



HAL
open science

The crystal structure of a long form tRNase Z from *S. cerevisiae* and study of its interactome

Miao Ma

► **To cite this version:**

Miao Ma. The crystal structure of a long form tRNase Z from *S. cerevisiae* and study of its interactome. Structural Biology [q-bio.BM]. Université Paris-Saclay, 2016. English. NNT : 2016SACLS353 . tel-01940586

HAL Id: tel-01940586

<https://theses.hal.science/tel-01940586v1>

Submitted on 30 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2016SACLS353



THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
"FACULTÉ DE PHARMACIE DE CHÂTENAY-MALABRY"

ECOLE DOCTORALE N° 569
Innovation thérapeutique : du fondamental à l'appliqué (ITFA)

Biochimie et Biologie structurale

Par

Mlle Miao MA

The crystal structure of a long form tRNase Z from *S. cerevisiae*
and study of its interactome

Thèse présentée et soutenue à Orsay, le 24 Novembre 2016:

Composition du Jury :

M. Minard Philippe	Professeur Université Paris-Sud (I2BC)	Président
Mme. Mayer Claudine	Professeure Université Paris-Diderot (Institut Pasteur)	Rapporteuse
M. Giegé Philippe	Directeur de recherche (Institut de Biologie Moléculaire des Plantes, CNRS)	Rapporteur
M. Condon Ciarán	Directeur de recherche (Institut de Biologie Physico-Chimique, CNRS)	Examineur
Mme. Gallay Inès	Ingénieur de recherche (I2BC, CNRS)	Examinatrice
M. Van Tilbeurgh Herman	Professeur Université Paris-Sud (I2BC)	Directeur de thèse

Titre : La structure cristalline d'une forme longue tRNase Z de la levure et l'étude de son interactome

Mots clés : RNase Z, TRZ1, ELAC2, NUC1, ENDOG, Glucose-6-phosphate mutarotase

Résumé : Trz1 chez la levure est responsable du clivage endonucléolytique à l'extrémité 3' au cours du processus de maturation des ARNt. Trz1 appartient à la famille des RNases de type β -lactamase, caractérisée par la présence d'un motif de séquence HxHxDH qui est impliqué dans la fixation des ions de zinc catalytiques. La famille des RNase Z est partagée en deux sous-familles de longueur de séquence différente : les formes courtes (250-400 acides aminés) et les formes longues qui sont approximativement le double en taille. Les structures cristallines des enzymes RNase Z de forme courte ont montré qu'ils sont actifs comme des homodimères. Une sous-unité englobe les ARNt substrat en utilisant un bras en saillie et l'autre fournit le site catalytique. Nous présentons ici la structure cristalline de Trz1, la première pour une RNase Z de forme longue. Trz1 est organisé en deux domaines reliés par un long peptide charnière. Chaque domaine est composé d'un repliement de type β -lactamase, et est caractérisé par la présence de plusieurs extensions

peptidiques. Le domaine N-terminal a perdu les résidus catalytiques au cours de l'évolution, mais il a maintenu le bras long qui est important pour le clivage de l'ARNt; tandis que c'est l'inverse pour le domaine C-terminal. L'interaction des deux domaines dans le monomère Trz1 mime le dimère de la forme courte de RNase Z et l'interaction des deux domaines est renforcée par les extensions. À partir d'études protéomiques, on avait observé que Trz1 forme un complexe ternaire d'une part avec NUC1, une nucléase mitochondriale impliquée dans l'apoptose, et d'autre part avec une mutarotase (codée par YMR099C). Nous avons purifié le complexe ternaire Trz1/NUC1/mutarotase et nous avons montré que la formation du complexe inhibe l'activité de Nuc1. Trz1/NUC1/mutarotase forme in vitro un hétérohexamère très stable en solution. À partir de nos données SAXS et MALLS nous proposons que l'homodimère NUC1 est au centre du complexe et que chaque sous-unité interagit avec une copie de Trz1 et une copie de mutarotase.

Title : The crystal structure of a long form tRNase Z from *S. cerevisiae* and study of its interactome

Keywords : RNase Z, TRZ1, ELAC2, NUC1, ENDOG, Glucose-6-phosphate mutarotase

Abstract: Yeast Trz1 is responsible for the endonucleolytic cleavage at the 3'-end during the maturation process of tRNAs. Trz1 belongs to the family of β -lactamase type RNases characterized by the presence of a HxHxDH sequence motif that is involved in coordination of the catalytic Zn-ions. The RNase Z family consists of two subfamilies: the short forms (between 250 and 400 residues), and the long forms which are approximately double in size. A few crystal structures of short form RNase Z enzymes showed that they are active as homodimers. One subunit embraces the substrate tRNA using a protruding arm and the other provides the catalytic site. We here present the crystal structure of Trz1, the first of a long form RNase Z. Trz1 is organized in two domains connected by a large linker. Each domain is composed of a β -lactamase type fold, and characterized by the presence of several extensions. The N-terminal domain has lost its

catalytic residues during evolution, but maintains the flexible arm that is important for tRNA cleavage; while it is the other way around of the C-terminal domain. The interaction of the two domains in the Trz1 monomer mimics the dimer of RNase Z short form; and the interaction of the two domains is reinforced by the extensions. From the proteomics studies, it is observed that Trz1 forms a ternary complex with NUC1, a mitochondrial nuclease involved in apoptosis, and with a mutarotase (encoded by YMR099C). We purified the ternary Trz1/Nuc1/mutarotase complex and show that complex formation inhibits Nuc1 activity. In vitro Trz1/Nuc1/mutarotase forms a very stable heterohexamere in solution. From our SAXS and MALLS data we propose that the Nuc1 homodimer is at the centre of the complex and that each subunit interacts with one copy of Trz1 and mutarotase.



Acknowledgement

I would like to express my gratitude to all the people who contribute to my thesis work and everyone helped me in any aspects during this four years.

First I would like to express my sincere thanks to the members of the juries who accepted to judge my thesis: The president of juries Philippe Minard; the reporters Claudine Mayer and Philippe Giegé; and the examiners Ciarán Condon and Ines Gallay.

I'd like to take this chance to express my great appreciation to my thesis director Herman van Tilbeurgh. I feel so lucky to have you as my thesis director, your professional, insightful and judgmatic view guide me to the accomplishment of my thesis. Your grasp on the main line of the project kept me on the right track and not fell into the repeating cycles when there was a problem. And your always ready-to-listen attitude make me overcome the worry on the communication barrier with others concerning the language. I couldn't appreciate more your tolerance when I met difficulties during writing, and the quick response and patience on the many rounds of correction.

My heartfelt appreciation on Ines Gallay co-directing on this project: it is your expertise not only on the crystallography but also on this project helped me making progress all through the thesis work. It was you opened the door and take me by the hand to teach me how to solve a structure step by step. Not only in work but also in other aspects you are always so nice to me and making me feel warm from the bottom of my heart.

My thesis work couldn't be completed without my dear other colleagues from the lab: I want to thank Nouredine Lazar for the always-on-line help on the biochemistry experiments, and the suggestions and reminders on the manner of working in a team for example asking for help proactively light me up when I was hesitating. Great gratitude to Dominique Durand for the patient explanation on both theory and practice on SAXS and for the always smiling kindness to me. I feel so lucky that I had the chance to spend my thesis in such a pleasant lab that everyone is generous and kind to share their knowledge and donate help in need: Dominique Liger and Bruno almost explained me all the aspects of enzymology involved in this work, ranging from the basic theory to the manipulation of the equipment, and always ready to answer my questions with their excellent English. Sylvie and Sophie don't supervise me directly but provide me suggestions when I need. Those talks with Karine raised me up when I was down. Stéphanie, Patrice, Antoine, Andrea, Wenhua and Dyana, Stéphane, Raoudh, Louisa, Théo: Thank you all for all the help to me.

I would also like to express my sincere appreciation to my collaborator Ciaran Condon, who was also the jury of mid-thesis, for the delicate tutorial on tRNA processing experiments, for the patient and quick response for all my questions, and for the correction of part of my thesis, etc. It was such an enjoyable experience working in your lab. I would also want to express my sincere thanks to the collaborators Jean-lepault for the electron microscopy, Yves Mechulam and Marc Graille for mutarotase activity assays, my thesis wouldn't be completed without your help.

I would also like to express my gratitude cordially to Agathe and Marielle for teaching me working with alpha-repeat proteins and the phage, I grew from zero to now able to proceed independently. The clear planning and explanations are prerequisites for me work efficiently. I would also want to thank Prof. Philippe Minard for those conversations not too long but full of essence on that part of work. I also got warm support from other members of the MIP team: Margarida, Thibault, Magali, Agnès, Sebastien and Corentin, thank you for the happy hours we spent together.

These beautiful memories with you would be the treasure of my entire life.

Thanks to China Council Scholarship for providing me the scholarship to conduct my PhD study.

Last but not the least I'd like to thank my family for the unconditional full support to me; and my dear friends who brings me so much joy in life.

Table of Contents

ABBREVIATIONS.....	- 1 -
INTRODUCTION.....	- 1 -
I. RNase Z.....	- 3 -
I.1 tRNA 3'-processing.....	- 3 -
I.1.1 tRNA.....	- 3 -
I.1.2 tRNA maturation.....	- 4 -
I.2 RNase Z.....	- 9 -
I.2.1 RNase Z belongs to the metallo- β -lactamase superfamily.....	- 10 -
I.2.2 Isoforms and distribution of RNase Z proteins.....	- 11 -
I.2.3 RNase Z for cell viability.....	- 13 -
I.2.4 RNase Z ^S structures.....	- 13 -
I.2.5 Structure of RNase Z in complex with product tRNA.....	- 19 -
I.2.6 Structure of RNase Z bound to precursor tRNA.....	- 22 -
I.2.7 The homology motifs affecting the tRNA catalysis/binding other than active site residues.....	- 24 -
I.2.8 RNase Z ^S : a cooperative enzyme.....	- 26 -
I.2.9 Prediction of the structure of an RNase Z ^L in complex with tRNA.....	- 28 -
I.2.10 RNase Z properties with respect to the organisms.....	- 28 -
I.2.11 Substrates of RNase Z.....	- 31 -
I.2.12 RNase Z ^L and diseases.....	- 33 -
II. Trz1/Nuc1/mutarotase complex.....	- 37 -
II.1 Identification of a Trz1/Nuc1/mutarotase protein complex.....	- 37 -
II.2 Localization of the subunits of the complex.....	- 37 -
III. Nuc1.....	- 39 -
III.1 Nuc1 is a major mitochondrial nuclease.....	- 39 -
III.2 Structures of Nuc1 homologs.....	- 41 -
III.3 Dual role of Nuc1.....	- 43 -
III.3.1 Apoptosis under high respiratory conditions.....	- 43 -
III.3.2 Nuc1 prevents cell death in glucose medium.....	- 45 -
IV. Glucose-6-phosphate mutarotase.....	- 46 -
IV.1 YMR099C encodes a glucose-6-phosphate mutarotase.....	- 46 -

V.	Objectives of the thesis	- 49 -
	RESULTS.....	- 51 -
I.	Crystal Structure of Trz1	- 53 -
I.1	Structure determination of Trz1	- 53 -
I.1.1	Purification of Trz1	- 53 -
I.1.2	Crystallization and optimization of Trz1	- 54 -
I.1.3	Diffraction and data processing of native Trz1.....	- 57 -
I.1.4	Structure determination of Selenomethionyl Trz1	- 57 -
I.2	Analysis of the crystal structure of Trz1	- 69 -
I.2.1	Overall structure.....	- 72 -
I.2.2	The N-terminal domain.....	- 74 -
I.2.3	The C-terminal domain	- 76 -
I.2.4	Comparison of the N- and C-terminal domains.....	- 81 -
I.2.5	The interaction of N- and C- terminal domain.....	- 85 -
I.2.6	Conclusion	- 86 -
II.	Reconstitution of Trz1/Nuc1/mutarotase Complexes.....	- 88 -
II.1	Purification of the individual Trz1, Nuc1and mutarotase proteins.....	- 88 -
II.1.1	Purification of Trz1	- 88 -
II.1.2	Purification of Mutarotase	- 89 -
II.1.3	Expression and purification Nuc1	- 89 -
II.2	Production of the complexes from Trz1, Nuc1 and mutarotase.....	- 91 -
II.2.1	Trz1 and mutarotase.....	- 91 -
II.2.2	Trz1/Nuc1 binary complex	- 92 -
II.2.3	Nuc1/Mutarotase binary complex.....	- 95 -
II.2.4	Reconstitution of the ternary Trz1/Nuc1/mutarotase complex.....	- 96 -
II.2.5	Purification of the Trz1/Nuc1/mutarotase complex from a polycistronic plasmid	- 99 -
II.3	Stoichiometry of the binary and tertiary complexes	- 101 -
II.4	Assemblage of the complex	- 104 -
III.	Enzymatic activities of the protein complexes.....	- 105 -
III.1	RNase Z activity and the effects of Nuc1/mutarotase	- 105 -
III.1.1	Substrates for measuring the RNase Z activity.....	- 105 -
III.1.2	RNase Z activity assay.....	- 107 -
III.2	Nuc1 nuclease activity and the effects of Trz1/mutarotase.....	- 111 -

III.2.1	Nuc1 nuclease activity towards dsDNA	- 111 -
III.2.2	Nuc1 nuclease activity towards fluorescent hairpin DNA.....	- 114 -
III.3	Mutarotase activity	- 116 -
III.4	Summary of the activity assays	- 118 -
IV.	Crystallization of the binary and ternary complexes	- 119 -
IV.1	Crystallization of the Trz1/Nuc1/mutarotase complex.....	- 119 -
IV.2	Crystallization of the Trz1/Nuc1 binary complex.....	- 121 -
IV.3	Crystallization of the Nuc1/mutarotase binary complex	- 121 -
V.	Structure study of Trz1/Nuc1/mutarotase and related complexes by SAXS..	- 125 -
V.1	Structural parameters and oligomerization state	- 125 -
V.2	Modeling of individual proteins using SAXS data	- 126 -
V.3	Analysis of the complexes in solution by SAXS	- 128 -
V.3.1	Analysis of the NM complex.....	- 128 -
V.3.2	Analysis of the ZNM complex.....	- 130 -
V.3.3	The model of ZN complex	- 130 -
V.3.4	Highlights from the SAXS results	- 131 -
	CONCLUSION & DISCUSSION	- 133 -
I.	Exploration of the RNase Z ^L structure	- 135 -
I.1	A model of Trz1 bound to its tRNA substrate	- 135 -
I.2	Catalytic mechanism of Trz1	- 138 -
I.3	Mutations related to prostate cancer susceptibility	- 139 -
I.4	Mutations related to hypertrophic cardiomyopathy	- 140 -
II.	Biological significance of Trz1/ Nuc1/mutarotase ternary complex model ...	- 142 -
	MATERIALS & METHODS.....	- 147 -
	ANNEX.....	- 166 -
	REFERENCE	- 172 -

ABBREVIATIONS

BME	β -Mercaptoethanol
IPTG	Isopropyl-1- β -D-thiogalactopyranoside
Mr	Molecular weight
SEC	Size exclusion chromatography
CV	Column volume
DTT	dithiothreitol
k_M	Michaelis constant
k_{cat}	catalytic constant
k_{ob}	observed constant
Glc6P	Glucose-6-phosphate
NADP	Nicotinamide adenine dinucleotide phosphate
Amp	Ampicillin
Kan	kanamycin
dNTP	Deoxynucleotide
dGMP	Deoxyguanosine monophosphate
SeMet	selenomethionine
SAD	single-wavelength anomalous dispersion
Rh	hydrodynamic radius
r.m.s.d	root-mean-square deviation

INTRODUCTION

I. RNase Z

I.1 tRNA 3'-processing

I.1.1 tRNA

Transfer RNAs (tRNA) carry specific amino acids at their 3'-end, which are “transferred” to the nascent peptide during translation. They also play a pivotal role in decoding the genetic 3-nucleotide message (codons) of the mRNA - by base-pairing with its anti-codon sequence. Together with the ribosome and many regulating protein factors, it is an essential element of protein synthesis.

Mature tRNA generally contains between 70 and 90 nucleotides, and its secondary structure is usually visualized as a “cloverleaf” (Figure 1). It is characterized by several features: 1) the acceptor stem containing 7 base pairs situated between the 5'- and 3'-ends; 2) the D arm, a 4-6 base pair loop that always contains a modified base, dihydrouridine; 3) the anticodon arm, presenting the anti-codon triplet for base-pairing with the genetic codon in the mRNA template during translation; 4) the variable loop (V loop), whose length differs from 3 to 21 nucleotides; 5) the T arm (TΨC), which is a loop region with a thymidine-pseudouridine-cytosine motif.

The “discriminator nucleotide”, the only unpaired nucleotide at the 3'-end of acceptor stem, is recognized by the CCA-adding enzyme (CCAse, nucleotidely-transferase) to start the synthesis of a CCA-triplet. The hydroxyl-group at the end of the CCA triplet is the attachment site of the specific amino-acid. The coupling between an amino acid and its specific tRNA is carried out by a dedicated aminoacyl-transferase. The discriminator nucleotide is an essential element for accurate aminoacylation (Redko et al. 2007; Vogel et al. 2005) .

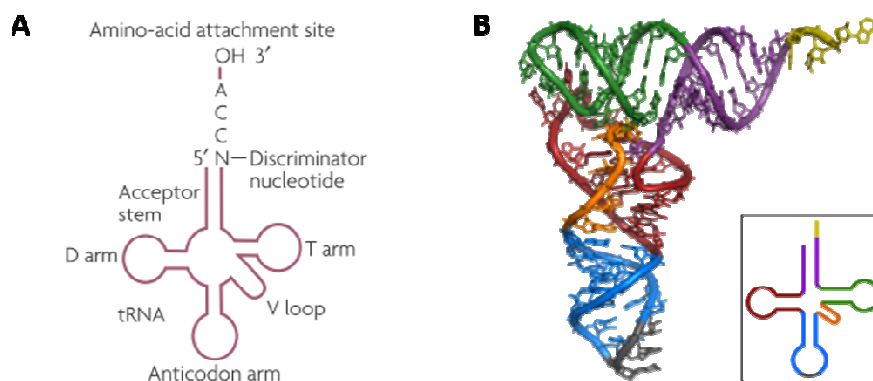


Figure 1 transfer RNA. (A) Features of mature transfer RNA (Redko et al., 2007). (B) Cartoon representation of tertiary structure of tRNA.

I.1.2 tRNA maturation

All tRNAs are transcribed as precursors with 5'- and 3'- extensions and require maturation to become functional. This process includes trimming of the 5' and 3'-extensions by endo- and/or exonucleases, addition of the CCA-triplet by tRNA nucleotidyl-transferase in many tRNAs, splicing of the introns and modifications of certain nucleotides. A specific amino acid is added to the last nucleotide of the CCA sequence by a dedicated tRNA aminoacyl synthetase, yielding an aminoacyl tRNA ready to engage in protein synthesis (Figure 3) (Hartmann et al. 2009; Späth et al. 2007). Besides these common maturation steps, tRNAs from yeast were also found to undergo cellular trafficking: export from the nucleus to cytoplasm, then retrograde transport to the nucleus, and then re-export to the cytoplasm (Phizicky & Hopper 2015; Hopper et al. 2010).

I will further focus my introduction of tRNA maturation process on the enzymes that are responsible for the trimming reactions. All these enzymes are RNases who are responsible for the cleavage reactions of the 5' and 3'-extensions. The 5'-leader sequence of pre-tRNA is ubiquitously trimmed by RNase P; whereas the 3'-trailer processing is often more complicated. The trimming processes may vary from organism to organism. In general, based on the type of ribonucleases involved, two pathways are referred: the endonucleolytic pathway, which trims the 3'-trailer from 5' to 3', and the exonucleolytic pathway trimming from 3' to 5'.

I.1.2.1 tRNA 5'- processing by RNase P

Ribonuclease P (RNase P), identified in the 1970s, is an essential endonuclease responsible for processing the 5'-leader of precursor tRNAs. RNase P enzymes exist in all domains of life and in nearly all species, and found as two types: RNA-based RNase P, and protein-only RNase P (PRORP) (Howard et al. 2013).

The RNA-based RNase Ps are ribonucleoprotein (RNP) complexes containing one conserved catalytic RNA subunit and a variable number of protein subunits (between one and ten) depending on the organism. They exist in bacteria, archaea, and many eukaryotes. In contrast, protein-only RNase Ps are found only in eukaryotes, especially mitochondrial/chloroplast RNase Ps and some nuclear RNase Ps are found to be this type,

although most eukaryotes nuclear RNase Ps are RNP based (Howard et al., 2013; Klemm et al., 2016; Salinas-Giegé, et al., 2015). RNase P enzymes hydrolyze a specific phosphodiester bond in a pre-tRNA by using divalent metal ions, resulting in the formation of a mature 5'-end containing a phosphate and a leader with 3' hydroxyl group (Figure 2). The hydrolytic and kinetic mechanisms of the two types of RNase P enzymes are similar, but the features unique to the RNA-based or PRORP enzymes are consistent with different evolutionary origins (Klemm et al. 2016; Howard et al. 2013).

5'-leader processing of precursor tRNA by RNase P usually happens prior to 3'-trailer processing in normal tRNA processing (Hartmann et al., 2009; Maraia & Lamichhane, 2011; Pellegrini, et al., 2003; Rossmannith, 2012).

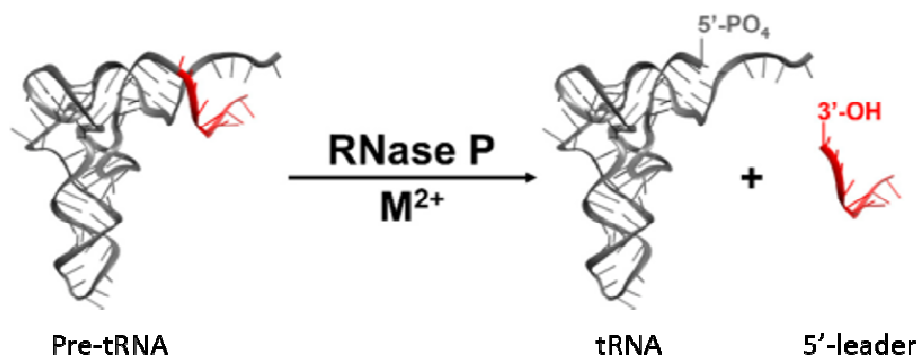


Figure 2 RNase P catalyzes of 5' leader sequences from precursor tRNAs using divalent metal ions (Klemm et al. 2016).

I.1.2.2 tRNA 3'- processing in prokaryotes

Whether the 3'-trailer of pre-tRNA is cleaved endonucleolytically or exonucleolytically greatly depends on how the CCA-triplet are added to the tRNA. For the CCA non-encoded tRNAs, the 3'-trailer is cleaved by RNase Z endonucleolytically right after the discriminator nucleotide in a single step (Figure 4A). The processing of tRNA from *Thermotoga maritima* forms an exception to this observation since its gene encodes the CCA triplet, but it is trimmed by RNase Z just after the CCA sequence (Figure 4A ii).

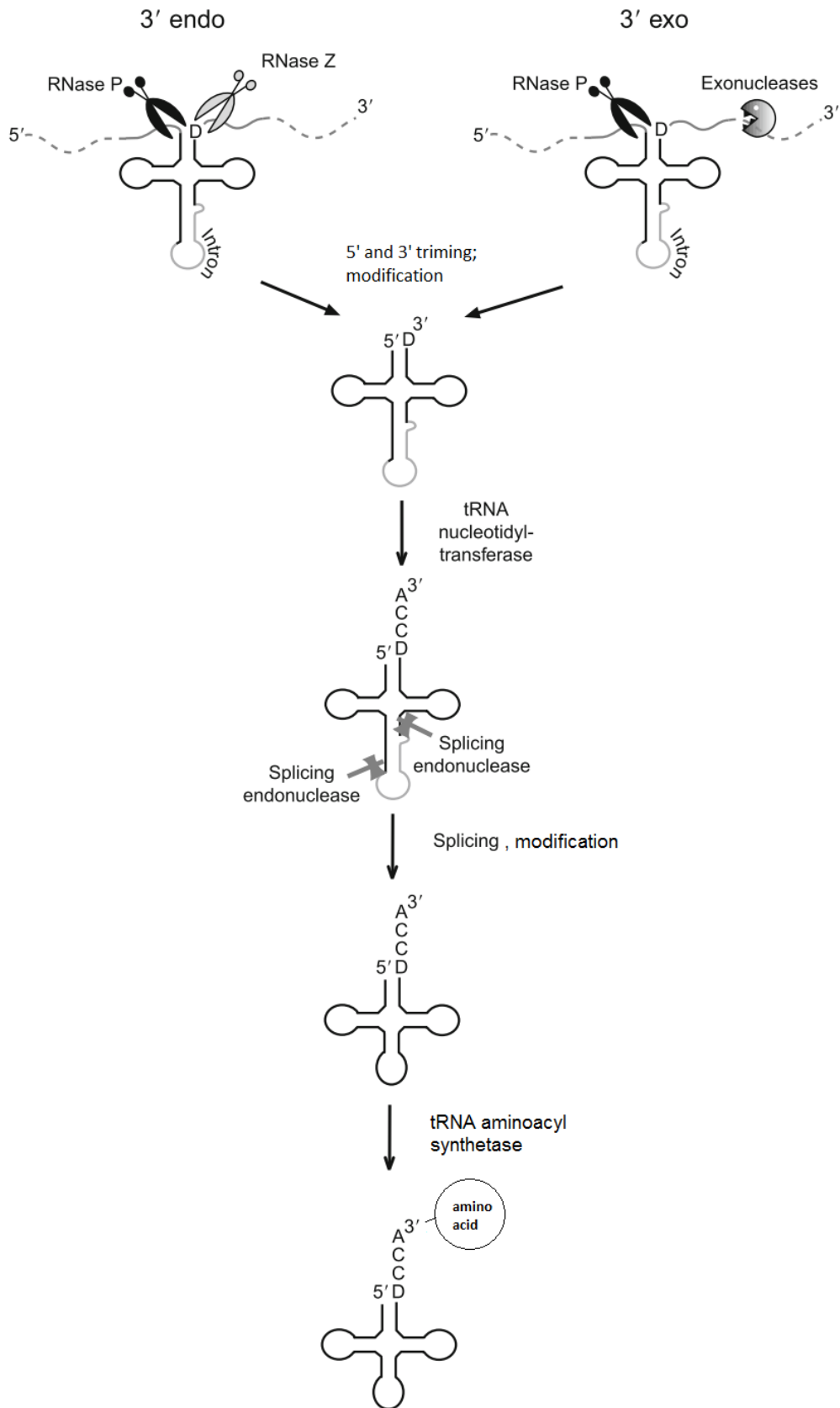


Figure 3 tRNA maturation.

(Adapted from (Hopper et al. 2010; Hartmann et al. 2009; Späth et al. 2007).

For the majority of the CCA-encoded tRNAs, the 3'-trailer is processed exonucleolytically in a multiple-step reaction. In *E. coli*, the 3'-end is firstly cleaved by the endonuclease RNase E, followed by RNase P trimming of the 5'-leader. The 3'-end is then further processed by several exoribonucleases such as RNase PH and RNase T. In the gram positive bacterium *Bacillus subtilis*, the main exoribonuclease involved in 3'-processing is RNase PH, assisted by other RNases such as RNase R, PNPase, YhaM (Redko et al. 2007). In bacteria and archaea, whether CCA-triplet is encoded in tRNA genes differs from organism to organism. For example, CCA is encoded in all types of tRNA genes in *E. coli*, consistent with the fact that the 3'-processing of tRNA in *E. coli* involves various exonucleases other than the endonuclease RNase Z. One third of *Bacillus subtilis* tRNAs are CCA non-encoded, therefore for these tRNAs the 3'-trailers are processed by RNase Z; those tRNA with an encoded CCA at the 3'-end are digested in multistep as described above (Hartmann et al. 2009). The CCA triplet is found to significantly suppress RNase Z activity (Pellegrini et al. 2003).

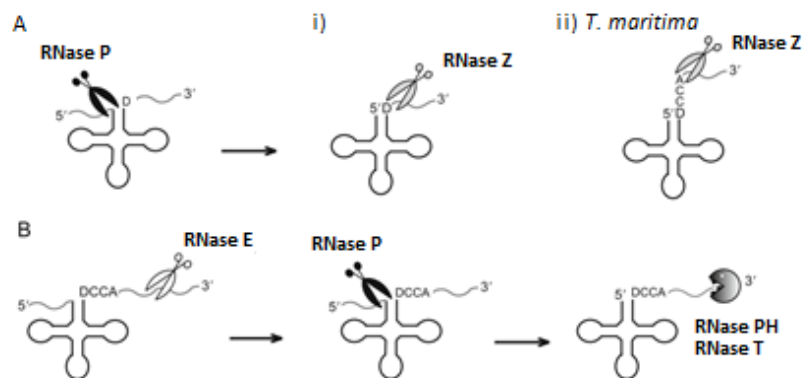


Figure 4 tRNA 3'-processing in prokaryotes. (A) Endonucleolytic pathway; (B) Exonucleolytic pathway (Hartmann et al, 2009).

I.1.2.1 tRNA 3'-processing in eukaryotes

Since none of the tRNAs from eukaryotes are CCA encoded, the endonucleolytic processing of the 3'-ends is executed by RNase Z followed by 5'-leader processing by RNase P, similar to processing of prokaryotic CCA-less tRNAs (Figure 3 “3'-endo”). This 5'-before-3' order appears to be the major pathway for tRNA processing (Kufel & Tollervey 2003).

But alternative pathways involving exonucleases and other protein complexes were also reported. 5'-processing is prior to 3'-processing or in reverse order is under the regulation of the protein La, which has RNA chaperon-like activity and differentiate the length of RNA. Nascent tRNA transcripts released from the RNA polymerase III have variable length 3'-

poly-U extensions. Those whose 3' extensions are of sufficient length associate stably with the protein La and are then processed by RNase P at their 5'-end followed by RNase Z cleavage at the 3'-end (Figure 5, right branch). In contrast, those tRNA 3'-oligo extensions are too short to bind the La protein are processed in an alternative way: the 3'-end is first cleaved by exonuclease such as Rex1p (possibly others), and then the 5'-end is processed by RNase P (Figure 5, left branch). In the cases of structural impaired tRNA precursors, the processing happens by neither of the two pathways described above. Instead, their processing is under nuclear surveillance by alternative pathway mechanisms involving exonucleases such as Rrp6p or others (Maraia & Lamichhane 2011).

Independent of which pathway the 3'-trailer is processed, RNase Z is the major endonuclease responsible for tRNA 3'- processing in bacteria, archaea, eukaryotes.

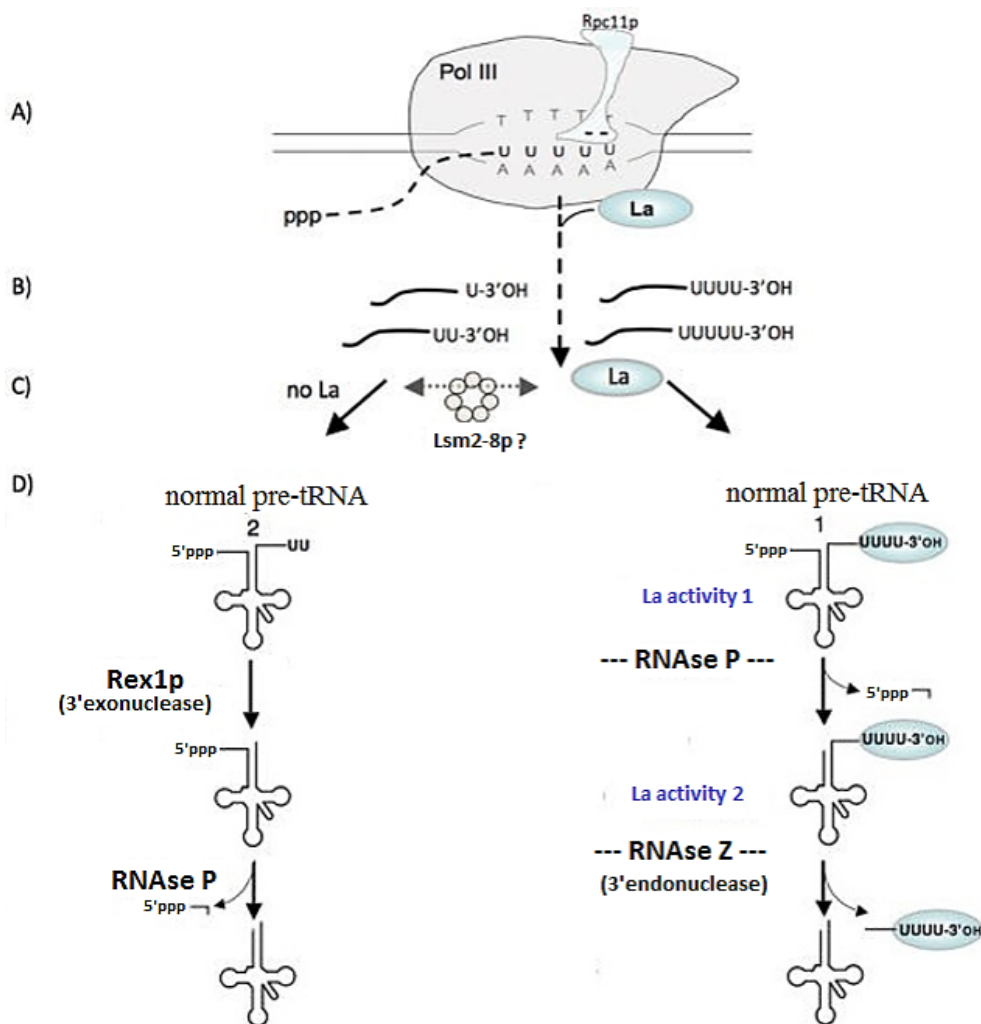


Figure 5 Alternative pathways for tRNA 3'-processing in eukaryotes (adapted from (Maraia & Lamichhane 2011)).

I.2 RNase Z

The discovery of tRNA 3'-end processing goes back to 1979, in which a single step endonucleolytic cleavage reaction was detected in assays of silkworm alanine tRNA processing by *Xenopus laevis* germinal vesicle extracts (Garber & Gage 1979; Hagenbüchle et al. 1979). The 3'-end endonucleolytic processing activity was subsequently characterized in yeast, *X. laevis*, *Drosophila melanogaster*, wheat, *Aspergillus nidulans*, potato and in mammals. However, the protein and DNA sequence of the responsible endonuclease, RNase Z, was not assigned until 2002. The first 3'-end type endonucleases whose sequences were identified were those from *Arabidopsis thaliana* and *Methanocaldococcus jannaschii*, constituting a breakthrough in the investigation on tRNA 3'-processing (Schiffer 2002). Based on the protein sequence of the found enzymes, homologues of RNase Z were identified throughout all three domains of life including bacteria, archaea, and eukaryotes (Condon & Putzer 2002; Pellegrini et al. 2003; Vogel et al. 2005; Ceballos & Vioque 2007).

Table 1 Uniform Nomenclature of RNase Z proteins

Organism	Protein accession no.	Old names		Function		New name (protein/gene)
		Gene	Protein	<i>In vivo</i>	<i>In vitro</i>	
Bacteria						
<i>E. coli</i>	P0A8V0	elaC	ZiPD EcoZ Elac1 Arylsulfatase RNase BN	Mutant made, none detected	tRNA processing, bpNPP	EcoTrz/trz
<i>B. subtilis</i>	P54548	yqjk rnz	YqjK	Processing of CCA-less tRNA precursors	Processing of CCA-less tRNA precursors	BsuTrz/trz
<i>T. maritima</i>	NP_228673	-	-	ND	tRNA processing	TmaTrz/trz
Archaea						
<i>M. jannaschii</i>	Q58897	-	-	ND	tRNA processing	MjaTrz/trz
<i>H. volcanii</i>	NA	-	-	ND	tRNA processing	HvoTrz/trz
<i>P. furiosus</i>	Q8U182	-	-	ND	tRNA processing	PfuTrz/trz
<i>T. acidophilus</i>	Q9HJ19	-	-	ND	tRNA processing	TacTrz/trz
<i>P. aerophilus</i>	Q8ZTJ7	-	-	ND	tRNA processing	PaeTrz/trz
Eukarya						
<i>S. cerevisiae</i>	P36159	YKR079c	-	tRNA processing*	tRNA processing	Trz1p, SceTrz/TRZ1
<i>C. elegans</i>	O44476	hoe-1 (Elac2 homolog)	-	RNAi: slow-growing, sterile	ND	CelTrz/trz-1
<i>D. melanogaster</i>	Q8MKW7	Jhl-1	3'-tRNase ELAC2	RNAi: tRNA processing (nucleus and mitochondrion)	tRNA processing	DmeTrz/Trz-1
<i>H. sapiens</i>	Q9H777	ELAC1	ELAC1	ND	tRNA processing	HsaTrz1/TRZ1
	Q9BQ52	ELAC2	ELAC2	ND	tRNA processing	HsaTrz2/TRZ2
<i>A. thaliana</i>	Q8LGU7	At1g74700	-	ND	tRNA processing	AthTRZ1/TRZ1
	Q8L633	At2g04530	-	ND	tRNA processing	AthTRZ2/TRZ2
	AAL49818	At1g52160	-	ND	tRNA processing	AthTRZ3/TRZ3
	NP_188247	At3g16260	-	ND	ND	AthTRZ4/TRZ4

(Vogel et al. 2005; Skowronek et al. 2014)

Since several proteins were found to have tRNA 3'-processing activity and before the name "RNase Z" was introduced, a uniform nomenclature of RNase Z as Trz was proposed

(listed in Table 1) (Vogel et al. 2005). *E. coli* RNase Z is encoded by the *elaC* gene, and the corresponding protein was used to be named ZiPD (zinc-dependent phosphodiesterase) and also RNase BN. For simplicity, we use “RNase Z” for general discussion; when referring to RNase Z from certain organisms the adapted organism-specific nomenclature will be used in this work, for example RNase Z from *B. subtilis* is referred as BsuTrz (Yqik). My thesis work concerns the RNase Z protein from *Saccharomyces cerevisiae* which will be referred to as Trz1.

I.2.1 RNase Z belongs to the metallo- β -lactamase superfamily

Before RNase Z was functionally characterized, the sequence of the *elaC* gene from *E. coli*, renamed as RNase Z afterward, assigned it to the metallo- β -lactamase (MBL) family.

I.2.1.1 The metallo- β -lactamase superfamily

The metallo- β -lactamase (MBL) family is characterized by a structural $\alpha\beta\beta\alpha$ -fold present in the metal-containing β -lactamases (MBL), which are hydrolytic enzymes employed by bacteria to resist β -lactam antibiotics. Members of the MBL family process a wide variety of substrates, most of which possess an ester linkage and a negative charge. The family is subdivided into four subclasses, the largest is class B β -lactamases (substrates: β -lactams), members of which are highly diverged proteins covering at least 16 identified biological functions found in all three domains of life. Apart from class B β -lactamases, the family also includes glyoxalase II (substrates: S-D-lactoylglutathiones), cAMP phosphodiesterases (substrate: cAMP), and the β -CASP proteins (derived from metallo- β -lactamase, CPSF, Artemis (ubiquitous DNA crosslink repair enzyme), SNM1, Pso2) acting on nucleic acids (Hartmann et al. 2009). Members of β -CASP subgroup as RNase J1, Int11, and CPSF-73 are ribonucleases: RNase J1 is involved in *thrS* mRNA leader and 16S rRNA processing; Int11 is a subunit of the integrator complex involved in snoRNA 3'-end processing; CPSF-73 is a subunit of the human cleavage and polyadenylation specificity factor (CPSF), and also found to be the pre-mRNA 3'-end-processing endonuclease. The structure of human CPSF-73 shows it contains the β -lactamase domain and a β -CASP domain (Aravind, 1999; Bebrone, 2007; Carfi et al., 1995; Daiyasu, et al., 2001; Hartmann et al., 2009; Redko et al., 2007).

The HxHxDH His-motif, which characterizes the members of this metallo- β -lactamase family, provides the ligands for metal coordination. The metal content of the members of this family includes two zinc, iron, or manganese ions. All members have the same β -lactamase

core domain, consisting of a parallel β -sheet sandwiched by α -helices. Substrate-recognition domains insert into this basic core domain, allowing the recognition of specific substrate and the accommodation to other functions, for example the flexible arm in the RNase Z family proteins for the binding of tRNA (Frazao et al., 2000; Daiyasu et al., 2001; Garauet al., 2005).

I.2.1.2 RNase Z belongs to the metallo- β -lactamase superfamily

RNase Z belongs to the zinc-dependent metallo- β -lactamase superfamily, characterized by the presence of a highly conserved Zn-coordinating signature motif HxHxDH (Figure 6) and the β -lactamase fold of its structure (details later).

I.2.2 Isoforms and distribution of RNase Z proteins

There exist two types of RNase Z: a short form (RNase Z^S), generally 280–360 amino acids in length; and a long form (RNase Z^L), which is approximately twice as long, and considered to be derived during evolution from RNase Z^S (Figure 6; Figure 7).

The structure of the short form of RNase Z displays the conserved the β -lactamase core, including the HxHxDH signature motif (motif II) coordinating zinc and forming the catalytic center. The unique feature that distinguishes RNase Z proteins from other β -lactamase family members is the presence of a “flexible arm” also called “exosite” which protrudes from the core, and is important for accommodating the tRNA substrate (details are presented further down in the manuscript). Several conserved motifs were defined throughout the RNase Z family proteins including (Figure 7): a PxKxRN loop and motif I involved in modulation of the RNase Z activity and cleavage specificity; motif II (Histidine motif) and motifs III to motif V whose histidines and aspartates contribute to the active center; the HEAT and HST motifs that are also involved in catalysis; the exosite; and a Ax Dx motif conserved in many RNase Zs and situated close to the PxKxRN loop in the 3D-structure (Karkashon et al., 2007; Wang et al., 2012) (Zareen, 2006).

The long form of RNase Z is typically 750–930 amino acids long, possesses an internal sequence similarity between the N-terminal and C-terminal halves, and is therefore considered to be derived during evolution from a duplication of the short form (Figure 6). The C-terminal half of RNase Z^L shares high sequence similarity with RNase Z^S, revealed by retaining the homology motifs described in RNase Z^S (Figure 7). However, sequence analysis shows the exosite domain does not exist in the C-terminal half, and is only retained in the N-terminal half. The other significant difference with between the N- and C-terminal halves is that the HxHxDH motif together with the other motifs contributing to the active site are

absent in the N-terminal half. Therefore, the N-terminal half was considered no longer able to form an active site. However, in the region surrounding the histidine motif some amino-acid conservation has been retained (Figure 6B), and the region is termed a pseudo His-motif (Ψ -motif II).

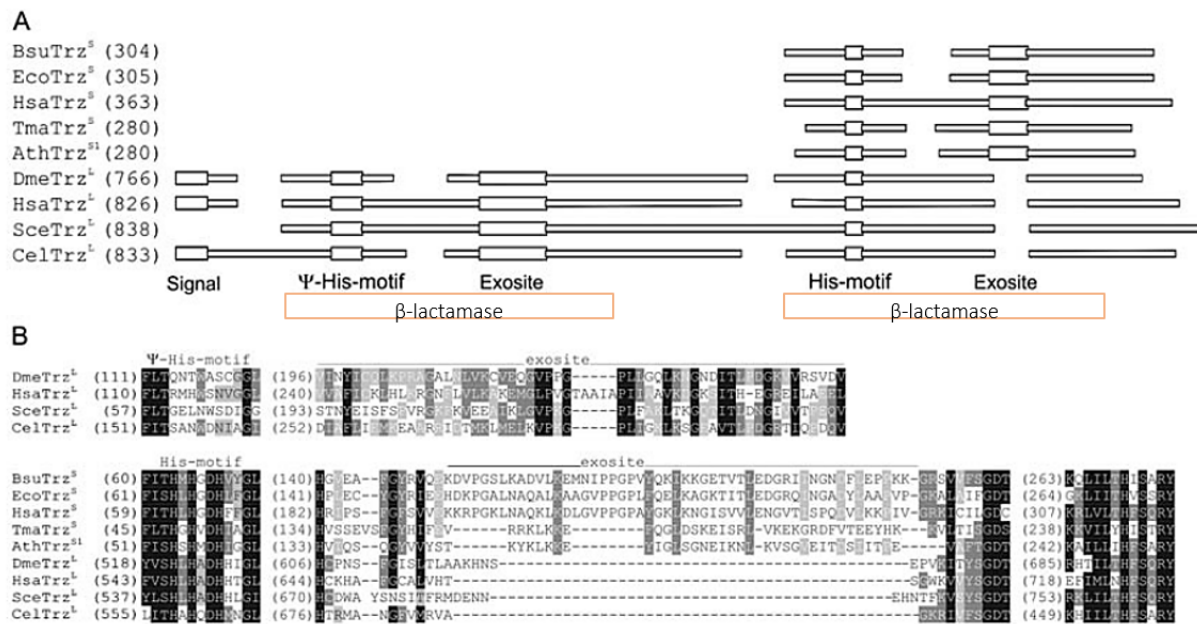


Figure 6 Primary sequence relations within short form and long form of RNase Z proteins (adapted from (Vogel et al. 2005).

(A) Schematic representation of representative RNase Z members drawn to scale. The relative position of functional motifs is indicated by rectangles. Breaks in the line represent gaps in the respective sequence. The length of the proteins is given in brackets. (B) Excerpt of a multiple alignment focusing on special features corresponding to the scheme above. More conserved amino acid residues are shaded in a darker tone. The amino acid numbering is given before the sequence in brackets. Metal coordinating residues are marked with an asterisk below the alignment.

The short form of RNase Z is present in all three domains of life; whereas the long form RNase Z^L is found exclusively in eukaryotes. Archaeal and bacterial organisms encode only a single short RNase Z protein. The only convincing exception is the archaeon *Haloquadratum walsbyi* in which two RNase Z paralogues have been reported. Generally, eukaryotes have both forms RNase Z. For example, Homo sapiens has one RNase Z^L and one RNase Z^S. The plant *A. thaliana* has two RNase Z^L and two RNase Z^S enzymes. Exceptions are *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, and *Drosophila melanogaster* have only single RNase Z^L enzyme, but lack the short variant. (Redko et al. 2007; Vogel et al. 2005).

I.2.3 RNase Z for cell viability

RNase Z is essential for *B. subtilis* cell life. In a *B. subtilis* mutant strain for which the chromosomal tRNase Z gene was put under control of an IPTG-dependent promoter, the removal of IPTG from the culture medium resulted in complete growth inhibition (Pellegrini et al. 2003). However, a comprehensive characterization of an *E. coli* RNase Z (elaC) deletion strain revealed wild-type-like viability and unaffected transcriptome and proteome profiles (Schilling et al. 2005). Deletion of the TRZ1 gene, encoding the RNase Z, is lethal for *S. cerevisiae*, although yeast can tolerate altered expression of TRZ1 (Chen et al. 2005). *Schizosaccharomyces pombe* encodes two RNase Z^L genes, both are essential for cell growth (Zhao et al. 2009).

I.2.4 RNase Z^S structures

Four RNase Z^S structures had been solved, three of them are from bacteria, which are *B. subtilis* (Li de la Sierra-Gallay et al. 2005), *E. coli* (Kostelecky et al. 2006) and *T. maritima* (Ishii et al. 2007) resolved at 2.1 Å, 2.9 Å and 1.97 Å, resolution respectively (PDB code 1y44, 2cbn, 2e7y, respectively). The only eukaryotic RNase Z^S whose structure was solved was from human at 1.7 Å (PDB code: 3zwf).

I.2.4.1 Free RNase Z^S structures from *B. subtilis*

The first crystal structure of RNase Z was from *B. subtilis* solved in 2005 (Figure 8). The protein is present in the crystal as a head to head homodimer. Each monomer consists of a β -lactamase domain, made of two opposing seven-stranded β sheets on each side covered by three α -helices forming a $\beta\beta\beta\alpha\beta\alpha\beta$ motif. The first four β strands are antiparallel (β 7, β 8, β 9, β 12 and β 16, β 1, β 2, β 3) and the next three are parallel in each of the β -sheets. The characteristic element distinguishing RNase Z from other members of the Zn-dependent metallo- β -lactamase family is a flexible arm (called “exosite”) inserted between the third and fourth β strand (β 9 and β 12) and protruding from the core β -lactamase domain (Figure 8B). This flexible arm is essential in substrate recognition (see below).

The major interface between the two monomers is contributed by the central α -helices (α 1, α 2, α 3), and β 1- β 2 loop of each monomer. The main hydrogen-bond contacts (<3.3 Å) are between residues Ala41, Gln43, His44, Ser103 and Thr107 of one monomer and Asn18, Glu40 and Tyr70 of the other. Reciprocal inter-subunit hydrogen bonds are formed between Asn18–His44, Glu40–Gln43, Glu40–Ala41, and between Tyr70 and both Thr107 and Ser103.

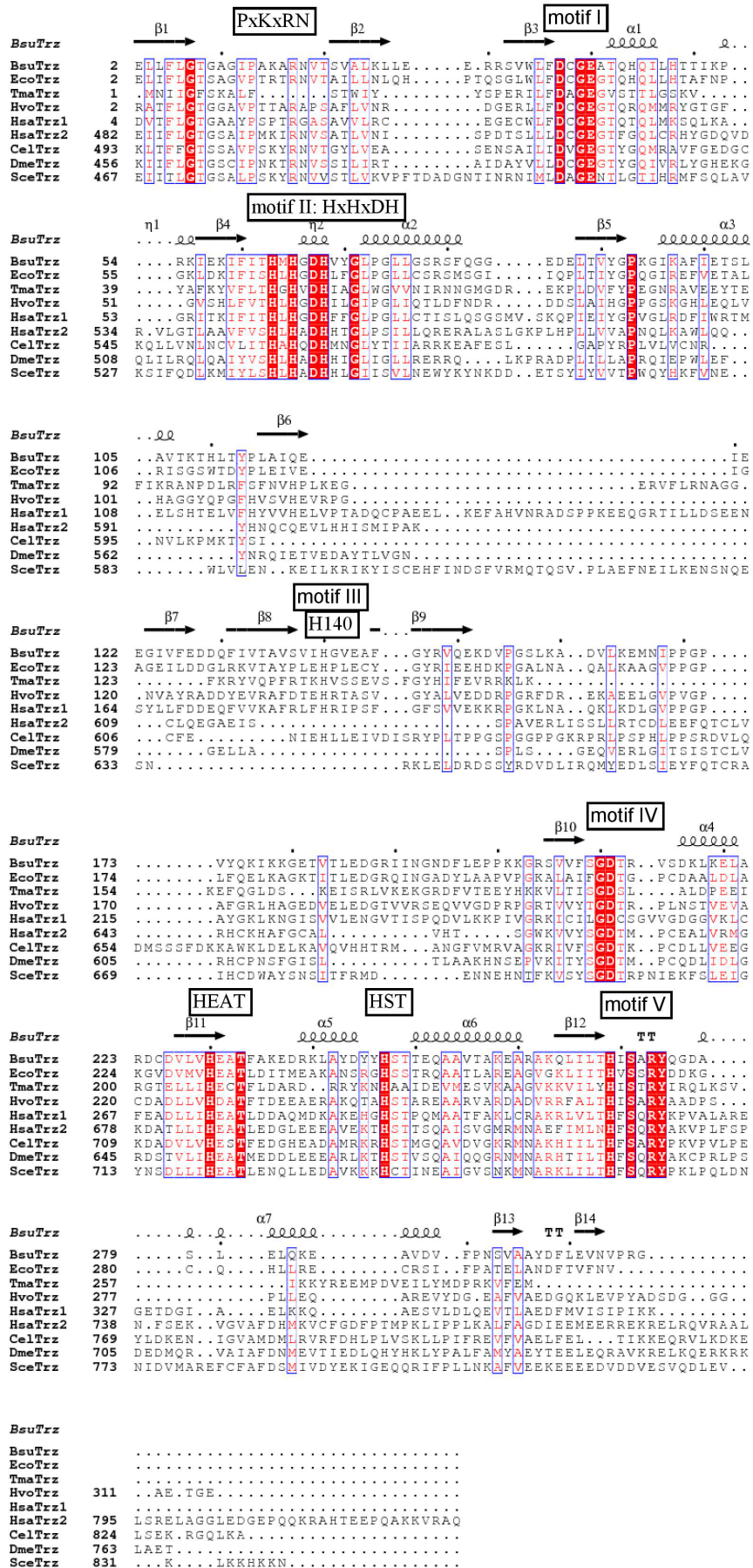


Figure 7 Sequence alignment of RNase Z^S and the C-terminal halves of RNase Z^L.

The conserved motifs modulating the activity are labelled above the sequences. Secondary structure elements as extracted from the crystal structure of *B. subtilis* RNase Z^S (PDB code 1y44) are superposed onto the sequences. RNase Z^S from *B. subtilis*, *E. coli*, *T. maritima*, *H. volcanii* and human, RNase Z^L from human, *C. elegans*, *D. melanogaster*, and *S. cerevisiae* are represented. Sequences were aligned by Blast, and colored by Esprit (Robert & Gouet 2014). The motif III is not well aligned.

The two monomers in the crystallographic structure are not identical in the *B. subtilis* RNase Z structure. The flexible arm is only resolved in one subunit (subunit B) but not in the other (subunit A); the two zinc ions and a bound phosphate are only found in the A subunit and not in subunit B. The A subunit has a well-structured $\alpha 7$ helix located between the first and second $\alpha\beta$ repeats of the same sheet which is $\beta\beta\beta\downarrow\beta\alpha\beta\downarrow\alpha\beta\alpha\beta$, which is important for substrate recognition, but is disordered in subunit B.

