, as will be described in the following sections 3.3.4 and 3.3.6.

3.3.4 Crystal Contacts in the *Mj*RelBE Protein Crystals

As in almost all protein crystals the volume occupied by solvent atoms and by protein atoms is about 50/50. In this particular case the Matthews coefficient (that describes the protein to solvent ratio) is 2.67 $Å^3$/Da, which corresponds to 54.04% solvent content. This means that most of the protein surroundings in the crystal are water and other solvent molecules (buffer, salt, precipitants used for crystallization, molecules of the cryo solution…), just as it is in solution. The crystal packing of the *Mj*RelBE crystals allows for crystal contacts with six molecules, four of which are unique contacts. The full contact listings can be found in Appendix 7.5.4.

The crystal contacts are not symmetrical with respect to the NCS related heterodimers. This is one of the reasons why there are some differences between the heterodimer chains. There are two particular examples of these differences caused by interactions with neighbouring heterotetramer molecules in the crystal. The first is the difference in the α2–β2 flexible loop. As shown in Figure 23, there are differences in temperature factors between the two *Mj*RelE chains that cause the poor definition of electron density of the loop in *Mj*$RelE_I$, as shown in Figure 24. When the surroundings of the loops are examined, it is apparent that the regions of crystal contacts involve these two loops and their surrounding but that this interaction is not symmetrical. The heterodimer II toxin loop is hydrogen bonded to a symmetry- related neighbour that helps the loop to adopt a unique

conformation. There are three hydrogen bonds formed, as can be seen in the contact tables and the Figure 25a. This influences both the loop in *Mj*RelE$_{II}$ and the residues of the β-strand β2 of *Mj*RelE$_{I}$ in the way that their B-factors are lover then the ones of the neighbouring residues. This, however, is not the case with the heterodimer I toxin α2–β2 flexible loop. In this loop the surroundings are mainly solvent molecules and therefore the loop is free to move around, hence the poor definition of the electron density. The only anchor points are the ending points of the loop and a weak hydrogen bond to the neighbouring toxin molecule *Mj*RelE$_{II}$. In the predominant conformation that has been modelled in this structure, the interaction between the *Mj*RelE$_{I}$ loop and the crystal neighbour is between Asn38 of *Mj*RelE$_{I}$ and Asp39 of the *Mj*RelE$_{II}$. This weak hydrogen bond is not enough to stabilize the loop of *Mj*RelE$_{I}$ but it straightens the already stabilized loop from the toxin *Mj*RelE$_{II}$ of heterodimer II.

A similar effect is observed in the C-terminal part of the toxin chains. This is the second example of crystal contact influence on the temperature factors in this structure. The terminal carboxylate of *Mj*RelE$_{I}$ Pro88 forms a salt bridge with Lys22 sidechain of the neighbouring *Mj*RelE$_{I}$ chain (Figure 25b) and at the same time forms another salt bridge with the second carboxyl oxygen to the sidechain of Lys30 of *Mj*RelB$_{I}$ from its own heterodimer. On the other hand Pro88' residue from heterodimer II has no crystal contacts and is pointing out from the protein into the solvent. This difference in crystal contacts is reflected in the B-factor values for the last five C-terminal residues of *Mj*RelE being up to 10 Å^2 higher in heterodimer II than in heterodimer I.

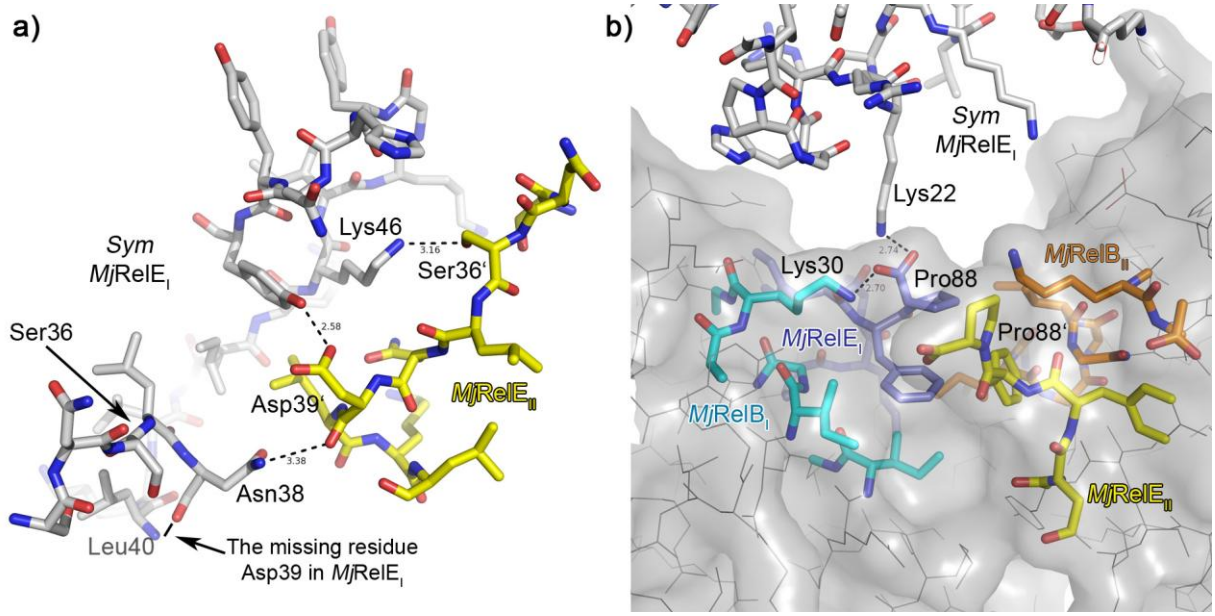

Figure 25 Crystal contacts. Chains of heterodimer I are coloured blue (*Mj*RelE$_I$) and cyan (*Mj*RelB$_I$), while the chains of heterodimer II are coloured yellow (*Mj*RelE$_{II}$) and orange (*Mj*RelB$_{II}$). Symmetry related crystal contact is coloured grey. **a)** The α2-β2 loop and the influence of crystal contacts on its flexibility. The *Mj*RelE$_{II}$ is stabilized by the crystal contact while *Mj*RelE$_I$ is freer to move. **b)** The C-termini of *Mj*RelE in the middle of the heterotetramer. *Mj*RelE$_I$ Pro88 is connected by a hydrogen bond to a symmetry related heterotetramer which stabilizes the proline position and enables a conformation in which the hydrogen bond formation to *Mj*RelB$_I$ Lys30 is possible. There are no neighbouring heterotetramer molecules in the surroundings of the *Mj*RelE$_{II}$ Pro88' and it is pointing out of the bulk of the protein complex, associated with. Therefore also the higher B-factors as can be seen in Figure 23.

3.3.5 Asymmetric Interactions Between Heterodimers I and II

In order to understand the structural differences it is prudent to first describe the resulting changes in the interactions and then to explain the cause of this asymmetrical behaviour between the heterodimers. This chapter will explaine effects that the chain geometries have on the heterotetramer formation and in later chapters the question why the chains in the heterodimers have adopted slightly different conformations will be answerd.

One interesting difference in heterodimer interactions is the *Mj*RelB Glu16 to *Mj*RelE Arg80 salt bridge. The packing between the chains in the C-terminal part of *Mj*RelE$_I$ and the beginning of helix α1' of *Mj*RelB$_{II}$ is very tight and the hydrogen bonds are directly interlinking the chains. In figure Figure 26 the difference in the interactions of the C-terminal part of RelE's with RelB's are depicted. The interaction between heterodimer I toxin and heterodimer II antitoxin is straight through residues Tyr84 and Arg80 of the toxin, with Glu16

of the heterodimer II antitoxin (Figure 26a). On the other hand, the interaction of the heterodimer II toxin with the heterodimer I antitoxin is entirely bridged by a water molecule positioned in the middle between antitoxin residue Glu16 and the toxin residues Arg80' and Tyr84' (Figure 26b).

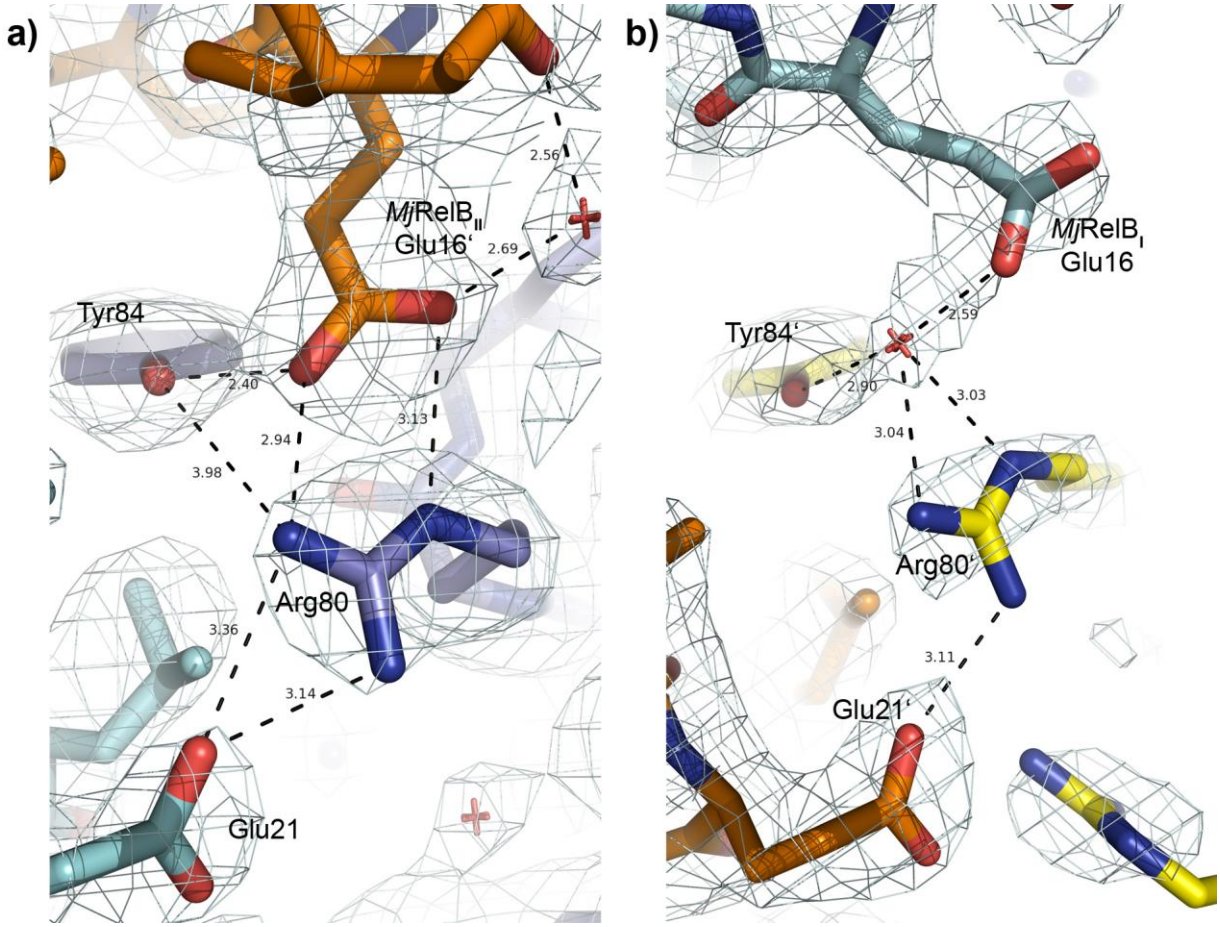

Figure 26 Asymmetric interactions between the heterodimers. Electron density is drawn at 1 sigma level and presented as a mesh around the residues. Possible hydrogen bonds are represented with dashed lines and the distances (in Angstroms) are given next to them. The red crosses symbolize water molecules. Chains of heterodimer I are coloured blue (*Mj*RelE$_{I}$) and cyan (*Mj*RelB$_{I}$), while the chains of heterodimer II are coloured yellow (*Mj*RelE$_{II}$) and orange (*Mj*RelB$_{II}$). **a)** The interaction of heterodimer I toxin with the antitoxin from heterodimer II **b)** Interaction of heterodimer II toxin with heterodimer I antitoxin mediated by a water molecule, unlike the supposed symmetrical interaction described in **a**.

Another difference in heterodimer-heterodimer interaction is the antitoxin Lys19 interaction with the toxin chain of the other heterodimer. As can be clearly seen in the electron density (Figure 27) the side chains of the antitoxin in the heterodimers adopt different conformations resulting in the different patterns of inter-chain interactions. In heterodimer II the *M*jRelB$_{II}$ Lys19' is stretched towards the *Mj*RelE$_{I}$ atoms Asp85NE and Lys81O (Figure 27a). Contrary

to this, Lys19 of the heterodimer I antitoxin is turned away from the heterodimer I toxin and is involved in an intra-chain salt bridge with the carboxyl group of Glu22 (Figure 27b). The discussed differences considered only salt bridges and hydrogen bonds but there are some differences also in the hydrophobic interactions.

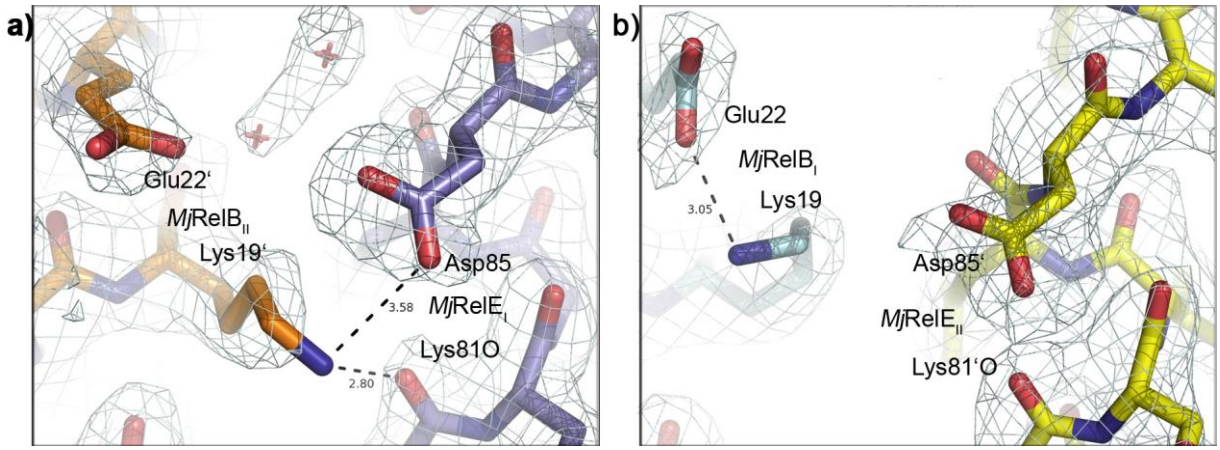

Figure 27 The asymmetric interactions of *Mj*RelB Lys19. Electron density is drawn at 1 sigma level and presented as a mesh around the residues. Possible hydrogen bonds are represented with dashed lines and the distances (in Angstroms) are given next to them. The red crosses symbolize water molecules. Chains of heterodimer I are coloured blue (*Mj*RelE$_I$) and cyan (*Mj*RelB$_I$), while the chains of heterodimer II are coloured yellow (*Mj*RelE$_{II}$) and orange (*Mj*RelB$_{II}$). **a)** The heterodimer II antitoxin residue Lys19' interacts with heterodimer I toxin residues Asp85 and Lys81 to possibly form hydrogen bonds **b)** the heterodimer I antitoxin conformation is inadequate for interaction with the heterodimer II toxin, the closest residue to heterodimer I residue Lys19 is Glu22 of the same chain.