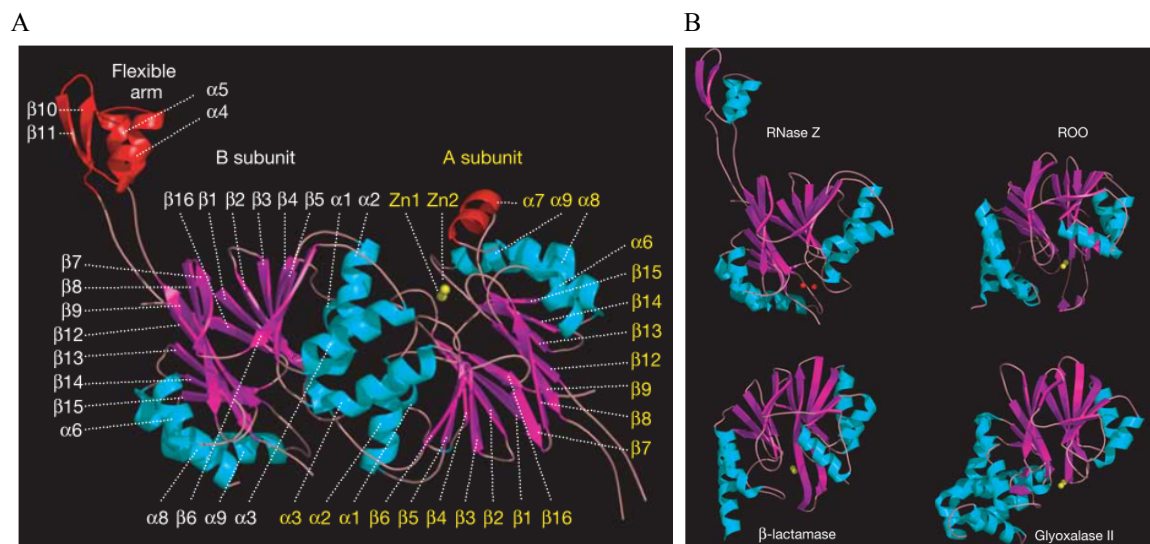


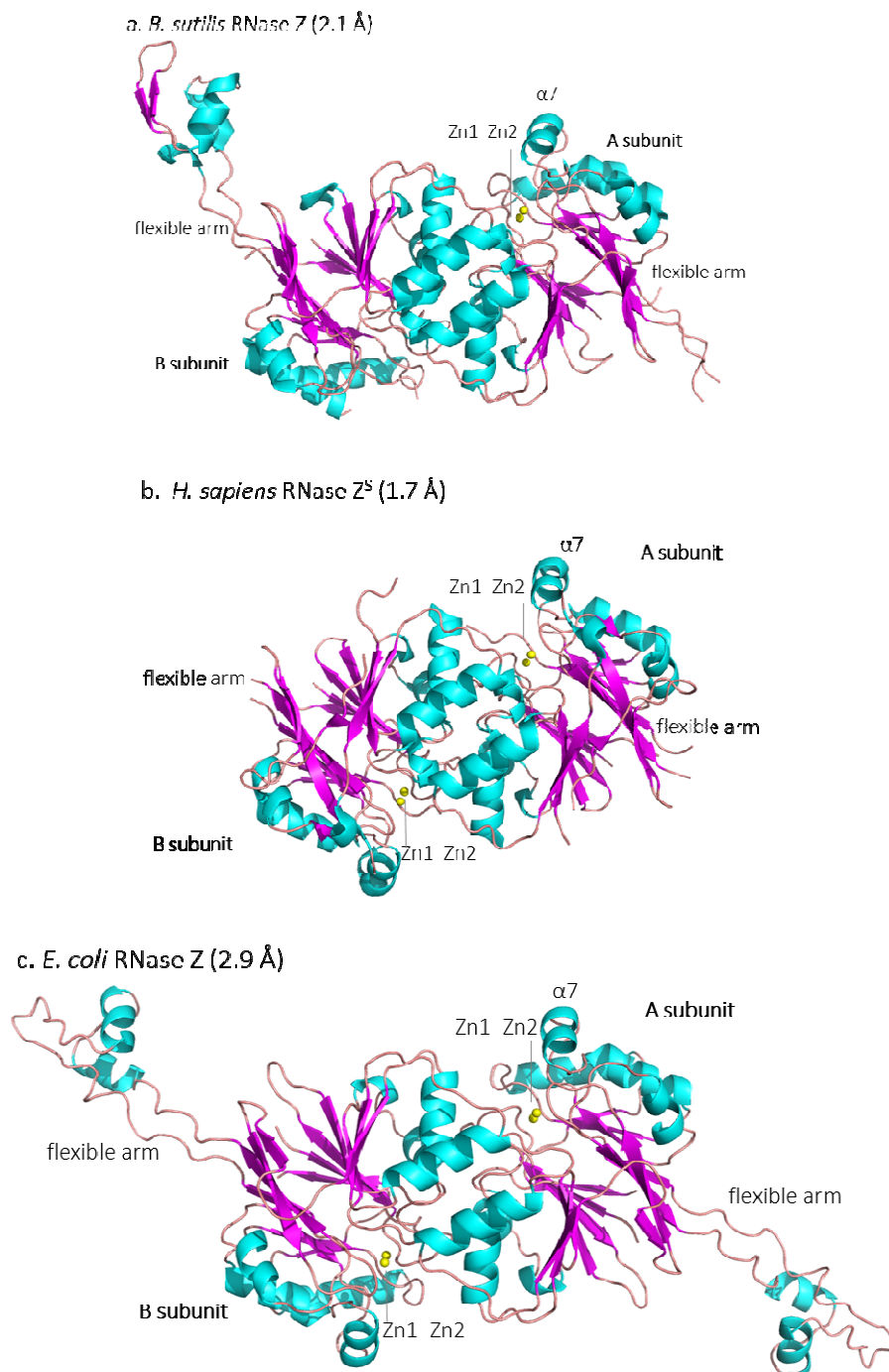
Figure 8 Structure of RNase Z holds a β -lactamase fold (Li de la Sierra-Gallay et al. 2005).

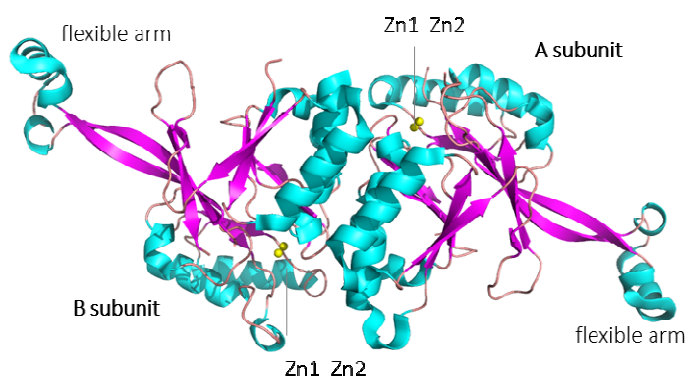
(A) Crystal structure of RNase Z from *Bacillus subtilis* (PDB code 1y44). (B) Structural comparison of members of the metallo- β -lactamase family.

I.2.4.2 Free RNase Z^S structures from other organisms

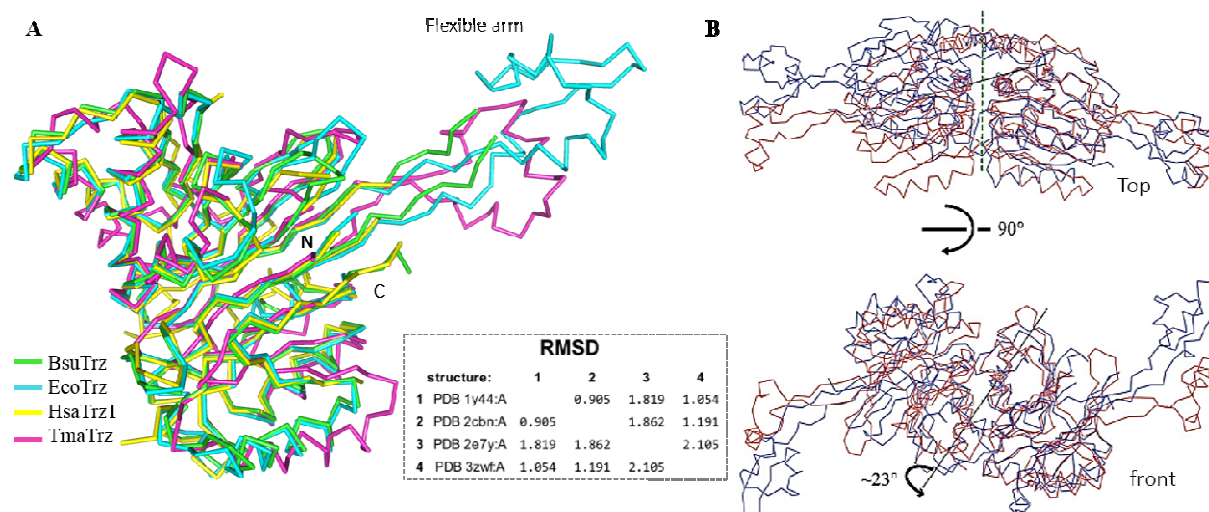
Three other RNase Z^S structures from *E. coli*, *T. maritima* and human confirm the presence of a β -lactamase domain and associate into the same dimeric configuration (Figure 9) (Li de la Sierra-Gallay et al. 2005; Ishii et al. 2007; Schilling et al. 2005; Ishii et al. 2005; Redko et al. 2007). All of them contain a flexible arm protruding from the β -lactamase core, although it is not resolved in the ELAC1 structure. Superposition of one subunit of EcoTrz, TmaTrz, and HsaTrz^S (ELAC1) structure with that of BsuTrz gives a r.m.s.d. of 0.905 Å, 1.819 Å, 1.054 Å, respectively (Figure 10A). It had been discussed afore that one subunit of BsuTrz dimer is active while the other is inactive. But in other three RNase Z^S structures, the

two subunits are identical. In the catalytic site two zinc ions are bound to each monomer, although the two zinc ions are with different occupancy in TmaTrz structure (Ishii et al. 2007), which explains that in a lower resolution TmaTrz structure (2.6 Å) only one zinc ion is found in each monomer (Ishii et al. 2005). No extra zinc was added during the purification or crystallization of BsuTrz and TmaTrz, meaning that the enzyme is capable to extract the zinc from the culture broth; whereas extra zinc was added in the crystallization solution of EcoTrz, indicating full occupancy of zinc in the crystals probably requires addition of the metal during crystallization.



d. *T. maritima* RNase Z (1.97 Å)**Figure 9 Structures of short form RNase Zs.**

Structures from (a) *B. subtilis*, (b) human, (c) *E. coli* and (d) *T. maritima* are represented in cartoons (PDB code: a. 1y44; b. 3zwf; c. 2cbn; d. 2e7y.), helix are colored in cyan and β -strand and loops are colored in purple.

**Figure 10 Comparison of structures of short form RNase Zs.**

(A) Superposition of subunit A of RNase Z^S structures from *B. subtilis*, *E. coli*, *T. maritima* and human (PDB code 1y44, 2cbn, 2e7y, 3zwf, respectively). Structures were aligned by Protein structure comparison service PDBeFold at European Bioinformatics Institute (<http://www.ebi.ac.uk/msd-srv/ssm>, E. Krissinel and K. Henrick) and represented as α -carbon chains. The N-terminal end (N) and C-terminal end (C) were labelled. The r.m.s.d. values of paired alignment are listed in inset. (B) Superposition of the TmaTrz dimer (red) and BsuTrz dimer (blue) (Ishii et al. 2007). The right subunits of the dimers were superimposed using the Secondary Structure Matching server (Krissinel & Henrick, 2004). The black line indicates the screw axis to superimpose one protomer of TmaTrz onto that of BsuTrz and the green dashed line indicates the pseudo-twofold axis of the TmaTrz dimer. The pseudo-twofold axis is parallel (top panel) and perpendicular (bottom panel) to the paper.

Only one of the flexible arms from the inactive subunit of the dimer was resolved in the BsuTrz structure. In the ELAC1 structure, neither flexible arms of the dimer were resolved, although two zinc ions are full loaded in each monomer. In EcoTrz and TmaTrz, both flexible arms in the two monomers were well structured (Figure 9c and 9d). A 15° difference of the

angles between the exosite and the main core body in the two subunits of TmaTrz are observed in the crystal structures, indicating the flexibility of the arms. The crystal contacts with adjacent dimers seem to be helpful for the stabilization of the flexible arms (Ishii et al. 2007). The presence of Zn at the active site and at least one structured arm indicate that the enzyme could adopt a fully active form without the binding of the tRNA substrate. As discussed further, the different conformational states of the two subunits of BsuTrz might relate to the cooperative nature of RNase Z (later in the manuscript) (Redko et al. 2007).

A comparison of the three bacterial RNase Z structures revealed differences in their dimer orientation, which may be related to the unique cleavage-site specificity of TmaTrz (Figure 10B) (Ishii et al. 2007).

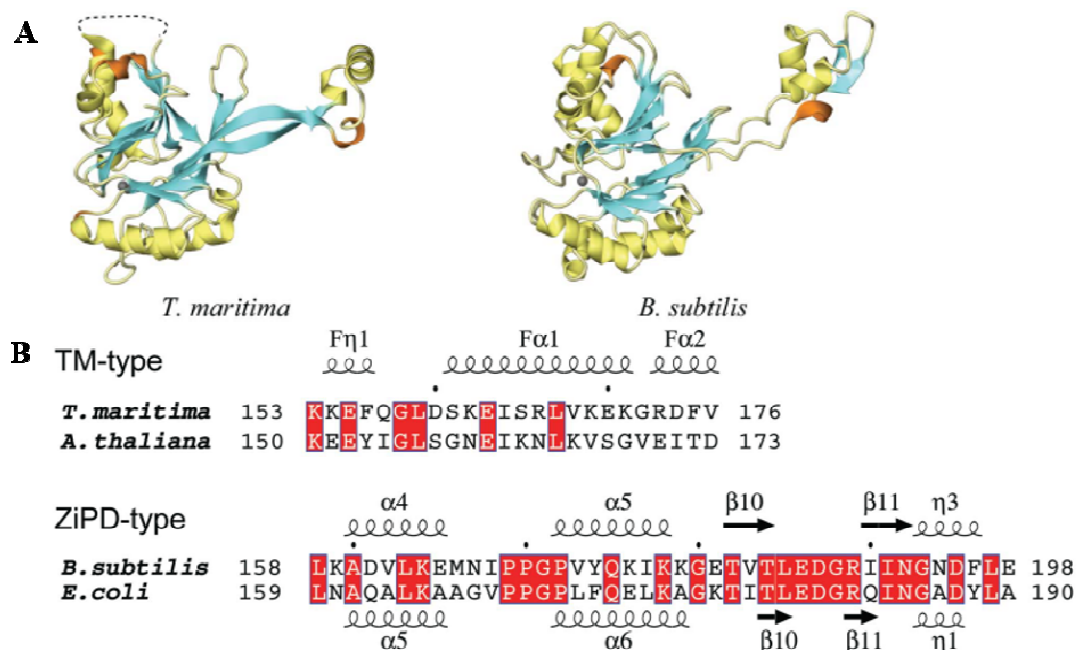


Figure 11 Two types of flexible arms of RNase Z [adapted from (Ishii et al. 2007)].

(A) Ribbon diagram displaying the overall monomeric structures of two RNase Zs. The α -helices, β -strands and 3_{10} -helices are colored yellow, cyan and orange, respectively. Zn ions are depicted as grey spheres. (B) Sequence alignment of the flexible arms of RNase Zs. The red boxes show the residues conserved in each type of RNase Z.

I.2.4.3 Two types of flexible arms are found in RNase Z family proteins

There exist two types of flexible arms among the RNase Z proteins: the ZiPD- and the TM-type (Figure 11). The ZiPD-type contains a glycine/proline-rich motif (GP-motif), and is about 50 amino acids in length (for example residues 151-203 in BsuTrz). The flexible arms of BsuTrz, EcoTrz, MjaTrz and HsaTrz1 (HsaTrz^S) are of this type according to the sequence conservation. The ZiPD-type flexible arm of BsuTrz (Li de la Sierra-Gallay et al. 2005) and EcoTrz (Kostelecky et al. 2006) contains two α -helices, two β -stands and one 3_{10} -helix (η)

(Figure 11). By contrast, in the TM-type of flexible arm, a cluster of 4-5 basic residues replaces the glycine/proline motif, and is shorter in length (23 residues in the case of *T. maritima*) (Schilling et al. 2005; Ishii et al. 2007). The structure of TmaTrz shows two long β -strands form the two “arms” of the exosite, holding the “fist” of the exosite consisting of one long α -helix and two short helices (one α -helix and one 3_{10} -helix). *Arabidopsis thaliana* has TM-type of exosite, and the same are found in many green plants RNase Z (Fan et al. 2011).

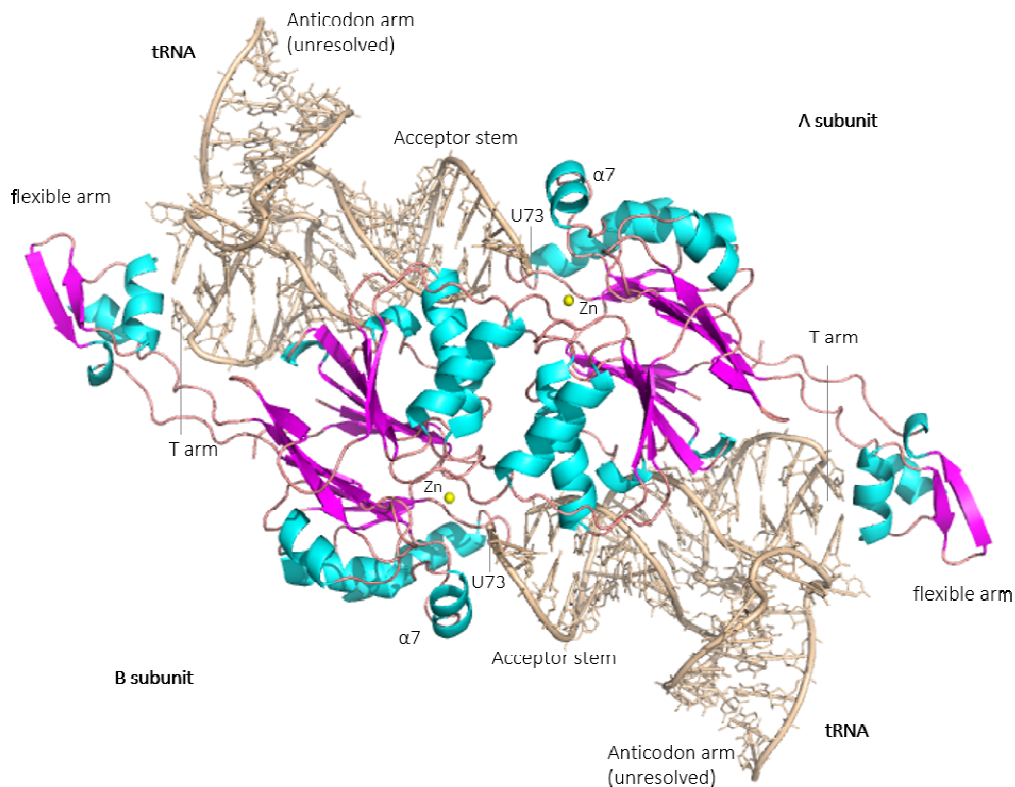


Figure 12 Structure of BsuTrz bound to tRNA (PDB code: 2FK6).

BsuTrz are shown in cartoon colored cyan (helices) and purple (β -strands and loops); and tRNAs are shown in sticks (golden), the T arm, acceptor stem and anticodon arm are labelled. Zinc ions are shown in yellow sphere.

I.2.5 Structure of RNase Z in complex with product tRNA

I.2.5.1 Overall structure

The first structure of an RNase Z-tRNA complex solved at 2.9 Å was that of BsuTrz bound to an tRNA^{Thr}. BsuTrz was present as a monomer in the asymmetric unit and complexed with one molecule of tRNA^{Thr}, but the functional dimer was created by crystallographic symmetry showing that two tRNAs are accommodated by one single dimer (Figure 12). The T arm and acceptor stem of the tRNA lie across the top surface of the dimer,

clamped between the flexible arm of one subunit and helix $\alpha 7$ of the other subunit where the 3'-end is processed. This configuration is consistent with the fact that the anticodon and D arms are not necessary for tRNA 3'-trailer cleavage (Nashimoto et al., 1999; Schiffer et al., 2001).

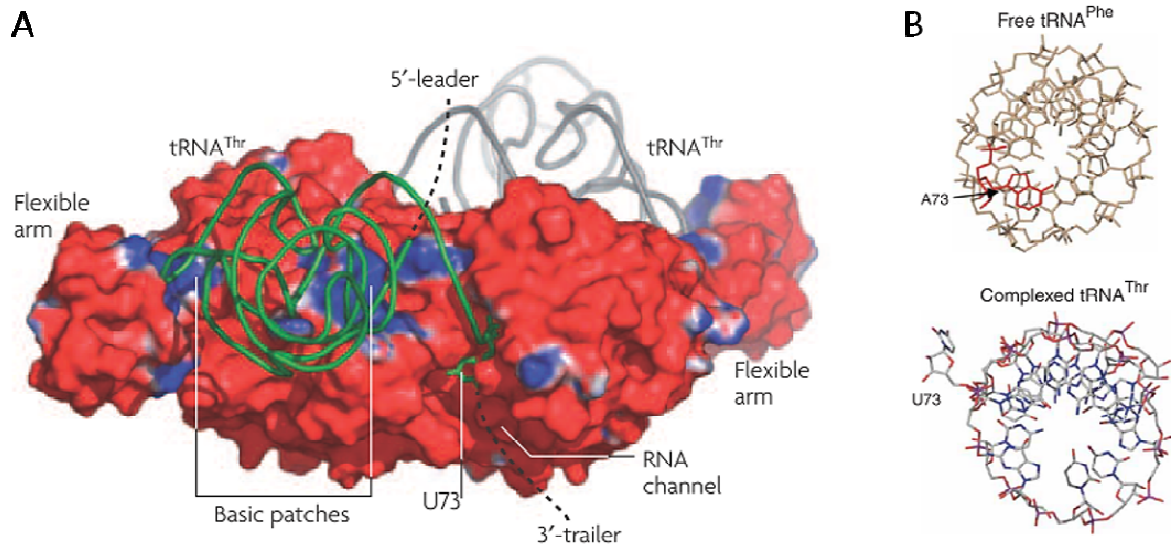


Figure 13 Mode of tRNA binding with RNase Z.

(A) Electrostatics surface representation of BsuTrz/tRNA complex. (B) Conformation change of tRNA from free to RNase Z-complexed form (Redko et al. 2007; Li de la Sierra-Gallay et al. 2006).

The region of BsuTrz that recognizes tRNA mainly involves two lysine/arginine-rich basic patches: one patch is situated on the “fist” of the flexible arm, and the other lies on the top surface of the same subunit (Figure 13A). On the tRNA side, most contacts with the RNase Z are established with the phosphodiester backbone, giving the possibility to bind a variety of tRNA species with different sequences. A few bases that are engaged in specific binding with RNase Z such as guanosine at position 1 and 19 are highly conserved among all tRNA sequences. In order to adapt a position to enter the catalytic site of RNase Z, the discriminator nucleotide U73 requires to be splayed out by 90° from the acceptor stem axis relative to its position in free tRNA where it stacks to the acceptor stem helix (Figure 13B)

The RNase Z-tRNA complex was solved using enzymes with a single mutation at the active site (H65A), in order to slow down the catalysis. But no nucleotide was found in the electron density of the crystal beyond the discriminator nucleotide (Uracil 73), indicating that the 3'-trailer of tRNA precursor used in crystallography had been cleaved during crystallogenesis, although only one zinc ion coordinated to the enzyme. Thus this structure is

considered to be an RNase Z in complex with its cleaving product, rather than with the precursor (Li de la Sierra-Gallay et al. 2006).

The channel that leads from the catalytic center to the exterior of the protein is just wide enough for providing access of a single RNA strand. This channel is proposed to be the exit of the 3'-trailer product that is cleaved off by the RNase Z, explaining why tRNA with a 5'-extension are not cleaved by RNase Z (5'-extensions actually inhibit RNase Z). Because the longer 5'-extension is, the more possibility that it would pair with 3'-trailer sequence, the double strand tRNA would be too big to pass through the channel of RNase Z.

I.2.5.2 The active center and catalytic mechanism

The active center of BsuTrz is made up of the conserved histidines and aspartates from motifs II to V, coordinating two zinc ions, bridged by one bound phosphate ion (Figure 14). The two zinc ions are coordinated by three histidines (His63, His65, His68) and one aspartate (Asp67) from the signature Histidine-motif (motif II), one His140 from motif III, one His269 from motif V, and further Asp211 from motif IV, which binds both zinc ions.

Furthermore, the position of His68 is further stabilized by hydrogen bonds with Asp37 from the motif I and Thr62. Mutation in Asp37 inactivates RNase Z (Minagawa et al., 2004). The phosphate ion is also bound by His247 from the HST motif, which is bound to Glu231 from the HEAT motif. Mutation of His247 or Glu231 significantly reduces catalytic activity in RNase Z^S (Li de la Sierra-Gallay et al. 2005; Minagawa et al. 2006; Späth et al. 2005). The glutamate from the HEAT motif and histidine from the HST motif act together to modulate substrate binding (Karkashon et al. 2007).

The positions of the Zn ions compared to that of the fortuitously cocrystallized phosphate ion in the active site led to the proposal of a plausible reaction mechanism (Figure 14C). Negatively charged Asp67 deprotonates a nearby water molecule, creating a hydroxide ion that carries out a nucleophilic attack of the phosphate ester linkage polarized by the two zinc ions. This breaks the phosphodiester bond between nt73 and nt74 of precursor tRNA. Then the 3' oxygen of the discriminator nucleotide (nt73) is reprotonated by another nearby water, resulting in a tRNA molecule with a 3'-hydroxy group (Redko et al. 2007).

The β -lactamase core of the A or B subunit of the RNase Z in tRNA free form are very similar to those in the complex with respective r.m.s. deviations of 0.51 Å and 1.19 Å for 245 C α positions. Upon superimposing the C α traces of a single subunit of the free and

complexed dimers, a slight rotation (1.75°) of the subunits relative to one another is noted (Redko et al. 2007).

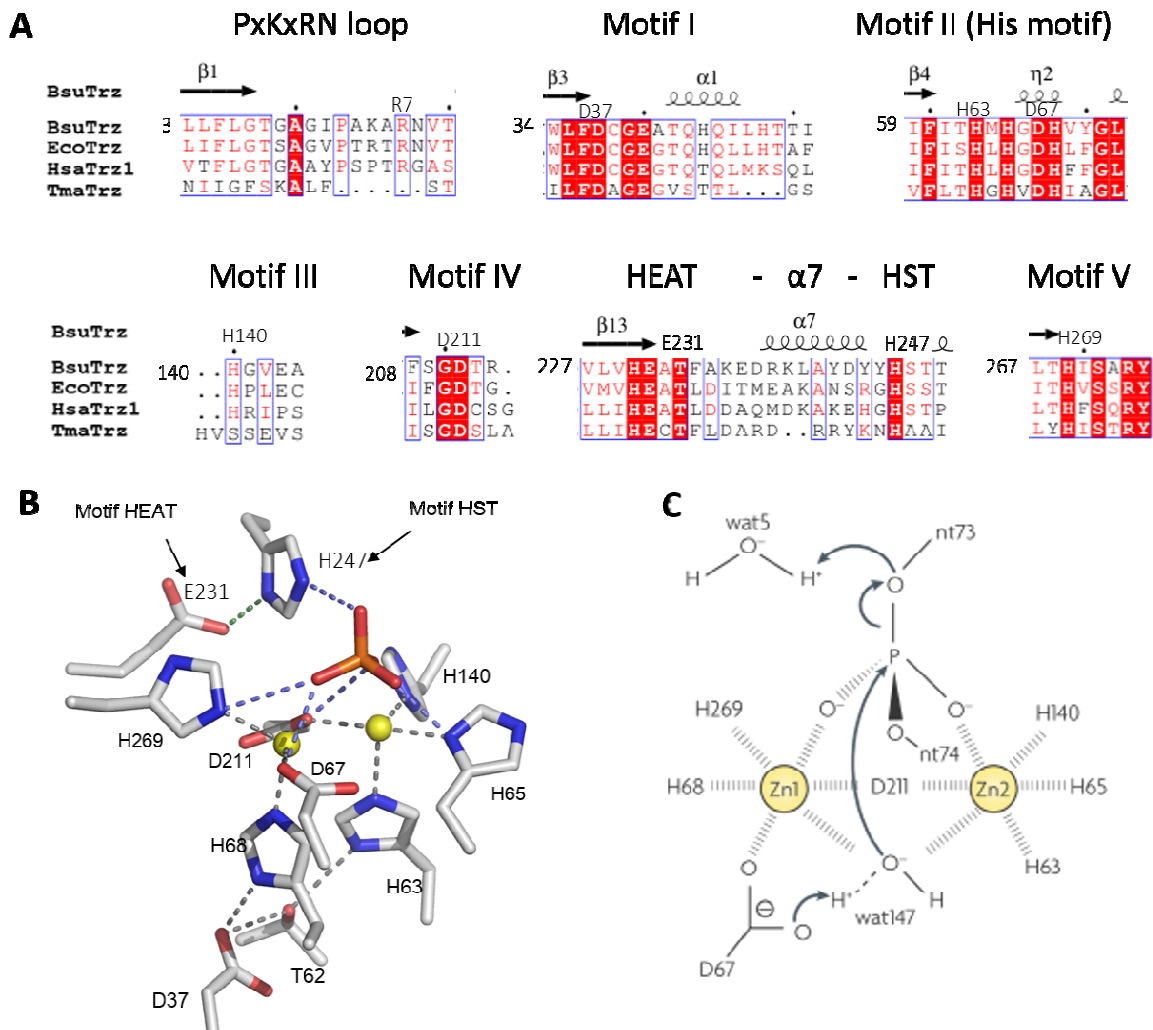


Figure 14 Active site of RNase Z and catalysis mechanism.

(A) Conserved motifs in RNase Z^S structures involved in catalysis and substrate binding. (B) The active site of BsuTrz (PDB code: 1y44, subunit A). (C) Presumed mechanism of catalysis (Redko et al. 2007).

I.2.6 Structure of RNase Z bound to precursor tRNA

The structure of the BsuTrz in complex with precursor tRNA^{Thr} was solved (Pellegrini et al. 2012). In order to prevent cleavage by RNase Z the discriminator and trailer UAA of the precursor tRNA were protected by 2'-O-methyl groups. The overall structure is very similar to the crystal structure of RNase Z in complex with its tRNA product (r.m.s.d. of 0.43 Å for 559 Cα positions of the dimer) (Figure 15). The most important finding for this structure is that it reveals a specific binding pocket for U74 adjacent to the catalytic center (Figure 16). The residues in contact with U74 are mainly from helix α2 of one subunit and binding

majority of tRNA, while the active site is from the symmetry subunit. The N3 hydrogen of uracil forms a hydrogen bond with both the side chain oxygen of Ser78 and the main chain oxygen of Gly74 (3.1 Å and 3.5 Å, respectively); at the mean time the O2 oxygen of U74 interacts with main chain oxygen of Gly77 (3.0 Å). On the other hand, the side chain amino group of Gln43 from helix α 1 is a potential hydrogen bond donor to the O4 oxygen of uracil (3.5 Å), while its side chain oxygen can accept a hydrogen bond from the Ser78 hydroxyl group (2.8 Å) (Figure 16 A-C).

Activity studies on a BsuTrz mutated on residues Ser78 and Gln43 shows that they are principal contributors to U74 binding. The Q43A and S78A mutations resulted in a 6- and 13-fold reduction of activity on the 74UAA76 tRNA substrate, respectively, but had little effect on a tRNA precursor with a CAA motif; the Q43A/S78A double mutant showed an almost complete lack of stimulation by U74 (100-fold reduced activity), confirming their role in U74 recognition (Pellegrini et al. 2012).

A potential uracil-binding pocket is present in the structures of human and EcoTrz, although they cannot fully accommodate the nucleotide in their current conformations due to steric clashes (Figure 16 D-F). The residues in the binding pocket are also conserved, Gln43 and Gly74 are conserved in both human and EcoTrz, while Ser78 is replaced by a functionally similar threonine in HsaTrz^S, and Gly77 is replaced by a cysteine in both organisms. TmaTrz is clearly very different: only Gly74 is conserved, and the U-pocket is not present.

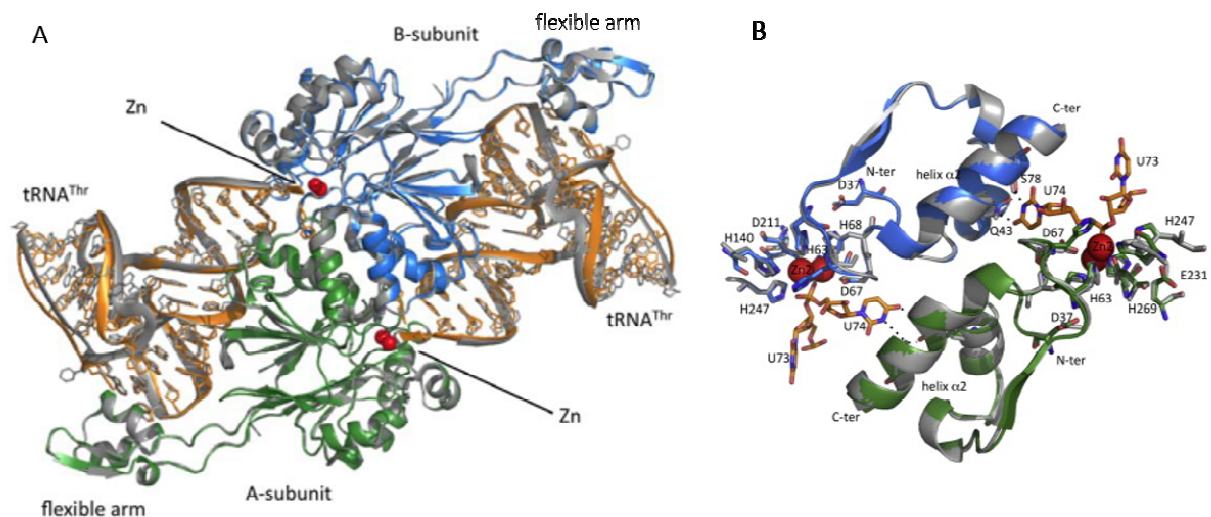


Figure 15 Superposition of BsuTrz bound to either pre- or product tRNAs.

(A) Full structure superposition. (B) The helix α 2. BsuTrz/tRNA precursor: Subunits A and B of RNase Z are shown in green and blue respectively; tRNA is shown in gold (PDB code: 4GCW). BsuTrz /tRNA product are shown in gray (PDB code: 2FK6). (Pellegrini et al. 2012).

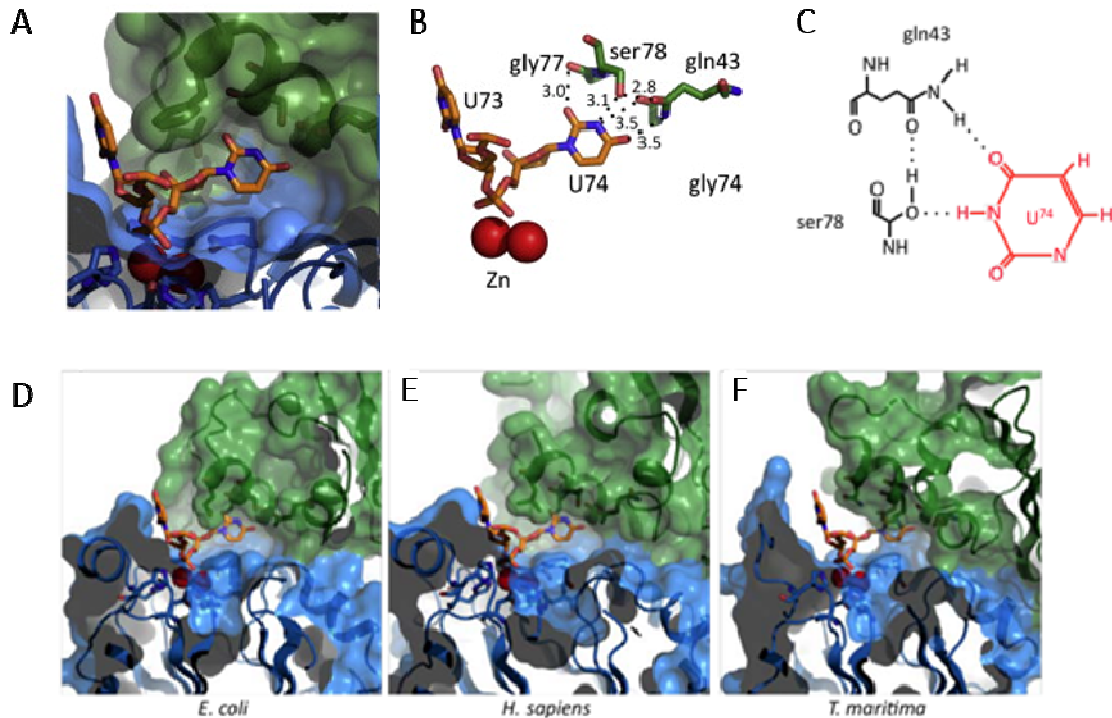


Figure 16 Uracil binding pocket of RNase Zs.

(A) Surface view of uracil binding pocket. The RNase Z subunit containing the active site is shown in blue, with Zn ions as red spheres. The subunit that binds most of the tRNA and provides the uracil-binding pocket is shown in green. (B) Amino acid residues interacting with U74. Potential interactions are shown as dotted lines. (C) Scheme representation of residues interacting with uracil. (D-F) Model of U74 bound to (D) *E. coli*, (E) *H. sapiens* and (F) *T. maritima* RNase Z.

I.2.7 The homology motifs affecting the tRNA catalysis/binding other than active site residues

Other than the conserved residues directly constitute the active site, several other conserved motifs were reported to affect the tRNA catalysis and/or binding. The conserved PxKxRN motif locates just ahead of the homology motif I, which is the amino side of the RNase Z^S or the beginning of the C-terminal half of RNase Z^L (Figure 7). Study on RNase Z^L from *D. melanogaster* (DmeTrz^LΔ23, mitochondrial signal peptide truncated) shows that the PxKxRN motif affects both the catalytic efficiency and substrate binding: alanine substitutions of residues from the PxKxRN loop and motif I from DmeTrz^L not only increase their k_M values 5 to 15-fold; but also unexpectedly diminish the k_{cat} values although the mutated residues do not directly contribute to the active center. For example, R448A of DmeTrz^LΔ23 from the PxKxRN loop leads to 1000-fold reduced catalytic activity; and a similar effect was observed for the G438A mutant. In contrast, alanine substitutions of active center residues only disrupt k_{cat} but have little effect on their k_M values (Zareen et al. 2005).

The PxKxRN motif of DmeTrz^L, in particular the two basic residues, also plays a role in tRNA cleavage site specificity (Zareen 2006).

The effect of PxKxRN loop and motif I could be explained by investigation on the structure of BsuTrz/tRNA complex (Figure 17) (Li de la Sierra-Gallay et al. 2006). The PxKxRN loop does not directly contribute to the active site, but is situated near the 3'-end of bound tRNA; on the other hand, the orientation of this loop and its interaction with His269 (motif V) clearly affects the orientation of the HEAT- α 7-HST motif; which in turn should influence substrate recognition.

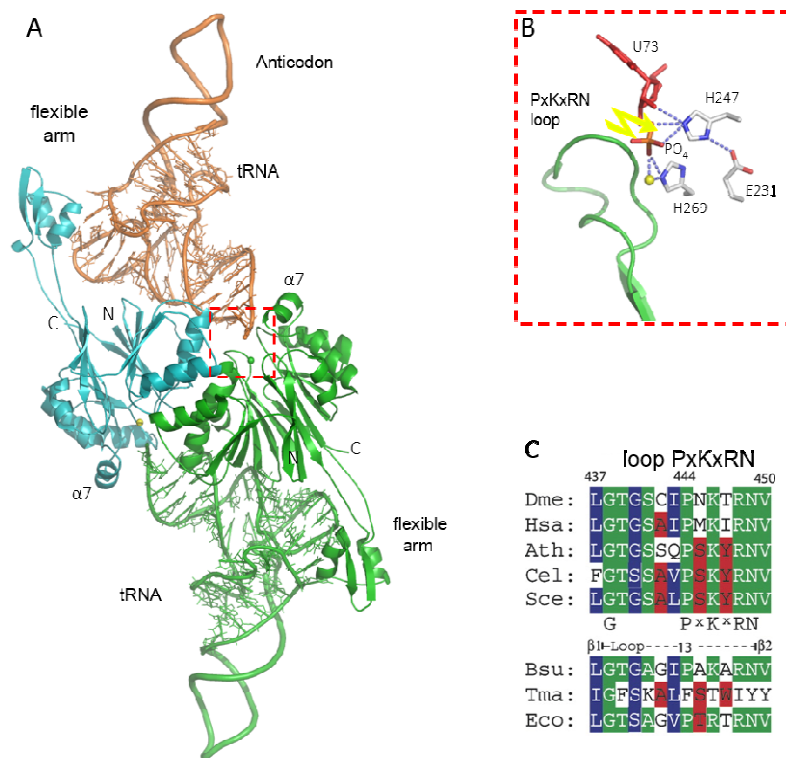


Figure 17 The conserved PxKxRN loop of RNase Z.

(A) Structure of BsuTrz in complex with tRNA (PDB code:2FK6) (Li de la Sierra-Gallay et al. 2006). Two subunits of BsuTrz are colored cyan and green, respectively. tRNA is in golden or green. The zinc is shown in sphere. (B) The PxKxRN loop and potential interacting residues. Phosphate ion is shown in orange and zinc in yellow sphere. The yellow flash indicates cleavage position. (C) Sequence conservation of the PxKxRN loop among RNase Z^S and C-terminal domain of RNase Z^L (Zareen, 2006).

Another more recently defined conserved motif Ax Dx near the C-terminus (residues 784 to 787) by sequence analysis (Figure 18) is proposed to affect cleavage site selection too, due to its proximity to the PxKxRN loop (Wang et al. 2012). The Ax Dx motif is conserved in the eukaryotic RNase Z^Ls and bacterial RNase Z^Ss except TM-type RNase Z^S found by sequence analysis, and forms a short loop that adopts the same conformation in the BsuTrz and HsaTrz^S structures; it has the potential to form hydrogen bonds with the PxKxRN loop and also to be

involved in cleavage. The AxDx motif also affects cell viability. Trz1 disrupted *S. pombe* strain has a temperature-sensitive growth defect which could be complemented by the wildtype gene, but an aspartate to alanine mutation in the AxDx motif impairs this ability (Zhao et al. 2009; Wang et al. 2012).

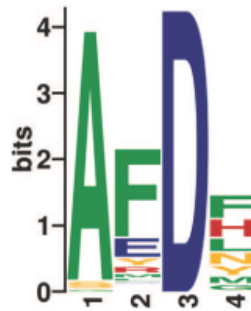


Figure 18 Sequence logos for the predicted AxDx motifs in RNase Z protein family (Wang et al. 2012).

The sequence logo of the AxDx motif is derived from 240 alignments including 178 RNase Z^Ls from 71 metazoans, 21 green plants, 77 fungi, and 62 RNase Z^Ss from 30 metazoans, 9 fungi, 2 bacteria and *M. brevicollis*. The sequence logos were created using WebLogo (<http://weblogo.berkeley.edu>). The height of each amino acid indicates the level of conservation at that position. Amino acids are colored as follows: red, basic; blue, acidic; orange, polar; and green, hydrophobic.

I.2.8 RNase Z^S: a cooperative enzyme

Unlike the RNase Z^L that displays typical *Michaelis–Menten* kinetics (Mohan et al., 1999), EcoTrz has sigmoidal saturation kinetics towards a chromogenic phosphodiester substrate (bpNPP) (Vogel et al., 2002), BsuTrz has similar sigmoidal kinetics with a pre-tRNA substrate (Redko et al. 2007). Therefore, RNase Z^S can be considered to be a cooperative enzyme for which binding of the first pre-tRNA increase either the affinity or the cleavage efficiency for the second substrate. Two possible explanations for the sigmoidal saturation kinetics were proposed based on the superposition of structures of the free and complex form of BsuTrz. One explanation relates to the relative positions of the two subunits that changes upon tRNA binding (Li de la Sierra-Gallay et al. 2006). The slight rotation (1.75°) of one subunit relative to the other subunit induced by tRNA binding could be sufficient to increase the binding or cleavage efficiency of a second substrate and give rise to a classical cooperative behavior. Another explanation is that RNase Z^S is a “mnemonic” enzyme, in which the enzyme is in an inactive ground state and undergoes a slow conformational change to the active state after binding its substrate (Redko et al. 2007).

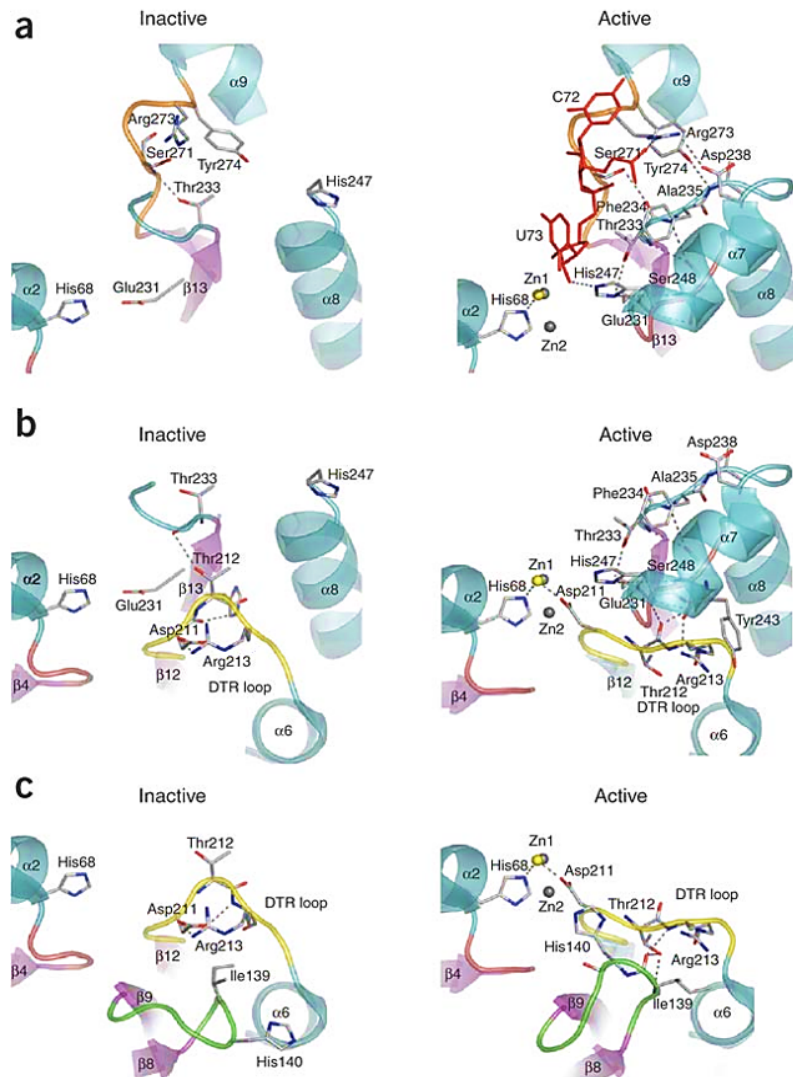


Figure 19 Major conformational changes of RNase Z induced by complex formation with tRNA (Li de la Sierra-Gallay et al. 2006).

(a) Key interactions between C72 and U73 from the tRNA bound to the B subunit (red) and the loop containing residues Ser271–Tyr274 (orange) and the $\alpha 7$ region (light blue) of the A subunit. Hydrogen bonds and metal coordination are shown with gray dotted lines. Positions of Zn^{2+} ions in the complex (yellow sphere) and in the 'active' A subunit of the free protein (gray spheres) are indicated. (b) Key interactions between the $\alpha 7$ helix region (light blue) and the DTR loop (residues Asp211–Arg213, yellow). (c) Key interactions between the DTR loop (yellow) and the $\beta 8$ – $\beta 9$ loop (green).

In the case of BsuTrz, the activation of the inactive B subunit is supposed to occur in five steps (Figure 19): (1) the sensing of tRNA presence by Arg273 (motif V), (2) the conveying of the message through hydrogen bonds between Arg273 and conserved residues, (3) the stabilization of the $\alpha 7$ helix, (4) a conformational change of the DTR loop (motif IV), and (5) in turn a restructuring of the neighboring loop which is involved in catalysis. In the absence of tRNA, the $\alpha 7$ helix is highly flexible. When tRNA is present, Arg273 acts as a bridge between the $\alpha 7$ and the tRNA substrate, thus sensing the presence of tRNA and initiating the activation, by a succession of hydrogen bond formation between the conserved residues. The stabilized

$\alpha 7$ helix then provokes a conformational change of its neighboring DTR loop, which subsequently causes a conformational change of its neighboring loop connecting strands $\beta 8$ and $\beta 9$, causing the conformational change of key residues His247, His140, and Asp211 from the catalytic center (Li de la Sierra-Gallay et al. 2006).

This cascade of events explains why the conformational change of the “mnemonic” enzyme after binding the first substrate is a slow process. The return to the ground state after product release, is also a relatively slow step. In this way, the enzyme can rapidly enter a new catalytic cycle, bypassing the activation step upon encounter of a new substrate molecule. At low substrate concentrations, mnemonic enzymes return to the inactive state at each cycle and have slower enzyme velocities, accounting for a sigmoidal saturation curve. Evidence that the affinity of metallo- β -lactamases for Zn is increased by up to three orders of magnitude in the presence of substrate is consistent with this model for RNase Z activation (Redko et al. 2007).

I.2.9 Prediction of the structure of an RNase Z^L in complex with tRNA

In the long form of RNase Z, a functional metal-binding site is present only in the C-terminal half, the N-terminal domain was therefore predicted to be inactive. This was supported experimentally by the fact that an ELAC2 construct lacking the C-terminal domain (ELAC2- Δ C) was not able to cleave the precursor tRNA while ELAC2- Δ N kept some tRNA processing activity (Takaku et al. 2003; Takaku 2004). On the other hand, only the N-terminal domain possesses the exosite. Therefore, although the RNase Z^S dimer could in principle catalyze two tRNA molecules simultaneously, RNase Z^L is capable of forming a complex with only a single tRNA substrate molecule at the time (Redko et al. 2007).

I.2.10 RNase Z properties with respect to the organisms

I.2.10.1 RNase BN/EcoTrz is both an endo- and exoribonuclease

In almost all organisms RNase Z acts solely as an endoribonuclease. However, EcoTrz and BsuTrz also display 3'-5' exoribonuclease activity against certain substrates.

RNase BN was originally discovered as an exonuclease essential for maturation of bacteriophage T4 tRNA precursors lacking a CCA-triplet; and was able to remove the 3'-terminal residues of synthetic mimics of these T4 tRNA precursors as mononucleotide (Asha

et al. 1983; Seidman et al. 1975). Therefore RNase BN was considered to be an exoribonuclease (Asha et al. 1983; Deutscher 1993). However, later on RNase BN was found to be actually the RNase Z homologue from *E. coli* (Ezraty et al. 2005), which was encoded by the *elaC* gene and initially discovered as zinc phosphodiesterase meaning it is also an endoribonuclease (Vogel et al. 2002). More recent work confirmed both the exo- and endoribonuclease activities of EcoTrz on certain substrates (Dutta & Deutscher 2009; Dutta & Deutscher 2010; Dutta et al. 2012): RNase BN displays exoribonuclease activity on model RNAs and tRNA precursor substrates in vitro and it acts as a dual exo- and endoribonuclease on tRNA precursors in vivo. The action on *E. coli* CCA-containing tRNAs precursors involves either a single endonucleolytic cleavage after the CCA sequence or exonucleolytic trimming to remove extra 3'- residues up to the CCA sequence, depending on the metal ion cofactor presence (Mg²⁺ or Co²⁺).

The exonuclease activity of RNase BN is strongly inhibited by the presence of CCA-triplet and 3'-phosphoryl group; while the endonuclease activity is not affected by the latter (Dutta & Deutscher 2009). The reason why CCA protects the tRNA from trimming by EcoTrz was investigated (Dutta et al. 2013). The adjacent C bases of the CCA sequence and an arginine residue within a highly conserved sequence motif in the channel leading to the RNase Z catalytic site are both required for the protective effect. When both of these determinants are present, CCA-containing RNAs in the channel are unable to move into the catalytic site; however, substitution of either of the C residues by A or U or mutation of Arg274 to Ala allows RNA movement and catalysis to proceed. These data define the mechanism for how tRNAs are protected against the promiscuous action of a processing enzyme (Dutta et al. 2013).

All *E. coli* tRNAs genes contain CCA sequence, 3'-tRNA processing in *E. coli* involves a handful of exonuclease which might be redundant. Surprisingly *E. coli* still contains the RNase Z homologue gene, although a mutant lacking this protein did not have an obvious growth phenotype (Schilling et al. 2004). Significantly different with other RNase Z proteins, RNase BN also plays a role in mRNA decay (Pervez & Kushner 2006). Absence of RNase BN can slightly lengthen mRNA half-life; and this effect is much more pronounced when RNase E is also missing. Therefore the physiological function of RNase BN is still not fully clarified.

I.2.10.2 The CCA isn't always an anti-determinant

The fact that CCA is an anti-determinant for RNase Z from *Drosophila* and pig had been described long ago (Mohan et al. 1999). Studies with DmeTrz^L showed that the presence of a CCA sequence in mature tRNA reduced *k_{cat}* ~80-fold with little effect on *K_M* compared to a natural tRNA precursor (whose 3'-trailer sequence after the nucleotide is not CCA) without affecting the *K_M* (Zareen, 2006). The CCA anti-determinant effect for EcoTrz exonuclease activity had been described in the last paragraph. Early studies on BsuTrz showed its activity was inhibited by the presence of CCA, too, where most of the inhibiting effect was due to the first C-residue (Pellegrini et al. 2003).

However, a more recent study shows BsuTrz activity is actually not inhibited by CCA but rather activated by a uracil at position 74 or 75 (Pellegrini et al. 2012). An uracil residue in position 74 stimulates activity of BsuTrz about 200-fold compared to any other residue C, G or A in this position. Therefore, BsuTrz is not negatively regulated by the CCA but rather positively regulated by the uracil, the activity of BsuTrz is just very low when position 74 is not a uracil. Most *B. subtilis* tRNA precursors naturally have a U in position 74 explaining why the system works *in vivo*. It was prospected that the human enzyme possibly functions in the same way because it also has a potential U-pocket (Pellegrini et al. 2012). EcoTrz was considered containing a similar potential U-pocket (Pellegrini et al. 2012), this seems consistent with the fact that substitution of either of the C residues of CCA by A or U would disrupt the protective mode and RNA were allowed to move in for catalysis to proceed, although in this case EcoTrz act as an exonuclease rather than the endonuclease BsuTrz (Dutta et al. 2013). Interestingly an uracil at position 75 or 76 will trigger the exonuclease activity of BsuTrz but can also stimulate BsuTrz although less than in position 74. With U in position 75 or 76, RNase Z cleaves as an endonuclease immediately before the U residue, and then surprisingly switches to exonuclease mode to remove the extra nucleotides one by one until it reaches the discriminator base (Pellegrini et al. 2012).

Clearly, further studies are required to unravel the effect of the CCA sequence with respect to RNase Z activities against certain substrates.

I.2.10.3 TmaTrz: a special case of RNase Z cleaves CCA-containing pre-tRNA

RNase Z specifically cleaves natural CCA-less tRNA precursors right after the discriminator nucleotide. However, *T. maritima* RNase Z presents a unique exception since it also recognizes CCA-containing tRNAs as substrate. In the bacterium *Thermotoga maritima*,

45 out of the 46 tRNA genes encode the CCA triplet (Minagawa et al. 2004). TmaTrz cleaves the 3'-trailer up to the CCA sequence (Figure 4A ii). For the single CCA-less tRNA species (tRNA^{Met}, UAG), TmaTrz cleaves after the discriminator base as other RNase Z proteins. Apparently, TmaTrz distinguishes between CCA-containing and CCA-less precursors (Minagawa et al. 2004).

Various elements of TmaTrz affect cleavage site selection: histidine and aspartate residues at the catalytic center; the dimer interface residues which affect relative orientation of the subunits; and the flexible arm (Minagawa et al. 2008). S31Q or T33Q mutations at the dimer interface shift the cleavage site to 1–2nt upstream; but double mutants restored the usual cleavage site explaining by cancel out the orientation change effect of single substitution.

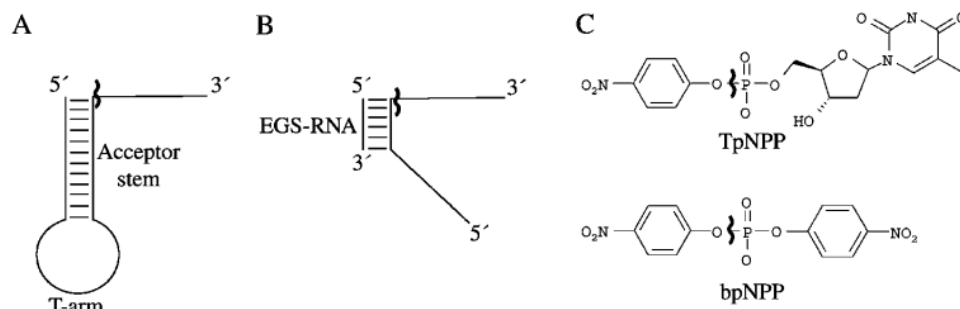


Figure 20 Other Substrate of RNase Z (Vogel et al. 2005). The RNase Z cleavage site is indicated by wavy lines.

I.2.11 Substrates of RNase Z

I.2.11.1 tRNA-like substrates

The minimal RNase Z substrate was shown to be a micro-pre-tRNA, which consists of the T-arm, the acceptor stem and a 3'-trailer (Figure 20A). It could be cleaved by the human and yeast RNase Z^L enzymes, but with low efficiency (Takaku 2004). The stem length is not arbitrary, it has to contain at least 11 base pairs to be recognized and cleaved. Reductions of the T stem may result in cleavage site displacement. Micro-helices containing only 7 base pairs are not substrates for mammalian RNase Z (Ceballos & Vioque 2007). The N-terminal domain of the long form of RNase Z is essential for the recognition of this minimum substrate; the RNase Z^S from human, *E. coli* and *T. maritima* do not process this substrate. This suggests that the N-terminal domain might be responsible for its broader spectrum of RNase Z^L substrate (Takaku 2004). In contrast, both AthTrz^{S1} and MjaTrz^S have little tolerance for

variations of the acceptor stem length. Wild-type tRNAs with seven base pairs in the acceptor stem are processed efficiently, but deletions or additions of more than one base pair in the acceptor stem drastically reduce cleavage efficiency; and the cleavage site in the acceptor stem variants is shifted (Vogel et al. 2005)(Schiffer et al., 2003).

I.2.11.2 Synthetic phosphodiesterase substrates

Another type of substrate for RNase Z are two synthetic phosphodiesterase substrates: thymidine-5-p-nitrophenyl phosphate (TpNPP) and bis (p-nitrophenyl) phosphate (bpNPP) (Figure 20C). These small chromogenic molecules are efficiently cleaved by EcoTrz but less efficiently by other RNase Z enzymes tested (AthTrz^{S1}, AthTrz^{S2}, DmeTrz^L, SceTrz^L, HvoTrz^S, PfuTrz^S) (Vogel et al. 2002). These chromogenic compounds are non-physiological substrates, but they are considered to be good model substrates for EcoTrz phosphodiesterase activity. EcoTrz with two zinc per monomer has two-fold cooperativity with the substrate bpNPP, indicating the allosteric mechanism is unrelated to sequential zinc incorporation (Vogel et al. 2002; Schilling et al. 2005). It is worth noting that none of the RNase Z^L have bpNPP cleavage activity (Vogel et al. 2002).

Like the tRNA 5'-processing enzyme RNase P, human RNase Z^L has the capacity to cleave hybrid RNAs consisting of a target RNA and an EGS (external guide sequence), which form a stem of at least 6 base pairs (Shibata et al. 2005) (Figure 20B). However, the short RNase Z proteins from human (ELAC1), *E. coli* and *T. maritima* could not cleave this substrate (Shibata et al. 2005). These findings suggest that the Trz^L does not require a complete tRNA structure for substrate recognition, indicating they have a broader substrate spectrum than the short RNase Z.

In fact, RNase Z^L can cleave any target RNA at any desired site under the direction of a small guide RNA (sgRNA) by recognizing a pre-tRNA-like or micro-pre-tRNA-like complex formed between the target RNA and the sgRNA (Ninomiya et al. 2015; Takaku 2004). Based on this property of RNase Z^L, Masayuki Nashimoto's group has developed an sgRNA-mediated gene silencing technology termed RNase Z^L-utilizing efficacious gene silencing (TRUE gene silencing), to investigate the gene regulatory network via RNase Z^L and sgRNA (Ninomiya et al. 2015).

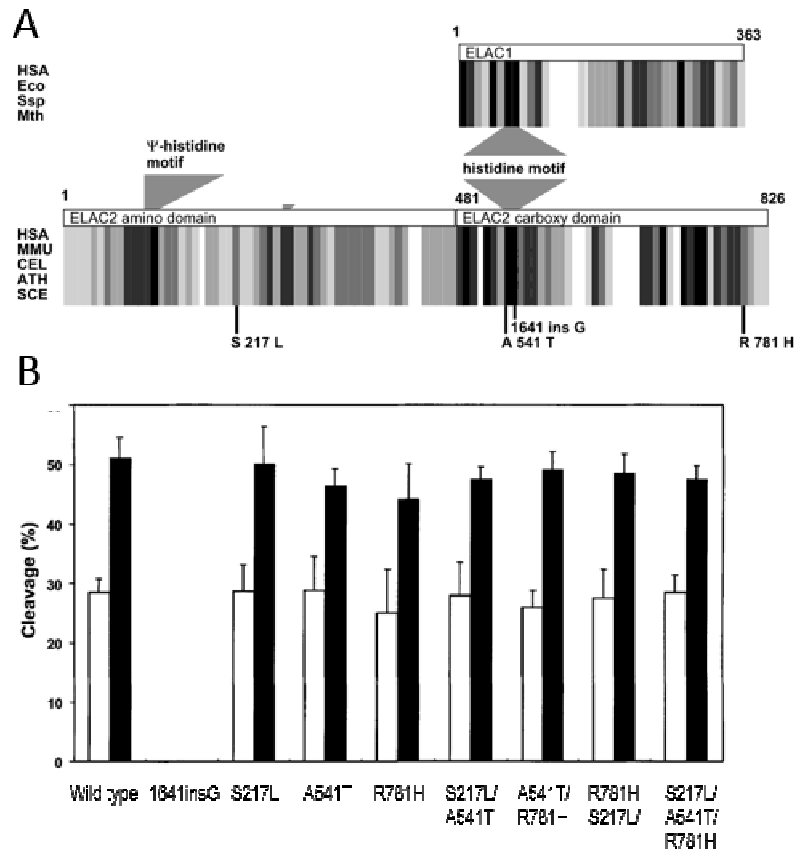


Figure 21 ELAC2 mutations associated with the occurrence of prostate cancer.

(A) Positions of the mutations in human ELAC2 and corresponding positions in RNase Z family proteins through sequence alignment. ELAC1 orthologs were selected from human (HSA), *E. coli* (Eco), the blue-green algae *Synechocystis sp.* (Ssp), and the archaeobacterium *Methanobacterium thermoautotrophicum* (Mth). ELAC2 orthologs were selected from human (HSA), mouse (MMU), *C. elegans* (CEL), *A. thaliana* (ATH) and *S. cerevisiae* (SCE). Positions of the pseudo-histidine motif, and the histidine motif are indicated by gray-shaded pointers (Tavtigian et al. 2001). (B) 3'-tRNA processing activities of enzyme variants associated with the occurrence of prostate cancer. Values of percent cleavage for low and high enzyme dosages were represented by open and filled bars, respectively (Takaku et al. 2003).

I.2.12 RNase Z^L and diseases

I.2.12.1 ELAC2 is a prostate cancer susceptible gene

The human RNase Z^L was initially identified as a prostate cancer gene (Tavtigian et al. 2001)(Takaku et al. 2003). When the human ELAC2 was purified and micro-sequenced, it was shown to correspond to the previously identified prostate cancer susceptibility gene. Eight ELAC2 variants harboring mutated residues associated with the occurrence of prostate cancer were tested for tRNA 3'-processing activity *in vitro* (Figure 21). Seven ELAC2 variants that contain one to three amino acid substitutions showed efficient tRNA 3'-processing activities, while one truncated variant, which lacked the C-terminal half region (early termination caused by a G (guanine) base insertion at 1641), had no activity. These data

suggest that prostate cancer might not be directly related to pre-tRNA processing activity (Korver et al. 2003; Minagawa et al. 2005).

Overexpression of RNase Z^L in tumor cells causes a delay in G2-M progression; and consistent with this function in mitosis, ELAC2 was found to physically interact with the γ -tubulin complex, which was speculated to interfere with cell division and lead to the tumorigenesis (Korver et al. 2003). In relation to this, the *Caenorhabditis elegans* homolog of human RNase Z^L plays a role in germline proliferation (Smith & Levitan 2004). These causal connections between mutations in RNase Z^L and prostate cancer suggest it could be a complex process (Ceballos & Vioque 2007).

I.2.12.2 RNase Z in mitochondria tRNA 3'-processing

1) Dual localization of RNase Z^L

Almost all RNases Z^L studied to date are predicted to have a mitochondrial or chloroplast targeting sequence at their N-terminus (Vogel et al. 2005). The only exception is the *S. cerevisiae* enzyme: Trz1 does not have a mitochondrial leading sequencing, but its dual location in both mitochondria and nucleus has been shown experimentally by *in vivo* fluorescence (Hazbun et al. 2003) and subcellular fractionation (Skowronek et al. 2014). Since RNases Z^S lack this signal, they are supposed to be nuclear or cytosolic enzymes. In those organisms where there is only a long version of the enzyme, as in *S. cerevisiae*, *D. melanogaster*, and *C. elegans*, it is supposed to be responsible for the processing of tRNAs in the nucleus, the mitochondria and the chloroplasts. For example, *Drosophila melanogaster* has only one RNase Z^L but alternative translation initiation yields two protein forms: the long one is mitochondrial, and the shorter one starting at the second AUG codon may localize in the nucleus or cytosol. It not only possesses both nuclear localization signals and mitochondrial targeting sequences, but can also account for the processing of nuclear and mitochondrial tRNAs *in vivo* (Xie & Dubrovsky 2015).

Human cells contain both a long and a short form of RNase Z (ELAC2 and ELAC1, respectively). ELAC2 localizes both in mitochondrion and nucleus in several human cell lines (HeLa, HT1299, 143B and 293). ELAC2 has a second methionine at amino acid position 16 which is considered to be an alternative translation start: the mutant protein lacking the first 15 amino acids localized exclusively to the nucleus, suggesting the N-terminal 15 amino acids between the first and the second ATG are indispensable for mitochondrial import. Instead,

flag-tagged ELAC1 was detected mainly in the cytoplasm and small amounts localized also in the nucleus, but did not colocalize with the mitochondrial staining (Brzezniak et al. 2011). Another report gave similar results in human cell line 293 (Rossmanith 2011): subcellular localization of human RNases Z^S fused to green fluorescent protein (EGFP) showed it is restricted only in the cytosol, whereas RNases Z^L localized to both nucleus and mitochondria; and the alternative translation initiation is responsible for the dual targeting of RNases Z^L . In Arabidopsis, four RNase Z proteins are encoded in the nucleus (AthTrz^{S1}, AthTrz^{S2}, AthTrz^{L1}, AthTrz^{L2}). AthTrz^{L1} is shared between nucleus and mitochondria, while another one (AthTrz^{L2}) is restricted to mitochondria (Canino et al. 2009; Salinas-Giegé et al. 2015).

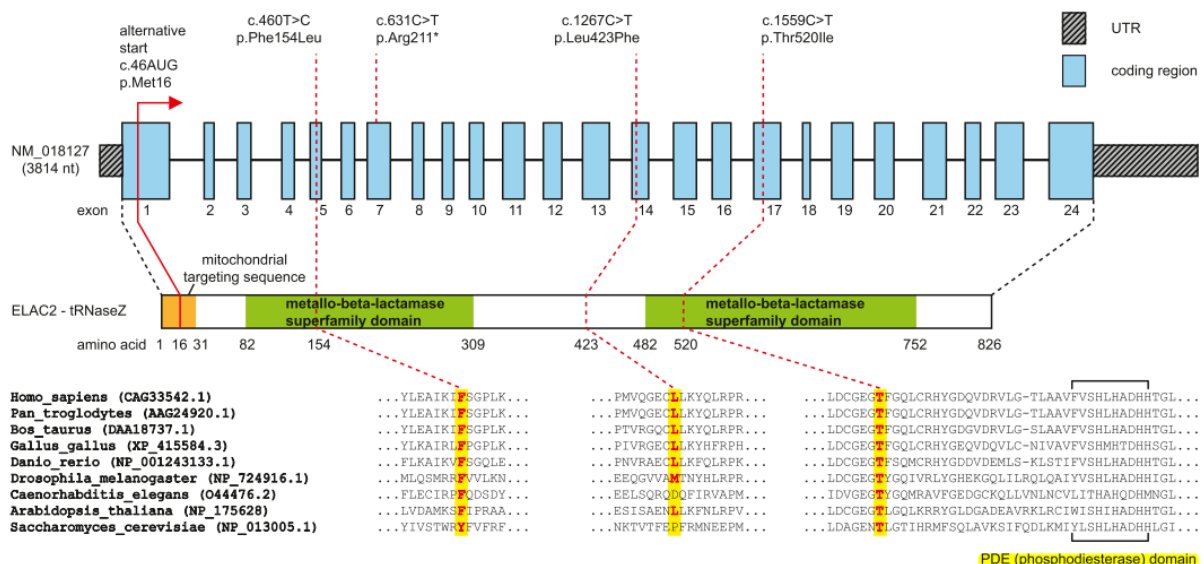


Figure 22 Hypertrophic cardiomyopathy-related mutations in ELAC2 (Haack et al. 2013).

Gene structure of ELAC2 with known protein domains of the gene product and localization and conservation of amino acid residues affected by mutations are shown. Intronic regions are not drawn to scale.

2) RNase Z^L associated with mitochondrial diseases

About 120 mutations have been described in mitochondrial tRNA genes that are associated with human diseases (Ceballos & Vioque 2007). These tRNA mutations cover a wide range of pathologies: myopathies, encephalopathies, cardiopathies, deafness, etc. The human mitochondrial genome encodes RNA components of its own translational machinery to produce the 13 mitochondrial-encoded subunits of the respiratory chain. Nuclear-encoded gene products are essential for all processes within the organelle, including RNA processing. Transcription of the mitochondrial genome generates large polycistronic transcripts punctuated by the 22 mitochondrial (mt) tRNAs that are conventionally cleaved by the RNase

P-complex and ELAC2 at the 5'-and 3'-ends, respectively (Haack et al. 2013)(Brzezniak et al. 2011). Not only the expression of tRNA would be affected by defect processing. In a more recent case ELAC2 mutations that cause a mitochondrial RNA processing defect were found to be associated with hypertrophic cardiomyopathy (Figure 22). These mutations actually impaired mitochondrial translation (Haack et al. 2013).

Using complementation experiments in yeast a model of a potential ELAC2 patho-mechanism was proposed (Figure 23) (Haack et al. 2013). Under normal conditions minor amounts of unprocessed ELAC2 substrates can be degraded by the RNA surveillance machinery. Reduced ELAC2 activity results in an accumulation of mitochondrial precursor mRNAs, which impair mitochondrial translation. The RNAi knockdown cells of *DmeTrz^L* display impaired mitochondrial polycistronic transcript processing, leading to impaired growth of mitotic and endoreplicating tissues. The precise mechanism of this phenomenon remains unclear (Xie & Dubrovsky 2015).

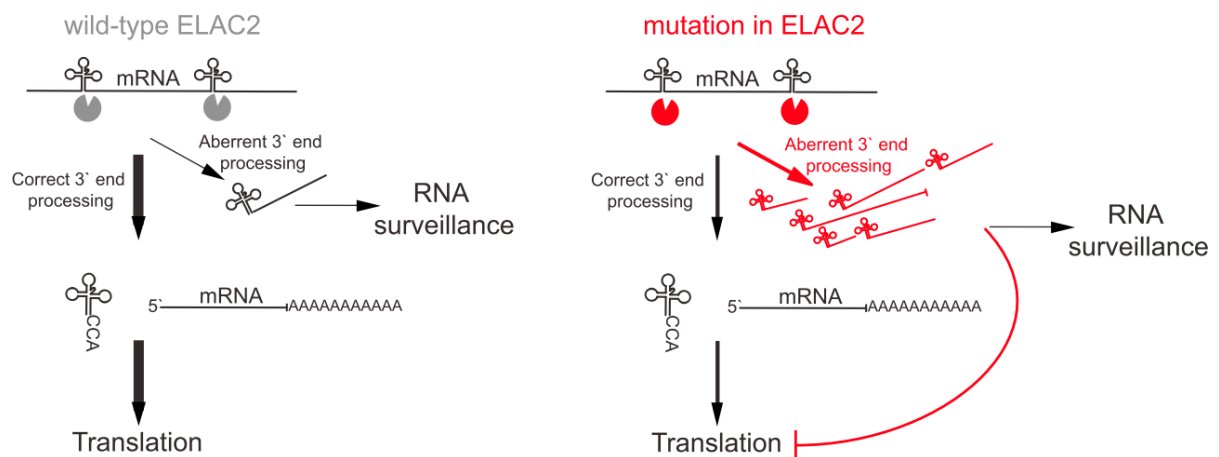


Figure 23 Model of potential ELAC2 patho-mechanism (Haack et al. 2013).

II. Trz1/Nuc1/mutarotase complex

II.1 Identification of a Trz1/Nuc1/mutarotase protein complex.

In a systematic proteomic search for protein complexes in yeast, TAP-tagged Trz1 was found to co-purify with two other proteins from the cell lysate. These two proteins were identified by mass spectrometry to be Nuc1 and mutarotase (ymr099c), respectively (Figure 24) (Hazbun et al. 2003). Interaction between either two of these three proteins were validated by in vivo co-precipitation in yeast: pairs of proteins were GST- and myc-tagged, followed by GST pull-down from a crude lysate and co-precipitate by western blot, confirming the association between Trz1/Nuc1, Nuc1/mutarotase, and Trz1/mutarotase (Figure 25) (Benschop et al. 2010).

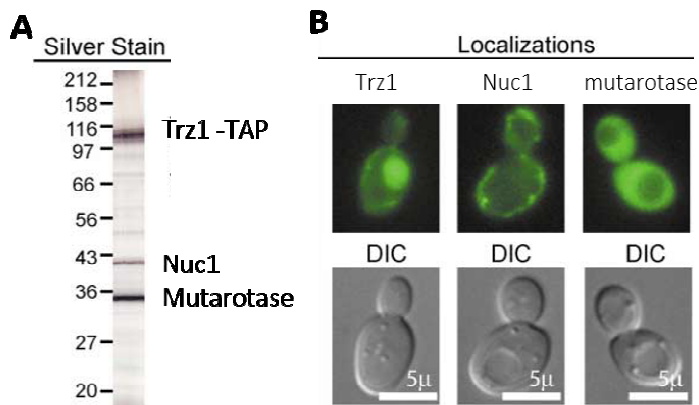


Figure 24

Figure 24 Trz1 copurifies and colocalize with Nuc1 and mutarotase (ymr099c) (adapted from Hazbun et al., 2003).

(A) purification of TAP-tagged Trz1 and subsequent identification of copurifying proteins from the cell lysate. The silver-stained gel of the eluate displays three dominant bands and mass spectrometry identifies YKR079c (Trz1), Nuc1, and YMR099c. (B) localization of YKR079c and Nuc1 fused to Venus and YMR099c fused to YFP. YKR079c localizes in the mitochondria and nucleus. Nuc1 localizes in the mitochondria, and YMR099c localizes in the nucleus and cytoplasm.

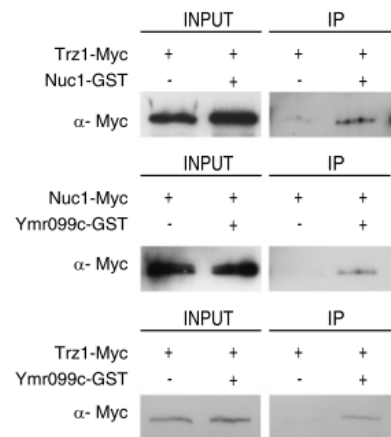


Figure 25

Figure 25 Coprecipitation of Trz1, Nuc1, and mutarotase (ymr099c) in pairs (Benschop et al., 2010).

Lysates from yeast strains (co)expressing a Myc-tagged protein, with or without a GST-tagged protein, were incubated with glutathione beads. After washing, interacting proteins were analyzed by SDS-PAGE and western blotting using anti-Myc antibodies.

II.2 Localization of the subunits of the complex

The biological significance of this complex is not understood. In budding yeast cells, fluorescence shows Trz1 localize in both nucleus and mitochondria (Figure 24B). Subcellular fractions from yeast extract further shows Trz1 localizes mainly in both nucleus and

mitochondria, and little in the cytosol, particularly in the matrix of mitochondria (Figure 26) (Skowronek et al. 2014).

Similar to human EndoG, Nuc1 was found to be predominantly localized in mitochondria, as demonstrated by immunofluorescence (Figure 27A) and in vivo fluorescence of a Nuc1-yEGFP fusion protein (Figure 27B). Upon H₂O₂ stimulus Nuc1 is concentrated in the nucleus. Further subcellular fractionation demonstrated that Nuc1 localizes at the inner membrane of mitochondria (Figure 27C-D) (Büttner, Eisenberg, et al. 2007). Mutarotase is found to localize all through the yeast cell except in the vacuole (Hazbun et al. 2003).

Although the fluorescence shows Trz1 and Nuc1 co-localization in the nucleus and mitochondria Nuc1 and mutarotase may co-localize in the nucleus or mitochondria, too (Figure 24; Figure 28), the localization of the complex remains unclear.

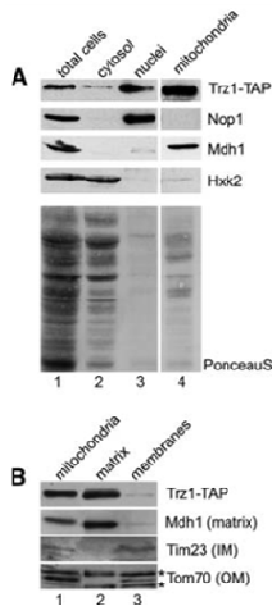


Figure 26

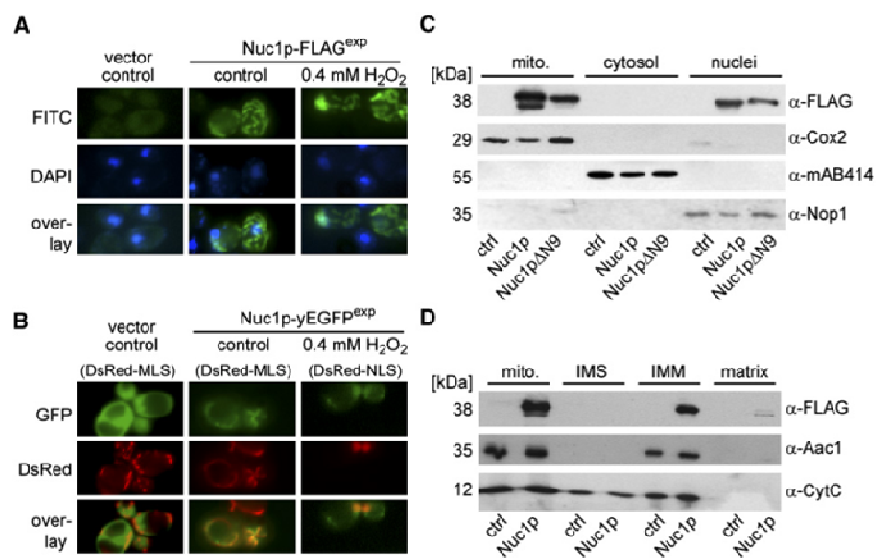


Figure 27

Figure 26 TAP-tagged Trz1 is localized in both nucleus and mitochondria (Skowronek et al. 2014).

(A) Subcellular fractions of yeast extracts (total cells, cytosol, mitochondria, and nuclei) were analyzed by Western blotting using peroxidase-anti-peroxidase antibodies (PAP). The same blot was probed against nuclear (Nop1), mitochondrial (Mdh1), and cytoplasmic (Hxk2) proteins using specific antibodies. Ponceau S staining is shown in the bottom panel as a loading control. (B) Localization of Trz1-TAP in yeast mitochondrial matrix. Alkali-extracted soluble and membranes fractions as well as total mitochondrial extract were analyzed by Western blotting using PAP antibodies for Trz1 and specific antibodies for markers of different mitochondrial compartments: Mdh1 (matrix, M), Tim23 (inner membrane, IM) and Tom70 (outer membrane, OM). Asterisks mark bands resulting from antibodies' cross-reactivity.

Figure 27 Nuc1 is localized in mitochondria and translocated to nucleus upon apoptosis induction.

(A) Immunofluorescence staining of cells expressing Nuc1p-FLAG after treatment with or without 0.4 mM H₂O for 20 hr on galactose. Cells were stained with anti-FLAG as primary and FITC-conjugated anti-mouse as secondary antibody. To visualize nuclei, cells were counterstained with DAPI. (B) Fluorescence microscopy of cells expressing Nuc1p-yEGFP after treatment with or without 0.4 mM H₂O₂ for 20 hr. Mitochondria and nuclei were visualized with mitochondrial marker DsRed Su1-69 or nuclear marker DsRed-NLS, respectively. (C) Immunoblot of mitochondrial, cytosolic, and nuclear fractions of cells expressing FLAG-tagged Nuc1p (lanes 2,

5, and 8), and vector control (lanes 1, 4, and 7). Blot was probed with antibodies against FLAG epitope, or Cox2p, or Nop1p, or with mAB414, a monoclonal antibody immunoreacting with a 55 kDa cytosolic protein. (D) Immunoblot of submitochondrial fractions of cells expressing FLAG-tagged Nuc1p (lanes 2, 4, 6, and 8) and vector control (lanes 1, 3, 5, and 7). Blot was probed with antibodies against FLAG epitope or cytochrome c as control for the inner-membrane space (IMS) or Aac1p as control for the inner-mitochondrial membrane (IMM).

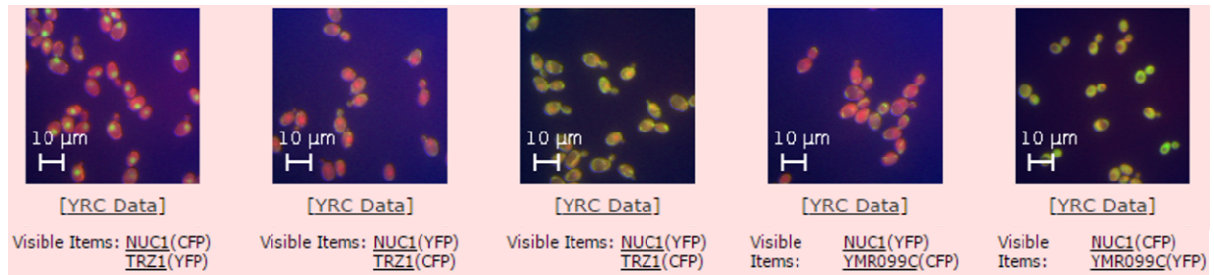


Figure 28 Co-localization of Trz1, Nuc1 and mutarotase (ymr099c). (Yeast Resource Center <http://www.yeastrc.org>).

III. Nuc1

III.1 *Nuc1 is a major mitochondrial nuclease*

The yeast EndoG (Nuc1) is an enzyme that accounts for more than 90% of total nuclease activity in mitochondria (Dake et al., 1988). It is located at the inner membrane of mitochondria (Büttner, Eisenberg, et al. 2007). The purified Nuc1 shows multiple activities including RNase activity on single-stranded, but not on double-stranded RNA; endonuclease activity on single- and double-stranded DNA; and a 5'- exonuclease activity on double-stranded DNA. Digestion of DNA by Nuc1 generates 5'-phosphorylated termini (Dake et al. 1988).

Nuc1, encoded by ORF YJL208c, displays significant sequence similarity to mammalian endonuclease G (EndoG) (42% identity and 62% similarity to human EndoG, Figure 29). Nuc1 is evolutionarily more closely related to human EndoG than to *C. elegans* EndoG (Figure 29B). Like mammalian EndoG, Nuc1 contains a potential mitochondrial localization sequence (MLS) (Büttner, Eisenberg, et al. 2007).

The catalytic properties of EndoG from several organisms were studied. Generally, EndoG enzymes are able to use various divalent metals as Mg, Mn, etc, depending on the origin of the protein; and its activity is inhibited by high salt concentrations (NaCl or KCl) (Kieper et al. 2010; Lin et al. 2012; Schäfer et al. 2004) (Loll et al. 2009). For example, upon the substrate double strand plasmid DNA the bovine EndoG prefers Mn^{2+} and Co^{2+} over Mg^{2+} , with an optimal Mg^{2+} concentration between 2.5–5.0 mM. The enzyme is most active at pH

6.0, at higher pH values it has the tendency to nick the DNA substrate rather than introducing double-strand breaks as judged by the accumulation of the open circular form of plasmid DNA (Figure 30) (Schäfer et al. 2004). For the yeast Nuc1, the optimal Mg^{2+} concentration was 2.5mM and the optimal pH is 7.5 for single strand DNA substrates (Kieper et al. 2010). Best reaction conditions for the CPS-6 (EndoG homologue from *Caenorhabditis elegans*) contained 2mM Mg^{2+} and a low salt buffer at pH 7. Similar to the bovine EndoG, CPS-6 binds preferentially to G-tract DNA in the optimum low salt buffer at pH 7 (Lin et al. 2012; Côté et al. 1989).

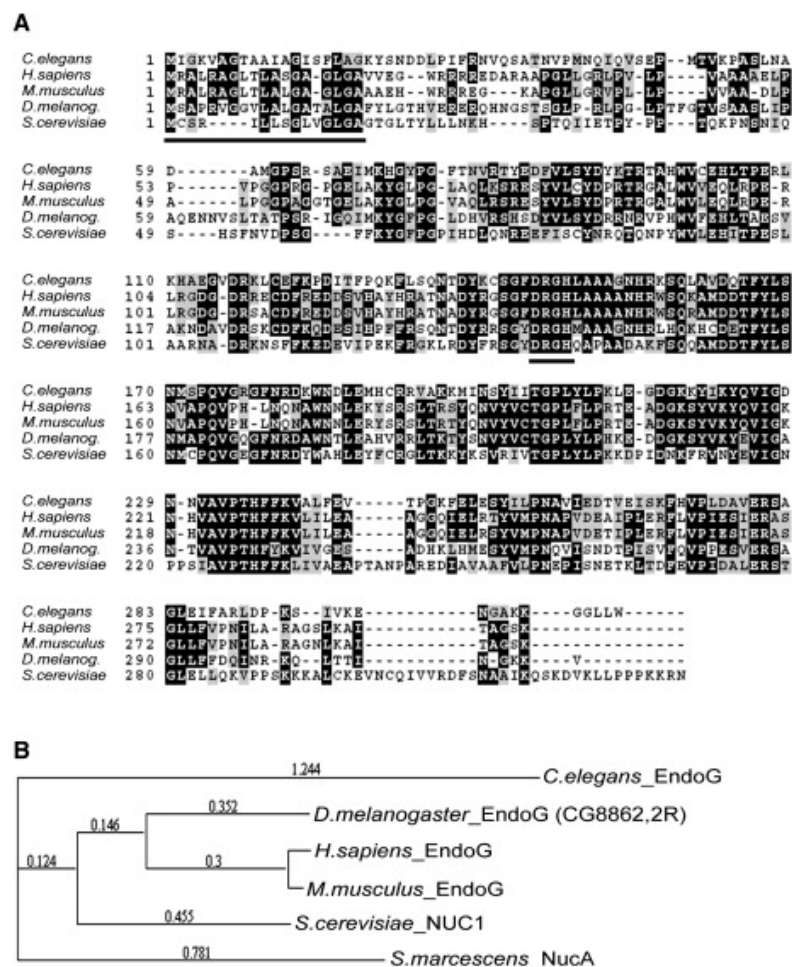


Figure 29 *S. cerevisiae* Nuc1 is a homolog of mammalian EndoG (Büttner, Eisenberg, et al. 2007). (A) Alignment of EndoG sequences from *C. elegans*, *H. sapiens*, *M. musculus*, and *D. melanogaster* with *S. cerevisiae* Nuc1. Black and gray boxes indicate amino acid identity and similarity, respectively. Putative N-terminal mitochondrial localization sequences and active site residues are marked with black bars. The active site “DRGH”-motif is highlight. (B) Phylogenetic tree of EndoG homologs aligned in (A) plus the bacterial nuclease NucA from *S. marcescens*. The numbers indicate evolutionary distance.

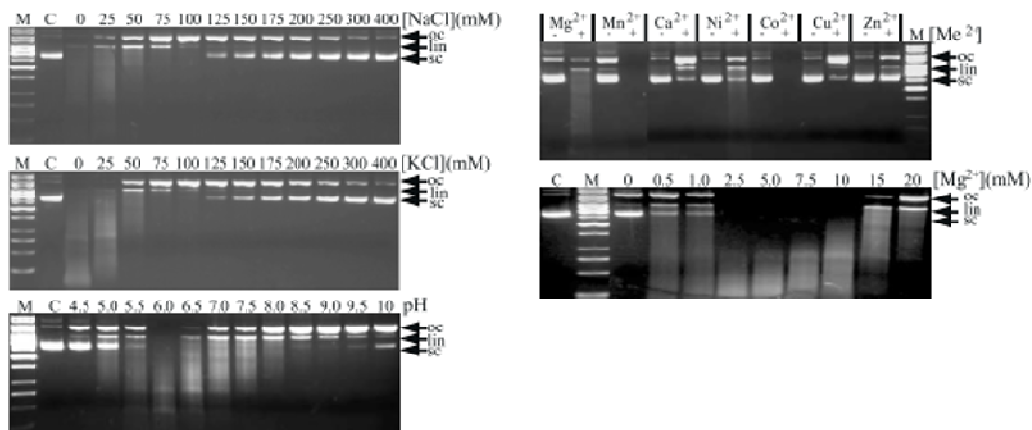


Figure 30 Biochemical properties of recombinant bovine EndoG (Schäfer et al. 2004). The plasmid DNA was digested into different forms (sc, oc, lin) by GST-tagged recombinant bovine EndoG in various conditions. sc, oc, lin: sc: supercoiled; oc: open circular; lin: linear.

III.2 Structures of *Nuc1* homologs

The EndoG family members belong to the large family of DNA/RNA nonspecific $\beta\alpha$ -Me-finger nucleases (Schäfer et al. 2004). The active site of EndoG nucleases is characterized by the DRGH-motif, which contains the catalytically important Histidine residue and structurally is part of the $\beta\alpha$ -Me-finger motif (Figure 31). The $\beta\alpha$ -Me-finger is found in many nucleases that exhibit globally distinct structures, including the caspase-activated DNases (Walker et al. 2002; Scholz et al. 2003).

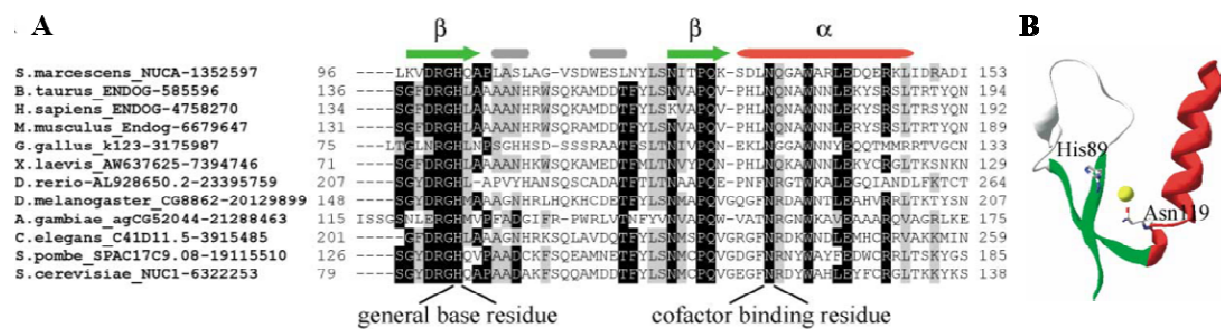


Figure 31 EndoG is a DNA/RNA non-specific nuclease belonging to the superfamily of $\beta\alpha$ -Me-finger nucleases (Schäfer et al. 2004).

(A) Sequence comparison of the active-site region of selected DNA/RNA non-specific nucleases. The DRGH-motif contains several catalytically important amino acid residues, in particular the general base histidine. The highly conserved asparagine residue in *Serratia* nuclease is the sole direct cofactor ligand, essential for divalent metal ion binding. (B) The active-site of *Serratia* nuclease shows the typical $\beta\alpha$ -Me-finger fold that is also found in a variety of other nucleases.

Structures of EndoG homologues from *Drosophila melanogaster*, *Caenorhabditis elegans*, and human were solved (Loll et al. 2009; Lin et al. 2012). They all show a homodimeric conformation, binding divalent metals. For example, the structure of *C. elegans* EndoG (CPS-6) is a homodimer coordinating one Mg ion in each monomer (Figure 32). It has a mixed $\alpha\beta$ topology with the two $\beta\alpha$ -metal finger nuclease motifs located distantly at the two sides of the dimeric enzyme (Figure 32A) (Lin et al. 2012).

A structure model of CPS-6-DNA complex reveals a positively charged DNA binding groove near the active site where Mg^{2+} is bound (Figure 32C). Mutagenesis of residues Phe¹²², Arg¹⁴⁶, Arg¹⁵⁶, and Phe¹⁶⁶ in the protein-DNA interface significantly reduced the DNA binding and cleavage activity of CPS-6, confirming that these residues are critical for CPS-6-DNA interactions (Figure 32D). These DNA binding residues and catalytic site residues are important for apoptosis promoting in *C. elegans* (Lin et al. 2012).

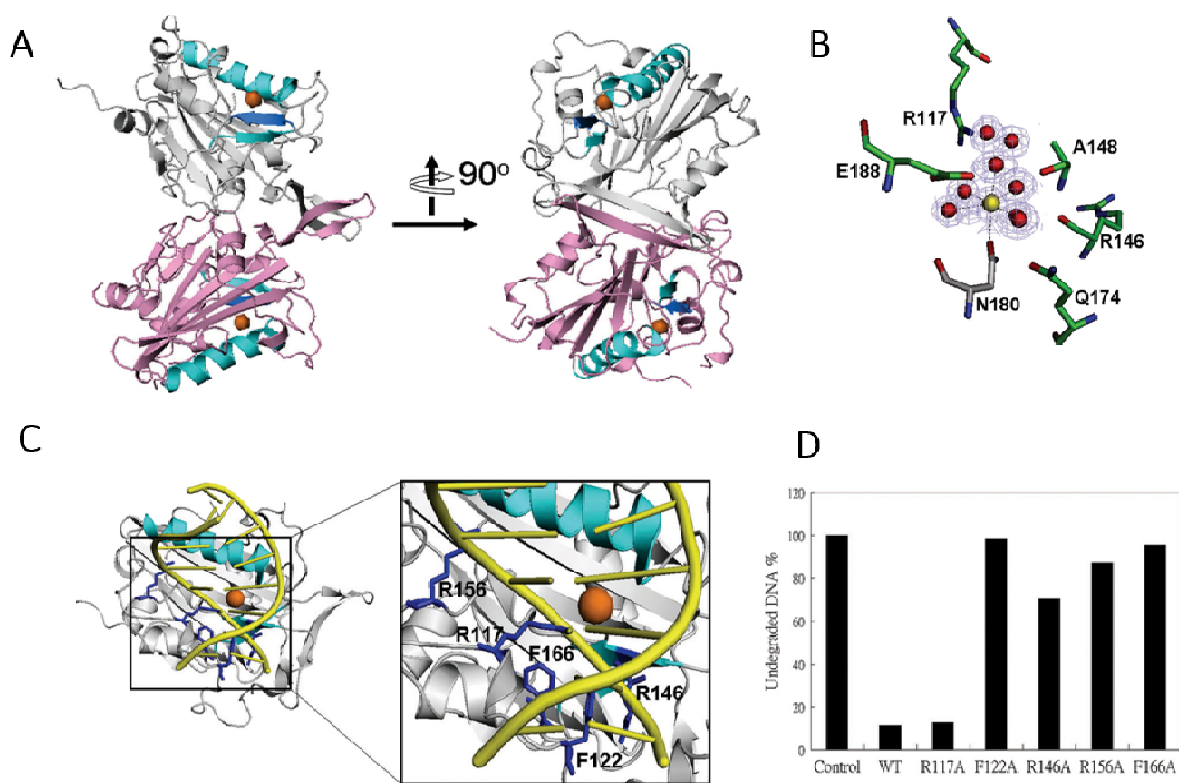


Figure 32 Crystal structure of CPS-6, the *Caenorhabditis elegans* EndoG (Lin et al. 2012).

(A) Over all structure of dimeric CPS-6 are presented (residues 63–305, H148A mutant). The structure was determined at 1.8 Å resolution (PDB code 3S5B). The two protomers are displayed in pink and gray, respectively. The $\beta\alpha$ -metal finger motifs are displayed in cyan with a Mg^{2+} ion (orange sphere) bound at the active site. (B) catalytic sites of CPS-6. (3) CPS-6-DNA complex model. It suggests that several basic and aromatic amino acid residues (in blue) including Phe¹²², Arg¹⁴⁶, Arg¹⁵⁶, and Phe¹⁶⁶ and Arg¹¹⁷ are located closely to the DNA backbones (yellow), likely involved in protein-DNA interactions. (D) the DNA digestion assays show that all the mutations in (C) are involved in DNA digestion or binding, except Arg¹¹⁷.

III.3 Dual role of Nuc1

EndoG has been suggested having a dual pro-death and pro-life role for the cell. Evidence that EndoG is responsible for apoptosis came from knockout experiments in *C. elegans* (Parrish et al. 2001; Li et al. 2001) as well as from numerous cell-culture studies (Arnoult et al. 2003; van Loo et al. 2001). In contrast, the pro-life role of mammalian EndoG was associated with mitochondrial DNA replication, DNA recombination, and cell proliferation (Low 2003)(Huang et al. 2006). We will concentrate for our introduction on the role of yeast EndoG: Nuc1.

III.3.1 Apoptosis under high respiratory conditions

Nuc1 was shown to be an apoptosis inducer under certain conditions by a comprehensive set of genetic (clonogenic survival test) and biochemical experiments (measurement of apoptotic markers) (Büttner, Eisenberg, et al. 2007). When cells are exposed to various apoptotic triggers, deletion of NUC1 (*Δnuc1*) clearly rescues the cells. Accompanied with the increased cell viability of *Δnuc1* under these conditions, suppression of apoptosis signals (including production of reactive oxygen species, phosphatidylserine externalization, and DNA fragmentation) was observed both in mammalian and budding yeast cells. Conversely, overexpression of Nuc1 promotes apoptosis. Importantly, a mutational study of catalytic residues of Nuc1 shows that its endonuclease activity is required for the promoting-death role (Büttner, Eisenberg, et al. 2007).

The mitochondrial permeability transition pore complex (PTPC) permeabilizes the mitochondria, releasing several mitochondrial proteins into the cytosol, which then turn into death effectors (Zamzami et al. 2005). The voltage-dependent anion channel VDAC in the outer-membrane and the AAC (yeast homologue of mammalian adenine nucleotide translocator [ANT]) in the inner- membrane are conserved components of the PTPC. In the Nuc1 mediated apoptosis, Nuc1 is found to physically interact with AAC, and the interaction is required for its promoting-death role. (Burhans & Weinberger 2007).

In addition, Nuc1 also physically interacts with the highly conserved histone H2B, one of the five main histone proteins involved in maintaining the structure of chromatin in eukaryotic cells (Wu et al. 2016). H2B phosphorylation is considered to be a universal prerequisite for apoptosis execution (Cheung et al. 2003). A mutation in H2B inhibited apoptosis in budding yeast and mammals by blocking its phosphorylation in cells exposed to various apoptosis stimuli (*i.e.* H₂O₂), and Nuc1-mediated apoptosis is attenuated in this H2B mutant. Nuc1-

dependent apoptosis is also attenuated by deletion of KAP123, a protein involved in nuclear import (Büttner, Eisenberg, et al. 2007) (Burhans & Weinberger 2007).

In contrast, deletion of AIF1 (apoptosis inducing factor, also excluded from mitochondria during apoptosis) or YCA1 (encoding budding yeast metacaspase Yca1p), does not alter Nuc1-dependant cell death. Furthermore, although Yca1p-dependent apoptosis was also attenuated by mutation in the PTPC, unlike Nuc1-dependent apoptosis, it is not attenuated by deletion of KAP123 or by mutation of H2B mentioned above. Therefore, Yca1p-dependent apoptosis is in a pathway separate from Nuc1 in budding yeast (Büttner, Eisenberg, et al. 2007). However, in the case of mammalian EndoG there has been contradictory and controversial arguments about its specific role in apoptosis. In some studies it is claimed that EndoG regulates apoptosis independently from caspase and AIF; while others suggest these proteins are required for EndoG-mediated cell death (Ekert & Vaux 2005; David et al. 2006; Irvine et al. 2005) (Burhans & Weinberger 2007).

A pathway of Nuc1-mediated cell death was proposed (Figure 33). In the normal state of the cells, Nuc1 localizes in the inner membrane of the mitochondria. Upon apoptosis induction, Nuc1 interacts with VDAC and is translocated to the cytosol. Another pro-apoptotic protein BNIP3 is also found to interact with VDAC to increase opening probabilities of mitochondrial permeability transition pores and induce mitochondrial release of EndoG (Zhang et al. 2014). In the cytosol, Nuc1 interacts with the karyopherin KAP123, which is involved in nuclear import, and transported to the nucleus. In the nucleus Nuc1 interacts with phosphorylated histone H2B and mediates cell death. Aif1p or Yca1p is not required in this process of Nuc1-dependent apoptosis (Büttner, Eisenberg, et al. 2007).

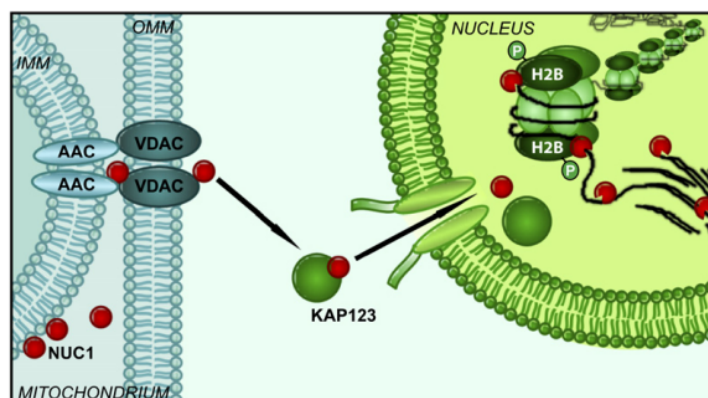


Figure 33 A Potential Pathway for Yeast EndoG-mediated cell death (Büttner, Eisenberg, et al. 2007).

More recently it was reported that overexpression of the tumor suppressor p53 from human in *S. cerevisiae* leads to growth inhibition and apoptotic cell death on minimal medium, and the p53-overexpression cell are susceptible to cell death after H₂O₂ induced cell death, and mainly mediated by Nuc1 (Palermo et al. 2013).

Similar to mammalian apoptosis, yeast apoptosis is under mitochondrial control: growth of cells on non-fermentable carbon sources (such as glycerol) potentiates mitochondrial mass and respiration, and therefore amplifies the impact of mitochondrial death effectors during apoptosis (Kroemer & Martin 2005) (Büttner, Eisenberg, et al. 2007).

III.3.2 Nuc1 prevents cell death in glucose medium

On the contrary with the pro-death effect of Nuc1 described above, Nuc1 has a pro-life effect on budding yeast cells under specific conditions: deletion of NUC1 inhibits apoptosis when cells grow in the usual glucose medium (fermentative growth) (Büttner, Eisenberg, et al. 2007).

When budding yeast cells were cultured in glucose, causing more frequent cell division, deletion of NUC1 reduces viability via a non-apoptotic mechanism, and causes growth defects that lead to accumulation of cells in the G2 phase of the cell cycle (Büttner, Eisenberg, et al. 2007). This is remarkably similar to the accumulation of proliferating mammalian cells in G2 when EndoG expression is knocked down by siRNA (Huang et al. 2006). As has been observed for other pro-death proteins in metazoans and budding yeast, these effects point to a highly conserved prolife role for EndoG, in addition to a conserved role in apoptosis (Burhans & Weinberger 2007).

The “vital” function of EndoG is also illustrated by the fact that depletion of EndoG selectively kills polyploid cells (Büttner, Carmona-Gutierrez, et al. 2007). Deletion of the yeast Endonuclease G gene causes the complete elimination of tetraploid cells during exponential growth. Consistently, conditional knockdown of mammalian EndoG selectively kills tetraploid but not diploid clones of the human HCT116 colon carcinoma cell line. So EndoG is important for the viability of polyploidy. In a yeast mutant lacking the NUC1 endo/exonuclease, mitochondrial DNA recombination was reduced (Zassenhaus & Denniger 1994).

Taken together, the Nuc1 role of pro-life or pro-death could be summarized as follows (Figure 34): in high respiratory activity cells, Nuc1 mediates apoptosis through its interaction with mitochondrial adenine nucleotide translocator AAC, the nuclear import protein KAP123, and phosphorylated histone H2B, a process that is independent from Aif1p or Yca1p. Under

the conditions of fermentative growth (glucose medium), Nuc1p promotes viability instead via unknown mechanisms (Burhans & Weinberger 2007).

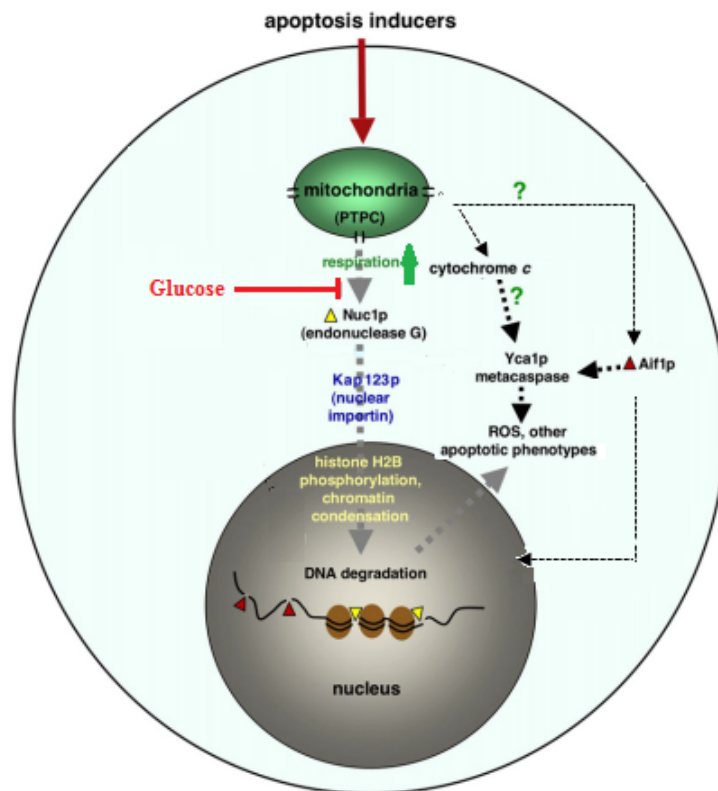


Figure 34 Nuc1 has dual lethal or vital role in Budding Yeast under apoptosis signal (adapted from (Burhans & Weinberger 2007)).

IV. Glucose-6-phosphate mutarotase

IV.1 *YMR099C* encodes a glucose-6-phosphate mutarotase

The *YMR099C* gene from *Saccharomyces cerevisiae* codes for a glucose-6-phosphate mutarotase (Glc6P mutarotase), a 34 kDa protein. It catalyzes the interconversion of the α to β anomers of glucose-6-phosphate, although the two anomers are in spontaneous equilibration in solution (Graille et al. 2006). The structure of Glc6P mutarotase is made of a β -sandwich of 22 β -strands decorated by two short α -helices. The strands are organized in four anti-parallel β -sheets arranged in parallel layers (Figure 35A). Glc6P mutarotase has the same fold as galactose mutarotase from *Lactococcus lactis*; superimposition of the two structure gives an r.m.s.d. of 2.2 – 2.6 Å over 240 C α positions, with 15%-17% sequence identity (Graille et al. 2006).

The structure of Glc6P mutarotase revealed the presence of a sulfate ion in the active site. This observation helped to discover the biochemical function of the enzyme. The sulfate ion is attached by an arginine clamp made by the side chains from two strictly conserved arginine residues (Figure 35C). This sulfate is ideally positioned to mimic the phosphate group of the hexose 6-phosphate substrate (Graille et al. 2006). The sulfate ion just apposes to a conserved binding pocket, which is bound to a glycerol moiety and characterized as the active site of the galactose mutarotase paralogous enzyme. The pocket of Glc6P mutarotase contains five strictly conserved residues: two histidine residues (His82 and His159), two acidic amino acids (Asp203 and Glu264) and Tyr161 (Figure 35C).

Glc6P mutarotase is not only active as a glucose-6-phosphate mutarotase, but is also able to mutarotate other D-Hexose-6-phosphates (Figure 37) whereas has not detectable affinity to glucose (Graille et al. 2006).

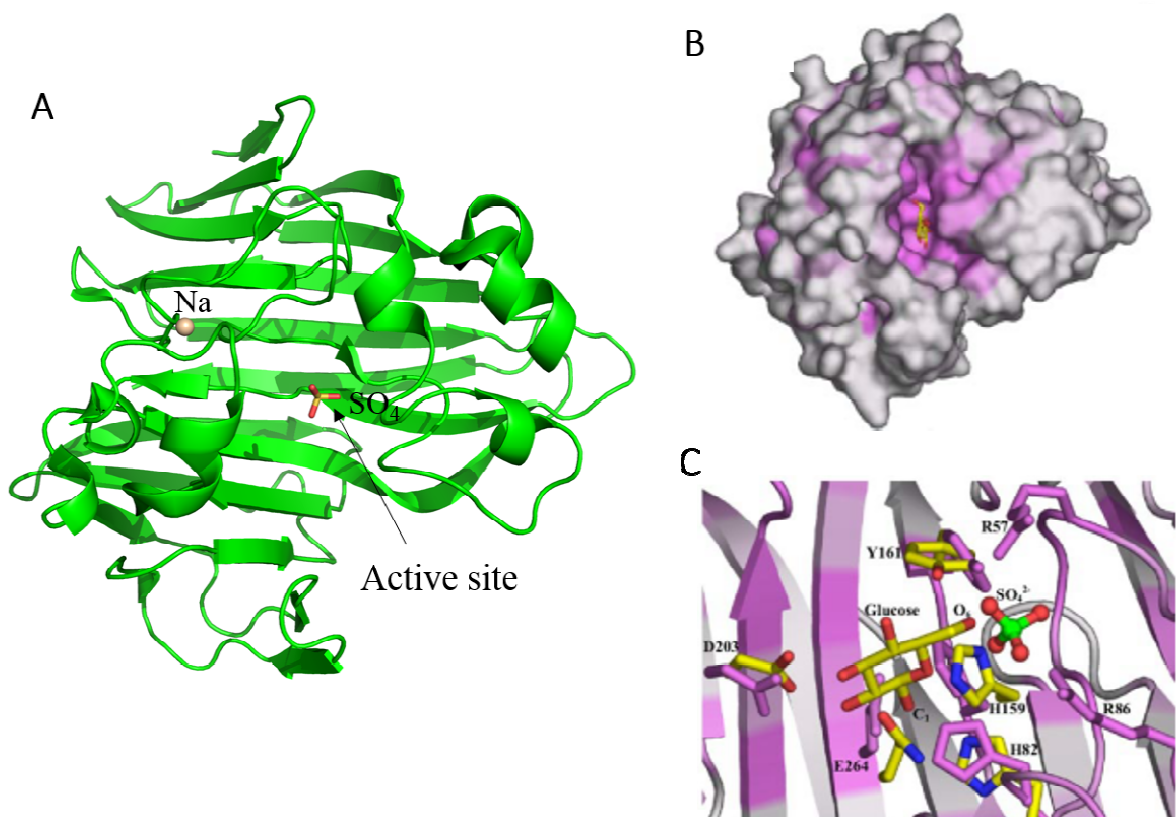


Figure 35 Structure of Glc6P mutarotase from *S. cerevisiae* (PDB: 2CIQ) (Graille et al. 2006).