The antitoxins of the two heterodimers interact with each other with their N-terminal parts (residues 8-24) by forming a mutual β-sheet with their β1 strands. Further they interact by hydrophobic contacts on the N-terminal parts of their α1 helices (residues 14-24).

There are also differences in the interactions between the two antitoxins that form the centre of this hetero-complex. As mentioned earlier (section 3.3.2) at the ends of the β1 strands, forming the mutual *Mj*RelB$_I$-*Mj*RelB$_{II}$ β-sheet, the Ser13' hydroxyl group, from *Mj*RelB$_{II}$, is involved in a hydrogen bond towards the peptide bond nitrogen of Phe10 from *Mj*RelB$_I$. However, this is not the case with Ser13 of *Mj*RelB$_{II}$ and Phe10' of *Mj*RelB$_{II}$.

3.3.6 Differences Between *Mj*RelBE Heterodimers I and II (Summary)

The two *Mj*RelBE heterodimers forming the heterotetramer are related by a non-crystallographic twofold symmetry axis and are consequently not entirely identical (Figure 23). The Cα traces correspond very well in the core parts of the proteins while in the flexible loops and terminal regions the temperature factors and conformation of the residues are governed by crystal contacts. In both toxin molecules *Mj*RelE$_I$ and *Mj*RelE$_{II}$, the flexible loop regions, residues 35-43, and the C-terminal regions, residues 83-88, show larger discrepancy in position and B factors that differ in these regions up to 10 Å^2, with molecule I being better defined in the loop region and molecule II better defined in the C-terminus. The loop 35-43 in *Mj*RelE$_{II}$ is better defined due to crystal contacts with residues 44-51 of a symmetry related molecule *Mj*RelE$_I$ whose loop 35-43 is more flexible because it is exposed to solvent. In a similar manner the C-terminus of *Mj*RelE$_I$ is restrained by crystal contacts with residues around Lys22 of a symmetry related *Mj*RelE$_I$ while the C-terminus of *Mj*RelE$_{II}$ is more solvent exposed.

As previously mentioned, another difference among the two NCS related heterodimers is the *Mj*RelB Glu16 to *Mj*RelE Arg80 interaction in heterodimer II that is mediated by a water molecule. This water molecule is in the centre of a hydrogen bond network formed between *Mj*RelB$_I$ (residue Glu16) and *Mj*RelE$_{II}$ (residues Arg80 and Tyr84) while in the other heterodimer, this water molecule is missing.

The interaction of *Mj*RelE$_I$ C-terminal Pro88 and *Mj*RelB$_I$ Lys30 is not found in heterodimer II as can be seen in Figure 25b. This is probably due to the stabilization effect of the symmetry related chain as explained above.

As showed above, the interactions, being intra-chain, inter-chain, heterodimer-heterodimer or even crystal contacts, can influence the structure, the temperature factors and finally the degree in how well a residue is defined in the crystal structure. We can separate this into interactions that are most probably present in solution and do influence the biological behaviour of the proteins, and to the interactions that are solely crystallization artefacts as are the interactions with the symmetry related molecules from the crystal. It is to be taken into account that some interactions are actually made possible through the presence of neighbouring molecules in the crystal lattice as they restrict movement of some residues or stabilize some residues by hydrogen bond formations. This is seen in the interaction of *Mj*RelE$_I$ Pro88 and *Mj*RelB$_I$ Lys30 that would hardly be possible if the proline had not been forced by a crystal contact to adopt a favourable conformation for hydrogen bond formation

with Lys22 of *Mj*RelB$_I$. To simplify this, it is enough to outline only the effects that are crystallization artefacts. First and outmost is the behaviour of the flexible loop α2-β2 in *Mj*RelE. The real conformation of this loop in solution is hardly constrained to one particular position as it does not form any hydrogen bonds with the rest of the more rigid part of the protein. Starting from Lys41 to Ser36 there is only one possible main-chain to main-chain hydrogen bond (Asn38 to Leu42) and the hydrophobic interaction of Leu37 side-chain. This is not enough to prevent flipping of the side-chains and movement of the main-chain.

3.3.7 Surface Charge of the *Mj*RelBE Heterotetramer

The charge distribution on the surface of proteins can be taken as an indicator for what binds to the protein and how. On the surface of *Mj*RelE positively charged regions can be observed (Figure 28a) that largely coincide with the binding cleft where the negatively charged antitoxin *Mj*RelB (Figure 28b) is positioned. This suggests that the cleft must be the part of the toxin that interacts with the target molecule. As it is already known that RelE cleaves mRNA, the positively charged cleft seems to perform a perfect fit for RNA binding.

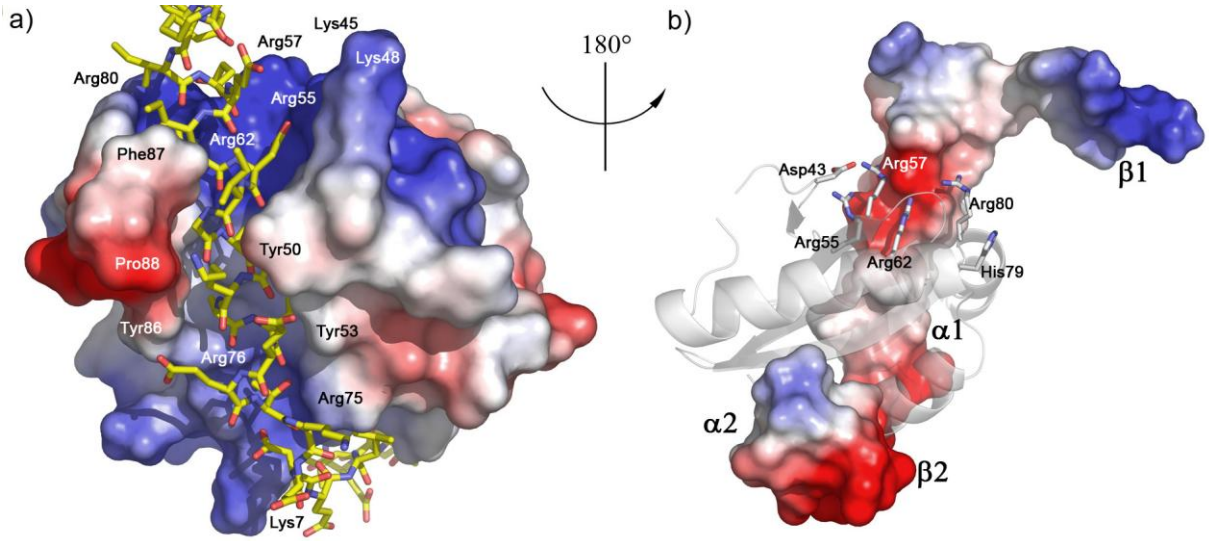

Figure 28 Surface representation. **a**) of *Mj*RelE with α1 of *Mj*RelB (yellow stick representation, O red, N blue) located in the deep and positively charged cleft of *Mj*RelE. Negative (red), positive (blue) and neutral (white) charges are indicated. The arginines on the far ends of the cleft, Arg55, Arg57, Arg62, Arg80 on the top side and Arg75, Arg76 on the bottom render the cleft positively charged. The tyrosines (Tyr50, Tyr53 and Tyr86), Pro88 and Phe87 that form both sides of the cleft are labeled. **b**) Surface electrostatics of the antitoxin *Mj*RelB and cartoon and stick representation of *Mj*RelE, grey. Secondary structure elements *Mj*RelB are marked.

Apart from the positive charge of the cleft, the residues at the sides like Phe87, Tyr53, Tyr50 and Pro88, could also take part in the binding of RNA ribose bases.

3.4 The relation of *Mj*RelBE to Other TA Systems

In previous sections the properties of *Ec*RelBE and properties and structure of *Mj*RelBE have been discussed. To understand the real implications of these, it is necessary to compare the known structures of TA systems with each other. The structure of the RelE family member *Ph*RelBE from *Pyrococcus horikoshii* should be almost identical, as they share the same family characteristics. This is, however not entirely so, as described in the following section.

3.4.1 Comparison of the *Mj*RelBE and *Ph*RelBE Structures

The antitoxin *Ph*RelB of *P. horikoshii* is longer (67 residues) than *Mj*RelB (52 residues), and its helix α2 and not the very long helix α1 is positioned in the cleft of the toxin *Ph*RelE (Figure 29). This is in constrast to *Mj*RelB that is bound with helix α1 to *Mj*RelE.

Starting from the C-termini of the antitoxins, they share similar features. They both have a short C-terminal helix (α2 in *Mj*RelB and α4 in *Ph*RelB) followed by a short β-strand (β2 in *Mj*RelB and β0 in *Ph*RelB) that participates in β-sheet formation with the toxin β1 strand. From there towards the N-termini the Cα traces of the antitoxins take different paths. *Ph*RelB forms a short helix α3 (not present in *Mj*RelB) and continues to a longer helix α2 positioned within the cleft formed by toxin *Ph*RelE that is wider than the cleft in *Mj*RelE.

Towards the N-termini the polypeptide chains of the antitoxins take opposite directions; the *Mj*RelB helix α1 is much longer than the cleft of *Mj*RelE and ends with N-terminal strand β1 whereas helix α2 of *Ph*RelB ends within the cleft and is connected by a loop with the long N-terminal helix α1.

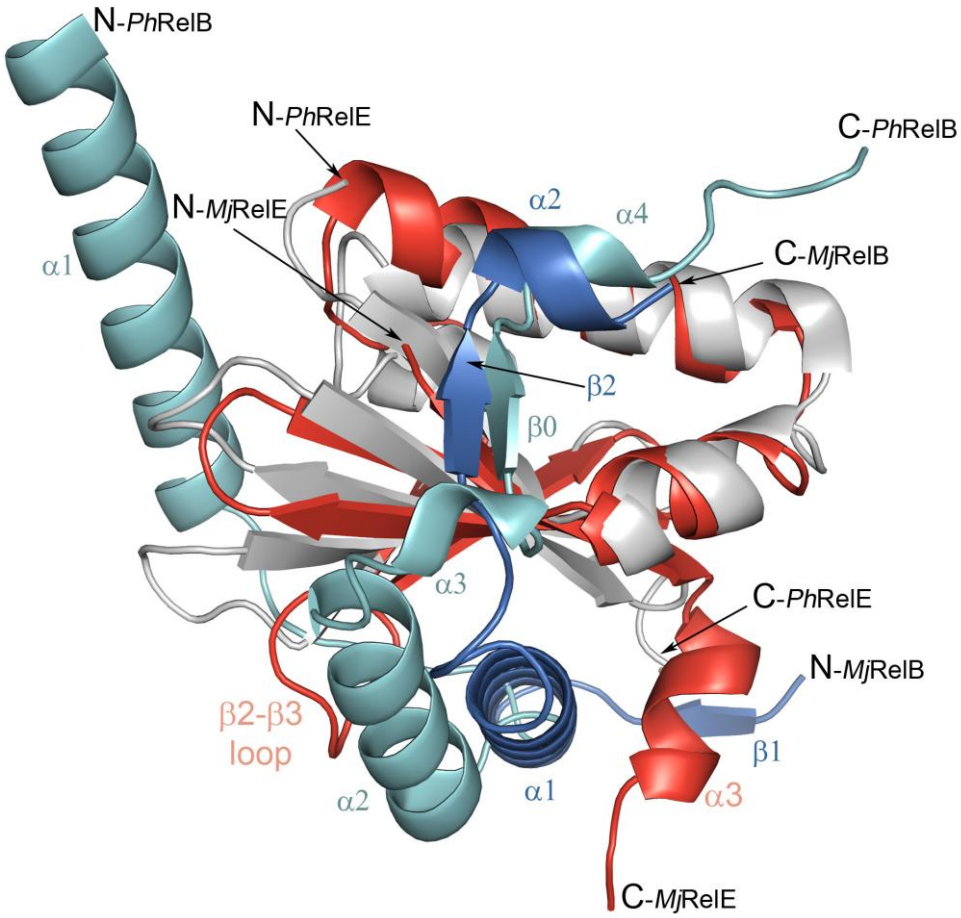

Figure 29 Superposition of the *Mj*RelBE (heterodimer II) and *Ph*RelBE heterodimers. The *Mj*RelE is coloured red and *Mj*RelB blue, *Ph*RelE is coloured grey and its antitoxin *Ph*RelB cyan. The antitoxin of *M. jannaschii* is tighter gripped by its toxin. β0 of *Ph*RelB and β2 of *Mj*RelB as well as α4 of *Ph*RelB and α2 of *Mj*RelB overlap fairly well while the rest of the antitoxin chains differ radically in their paths of wrapping around the cognate toxins.

The folds of the RelE toxins are fairly well conserved with root mean square deviation (r.m.s.d.) of 1.3 Å , the main differences occurring in the position of loop β2-β3 and helix α3 that form the cleft of *Mj*RelE while in *Ph*RelE the cleft is shallow and formed only by the bent β-sheet (Figure 29). In the *Mj*RelBE heterotetramer the contact between the two heterodimers is formed between the antitoxins and there is no $MjRelE_I$ – $MjRelE_{II}$ interaction, an arrangement that contrasts *Ph*RelBE where the two *Ph*RelE toxins share a common interaction surface (Figure 30).