(A) Overall structure of the Glc6p mutarotase. (B) Surface mapping of residue conservation of Glc6P mutarotase. Coloring from purple to grey shows higher to lower conservation. A glucose molecule shown in yellow stick comes from the superimposition with *L. lactis* galactose mutarotase-glucose complex, is in the putative active site pocket. (C) Active site of Glc6P mutarotase (colored as in (B)) superimposition with the *L. lactis* galactose mutarotase (E304Q mutant) (yellow). The sulfate bound to Glc6P mutarotase, and glucose bound to galactose mutarotase are shown in sticks. For clarity, only Glc6P mutarotase residues are labeled.

It should be noted that Glc6P mutarotase is one of the proteins induced in response to the DNA-damaging agent MMS (Lee et al. 2007), an observation that may be important to understand the biological meaning of the tertiary Trz1/Nuc1/mutarotase complex.

The glucose-6-phosphate mutarotase activity was quantified by a coupled -enzyme assay that I further also used in for my thesis work (Figure 36). Both spontaneous and mutarotase catalyzed formation of β -Glc6P is converted to D-glucono-1,5 6-phosphate in the presence of glucose -6-phosphate dehydrogenase; coupled to with generation of NADPH ($\lambda = 340$ nm) from NADP^+ , the increasing of the absorbance at 340nm are observed (Figure 36A). Therefore, a faster increasing of absorbance at 340nm indicates a faster formation of β -Glc6P resulted from Glc6P mutarotase activity. It is obviously seen that with the increasing concentration (lower than saturation) of Glc6P mutarotase, faster NADPH were generated, demonstrate the Glc6P activity of this enzyme (Figure 36B).

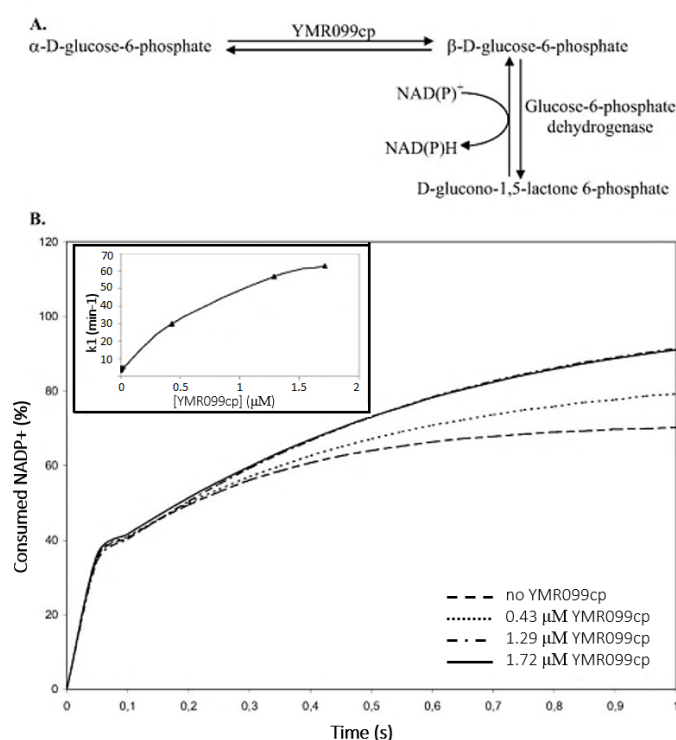


Figure 36

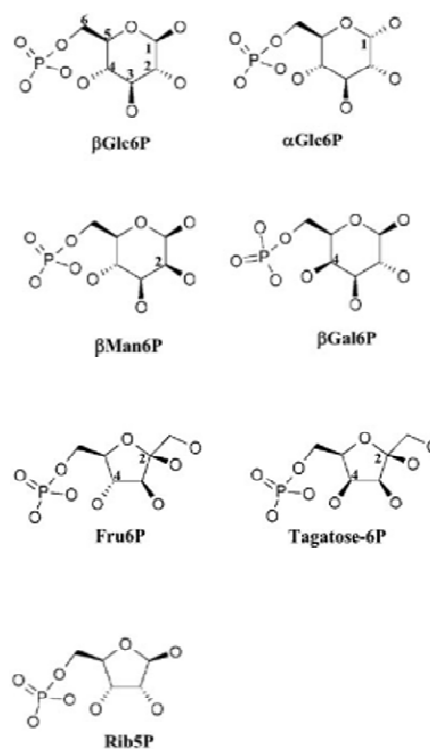


Figure 37

Figure 36 Glucose -6-phosphate activity assay (Graille et al. 2006).

(A) schematic representation of the coupled enzyme assay used to detect the Glc6P mutarotase activity. (B) Activity assay by stopped-flow experiments. The first phase corresponds the initial burst of β -Glc6P because of its presence in the racemic mixture, and hence is independent on Glc6P. In the second phase, faster generation of NAD(P)H reflecting higher velocity constant (k_1 , inset) for the conversion from α -Glc6P to β -Glc6P, which obtained by increasing concentration of Glc6P mutarotase.

Figure 37 Scheme representation of phosphosugars tested bound to Glc6P mutarotase (Graille et al. 2006).

V. Objectives of the thesis

Since the identification of RNase Z enzymes 15 years ago, several structures of the short form RNase Z from bacteria and human have been solved. Despite these accumulating data on the structure and function of RNase Z enzymes, knowledge about the structure and mechanism of the long form RNase Z is lacking. Sequence analysis of the long form RNase Z suggested the presence of two β -lactamase domains, each covering roughly one half of the sequence. The two halves of the RNase Z^L however are not equivalent with respect to the presence of catalytic and substrate binding sequence motifs. Despite the predictions on the overall structure of the long form RNase Z, a better understanding of the enzyme required the determination of its crystal structure. This was the first main objective of the thesis. This structure should allow us to analyze how the two β -lactamase domains are organized within the protein and provide information on the structural organization of the very long extensions that exist in the long form compared to the short form RNase Z sequences. It should also inform on why there is such an important sequence divergence between the N- and C-terminal domains. The latter question is strongly related to the mechanism by which the tRNA substrate is bound and trimmed by the asymmetric long form RNase Z compared to the dimeric symmetric short form. Comparison of structure should also inform on how RNase Z^L evolved from RNase Z^S. Furthermore, the human ELAC2 gene is connected with to a number of diseases. ELAC2 is a susceptible gene to prostate cancer; mutations in ELAC2 were also related to hypertrophic cardiomyopathy. The structural analogy between the yeast Trz1 and the mammalian RNase Z^L enzymes will allow constructing models which will provide a more rational basis for the molecular understanding of these genetic diseases.

Several global proteomic studies in yeast have established that RNase Z^L forms a ternary complex with two other enzymes Nuc1 and mutarotase. The biological significance of this complex is unknown. The structure and biochemical function of mutarotase were studied in the laboratory a few years ago. The second main objective of the thesis is to biochemically confirm the existence of the RNase Z^L/Nuc1/mutarotase complex and to study its 3D structure and its enzymatic properties. These data should help us in understanding its biological role.

RESULTS

I. Crystal Structure of Trz1

No crystal structure is currently available for a long form RNase Z. Although it was predicted that the long form would be composed of two β -lactamase domains, the sequence divergence of the N-terminal domain is considerable and precludes its confident modelling. In order to probe the differences between the short and long form RNase Z enzymes we decided to obtain structural information on the long form. We therefore set out to determine the crystal structure of Trz1 from *S. cerevisiae*, a long form RNase Z consisting of 838 amino acids.

I.1 Structure determination of Trz1

I.1.1 Purification of Trz1

The RNase Z gene from *S. cerevisiae* (TRZ1) was cloned into a pET-45 vector with a N-terminal fusion to His-tag and an enterokinase cleavage site. The best expression yield was obtained using the *E. coli* BL21-Gold host strain and induction was done at 15°C. Despite the low expression level, we attempted purification of the protein (Figure I-1 A, lane “T”).

His-tagged Trz1, which we will from now simply call Trz1, was purified in three steps: first an affinity chromatography step using a Ni-NTA resin, subsequently an ion-exchange step using a mono Q column and a final size exclusion chromatography purification step (experimental details are found in the legends of Figure I-1). Cells expressing the Trz1 construct were lysed in buffer 20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 5 mM BME plus 10% of glycerol. During the Ni-column affinity chromatographic step, Trz1 was majorly eluted using 100 mM of imidazole although at 20 mM it already was some leakage from the resin. The eluted Trz1 protein was contained a number of contaminants that could be removed by running the sample over an ion-exchange column. In the final gel filtration purification step, Trz1 eluted as a symmetric peak (Figure I-1 D-E). The purity of the sample is more than 95% (Figure I-1 F). In the last step Trz1 was solubilized in buffer made of 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM BME, which is the classic buffer used for the final gel filtration purification step all through this work unless stated otherwise. The final yield was poor, 0.2 mg of Trz1 starting from a 800 ml bacterial culture.

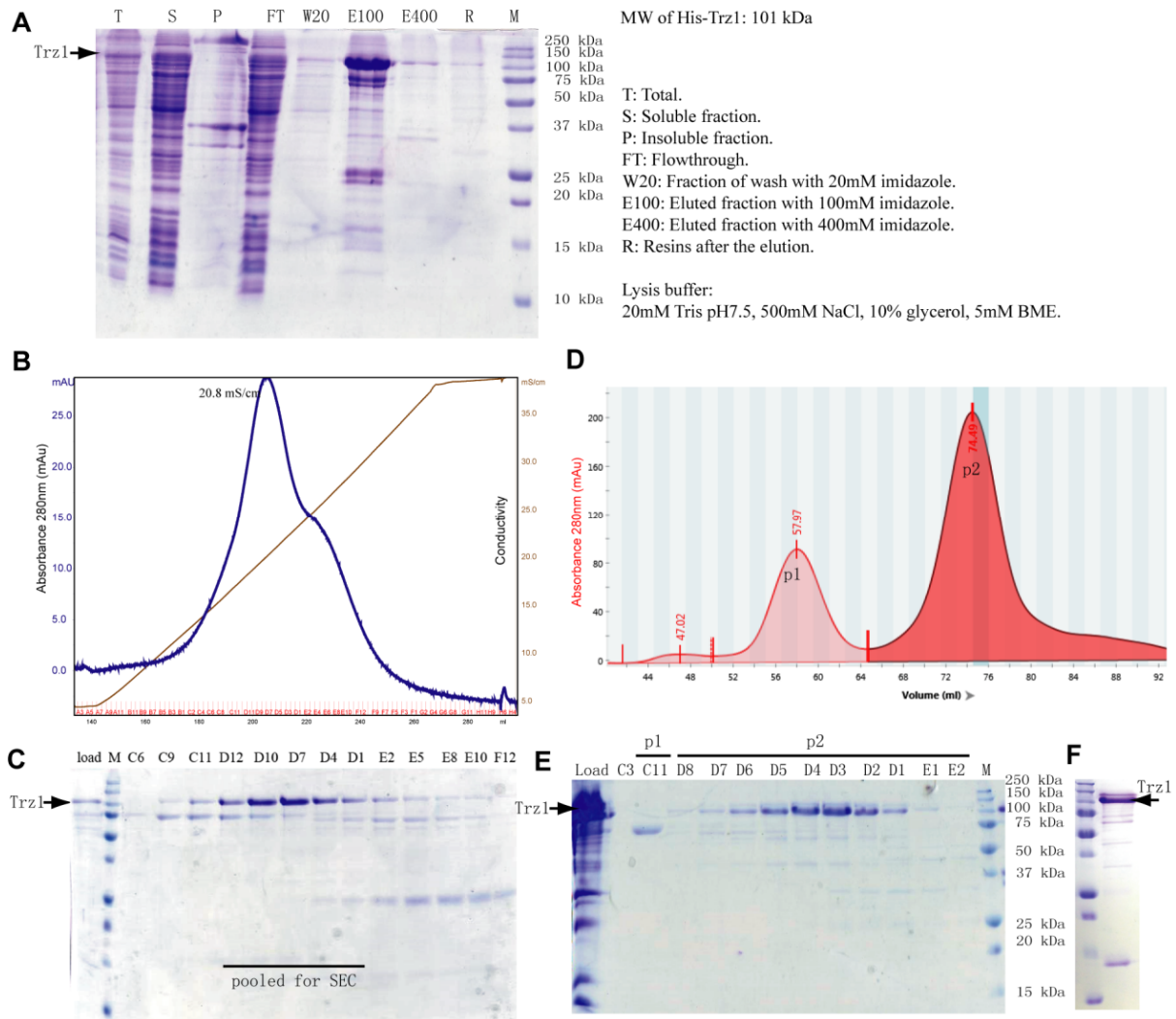


Figure I-1 Purification of Trz1. (A) SDS-PAGE of the fractions obtained by Ni-NTA affinity purification. (B) Ion change chromatography. Column: monoQ 5/50GL. Buffer low/high salt: 20 mM Tris-HCl, pH 7.5, 50 mM /500 mM NaCl, 10% glycerol, 10 mM BME. (D) Size exclusion chromatography(SEC). Column: Superdex 200 Hiload 16/60. Buffer: 20 mM Tris-HCl, pH7.5, 200 mM NaCl, 10 mM BME. (C) (E) SDS-PAGE of the eluted fractions of (B) and (D), separately. (F) SDS-PAGE of purified Trz1.

I.1.2 Crystallization and optimization of Trz1

Purified Trz1 was concentrated at 2.5 mg/ml or 5 mg/ml and protein solutions supplemented with or without 1mM of dGMP were used for crystallization. 100 nL of protein was mixed with 100 nL of the crystallization reagents from screening kits (Qiagen) at a 1:1 ratio using the sitting drop set up and a pipetting Cartesian Robot. Crystallization drops were inspected regularly using a Formulatrix visualization robot.





Crystals were obtained under many conditions, which had in common that they contained different type of PEGs as the main precipitant (Table I-1). Crystals were optimized by applying gradients of the reagents in the crystallization solution; by varying the pH of the

crystallization solution, the protein concentration and the crystal grow temperature. Various additives were tested. For the native protein, the phasing reagents I3C (5-Amino-2,4,6-triiodoisophthalic acid) and Gadolinium (Gd) were also tried, but lead to crystals that diffract only at low resolution. Both hanging drop (1-4 μ l) and sitting drop (200 nL) were used in crystal optimization.

We tested about 170 Trz1 crystals obtained from various conditions (Table I-1). In the case of Trz1, crystals that looked well-formed under the light microscope did not diffract. Figure I-2 shows an example of a good-looking crystal diffracting to 8 Å. In additive screening, more than 50% of the conditions gave crystals. Crystals of better size or shape did not improve diffraction quality. The dehydration technique (which consists of moving crystals gradually to higher-concentration precipitates every 3h) also led to weak diffracting crystals, maybe caused by mechanical damage during the transfer. For some crystallization conditions the small crystals from initial screening sitting drops diffracted better than bigger crystals optimized by hanging drops using home-made solutions. The crystals were usually mounted and flash frozen within two weeks after set up.

We also tried proteinase digestion of Trz1 aiming to optimize the construct of Trz1 for crystallization. But mass spectrometry of the main bands of the digest products showed that a cleavage site was present about in the middle of the N-terminal domain, therefore this approach was discarded (data not shown). *In situ* use of proteinases during the crystallization did not give any crystals.

Table I-1 Examples of crystallization conditions of native Trz1 and the diffraction

Buffer	Precipitant	Salt	Cryo*	Diffract resolution	crystal image	Freeze Day
0.1 M Hepes pH 7.5	25% PEG4K, 8% 2-propanol	0.1 M NaAc	-	7.8 Å		8th
0.1 M Hepes pH 7.5	25% PEG6K	0.1 M LiCl	-	7.8 Å	 	4th
0.1 M MES pH 6.5	24% PEG4K	0.1 M NaAc	EG	10 Å		12th
0.1 M MES pH 6.5	26% PEG12K	0.1 M NaAc	Gly /EG	9 Å		13th
0.1 M MES pH 6.5	30% PEG 2K mme	0.1 M NaAc	EG	4.6 Å, 9 Å		13th
0.1 M Sodium Cacodylate	30% PEG 8K	0.2 M NaAc	Gly /EG	8 Å	Good	13th
0.1 M MES pH 6.5	25% PEG 8K	-	Gly	Weak	Good	13th

*Cryo: cryoprotectant. EG: ethylene glycol; Gly: Glycerol.

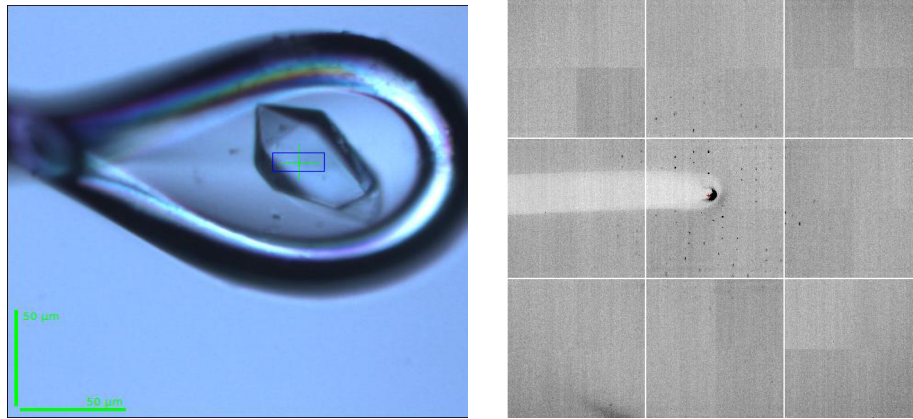


Figure I-2 A. Crystals of native Trz1 collected at Px2 (SOLEIL, France).

Left: Best look-like crystal (loop dimension: 0.05-0.1mm). Condition: 0.1M MES pH 6.5, 24% (w/v) PEG 4000, 0.1M NaAc. cryoprotectant: 20% ethylene glycol. Right: An example of diffraction pattern of the crystal shown on the left which diffracts about 8 Å resolution.

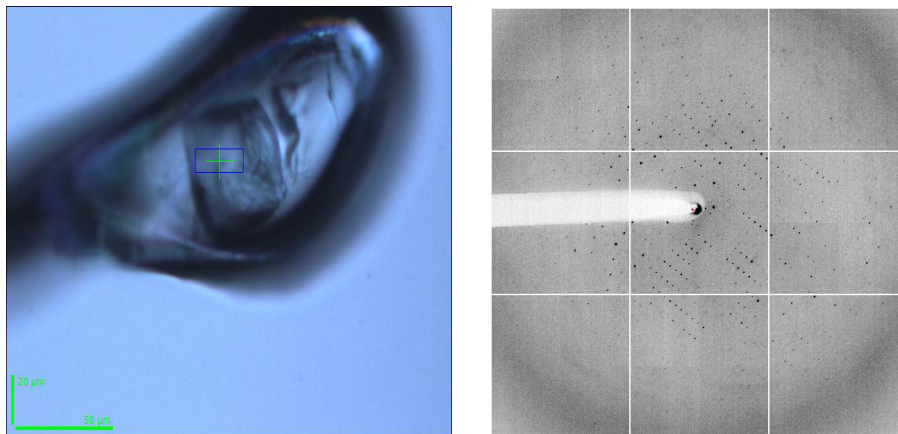


Figure I-2 B. Crystals of native Trz1 collected at Px2 (SOLEIL, France).

Left: Best diffracting crystal (loop dimension: 0.05-0.1 mm). Condition: 0.1 M MES, pH 6.5, 30% (w/v) PEG 2000 mme, 0.1 M NaAc. Cryoprotectant: 20% ethylene glycol. Right: Diffraction pattern of the crystal shown on the left, which diffracts about 4 Å resolution.

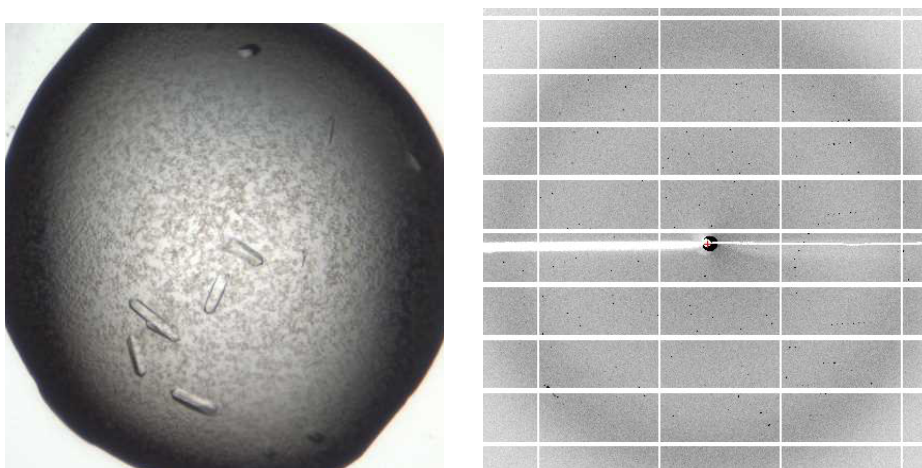


Figure I-3 Crystals of Selenomethionyl Trz1 collected at Px1 (SOLEIL, France).

Left: Best diffracting crystal of Selenomethionyl Trz1. Crystal dimension: up to about 60x120μm. Protein concentration: 5 mg/ml, plus 1 mM dGMP. Condition: 0.1 M Hepes, pH 7.5, 25% (w/v) PEG 8000. Cryoprotectant: 20% glycerol. Right: Diffraction pattern of Selenomethionyl Trz1 (crystal diffract to about 3 Å resolution).

I.1.3 Diffraction and data processing of native Trz1

The best crystals of native Trz1 were grown from 0.1 M Hepes, pH 7.8, 25% (w/v) PEG 6000, 0.1 M LiCl and diffracted up to 4.3 Å. Since molecular replacement, using search models like the existing structures of short form of RNase Zs or derived structures from Phyre2, did not provide a solution; and co-crystallization with the phasing reagent I3C (5-Amino-2,4,6-triiodoisophthalic acid) and Gadolinium (Gd) lead to low-resolution crystals, we directed our efforts towards the Se-methionine substituted version of Trz1.

It is worth noting that zinc was detected in the native crystals of Trz1 by measurement of fluorescence as a function of the energy of X-ray. Scattering factors f' and f'' were measured at the zinc threshold (Table I-2). Zinc ions were not explicitly added in the preparation of the protein, nor in the crystallization liquor, indicating Trz1 has a high affinity for zinc that was incorporated from the culture broth. This was also observed during the structure analysis of the short form of RNase Z from *B. subtilis* (Li de la Sierra-Gallay et al. 2005).

Table I-2 Measured scattering factors f' and f'' at zinc edge

	Energy (keV)	f'	f''
Native	12.65	-	-
Zinc	9.673	7.26	6.27

However, the anomalous signal from zinc was not strong enough to solve the phase problem. Then we turned to the Selenomethionyl Trz1. Crystals of Selenomethionyl substituted Trz1 diffracted up to 3.39 Å; which is much better than for the native crystals and we therefore used the same data set for solving and refining the structure. So here I present only the data processing from this anomalous dataset.

I.1.4 Structure determination of Selenomethionyl Trz1

Selenomethionyl Trz1 was purified using the same protocol as for the native protein. The protein was screened at 2.5 or 5 mg/ml plus 1mM of dGMP in the classical buffer A (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM BME). Crystals appeared under 18 different conditions all containing the main precipitate PEG and some conditions are the same as for the native crystals.

The best diffracting crystal was grown using a protein solution at 5mg/ml mixed with 1mM of dGMP from 0.1 M Hepes, pH7.5, 25% (w/v) PEG 8000 at 18 °C (Figure I-3). The crystal was cryoprotected with 20% glycerol in the crystallization liquor.

I.1.4.1 Diffraction and data collection

The presence of Se in the crystal was confirmed by an energy scan. Then three anomalous datasets at Se peak wavelength were collected at the beamline PROXIMA 1 at Synchrotron SOLEIL. The diffraction patterns were recorded on the Pilatus detector. The collection strategy was the same for the three datasets: 2000 images collected using a rotation step of 0.2°. The best dataset was collected at a nominal resolution of 3 Å. An example of a diffraction pattern is shown in Figure I-3.

We used the best dataset for SAD phasing and the resolution.

I.1.4.2 Image processing for phasing

Datasets were processed by the XDS and XSCALE programs (Kabsch 2010). The processing statistics of best data are shown in Table I-3. The spacegroup is P4₁ and the parameters of the unit cell are $a = 136.07$ Å, $b = 136.07$ Å, $c = 115.88$ Å, $\alpha = \beta = \gamma = 90^\circ$. The resolution was cut off at 3.39 Å (Figure I-4). The signal to noise was $I/\sigma=2.62$ in the last resolution shell and the completeness of the data set was 98.7%. The R-means is 20.8 % (84.7 % for the last shell).

LATTICE- CHARACTER	BRAVAIS- LATTICE	QUALITY OF FIT	UNIT CELL CONSTANTS (ANGSTROM & DEGREES)						REINDEXING TRANSFORMATION													
			a	b	c	alpha	beta	gamma														
* 44	aP	0.0	116.0	136.1	136.1	90.0	90.0	90.0	0	0	1	0	1	0	0	0	0	0	1	0	0	0
* 31	aP	0.0	116.0	136.1	136.1	90.0	90.0	90.0	0	0	1	0	-1	0	0	0	0	0	-1	0	0	0
* 25	mC	0.0	192.5	192.5	116.0	90.0	90.0	90.0	1	1	0	0	-1	1	0	0	0	0	0	1	0	0
* 20	mC	0.0	192.5	192.5	116.0	90.0	90.0	90.0	-1	-1	0	0	-1	1	0	0	0	0	0	-1	0	0
* 33	mP	0.0	116.0	136.1	136.1	90.0	90.0	90.0	0	0	1	0	1	0	0	0	0	0	1	0	0	0
* 34	mP	0.0	116.0	136.1	136.1	90.0	90.0	90.0	0	0	-1	0	0	-1	0	0	0	-1	0	0	0	0
* 35	mP	0.0	136.1	116.0	136.1	90.0	90.0	90.0	-1	0	0	0	0	0	-1	0	0	-1	0	0	0	0
* 32	oP	0.0	116.0	136.1	136.1	90.0	90.0	90.0	0	0	1	0	1	0	0	0	0	0	1	0	0	0
* 23	oC	0.0	192.5	192.5	116.0	90.0	90.0	90.0	1	1	0	0	-1	1	0	0	0	0	0	1	0	0
* 21	tP	0.0	136.1	136.1	116.0	90.0	90.0	90.0	1	0	0	0	0	1	0	0	0	0	0	1	0	0
14	mC	188.6	178.9	178.9	136.1	90.0	90.0	80.9	1	0	1	0	1	0	-1	0	0	1	0	0	0	0
13	oC	188.6	178.9	178.9	136.1	90.0	90.0	80.9	1	0	1	0	1	0	-1	0	0	1	0	0	0	0
4	hR	188.6	178.9	178.9	224.8	97.3	82.7	114.9	-1	0	1	0	0	1	-1	0	-1	-1	-1	0	0	0

Figure I-4 A. Indexing of the diffraction pattern. The highlighted Bravais lattice is selected because of higher crystallographic symmetry and low penalty score (quality of fit).

RESOLUTION LIMIT	NUMBER OF REFLECTIONS			COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR expected	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano	
	OBSERVED	UNIQUE	POSSIBLE										
9.98	17264	2241	2258	99.2%	5.5%	5.7%	17263	31.20	5.9%	99.8*	89*	3.532	1054
7.13	30922	3942	3942	100.0%	7.8%	8.0%	30922	22.02	8.4%	99.7*	77*	2.410	1905
5.84	39441	5027	5027	100.0%	14.4%	15.0%	39441	13.60	15.4%	99.1*	62*	1.645	2445
5.07	48574	5985	5985	100.0%	16.0%	17.0%	48574	12.72	17.0%	99.0*	47*	1.332	2927
4.54	53089	6776	6776	100.0%	14.9%	16.2%	53088	13.29	16.0%	99.0*	39*	1.176	3319
4.15	58468	7495	7495	100.0%	20.5%	21.4%	58468	10.51	21.9%	98.4*	24*	1.004	3676
3.84	64171	8105	8105	100.0%	34.9%	34.6%	64171	6.81	37.4%	96.4*	15*	0.946	3984
3.59	66564	8749	8749	100.0%	52.1%	50.6%	66562	4.31	56.0%	92.4*	18*	0.966	4306
3.39	68577	9125	9241	98.7%	78.9%	81.1%	68397	2.62	84.7%	84.1*	9	0.830	4374
total	447070	57445	57578	99.8%	19.5%	20.0%	446886	10.22	20.8%	99.2*	44*	1.264	27990

Figure I-4 B. Results of data processing cutting off at 3.39 Å by program XDS: Subset of intensity data as function of resolution.

Table I-3 Data collection and Processing Statistics

Data Collection and Processing Statistics	
Data Collection	
	SetMet
Beamline	Proxima-1, Soleil
Wavelength (Å)	0.979180
Resolution (Å)	50.0 – 3.39 (3.59- 3.39)
Space group	P41
Cell Parameter	$a = 136.07 \text{ Å}, b = 136.07 \text{ Å}, c = 115.88 \text{ Å}$ $\alpha = 90^\circ, \beta = 90^\circ, \gamma = 90^\circ$
Observed Reflections	447070 (68577)
Unique Reflections	57445 (4374)
Completeness (%)	99.8 (98.7)
I/σ	10.22 (2.62)
R-meas (%)	20.8 (84.7)

*Values in parentheses refer to the highest resolution shell.

I.1.4.3 Estimation of cell contents by Matthew's coefficient

In order to estimate the number of copies in the asymmetric unit we calculated the Matthews coefficient, V_m , using the equation:

$$V_m = \frac{V}{M \cdot Z}$$

In which V is the volume of the asymmetric unit in Å³; M is the molecular weight of the protein in daltons and Z is the number of molecules in the asymmetric unit. For protein crystals the mean value of V_m is 2.69 Å³/Dalton and the most common value is 2.34 Å³/Dalton which corresponds to a solvent content of ~47% (Kantardjieff & Rupp 2003).

By using the <Matthews coefficient> program in the CCP4 suit programs, we estimated that there probably were two copies of Trz1 (Mr 100 kDa) per asymmetric unit which would correspond to a $V_m = 2.75 \text{ \AA}^3/\text{Da}$ and a solvent content of 55.25%.

Table I-4 Matthews coefficient

Nmol/asym	Matthews Coeff	\bar{v}_s solvent	P(3.39)	P(tot)
1	5.49	77.62	0.01	0.01
2	2.75	55.25	0.92	0.87
3	1.83	32.87	0.06	0.12
4	1.37	10.50	0.01	0.00

I.1.4.4 Determination of Selenium sites with the program SHELX C/D/E

We used the program SHELX C/D/E (Sheldrick 2010; Sheldrick 2008) for location of the Se sites. The program SHELXC provides a statistical analysis of the input data, estimates the marker-atom structure factors F_A and the phase shifts α , and sets up the files for the other two programs. SHELXD is used for solving the substructure (i.e. locating the marker-atoms). And SHELXE provides iterative phase improvement by density modification (Usón & Sheldrick, 1999; Sheldrick et al., 2001; Schneider & Sheldrick, 2002).

1) SHELXC to process reflections with anomalous signal

Since the Trz1 sequence contains 18 methionines and assuming there are two copies of Trz1 in the asymmetric unit, 36 selenomethionines were expected per asymmetric unit.

<xdsconv.py> was used to convert the file containing reflection information (pos1_coll.hkl) into a format readable by the program SHELXC, where the atom type and number of atom were designated to Se and 36, respectively. SHELXC read 447070 reflections containing 29376 unique reflections up to resolution 3.389 Å from the file, and 170.6 Friedel pairs were used on average for local scaling.

➤ SHELXC output

i) a file trzl.hkl containing native data h, k, l , the intensity F_{obs}^2 , and $\sigma (F_{\text{obs}}^2)$, which can also be used later by SHELXE for density modification;

ii) a file trzl_fa.hkl containing h, k, l , $\|F(hkl)\| - |F(-h-k-l)|$ (Bijvoet differences), and the estimate phase shift α . This file is used by SHELXD to determine substructure coordinates solution. The α estimate is not used by SHELXD, but only used by SHELXE for calculating initial phase for the protein structure, as

$$\varphi_T(hkl) = \varphi_A(hkl) + \alpha(hkl)$$

where φ_T is the whole protein structure phase, φ_A is the reference phase which is the phase of the substructure of the heavy-atom, and the estimate α from observed anomalous signal.

iii) a file `trzl_fa.ins` containing the cell, symmetry and instructions for running SHELXD.

From the log of SHELXC we have statistics of anomalous signal versus resolution (Figure I-5). The `N(data)` shows the number of read-in reflections. `<I/sig>` shows the average intensity against the background. `<d''>` is the average anomalous difference ($d'' = I(+) - I(-)$), `<d''/sig>` gives the strength of the anomalous signal, it should high in the inner shell and asymptote to the random noise (zero anomalous signal) value of 0.8 in the outer shell if the data were processed properly. In our case the anomalous signal is very good until 4.6 Å resolution (Figure I-5). But we could use the data cut-off until 3.9 Å because the `<d''/sig>` is good enough.

Resl	Inf	8.0	6.0	5.0	4.8	4.6	4.4	4.2	4.0	3.8	3.6	3.39
<code>N(data)</code>		2290	3095	3872	1191	1403	1671	2025	2424	2958	3684	4763
<code>Chi-sq</code>		0.98	1.18	1.25	1.29	1.26	1.39	1.45	1.49	1.63	1.39	1.16
<code><I/sig></code>		39.7	21.6	17.8	18.3	19.2	17.5	13.8	10.9	8.1	5.8	3.7
<code>%Complete</code>		99.0	99.9	99.9	100.0	99.9	99.9	100.0	100.0	99.9	100.0	99.2
<code><d''/sig></code>		3.12	1.96	1.47	1.35	1.28	1.24	1.12	1.06	1.02	1.06	0.97
<code>CC(anom)</code>		83.1	59.1	37.5	30.8	28.6	18.6	10.9	8.6	0.9	13.7	3.3

Figure I-5 Partial log file of SHELXC (*shelxc.log*). Resl: resolution. Each column shows the statistics (`<I/sig>`, completeness, completeness, etc..) as function to the resolution.

2) Finding the substructure of the Se-sites by SHELXD

The positions of the Se atoms in the asymmetric unit were located using the SHELXD program. To determine a correct solution: *i*) the correlation coefficient (CC) between the observed and calculated E values (the normalized structure factor to avoid scaling problems in SHELX C/D/E) usually enables correct solutions to be identified unambiguously. *ii*) the value of `CC(weak)`, the correlation coefficient based on the reflections not used in the dual-space recycling, is also a good check.

The input file for SHELXD (Figure I-6 A), as was done by SHELXC, was modified in order to run SHELXD under instruction. The resolution cut-off was modified (line “SHEL”) to run SHELXD four time to reach higher resolution but still sufficient anomalous signal step by step: 5.4 Å, 4.9 Å, 4.4 Å until 3.9 Å were tried. We instructed the program to look for 36 heavy atoms (line “FIND”). We kept the minimum distance between atoms (MIND) at default value. Other parameters we modified to re-run the cycles were: “NTRY” (number of trials)

from 200 to 1500, and “ESEL” (minimum $|E|$ value used in peak picking during dual-space recycling) from 2.0 to 1.0.

```

TITL trzl_fa.ins SAD in p41
CELL 0.98000 136.00 136.00 115.82 90.00 90.00 90.00
LATT -1
SYMM -Y, X, 1/4+Z
SYMM -X, -Y, 1/2+Z
SYMM Y, -X, 3/4+Z
SFAC Se
UNIT 576
SHEL 999 3.9
PATS
FIND 36
MIND -3.5 1
ESEL 1.4
NTRY 1000
SEED 1
HKLF 3

```

Figure I-6 A. An example of SHELXD input file trzl_fa.ins.

Record Type	No.	Atom type	chain type	No.	x	y	z	occupancy	B-factor
HETATM	1	S	HAT	1	106.963	132.613	9.787	1.000	20.00
HETATM	2	S	HAT	2	64.003	90.391	-6.086	0.993	20.00
...									
HETATM	31	S	HAT	31	110.962	125.377	20.766	0.516	20.00
HETATM	32	S	HAT	32	110.305	173.538	35.475	0.489	20.00
HETATM	33	S	HAT	33	120.372	144.611	15.782	0.420	20.00
HETATM	34	S	HAT	34	82.134	123.622	4.383	0.349	20.00
...									
HETATM	49	S	HAT	49	33.931	113.417	-12.769	0.142	20.00
HETATM	50	S	HAT	50	103.998	166.106	-4.413	0.111	20.00
END									

Figure I-6 B. Output file of SHELXD (partial): trzl_fa.pdb. The sharp fall-off in the refined occupancy between site n° 33 and n° 34 is indicated in red square.

To interpret the result of SHELXD, we checked the statistics of CC, CC(weak), and PATFOM. The best solution proposed by SHELXD contains 50 sites, of which the correlation coefficient (CC All) and CC (Weak) values of 47.04% and 23.78%, respectively.

It is also important to find the number of true Se sites among all the sites listed. A sharp fall-off in the refined occupancy between the last true site and the first noise peak usually indicates a good solution (Sheldrick 2010). We found there was a sharp fall-off between site n° 33 and n° 34 (Figure I-6 B), so we judged that site n° 33 was the last true site. Thus we restricted to the first 33 sites for the input for SHELXE.

3) Choice of the hand by SHELXE

SHELXE is the Phasing, Density Modification and Model Building module of the SHELX C/D/E suite. In our case, the data was not good enough to perform phasing and automatically model building using SHELXE. But it was very useful to quickly decide the correct hand.

SHELXE read native data from file trzl.hkl; read estimate anomalous difference angle α from trzl_fa.hkl; and read substructure of heavy atom coordinates from trzl_fa.res (SHELXD output). The command line used for SHELXE was

```
Shelxe trzl trzl_fa -h -s0.54 -m20
```

The option -h was used because the native data (trzl.hkl) does include the Se anomalous signal; -s0.54 assumes the solvent content of 54%; and -m20 is the cycles of density modification. We used various value of the solvent content and the cycles number to have the solution.

It is impossible for SHELXE to distinguish the substructure from its enantiomer with the anomalous data, and there is a 50 % chance that the coordinates in trzl_fa.res are inverted in the correct substructure. Therefore, SHELXE were run a second time with the option “-i” in the command line.

Comparing the solutions in both hands (Figure I-7), the solution with original hand has higher values of <mapCC> (map correlation coefficient) and <Pseudo-free CC> (pseudo-free correlation coefficient, 10% of the reflections left out at random in the calculation of a density modified map) than that of the inverted hand, thus we took the solution with original hand as the correct one.

Mean weight and estimated mapCC as a function of resolution											
d	inf	- 7.37	- 5.83	- 5.09	- 4.62	- 4.28	- 4.03	- 3.83	- 3.66	- 3.52	- 3.39
<wt>	0.586	0.564	0.547	0.595	0.617	0.611	0.581	0.550	0.525	0.495	
<mapCC>	0.778	0.784	0.769	0.819	0.841	0.846	0.839	0.811	0.799	0.758	
N	2942	2949	2924	2963	3001	2917	2917	2988	2913	2862	
Pseudo-free CC = 61.83 %											

Figure I-7 A. Mean weight (<wt>) and estimated mapCC as a function of resolution in the solution with original hand obtained by SHELXE.

Mean weight and estimated mapCC as a function of resolution											
d	inf	- 7.37	- 5.83	- 5.09	- 4.62	- 4.28	- 4.03	- 3.83	- 3.66	- 3.52	- 3.39
<wt>	0.276	0.263	0.327	0.323	0.355	0.408	0.439	0.448	0.437	0.423	
<mapCC>	0.378	0.367	0.468	0.456	0.512	0.622	0.705	0.711	0.704	0.671	
N	2942	2949	2924	2963	3001	2917	2917	2988	2913	2862	
Pseudo-free CC = 42.28 %											

Figure I-7 B. Mean weight (<wt>) and estimated mapCC as a function of resolution in the solution with inverted hand obtained by SHELXE.

4) Identification of the Se- sites that are related by non-crystallographic symmetry

Since there are two copies of Trz1 in the asymmetric unit, we attempted to find the non-crystallographic symmetry (NCS) operator between them. The identification of this operator is crucial if we want to use non crystallographic symmetry averaging.

We used the program <Profess> from *CCP4suite* (Kevin Cowtan, York) to identify NCS related atoms from a list of heavy atom positions. It assembles atom triplets into similar triangles applying any necessary symmetry operators.

In the input 50 Se-sites, 11 pairs were determined to be NCS-related; and these sites were then used for phasing and electron density map construction described below.

I.1.4.5 Phasing and Electron density map calculation

These 22 sites (pdb format) together with the initial MTZ file containing reflection information obtained from *XDS* program (pos1_coll.mtz) single-wavelength anomalous dispersion were fed to the program <Phaser SAD pipeline> (*CCP4suites*) (McCoy et al. 2007). In the program “Mode for experimental phasing” was set to “Single-wavelength anomalous dispersion” (SAD) and F(+), SIGF(+), F(-) and SIGF(-) option were used. The <Phaser SAD pipeline> program output a mtz file with experimental phases (trz1_Se.mtz) and a pdb file containing a new set of sites containing the refined positions of input sites plus some more new ones.

The file containing the new set of sites (pbd format) was input back to the program <Profess> again to be optimized. Similar as above the Se-sites non-related to NCS operator and/or low occupation were removed. After several refinement cycles, 5 sites not related by NCS or having too low occupancy were removed. The final set for phasing contained 36 sites.

These 36 sites were used in phase improvement with the program <Parrot> (*CCP4*) (Kevin Cowtan, York). Provided with an MTZ file containing the structure factor magnitudes and phase probability distributions obtained from <Phaser>, The program <Parrot> performs

phase improvement to improve the phase estimates (or phase probability distributions) obtained from experimental phasing, and thereby improve the electron density map.

I.1.4.6 Model building, refinement, and validation

Using the phases obtained by <Parrot>, and the sequence of Trz1, an initial model was built with the automated model building program <Buccaneer> (CCP4) (Cowtan 2006). The program was run with following options: built selenomethionine, 10 cycles of building/refinement (with 6 internal cycles of building for the first Buccaneer cycle and 4 internal cycles of building for the subsequent cycles).

The model after the first Buccaneer cycle contained 1414 residues built in 76 fragments, the longest having 88 residues. 1198 residues of them were sequenced, after pruning 725 residues were uniquely allocated to 2 chains. This model was built with a completeness of residues of 51.3% . The REFMAC refinement of this model improve both R and R_{free} factors (R/R_{free} factor initial = 0.4416/0.4938 and R/R_{free} factor final = 0.3613/0.4188)

The final model done by Buccaneer contained 1587 residues from 54 fragments, the longest having 145 residues. 1334 residues were sequenced, after pruning 1214 residues uniquely allocated to 3 chains. The final model was built with a completeness of residues 76.5 %.

The non-sequenced residues (named UNK by Buccanner) were manually rename to alanines and then replaced by the correct residues of the Trz1 sequence, a process guided by the electron density maps (using the software <Coot> (Emsley & Cowtan 2004)); and each cycle of modification of the model was verified with the program <Refmac5> (CCP4) (Vagin et al. 2004) .

To identify errors or missing information in the constructed model, difference electron density maps are calculated. Two types of electron density maps are calculated:

$$2F_{obs} - F_{calc} \quad \text{and} \quad F_{obs} - F_{calc}$$

Where $2F_{obs} - F_{calc}$ represent the electron density of the molecule, and $F_{obs} - F_{calc}$ is the density map that highlights the residual differences between the electron density calculated from the model and the observed electron density.

To assess the quality of the Trz1 model, the structure factors calculated from the model (F_{calc}) are compared with the observed structure factors (F_{obs}). This agreement is quantified by the R factor, defined as

$$R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

the R_{free} factor is defined by the same equation as R factor but calculated for 5% of the randomly selected reflections that are set aside at the start of the refinement procedure and never included during the refinement. The R_{free} factor measures in a non-biased manner how well the current atomic model accounts for the experimental data; whereas the R factor measures how well the entire data set which generated the model are predicted in the atomic model. At the end of the refinement, we arrived at $R / R_{free} = 19.27\% / 23.8\%$ values, which indicate that the refinement was satisfactory for a 3.4 Å resolution data set (Table I-5).

The stereochemical quality of the model is judged by the average root mean square deviations (r.m.s.d.) of the bond length and angles in the model. Ideal stereochemical values are derived from high resolution small molecule structures. In our final model of Trz1, the r.m.s.d. values for bond length and bond angles are 0.0145 Å and 1.9825°, respectively (Table I-5).

Table I-5 Refinement Statistics

Refinement Statistics	
Resolution (Å)	3.39
R (%)	19.27
R_{free} (%)	23.8
r.m.s.d. bond length (Å)	0.0145
r.m.s.d. bond angles (°)	1.9825
Ramachandran plot	
Favored regions	85 %
Allowed regions	9 %
Outlier	5 %

The stereochemical quality of the model is also judged by the aid of the Ramachandran plot, representing the experimental main chain conformation angles ϕ and ψ . The combination of ϕ and ψ angles are limited by the constraints of the peptide bond and the secondary structures of the protein. The favorable regions regarding ϕ and ψ angles and conformationally unrealistic regions of the model are visualized by the Ramachandran diagrams. In our final model 85 % of residues are in favored regions and 9 % are in allowed regions, indicating a correct stereochemistry of the peptide chain.

Table I-7 Programs used to go from diffraction data to the atomic structure model

No.	Step	program	output
1	Image processing	XDS	Reflection information (Cell parameter, space group, etc.)
2	Se sites searching	Shelx C/D/E	33 sites used of total 50 found sites; original hand
3	Identify the Se atoms related with NCS	Professs (CCP4)	22 sites
4	Search of new sites and refinement of its positions, occupancy and Phasing	Phaser SAD Pipeline (CCP4)	Experimental phases from 41 Se sites
5*	Identify the Se atoms related with NCS	Professs (CCP4)	5 sites NOT related with NCS or low occupation removed, resulting 36 sites
6	Density Modification	Parrot (CCP4)	Improved phases
7	Model building	Buccaneer (CCP4)	1587 residues built in 54 fragments
8	Model building and refinement	Coot Refmac5 (CCP4)	3D Structure of Trz1 were built in two chains
9**	Validation	Coot	-

*go back to step No.3 & 4 several times until satisfied with the sites.

** go back to step N°8 several times until satisfied with the model.

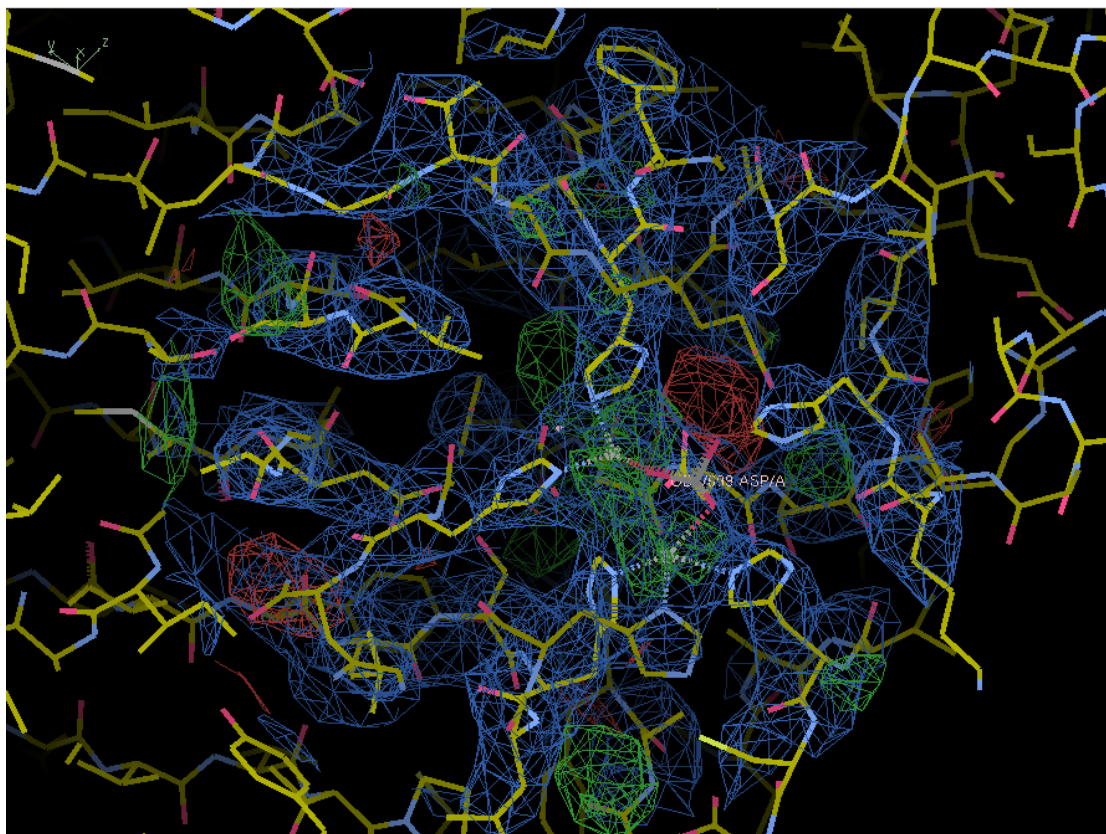


Figure I-8 Constructed electron density map with Trz1 model ($\sigma = 2.0$). Carbon atoms are colored yellow, nitrogen's blue, oxygen's red. $2Fo-Fc$ density are in dark blue, $Fo-Fc$ positive density are in green and negative density red.

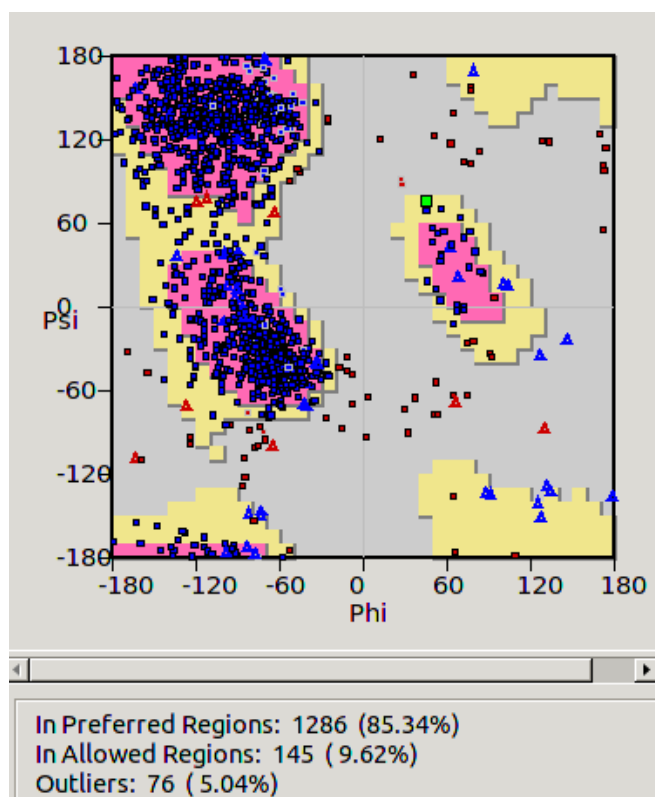


Figure I-9 Ramachandran diagrams of the refined Trz1 structure at 3.39 Å. One square represent one residue. The blue ones indicating they're in preferred regions, and the red are in outlier. Glycines are represented as triangles.

1.2 Analysis of the crystal structure of Trz1

The current model includes 766 residues (covering regions 1-204, 246-623, 636-818; and an extra Phenylalanine at the N-terminus introduced from the affinity tag and enterokinase cleavage site), two Zn^{2+} ions and a phosphate ion. Most of the protein chain has well defined electron density, but some peptide regions are missing in the final structure. The regions contained between residues 205-245, 624-635 and 819-838 were not visible in the electronic density maps and absent in the final molecular model. The missing regions are not the same for the two copies in the asymmetric unit and their absence is likely due to their mobility in the crystal.

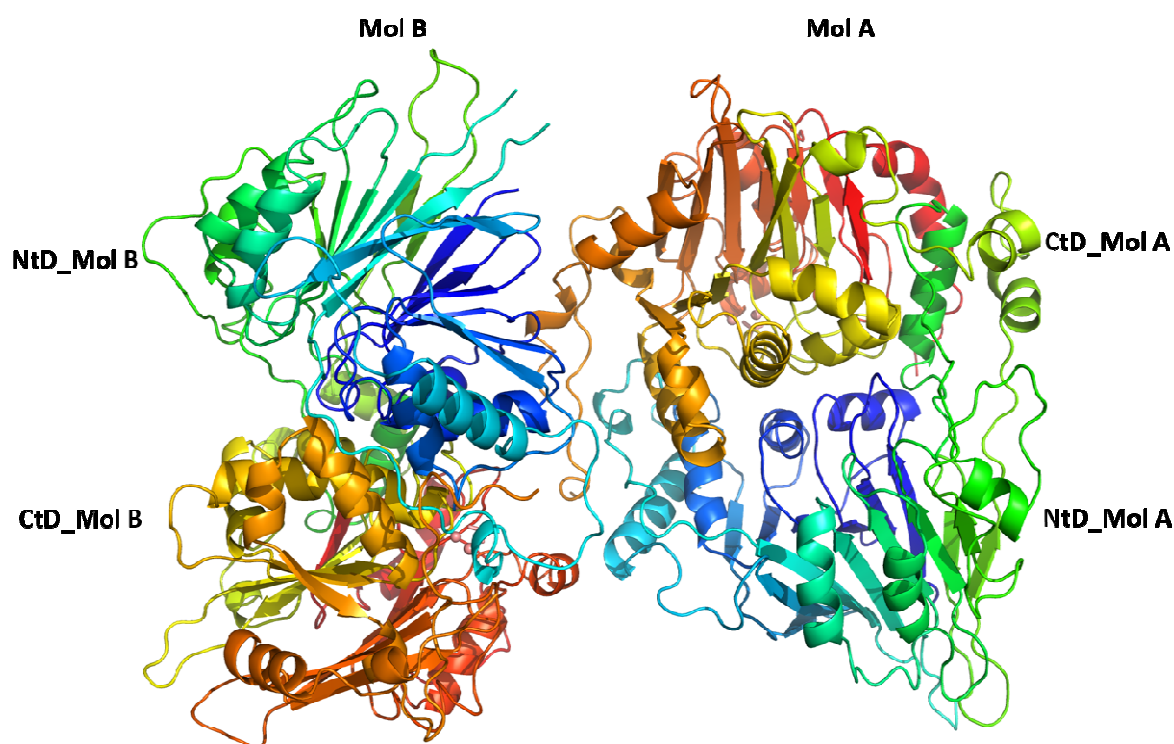


Figure I-10 Cartoon representation of the two copies of Trz1 in the asymmetric unit.