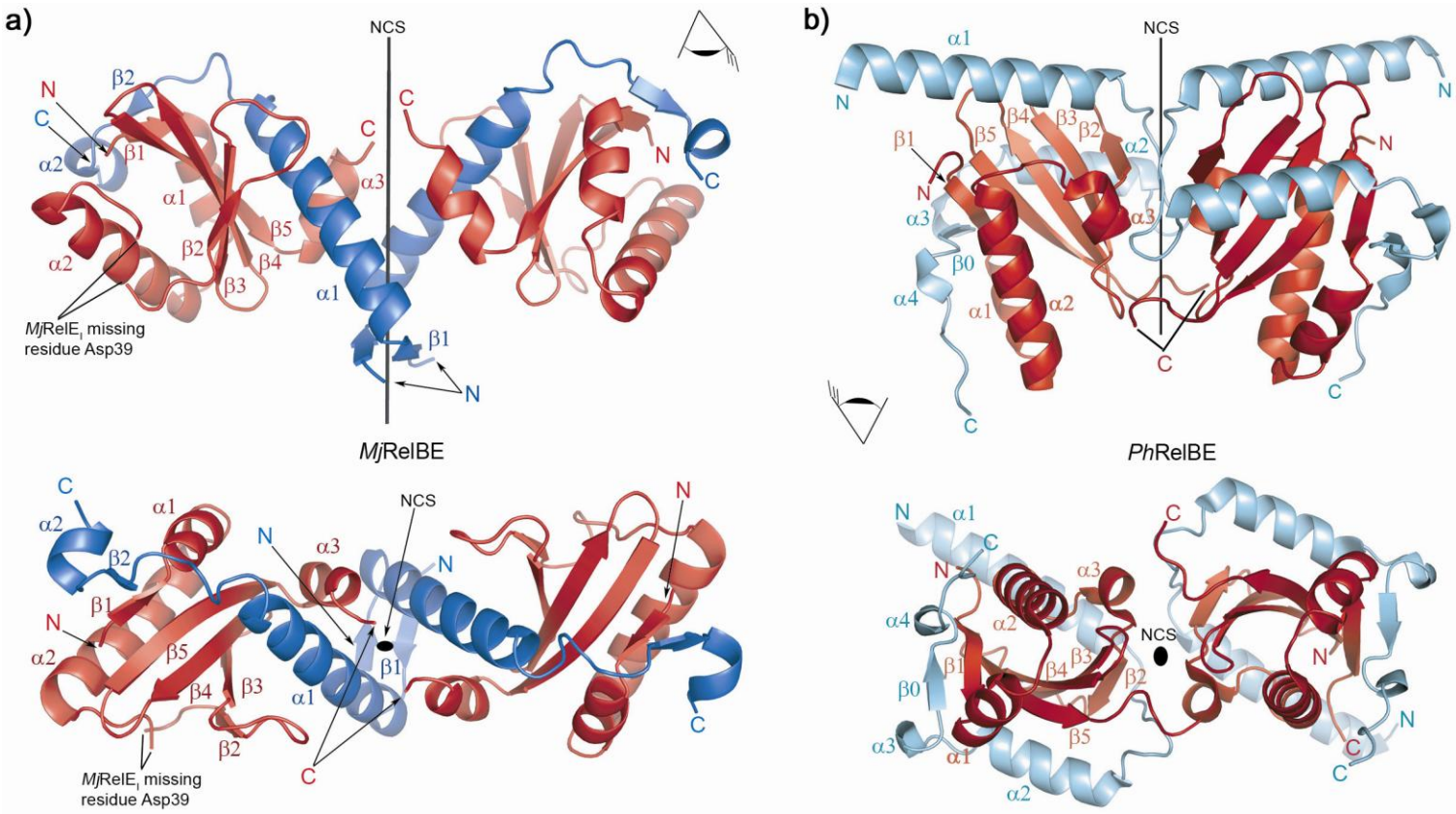

Figure 30 Comparison between the *Mj*RelBE and *Ph*RelBE heterotetramers. **a)** Cartoon representation of the *M. jannaschii* *Mj*RelE (red) *Mj*RelB (blue) heterotetramer. Helices α1 and α2 of *Mj*RelE form the U-shaped base on which the β-sheet is placed. The long helix α1 of the antitoxin *Mj*RelB is tightly gripped by the antiparallel β-sheet, C-terminal helix α3 and loop connecting β2 and β3 of *Mj*RelE. The rest of the antitoxin is tightly wrapped around the toxin. The $MjRelB_I$ strand β2 is involved in intermolecular β-sheet formation with $MjRelE_I$ and the $MjRelB_I$ strand β1 forms an intermolecular β-sheet with β1' of $MjRelB_{II}$, the antitoxin from the second heterodimer related by non-crystallographic symmetry (NCS) to $MjRelBE_I$. The heterodimer-heterodimer interaction is primarily stabilized through *Mj*RelB-*Mj*RelB interactions. **b)** The *P. horikoshii* antitoxin *Ph*RelB (cyan) wraps around the toxin *Ph*RelE (red). The *Ph*RelBE heterotetramer form is very different from that of the *Mj*RelBE heterotetramer.

3.4.2 Comparison to Other TA Toxins

So far, the three-dimensional structures of εζ, CcdAB, MazEF, Kis/Kid, YefM/YoeB and RelBE toxin-antitoxin systems have been determined. When the known structures are compared to each other the family property of each toxin is obvious. The structural properties of the CcdB and ζ toxin superfamilies are discussed in the introduction and the structures of the RelE toxin superfamily are elaborated in the Introduction and Result chapter. A structure of the structurally yet undetermined Doc toxin superfamily is awaited, as the crystallization of the inactive variant of Doc has succeeded (Garcia-Pino *et al.*, 2008).

From all the notes in the chapters above, it can be seen that the same fold does not definitely lead to the same target molecule preference. This can be illustrated by the example of MazF and CcdB which share the same CcdB toxin superfamily fold but not the same target molecule specificity (described in the Introduction chapter Table 3).

Figure 31 depicts the representatives of the various TA system families. The MazF toxin superfamily has three structural representatives which fit very well to each other. The RelE superfamily is represented with two members, the RelE toxin family (represented by *Mj*RelBE heterodimer) and YoeB toxin in complex with its antitoxin YefM. The last one is interesting because the N-terminal part of the antitoxin resembles the fold of the antitoxin from another superfamily – the Phd superfamily antitoxins (Gazit and Sauer, 1999).

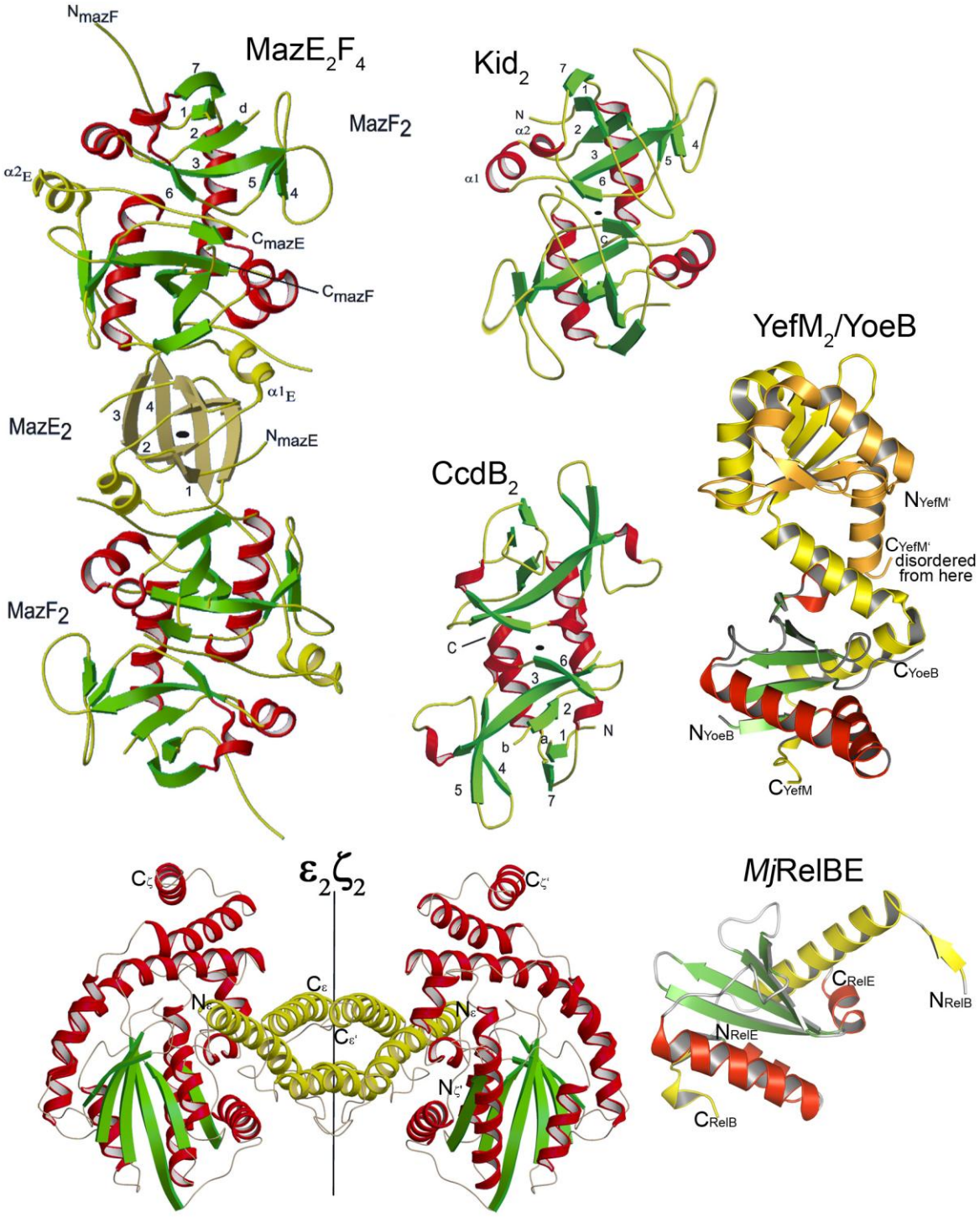

Figure 31 The known structures of TA systems. The toxins are coloured according to the secondary structure elements, green (β-strands) and red (α-helices). The antitoxins are thoroughly coloured yellow. In the upper panel the structures of the members of the CcdB superfamily are shown. The MazEF heterohexamer (MazF2-MazE2-MazF2) on the left is compared to Kid and CcdB dimers. The toxin configuration is identical as can be shown when the numbered secondary elements are compared. In the lower panel the representatives of the εζ and RelE superfamilies are depicted. YefM/YoeB shares the N-terminal part of the antitoxin fold with the Phd superfamily antitoxins but the toxin fold of RelE superfamily (see Table 1).

3.4.3 Relation of the *Mj*RelBE Structure to its Superfamily Neighbors and Their Relation to Endoribonucleases

*Mj*RelE from *M. jannaschii* is the second structurally characterized member of the RelE toxin superfamily. The structure of *Mj*RelBE differs from *P. horikoshii Ph*RelBE in the heterodimer arrangement, the depth of the antitoxin binding cleft and the positions of some crucial residues that are similar in *Mj*RelE and YoeB but not in *Ph*RelE. Structure and amino acid sequence superimpositions of the RelE toxins from *M. jannaschii*, *P. hiroschii*, *E. coli*, and toxin YoeB shows that there are several highly conserved residues, Arg57, Arg62 (essential for toxic function), Tyr84 and Asp43 (Glu46 in YoeB) (RelE numbering, Figure 32 and Figure 33). Other residues are type conserved.

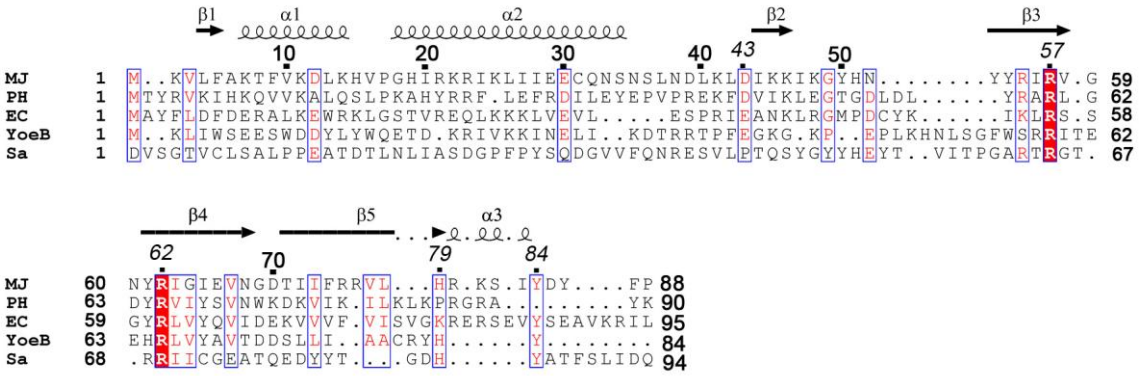

Figure 32 Sequence alignment of *Mj*RelE (MJ), *Ph*RelE (PH), *Ec*RelE (EC), YoeB (YoeB) and RNase Sa (Sa). Conserved residues in all sequences are coloured red and other conserved residues in red and white. Secondary structure elements of *Mj*RelE are indicated above the sequences.

These toxins share low sequence identity but a common fold (Table 8). The r.m.s.d of the superimposed Cα atom positions obtained by the TOP program (Lu, 2000) is 1.3Å for *Mj*RelE - *Ph*RelE (26.7 % sequence identity for 60 overlapping residues), 1.8 Å for *Mj*RelE - YoeB (13.4 % sequence identity for 67 overlapping residues) and 1.5 Å for *Ph*RelE - YoeB (15.2 % sequence identity for 66 overlapping residues). The r.m.s.d for the superimposition of the toxins with fold-related ribonucleases is 1.8 Å for *Mj*RelE - RegB (34 residues overlapping), 1.9 Å for *Mj*RelE - RNase Sa (33 residues overlapping) and 1.5 Å for *Mj*RelE - ColicinE5 (31 residues overlapping).