The two copies are labelled as “Mol A”/”Mol B”, and coloring scheme changes from blue (N-terminus) to red (C-terminus) for each copy. The N- and C-terminal domains are labeled as NtD, CtD, respectively.

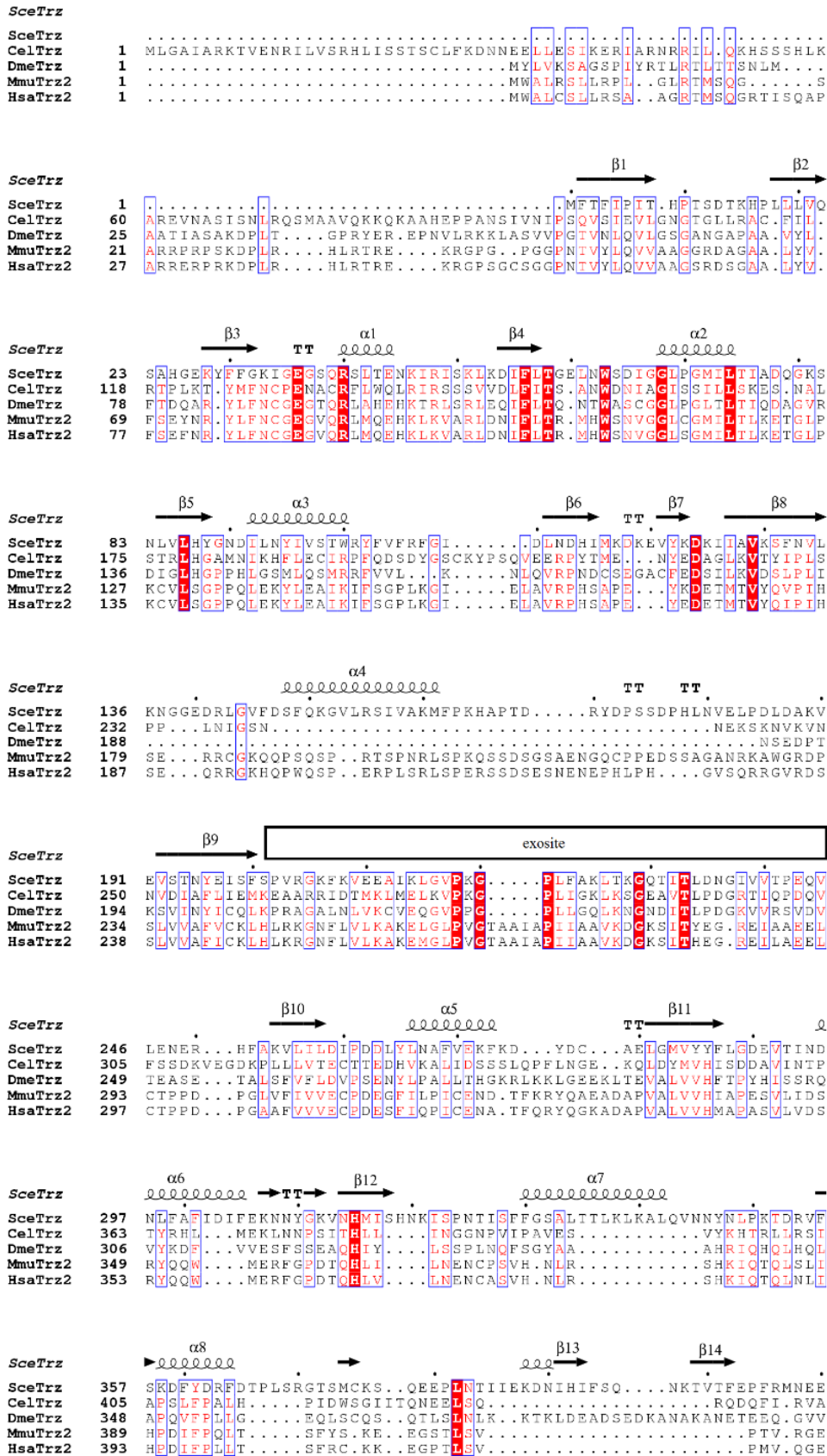


Figure I-11 Sequence alignment of the long form of RNase Z.

Secondary structure elements as extracted from the Trz1 crystal structure are superposed onto the sequences. RNase Z^L from *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Mmu musculus* (Mouse) and *Homo sapiens* are represented.

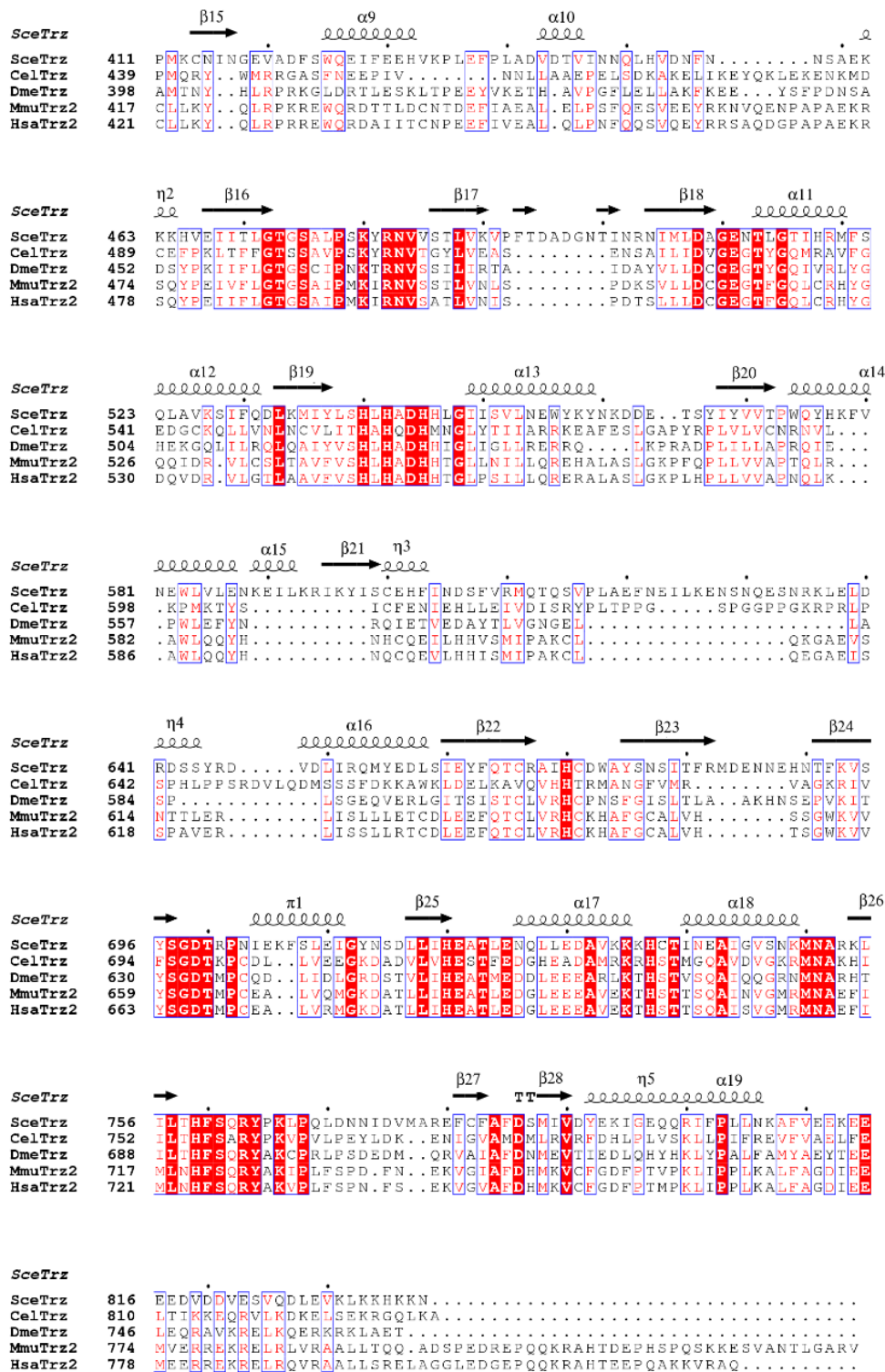


Figure I-12 (continue) Sequence alignment of the long form of RNase Z.

I.2.1 Overall structure

The crystals contain two copies of Trz1 per asymmetric unit, which adopt a head to tail orientation (Figure I-11). Superimposition of the two copies gives an r.m.s.d. of 0.522 Å, meaning that they are almost identical. Non-crystallographic symmetry averaging (NCS) was applied at the early cycles of the refinement (option “NCS” in the program *Refmac5*) and abandoned in later cycles in order to allow for structural differences between the two copies. As shown by gel filtration and SAXS (later), Trz1 forms a monomer in solution and the non-crystallographic dimer can therefore be considered as a crystallization artifact. There was no residual electron density for dGMP, a compound that was added in the crystallization liquor, indicating this ligand does not bind under the crystallization conditions.

Trz1 consists of two main domains connected by an extensive linker. The N-terminal half is contained between residues 1 and 404 and the C-terminal half between residues 466 and 838. The start of the connecting linker between the two halves of Trz1 (residues 405 to 465) forms an elongated peptide stretch between residues 405 and 423 including a β -strand (residues 414 to 417) that lies against the N-terminal domain, followed by an α -helical bundle (residues 425 to 451) and another irregular peptide stretch that interacts with the C-terminal domain (Figure I-13).

Both the N-terminal and C-terminal halves of Trz1 possess a β -lactamase core domain, made up of two opposing seven (or six)-stranded β -sheets covered by α -helices on each side, a so called “sandwich” fold. The crystal structure of the short form RNase Z showed that its β -lactamase core is supplemented by a flexible arm that is involved in the binding of the tRNA substrate (Li de la Sierra-Gallay et al. 2005). Sequence alignment shows that for the long form RNase Z, the flexible arm is only present in the N-terminal domain and not in the C-terminal domain (Figure I-12). The two domains of Trz1 are interacting considerably, and their positions mimic the arrangement of the two subunits of the dimeric short RNase Z form. Each β -lactamase core of Trz1 has several protruding regions, which are not existing in the short form RNase Z and contribute to the interaction between the two domains.

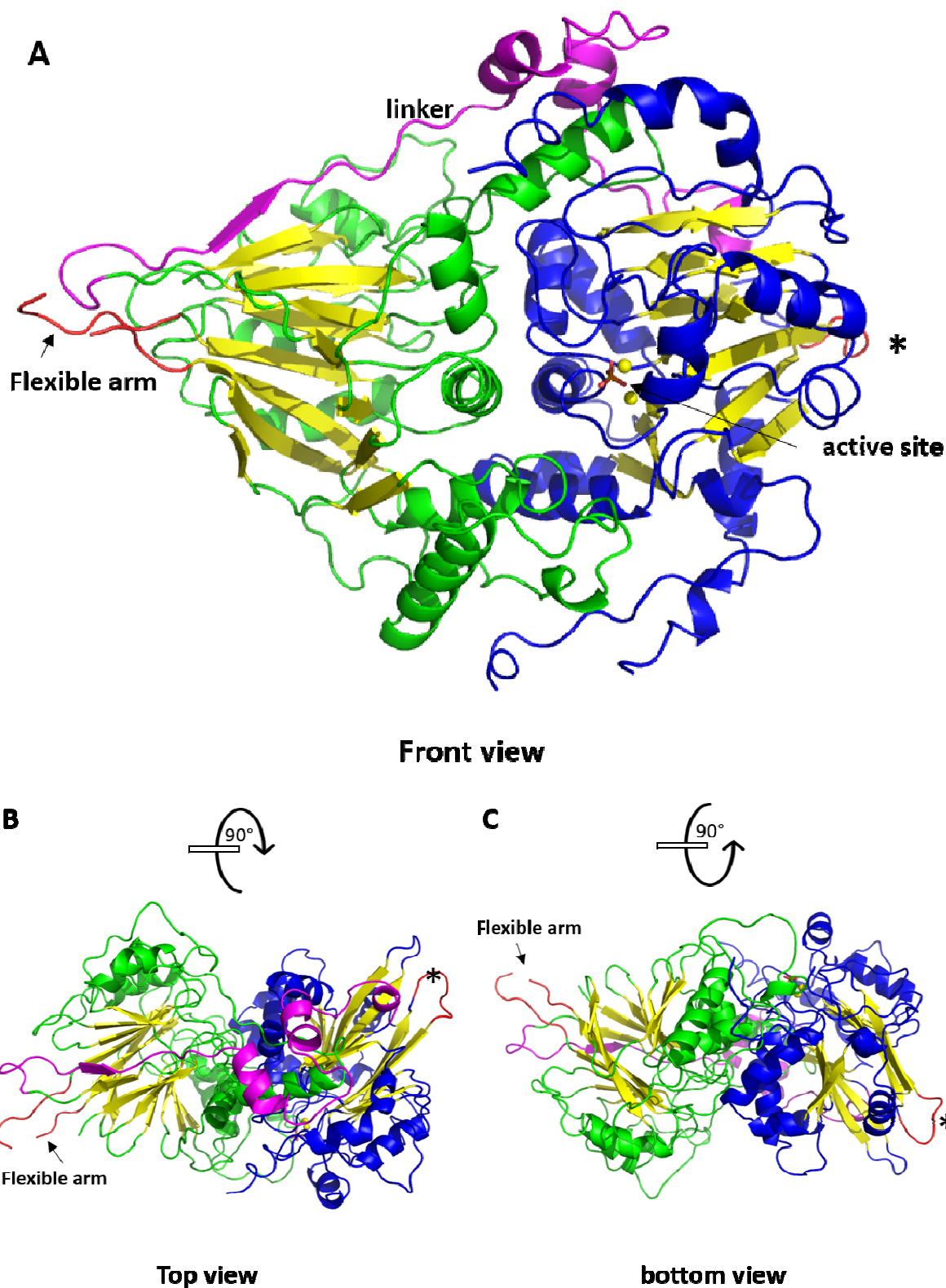


Figure I-12 Cartoon presentation of the yeast Trz1 structure, representing a long form RNase Z. The N-terminal domain is shown in green, the C-terminal domain is shown in blue, β -sheets of the both β -lactamase domains are shown in yellow. The linker is shown in purple. Two zinc ions are shown in yellow spheres; and the phosphate ion is shown in sticks (marking also the active site). The disordered flexible arm protruding from in the N-terminal domain is in red and indicated with an arrowhead, and the loop replacing the flexible arm in the C-terminal domain is labelled with an asterisk. (A) Front view; (B) Top view; (C) bottom view.

I.2.2 The N-terminal domain.

The primary sequences of the N-terminal domains of RNase Z^L proteins are much less conserved than that of the C-terminal domain compared to the short RNase Z^S forms. It was therefore important to confirm the previously formulated hypothesis, based on moderate sequence analogy, claiming that their N-terminal domains would also belong to the β -lactamase fold family (Redko et al. 2007). Superposition of the N-terminal domain of Trz1 with the BsuTrz structure (PDB code: 1Y44) yielded a r.m.s.d. of 2.5 Å for 210 aligned C α positions and 11.9% sequence identity. Our present structure of Trz1 therefore confirms this hypothesis.

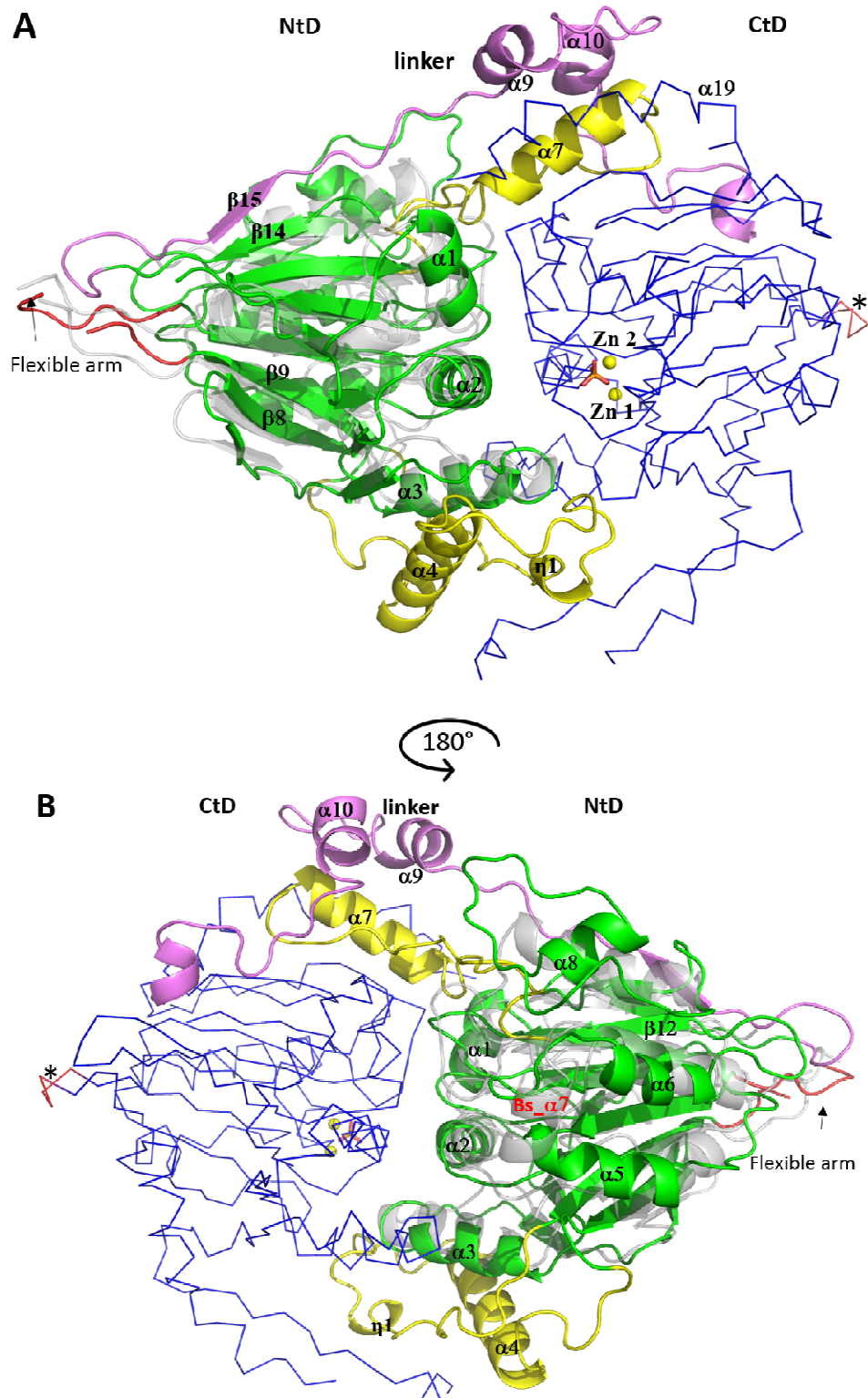
The β -lactamase core of N-terminal domain of Trz1 consists of two opposing seven-stranded β -sheets on each side covered by three α -helices forming a $\beta\beta\beta\alpha\beta\alpha\beta$ motif, assembling as a “sandwich”. In each sheet the first four β -strands (β 14, β 1, β 2, β 3; β 7, β 8, β 9, β 10) are antiparallel, and the last three (β 4, β 5, β 6; β 11, β 12, β 13) are parallel and connected by the α -helices. The linker between the two domains is preceded by a β -strand (β 15) that is antiparallel to β 14 of the β -lactamase core of the N-terminal domain (Figure I-14). The flexible arm is inserted between the third and fourth β -strand of the “outside” sheet, at the same position as the arm in RNase Z^S structures. However, the second substrate recognition element, the “ α 7” helix of BsuTrz, is absent in the N-terminal domain (Figure I-14 B).

The N-terminal domain of Trz1 further contains two long extensions emanating from the β -lactamase core that are not present in the RNase Z^S structures (Figure I-14). These extensions reach out towards and interact extensively with the C-terminal domain. The first extension between strands β 8 and β 9 (Ex-I) comprises residues 136 to 191. This region contains a long helix (α 4, residues 147-162) and a short helical turn (η 1) while the rest of the connection lacks regular secondary structure. Interestingly, multiple sequence alignment of RNase Z^L homologs shows this extension is present in human and mouse but its sequence is poorly conserved. The length of the extensions is by much shorter in RNase Z^L sequences of *C. elegans* and fruit fly (Figure I-14 C).

The second-long connection of the N-terminal domain is contained between β 12 and β 13 (Ex-II, residues 319 to 390). This connection contains a helix (residues 330-345; α 7) that lies against the core of the C-terminal β -lactamase domain. It is covered by the helices of the inter-domain linker region (α 9 and α 10) and by the C-terminal helices. Together these helices form a bundle like subdomain. The remainder of extension Ex-II (residues 353-390) mainly

interacts with the N-terminal β -lactamase core. The Ex-II extension contributes considerably to the interaction surface between the two domains.

We conclude that despite very low sequence similarity with the short form RNase Z, the N-terminal domain of Trz1 adopts a β -lactamase fold and is decorated by long extensions that contribute significantly to the interaction with the C-terminal domain.



C

Extension-I

$\alpha 4$

<i>SceTrz</i>			o o o o o o o o o o o o o o o o			TT	TT	
SceTrz	136	KNGGEDRL	G	VFDSFQKGVLRSIVAKMFPKHAPTD.....RYDPSSDPHLNVELPDLDKAV				190
CelTrz	232	PP...LNI	G	SN.....NEKSKNVKVN				249
DmeTrz	188NSEDPT				193
MmuTrz2	179	SE...RRC	G	KQQPSQSP..RTSPNRLSPKQSSDSGSAENGQCPEPSSAGANRKAWRDP				233
HsaTrz2	187	SE...QRR	G	KHQPWQSP..ERPLSRSLSPERSSDSESNENEPHLPH....GVSQRRGVRDS				237

Figure I-13 The N-terminal domain of Trz1 compared to BsuTrz^S.

The β -lactamase core of the N-terminal domain is shown in green and the extensions are in yellow; the linker of N- and C-terminal domain is shown in purple. The disordered flexible arm in the N-terminal domain is in red and indicated with an arrowhead, and the loop replacing the flexible arm in the C-terminal domain is labelled with an asterisk. The C-terminal domain is shown in blue ribbon for simplicity; and two zinc ions (yellow spheres) and a phosphate ion (orange ribbon) coordinating with the C-terminal domain are labelled. BsuTrz^S is shown in transparent gray (PDB code 1y44, chain A) (A) front view. (B) back view (rotate 180° along axes y). The “ $\alpha 7$ ” from BsuTrz^S is labeled in red.

I.2.3 The C-terminal domain

The sequence of the C-terminal half of Trz1 is better conserved than that of the N-terminal half compared to RNase Z^S (Figure I-12 (continue)). The 3-D structure of the C-terminal domain has the predicted β -lactamase fold and superposes well onto the structure of BsuTrz (PDB code 1Y44, r.m.s.d. = 2.1 Å for 227 aligned C α positions and 25% sequence identity, Q score = 0.41) (Figure I-15), and to the human ELAC1 (PDB code 3ZWF, r.m.s.d. = 2.1 Å for 218 aligned C α positions and 25% sequence identity, Q score = 0.39). The histidine signature motif (motif II) together with residues in other motifs (motifs III to V and HEAT, HST motifs) that characterize the active center in short form RNase Z are all situated in the C-terminal domain. The crystal structure revealed also the presence of two zinc ions and a phosphate ion as was the case for the crystal structure of short RNase Z from *B. subtilis* and *H. sapiens* (PDB code 2FK6 and 3ZWF, respectively). The C-terminal domain of Trz1 therefore carries the catalytic activity.

I.2.3.1 The β -lactamase core

The β -lactamase core of the C-terminal is also the sandwich fold: two opposing sides of β -sheets covering by big helices in each side (Figure I-15). The inner sheet has seven β -strands as that of BsuTrz, with the first four ($\beta 28$, $\beta 16$, $\beta 17$, $\beta 18$) antiparallel, and the last three parallel ($\beta 19$, $\beta 20$, $\beta 21$). The outer sheet has only six β -strands instead of seven, where the first β -strand is lost in Trz1, leaving only three antiparallel β -strands ($\beta 22$, $\beta 23$, $\beta 24$) (labelling “the lost β -strand in Trz1” in Figure I-14). The position of the lost β -strand is taken by $\alpha 16$ from the long extension (Ext-III) which is not present in RNase Z^S. This long

extension is about sixty residues long in Trz1 (residues 600 to 659), including one big α -helix (α 16) and two small η -helical segments (η 3, η 4). This extension will be further discussed in chapter I.2.3.4.

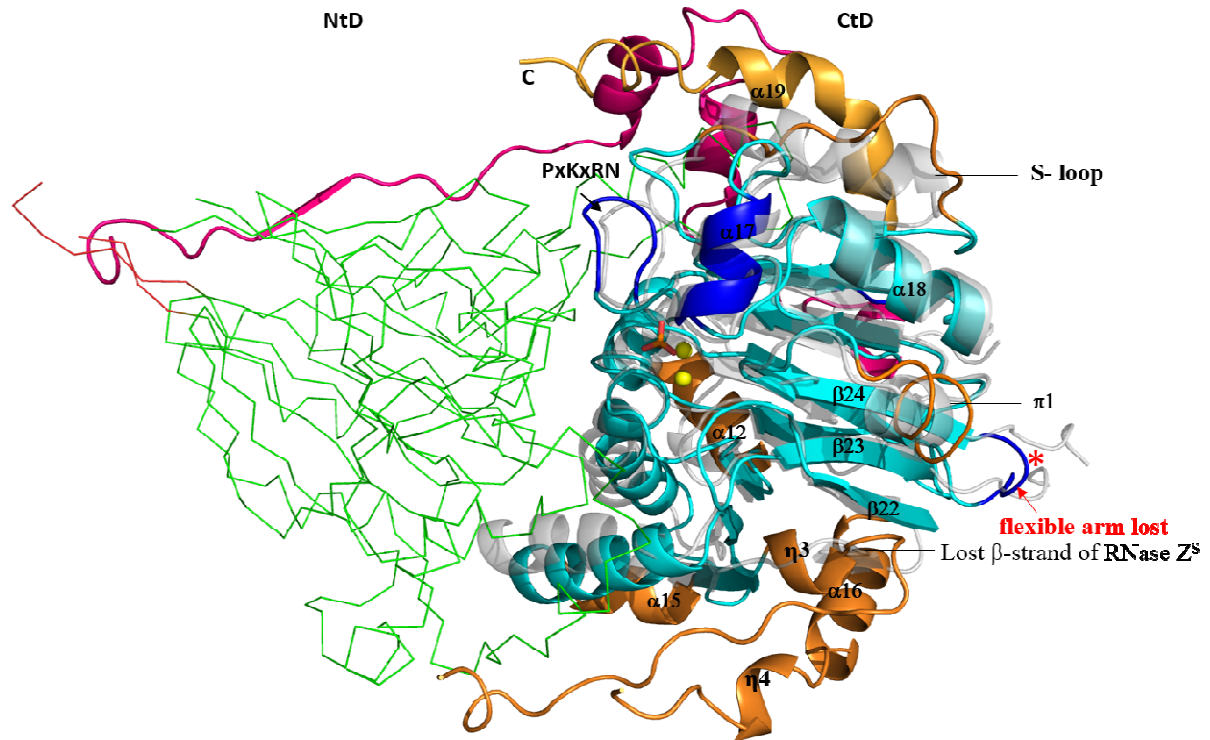


Figure I-14 The C-terminal domain of Trz1 superposed onto BsuTrz^S.

The β -lactamase core of the C-terminal domain is shown in cyan and the extensions are in orange. The loop replacing the flexible arm in BsuTrz^S is labelled with asterisk. The conserved α 17 and “PxKxRN” loop are colored in blue. The N-terminal domain is shown as a ribbon; and the linker region is shown in purple. BsuTrz^S is shown in transparent gray (PDB 1y44, chain A).

The difference the β -lactamase core between that of C-terminal domain and RNase Z^S lies in not only the β -sheets, but also the α -helices covering the sheets. For the α -helices covering the inner β -sheets, Trz1 has an α -helix (α 12, residue 522-532) before the β -strand β 19, which is substituted in BsuTrz by a short three residues η -helical turn. In addition, Trz1 has an exclusive short α -helix (α 15, residues 588-593) perpendicular to strand β 21 (Figure I-15). For the α -helices covering the outer β -sheets, instead of three big α -helices in BsuTrz, only one equivalent α -helix is found in Trz1 (α 18). The other two are substituted to a π -helix (π 1) and a S-shaped loop (labelled as S-loop in Figure I-15), respectively.

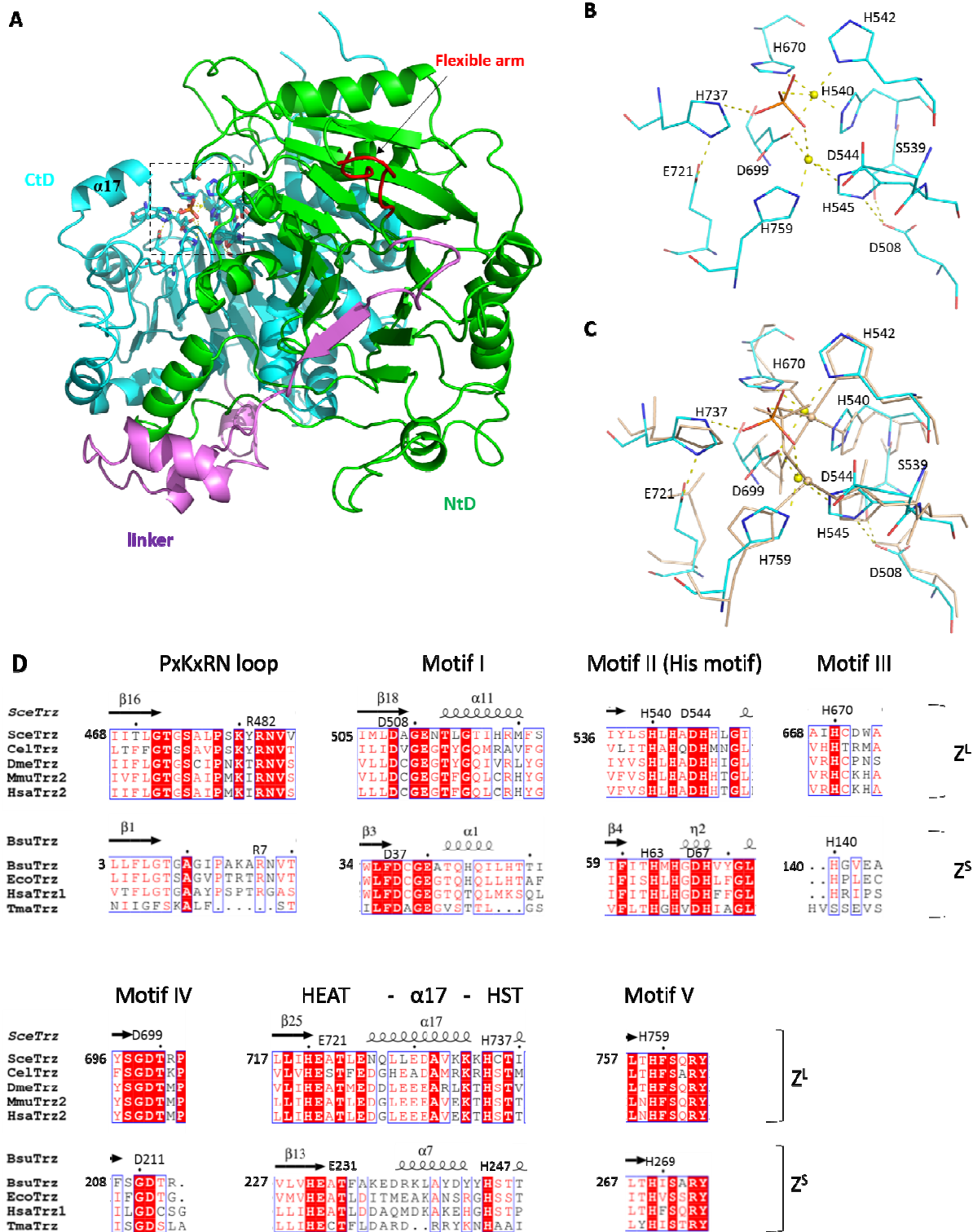


Figure I-15 Active site of Trz1.

(A) The active site is found in the C-terminal domain of Trz1. The N-terminal halve (NtD, green), C-terminal halve (CtD, cyan) and the linker (purple) of Trz1 are labeled; and the position of the disordered flexible arm is shown in red. The active center is marked with dashed line square. (B) Zoom on the active site. The active site residues originate from various motifs and coordinate two zinc (yellow sphere) and a phosphate (orange stick) ion. (C) Superposition of the active sites from Trz1 and BsuTrz (gray). For simplicity only Trz1 residues are labeled. (D) Sequence alignment of the conserved motifs of long/short RNase Z found in the active site, shown in (A) to (C).

I.2.3.2 The active center

Inspection of the residual electron density in the active site region of the Trz1 C-terminal domain reveals the presence of two Zn ions at a distance of 3.4 Å that are bridged by a bound phosphate ion (Figure I-16). The two Zn ions are further coordinated by five histidines (His540, His542, His545, His670 and His759) and two aspartates (Asp544, Asp699), one of which (Asp699) bridges both Zn ions. His545 is also hydrogen bonding to Asp508 in the motif I, which also itself is bound to Ser539 from the Histidine motif II. Mutations in Asp508 inactivate BsuTrz and DmeTrzL (Li de la Sierra-Gallay et al. 2005; Zareen 2006), indicating its key role in maintaining the active conformation of the catalytic site.

The phosphate ion is also bound by His737 from the “HST” (“HCT”) motif, and His737 is interacting with Glu721 in the “HEAT” motif. Mutation of His737 or Glu721 greatly reduce catalysis activity in the long form of RNase Z from *D. melanogaster* (Karkashon et al. 2007) and RNase Z^S (Späth et al. 2005; Minagawa et al. 2006; Vogel et al. 2004).

These conserved residues and their coordination with Zn and phosphate ion clearly delineate the active centre in the C-terminal domain. The coordination sphere of the two Zn ions in the Trz1 C-terminal domain is perfectly superposed to that of BsuTrz (Figure I-16 C).

I.2.3.3 Conservation motifs involved in tRNA binding

In the structure of BsuTrz in complex with tRNA^{Thr}, the BsuTrz dimer binds two tRNAs (PDB code 2FK6, Li de la Sierra-Gallay et al. 2006). The tRNA is clamped by the flexible arm of one subunit, and further fixed by the $\alpha 7$ helix, linking the “HEAT” and “HST” motifs, from the other subunit. The cleavable 3'-end of the tRNA substrate is presented to the active site of the opposing subunit. It was proposed that the key to catalytic site activation is the stabilization and orientation of helix $\alpha 7$ and its flanking regions (Li de la Sierra-Gallay et al. 2006). The $\alpha 7$ helix in the inactive subunit of BsuTrz apo-dimer was disordered in one subunit (PDB code 1y44, chain A) but had well defined electron density both in the opposing subunit (PDB code 1y44, chain B) and in the BsuTrz/tRNA complex (PDB code 2FK6) (Li de la Sierra-Gallay et al. 2005; Li de la Sierra-Gallay et al. 2006).

The “PxKxRN” loop (residues 478 to 483 in Trz1) is another element substantial for activity and substrate binding (Zareen 2006). Study on RNase Z^L from *D. melanogaster* (DmeTrz^L Δ 23, mitochondrial signal peptide truncated) showed the PxKxRN loop affects both the catalytic efficiency and substrate binding. For example, substitution of Arg448 of DmeTrz^L Δ 23 (equivalent with Arg482 of Trz1) in the loop leads to 1000 fold reduced

catalytic activity; similar effect was observed in substitution of Gly438 (equivalent Gly472 of Trz1). And Lys448 and Asn449 (DmeTrz^LΔ23 numbering) affect both substrate catalytic and binding observed by changes in both k_{cat} and k_M (Zareen 2006).

The superposition of the Trz1 C-terminal domain onto the subunit of BsuTrz which is in the active conformation shows that the $\alpha 17$ helix (residues 729 to 735, equivalent to $\alpha 7$ of BsuTrz) and the “PxKxRN” loop are well superimposed (Figure I-15). However, the flexible arm that binds the tRNA substrate in BsuTrz is clearly absent in the C-terminal domain of Trz1 and is replaced by a short loop connecting the two strands $\beta 23$ and $\beta 24$.

I.2.3.4 The extensions in the C-terminal domain

The C-terminal domain of Trz1 contains a region protruding from the β -lactamase core of sixty residues (residues 600-659, Ex-III). The region contains irregular peptide stretches, one big α -helix and two small η -helical segments, situated close to the protruding region from the N-terminal domain. This domain mainly interacts with the C-terminal domain core. The lacking electron density of residues 625 to 635 indicates that this region is partly mobile. The length and sequence of this protruding region seem to vary considerably among the eukaryotic RNase Z^L sequences (Figure I-17).

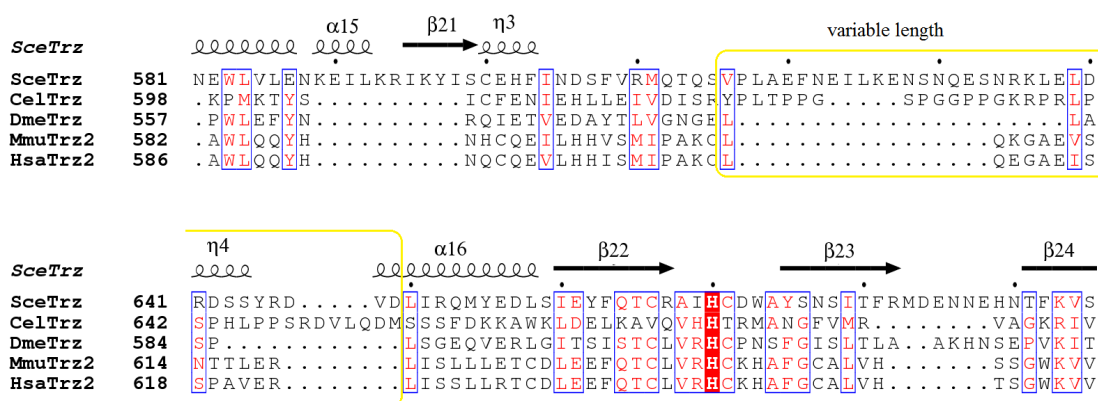


Figure I-16 The long extension (Ex-III) in the C-terminal domain of RNase Z^L.

The length and sequence of the Extension -III (residues 600 to 659 in Trz1 numbering) varies considerably among eukaryotic RNase Z^L homologs. The variable regions (about residues 625 to 647 of Trz1) are surrounded by a yellow square.

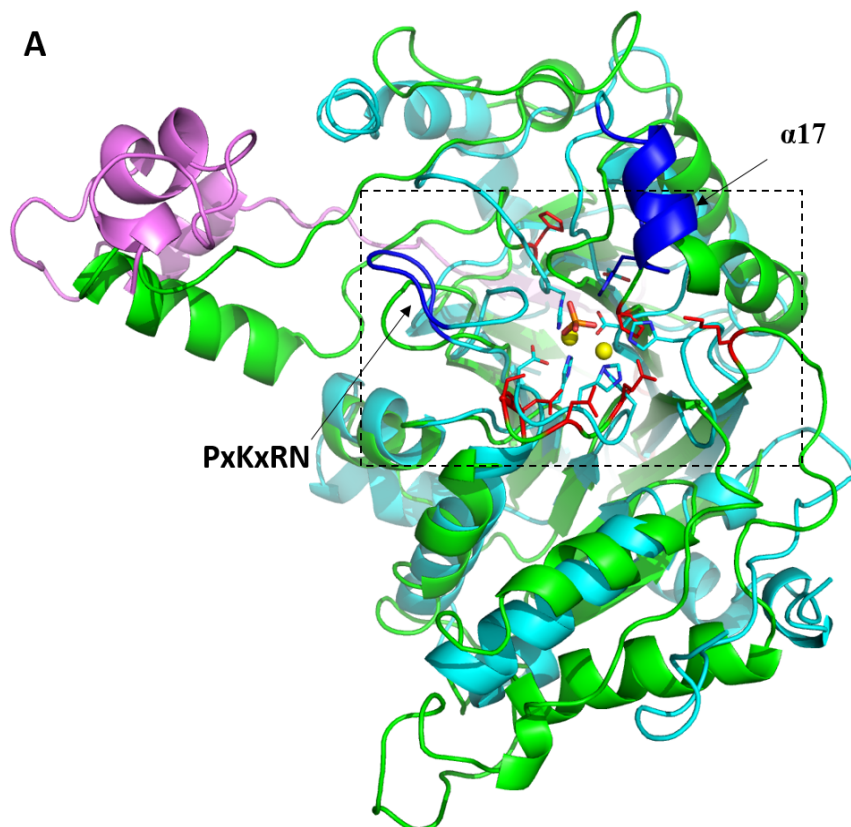
I.2.3.5 The “tail” of RNase Z^L

The structure of Trz1 has a long helical tail at its C-terminus (residues 792 to 818) compared to the RNase Z^S forms. These helices are in contact with helices from the extension of the N terminal domain ($\alpha 7$) and the linker ($\alpha 9$, $\alpha 10$). Together these helices and loops are

forming a small helical bundle and increase the interaction surface between the two β -lactamase domains. The residues beyond 818 are disordered in the crystal. All eukaryotic RNase Z^L sequences have this C-terminal tail and in human and mouse the tail is much longer (Figure I-12 (continue)).

I.2.4 Comparison of the N- and C-terminal domains

As stated before, the C-terminal domain of Trz1 shares higher sequence similarity with the RNase Z^S homologs than the N-terminal domain. We can now establish a structural comparison between the N- and C-terminal halves of long form RNase Z enzymes (Figure I-18). Structural superposition of the N- and C-terminal domain of Trz1 yielded an r.m.s.d. of 3.3 Å for 228 aligned C α positions and 13% sequence identity (Q score = 0.174). The sequence extensions of both domains compared to the short form RNase Z are concentrated in the linker and in extra domains whose functional role, if any, remains to be established.



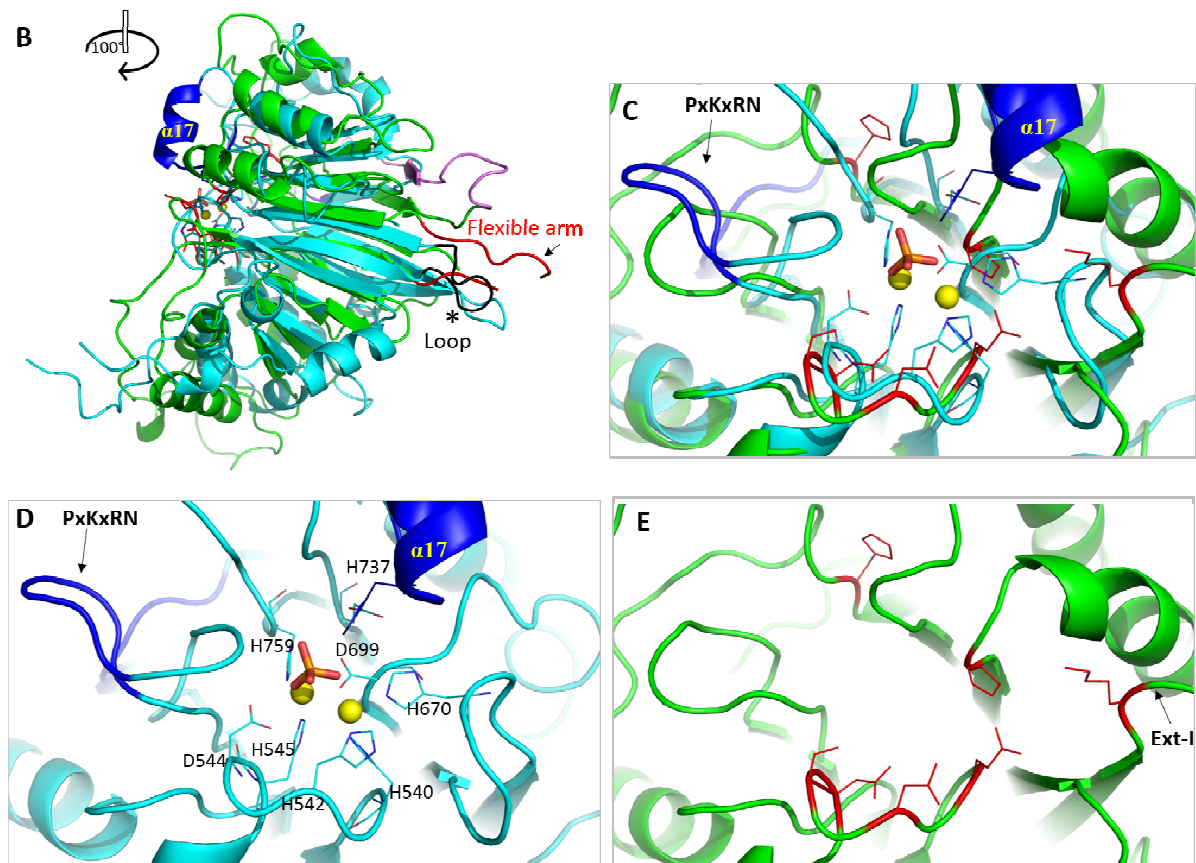


Figure I-17 Comparison between the N- and C-terminal domains of Trz1.

(A) Superposition of the N- and C-terminal halves of Trz1. The N-terminal domain (green), C-terminal domain (cyan) and the linker (purple) are labelled. The conserved tRNA recognition elements $\alpha 17$ and PxKxRN loop from the C-terminal domain are shown in blue. The residues involved in the active center from the C-terminal domain (blue) and the superimposed region from the N-terminal domain (red) are shown in stick. Zinc ions are shown in yellow spheres and the phosphate ion in stick. (B) Rotate view of (A) by y axis over 100° . The flexible arm from the N-terminal domain is shown in red; and the loop at the corresponding position from the C-terminal domain is in black (labelled with asterisk). (C) Zoom in the dashed line square of (A). (D) Representation of the C-terminal domain in the dashed line square of (A). (E) Representation of the N-terminal domain in the dashed line square of (A). (F) Structure based sequence alignments of the conserved motifs in C- and N-terminal domains.

Even though the N- and C-terminal domains of Trz1 have a common core structure, their functional roles are clearly distinct. The sequence and structure of the N-terminal domain of Trz1 lack the various motifs (II to V) that form the active centre network in the RNase ZS structures. Only the flanking sequences, the so-called ψ -motifs, are present in the N-terminal domain. Comparison of the sequences and overlay of the structure of N-terminal and C-terminal domain clearly shows the essential residues are lost in the N-terminal domain (Figure I-18).

In the ψ -motif II, the central HxHxDH signature motif is lost but the backbone conformation has been maintained (Figure I-18 B). The same holds for motif I: Asp508 in the

motif I which helps to stabilize His545 in the Histidine motif II is lost in the N-terminal sequence, but the backbone of this region is well superposed. Motifs III to V do not have equivalent ψ -motifs in the N-terminal domain, because the presence of the extensions I (residues 136-191) and II (residues 319-390) completely change the conformation of the corresponding regions. In the case of motif IV loop, where Asp699 locates, no conserved motif is found in this region. Interesting that following the motif IV loop (N-terminal domain) there's a major α -helix in the C-terminal; whereas the corresponding non- conserve loop in the C-terminal domain) is followed by a π -helix.

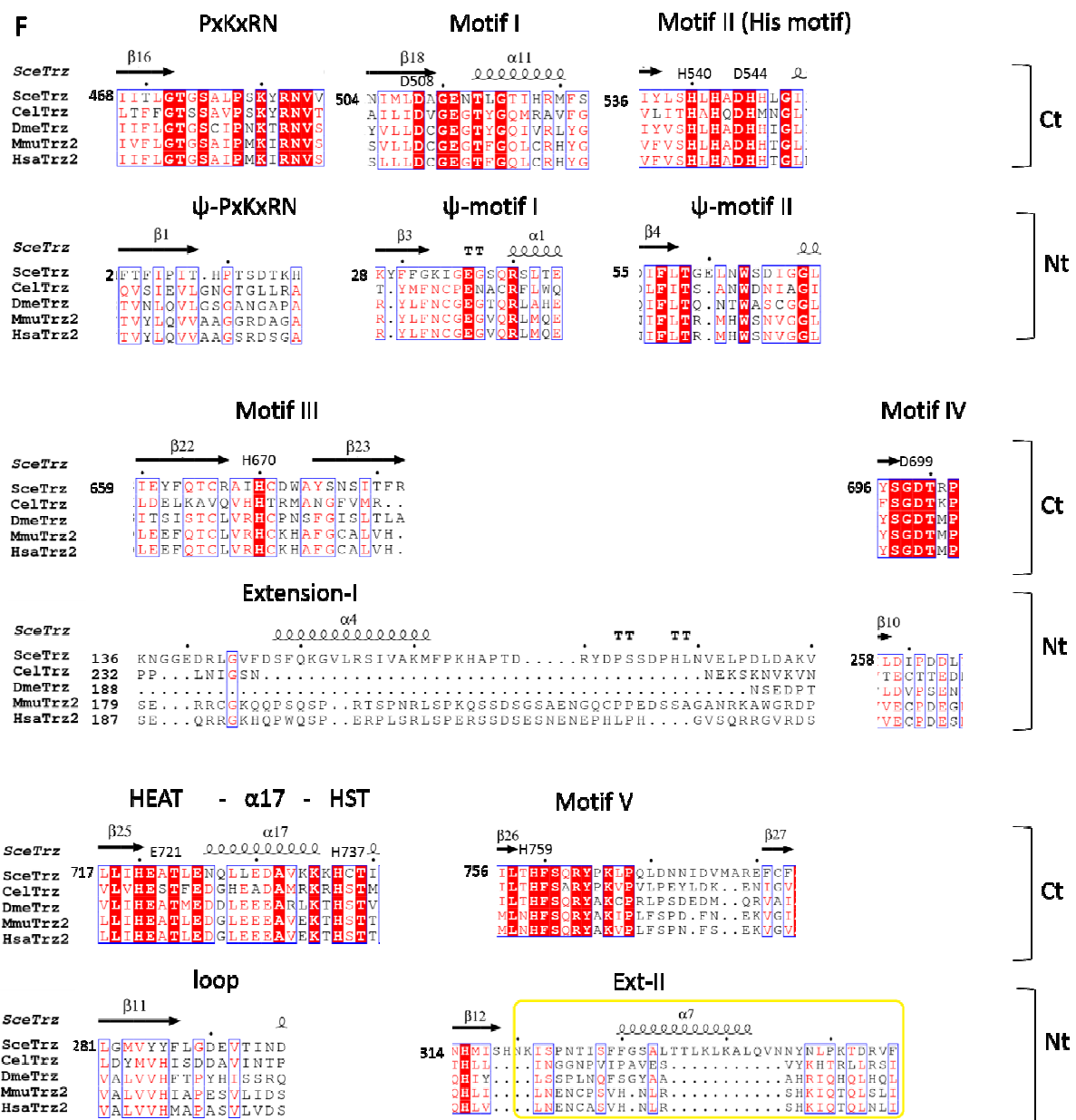


Figure I-18 (continue) Comparison between the N- and C-terminal domains of Trz1.

Table I-6 Comparison of the secondary structures of N- and C-terminal domains of Trz1
by superimposition

N-terminal domain					v.s.	C-terminal domain				
notation	SS*	length	start	end		notation	SS*	length	start	end
	β1	7	PHE 2	THR 8	<->		β16	6	GLU 467	GLY 472
	β2	5	LEU 18	GLN 22	<->		β17	5	SER 486	LYS 490
	β3	5	LYS 28	GLY 32	<->		β18	6	ASN 504	ALA 509
	α1	7	GLN 39	ASN 45	<->		α11	10	ASN 512	PHE 521
					x		α12 ⁽²⁾	11	SER 522	ASP 532
	β4	4	ASP 54	LEU 57	<->		β19	5	LEU 533	TYR 537
	α2	9	GLY 67	THR 75	<->		α13	13	GLY 548	ASN 560
	β5	5	LEU 84	TYR 88	<->		β20	5	ILE 568	THR 572
	α3	11	ASP 91	ARG 101	<->		α14	16	PRO 573	ASN 588
					x		α15 ⁽³⁾	6	ASN 588	LYS 593
	β6	5	LEU 111	ILE 115	<->		β21	5	ILE 595	SER 599
					x	Ex-III (resi. 600-659)	η3		CYS 600	PHE 603
							α16		VAL 648	SER 659
							η4		ARG 641	SER 644
	β7	3	VAL 121	LYS123	x					
	β8	9	ILE 127	LEU 135	<->		β22	8	ILE 660	ARG 667
Ex-I (resi.136-191)	α4	16	ASP 147	PHE 162	x					
	η1	3	PRO 177	LEU 179						
	β9	9	VAL 192	PHE 200	<->		β23	8	TYR 675	ARG 682
exosite			SER 201	ARG 250	x		loop, no exosite			
	β10	5	LYS 254	LEU 258	<->		β24	7	THR 691	SER 697
	α5	12	ASP263	LYS274	<->		π1	10	ASN 703	GLY 712
	β11	7	LEU 281	PHE 287	<->		β25	4	LEU 717	HIS 720
	Loop, no helix				x		α17	9	LEU 728	LYS 736
	α6	9	ASN 295	ASP 303	<->		α18	12	THR 739	MSE 750
	β12	5	ASN 314	SER 318	<->		β26	4	LYS 754	LEU 757
Ex-II (resi. 319 to 356)	α7	15	PHE 329	GLN 343	x					
	α8 ⁽¹⁾	8	LYS 358	ASP 365	x					
	β13	3	ILE 391	ILE 393	<->		β27	3	PHE 781	PHE 783
	β14	4	THR 399	PHE 402	<->		β28	3	MSE 788	VAL 790
Linker (resi.406 to 465)	β15		CYS 414	ASN 417		Tail (resi.792-end)	α19	16	Glu 793	PHE 809
	α9		SER 424	VAL 433						
	α10		ASP 442	ILE 447						
	η2		Glu 461	HIS 465						

*SS: secondary structure. The elements are colored the same as in the figure I-18. <->: Equivalent regions of the two domains by superimposition; x: The region could find its equivalence in the other domain.

(1) α8: shorter and different orientation with that of BsuTrz (PDB code: 1y44, residues 277-290, no special notation about the biological function), followed by the extension-II of Trz1.

(2) α12: a longer helix than that of BsuTrz; no equivalent helix but only a loop in the C-terminal domain of Trz1.

(3) α15: a short helix solely in the C-terminal domain of Trz1, replaced by a loop in N-terminal domain of Trz1 and BsuTrz.

The HEAT (Glu721) -α17- HST (His737) motif that is essential for tRNA binding and catalysis, only exists in the C-terminal domain and is replaced in the N-terminal domain by a loop (residues 278 to 295). Also the PxKxRN motif that affects tRNA binding (and also catalysis), is also absent in the N-terminal domain (Figure I-18 A and C). The corresponding loop in the N-terminal domain has a different conformation (Figure I-18 B).

Another important functional difference between the N- and C-terminal domains resides in the exosite, which is present in the N-terminal but not in the C-terminal half. Although the exosite is disordered in the Trz1 crystal structure, superposition of the Trz1 N-terminal

domain onto the RNase Z^S structures showed that it is emanating from exactly the same position on the surface in both RNase Z forms. This arm clamps the tRNA substrate in BsuTrz (Li de la Sierra-Gallay et al. 2006) and our structure shows that (only) the N-terminal half of Trz1 carries this function. Parallel to this, the N-terminal has lost its Zn binding capacity and the two halves of Trz1 are not exchangeable from a mechanistic point of view.