Table 8 Structure superimposition quality of RelE toxin superfamily

Toxin	Sequence identity (%)*	Overlapped residues	r.m.s.d. (Å)	Compared to toxin
*Mj*RelE	26.7	60	1.3	*Ph*RelE
*Mj*RelE	13.4	67	1.8	YoeB
*Ph*RelE	15.2	66	1.5	YoeB
*Mj*RelE	14.7	34	1.8	RegB**
*Mj*RelE	3.2	31	1.5	ColicinE5

*with the matched chain region
**for RegB the best fit model was compared as the structure was determined with NMR

The conserved residues encompass Arg55, Arg57 and Arg62 on the surface of the central β-sheet of *Mj*RelE (Table 9, Figure 33). Mutational studies (Kamada and Hanaoka, 2005; Takagi *et al.*, 2005) are confirming that the positively charged cleft of the toxin is essential for the toxic effects of RelE and its homologs.

Table 9 Conserved residues in the RelE toxin superfamily and endoribonucleases.

***Ec*RelE**	***Ph*RelE**	***Mj*RelE**	**YoeB**	**RNase Sa**
Glu40	Asp44	Asp43	Glu46	Glu54
	Arg58	Arg55		
		Arg57	Arg59	Arg65
Arg56	Arg65	Arg62	Arg65	Arg69
		His79	His83	His85
Arg81	Arg85	Arg80		
Tyr86		Tyr84	Tyr84	Tyr86

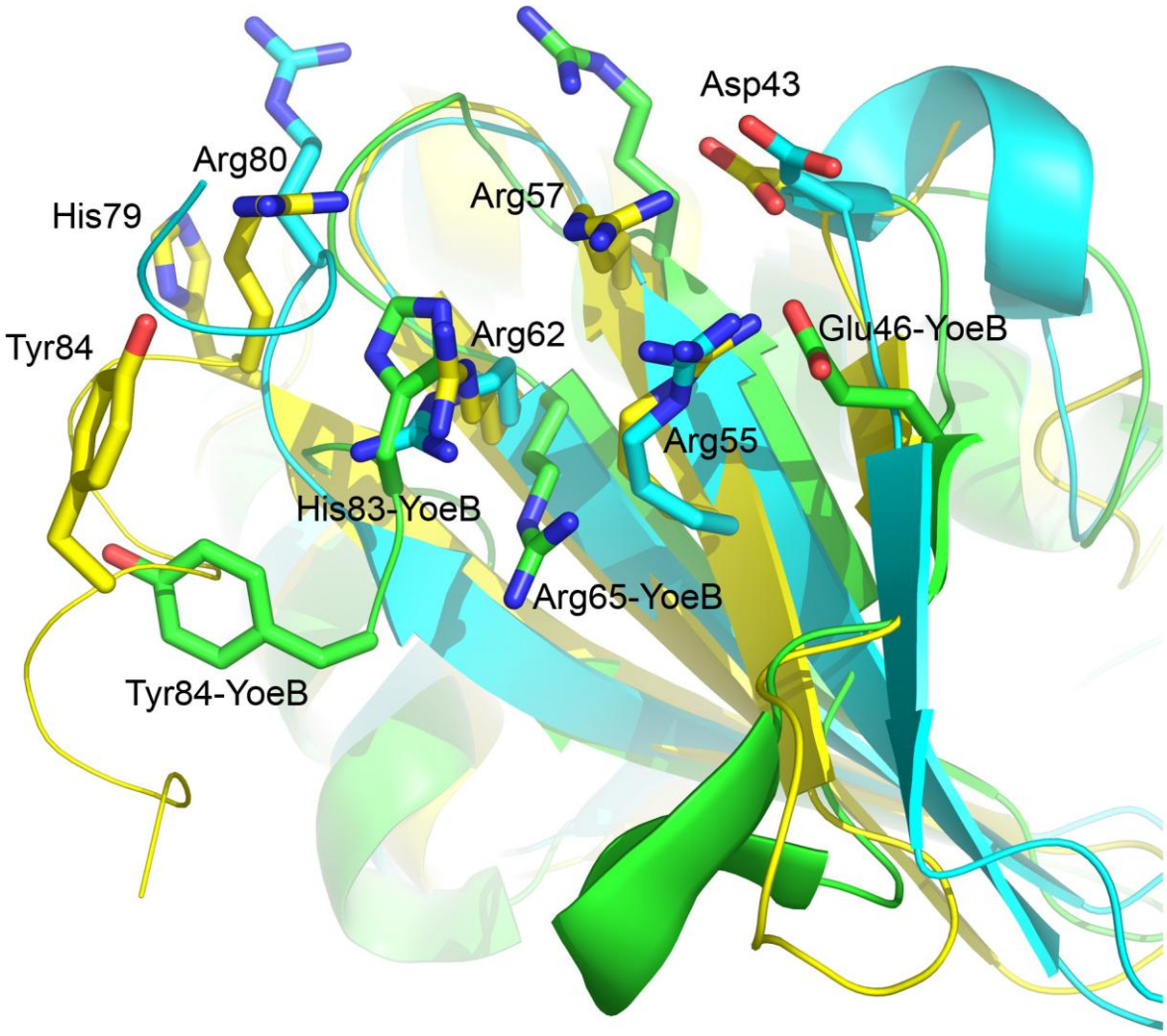

Figure 33 Superposition of the related toxin/ribonuclease structures. *P. horikoshii Ph*RelE (blue), *M. jannaschii Mj*RelE (yellow) and YoeB (green). The residues likely to be involved in RNA binding and hydrolysis are in stick representation and numbered by *Mj*RelE numbering if not otherwise noted. All three structures have Arg62 at the same place while Glu46 of YoeB is replaced by nearby Asp43 in *Ph*RelE and *Mj*RelE. His79 is moved out from the cleft in *Mj*RelE and totally absent in the sequence of *Ph*RelE. Tyr84 is also missing in *Ph*RelE but present in *Mj*RelE and YoeB.

One crucial residue is Arg62 in RelE of *M. jannaschii* and Arg56 of *E. coli* (Table 9) that corresponds to Arg65 of YoeB and Arg65 of *Ph*RelE that probably serves as phosphate binding residue and would stabilize the negative charge of the 2',3'- cyclic phosphate intermediate as in ColicinE5 (Lin *et al.*, 2005). The same arginine is also conserved among toxins Txe, YhaV and HigB. In both cases, *M. jannaschii Mj*RelE and YoeB, mutation of this Arg renders the toxin inactive (this work and Kamada *et al.* 2005).

During the course of this research the occurrence of spontaneous mutation in *Mj*RelE of Arg62 to Ser62 or Lys62 inactivated the toxin, and hence it was possible to clone it

alone into another vector and to express it without its cognate antitoxin *Mj*RelB. All trials failed to revert this mutation and to clone wild type *Mj*RelE together with the antitoxin *Mj*RelB. The transformation of cells was never successful even though the vectors were correctly cloned, the insert was present (as confirmed by PCR and endonuclease digestion tests), and the positive controls transformed correctly. We have learned that Gerdes group in Odense/Denmark (personal communication by Susanne Christensen) encountered the same problem. This is a puzzling effect because we had no problems to clone wild type *Ec*RelBE of *E. coli* in expression strains. The toxin *Mj*RelE is evidently more toxic to the cells because even basal expression of the protein caused cell stasis or death.

The mechanism of RNA cleavage by endoribonucleases includes an acid/base catalytic step where the general acid/base pair consists of Glu and His, but it is possible that an Arg/Asp pair may take this catalytic role (Saïda and Odaert, 2007). In the tRNase ColocinE5, the proposed mechanism for RNA cleavage involves Asp46 and Arg48 (Lin *et al.*, 2005) and does not require any His.

Comparison of the structures of YoeB and *Mj*RelE (Figure 33) shows that the residues denoted to be crucial or important for the function as an endoribonuclease are conserved between YoeB and *Mj*RelE. The equivalents would be Glu46, Arg59, Arg65, His83 of YoeB and Asp43, Arg57, Arg62, His79 of *Mj*RelE, respectively, Table 9. His83 and Arg65 of YoeB are likely to be involved in phosphate binding (Kamada and Hanaoka, 2005) and may serve as general acid in the catalytic reaction. The most prominent differences among the conserved residues are the orientations of *Mj*RelE His79 and YoeB His83, and substitution of Glu46 in YoeB by Asp43 in *Mj*RelE (Asp44 in *Ph*RelE) that are located in both RelE at the beginning of strand β2 of the antiparallel β-sheet. In the YoeB structure His83 (equivalent to *Mj*RelE His79) is positioned at 13 Å (Cα to Cα distance) from Glu46 while His79 in *Mj*RelE is positioned 15.2 Å from Asp43 that corresponds to Glu46 in the YoeB structure. Further, toxin YoeB (84 residues) is shorter than *Mj*RelE, and since His83 is at the flexible C-terminus of YoeB, it can be repositioned for binding to the RNA chain. On the other hand, a major conformational change for the equivalent histidine (His79) in *Mj*RelE would be required to occupy the same position as His83 in YoeB. The same applies for Tyr84 of *Mj*RelE located on helix α3 that corresponds to Tyr84 of YoeB.

It is noticeable that His83 present in YoeB and His79 in *Mj*RelE is missing in the sequence of *E. coli* and *P. horikoshii* RelE. Since His83 is essential for the toxicity of YoeB (Kamada and Hanaoka, 2005), this suggests that the *E.coli* and *P. horikoshii* RelE are missing a residue (possibly supplemented by the ribosome) to form an active ribonuclease.

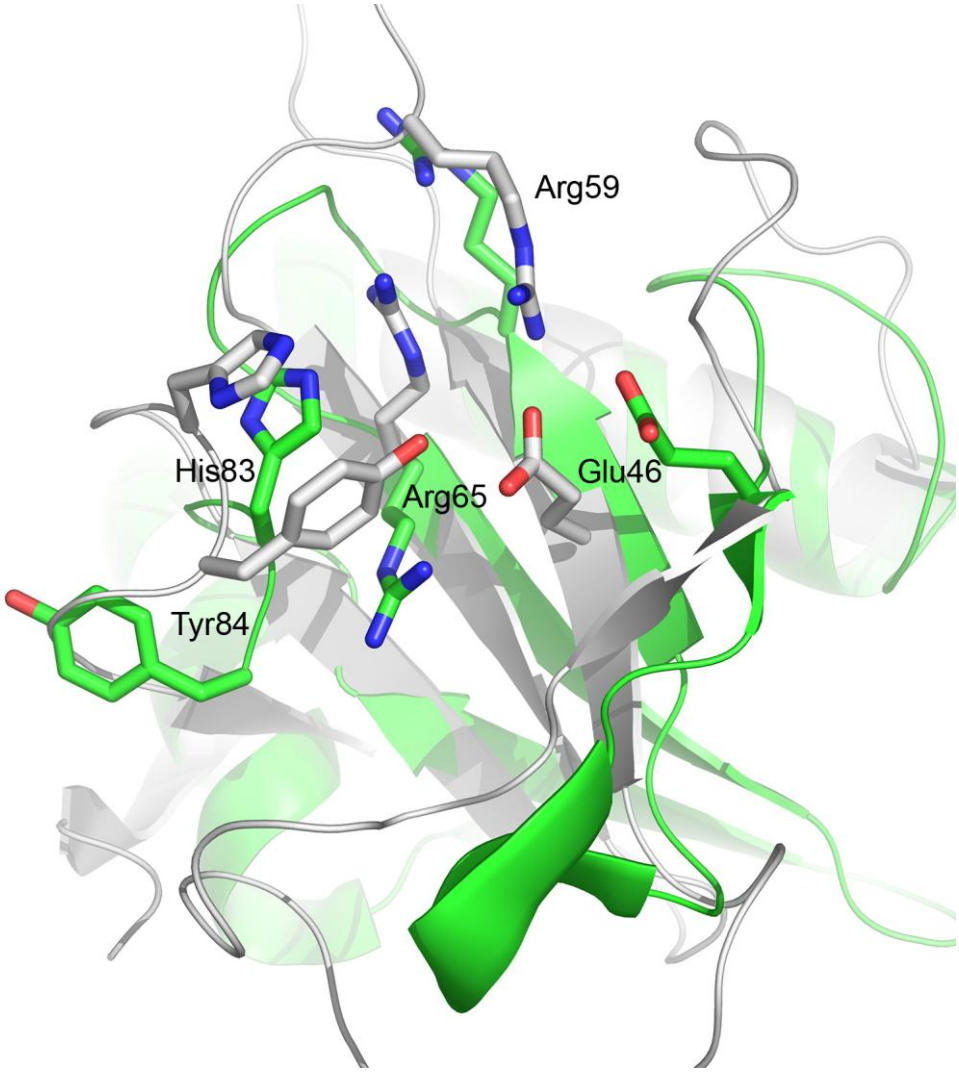

Figure 34 YoeB (green) superimposed with RNase Sa (grey). Numbering according to YoeB and orientation of the proteins as in Figure 33. The vital residues Arg59, Arg65, Glu46, His83 and Tyr84 are close, sometimes verlapping. The positions of Glu46 and His83 are shifted in YoeB relative to RNase Sa but the distance between their Cα atoms is nearly identical. His83 is located at the flexible C-terminus.

As noted in Table 9, Arg62 in *Mj*RelE and Arg65 in *Ph*RelE are at the same location and essential for the activity of these toxins. In YoeB, Glu46 and Arg65 are required for toxicity (Kamada and Hanaoka, 2005). Glu46 is near the position where Asp43 is located in *Mj*RelE, and Arg65 is placed exactly as the essential Arg62 in the structures of *Mj*RelE and *Ph*RelE (Figure 33), suggesting that residues Asp43/Arg62 in *Mj*RelE and Asp44/Arg65 in *Ph*RelE could be the most probable catalytic residues for RNA cleavage. YoeB, a purine-specific ribonuclease (Kamada and Hanaoka, 2005), has a "classical" catalytical site as shown in the comparison with RNase Sa (Figure 34). YoeB residues Arg59 and Arg65 overlap with the arginines from RNase Sa while His83 and Glu46 of YoeB are at similar distances compared to the positions of the equivalent His and Glu in RNase Sa. The distances

between these residues are 10.4 Å in YoeB and 9.9 Å (measured from Cγ to Cγ) in RNase Sa. His83 and Tyr84 of YoeB are placed at the C-terminus of the polypeptide chain, and their positions vary in the antitoxin bound and free forms (Kamada and Hanaoka, 2005), indicating that their distance to Glu46 most probably changes upon RNA binding (Kamada and Hanaoka, 2005).

In ColicinE5 and RegB the active site is located at different places of the toxins β-sheet and surrounding helices (Figure 35). Both of these proteins have endoribonuclease activity as YoeB does. Therefore the position of the catalytic diad is not restricted to one place on the surface of the cleft but can be placed at another part of the β-sheet, depending on the specificity of the endoribonuclease. The substrates for ColicinE5 are tRNAs coding for tyrosine, histidine, glutamic and aspartic acid, and are preferentially cleaved at anticodons while RegB cleaves in the Shine-Dalgarno sequence GGAG. Both enzymes have, besides the preferential target sequence, the requirement for a particular secondary structure of tRNA in the vicinity of the cleavage site (Lebars *et al.*, 2001; Lin *et al.*, 2005), which has not yet been established for YoeB or RelE.

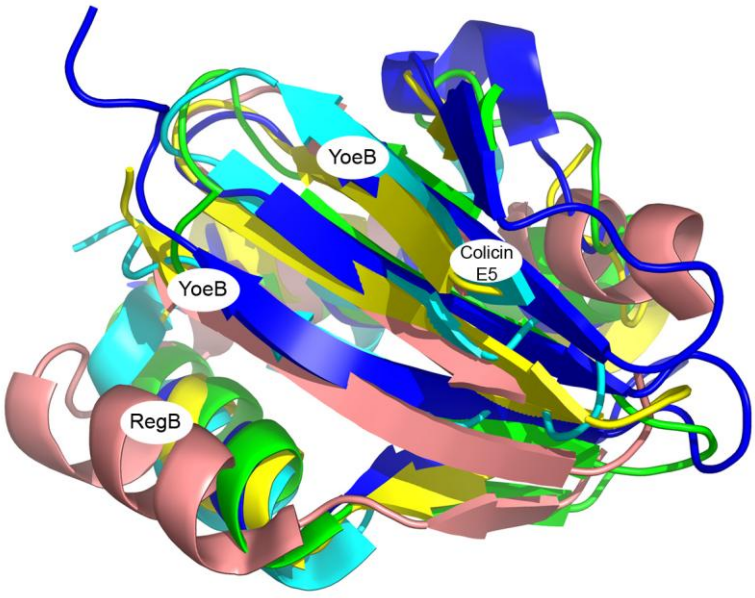

Figure 35 Superimposition of the central β-sheet from *M. jannaschii Mj*RelE (yellow), *P. horikoshii Ph*RelE (blue), YoeB (green), ColicinE5 (cyan) and RegB (pink). The names of the corresponding proteins are placed where the presumed active centre of the named protein is located. In RegB and ColicinE5 the general acid and base amino acids are placed on the neighbouring residues in the sequence while in YoeB they are placed far apart in sequence, the general acid on the C-terminus and the general base on strand β5.

4 Summary and Outlook

In the last decades a number of plasmid stabilization systems were discovered. Although difficult to study because they exhibit high toxicity to host cells, their function became clear after the toxin – antitoxin interplay was elucidated. They play an important part in stable maintenance and inheritance of plasmids and, as their function was now clear, only new members of similar traits were expected to be discovered. As the number of sequenced genomes grew, the number of putative chromosomally encoded TA systems rouse quickly. These new members of the TA superfamilies shared the common features of the already known TA systems. They were comprised of a toxin that is counteracted by tight binding of an upstream encoded antitoxin. These two proteins form a stable non-toxic toxin – antitoxin complex but the antitoxin is susceptible to protease degradation. If the antitoxin pool is not replenished the toxin is free to act upon its target, poisoning the cell and leading to cell stasis or even cell death. The delay in antitoxin production may be caused due to plasmid loss, in cases when the TA system is plasmid borne or inhibition of protein synthesis in cases of chromosomally encoded TA systems. The discovery of these chromosomally encoded TA systems was surprising as it contradicted the idea of the TA systems being the failsafe of plasmid inheritance. What was their purpose if they are not coded on the plasmid but on the chromosome? Later theories suggested that these systems are activated during cell starvation in an act of "Bacterial apoptosis". The current opinion is that these systems represent an advantage to the cell harbouring the system as they function like an emergency brake that holds the cell metabolism and redirects it to a new pathway. The advantage lies in a swift transfer to minimal protein production targeted for cell survival in stress conditions.

The known TA systems share little sequence identity, even within a toxin superfamily. Although bioinformatic methods have advanced and new members of the TA system toxins are identified only on their primary sequence, the real connections between the proteins can be realized only through the determinations of their three-dimensional structures. This was also the case with the here studied TA system from *Methanococcus jannaschii*.

The research on these TA systems is giving good insights in their function and predictions on further findings can be made.

There are new TA system reported but till date they have all fallen in one of the above mentioned four (or possibly five) TA superfamilies. We shell await for the new structure of the Phd/Doc TA system in order to judge the structural relations of this

superfamily to the TA superfamilies that already have their structural representatives. For now it seems that every toxin superfamily evolved on its own to form a TA system. The antitoxins followed the development but their relation with the toxins is subjective to changes. It is noticeable how the YefM/YoeB TA system shares the toxin with the RelE superfamily and the antitoxin with the Doc toxin superfamily. During development it was possible that switching of partners did occur, as has been suggested above.

Comparing the endoribonucleases and the toxins some conclusions can be drawn. The cleft shape, the arrangement of secondary structure elements and the positioning of positively charged residues are the essential ingredients for RNA binding. This common fold between the TA systems and other endoribonucleases is associated with the substrate (RNA) and therefore wide spread. The positioning of the residues responsible for RNA cleavage differs regardless of the common fold.

The structure determination of the *M. jannaschii Mj*RelBE protein complex was successful and there were more surprises waiting after the structure was refined. This structure represents the bridge between the RelE superfamily members *Ph*RelE and YoeB because it possesses the characteristics of both (see section 3.4.3). It did raise the question again whether the RelE toxin really needs another component, like the ribosome, in order to be active. Unfortunately the toxicity of the *Mj*RelE made it harder to work with the protein and only the inactive Arg62Ser variant was available for studies.

The structure determination of the *Ec*RelBE protein complex was unfortunately not possible as the major bottleneck in protein crystallography, the crystallization of proteins itself, has prevented this aim to be achieved. Despite all efforts no crystals of the *Ec*RelBE were obtained. As this *Ec*RelE protein was expressed in native form (which would mean it was less toxic than the *Mj*RelE, as discussed in section 3.4.3), some biochemical studies were carried out with it only to confirm previously published results by the Gerdes group (Christensen and Gerdes, 2003).It can be expected that the *Ec*RelE structure will be very near to the already determined structures of *Ph*RelE and *Mj*RelE, or better said it will be in the middle of these two structures as this can be predicted when looking at the primary sequence comparison (Figure 32).

TA systems have been studied for a number of years now and some applications have been proposed or already applied. It is left to further studies to develop an application many have waited for, the development of new antimicrobial drug based on the results of the research on TA systems.

5 Zusammenfassung

Die Toxin - Antitoxin (TA) Systeme werden von Chromosomen oder Plasmiden verschiedener Bakterienstämme kodiert. Die Rolle der Toxin - Antitoxin Systeme in Apoptose- ähnlichen Prozessen bei Bakterien, sowie der Zusammenhang mit der Stressantwort auf Nährstoffmangel, ist Gegenstand aktueller Forschung. So wurden während Untersuchungen an Toxin - Antitoxin Systemen in Bakterien auch die Proteine RelE und RelB entdeckt. Das *relB* Gen zog zuerst die Aufmerksamkeit auf sich, da Zellen, deren *relB* Gen im *E. coli* Chromosom Fehler aufwies, ein eigentümliches Verhalten nach Nährstoff-Entzug aufwiesen. Diese Zellen verlangsamten nach einem Aminosäurenentzug ihre Proteinbiosynthese, wie es für *Wild Typ* Zellen zu erwarten wäre, setzten aber nach etwa 10 Minuten die Biosynthese fort. Der Proteinkomplex, der von den aufeinander folgenden Genen *relB* und *relE* kodiert wird, wurde charakterisiert und gefunden, dass das Protein RelE als Toxin und RelB als sein spezifisches Antitoxin wirken.

Das RelBE System ähnelt den schon bekannten Toxin-Antitoxin Systemen (TA Systeme). Sie alle haben ähnliche Charakteristiken: a) sie bestehen aus zwei aufeinander folgenden Genen, b) das Gen-Produkt des einen ist ein langlebiges toxisches Protein und das des anderen ein kurzlebiges (Proteolyse- empfindliches) antitoxisches Protein, c) das Gen des Antitoxins liegt vor dem Gen des Toxins, d) beide Proteine werden gleichzeitig exprimiert (produziert), e) das Antitoxin wird im Überschuss zum Toxin produziert, f) das Antitoxin wird von Proteasen verdaut, h) das Operon wird entweder vom Antitoxin allein oder vom Antitoxin – Toxin Komplex autoreguliert.

In dieser Arbeit wurde die Struktur des *M. jannaschii Mj*RelBE Protein Komplexes ermittelt. Diese wies zu einem Zusammenhang mit den Endoribonucleasen, da sie eine sehr eng verwandte Faltung haben. Aus der Struktur konnte auch hergeleitet werden, weshalb die spontan entstandene Mutante *Mj*RelE(R62S) nicht toxisch für die *E. coli* Zellen ist. Diese Aminosäure nimmt eine Schlüsselrolle bei der Spaltung von mRNA an und hat die gleiche Rolle in den strukturverwandten Endoribonucleasen YoeB (gehört zur RelE toxin superfamilie) und RNase Sa.

6 Bibliography

Abrahams, J.P. and Leslie, A.G. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D Biol Crystallogr*, **52**, 30-42.

Alberts, B., Bray, D, Lewis, J, Raff, M, Roberts, K and Watson, J.D. (1994). *Molecular biology of the cell*. Garland Science, New York.

Alonso, J.C., Balsa, D, Cherny, I, Christensen, S.K, Francuski, D, Gazit, E, Gerdes, K, Hitchin, E.M, Teresa, M.M, Nieto, C, Overweg, K, Pellicer, T, Saenger, W, Welfle, H, Welfle, K and Wells, J. (2007). Bacterial Toxin-Antitoxin Systems as Targets for the Development of Novel Antibiotics. In R.A. Bonomo and M.E. Tolmasky (ed.), *Enzyme-Mediated Resistance to Antibiotics: Mechanism, Dissemination, and Prospects for inhibition*. ASM Press, Washington, D.C, pp. 313-329.

Altschul, S.F., Madden, T.L, Schäffer, A.A, Zhang, J, Zhang, Z, Miller, W and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, **25**, 3389-3402.

Anantharaman, V. and Aravind, L. (2003). New connections in the prokaryotic toxin-antitoxin network: relationship with the eukaryotic nonsense-mediated RNA decay system. *Genome Biol*, **4**, R81.

Andrade, M.A., Chacón, P, Merelo, J.J and Morán, F. (1993). Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng*, **6**, 383-390.

Bahassi, E.M., Salmon, M.A, Van Melderen, L, Bernard, P and Couturier, M. (1995). F plasmid CcdB killer protein: ccdB gene mutants coding for non-cytotoxic proteins which retain their regulatory functions. *Mol Microbiol*, **15**, 1031-1037.

Bennett-Lovsey, R.M., Herbert, A.D, Sternberg, M.J.E and Kelley, L.A. (2008). Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins*, **70**, 611-625.

Bhardwaj, A.. (2005).Biophysical and Structural Characterization of Site Specific β Recombinase and RelBE Toxin –Antitoxin Systems, *PhD thesis*, FU-Berlin, Berlin.

Bond, C.S.. (2003). TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics*, **19**, 311-312.

Bradford, M.M.. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, **72**, 248-254.

Brantl, S.. (2002). Antisense RNAs in plasmids: control of replication and maintenance. *Plasmid*, **48**, 165-173.

Bravo, A., de Torrontegui, G and Díaz, R. (1987). Identification of components of a new stability system of plasmid R1, ParD, that is close to the origin of replication of this plasmid. *Mol Gen Genet*, **210**, 101-110.

Brünger, A.T., Adams, P.D, Clore, G.M, DeLano, W.L, Gros, P, Grosse-Kunstleve, R.W, Jiang, J.S, Kuszewski, J, Nilges, M, Pannu, N.S, Read, R.J, Rice, L.M, Simonson, T and Warren, G.L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*, **54**, 905-921.

Bult, C.J., White, O, Olsen, G.J, Zhou, L, Fleischmann, R.D, Sutton, G.G, Blake, J.A, FitzGerald, L.M, Clayton, R.A, Gocayne, J.D, Kerlavage, A.R, Dougherty, B.A, Tomb, J.F, Adams, M.D, Reich, C.I, Overbeek, R, Kirkness, E.F, Weinstock, K.G, Merrick, J.M, Glodek, A, Scott, J.L, Geoghagen, N.S and Venter, J.C. (1996). Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii. *Science*, **273**, 1058-1073.

Buts, L., De Jonge, N, Loris, R, Wyns, L and Dao-Thi, M. (2005). Crystallization of the C-terminal domain of the addiction antidote CcdA in complex with its toxin CcdB. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, **61**, 949-952.

Cegłowski, P., Boitsov, A, Chai, S and Alonso, J.C. (1993). Analysis of the stabilization system of pSM19035-derived plasmid pBT233 in Bacillus subtilis. *Gene*, **136**, 1-12.

Christensen, S.K. and Gerdes, K. (2003). RelE toxins from bacteria and Archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. *Mol Microbiol*, **48**, 1389-1400.

Christensen, S.K., Mikkelsen, M, Pedersen, K and Gerdes, K. (2001). RelE, a global inhibitor of translation, is activated during nutritional stress. *Proc Natl Acad Sci U S A*, **98**, 14328-14333.

Christner, B.C., Morris, C.E, Foreman, C.M, Cai, R and Sands, D.C. (2008). Ubiquity of biological ice nucleators in snowfall. *Science*, **319**, 1214.

Collaborative Computational Project, N.4.. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallographica Section D*, **50**, 760-763.

Emsley, P. and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-2132.

Engelberg-Kulka, H. and Glaser, G. (1999). Addiction modules and programmed cell death and antideath in bacterial cultures. *Annu Rev Microbiol*, **53**, 43-70.

Fire, A., Xu, S, Montgomery, M.K, Kostas, S.A, Driver, S.E and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, **391**, 806-811.