Table 1-7 shows the comparison of the secondary structure of the N- and C- terminal domain by superimposition. The equivalent elements of the two domains (i.e. $\beta 1$ and $\beta 16$) and the characteristic regions of solely domain (i.e. Ex-I from the N-terminal and Ex-III from the C-terminal domain) are clearly listed.

I.2.5 The interaction of N- and C- terminal domain

The N-terminal and C-terminal domains of Trz1 are interacting tightly in a way that is very similar as for the two subunits of the short form RNase Zs. The structure of Trz1 superposes well onto the dimeric RNase Z^S, as illustrated for BsuTrz (Figure I-19 A). The major interface of the two domains of Trz1 is contributed by three major α -helices, which are $\alpha 1/\alpha 11$, $\alpha 2/\alpha 13$ and $\alpha 3/\alpha 14$ from the N-/C-terminal domain, respectively (Figure I-19 B).

As mentioned before, several long extensions from the two domains of Trz1 contribute significantly to their interface. These regions together with the linker are presumably important for maintaining the two halves of Trz1 in the correct configuration (Figure I-19 C-E). In order to study the biochemical behavior of the two Trz1 halves we cloned and attempted to overexpress and purify them. If successful, these individual domains could help us in better defining the interactions between Trz1 and Nuc1/mutarotase. Unfortunately, our constructs either yielded very poor expression levels for the N-terminal half or had a strong tendency to aggregate for the C-terminal half. Monomeric species have not been reported in the literature for RNase Z^S forms and the protein might not be stable or soluble as a monomer. The much larger interaction surface between the two halves of Trz1 may render their individual expression even more precarious.

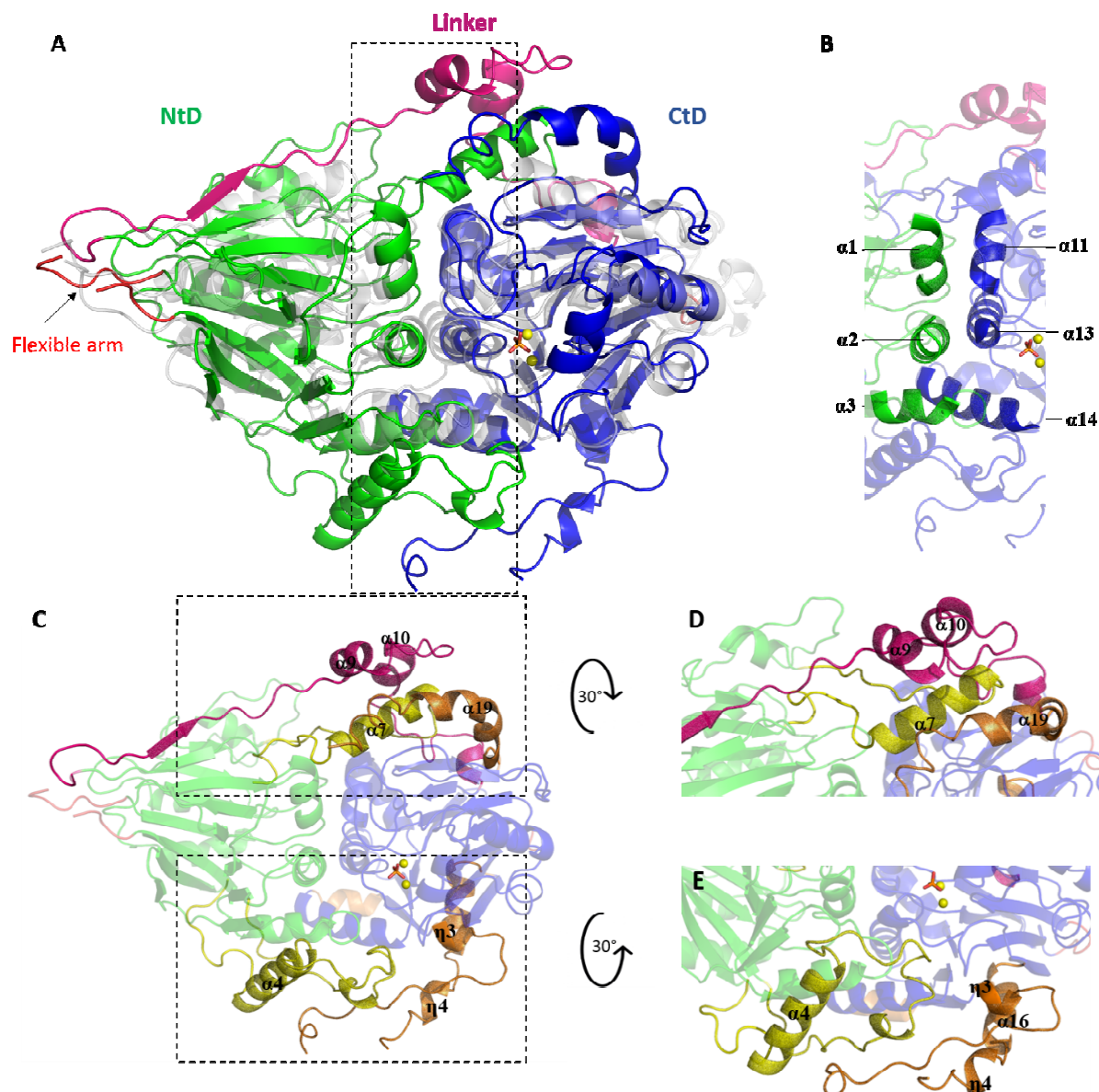


Figure I-18 Interface between the N- and C-terminal domains of Trz1.

(A) Trz1 superimposed onto the dimeric short RNase Z from *B. subtilis* (PDB 1y44, transparent gray). The N-terminal (green) and the C-terminal (blue) halves and the linker (purple) are shown. (B) The α -helices of the interface of the two domains are similar to those at the dimer interface of short RNase Z. (C) The linker and protruding regions of N- and C-terminal domain reinforce the interaction of the two domains. The protruding regions from N- (yellow) and C-terminal (orange) domains are shown; the linker is shown in purple. The rest of the structure is shown in 50% transparency. Zinc ions are shown in yellow spheres and the phosphate ion is shown in stick. (D) Higher dashed square of representation (C) rotated out of the screen over 30° . (E) Lower dashed square of representation (C) rotated into the screen 30° .

I.2.6 Conclusion

In conclusion, the structure of Trz1 consists of two domains with the same fold whose association mimics the structure of the dimeric short form RNase Z. The N-terminal domain has no Zn bound and the residues crucial for catalytic activity are absent. However, the N-

terminal domain has a protruding arm that is likely involved in clamping of the tRNA substrate. The C-terminal domain has a properly formed active site responsible for the cleavage reaction, but the protruding arm is replaced by a short loop, that is certainly too short to embrace the tRNA substrate. The tRNA substrates are likely bound and processed in a very similar in the long and short RNase Z forms (see discussion). As suggested by others, our structural data confirm that the long form RNase Z probably originated through gene duplication and divergent evolution of a single subunit.

II. Reconstitution of Trz1/Nuc1/mutarotase Complexes

As explained in the introduction of this thesis, Trz1, Nuc1 and mutarotase were reported to form a ternary complex (Hazbun et al. 2003), but some mixtures of pairs of these proteins also form binary complexes (Benschop et al. 2010). So we set out to overproduce the ternary and binary complexes with the aim to characterize their structure and biochemical properties.

II.1 Purification of the individual Trz1, Nuc1 and mutarotase proteins.

In order to overproduce the complexes, firstly we needed to produce individual recombinant Trz1, Nuc1 and mutarotase in good yields. Trz1 and mutarotase are fused with the His-tag in the construct at the N-end and C-end, respectively. So Nuc1 was designed to fuse with Strep-tag in order to facilitate studying the interaction with other two proteins (see details in chapter II.1.3). Nuc1 was in mitochondrial localization signal and the flanking sequence truncated form, and Trz1 and mutarotase were in full length, whose molecular weight (including the tag) are 33.8 kDa, 101 kDa and 34.8 kDa, respectively (Table II-1).

Table II-1 Molecular weights and various tags of the proteins.

Protein	Tag	Tag position	Mr*
Trz1	His-tag	N-terminal	101 kDa
Nuc1**	Strep-tag	N-terminal	33.8 kDa
Mutarotase	His-tag	C-terminal	34.8 kDa

* Mr includes the size of protein, tag, and cleavage sites. **Mr of Nuc1 shown here is the Δ N43 truncated form.

II.1.1 Purification of Trz1

Trz1 was produced the same way as for the crystallization described in chapter I.1.1.

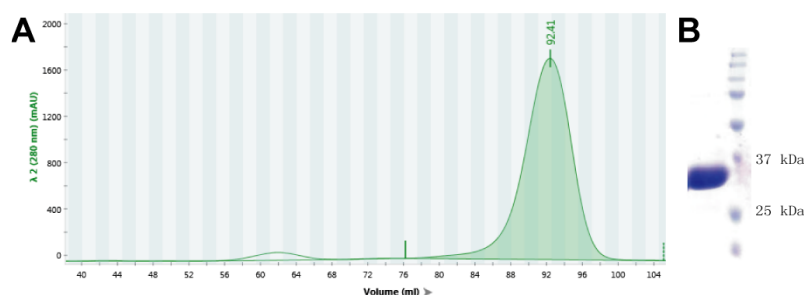


Figure II-1 purification of mutarotase.

(A) Size exclusion chromatography of pooled affinity chromatography fractions of mutarotase. Column: Superdex 200 Hiload 16/60. Buffer: 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM BME. (B) SDS-PAGE of purified mutarotase.

II.1.2 Purification of Mutarotase

His-tagged Mutarotase was expressed in a BL21-Gold *E. coli* strain and induced at 37 °C using a plasmid construct available in the lab (Graille et al. 2006). Mutarotase was highly expressed in soluble form and could be purified with high yields using Ni-NTA affinity and size exclusion chromatography (Figure II-1).

II.1.3 Expression and purification Nuc1

The NUC1 gene contains a N-terminal mitochondrial localization peptide (MLS). The MLS-truncated Nuc1 protein was reported to be poorly expressed in *E. coli* (Kieper et al. 2010). We cloned three different forms of NUC1, the full-length protein, the MLS-truncated construct (Δ N24, with the N-terminal 24 amino acids deleted), and the Δ N43 (with the N-terminal 43 amino acids deleted) into the pET-24a (+) vector. The NUC1 gene was fused either to a His-tag or to a Strep-tag encoding sequence and subsequently these constructs were transformed in BL21-Gold or BL21-CodonPlus cells. The induction was done either at 37 °C or at 15 °C, and the total and soluble fractions of the cells were analyzed by SDS-PAGE. Further detection by western blot showed only the Strep-NUC1 Δ N43 construct yielded expressed soluble protein using CodonPlus cells induced at 37 °C or 15 °C and Gold cells induced at 15 °C (Figure II-2). No expression was obtained for the full length protein and only trace expression amounts for the Δ N24 construct, indicating that the N-terminal mitochondrion targeting signal and its flanking region might corrupt protein expression in *E. coli*.

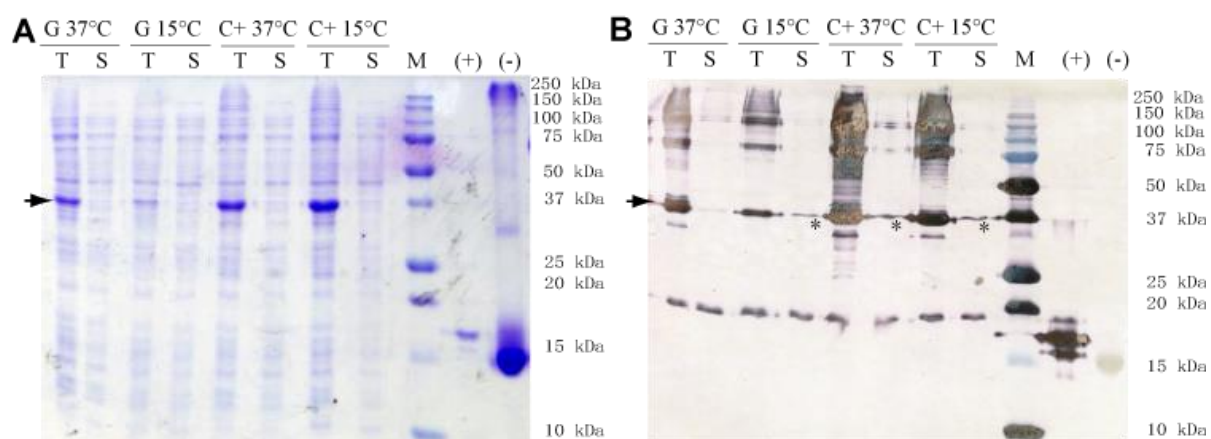


Figure II-2 Expression test of Strep-Nuc1 Δ 43 using different expression strains and temperatures.

G: BL21-Gold. C+: BL21-CodonPlus. Strep-Nuc1Δ43 bands are indicated by arrows. (A) SDS-PAGE of the fractions of cell lysate. T: Total fraction. S: Soluble fraction. (B) Western blot of the SDS-PAGE shown in A revealed with Strep-Tactin-HRP conjugate. The stars indicate soluble Strep-Nuc1Δ43. (+)/ (-) are other Strep-tagged/His-tagged proteins used as positive/negative control.

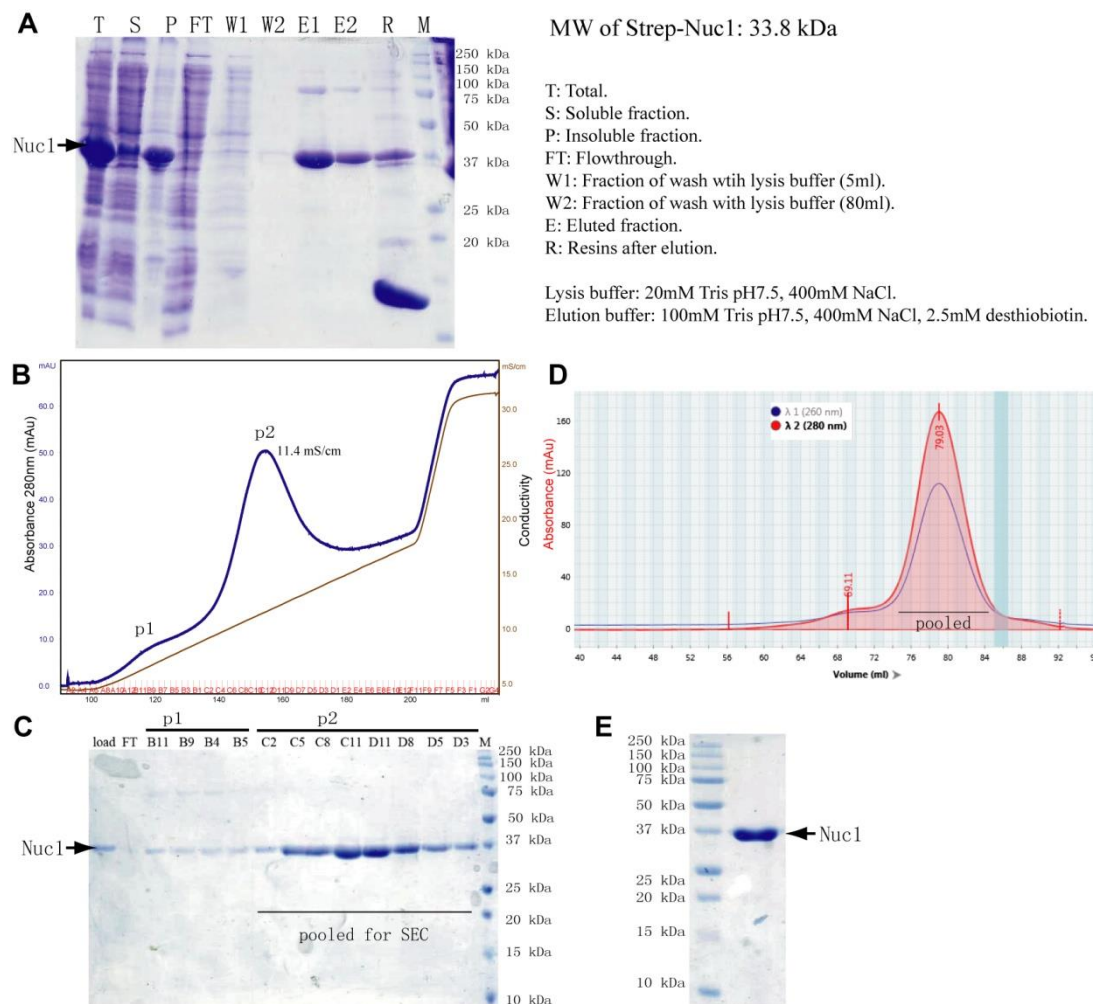


Figure II-3 Purification of Nuc1.

(A) SDS-PAGE of the fractions eluted from the Strep-Tactin affinity column. (B) Ion change chromatography of the fractions “E1” and “E2” from (A). Column: monoS 5/50GL. Buffer low/high salt: 20 mM Tris-HCl, pH 7.5, 50 mM /500 mM NaCl, 10% glycerol, 10 mM BME. (C) SDS-PAGE of the eluted fractions from (B). (D) Size exclusion chromatography(SEC) of pooled fraction (“C2” to “D3”, labelled in (C)) from in (B). Column: Superdex 200 Hiload 16/60. Buffer: 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM BME. (E) SDS-PAGE of purified Nuc1.

For high quantity Nuc1 protein production we exclusively used the Strep-Nuc1ΔN43 construct expressed in BL21-CodonPlus cells with induction at 15 °C. By accident we found that induction at an O.D.₆₀₀ of 0.9 provided the best yields for Nuc1 expression. Although Strep-Nuc1ΔN43 in BL21-CodonPlus cell was found to be the best expression condition we were still left with about 80% of Nuc1 in the insoluble fraction (Figure II-3A, lanes T and S). Increase of the NaCl up to 400 mM in the lysis buffer improved considerably the amount of

Nuc1 in the soluble fraction. We will from now on use the Nuc1 nomenclature for the Nuc1 Δ N43 protein construct.

Strep-tagged Nuc1 was purified using a Strep-Tactin affinity resin, followed by size exclusion chromatography (Figure II-3). SDS-PAGE of the gel filtration fractions showed that Strep-Tactin purified Strep-Nuc1 contained a protein contaminant of about 75 kDa, which is approximately the size of Nuc1 dimer and could therefore not be removed by size exclusion chromatography. Therefore, an extra Mono S ion-exchange chromatography step was introduced. The Nuc1 elution partially overlaps with the contaminant leading to a drop in the final protein purification yield (p1 in Figure II-3 B-C). We could yield Nuc1 about 0.35 mg from a 800 ml bacterial culture.

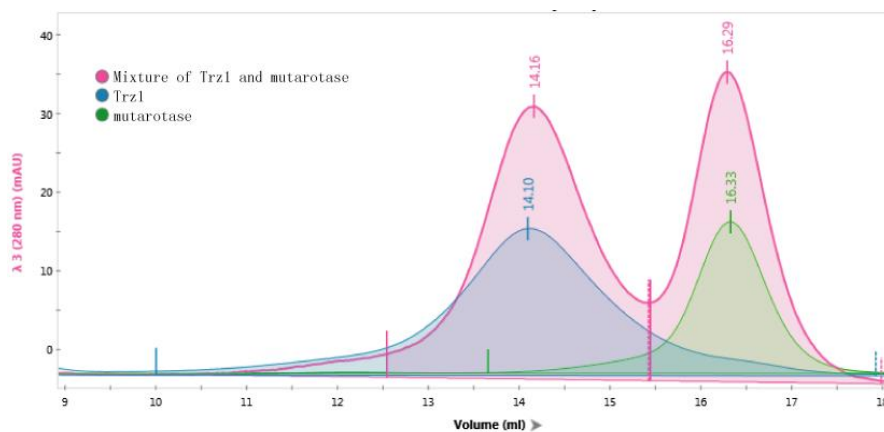


Figure II-4 Overlay of size exclusion chromatograms of a Trz1-mutarotase mixture and the individual proteins.

The elution curve of the Trz1-mutarotase mixture are in pink; and elution curve of individual Trz1 are in blue, and mutarotase in green.

II.2 Production of the complexes from Trz1, Nuc1 and mutarotase.

II.2.1 Trz1 and mutarotase

We tested the interaction between Trz1 and mutarotase by gel filtration. Purified Trz1 and mutarotase proteins were mixed at equal molar concentration (20 μ M), kept on ice for 30 min and subsequently injected to size exclusion chromatography at 4 °C. The elution profile of the mixture contained two peaks, with an elution volume of 14.16 ml and 16.29 ml, respectively. Their elution volumes and SDS-PAGE profiles suggest that they correspond to the individual Trz1 and mutarotase proteins respectively (Figure II-4), providing no evidence for the formation of a stable Trz1/mutarotase complex. Mixing both proteins in buffers with

various salt concentrations (100 mM or 200 mM NaCl) did not change the size exclusion elution profile (data not shown). Obviously, in disagreement with the reported co-IP data, Trz1 does not interact with mutarotase, at least in our *in vitro* conditions.

II.2.2 Trz1/Nuc1 binary complex

Our first strategy to obtain the binary Trz1/Nuc1 complex was to express the two proteins from co-transfected cells (Figure II-5). Strep-tagged Nuc1 co-eluted with His-Trz1 from the Strep-Tactin affinity purification (Figure II-5 A), with an excess of Nuc1 due to higher expression level in the cells. Trz1 and Nuc1 also co-eluted in the same peak in the followed by size exclusion chromatography (p1 of Figure II-5 B and C), and extra Nuc1 eluted individually in another peak (p2). The Trz1/Nuc1 complex containing a contamination protein from peak1 was further purified by a Ni-NTA affinity purification step (Figure II-5 D) and a subsequently gel filtration (Figure II-5 E and F). Only 0.17 mg of Trz1/Nuc1 complex yielded from an 800 ml bacterial culture, therefore this method was finally abandoned.

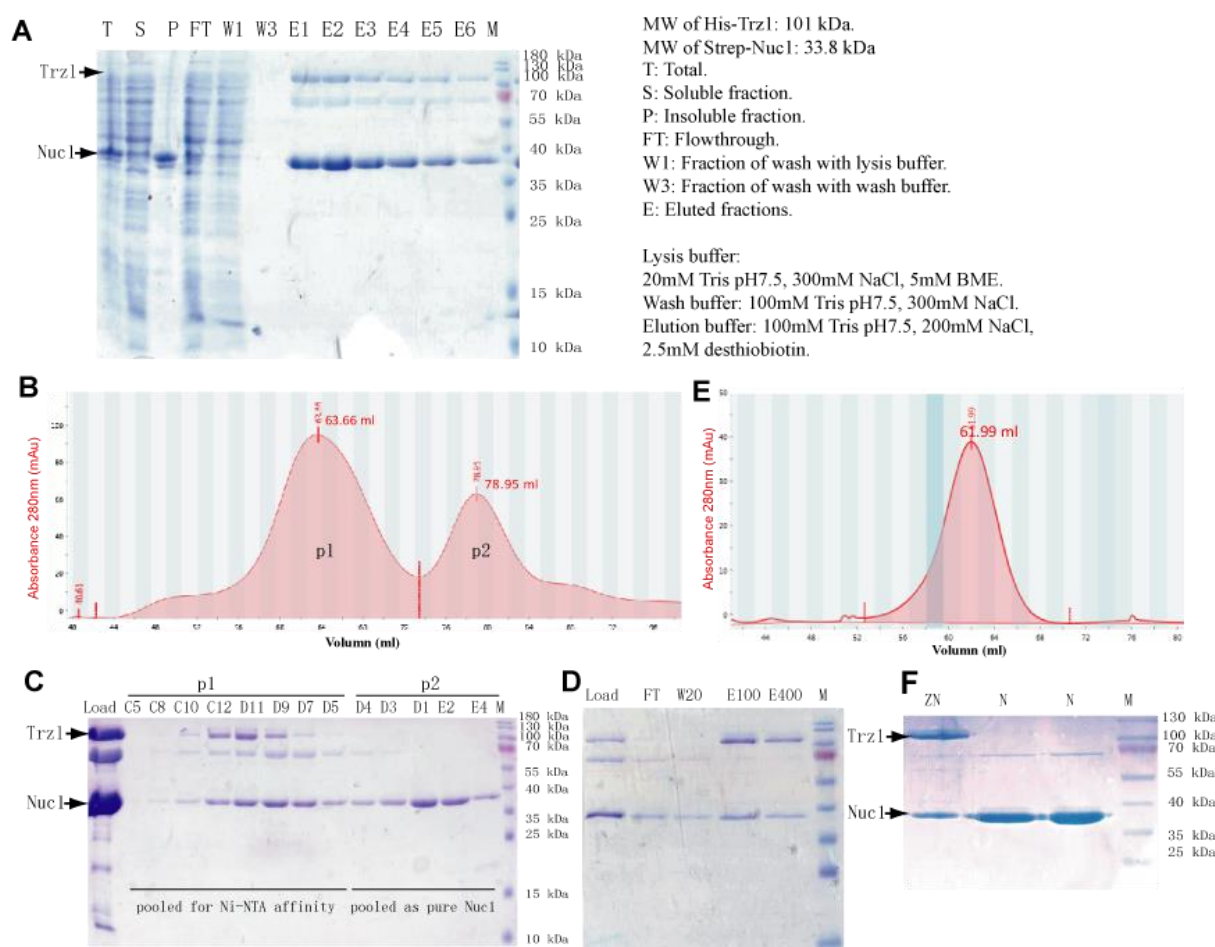


Figure II-5 Trz1/Nuc1 Purification from co-transfected cells.

The plasmids containing the Trz1 and Nuc1 ORFs are co-transfected into BL21-Gold cells. (A) SDS-PAGE of the fractions of the Strep-Tactin affinity purification (B) Size exclusion chromatography of elutions (fraction "E1" to "E6") from affinity purification step. Column: Superdex 200Hiload 16/60. Classic buffer A. (C) SDS-PAGE of the peak fractions of (B). (D)SDS-PAGE of the fractions of the Ni-NTA affinity purification of p1. Buffer based on classic buffer A. (E) SEC of the elution from Ni-NTA affinity. Column: Superdex 200 Hiload 16/60. Classic buffer A. (F) Purified Trz1/Nuc1 (ZN) and Nuc1 (N), pooled from (E) and p2 of (B), separately.

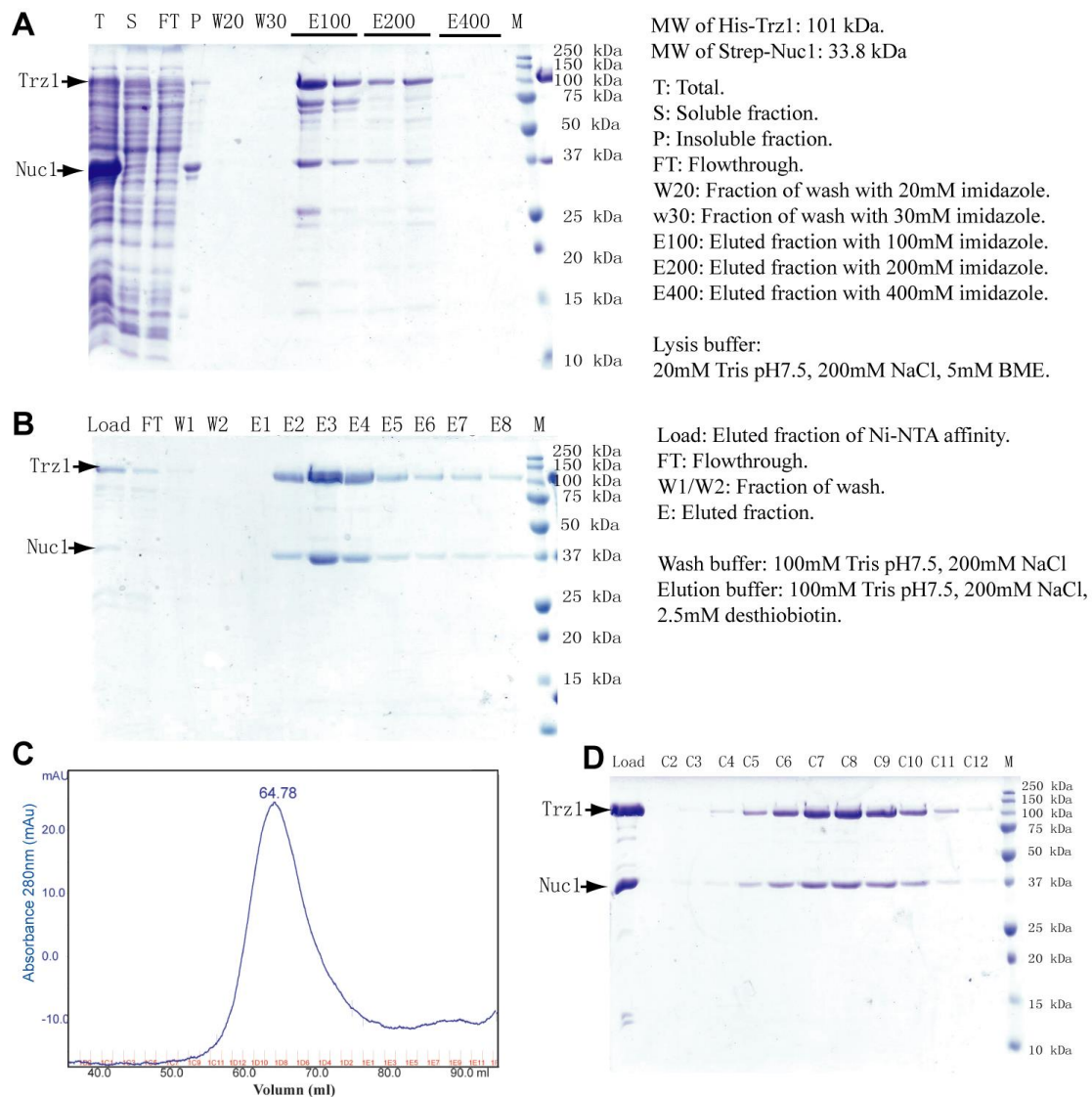


Figure II-6 Purification of the Trz1/Nuc1 complex.

Trz1 was expressed in BL21-Gold and Nuc1 was expressed in BL21-CodonPlus cells. Cells resuspensions were mixed and sonicated together. (A) SDS-PAGE of the fractions of the Ni-NTA affinity purification. (B) SDS-PAGE of the fractions of the Strep-Tactin affinity purification after the Ni-NTA purification step. (C) Size exclusion chromatography of the pooled fractions after affinity chromatography. Column: Superdex 200 Hiload 16/60. Buffer: 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM BME. (D) SDS-PAGE of the peak fractions of (C).

In a second strategy suspensions of the cells, which individually expressed Trz1 and Nuc1 in their preferred strains (BL21-Gold for Trz1 and CondonPlus for Nuc1) were mixed

and sonicated. In this experiment His-tagged Trz1 was used to bait strep-Nuc1 using Ni-NTA affinity resin, and upon imidazole washing of the column Trz1 and Nuc1 were co-eluted (Figure II-6 A). The eluate was further purified by a second affinity purification by Strep-Tactin resin, gaining a rather pure Trz1/Nuc1 complex (Figure II-6 B). Subsequently size exclusion chromatography showed a symmetric peak of the complex and Trz1 and Nuc1 were clearly in a 1:1 ratio (Figure II-6 C and D). With this strategy, we obtained about 0.2 mg of Trz1/Nuc1 binary complex from 800 ml cultures of each protein.

Therefore, reasoning order to obtain the highest yields we abandoned the two strategies described above and decided to prepare the Trz1/Nuc1 binary complex by mixing individually purified Trz1 and Nuc1 in a 1:1 ratio and co-elute in gel filtration (data not shown).

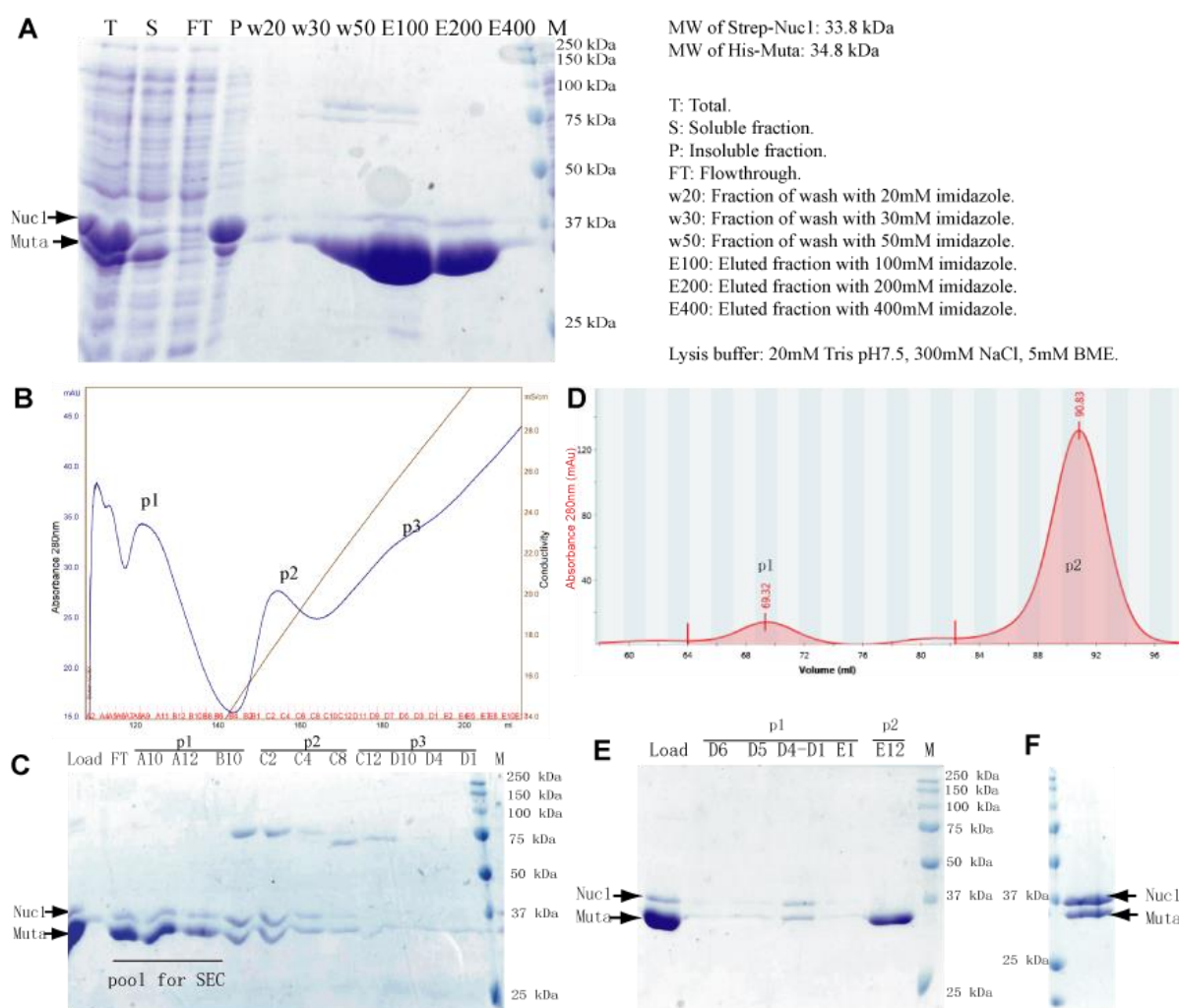


Figure II-7 Purification of the Nuc1/mutarotase complex.

Nuc1 was expressed in BL21-CodonPlus cells; mutarotase was expressed in BL21-Gold cells. Resuspensions of Nuc1-expressing and mutarotase-expressing cells were mixed at a ratio of 6:1. (A) SDS-PAGE of the fractions of the Ni-NTA affinity purification. (B) Ion change chromatography of the pooled fractions of the affinity chromatography step. Column: mono Q 5/50 GL. Buffer Low/high salt: 20 mM Tris-HCl, pH 7.5, 50 mM / 1 M

NaCl, 10% glycerol, 10 mM BME. (D) Size exclusion chromatography. Column: Superdex 200 HiLoad 16/60. Classic buffer A (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM BME). (C) (E) SDS-PAGE of the fractions of (B) and (D), separately. (F) SDS-PAGE of purified Nuc1/Mutarotase pooled from peak 1 of (D).

II.2.3 Nuc1/Mutarotase binary complex

We previously noticed that the majority of the individually expressed Nuc1 protein ended up in the insoluble fraction. We therefore attempted to “solubilize” Nuc1 by mixing and lysing the suspension of cells expressing Strep-Nuc1 together with cells expressing His-mutarotase. After cell lysis we attempted to purify the complex either by Strep-Tactin or by Ni-NTA chromatography (Figure II-7 A; Figure II-8 A).

The strategy shown in used a Ni-NTA column for affinity purification: His-mutarotase eluted in large excess together with Strep-tagged Nuc1, although cells expressing His-mutarotase versus cells expressing Strep-tagged Nuc1 were already mixed in a volume ratio of 1:6 (Figure II-7 A). A Mono Q ion-exchange chromatography of the eluate partially separated the Nuc1/mutarotase complex from the contaminating protein (p1 of Figure II-7 B and C), and with the subsequent size exclusion chromatography with Superdex200 column we obtained pure Nuc1/mutarotase binary complex (p1 of Figure II-7 D and E; Figure II-7 F). Although the molecular weight of Strep-Nuc1 (33.8 kDa) is slightly smaller than that of the His-Mutarotase (34.8 kDa), it migrates a bit slower on the SDS-PAGE. We yielded poorly 0.1 mg of Nuc1/mutarotase complex from 1.6 L bacterial culture expressing Nuc1 and 250 ml culture expressing mutarotase.

For a second strategy, we increased the volume ratio of cells expressing Nuc1 to cells expressing mutarotase to 12:1; and used a Strep-Tactin column in the affinity purification step. SDS-PAGE of the purified fractions suggests Nuc1 and Mutarotase are present in a 1:1 ratio in the eluate as estimated from the intensity of the bands (Figure II-8 A). Subsequent size exclusion chromatography with the Superdex 75 column showed Nuc1 and mutarotase co-eluted in a sharp peak, and the contaminant created a “shoulder” at the start of the peak (Figure II-8 B and C). We pooled the fractions which were not contaminated, and obtained relatively pure Nuc1/mutarotase complex (Figure II-8 D), at a yield of about 0.3 mg purified complex starting from by 1.6 L of culture expressing Nuc1 and 125 ml of culture expressing mutarotase.

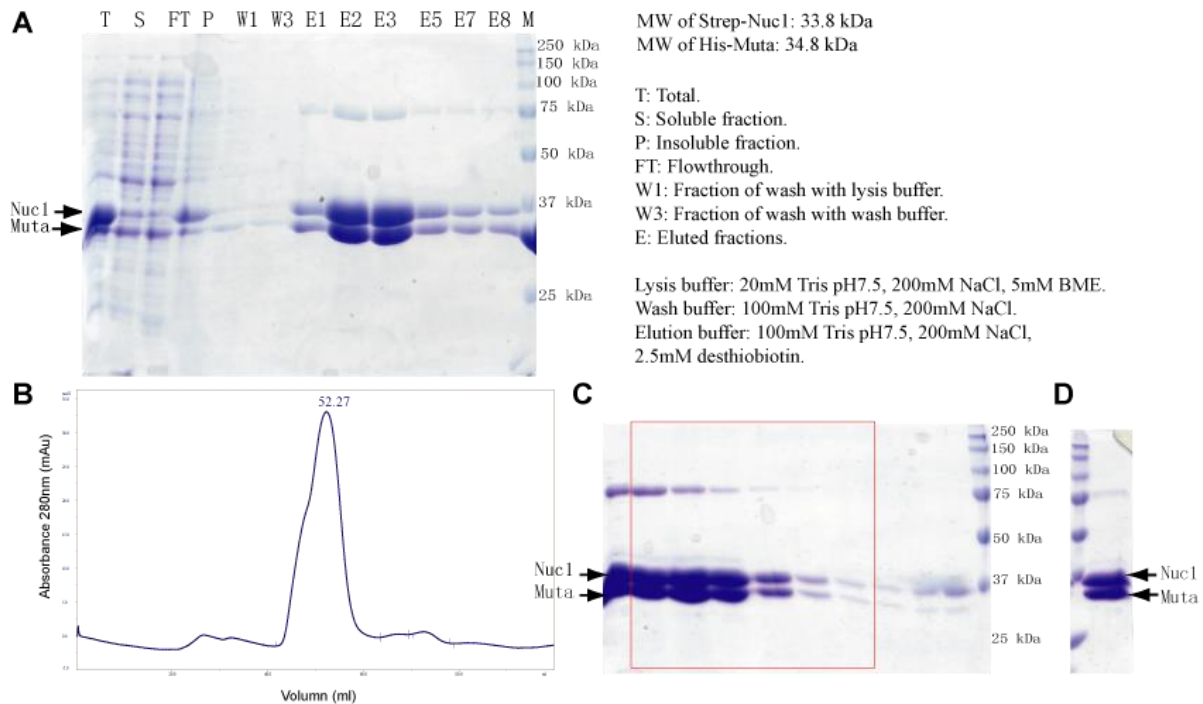


Figure II-8 Purification of the Nuc1 / mutarotase complex.

Nuc1 was expressed in BL21-CodonPlus cells; mutarotase was expressed in BL21- Gold cells. Resuspensions of Nuc1-expressing and mutarotase-expressing cells were mixed at a ratio of 12:1. (A) SDS-PAGE of the fractions of the Strep-Tactin affinity purification. (B) and (C) Size exclusion chromatography and SDS-PAGE of the corresponding fractions. Column: Superdex 75 Hiload 16/60. Buffer: 20mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM BME. Red square: pooled fractions. (D) SDS-PAGE of the purified Nuc1/mutarotase complex.

II.2.4 Reconstitution of the ternary Trz1/Nuc1/mutarotase complex

We chose to reconstitute the ternary Trz1/Nuc1/mutarotase complex by mixing purified Trz1 and binary Nuc1/mutarotase complex. We first mixed a suspension of Nuc1- and mutarotase- expressing cells in a volume ratio 4:1, lysed the cell mixture and ran the lysate on a Strep-Tactin resin to fix the binary Nuc1/mutarotase complex (lane “Rnm” in Figure II-10 A). Crudely purified Trz1 (Figure II-9) was then added to the resin and incubated for 30 min at 4 °C. After elution of the column using desthiobiotin we found His-tagged Trz1 in the eluted fractions together with Nuc1 and mutarotase, suggesting that Trz1 remained bound to the binary Nuc1/mutarotase complex (lane “E” in Figure II-10 A). The eluted fractions were then concentrated and injected onto a size exclusion column. The chromatogram was composed of two peaks, the first eluting peak (p1 of Figure II-10 B) containing Trz1, Nuc1 and mutarotase (Figure II-10 C) and the slower eluting peak (p2) containing mainly Nuc1 and mutarotase. To prepare high quantities of the ternary complex we increased the amount of added Trz1 in order to saturate the binary Nuc1/mutarotase complex and purified the mixture by ion exchange (Figure II-10 D and E), and subsequent size exclusion chromatography.

SDS-PAGE of the major peak in the size exclusion chromatogram shows that i) the three proteins co-eluted in an estimated 1:1:1 ratio (Figure II-10 F and G); and ii) that the putative Trz1/Nuc1/mutarotase ternary complex is relatively pure.

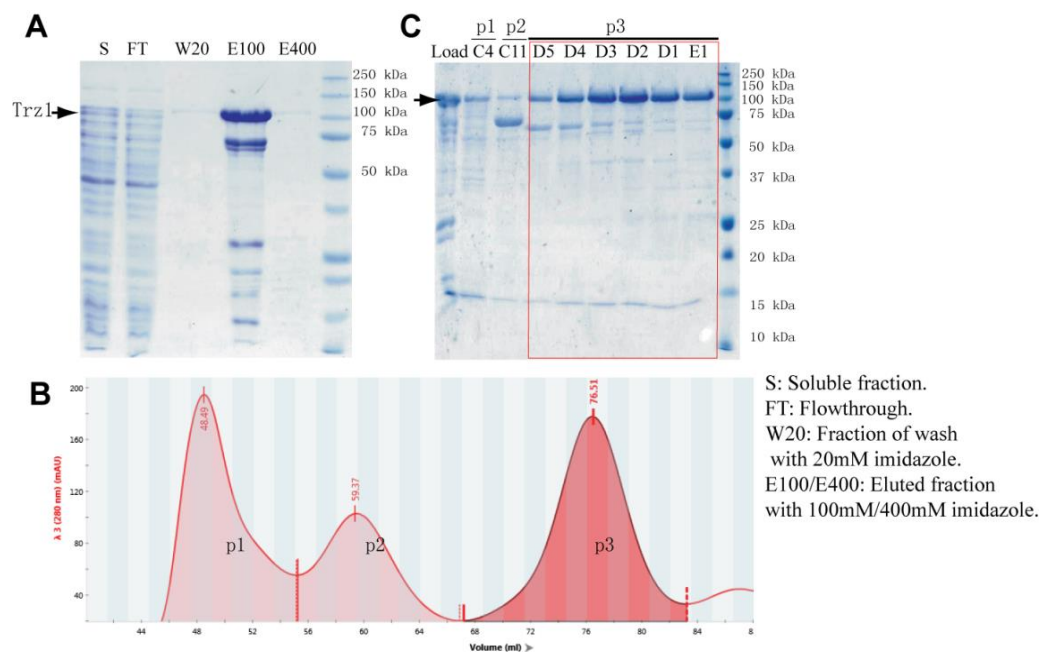


Figure II-9 Crude purification of Trz1.

(A) SDS-PAGE of the fractions of Ni-NTA affinity purification. Lysis buffer: 20mM Tris-HCl, pH 7.5, 400 mM NaCl, 5 mM BME. (B) Size exclusion chromatography. Column: Superdex 200 Hiload 16/60. Buffer: 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM BME. (C) SDS-PAGE of the eluted fractions of (B). The red box indicates pooled fractions.

The ternary complex can also be reconstituted by adding purified mutarotase to the separately purified Trz1/Nuc1 complex but this strategy produces lower yields and was abandoned (data not shown).

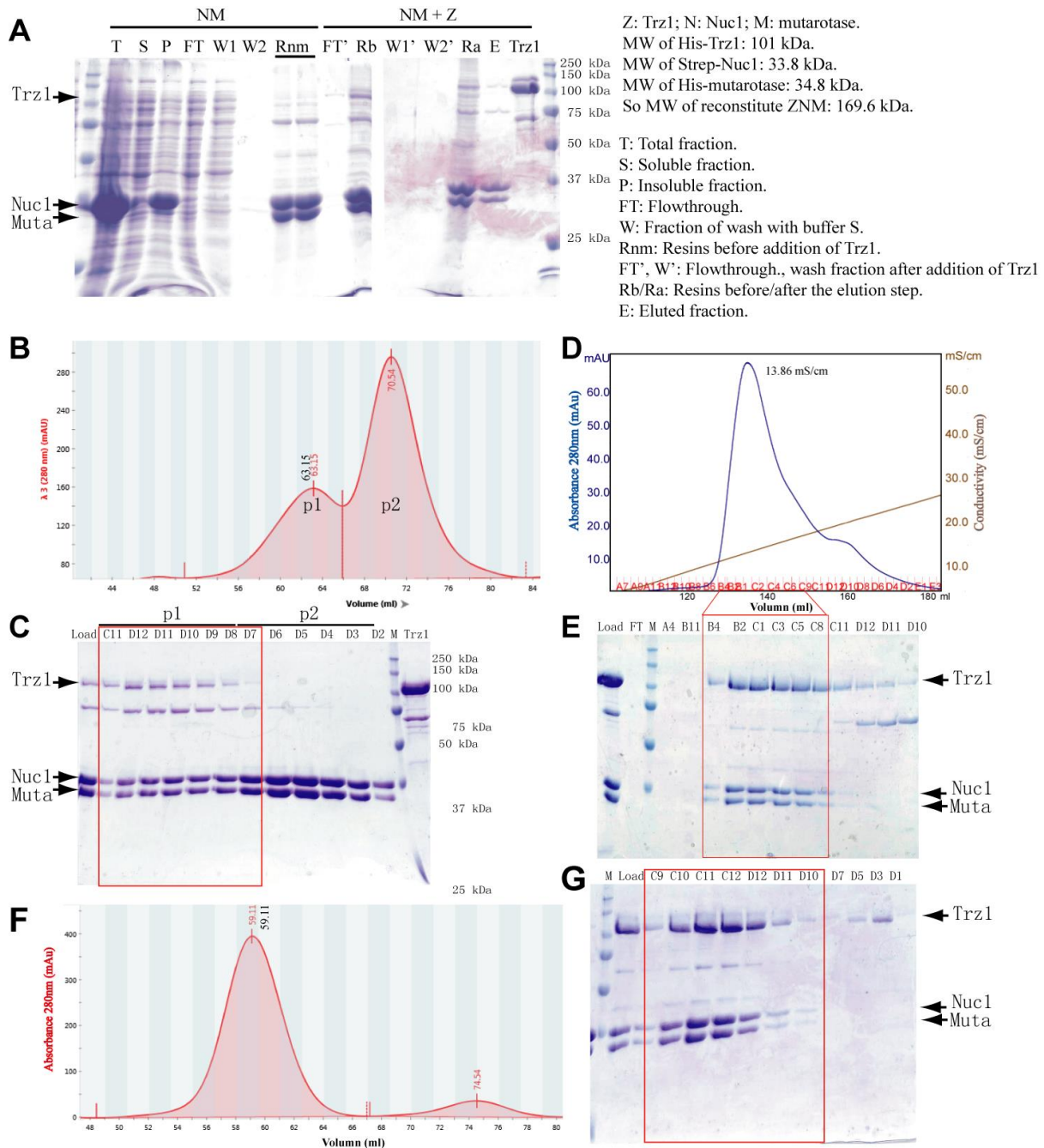


Figure II-10 Reconstitution of the ternary Trz1/Nuc1/mutarotase complex.

(A) SDS-PAGE of the fractions of the Strep-Tactin affinity purification. Lysis buffer: 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 5 mM BME. Buffer S: 100 mM Tris-HCl, pH 7.5, 200 mM NaCl. Buffer E: buffer S in addition with 2.5 mM desthiobiotin. Trz1 is from crude purification. (B) (C) Size exclusion chromatography with buffer A and corresponding fractions on SDS-PAGE. Column: Superdex 200 Hiload 16/60. (D) (E) Ion change chromatography and corresponding fractions on SDS-PAGE. Column: mono Q 5/50 GL. Buffer low/high salt: 20 mM Tris-HCl, pH 7.5, 50 mM / 500 mM NaCl, 10% glycerol, 10 mM BME. (F) Size exclusion chromatography of pooled fraction from mono Q. Same column and buffer with (B). (G) SDS-PAGE of corresponding fractions of (F). The red box indicates pooled fractions for the next purification step.

II.2.5 Purification of the Trz1/Nuc1/mutarotase complex from a polycistronic plasmid

Due to the complexity and low yield of purification protocols of the ternary complex, we designed a polycistronic plasmid harboring ORFs coding for Trz1, Nuc1 and mutarotase. The proteins express and fold independently, and form the complex naturally inside the cells.

II.2.5.1 Design of the polycistronic plasmid

Our results show that the binary complexes of Trz1/Nuc1, Nuc1/mutarotase are well formed and can be purified, while Trz1 and mutarotase do not form a complex *in vitro*. These first results suggest Nuc1 plays key role for the ternary complex formation. We therefore arranged the ORFs of the three partner in the following order from 5' to 3' in the polycistronic plasmid: Trz1-Nuc1-mutarotase. The Nuc1 ORF was the $\Delta N43$ construct with a N-terminal Strep-tag, the mutarotase ORF is fused with the His-tag at the C-terminal, whereas the Trz1 ORF was tag-free (Figure II-11). RBS sequences were inserted manually between each two ORFs. The polycistronic sequence was inserted between the enzyme restriction sites NdeI-XhoI of the pET21a(+) plasmid. Several nucleotides were mutated to remove undesired restriction enzyme sites; we also avoided to have repeats of any restriction enzyme sites at the 5' or 3' ends of the ORFs so that the plasmid will also be suitable for sub-cloning Trz1-Nuc1 or Nuc1-mutarotase binary complexes if necessary. The full sequence is attached in the supplementary materials chapter.

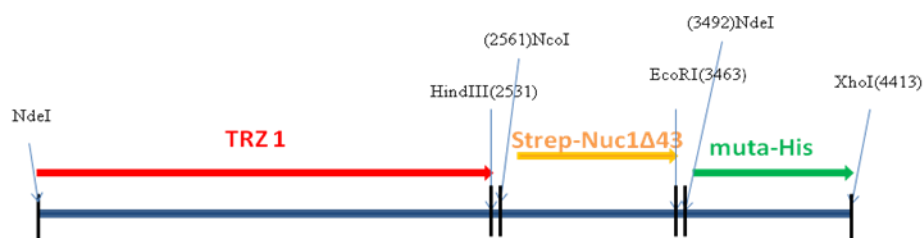


Figure II-11 Schematic presentation of the ORFs coding for Trz1, Nuc1, and mutarotase in the polycistronic plasmid designed for overexpression of the ternary complex.

II.2.5.2 Purification of the ternary Trz1/Nuc1/mutarotase complex with the polycistronic plasmid

The polycistronic plasmid harboring Trz1/Nuc1/mutarotase was transfected into BL21-Gold and the protein expression was induced at 15 °C. The protein expression levels of the three proteins corresponded to those observed for the individual constructs: Trz1 expressed poorly as usual, Nuc1 was expressed well but only partly soluble, and mutarotase was highly

expressed. Using the His-tag on mutarotase, cell lysates were first purified by affinity purification on a Ni-NTA resin. Three main proteins were found in the eluted fractions (Figure II-12 A), indicating that the ternary complex had been properly constituted inside the cells. Pooled elution fractions from Ni-NTA affinity were loaded to a mono Q column. Excess of mutarotase passed in the flowthrough, whereas the ternary complex bound to the column and was eluted as one single peak using a salt gradient (Figure II-12 B and C). The final purification step of the pooled mono Q fractions by size exclusion chromatography gave a single symmetric peak eluted at 62.4 ml, and SDS-PAGE of the fractions clearly revealed the presence Trz1, Nuc1, mutarotase (Figure II-12 D and E).