Fitzgerald, D.. (1996). Why toxins!. *Semin Cancer Biol*, **7**, 87-95.

Garcia-Pino, A., Dao-Thi, M, Gazit, E, Magnuson, R.D, Wyns, L and Loris, R. (2008). Crystallization of Doc and the Phd-Doc toxin-antitoxin complex. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, **64**, 1034-1038.

Gazit, E. and Sauer, R.T. (1999). The Doc toxin and Phd antidote proteins of the bacteriophage P1 plasmid addiction system form a heterotrimeric complex. *J Biol Chem*, **274**, 16813-16818.

Gerdes, K. and Wagner, E.G.H. (2007). RNA antitoxins. *Curr Opin Microbiol*, **10**, 117-124.

Gerdes, K., Bech, F.W, Jørgensen, S.T, Løbner-Olesen, A, Rasmussen, P.B, Atlung, T, Boe, L, Karlstrom, O, Molin, S and von Meyenburg, K. (1986). Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the E. coli relB operon. *EMBO J*, **5**, 2023-2029.

Gerdes, K., Christensen, S.K and Løbner-Olesen, A. (2005). Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol*, **3**, 371-382.

Gerdes, K., Gultyaev, A.P, Franch, T, Pedersen, K and Mikkelsen, N.D. (1997). Antisense RNA-regulated programmed cell death.. *Annual Review of Genetics*, **31**, 1-31.

Gerdes, K., Rasmussen, P.B and Molin, S. (1986). Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proc Natl Acad Sci U S A*, **83**, 3116-3120.

Gotfredsen, M. and Gerdes, K. (1998). The Escherichia coli relBE genes belong to a new toxin-antitoxin gene family. *Mol Microbiol*, **29**, 1065-1076.

Gouet, P., Courcelle, E, Stuart, D.I and Métoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305-308.

Graille, M., Mora, L, Buckingham, R.H, van Tilbeurgh, H and de Zamaroczy, M. (2004). Structural inhibition of the colicin D tRNase by the tRNA-mimicking immunity protein. *EMBO J*, **23**, 1474-1482.

Hargreaves, D., Santos-Sierra, S, Giraldo, R, Sabariegos-Jareño, R, de la Cueva-Méndez, G, Boelens, R, Díaz-Orejas, R and Rafferty, J.B. (2002). Structural and functional analysis of the kid toxin protein from E. coli plasmid R1. *Structure*, **10**, 1425-1433.

Hartmann, R.K., Heinrich, J, Schlegl, J and Schuster, H. (1995). Precursor of C4 antisense RNA of bacteriophages P1 and P7 is a substrate for RNase P of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 5822-5826.

Hayes, F.. (2003). Toxins-Antitoxins: Plasmid Maintenance, Programmed CellDeath, and Cell Cycle Arrest.. *Science*, **301**, 1496 - 1499.

Hazan, R., Sat, B, Reches, M and Engelberg-Kulka, H. (2001). Postsegregational killing mediated by the P1 phage "addiction module" phd-doc requires the Escherichia coli programmed cell death system mazEF. *J Bacteriol*, **183**, 2046-2050.

Hooft, R.W., Vriend, G, Sander, C and Abola, E.E. (1996). Errors in protein structures. *Nature*, **381**, 272.

Kabsch, W.. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, **32**, 922-923.

Kabsch, W.. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography*, **26**, 795-800.

Kabsch, W. and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577-2637.

Kamada, K. and Hanaoka, F. (2005). Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Mol Cell*, **19**, 497-509.

Kamada, K., Hanaoka, F and Burley, S.K. (2003). Crystal structure of the MazE/MazF complex: molecular bases of antidote-toxin recognition. *Mol Cell*, **11**, 875-884.

Kamphuis, M.B., Bonvin, A.M.J.J, Monti, M.C, Lemonnier, M, Muñoz-Gómez, A, van den Heuvel, R.H.H, Díaz-Orejas, R and Boelens, R. (2006). Model for RNA binding and the catalytic site of the RNase Kid of the bacterial parD toxin-antitoxin system. *J Mol Biol*, **357**, 115-126.

Kamphuis, M.B., Monti, M.C, van den Heuvel, R.H.H, López-Villarejo, J, Díaz-Orejas, R and Boelens, R. (2007). Structure and function of bacterial kid-kis and related toxin-antitoxin systems. *Protein Pept Lett*, **14**, 113-124.

Kedzierska, B., Lian, L and Hayes, F. (2007). Toxin-antitoxin regulation: bimodal interaction of YefM-YoeB with paired DNA palindromes exerts transcriptional autorepression. *Nucleic Acids Res*, **35**, 325-339.

Kelley, L.A., MacCallum, R.M and Sternberg, M.J. (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J Mol Biol*, **299**, 499-520.

Kibbe, W.A.. (2007). OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res*, **35**, W43-6.

Kissinger, C.R., Gehlhaar, D.K and Fogel, D.B. (1999). Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr D Biol Crystallogr*, **55**, 484-491.

Krissinel, E. and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, **372**, 774-797.

Laemmli, U.K.. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

Lebars, I., Hu, R.M, Lallemand, J.Y, Uzan, M and Bontems, F. (2001). Role of the substrate conformation and of the S1 protein in the cleavage efficiency of the T4 endoribonuclease RegB. *J Biol Chem*, **276**, 13264-13272.

Leslie, A.G.W.. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, **26**, .

Lewis, K.. (2000). Programmed death in bacteria. *Microbiol Mol Biol Rev*, **64**, 503-514.

Lin, Y., Elias, Y and Huang, R.H. (2005). Structural and mutational studies of the catalytic domain of colicin E5: a tRNA-specific ribonuclease. *Biochemistry*, **44**, 10494-10500.

Lioy, V.S., Martín, M.T, Camacho, A.G, Lurz, R, Antelmann, H, Hecker, M, Hitchin, E, Ridge, Y, Wells, J.M and Alonso, J.C. (2006). pSM19035-encoded zeta toxin induces stasis followed by death in a subpopulation of cells. *Microbiology*, **152**, 2365-2379.

Loris, R., Dao-Thi, M.H, Bahassi, E.M, Van Melderen, L, Poortmans, F, Liddington, R, Couturier, M and Wyns, L. (1999). Crystal structure of CcdB, a topoisomerase poison from E. coli. *J Mol Biol*, **285**, 1667-1677.

Lu, G.. (2000). TOP: a new method for protein structure comparisons and similarity searches. *Journal of Applied Crystallography*, **33**, 176-183.

Madl, T., Van Melderen, L, Mine, N, Respondek, M, Oberer, M, Keller, W, Khatai, L and Zangger, K. (2006). Structural basis for nucleic acid and toxin recognition of the bacterial antitoxin CcdA. *J Mol Biol*, **364**, 170-185.

Masuda, Y., Miyakawa, K, Nishimura, Y and Ohtsubo, E. (1993). chpA and chpB, Escherichia coli chromosomal homologs of the pem locus responsible for stable maintenance of plasmid R100. *J Bacteriol*, **175**, 6850-6856.

McCoy, Airlie J,, Grosse-Kunstleve, Ralf W,, Adams, Paul D,, Winn, Martyn D,, Storoni, Laurent C, and Read, Randy J,. (2007). Phaser crystallographic software. , **40**, 658-674.

Meiler, J. and Baker, D. (2003). Coupled prediction of protein secondary and tertiary structure. *Proc Natl Acad Sci U S A*, **100**, 12105-12110.

Meinhart, A., Alonso, J.C, Sträter, N and Saenger, W. (2003). Crystal structure of the plasmid maintenance system epsilon/zeta: functional mechanism of toxin zeta and inactivation by epsilon 2 zeta 2 complex formation. *Proc Natl Acad Sci U S A*, **100**, 1661-1666.

Metzger, S., Dror, I.B, Aizenman, E, Schreiber, G, Toone, M, Friesen, J.D, Cashel, M and Glaser, G. (1988). The nucleotide sequence and characterization of the relA gene of Escherichia coli. *J Biol Chem*, **263**, 15699-15704.

Moretti, S., Armougom, F, Wallace, I.M, Higgins, D.G, Jongeneel, C.V and Notredame, C. (2007). The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. *Nucleic Acids Res*, **35**, W645-8.

Murshudov, G.N., Vagin, A.A and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, **53**, 240-255.

Navaza, J,. (1994). AMoRe: an automated package for molecular replacement. , **50**, 157-163.

Nordström, K. and Austin, S.J. (1989). Mechanisms that contribute to the stable segregation of plasmids. *Annu Rev Genet*, **23**, 37-69.

Ogura, T. and Hiraga, S. (1983). Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proc Natl Acad Sci U S A*, **80**, 4784-4788.

Ogura, T. and Hiraga, S. (1983). Partition mechanism of F plasmid: two plasmid gene-encoded products and a cis-acting region are involved in partition. *Cell*, **32**, 351-360.

Otwinowski, Z. and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology*, **276**, 307-326.

Pandey, D.P. and Gerdes, K. (2005). Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res*, **33**, 966-976.

Pedersen, K. and Gerdes, K. (1999). Multiple hok genes on the chromosome of Escherichia coli. *Molecular Microbiology*, **32**, 1090-1102.

Pedersen, K., Zavialov, A.V, Pavlov, M.Y, Elf, J, Gerdes, K and Ehrenberg, M. (2003). The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell*, **112**, 131-140.

Potter, H.. (1988). Electroporation in biology: methods, applications, and instrumentation. *Anal Biochem*, **174**, 361-373.

Rice, K.C. and Bayles, K.W. (2003). Death's toolbox: examining the molecular components of bacterial programmed cell death. *Mol Microbiol*, **50**, 729-738.

Sambrook, J., Fritsch, E and Maniatis, T. (1989). *Molecular Cloning - A Laboratory Manual*. Cold Spring Harbor Press, Cold Spring Harbor, New York.

Saïda, F. and Odaert, B. (2007). RNA recognition and cleavage by sequence-specific endoribonucleases. *Protein Pept Lett*, **14**, 103-111.

Schmidt, O., Schuenemann, V.J, Hand, N.J, Silhavy, T.J, Martin, J, Lupas, A.N and Djuranovic, S. (2007). prlF and yhaV encode a new toxin-antitoxin system in Escherichia coli. *J Mol Biol*, **372**, 894-905.

Sevin, E.W. and Barloy-Hubler, F. (2007). RASTA-Bacteria: a web-based tool for identifying toxin-antitoxin loci in prokaryotes. *Genome Biol*, **8**, R155.

Shyu, A., Wilkinson, M.F and van Hoof, A. (2008). Messenger RNA regulation: to translate or to degrade. *EMBO J*, **27**, 471-481.

Sleytr, U.B., Egelseer, E.M, Ilk, N, Pum, D and Schuster, B. (2007). S-Layers as a basic building block in a molecular construction kit. *FEBS J*, **274**, 323-334.

Takagi, H., Kakuta, Y, Okada, T, Yao, M, Tanaka, I and Kimura, M. (2005). Crystal structure of archaeal toxin-antitoxin RelE-RelB complex with implications for toxin activity and antitoxin effects. *Nat Struct Mol Biol*, **12**, 327-331.

Terwilliger, T.C.. (2003). Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr D Biol Crystallogr*, **59**, 38-44.

Tian, Q.B., Ohnishi, M, Tabuchi, A and Terawaki, Y. (1996). A new plasmid-encoded proteic killer gene system: cloning, sequencing, and analyzing hig locus of plasmid Rts1. *Biochem Biophys Res Commun*, **220**, 280-284.

Tomizawa, J., Itoh, T, Selzer, G and Som, T. (1981). Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proc Natl Acad Sci U S A*, **78**, 1421-1425.

Tony J Greenfield, Keith E Weaver,. (2000). Antisense RNA regulation of the pAD1 par post-segregational killing system requires interaction at the 5' and 3' ends of the RNAs. *Molecular Microbiology*, **37**, 661-670.

Tsuchimoto, S., Ohtsubo, H and Ohtsubo, E. (1988). Two genes, pemK and pemI, responsible for stable maintenance of resistance plasmid R100. *J Bacteriol*, **170**, 1461-1466.

Uzan, M., Favre, R and Brody, E. (1988). A nuclease that cuts specifically in the ribosome binding site of some T4 mRNAs. *Proc Natl Acad Sci U S A*, **85**, 8895-8899.

Vonrhein, C., Blanc, E, Roversi, P and Bricogne, G. (2007). Automated structure solution with autoSHARP. *Methods Mol Biol*, **364**, 215-230.

Wagner, E.G.H. and Simons, R.W. (1994). Antisense RNA control in bacteria, phages, and plasmids. *Annual Review of Microbiology*, **48**, 713.

Weaver, K.E.. (2007). Emerging plasmid-encoded antisense RNA regulated systems. *Curr Opin Microbiol*, **10**, 110-116.

Wilson, D.N. and Nierhaus, K.H. (2005). RelBE or not to be. *Nat Struct Mol Biol*, **12**, 282-284.

Winn, M.D., Isupov, M.N and Murshudov, G.N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr*, **57**, 122-133.

Yamamoto, T.M., Gerdes, K and Tunnacliffe, A. (2002). Bacterial toxin RelE induces apoptosis in human cells. *FEBS Lett*, **519**, 191-194.

de la Cueva-Méndez, G., Mills, A.D, Clay-Farrace, L, Díaz-Orejas, R and Laskey, R.A. (2003). Regulatable killing of eukaryotic cells by the prokaryotic proteins Kid and Kis. *EMBO J*, **22**, 246-251.

7 Appendices

7.1 Tricine Gels

The gels consist of three layers of different acrylamide concentrations (Table 10).
First the gel buffer is prepared as following:

Gel buffer: 0.3% SDS; 3 M Tris pH 8.45
(36.3 g Tris; 0.3 g SDS; titrate to pH 8.45 with HCl and fill up to 100 ml)

Table 10 The compositions of tricine gel layers

		4% Stack gel	10% Sepparation gel	16% Sepparation gel
Acrylamid	(ml)	1.66	1.66	8.25
Gel buffer	(ml)	3.13	1.66	5.00
Glycerin	(ml)	-	-	1.66
ddH2O	(ml)	7.7	1.68	0.09
Total volume	(ml)	12.5	5	15

For the 10x15 cm glass plates the following volumes are prepared and poured in the order given.

5ml 16% Sepparation gel
+5 µl Temed
+20 µl APS

2ml 10% Sepparation gel
+5 µl Temed
+10 µl APS

2ml 4% Stack gel
+5 µl Temed
+20 µl APS

The gel is run with 30 V till the samples enter the first separation layer and then at 150 V for 120 minutes. The buffers for the cathode and anode are given below:

Cathode buffer: 0.1% SDS; 0.1 M Tricin; 0.1 M Tris pH 8.25
(3 g Tris; 4.5 g Tricin; 0.25 g SDS; titrate to pH 8.25 with HCl and fill up to 250 ml)

Anode buffer: 0.2 M Tris pH 8.9
(24.2 g Tris; titrate to pH 8.9 with HCl and fill up to 1000 ml)

7.2 PDB codes of relevant proteins mentioned in the dissertation

Table 11 PDB entry codes of the for this work relevant proteins.

Protein(s)	**PDB entry code**
*Mj*RelBE	3bpq
*Ph*RelBE	1wmi
YoeB	2a6s
RNase Sa	1rgg
ColicinE5	2a8k
RegB	2hx6
Ribonuclease T1	2rnt
Colicin D	1v74
YefM/YoeB	2a6q

7.3 Crystal Properties and Data Collection for the trimethylleadacetate ($C_5H_{11}O_2Pb$) socked crystals

The trimethylleadacetate ($C_5H_{11}O_2Pb$) socked crystals diffracted up to 2.6 Å but data were cut off at 2.8 Å due to poor data quality in the higher resolution shells. The data collection statistics are given in comparison to the native dataset in the following table (Table 12). The cell constant difference between this derivative and native datasets as well the high mosaicity, are the reason why this derivative was not used in the phase determination.

Table 12 Data collection statistics for native and derivative crystals.

Data collection

	Native	Trimethylleadacetate
Space group	$P2_1$	$P2_1$
Cell parameters		
a, b, c (Å);	52.79 57.99 58.75	51.84 58.32 60.01
α, β, γ (°)	90.00 92.30 90.00	90.00 91.41 90.00
Wavelength (Å)	0.9490	0.9501
Resolution limits (Å)	50.0-2.13 (2.13-2.18)	50.0-2.80 (2.80-2.88)
Total reflections	232747	344455
Unique reflections	20062	20727
R sym[a]	0.050 (0.42)	0.050 (0.44)
I / σ(I)	27.03 (3.13)	11.90 (4.52)
Completeness (%)	99.6 (98.9)	97.2 (93.1)
Redundancy	4.2 (4.1)	6.9 (6.7)
Mosaicity (°)	0.39	1.38

a. *R*sym = Σhkl Σj|Ij-<*I*>|/ΣhklΣjIj, where <*I*> is the mean intensity of reflection *hkl*.

7.4 DNA and Amino Acid Sequences of the Studied Proteins

*Mj*RelB

1 atg agg ctc aaa aag aga ttt aaa aaa ttt ttc atc agc aga aaa 45
1 Met Arg Leu Lys Lys Arg Phe Lys Lys Phe Phe Ile Ser Arg Lys 15

46 gaa tat gaa aag att gag gaa att tta gat att ggc ttg gct aaa 90
16 Glu Tyr Glu Lys Ile Glu Glu Ile Leu Asp Ile Gly Leu Ala Lys 30

91 gct atg gag gaa aca aaa gat gat gaa tta ttg act tat gat gaa 135
31 Ala Met Glu Glu Thr Lys Asp Asp Glu Leu Leu Thr Tyr Asp Glu 45

136 ata aag gaa tta ttg gga gat 156
46 Ile Lys Glu Leu Leu Gly Asp 52

*Mj*RelE – Uniprot accession number: Q58503

1 atg aaa gtg tta ttt gct aaa aca ttt gtt aag gat tta aag cat 45
1 Met Lys Val Leu Phe Ala Lys Thr Phe Val Lys Asp Leu Lys His 15

46 gtt cca ggg cat ata aga aaa aga ata aag cta ata att gaa gaa 90
16 Val Pro Gly His Ile Arg Lys Arg Ile Lys Leu Ile Ile Glu Glu 30

91 tgt caa aat tct aac tca tta aat gat tta aag tta gat att aag 135
31 Cys Gln Asn Ser Asn Ser Leu Asn Asp Leu Lys Leu Asp Ile Lys 45

136 aaa ata aag ggc tat cac aat tat tat agg att aga gta gga aat 180
46 Lys Ile Lys Gly Tyr His Asn Tyr Tyr Arg Ile Arg Val Gly Asn 60

181 tat aga ata ggt att gag gtt aat gga gat acg att att ttt aga 225
61 Tyr Arg Ile Gly Ile Glu Val Asn Gly Asp Thr Ile Ile Phe Arg 75

226 aga gta ttg cat aga aaa agc ata tat gat tat ttc cca 264
76 Arg Val Leu His Arg Lys Ser Ile Tyr Asp Tyr Phe Pro 88

*Ec*RelB - Uniprot accession number: P0C079

1 ATG GGT AGC ATT AAC CTG CGT ATT GAC GAT GAA CTT AAA GCG CGT 45
1 Met Gly Ser Ile Asn Leu Arg Ile Asp Asp Glu Leu Lys Ala Arg 15

46 TCT TAC GCC GCG CTT GAA AAA ATG GGT GTA ACT CCT TCT GAA GCG 90
16 Ser Tyr Ala Ala Leu Glu Lys Met Gly Val Thr Pro Ser Glu Ala 30

91 CTT CGT CTC ATG CTC GAG TAT ATC GCT GAC AAT GAA CGC TTG CCG 135
31 Leu Arg Leu Met Leu Glu Tyr Ile Ala Asp Asn Glu Arg Leu Pro 45

136 TTC AAA CAG ACA CTC CTG AGT GAT GAA GAT GCT GAA CTT GTG GAG 180
46 Phe Lys Gln Thr Leu Leu Ser Asp Glu Asp Ala Glu Leu Val Glu 60

181 ATA GTG AAA GAA CGG CTT CGT AAT CCT AAG CCA GTA CGT GTG ACG 225
61 Ile Val Lys Glu Arg Leu Arg Asn Pro Lys Pro Val Arg Val Thr 75

226 CTG GAT GAA CTC 237
76 Leu Asp Glu Leu 79

*Ec*RelE – Uniprot accession number: P0C077

1 ATG GCG TAT TTT CTG GAT TTT GAC GAG CGG GCA CTA AAG GAA TGG 45
1 Met Ala Tyr Phe Leu Asp Phe Asp Glu Arg Ala Leu Lys Glu Trp 15

46 CGA AAG CTG GGC TCG ACG GTA CGT GAA CAG TTG AAA AAG AAG CTG 90
16 Arg Lys Leu Gly Ser Thr Val Arg Glu Gln Leu Lys Lys Lys Leu 30

91 GTT GAA GTA CTT GAG TCA CCC CGG ATT GAA GCA AAC AAG CTC CGT 135
31 Val Glu Val Leu Glu Ser Pro Arg Ile Glu Ala Asn Lys Leu Arg 45

136 GGT ATG CCT GAT TGT TAC AAG ATT AAG CTC CGG TCT TCA GGC TAT 180
46 Gly Met Pro Asp Cys Tyr Lys Ile Lys Leu Arg Ser Ser Gly Tyr 60

181 CGC CTT GTA TAC CAG GTT ATA GAC GAG AAA GTT GTC GTT TTC GTG 225
61 Arg Leu Val Tyr Gln Val Ile Asp Glu Lys Val Val Val Phe Val 75

226 ATT TCT GTT GGG AAA AGA GAA CGC TCG GAA GTA TAT AGC GAG GCG 270
76 Ile Ser Val Gly Lys Arg Glu Arg Ser Glu Val Tyr Ser Glu Ala 90

271 GTC AAA CGC ATT CTC 285
91 Val Lys Arg Ile Leu 95

7.5 Contact tables

7.5.1 Heterodimer I

$RelE_I$ (chain B) and $RelB_I$ (chain A)

Hydrogen bonds

	Residue1	Distance (Å)	Residue2	Salt bridge
1	B:ARG 57[NH2]	2.82	A:GLU 21[OE1]	+
2	B:ARG 57[NE]	3.07	A:GLU 21[OE1]	+
3	B:ARG 80[NH1]	3.14	A:GLU 21[OE2]	+
4	B:ARG 80[NH2]	3.36	A:GLU 21[OE2]	+
5	B:ARG 55[NH1]	3.86	A:ASP 25[OD1]	+
6	B:LYS 45[NZ]	3.55	A:ASP 25[OD1]	+
7	B:LYS 45[NZ]	2.57	A:ASP 25[OD2]	+
8	B:PRO 88[O]	2.70	A:LYS 30[NZ]	+
9	B:ARG 76[NH1]	2.85	A:ALA 31[O]	
10	B:ARG 75[NH1]	3.07	A:THR 35[O]	
11	B:ARG 75[NH1]	3.04	A:THR 35[OG1]	
12	B:LYS 7[N]	3.50	A:ASP 38[OD2]	
13	B:THR 8[N]	2.84	A:ASP 38[OD2]	
14	B:THR 8[OG1]	2.60	A:ASP 38[OD2]	
15	B:LYS 7[N]	2.88	A:GLU 39[O]	
16	B:PHE 5[N]	2.74	A:LEU 41[O]	
17	B:PHE 5[O]	2.67	A:LEU 41[N]	
18	B:GLU 29[OE2]	2.76	A:TYR 43[OH]	

7.5.2 Heterodimer II

$RelE_{II}$ (chain D) - $RelB_{II}$ (chain C)

Hydrogen bonds

	Residue1	Distance (Å)	Residue2	Salt bridge
1	D:ARG 57[NE]	3.54	C:GLU 21[OE1]	+
2	D:ARG 57[NH1]	2.66	C:GLU 21[OE1]	+
3	D:ARG 57[NH1]	2.97	C:GLU 21[OE2]	+
4	D:ARG 80[NH1]	3.11	C:GLU 21[OE2]	+
5	D:ARG 80[NH2]	3.81	C:GLU 21[OE2]	+
6	D:LYS 48[NZ]	3.72	C:GLU 22[OE2]	+
7	D:ARG 55[NE]	3.90	C:ASP 25[OD1]	+
8	D:LYS 45[NZ]	3.03	C:ASP 25[OD2]	+

9	D:ARG 76[NH2]	2.83	C:GLU 34[OE2]	+
10	D:ARG 76[NE]	3.14	C:ALA 31[O]	
11	D:TYR 50[OH]	3.71	C:GLU 33[OE2]	
12	D:ARG 76[NH2]	2.83	C:GLU 34[OE2]	
13	D:ARG 75[NH1]	2.88	C:THR 35[O]	
14	D:ARG 75[NH1]	3.25	C:THR 35[OG1]	
15	D:LYS 7[N]	3.61	C:ASP 38[OD1]	
16	D:THR 8[OG1]	2.68	C:ASP 38[OD1]	
17	D:THR 8[N]	3.15	C:ASP 38[OD1]	
18	D:LYS 7[N]	3.14	C:GLU 39[O]	
19	D:PHE 5[N]	2.75	C:LEU 41[O]	
20	D:TYR 53[OH]	3.01	C:LYS 36[NZ]	
21	D:PHE 5[O]	2.70	C:LEU 41[N]	
22	D:VAL 3[O]	3.85	C:TYR 43[N]	

7.5.3 Heterodimer to Heterodimer Interactions

$RelB_I$ - $RelB_{II}$ (chain A - chain C)

Hydrogen bonds

	Residue1	Distance (Å)	Residue2	Salt bridges
1	A:ILE 12[O]	2.79	C:PHE 10[N]	
2	A:PHE 10[O]	3.23	C:ILE 12[N]	
3	A:PHE 10[O]	3.71	C:SER 13[N]	
4	A:LYS 8[O]	3.77	C:SER 13[OG]	
5	A:ILE 12[N]	3.14	C:PHE 10[O]	
6	A:PHE 10[N]	2.65	C:SER 13[OG]	

$RelE_I$ - $RelB_{II}$ (chain B - chain C)

Hydrogen bonds

	Residue1	Distance (Å)	Residue2	Salt bridges
1	B:ARG 80[NE]	2.90	C:GLU 16[OE1]	+
2	B:ARG 80[NE]	3.13	C:GLU 16[OE2]	+
3	B:ARG 80[NH2]	2.94	C:GLU 16[OE1]	+
4	B:ASP 85[OD1]	3.58	C:LYS 19[NZ]	+
5	B:TYR 84[OH]	2.40	C:GLU 16[OE1]	
6	B:LYS 81[O]	2.80	C:LYS 19[NZ]	

7.5.4 Crystal contacts

Above each table the coordinates of the symmetry related molecule is given.
$RelB_I$ is chain A
$RelE_I$ is chain B
$RelB_{II}$ is chain C
$RelE_{II}$ is chain D

A: x, y, z B: -x,y-1/2,-z+1
Hydrogen bonds

	Residue1	Distance (Å)	Residue2	Salt bridges
1	A:ASP 37[OD1]	2.70	B:HIS 79[NE2]	+
2	A:ASP 37[OD1]	2.74	B:ASN 60[ND2]	

B: x, y, z D: x-1,y,z

	Residue1	Distance (Å)	Residue2	Salt bridges
1	B:ASN 38[ND2]	3.38	D:ASP 39[O]	
2	B:LYS 46[NZ]	3.16	D:SER 36[OG]	
3	B:TYR 54[OH]	2.58	D:ASP 39[OD2]	

A: x, y, z A: -x,y-1/2,-z+1

	Residue1	Distance (Å)	Residue2	Salt bridges
1	A:GLU 45[OE2]	3.49	A:LYS 9[NZ]	+
2	A:GLU 45[OE1]	3.34	A:LYS 9[NZ]	+
3	A:GLU 39[OE1]	2.72	A:LYS 9[NZ]	+
4	A:LEU 40[O]	2.92	A:LYS 9[NZ]	

C: x, y, z D: -x+1,y-1/2,-z

	Residue1	Distance (Å)	Residue2	Salt bridges
1	C:ASP 38[O]	3.38	D:LYS 22[NZ]	
2	C:LEU 40[O]	3.43	D:ARG 23[NH1]	
3	C:GLU 45[OE2]	2.84	D:ARG 23[NH1]	
4	C:GLU 45[OE1]	2.74	D:ARG 23[NH2]	

7.6 Purification of the *Mj*RelBE protein complex

The protocol for purification is given in the Materials and Methods section. The *Mj*RelBE protein complex elutes at high salt concentration from the heparin column (Figure 36a) in high purity. The gel filtration (Figure 36b) step is used only to "polish" the already "crystallisation" clean sample.