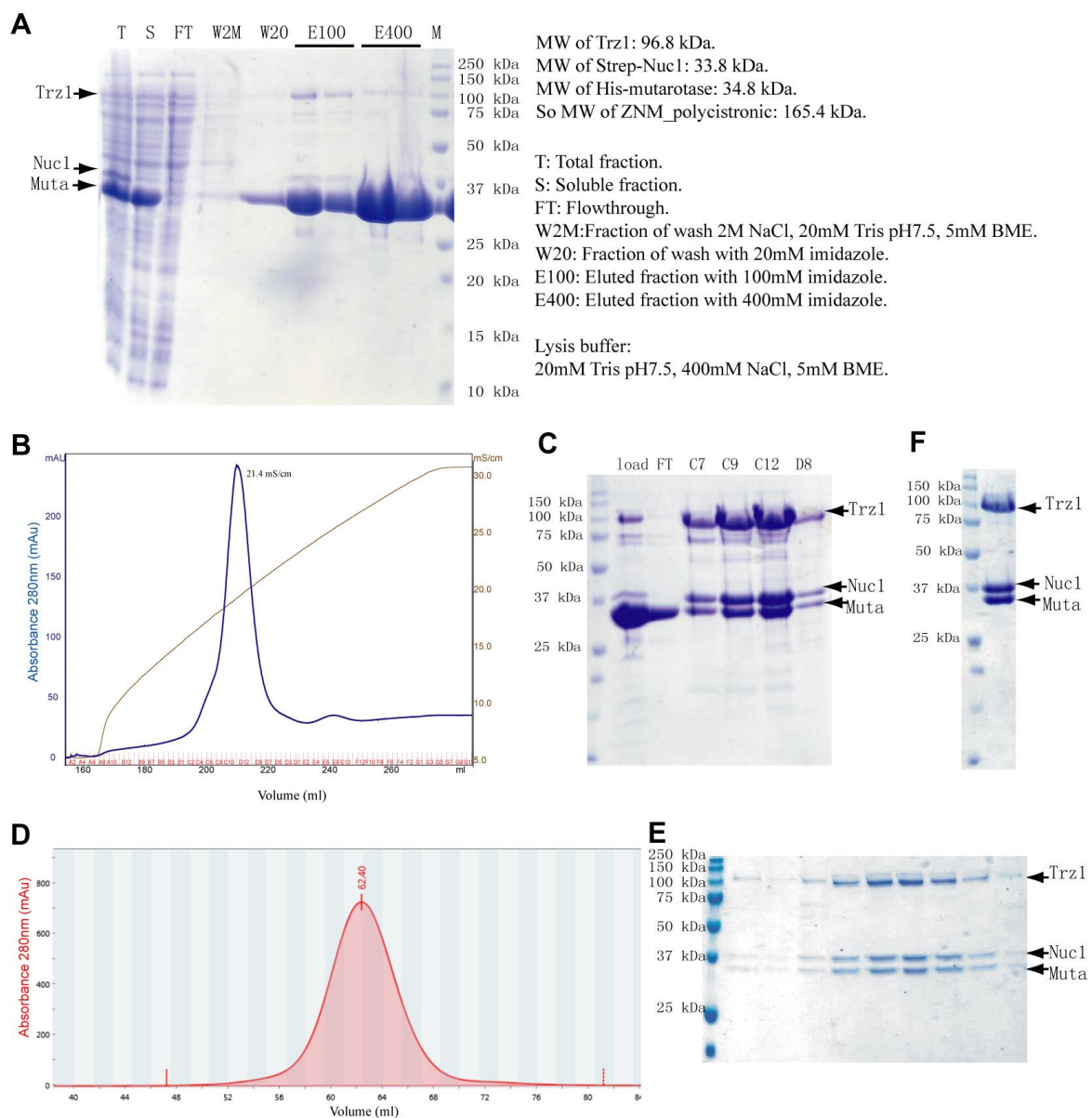


Figure II-12 Purification of Trz1/Nuc1/mutarotase complex expressed from the polycistronic plasmid.

(A) SDS-PAGE of the fractions of the Ni-NTA affinity purification. (B) Ion change chromatography. Column: mono Q 5/50GL. Buffer low/high salt: 20 mM Tris-HCl, pH 7.5, 100 mM / 500 mM NaCl, 10% glycerol, 10 mM BME. (D) Size exclusion chromatography. Column: Superdex 200 HiLoad 16/60. Buffer: 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM BME. (C) (E) SDS-PAGE of the eluted fractions of (B) and (D), separately. (F) SDS-PAGE of purified Trz1/Nuc1/mutarotase.

The ternary complex purified by this simple 3-step protocol is purer (Figure II-12 F vs Figure II-10 G), and the yield is much higher than that of the reconstituted complex (1.4 mg from a 800 ml bacterial culture with optimized protocol). The polycistronic plasmid greatly facilitates and improves the quality of the purification of the complex.

II.3 Stoichiometry of the binary and tertiary complexes

The gel filtration columns were calibrated by several standard proteins (see material and methods). Since all size exclusion chromatograms were collected using the same calibrated column and using the same buffers, the elution volumes informed us on the molecular weight in solution of the various proteins and their complexes.

Overlay of the size exclusion chromatograms of the various preparations of the complexes and individual proteins clearly reveal important differences in the elution volumes of the peaks (Figure II-13). Trz1 (blue curve), Nuc1 (purple curve) and mutarotase alone (green curve) eluted at 14.1 ml, 14.9 ml and 16.1 ml, respectively (Figure II-13 A). These values correspond to the following molecular weights: 109 kDa for Trz1 (theoretical Mr 101 kDa), 69 kDa for Nuc1 (theoretical Mr 33.8 kDa) and 37 kDa for mutarotase (theoretical Mr 34.8 kDa). Trz1 and mutarotase are hence monomers in solution and Nuc1 forms homodimers confirming what was previously known for the quaternary structure of these proteins (Table II-2).

The binary Nuc1/mutarotase complex (NM, Figure II-13 B) eluted at 13.8 ml, corresponding to a Mr in solution of 131 kDa, close to the value of a hetero-tetramer composed of two Nuc1 and two mutarotase subunits (theoretical Mr 137.2 kDa). The Trz1/Nuc1 binary complex (ZN, Figure II-13 C) eluted at 12.6 ml, corresponding to a Mr in solution of 241 kDa, close to the value of a hetero-tetramer composed of two Nuc1 and two Trz1 subunits (theoretical Mr 269.6 kDa). The ternary Trz1/Nuc1/mutarotase complex (ZNM, Figure II-13 A) eluted at 11.7 ml, which represents a shift of 2.4 ml compared to the elution volume of Trz1 alone, and corresponding to a Mr of 395 kDa. The theoretical masses of a Trz1/Nuc1/mutarotase complex is 165.4 kDa, therefore the mass of the ternary complex in solution could correspond to a heterohexamers. Since Nuc1 forms a homodimer and interacts

both with Trz1 and mutarotase, we propose that the Trz1/Nuc1/mutarotase assembles as a dimer of hetero-trimers (Table II-2). The molecular weights of all the complexes were also measured from their elution volumes on a larger calibrated column Superdex 200 Hiload 16/60, leading to similar Mr values (Table II-3).

Table II-2 Experimental molecular weights of the complexes by elution volume from column Superdex 200 10/300 GL

Protein	Mr theor.	Ve (ml)	Mr exp.	Oligmerization	assembly	annotation
Z	101 kDa	14.1 ml	109 kDa	monomer	Z	-
N	33.8 kDa	14.96 ml	69 kDa	dimer	N ₂	-
M	34.8 kDa	16.1 ml	37 kDa	monomer	M	-
ZN	134.8 kDa	12.62 ml	241 kDa	tetramer	(ZN) ₂	dimer of heterodimers
NM	68.6 kDa	13.76 ml	131 kDa	tetramer	(NM) ₂	dimer of heterodimers
ZNM (poly)*	165.4 kDa	11.7 ml	395 kDa	hexamer	(ZNM) ₂	dimer of heterotrimers

*ZNM(poly) indicates ZNM purified from polycistronic plasmids, in which Trz1 is tag-free and 96.8 kDa. Ve: elution volume from the column. Mr theo.: theoretical molecular weight calculated by protein sequence. Mr exp.: experimental molecular weight calculated from the elution volume by calibrated column.

Table II-3 Experimental molecular weights of the complexes by elution volume from column Superdex 200 Hiload 16/60.

Protein	Mr theor.	Ve (ml)	Mr exp.	Oligmerization	assembly	annotation
Z	101 kDa	76.21 ml	108 kDa	Monomer	Z	-
N	33.8 kDa	78.95 ml	87 kDa	dimer	N ₂	-
M	34.8 kDa	91.40 ml	34 kDa	monomer	M	-
ZN	134.8 kDa	63.66 ml	280 kDa	tetramer	(ZN) ₂	dimer of heterodimers
NM	68.6 kDa	70.05 ml	172 kDa	tetramer	(NM) ₂	dimer of heterodimers
ZNM (poly)*	165.4 kDa	60.26 ml	364 kDa	hexamer	(ZNM) ₂	dimer of heterotrimers

Legends: The same as those of Table II-2

Table II-4 Dimension and experimental molecular weights of the complexes by SEC-MALLS

Protein	Mr theor.	Rh (nm)	Mr exp.	Oligmerization
Z	101 kDa	3.775	98.639 kDa	monomer
M	34.8 kDa	2.537	33.893 kDa	monomer
NM	68.6 kDa	4.692	150.443 kDa	tetramer
ZNM (poly)*	165.4 kDa	6.441	321.440 kDa	hexamer

*ZNM(poly) indicates ZNM purified from polycistronic plasmids, in which Trz1 is tag-free and 96.8 kDa.

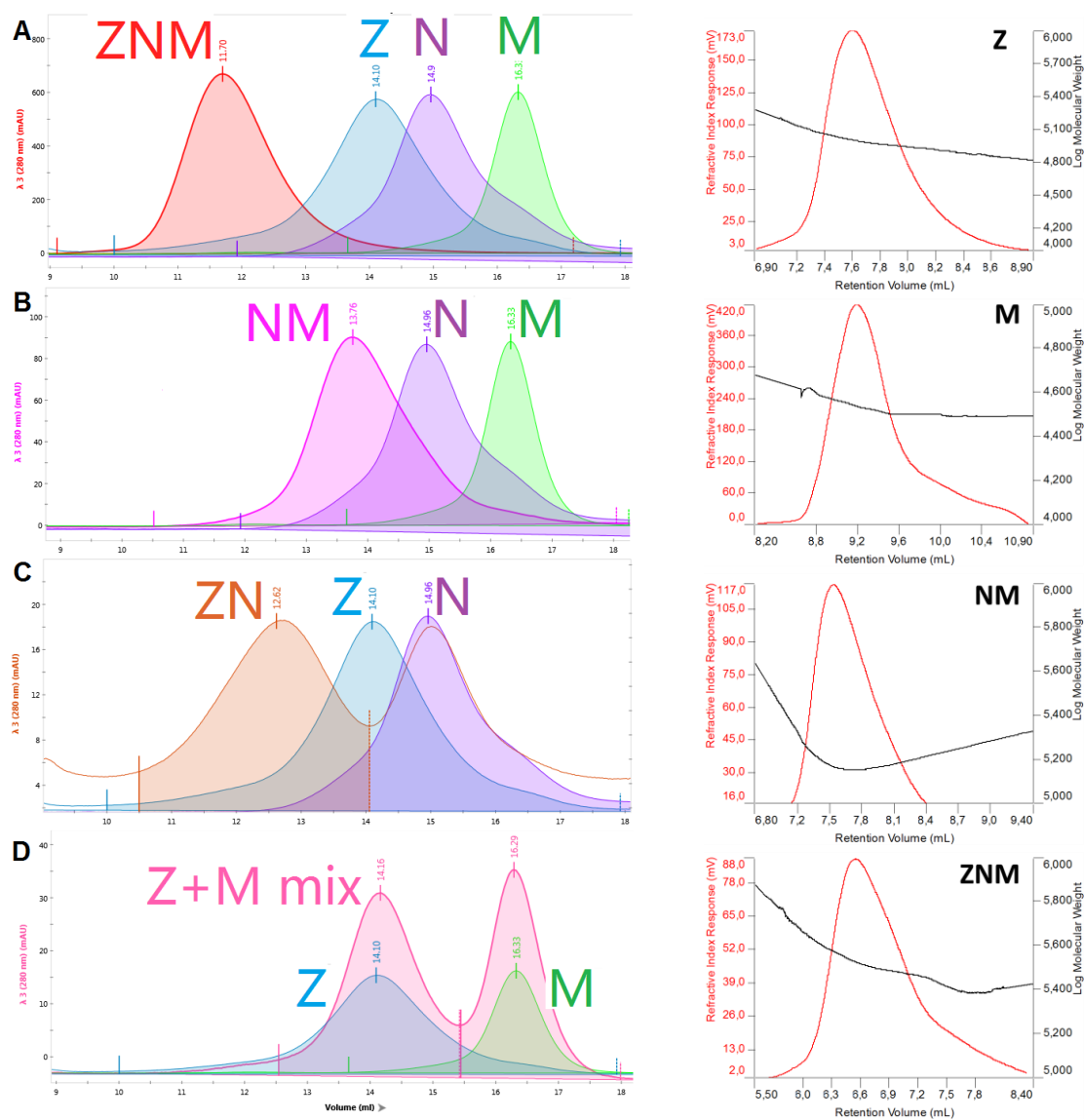


Figure II-13

Figure II-14

Figure II-13 Overlay of the size exclusion chromatograms of Trz1, Nuc1, mutarotase and their complexes. (A) Overlay of size exclusion chromatograms of the ternary Trz1/Nuc1/mutarotase complex (ZNM) and that of the individual Trz1 (Z), Nuc1 (N) and mutarotase (M) proteins. The curve of ZNM are shown in red; the curves of Z, N, M are shown in blue, purple, and green, respectively, and coloring the same in the following figures. (B) Overlay of size exclusion chromatograms of the binary Nuc1/mutarotase complex (NM, pink curve) and that of the individual Nuc1 (N) and mutarotase (M) proteins. (C) Overlay of size exclusion chromatograms of a ternary Trz1/Nuc1 complex (ZN, orange curve) and that of the individual Trz1 (Z) and Nuc1 (N) proteins. (D) Overlay of size exclusion chromatograms of a Trz1-mutarotase mixture (light pink curve) and that of the individual Trz1 (Z) and mutarotase (M) proteins. (D) Overlay of size exclusion chromatograms of a Trz1-mutarotase mixture (pink) and the individual Trz1 and mutarotase proteins.

Figure II-14 SEC-MALLS analysis of Trz1, Nuc1, mutarotase and related complexes.

Purified proteins samples were injected to the SEC-MALLS system, and data were analyzed with software supplied by the manufacturer.

The oligomerization state of the complexes was further measured by size exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS) (Figure II-14). The latter detection method provides an independent estimated of the molecular weight of the

species in solution. The Mr of Trz1 and mutarotase were estimated as 98.64 kDa and 33.89 kDa, respectively, confirming they are monomeric in solution. The measured Mr of the ternary Trz1/Nuc1/mutarotase (ZNM) complex was 321.44 kDa, which is approximately double of theoretic value for the heterotrimer (165.4 kDa), indicating ZNM complex is a hexamer in solution. The same held for the binary Nuc1/mutarotase (NM) complex: the Mr obtained from SEC-MALLS was 150.44 kDa, compared to a theoretical Mr of 68.6 kDa for the hetero-dimer, we drew the conclusion it is a hetero-texamer (Table II-4). We didn't obtain a nice signal of the Trz1/Nuc1 (ZN) complex and Nuc1 caused by insufficient quantity of the samples.

II.4 Assemblage of the complex

The gel filtration and SEC-MALLS data obtained of the various complexes provide us with a first glimpse of the quaternary organization of the ternary complex (Figure II-15). First, since we showed that the binary Trz1/Nuc1 and Nuc1/mutarotase complexes could be purified by gel filtration, they both must be stable. Secondly, Trz1 could not be co-purified with mutarotase under similar biochemical conditions as for the other two, suggesting they do not form a stable complex in vitro. We therefore propose that the Nuc1 dimer forms the kernel of the complex and bridges the Trz1 and mutarotase subunits. We cannot exclude that Trz1 and mutarotase interact in the context of the ternary complex, but if these contacts exist they must be less stable than those among the other subunits.

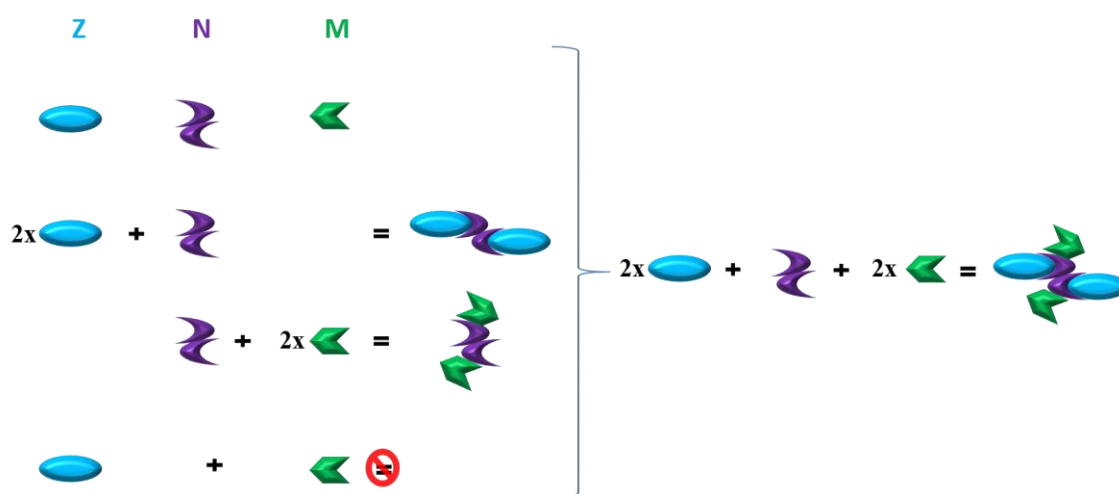


Figure II-15 Schematic model of the assemblage of the Trz1/Nuc1/mutarotase complex.

III. Enzymatic activities of the protein complexes.

III.1 RNase Z activity and the effects of Nuc1/mutarotase

Since the three proteins forming the ZNM complex are each endowed with enzymatic activities, we investigated whether these activities were affected compared to those of the individual proteins. We started by measuring the pre-tRNA processing RNase Z activity of Trz1, and then explored whether the presence of Nuc1 or mutarotase would affect this activity.

III.1.1 Substrates for measuring the RNase Z activity.

We decided to use previously characterized tRNA-precursors tRNA^{Ser} from *S. cerevisiae* and pre-tRNA^{Thr} from *B. subtilis* as substrates for the RNase Z activity assay, respectively. These precursor tRNAs were synthesized by T7 RNA polymerase *in vitro* (Pellegrini et al. 2003; Skowronek et al. 2013).

III.1.1.1 Templates for *in vitro* transcription

The templates for *in vitro* transcription were amplified from chromosomal DNA (for yeast tRNA^{Ser}) or plasmid pHMT1 (for *B. subtilis* tRNA^{Thr}) by PCR (see methods). The upstream primers contained a T7 RNA polymerase promoter; and the downstream primers were designed to include the 3'-trailer sequence of the pre-tRNA. The PCR products were examined by agarose gel electrophoresis (Figure III-1).

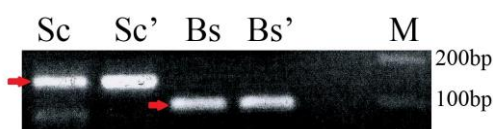


Figure III-1 PCR product of pre-tRNA as templates for *in vitro* transcription. Sc: tRNA^{Ser} from *S. cerevisiae*. Bs: tRNA^{Thr} from *B. subtilis*. A second equivalent PCR were set up in parallel (labelled with a prime) for each species. Expected products were indicated by red arrows. 6 μ l of 50 μ l was loaded on 2% agarose gel and electrophoresis in TBE buffer.

III.1.1.2 *In vitro* transcription

Precursor tRNAs were transcribed *in vitro* from the PCR templates by T7 RNA polymerase. RNase Z generally removes the 3'-end of pre-tRNAs after 5'-end processing by RNase P (Hartmann et al. 2009), so we prepared 5'-processed substrate mimics for optimal RNase Z cleavage (Table III-1) by making pre-tRNAs with a 5' monophosphate group. This

was done by using 8-fold higher concentrations of GMP (4 mM) than GTP (0.5 mM was used for the G/A/C nucleotides). The pre-tRNA was uniformly labeled using a mixture of $\alpha^{32}\text{P}$ -labelled UTP at high concentration and non-labelled UTP at low concentration (12 μM).

Transcribed pre-tRNA was visualized by agarose gel electrophoresis (data not shown), and unincorporated nucleotides were removed (See methods). Yeast pre-tRNA^{Ser} was 136 nt and *B. subtilis* pre-tRNA^{Thr} was 83 nt in length; and 3'-trailer removal yielded mature tRNAs of 82 nt and 73 nt, respectively (Figure III-2, Table III-2).

Table III-1 tRNA *in vitro* transcription system

PCR product (template, 0.1 μg)	2 μl
DTT (100 mM)	2.5 μl
Mix ACG (2.5 mM)	5 μl
U (60 μM)	
UTP - $\alpha^{32}\text{P}$ (high [c])	3 μl
GMP 50 mM	2 μl
RNase inhibitor	1.5 μl
Buffer 5x	5 μl
T7 RNA polymerase	2 μl
H ₂ O	5 μl
Final Volume	25 μl

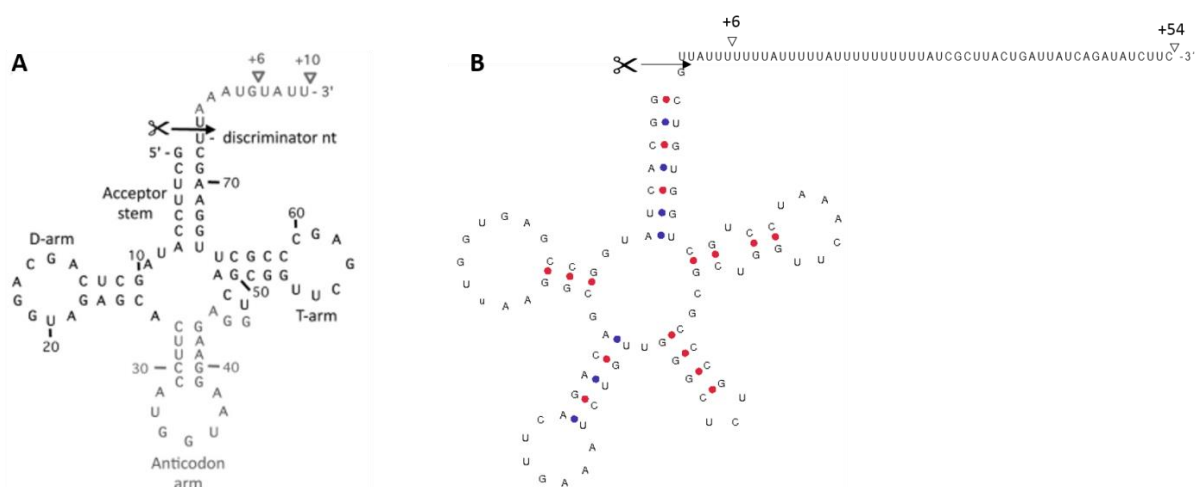


Figure III-2 pre-tRNAs. (A) pre-tRNA Threonine from *B. subtilis* (Pellegrini *et. al*, 2012); (B) pre-tRNA Serine from *S. cerevisiae*. The RNase Z cleavage site following the discriminator nucleotide is marked by an arrow and scissors symbol.

Table III-2 Lengths of tRNA species

type	source	Sc tRNA ^{Ser}	Bs tRNA ^{Thr}
Precursor tRNA	<i>in vitro</i> transcription	136 nt	83 nt
Mature tRNA	RNase Z cleavage	82 nt	73 nt
3'-trailer	RNase Z cleavage	54 nt	10 nt

III.1.2 RNase Z activity assay

III.1.2.1 Trz1 activity

First we needed to confirm that yeast tRNA^{Ser} prepared by *in vitro* transcription could be processed by Trz1, to show that our recombinant Trz1 was active (Figure III-3 A). We used a reaction buffer optimized for BsuTrz (40 mM Tris-HCl, pH 8.4, 2 mM MgCl₂, 2 mM KCl, 2 mM DTT (Pellegrini et al. 2003)), that we adapted for the test with Trz1 (50 mM Tris-HCl, pH 7.1, 5 mM MgCl₂, 5 mM KCl, 2 mM DTT (Skowronek et al. 2013)).

A constant amount of radioactively labeled precursor yeast tRNA^{Ser} was incubated with increasing amounts of Trz1 at 37 °C for 30 min, the mixture was then separated by urea gel and analyzed by autoradiography. Each lane represents a reaction mixture with a 10-fold increase of enzyme concentration. Precursor tRNA (136 nt) was clearly processed into mature tRNA (82 nt) and a 3'-trailer fragment (54 nt). An RNA species longer than 82 nt (marked with a star in the figure) corresponded to a smaller pre-tRNA species, caused by fall-off/arrest of T7 RNA polymerase transcription 6 nt after the tRNA structure (Pellegrini et al. 2003). Therefore, transcripts used as substrates were in fact a mixture of 136 nt and 88 nt fragments (82 nt plus 6 nt), so the mature tRNA cleavage product species (82 nt) was used as marker to follow the activity (Red frame in Figure III-3).

The autoradiograph of the gel electrophoretic profiles of the reaction mixtures revealed that 0.1 pmol (per 5 µl reaction) of Trz1 enzyme caused no observable hydrolysis while 1 pmol processed more than half of the substrate (Figure III-3, lanes 1-3). However, increasing the amount of Trz1 up to 10 pmol caused a smear in the lane (lane 4), which might be because some tRNA are trapped due to the high concentration of protein, and loss of signal because some of the radioactivity is bound up in the smear. BsuTrz also efficiently cleaved yeast pre-tRNA, and increasing enzyme concentration also increased the proportion of mature tRNA in the reaction mixture (lanes 6-8).

We also measured the Trz1 activity on the pre-tRNA^{Thr} substrate from *B. subtilis*, and the same trend as for the yeast pre-tRNA substrate was observed (Figure III-3 B). The 3'-processed fragment was not observed because it was only 10 nt long and had therefore run out of the gel. From these experiments we concluded that our recombinant Trz1 enzyme efficiently removed the 3'-trailer from both *S. cerevisiae* and *B. subtilis* precursor tRNAs.

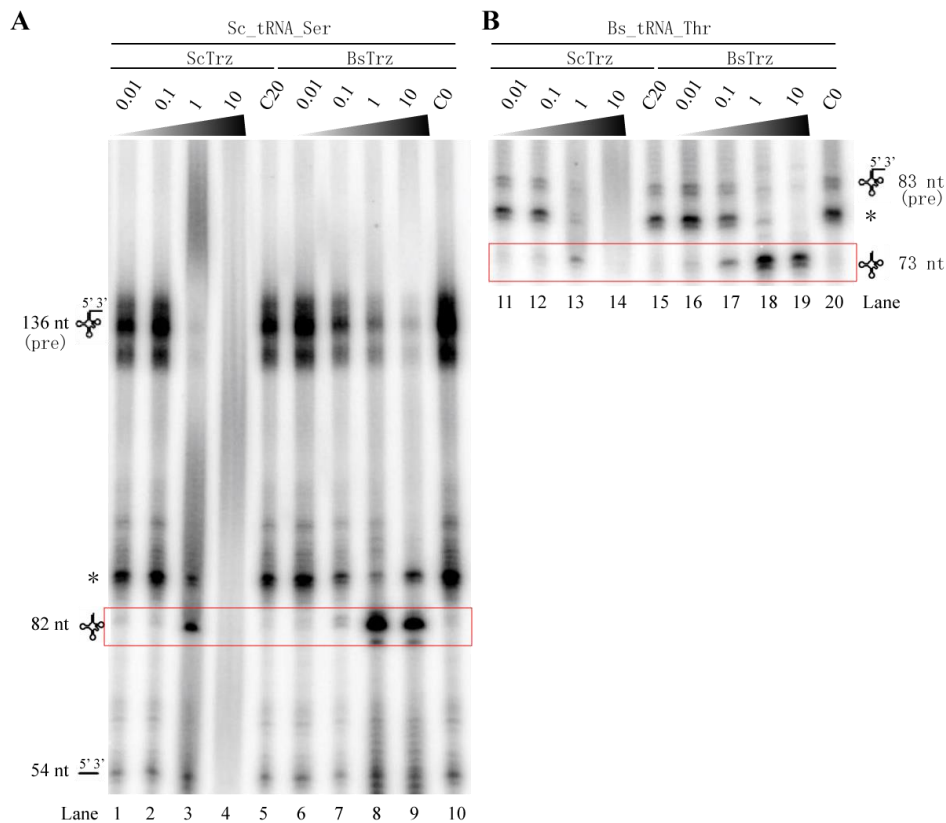


Figure III-3 Processing of tRNA from (A) *S. cerevisiae* and (B) *B. subtilis* by Trz1 (ScTrz) and BsuTrz *in vitro*. Precursors tRNAs with 54 and 10 nucleotide (nt) trailer sequences, respectively, were synthesized by T7 RNA polymerase *in vitro*. Precursor tRNAs (pre) and the cleavage product are symbolized and indicated with corresponding sizes at the side of the autoradiograms. The triangle indicates increasing amounts of RNase Z (0.01, 0.1, 1, 10 pmol, respectively) added per 5 μ l reaction. Radioactively labeled tRNA substrates were incubated with RNase Zs at 37 $^{\circ}$ C for 20 min. C20: The control which reaction buffer is used instead of enzyme. C0: The control same as C20 except the reaction was stopped immediately without incubation for 20 min. The RNA species indicating by asterisk are due to pre-mature transcription pausing of T7 RNA polymerase approximately 6 nt downstream of mature tRNAs.

III.1.2.2 Trz1 activity in presence of Nuc1 and mutarotase.

We then wanted to figure out whether the RNase Z 3'-processing activity towards the pre-tRNA substrate would be affected when Trz1 is in complex with Nuc1 and mutarotase.

We mixed the substrate with various amounts of complexes, while keeping the Trz1 concentrations at a constant 0.6 pmol in each reaction mixture. We varied the Nuc1 or mutarotase concentrations between 0.06 pmol and 6 pmol (Figure III-4). To reconstitute the complexes, the three enzymes were mixed in various combinations and incubated on ice at least 30 min before activity test.

0.6 pmol of Trz1 alone clearly cleaved the 82 nt mature tRNA species (Figure III-4A, lanes 2, 6, 10 and 14, multiple duplicate lanes were set up in order to facilitate comparison with other lanes). Addition of increasing amounts of Nuc1 (from 0.06 pmol to 0.6 pmol) did

not seem to affect the amount of mature tRNA generated by Trz1 (lanes 3-4) compared to Trz1 alone (lane 2). The same observation held for Trz1 in the presence of both Nuc1 and mutarotase (lanes 6-8): the ternary Trz1/Nuc1/mutarotase complex had the similar activity as Trz1 alone (lanes 6 and 8). Logically the addition of mutarotase alone did not change Trz1 activity either (lanes 10-12). The quantification of the mature tRNA (82 nt) shows that the amounts produced by either Trz1 alone or by the Trz1-containing complexes are of the same order of magnitude (Figure III-4 B). At 6 pmol of Nuc1 or mutarotase in combination with Trz1, the bands for pre-tRNA disappeared into a smear at the top of the lanes probably due to complex formation with the substrates, so we were not convinced by the slightly more intense of mature tRNA lower in the lanes (lanes 5, 8 and 11).

Designed as negative control, the tRNase activities of Nuc1 and mutarotase alone were tested at the highest concentrations that were used in their combination with Trz1 (6 pmol). Logically mutarotase alone did not cleave pre-tRNA (lane 16), although minute quantities of mature tRNA were observed which were probably caused by a previously described low-level auto-cleavage of the substrates. Interestingly, Nuc1 digested the pre-tRNA into multiple fragments (lane 15), explained by the Nuc1 DNA/RNA non-specific endonuclease activity (Schäfer et al. 2004). More interestingly, the tRNase activity of Nuc1 was inhibited by mutarotase (lane 15 comparing to lane 17), consistent with our finding that also the DNase Nuc1 activity is inhibited by mutarotase (details later in the chapter III.2). The same held for Trz1: the fragmentation of tRNA caused by Nuc1 was also lost when Trz1 and Nuc1 were both present (lanes 3-5).

Globally the activities of Trz1 and Trz1 partners were similar when we used the *B. subtilis* pre-tRNA as substrate (Figure III-4C).

From these findings we conclude that Trz1 3'-processing of pre-tRNA activity is not significantly affected by Nuc1 or mutarotase in the Trz1/Nuc1 binary or Trz1/Nuc1/mutarotase ternary complexes.

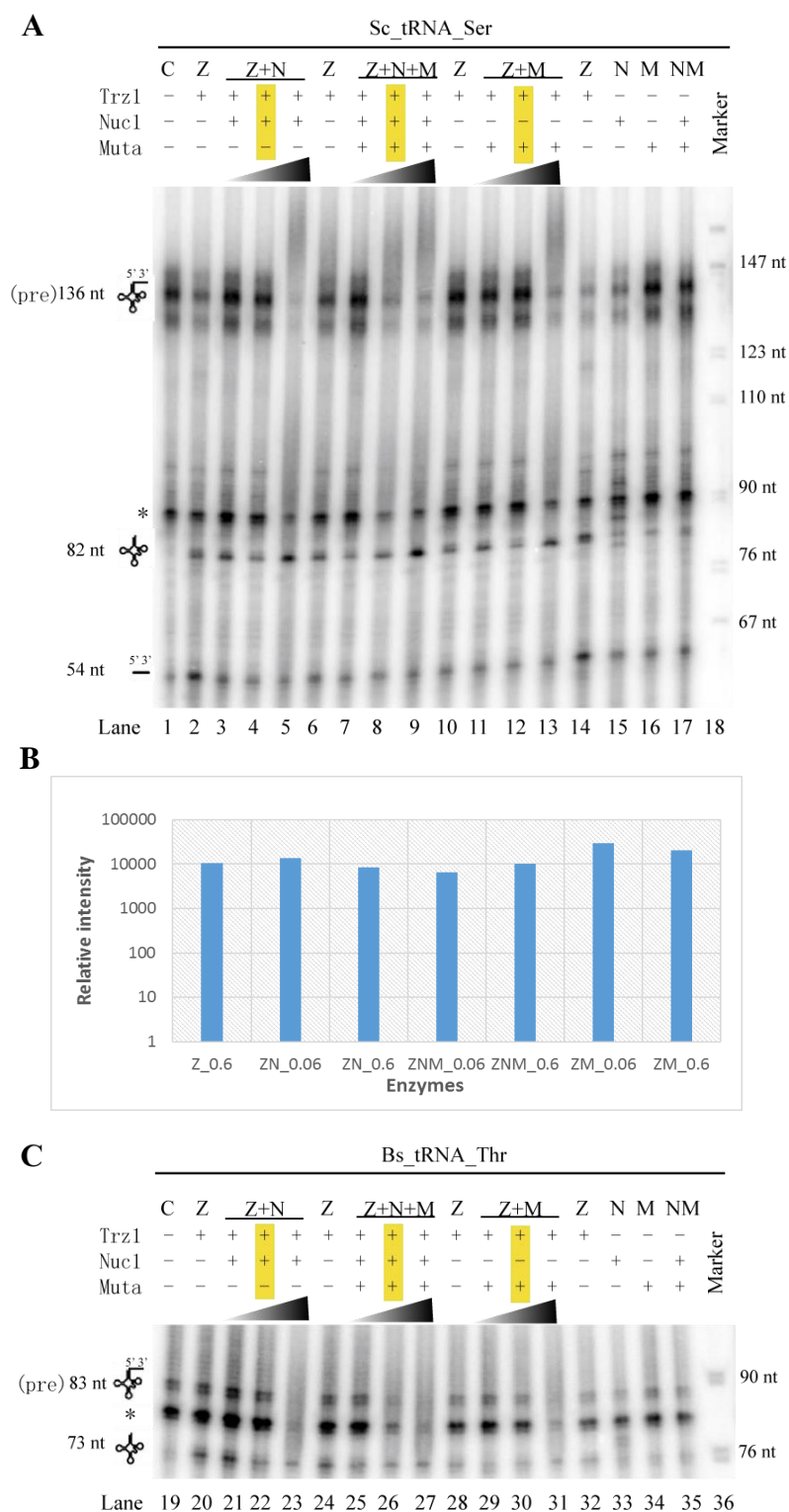


Figure III-4 Trz1 processing tRNAs in presence of titrated Nuc1, mutarotase or both.

Precursor tRNA from (A) *S. cerevisiae* (Sc_tRNA_Ser) or (C) *B. subtilis* (Bs_tRNA_Thr) were processed by various combinations of enzymes in vitro. Precursor tRNAs (pre) and the cleavage products are symbolized and with their corresponding size next to the autoradiograms. Radioactively labeled precursor tRNA substrates were incubated with enzymes at 37 °C for 20 min. Z: Trz1; N: Nuc1; M(muta): mutarotase. Trz1: constant 0.6 pmol. The triangle indicates increasing amounts of Nuc1 or mutarotase, with 0.06, 0.6, 6 pmol added per 5 µl reaction. The yellow frame indicates various enzymes at the same molar concentration. N, M, NM in lanes 15-17, 33-35: 6 pmol. C: control, reaction buffer was added in the system instead of enzymes. Marker: pBR 322 digested by MspI and 5'-P³² labeled. The RNA species indicating by asterisk are due to pre-mature transcription pausing of

T7 RNA polymerase approximately 6 nucleotides downstream of mature tRNAs. Experiment were repeated twice. (B) Histogram showing quantification of the cleavage product (82 nt) shown in (A). The intensity of the 82 nt bands from all lanes were measured by Image J, then the obtained value of the control lane was subtracted from values of the other lanes, and shown in relative values. The relative intensity of Trz1 alone (Z_6) is the average of that of lanes 2, 6 and 10.

III.2 Nuc1 nuclease activity and the effects of Trz1/mutarotase

III.2.1 Nuc1 nuclease activity towards dsDNA

Nuc1 has nuclease activity on DNA or RNA and we wanted to test the activity of Nuc1 alone versus that of the Nuc1 binary and ternary protein complexes. We used the plasmid pUC18 dsDNA as substrate as described for yeast Nuc1 or its homolog EndoG, and used similar reaction conditions as reported (37 °C in buffer 50 mM Tris-HCl, pH 7.0, 5 mM MgCl₂, 1 mM DTT) (Wu et al. 2009; Dake et al. 1988; Schäfer et al. 2004). Freshly extracted pUC18 contained mainly supercoiled DNA (Figure III-5A, Lane “0”), which were digested by Nuc1 first into nicked open circular and then into linear forms, and then the plasmid was finally degraded. Starting from 99% of supercoiled DNA at T₀ (sc), after one minute about 30% converted to open circular (oc) and trace of linear (line) form; after 10 minutes about 90% of DNA was converted in open circular form and linear form reached its intensity peak. Afterwards the open circular was cleaved to linear, and linear was consumed by degradation shown as loss of band intensity and a smear on the agarose gel. This indicates that, like its mammalian homolog EndoG (Schäfer et al. 2004; Wu et al. 2009; Ohsato et al. 2002), Nuc1 tends to nick the dsDNA rather than creating double strand break, judging from the accumulation of open circular form of DNA. Increasing concentrations of Nuc1 accelerated the breakdown of dsDNA (Figure III-5B). These results proved that our recombinant Nuc1 could efficiently digest dsDNA in our reaction system, and the conversion of dsDNA is Nuc1 dependent.

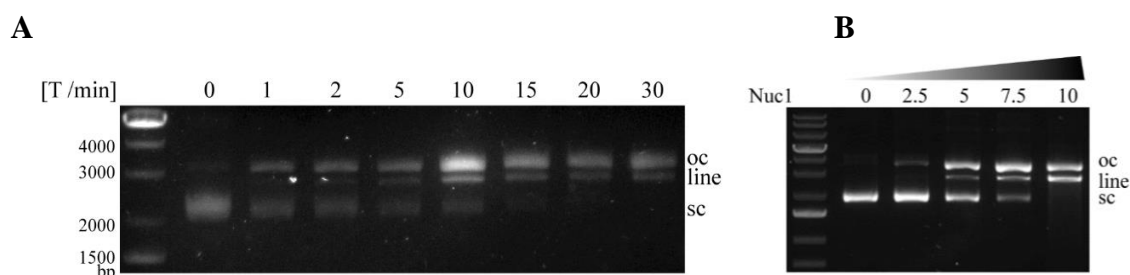


Figure III-5 Kinetics of Nuc1 nuclease activity.

(A) Conversion of dsDNA by Nuc1 against time. Substrate: plasmid pUC18 (2686 bp), converting from supercoiled (sc) to nicked open circular (oc) and linear (line). 50 µl of mixture containing enzyme and substrates

in reaction buffer were incubated at 37 °C, at each time point 5 µl aliquots (containing 6 pmol enzyme and 300 ng dsDNA) were removed from the mixture and stopped by addition of 1 µl loading dye containing 60 mM EDTA. Reaction buffer: 50 mM Tris-HCl, pH 7.0, 5 mM MgCl₂, 1 mM DTT. Samples were electrophoretically separated over a 1% (w/v) agarose gel and stained with SYBR. (B) Conversion of dsDNA by Nuc1 against enzyme concentration. Enzyme unit: pmol. 1 µl of enzyme in various concentrations was added to 9 µl pre-mixture containing 200 ng dsDNA in reaction buffer, and incubated at 37 °C for 30 min. Reaction buffer is the same as in (A).

We then investigated whether Nuc1 DNase activity would be affected by complex formation with Trz1 and mutarotase. We reconstituted the Nuc1 binary and ternary complexes to compare their DNase activity with that of Nuc1 alone (Figure III-6). The Nuc1 concentration was kept at a constant 5 pmol (per 50 µl reaction), and the protein partner(s) was added at increasing concentrations (1.25 pmol, 2.5 pmol, 5 pmol and 10 pmol) and the mixtures were incubated on ice for at least 30 min to reconstitute the complex. The enzymes were then mixed with dsDNA and incubated on 37 °C for 30 min.

Degradation of dsDNA accelerated with increasing amounts of Nuc1 (Figure III-6, lanes 2-5). Surprisingly, addition of mutarotase completely inhibited DNase activity (lanes 6-9), even at only 1:4 mutarotase to Nuc1 ratio (lane 6). Consistent with this observation, the tertiary Trz1/Nuc1/mutarotase complex was also devoid of DNase activity (lanes 14-17). On the other hand, the presence of Trz1 inhibited Nuc1 activity about 2-fold: 5 pmol of Trz1/Nuc1 had approximately the same activity as 2.5 pmol of Nuc1 (lanes 10-12).

From these observations we concluded that Nuc1 DNase activity is markedly suppressed in both the Nuc1/mutarotase binary and Trz1/ Nuc1/mutarotase ternary complexes; and partially suppressed in the Trz1/Nuc1 complex.

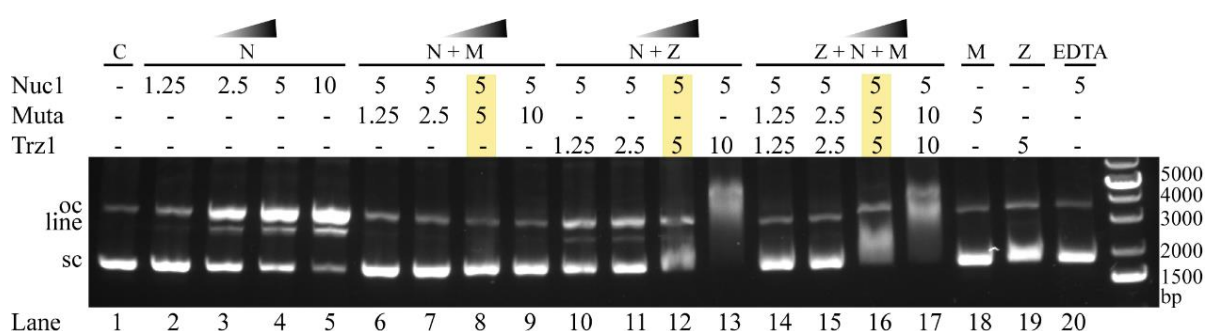


Figure III-6 DNase activity of Nuc1 in presence of Trz1, mutarotase or both.

Enzyme unit: pmol. Enzymes with various combinations were incubated on ice for 1 hour to reconstitute the complex. Then the 3 µl enzymes were added to 7 µl pre-mixture containing 200 ng dsDNA and reaction buffer, and incubated at 37 °C for 30 min. C: control, reaction without enzyme. EDTA: 10 mM EDTA added in the reaction. Reaction buffer: 50 mM Tris-HCl, pH 7.0, 5 mM MgCl₂, 1 mM DTT. The lanes in yellow frame indicate enzymes are in equal molar. Experiment were repeated four times.

We further performed the kinetic assays with gel-filtration-purified complexes instead of reconstituted complexes (Figure III-7). From the gels we could tell that with equal amount of enzyme (50 pmol) and substrate (2 μ g) in the 100 μ l reaction mixture, Nuc1 efficiently digested dsDNA starting at 1 min, only 5% of DNA remained in supercoiled form at 30 min (Figure III-7 A); Trz1/Nuc1 started to digest dsDNA at 5 min, but the DNA pattern evolved very slowly until the end (Figure III-7 B); For Nuc1/mutarotase and the ternary complex, dsDNA was hardly digested from the beginning to the end (Figure III-7 C and D). This further proves that Trz1/Nuc1 is less active than Nuc1 on dsDNA, whereas Nuc1/Mutarotase and Trz1/Nuc1/Mutarotase almost have essentially no activity on dsDNA.

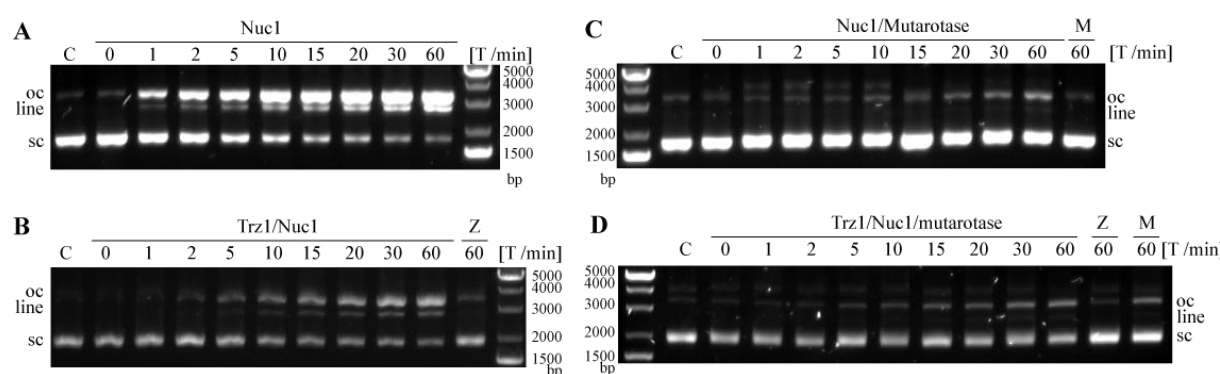


Figure III-7 Kinetics of DNase activity of Nuc1 and related complexes.

Z: Trz1; N: Nuc1; M: mutarotase. Binary complex ZN and NM was pre-purified by gel filtration, and ternary complex ZNM was reconstituted by mixing Z and NM. 100 μ l reaction mixture containing 50 pmol of enzyme and 2 μ g plasmid DNA (pUC18) in reaction buffer (50 mM Tris-HCl, pH 7.0, 5 mM MgCl₂, 1 mM DTT) were incubated at 37 °C. 10 μ l aliquots (containing 5 pmol enzyme and 200 ng dsDNA) were taken at different time intervals, and stopped by addition of 2 μ l loading dye containing 60 mM EDTA. Samples were electrophoretically separated over a 1% (w/v) agarose gel and stained with SYBR. C: control, protein buffer A is used instead of enzyme.

We noticed that enzyme combinations containing 10 pmol of Trz1 led to smearing and slower migration of the dsDNA (Figure III-6, lanes 13 and 17). We performed a gel migration test to see if this was due to Trz1 only. Various amounts of Trz1 were mixed with dsDNA for 30 min, and treated in three ways: I) immediately adding of the loading dye; II) incubating at 37 °C for 30 min and then adding loading dye, which was the same treatment as in the Nuc1 DNase activity assay; III) the same treatment as in II) but the sample was boiled and centrifuged (Figure III-8). A similar smear as in the Nuc1 activity assay was observed when Trz1 was at 10 pmol in conditions I and II; and this smear effect could be eliminated by boiling and centrifuging the sample before analysis (condition III). This indicates that the smear was caused by trapping of dsDNA by Trz1, precluding the use of higher Trz1 concentrations in the DNase activity tests.

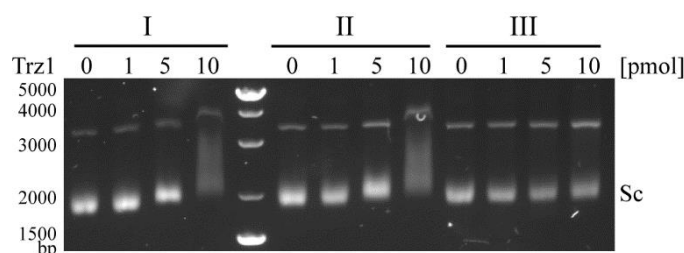


Figure III-8 Trz1 affects dsDNA migration.

1 μ l of Trz1 in various concentration plus 2 μ l protein buffer A was added to 7 μ l mixture containing 200 ng dsDNA and reaction buffer, and incubate on ice for 30 min, which is the same reaction system as Nuc1 DNase activity assay. I. 2ul loading dye immediately added; II. incubation at 37 $^{\circ}$ C for 30 min and 2 μ l of loading dye added, mimicking Nuc1 DNase activity assay; III. incubation at 37 $^{\circ}$ C for 30 min and 2 μ l of loading dye added, then boiled for 5 min and centrifuged at 13000 rpm for 1 min. Samples were electrophoretically separated over a 1% (w/v) agarose gel and stained with SYBR.

III.2.2 Nuc1 nuclease activity towards fluorescent hairpin DNA

To better quantify the effect of Trz1 and mutarotase on Nuc1 activity, we synthesized a fluorescent beacon substrate as described in (Kieper et al. 2010). The molecular beacon was labeled at the 5'-end with 6-FAM (6-Carboxyfluoresceine) and at the 3'-end with BHQ-1 (Black Hole Quencher-1), and forms a hairpin structure (Figure III-9 A). Upon cleavage by Nuc1 at several sites, the quencher BHQ-1 becomes separated from the fluorescent dye 6-FAM, allowing fluorescence release and signal detection. This artificial substrate is a convenient way to monitor the activities in real time.

Preliminary experiments showed that at 50 nM substrate concentration, Nuc1 induced fluorescence increased significantly faster at 5 μ M than at 1 μ M enzyme concentration (Figure III-9 B).

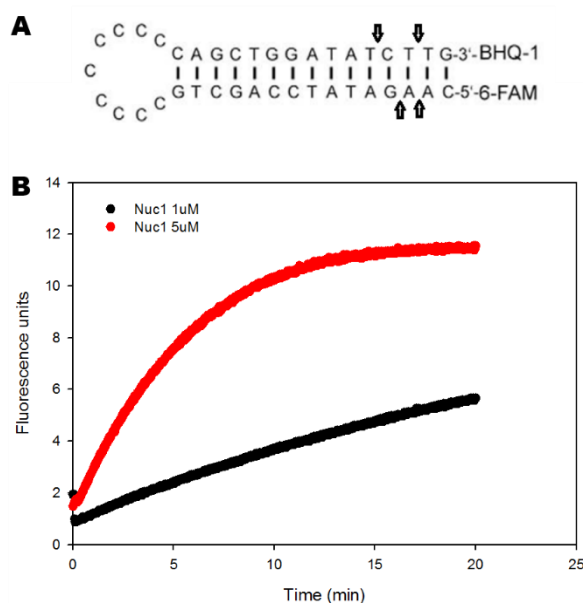
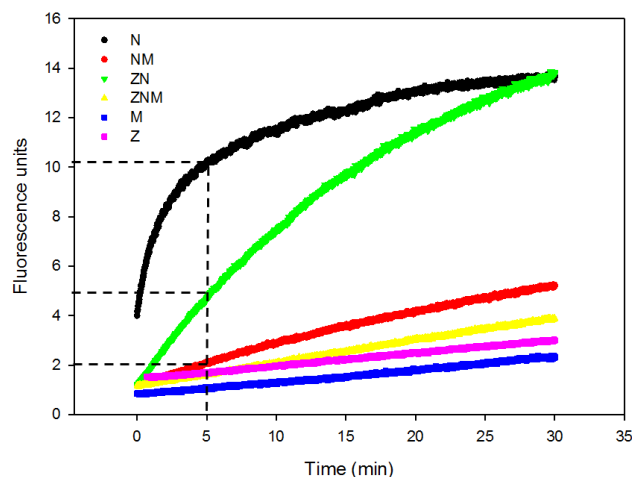


Figure III-9 Nuclease activity of Nuc1 towards fluorescent hairpin DNA.

(A) Fluorescent molecular beacon used as substrate for Nuc1. The arrows indicate Nuc1 cleavage sites. (B) Progress curves (fluorescence units vs. time) for the cleavage of the substrate (50 nM) by Nuc1 at 1 μ M or 5 μ M, respectively. Reaction volume: 200 μ l. buffer: 50 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 1 mM DTT. 25 °C. Experiments were repeated twice.

**Figure III-10 Nuclease activity of Nuc1 and related complexes against the fluorescent hairpin DNA.**

Progress curves (fluorescence units vs. time) for the cleavage of fluorescent hairpin DNA (50 nM) by Nuc1 and related complexes at 5 μ M. Reaction buffer: 50mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 1 mM DTT. 25 °C. Experiments were repeated three times.

Then we performed activity assays with the three Nuc1-containing complexes: Trz1/Nuc1, Nuc1/mutarotase, Trz1/Nuc1/mutarotase. The complexes were reconstituted by mixing single proteins in equimolar concentrations and incubated on ice for more than 30 min. With 50 nM substrate, activities of various complexes at 5 μ M in a 200 μ l reaction mixture were measured (Figure III-10). The progress curves of the fluorescence signal were recorded. Nuc1 alone had the highest activity (black curve) and was inhibited when present in the Trz1/Nuc1 binary complex (green curve). For the other binary complex, Nuc1/mutarotase, the rate of fluorescence increase dropped much more (red curve) and for the ternary Trz1/Nuc1/mutarotase complex, the fluorescence increase was even slower (yellow curve). As the curves of various samples could not be fitted with a uniform equation for comparing their catalysis constant, we simply took the time point 5 min to quantitatively compare the activities of various enzyme combinations: Nuc1 cleaved approximately 70% of substrate; the Trz1/Nuc1 binary complex cleaved about 36% (5 of total 14 fluorescence units) and the Nuc1/mutarotase and the ternary Trz1/Nuc1/mutarotase complex less than 14%.

This indicates Nuc1 nuclease activity is partially inhibited by Trz1; but totally inhibited by mutarotase in the presence or absence of Trz1, which is consistent with our previous findings using plasmid dsDNA as substrate.

III.3 Mutarotase activity

We finally also tested whether the mutarotase activity would be affected by the formation of complexes with Trz1 and Nuc1. Mutarotase catalyzes the conversion of α and β glucose-6-phosphate (Glc6P), although in itself this is a spontaneous process. We used a coupled-enzyme assay to measure mutarotase activity (Graille et al. 2006). Glucose-6-phosphate dehydrogenase converts β -Glc6P (but not α -Glc6P) into the keto form and the H^+ released is used to generate NADPH from $NADP^+$. Both spontaneous and mutarotase-catalyzed formation of β -Glc6P is converted to D-glucono-1,5 6-phosphate in the presence of glucose-6-phosphate dehydrogenase; coupled to the generation of NADPH ($\lambda = 340$ nm) from $NADP^+$, and increasing absorbance at 340 nm can be measured (Figure III-11). Therefore, a faster increase of absorbance at 340 nm indicates a faster formation of β -Glc6P resulting from Glc6P mutarotase activity.