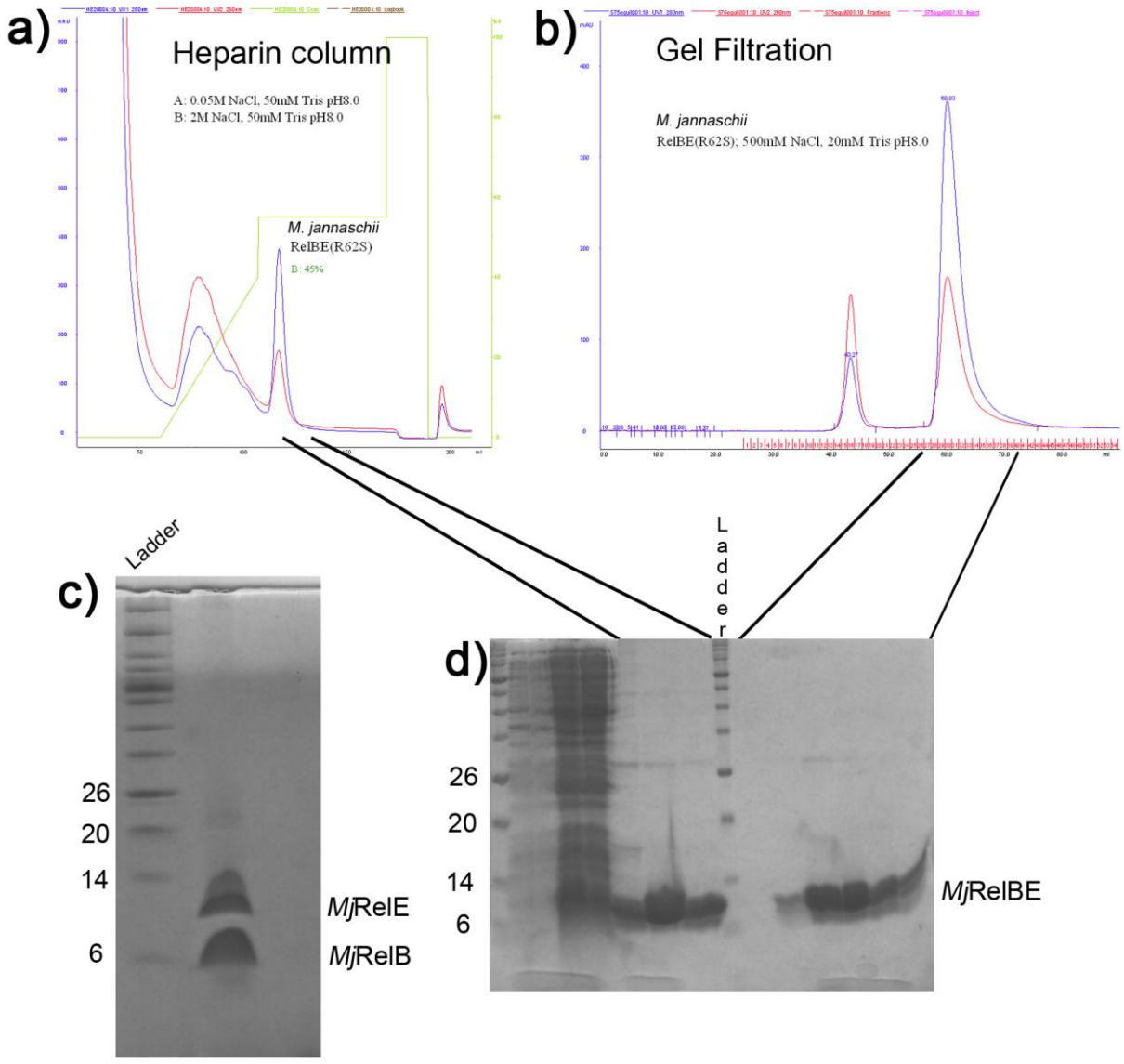

Figure 36 Purification of *Mj*RelE and *Mj*RelB. **a)** Chromatogram after the Heparin column. *Mj*RelBE elutes at almost 1 M NaCl (45% of buffer B). **b)** Gel filtration on a S75 column with a clean elution peak. **c)** Tricine gel of a purified sample of *Mj*RelBE. **d)** SDS-PAGE of the fractions from HE20 and S75 columns.

7.7 Purification of the *Ec*RelBE protein complex

The purification protocols are given in the Materials and Methods section. The *Ec*RelE and *Ec*RelB protein bands in the SDS-PAGE are easily distinguishable (Figure 37b), unlike the *M. jannaschii* proteins (Figure 6 in section 3.1.1.2). Purification of the non-tagged *Ec*RelBE was difficult as its binding to heparin is far from the specificity the *M. jannaschii* proteins exhibited. Alternative purification solutions were necessary but none of them led to more than 90% pure protein.

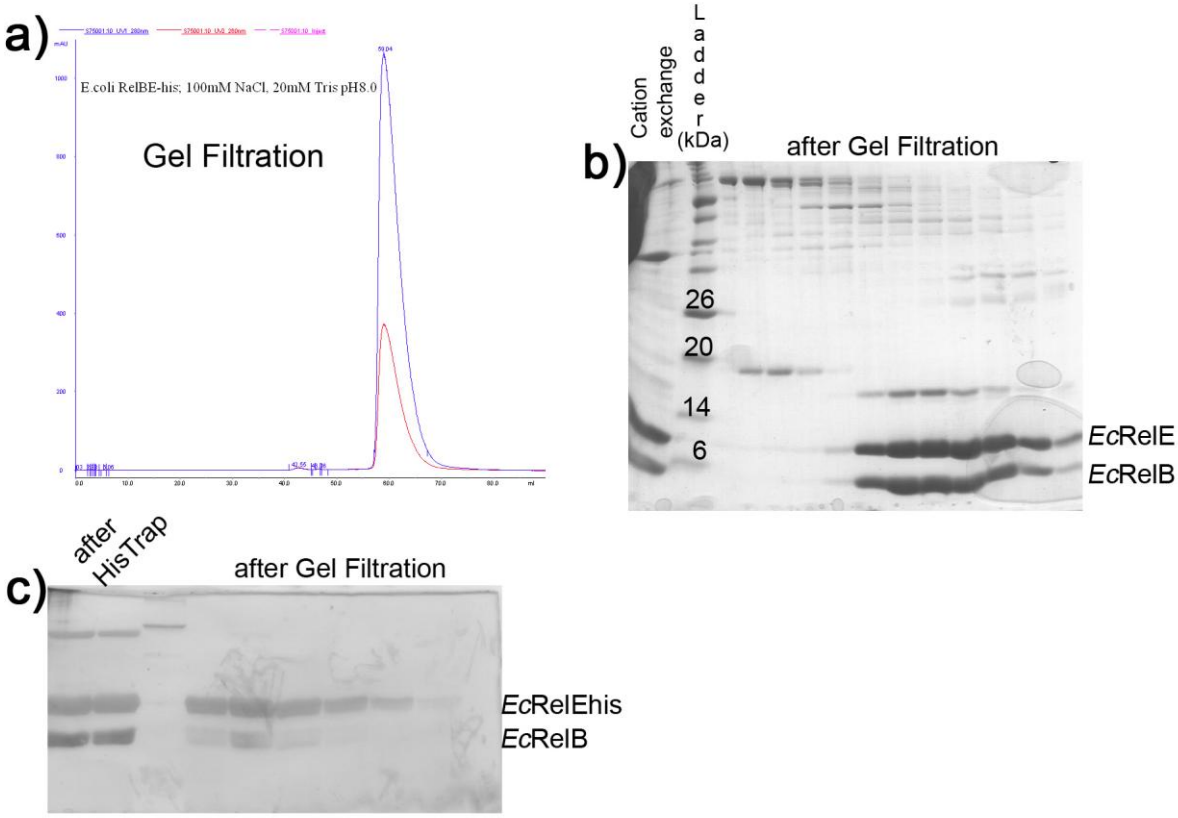

Figure 37 Purification of *Ec*RelBE. **a)** Gel filtration chromatogram of *Ec*RelBE-his. **b)** SDS-PAGE from the *Ec*RelBE purification. The first step is AS precipitation, followed by cation exchange on a HS column and finalized with gel filtration on a S75 size exclusion column. **c)** Silver stained SDS-PAGE from the purification of *Ec*RelBE-his. After affinity column (HisTrep) only a gel filtration step on a S75 size exclusion column is necessary to purify the protein complex.

7.8 Sequence alignments of the RelE superfamily members

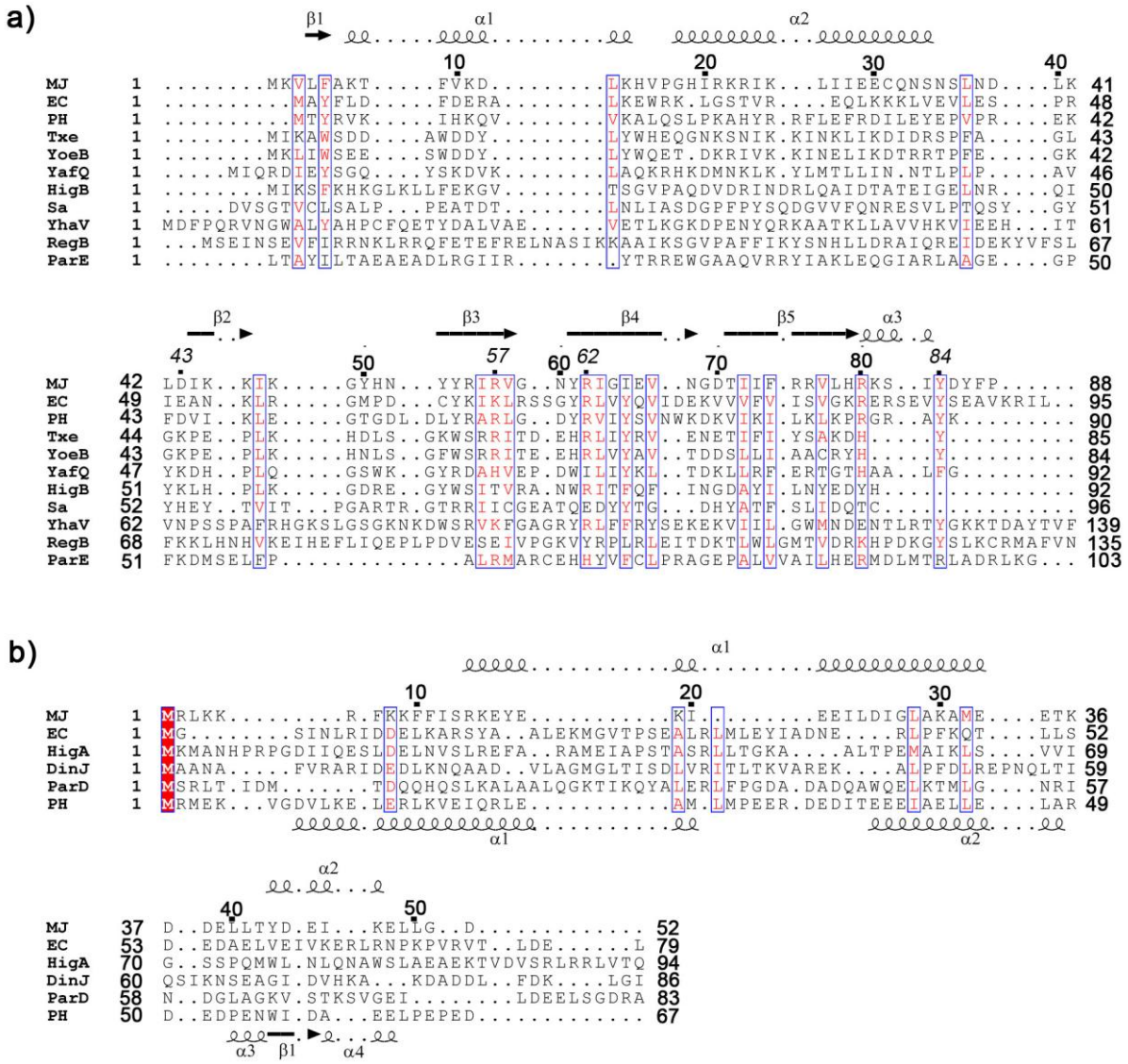

Figure 38 Alignment. **a)** Alignment of the RelE superfamily toxins. RelE from MJ *(M. jannaschii)*, EC (*E.coli*), PH (*P. horikoshii*), Txe (*E. faecium*), HigB (plasmid *Rts1*), ParE (plasmid *RK2*), RNase Sa (*S. aureofaciens*) and YoeB, YhaV, YafQ (all thre from *E. coli* chromosome). Conserved residues among RelE family members numbered in cursive font. Most noticeable are the conserved basic residues in the sequences. Secondary structure elements of *Mj*RelE are indicated above the sequences. **b)** Alignment of the antitoxins from RelE MJ *(M. jannaschii)*, EC (*E. coli*), HigA (plasmid *Rts1*), DinJ (*E. coli*), ParD (plasmid *RK2*) and PH (*P. horikoshii*) all with little sequence identity. Secondary structure elements of *Mj*RelB and *Ph*RelB are indicated above and below the sequences respectively.

8 Curriculum vitae

„Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten“

9 Publications

Alonso, J.C., Balsa, D, Cherny, I, Christensen, S.K, Francuski, D, Gazit, E, Gerdes, K, Hitchin, E.M, Teresa, M.M, Nieto, C, Overweg, K, Pellicer, T, Saenger, W, Welfle, H, Welfle, K and Wells, J. (2007). Bacterial Toxin-Antitoxin Systems as Targets for the Development of Novel Antibiotics. In R.A. Bonomo and M.E. Tolmasky (ed.), *Enzyme-Mediated Resistance to Antibiotics: Mechanism, Dissemination, and Prospects for inhibition.* ASM Press, Washington, D.C, pp. 313-329.

Francuski D. and Saenger W. Crystal Structure of the Antitoxin-Toxin Protein Complex RelB-RelE from *Methanococcus jannaschii*, *Manuscript in preparation*

Francuski D., Rossmann M. and Saenger W. Crytsal Structure of Lactaldehyde Dehydrogenase from *Esherichia coli* and Inhibition by Taurocholate, *Manuscript in preparation*