We measured the conversion of α -Glc6P into β -Glc6P with varying concentrations of mutarotase and the related binary and ternary complexes using stopped-flow. 30 μ M equilibrated Glc6P were mixed with 75 units glucose-6-phosphate dehydrogenase from *L. mesenteroides* and various concentrations of mutarotase (or its complexes) in equal volume while 1 mM $NADP^+$ was present in both syringes (Figure III-11 B).

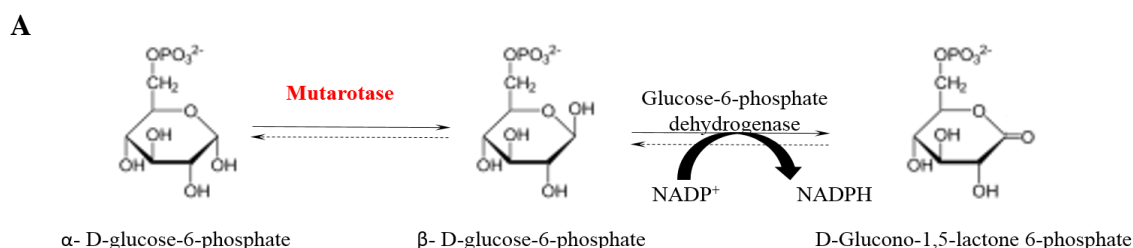
The absorbance progress curve revealed two distinct phases (Figure III-12 A). The first burst phase corresponds to the consumption of β -Glc6P initially present in the equilibrated racemic mixture; the subsequent slower absorbance increase depends on the rate of β -Glc6P formation catalyzed by mutarotase. The progress curve was fitted by a double exponential function:

$$f(t)_{absorbance} = a_1 * \exp(-k_1 t) + a_2 * \exp(-k_2 t) + c$$

in which the k_1 constant corresponds to the Glc6P dehydrogenase reaction, remaining constant because the enzyme is always in large excess; and the constant k_2 describes the second phase which depends on the mutarotase activity.

The progress curves for various mutarotase/complexes concentrations were fitted according to the above equation and the values for k_2 were extracted (Table III-3, Figure III-12 B). The data show that the value of k_2 is roughly proportional to the mutarotase concentration. When present at the same concentrations (27 nM or 54 nM), analysis of the kinetic curves for mutarotase, Nuc1/mutarotase and Trz1/Nuc1/mutarotase yielded the same

value for k_2 , indicating that complex formation does not seem to affect the catalytic capacity of mutarotase. Nuc1 and Trz1 alone did not display any mutarotase activity. These findings indicate that mutarotase activity on Glc6P is not affected by complex formation with Trz1 and Nuc1.



B

Chamber 1 (Substrates)	Chamber 2 (Enzymes)
α/β – glucose-6-phosphate (equilibrated) 30 μ M	glucose-6-phosphate dehydrogenase (75 units/ml - excess)
NADP ⁺ 1mM	NADP ⁺ 1mM
--	Mutarotase*

* Variable factor.

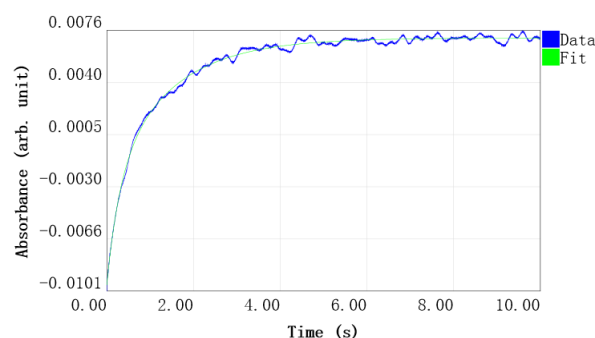
Figure III-11 Coupled enzyme assay for measuring mutarotase activity (adapted from Marc G. *et al*, 2006). (A) Scheme representation of the coupled- enzyme reaction. (B) Experiment setup of the Stopped-flow.

Table III-3 Observed catalytic constant k_2 of mutarotase-complexes in mutarotase activity assay

Enzyme	Concentration (nM)	k_2 (min ⁻¹)*	SD
ctrl	0	3.992	0.019799
Mutarotase	27	26.60245	1.366156
	54	42.29581	3.136667
Nuc1/mutarotase	27	27.55326	0.980691
	54	48.70143	2.444533
Trz1/Nuc1/mutarotase	27	26.55427	0.838572
	54	44.62805	2.283217
Nuc1	54	4.415775	0.239222
Trz1	54	4.32861	0.258375

* k_2 was the mean value of at least four times repeat. SD: standard deviation.

A



B

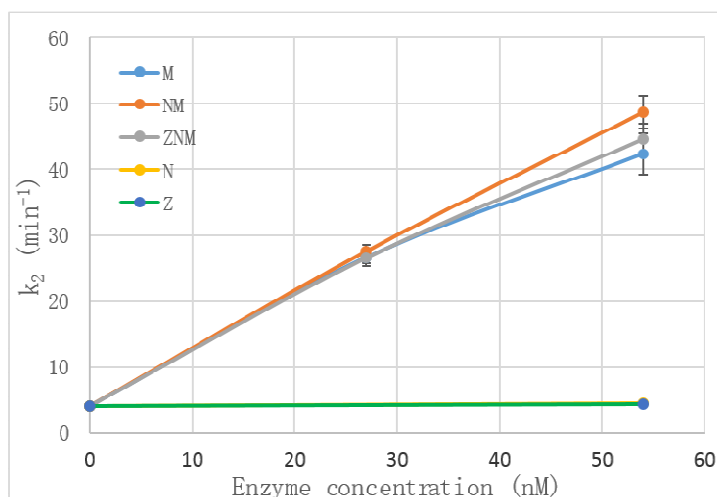


Figure III-12 Mutarotase and containing complex activity assay.

Experiments were performed using the Stopped-flow at 25 °C in buffer 50 mM imidazole, pH 7.5, 50 mM KCl, 8 mM MgSO₄. Reactions were initiated by mixing equal volumes of 30 μM equilibrated Glc6P and 1mM NADP⁺ with 75 units glucose-6-phosphate dehydrogenase, 1 mM NADP⁺, and various concentration of mutarotase or containing complexes. (A) Profile of the assay when mutarotase is 54 nM. The curve is fitted by double exponential function. (B) Plotted k_2 in function of enzyme concentration, data from Table III-3. Z: Trz1; N: Nuc1; M: mutarotase.

III.4 Summary of the activity assays

The results of activity assays towards each substrate of Trz1, Nuc1 and mutarotase demonstrate that Trz1 or mutarotase activity are not affected when they form complexes with the other two proteins; whereas Nuc1 nuclease activity is significantly inhibited in both binary complexes and the ternary complex.

IV. Crystallization of the binary and ternary complexes

Since we've confirmed the existence of Trz1, Nuc1 and mutarotase related complexes and able to purified them in large quantities, studied their enzyme activities towards various substrates, surely we're interested in solving the structures of the complexes. We performed crystallization experiment aiming to resolve the structures of the complexes.

Initial crystallization screenings of the complexes were performed by the Cartesian robot. 100 nL of protein samples were mixed with 100 nL of the crystallization cocktail solution (Qiagen Protein Crystallization Kits) as sitting drops on the TTP 96 well plates. Screening plates were stored at 18 °C or 6 °C and inspected regularly using a Formulatrix visualization robot. All the initial screenings were using freshly prepared protein samples, while optimization of the conditions depends. Proteins were always in classic buffer unless stated otherwise.

IV.1 Crystallization of the Trz1/Nuc1/mutarotase complex

The Trz1/Nuc1/mutarotase complex used for crystallization screening was purified from the polycistronic plasmid as shown in previous chapter II.5.2. Fresh proteins at three different protein concentrations (4 mg/ml, 7.5 mg/ml and 15 mg/ml) solubilized in buffer A were screened with various crystallization kit. The screening plates were incubated at 18 °C or 6 °C (Table IV-1 and Table IV-2).

Considering Nuc1 binds Mg^{2+} , 1 mM of $MgCl_2$ was added to the proteins. 1 mM of glucose-6-phosphate (Glc6P), which is a substrate of mutarotase, was also used as an option.

Table IV-1 Crystallization conditions for the Trz1/Nuc1/mutarotase complex at 6°C.

6°C	Kits	[c](mg/ml)	[c] (mg/ml)
	Classic	5.5	--
	PEGs	5.5	--
	PEGs II	5.5	--
	pH clear	--	7.5
	pH clearII	5.5	--
	AmSO4	5.5	--
	Nucleix	--	7.5
	Protein complex	--	7.5

Mb classic I	--	7.5
Mb classic I	--	7.5 + MgCl ₂ + Glc6P
JCSG+	--	7.5 + MgCl ₂ + Glc6P

MgCl₂: 1 mM ; Glc6P: Glucose-6-phosphate, 1 mM.

Crystals were obtained for only in one condition: 22% (w/v) polyacrylic acid 5100 sodium salt, 0.2 M MgCl₂, 0.1 M Hepes pH 7.5; the protein was at 7.8 mg/ml mixed with 1 mM MgCl₂ and Glc6P (Figure IV-1). Crystals appears at the 2nd day and did not growing further. The crystals were too tiny to fish and diffract, but reproducing of the crystals failed either with commercial solutions from the Kit or homemade solutions containing gradient polyacrylic acid. Even macro-seeding with these tiny crystals as seeds didn't work, only precipitates formed around the tiny crystals.

Then we performed another batch of screening with lower salt concentration in protein buffer (100 mM instead of 200 mM NaCl), gaining only salt crystals (data not shown).

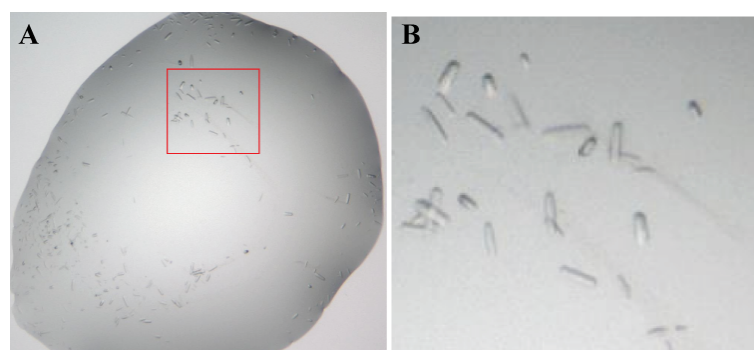


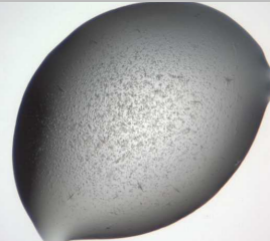
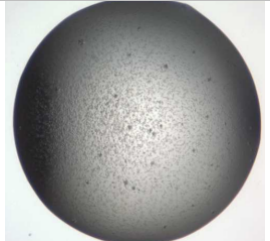
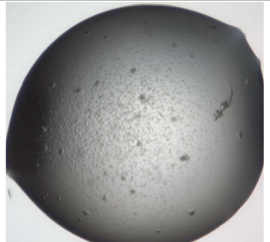
Figure IV-1 Image of hits from Trz1/Nuc1/mutarotase screening.

Protein: Trz1/Nuc1/mutarotase 7.8 mg/ml plus 1 mM MgCl₂ plus 1 mM Glc6P. 18 °C. Condition: 22% (w/v) polyacrylic acid 5100 sodium salt, 0.2 M MgCl₂, 0.1 M Hepes pH 7.5. (A) Full drop. (B) zoomed area of red square in (A).

Table IV-2 Crystallization conditions for the Trz1/Nuc1 complex at 18°C.

18 °C	Kit	[c] (mg/ml)	[c] (mg/ml)
ttp0335	Classic lite	5.4	2.7
ttp0336	PEGs	5.4	2.7
ttp0337	PEGs II	5.4	--
ttp0341	JCSG+	5.5	--
ttp0340	Mb classic I	5.5	--
ttp0363	Mb classic II	5.4	--
ttp0362	AmSO4	5.4	--

Table IV-3 Potential crystallization hits for the Trz1/Nuc1 complex

Pr.	[c](mg/ml)	Crystallization conditons	No.	image	Image date
Trz1/Nuc1	5.5 mg/ml	0.1 M bis-Tris pH 5.5, 2 M ammonium sulfate [ttp0341, JCSG+,G11]	1#		Day 5
Trz1/Nuc1	5.4 mg/ml	0.1 M sodium citrate pH 5.6, 2 M ammonium sulfate. [ttp0340, MbClass A09]	2#		Day 5
Trz1/Nuc1	5.4 mg/ml	0.5 M lithium chloride, 1 M sodium citrate pH 5.6. [ttp0340, MbClass B02]	3#		Day 5

IV.2 Crystallization of the Trz1/Nuc1 binary complex.

The crystallization conditions tried for the binary Trz1/Nuc1 are gathered in Table IV-3. Microcrystallines were obtained for several conditions (Table IV-4). We tried to reproduce and optimize these crystals by varying the pH and gradient of the main precipitant, either with fresh or defrozen protein. But crystals were either too small or could not be reproduced.

IV.3 Crystallization of the Nuc1/mutarotase binary complex

The crystallization of the Nuc1/mutarotase complex was tested at various concentrations (Table IV-5). Crystallization hits are listed in Table IV-6. Giant crystals grew after 13 days in conditions No.1-3 but they proved to be mutarotase crystals (Figure IV-2 A). SDS-PAGE analysis of some crystallization drops shows that one of the two proteins gets degraded during the experiment (Figure IV-2 B). Mass spectrometry of the bands in the lane 3 showed that the protein “b” was intact mutarotase and “c” was degraded Nuc1.

Crystals in condition No.4 appeared after 6 days, earlier than for the other conditions, and with a different shape. By optimizing this condition by hand in hanging drops, we

obtained crystals diffracted up to 7.5 Å. SDS-PAGE analysis of these crystals confirmed they contain both Nuc1 and mutarotase (Figure IV-3 A).

Table IV-4 Crystallization screening of the Nuc1/Mutarotase complex at 18°C.

18°C	Kit	[c] (mg/ml)	[c] (mg/ml)	[c] (mg/ml)
	PEGs	18.4	9.2	4.5
	PEGs II	--	9.2	4.5
	JCSG+	--	9.2	4.5
	MbClassic	--	9.2	4.5

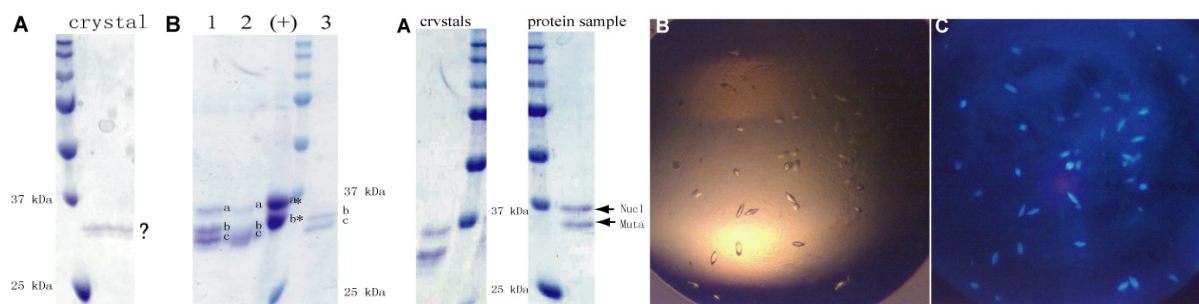


Figure IV-2

Figure IV-3




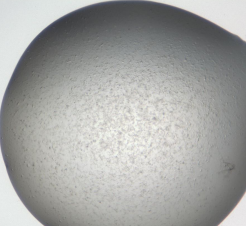
Figure IV-2 SDS-PAGE of (A) crystals (B) drops from Nuc1/mutarotase crystallization screening.

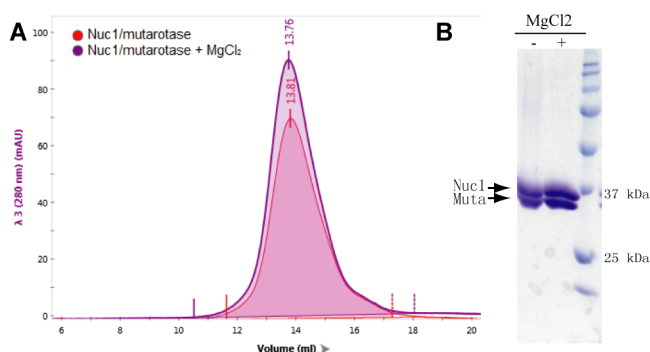
Protein: 9.2 mg/ml. (A) Crystals of 21 days from condition 0.1 M MES pH 6.5, 15% PEG 20K. (B) (+) Protein sample before crystallization. Lane 1, 2, 3: crystallization drops after 40 days. a*: Nuc1; b*: mutarotase. Protein marked with the same letter (a/b/c) in each lane migrate in the same speed. Condition of lane 1: 0.1 M MES pH 6.5, 25% PEG 6K; lane 2: 0.1 M Hepes pH 7.5, 15% PEG 10K. lane 3: 0.02 M bis-tris propane pH 7.0, 15% PEG 2K.

Figure IV-3 Crystals of Nuc1/mutarotase.

(A) Left: SDS-PAGE of the crystals fished from the drop; Right: protein sample before crystallization. (B) / (C) Hanging drop of optimization under microscopy / UV. drop volume: 2 µl. Protein: 5.5 mg/ml in buffer A. Condition: 0.1 M phosphate-citrate pH 5.02, 29% PEG 300. The crystallization plate was incubated at 18 °C.

Table IV-5 Hits of the crystallization screening of Nuc1/mutarotase

Pr.	[c](mg/ml)	Crystallization conditons	No.	image	appear date
Nuc1/mutarotase	18.4 /9.2	0.1M MES pH 6.5, 15% (w/v) PEG 20K. [PEGs C12]	1#		Day 13
Nuc1/mutarotase	9.2	0.02 M bis-tris propane pH 7, 15(w/v) PEG 2K. [MbClass D12]	2#		Day 13
Nuc1/mutarotase	9.2	0.025 M Potassium Phosphate, 12% PEG 8K, 10% MPD. [Mbclass G12]	3#		Day 35
Nuc1/mutarotase	9.2/4.6	0.1 M phosphate citrate pH 4.2, 40% (w/v) PEG 300. [JCSG+ C06]	4#		Day 6

**Figure IV-4 Nuc1/mutarotase in presence or absence MgCl₂.**

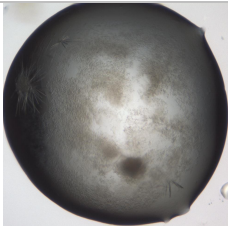
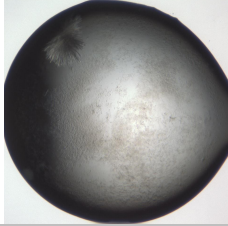
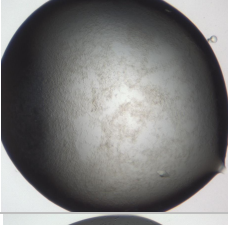
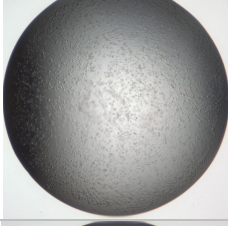
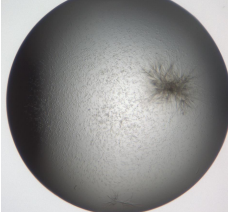
(A) Overlay of size exclusion chromatography. MgCl₂: 1 mM. Column: Superdex 200 10/300 GL Buffer A.

(B) SDS-PAGE of the fraction from the peak shown in (A).

Nuc1 is known to bind Mg²⁺, so we tested the effect of Mg²⁺ on the complex formation with mutarotase by gel filtration chromatography, expecting that the presence of Mg²⁺ would

stabilize Nuc1/mutarotase and prevent Nuc1 from degradation. Overlay of the curves of size exclusion chromatography in presence or absence of MgCl₂ showed no difference in the elution profiles (Figure IV-4). We performed a new crystallization screening of Nuc1/mutarotase complex in presence of MgCl₂ (Kit JCSG+ and PEGs). The potential crystallization hits are listed in (Table IV-7). The crystals from condition No.3 diffracted up to 7 Å. In coincidence, it's the same condition that in which best diffracting crystal of Nuc1/mutarotase without MgCl₂ (7.5 Å) grows, indicating MgCl₂ was not helpful for improve the diffraction of the crystal. Other hits were not optimized.

Table IV-6 Crystallization hits for the Nuc1/mutarotase complex in presence of MgCl₂

Pr.	[c] (mg/ml)	Crystallization conditons	No.	image	Image date	Appear date
Nuc1/muta + MgCl ₂	5.5	0.2 M potassium nitrate, 20% (w/v) PEG 3350. [ttp0462, JCSG+ A12]	1#		Day11	Day11
Nuc1/muta + MgCl ₂	5.5	0.2 M sodium thiocyanate, 20% (w/v) PEG 3350. [ttp0462, JCSG+ B02]	2#		Day11	Day11
Nuc1/muta + MgCl ₂	5.5	0.1 M tri-sodium acetate pH 4.5, 0.1 M bis-tris pH5.5, 25% PEG3350. [ttp0462, JCSG+ H03]	3#		Day11	Day1
Nuc1/muta + MgCl ₂	5.5	0.1 M phosphate citrate pH 4.2, 40% (w/v) PEG 300. [ttp0462, JCSG+ C06]	4#		Day21	Day 6
Nuc1/muta + MgCl ₂		0.16 M Calcium acetate, 0.08 M sodium cacodylate pH 6.5, 14.4% PEG 8000. [ttp0462, JCSG+ E11]	5#		Day11	Day2

V. Structure study of Trz1/Nuc1/mutarotase and related complexes by SAXS

Since we failed to obtain high-quality crystals of the binary or ternary complexes of Trz1, Nuc1 and mutarotase, we performed small angle X-ray scattering (SAXS) experiments on the complexes and the individual proteins to figure out how the partners assemble into complexes in solution.

V.1 Structural parameters and oligomerization state

We started by determining the structural parameters (R_g , radius of gyration, D_{max} , maximal extension) and molecular mass of the individual proteins and the complexes (Table V-1). For each sample, the radius of gyration (R_g) calculated from the *Guinier plot* is consistent with that calculated with the distance-distribution function ($P(r)$); and the molecular mass (MM) calculated by two methods are also consistent with each other, indicating the liability of the value of these parameters.

Table V-1 Scattering derived parameters: R_g , D_{max} , Molecular Mass, oligomerization state

	Z	N	M	ZN	NM	ZNM
Structural parameters						
R_g Guinier [Å]	32.6 ± 0.3	28.8 ± 0.4	20.4 ± 0.3	48.4 ± 1.1	39.6 ± 0.4	50.4 ± 0.4
R_g p(r) [Å]	33.0 ± 0.2	28.6 ± 0.3	20.5 ± 0.2	49.5 ± 0.3	40.0 ± 0.3	51.1 ± 0.3
D_{Max} [Å]	120	96	67	180	145	190
Molecular mass determination						
MM _{sequence} [kDa]	101.0	33.8	34.8	134.8	68.6	165.4
MM _{SAXS MoW} [kDa]	106 ± 10	71 ± 7	34 ± 3	272 ± 27	140 ± 14	363 ± 36
MM _{SAXS QR} [kDa]	103 ± 10	63 ± 6	32 ± 3	254 ± 25	129 ± 13	339 ± 34
Oligomerisation state						
	monomer	dimer	monomer	dimer	dimer	dimer

*Trz1 is His-Tagged, Nuc1 Strep-tagged and Mutarotase His-tagged. But the ZNM complex is purified from the polycistronic plasmid, where Z is tag-free.

For the individual proteins, the molecular weights and R_g values obtained from SAXS ascertain Trz1 and mutarotase are monomeric in solution while Nuc1 behaves as a homodimer. A dimeric form was also observed for the EndoG homologs from *C. elegans* (Bahi et al. 2006) and *D. melanogaster* (Loll et al. 2009). For the complexes, the molar mass values obtained by SAXS for the NM (139.9 kDa) and ZN (271.8 kDa) binary complexes suggest that they are composed of a Nuc1 dimer and two M or Z subunits, corresponding to calculated molecular

masses of 137.2 and 269.6 kDa respectively. The measured molar mass of the ternary ZNM complex (362.8 kDa) indicates that it forms a heterohexamer containing two copies of each subunit (calculated MM 330.8 kDa). These observations are in agreement with our MALLS and gel filtration measurements.

V.2 Modeling of individual proteins using SAXS data

The crystal structures of Trz1 (this work) and mutarotase (PDB code 2CIQ) are available for calculating the SAXS curves and comparing them to the experimental data. For Nuc1 a confident atomic model was generated by the webserver Phyre2 (Kelley et al. 2015) using the crystal structure of DmEndoG as template (PDB code 3ISM, 45% sequence identity).

As several residues are missing in the crystal structure of Trz1, the challenge is to generate a model of Trz1 containing all residues, which fits the experimental SAXS data. For the adjustment of the calculated scattering curve of Trz1 against the experimental data, we generated conformations of the parts which are missing in the crystal structure (the 35 N-terminal residues, the loops 205-245 and 624-635, and the 20 C-terminal residues). We therefore used the program AllosMod-FoXS, which generates the most probable conformations consistent with the input crystal structure. Figure V-1 shows three typical conformations of Trz1 that fit well the scattering data ($\chi = 1.25$; fit performed using the program CRY SOL). The superposition of these structures is consistent with a high mobility of the flexible arm in the N-terminal domain.

Since data from SAXS and MALLS indicate that Nuc1 forms a dimer in solution, we constructed a dimer model based on the Phyre2 model. The curve calculated on this model using program CRY SOL is in good agreement with experimental data ($\chi = 1.78$, Figure V-2).

The SAXS data for the mutarotase can also be satisfactorily described from the crystal structure (PDB code 2CIQ) and modeling of the missing 15 C-terminal residues using AllosMod-FoXS (Figure V-3, $\chi = 1.64$).

We conclude that the SAXS curves of the individual Trz1 and mutarotase proteins are in good agreement with their crystal structures. The SAXS data suggests also that Nuc1 adopts a dimer structure in solution similar to that of its fruit fly homolog.

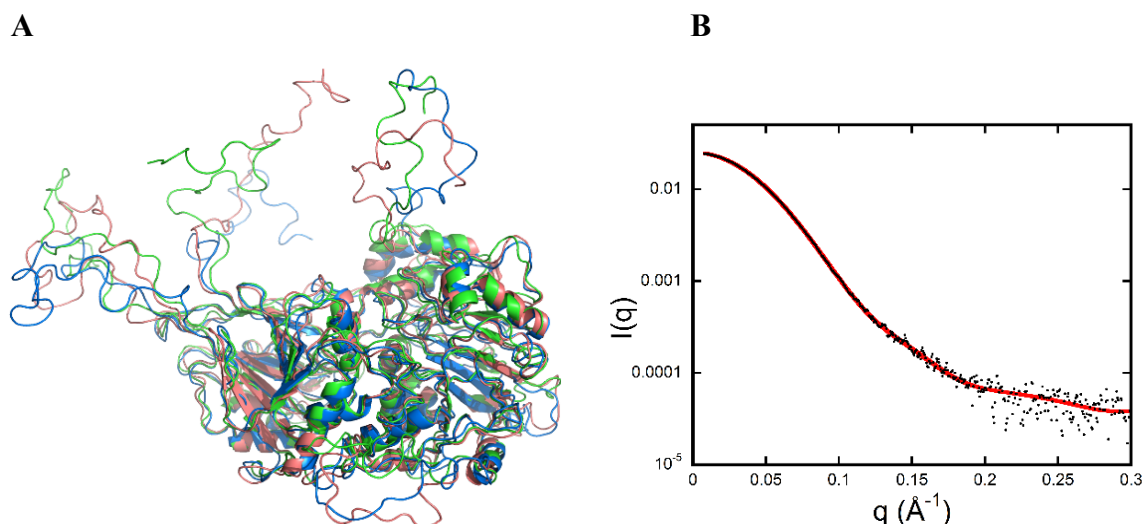


Figure V-1 SAXS analysis of Trz1. (A) Superimposition of three typical conformations of Trz1 obtained from scattering data using the program AllosMod-Foxs. (B) The calculated scattering curve for the pink model in (A) using the program CRY SOL (continuous red line); experimental scattering data are represented as black spots ($\chi = 1.23$).

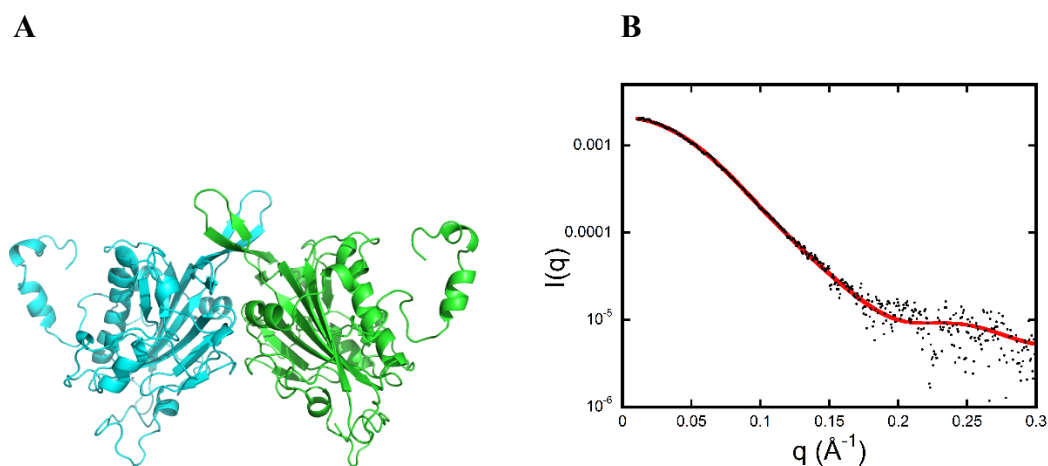


Figure V-2 SAXS analysis of Nuc1. (A) Model of the Nuc1 dimer generated by the webserver Phyre2 (Kelley et al. 2015) using the crystal structure of DmEndoG as template (PDB code 3ISM). (B) The calculated curve using the program CRY SOL (continuous red line) running on the model in (A) and the experimental scattering curve (black spots), $\chi = 1.78$.

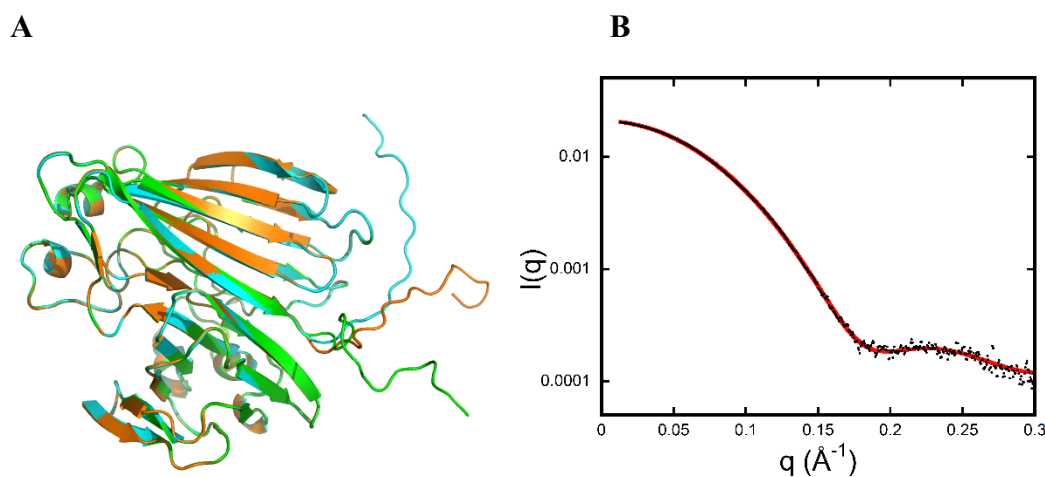


Figure V-3 SAXS analysis of mutarotase. (A) Superimposition of three typical conformations of mutarotase obtained from scattering data using the program AllosMod-Foxs. (B) The calculated curve using the program CRY SOL (continuous red line) running on the orange model in (A) and the experimental scattering data (black spots), $\chi = 1.64$.

V.3 Analysis of the complexes in solution by SAXS

V.3.1 Analysis of the NM complex.

In the absence of any biochemical or structural information about the interaction surfaces between Nuc1 and mutarotase, the first step of the analysis of SAXS curves collected on the NM complex consisted in restoring a low-resolution *ab initio* envelope of the complexes using the program DAMMIN. Since we have good reasons to assume that Nuc1 forms a symmetric dimer, we imposed a two-fold symmetry on the complex. A few tens of DAMMIN runs on the binary NM complex yielded very similar models (NSD = 0.68) all with a prolate shape (Figure V-4 C).

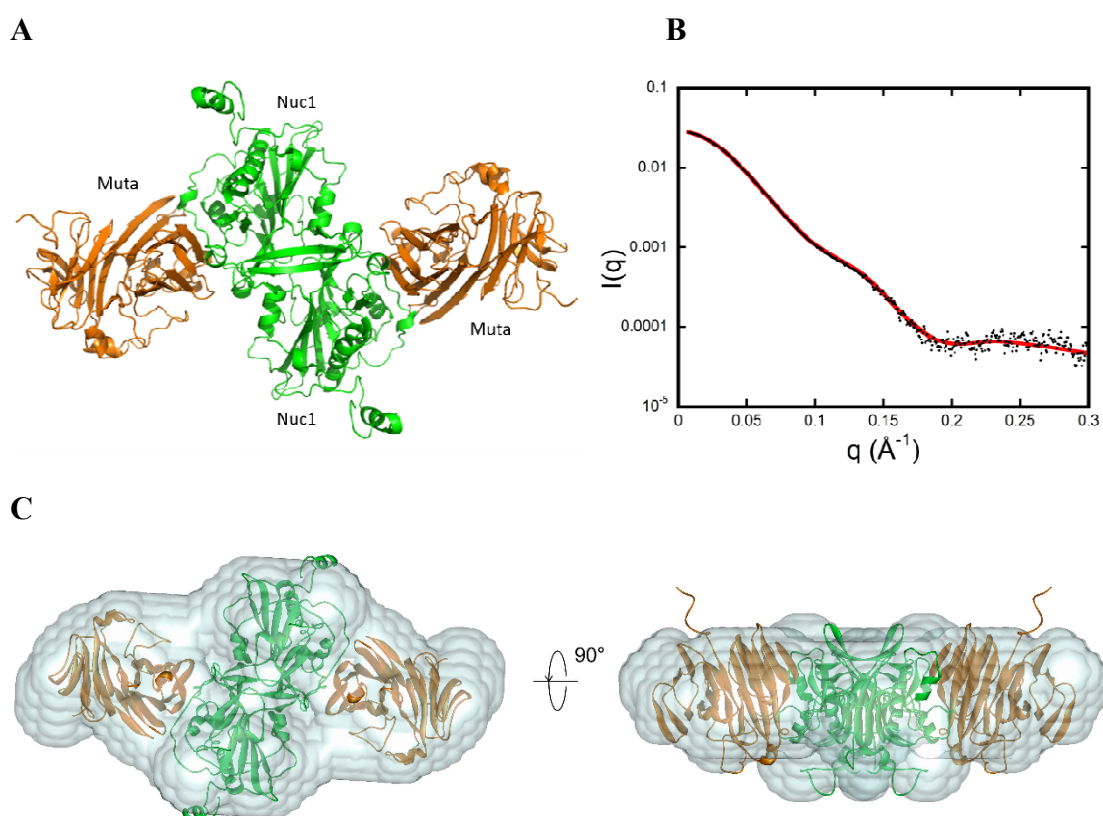


Figure V-4 SAXS analysis of the binary Nuc1/Mutarotase complex. (A) Putative atomic model of the NM binary complex built using the docking program ZDOCK and refined from SAXS data using the program SASREF. Nuc1 subunits are shown in green and mutarotase in orange. (B) The calculated curve (continuous red line) using the program running on the model in (A) and the experimental scattering curve (black spots), $\chi = 1.28$. (C) The atomic model fits well into a typical envelope (obtained by program DAMMIN).

We then attempted a second independent approach to obtain atomic structures. We reasoned that the Nuc1 dimer would be at the center of the NM heterotetrameric complex and the mutarotase subunits were docked onto the Nuc1 dimer (program ZDOCK, (Pierce et al. 2014)). Their positions were refined against the experimental scattering data using the program SASREF (Petoukhov & Svergun 2005) accounting for the two-fold symmetry. The curve calculated (by program CRY SOL) on this putative atomic model of NM shows excellent fitting to the experimental curve ($\chi = 1.28$, Figure V-4 A, B).

Models generated by the two approaches were centered at the origin and rotated to align with the same coordinate axis. The superimposition result, obtained without any fitting, revealed that the global shape of the two models is very similar. We realize that in absence of any biochemical and functional constraints, we are not in a position to obtain the precise atomic model of the complexes from SAXS showing a confident interacting surface. Our data however are able to inform on the stoichiometry and topology of the various complexes.

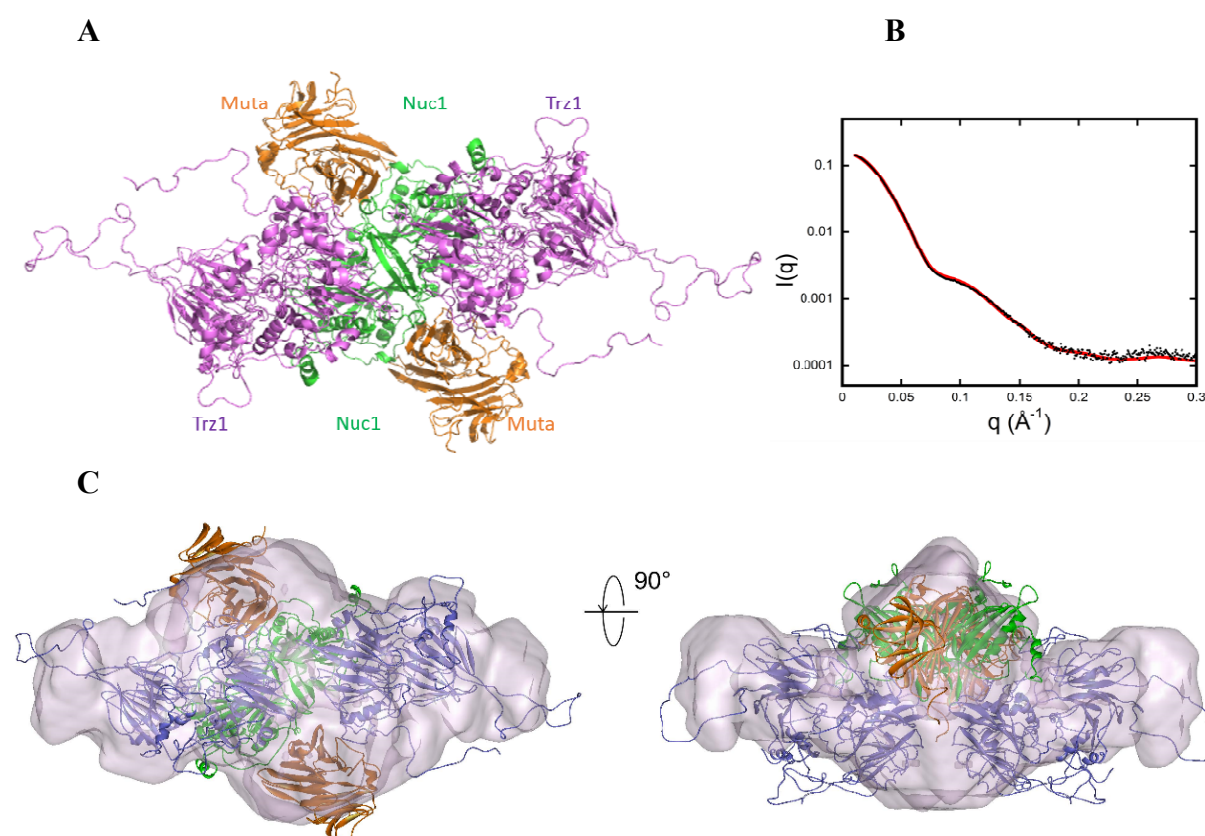


Figure V-5 SAXS analysis of the ternary Trz1/Nuc1/Mutarotase complex.

(A) Putative atomic model of ZNM ternary complex obtained using the program SASREF from the atomic models of the binary complex NM and Z. Nuc1 subunits are shown in green, mutarotase in orange and Trz1 in purple. (B) The calculated curve (continuous red line) on the model in (A) and the experimental scattering curve (black spots), $\chi = 7.62$. (C) Superimposition of the atomic model and a typical envelope (obtained by program DAMMIN).

V.3.2 Analysis of the ZNM complex.

Starting with the SAXS models of Trz1 and NM described above, an atomic model of the ternary ZNM complex was proposed in which the Trz1 subunits were positioned in close contact with the NM binary complex (program SASREF; Figure V-5 C, Right) and a two-fold symmetry was imposed. The agreement between the calculated and experimental scattering curves is not perfect ($\chi = 7.62$, Figure V-5 A, B). We could try to obtain a model better fitting (with a lower χ value) by modifying the conformation of disordered regions in Trz1 and by releasing the two-fold symmetry for these loops, but in the absence of experimental constraints, it remains elusive to determine precisely the relative orientations of the three partners. Moreover, the flexible N- and C-terminal regions and the many loops of Trz1 are an obstacle to an adequate atomic model building. We preferred not to overfit the data

A typical envelope (program DAMMIN) extracted from the experimental scattering data fitted satisfactorily with the atomic model of the complex (Figure V-5 C).

V.3.3 The model of ZN complex

Although we dispose of the crystal structure of Trz1 and a good model of Nuc1, we were not in a position to propose a solution model for the ZN binary complex due to insufficient quality of the SAXS data collected on this complex. We noticed that the calculated curve of the ZN complex which we extracted from the ZNM model fits relatively well with the experimental SAXS curve ($\chi = 4.94$, Figure V-6). This seems to indicate that the assemblage of ZN is similar in the binary or the ternary complex with mutarotase, meaning that addition of mutarotase to ZN does not cause a conformation change detectable by SAXS.

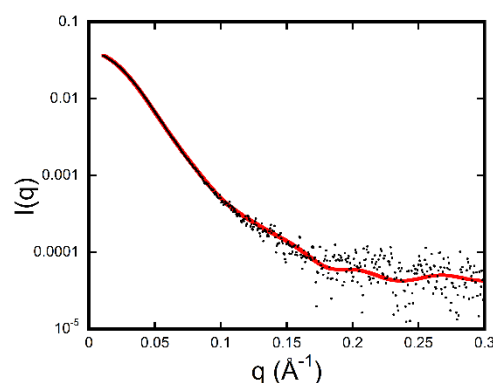


Figure V-6 Fitting of the ZN model derived from ZNM model against experimental data. The ZN is from the ZNM model where M is removed; and its fitting with the experimental data (black spots) is calculated still by the program CRY SOL. $\chi = 4.94$.

V.3.4 Highlights from the SAXS results

The inspection of the model of the complex highlights two main points:

- Trz1 does not extensively interact with the mutarotase (in agreement with the fact that they do not form a complex in absence of Nuc1).
- both the N- and C-terminal domain of Trz1 contribute to binding to Nuc1.

CONCLUSION & DISCUSSION

I. Exploration of the RNase Z^L structure

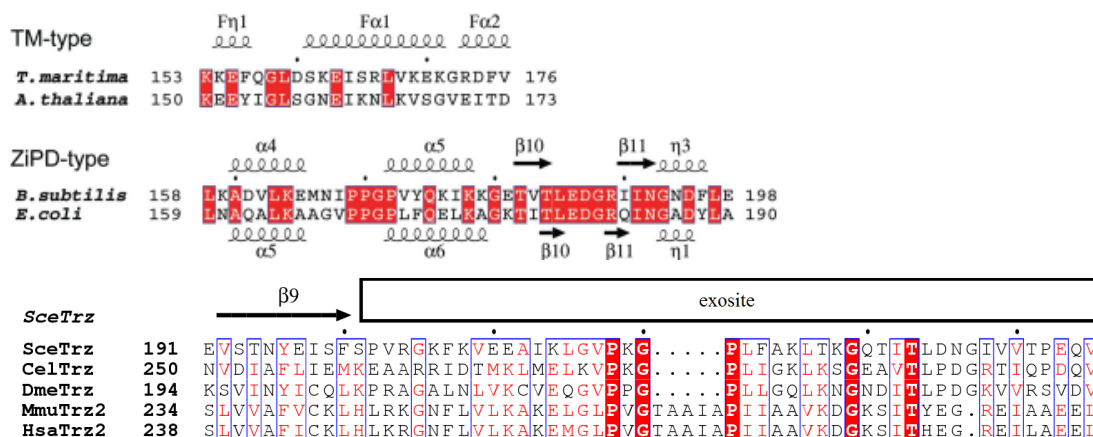
We resolved the crystal structure of Trz1, the first structure of a long form RNase Z. The structure shows it comprises two domains connected by a large linker. Each domain represents a β -lactamase fold, and the association of the two domains mimics the dimeric RNase Z^S. The 3D structure firmly established that the C-terminal domain holds the active site, but lost the flexible arm that is essential for tRNA substrate binding; and that for the N-terminal domain it is the other way around, the active site residues were mutated during evolution but it has an insertion that plays the role of the flexible arm. The active center in the C-terminal domain is constituted by conserved residues from homology motifs, coordinating two zinc ions and a phosphate ion and is also structurally very similar to the active site of the RNase Z^S. Furthermore, the features essential for substrate binding are also retained in the C-terminal. In addition, the structure of Trz1 revealed the presence of three long extensions protruding from the β -lactamase core that are exclusive for the long form RNase Z. These extensions are considerably increasing the interactions between the β -lactamase cores of the two domains.

1.1 A model of Trz1 bound to its tRNA substrate

In the crystal structure of the BsuTrz in complex with substrate tRNA^{Thr} the dimer binds two tRNAs. One subunit provides the main contacts with the tRNA substrate, while the other subunit both establishes binding and provides the catalytic site competent for tRNA cleavage. In order to better understand the structure/function relationship of RNase Z^L forms, we created a model of the Trz1/tRNA complex. The model is obtained by superimposing the structure of Trz1 onto the BsuTrz / tRNA^{Thr} complex form (PDB code: 2FK6). The model was satisfactory in the sense that the two Trz1 domains superpose very well with the BsuTrz dimer and that the position of the tRNA did not create any major clash with the Trz1 protein. This model allowed making solid predictions on the substrate-binding mode of Trz1 (Figure S2).

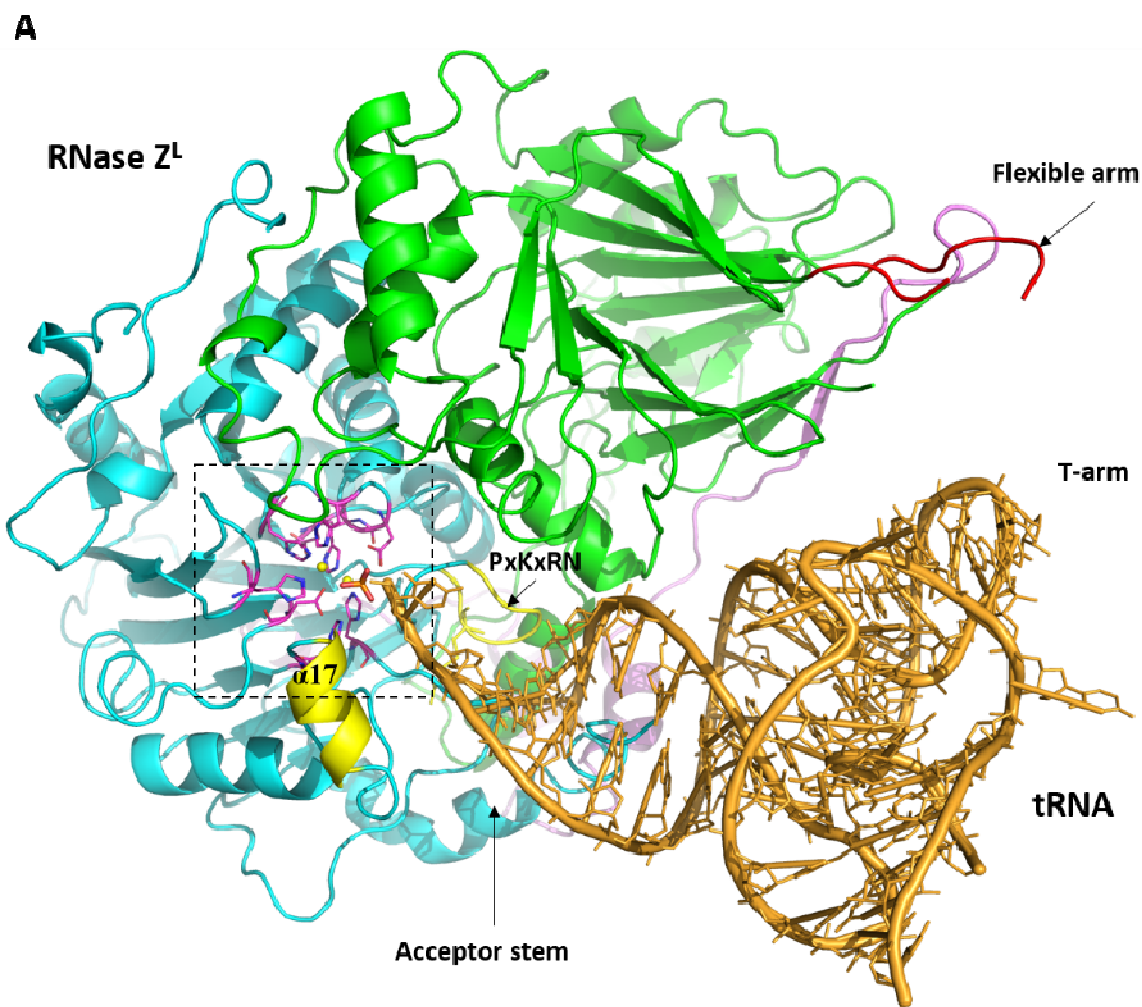
Although the structure and association of the two Trz1 domains strongly mimic the dimer of BsuTrz, the functional roles of the two domains are clearly separated. First, the flexible arm present in the N-terminal domain of Trz1, although disordered in absence of substrate tRNA, is well positioned to play the same role in tRNA clamping as for the BsuTrz.

The flexible arm is replaced by a short loop instead in the C-terminal domain, which is clearly too short to embrace the T arm of the tRNA. According to the sequence alignment, the flexible arm of Trz1 belongs to the ZiPD-type, which is the same as that of BsuTrz. In fact, sequence alignment shows many long form RNase Z (i.e. SceTrz^L, CelTrz^L, DmeTrz^L, MmuTrz^L, HsaTrz^L) have a ZiPD-type flexible arm in their N-terminal domain (Figure S1).



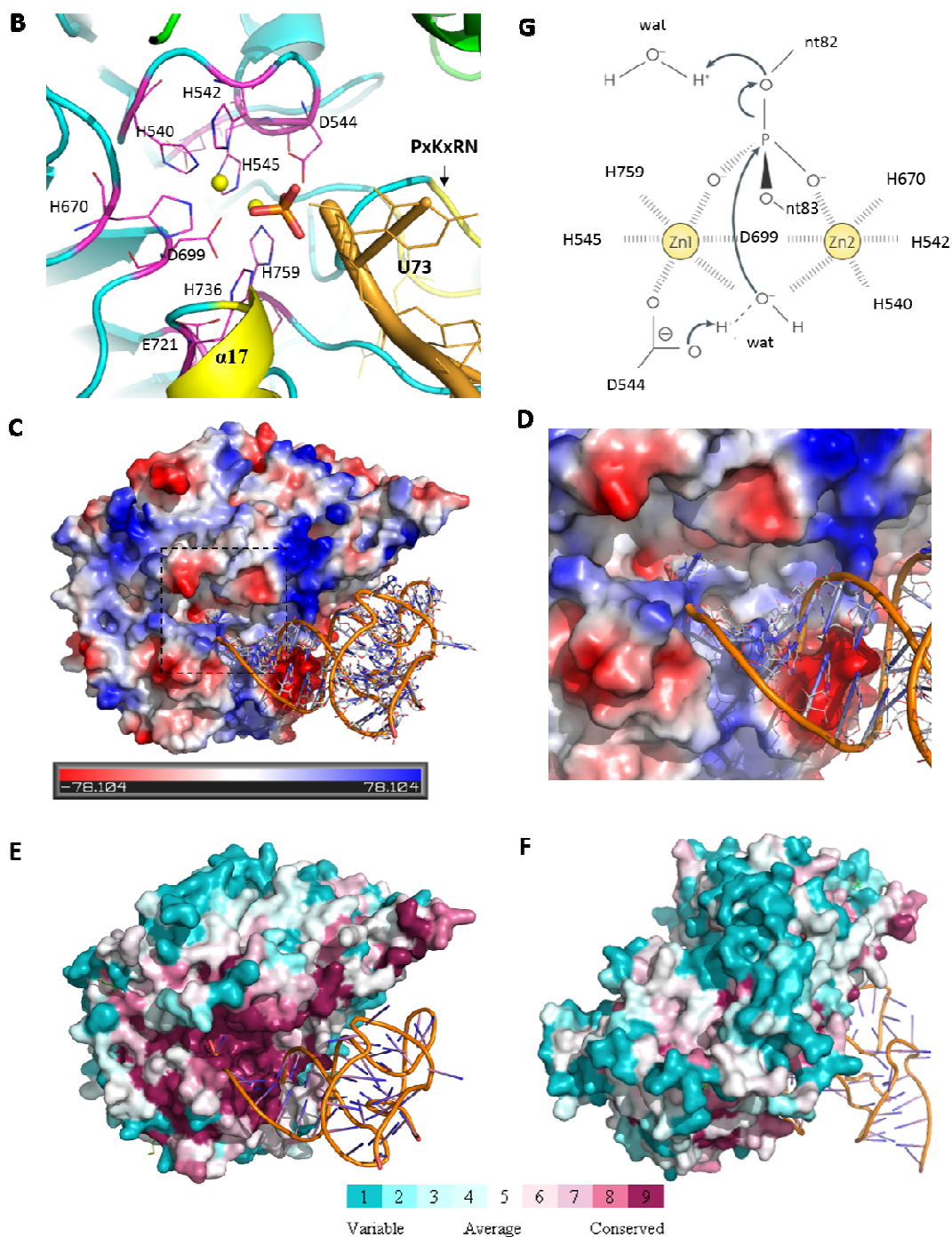
FigureS 1 Many RNase Z^L have the ZIPD-type of flexible arms. Sequence alignment of the flexible arm from SceTrz^L, CelTrz^L, DmeTrz^L, MmuTrz^L, HsaTrz^L are compared with the two types of exosite from RNase Z^S.

In our model the 3'-end of the tRNA substrate plunges exactly into the active center of the C-terminal domain and establishes contacts that are identical as those observed for the BsuTrz/tRNA complex (Figure S2 A and B). In addition, the acceptor stem of tRNA is fixed by the helix α 17 and the PxKxRN loop motif of the C-terminal domain. Our model of the Trz1/tRNA complex is also supported by the calculated electrostatic potential surface of Trz1, which clearly displays a positively charged cave surrounding the 3'-end of tRNA fitting into the active site pocket (Figure S2 C-D). A representation of the sequence conservation projected onto the surface showed this cave is highly conserved; whereas the remainder of the protein surface is not (Figure S2 E-F). We conclude that our model of the Trz1/tRNA complex obtained by simple superposition of Trz1 and the BsuTrz/tRNA complex structures seems very compatible with sequence and structure data of the long form of RNase Z.



FigureS 2 Model of Trz1 in complex with tRNA.

(A) The model is obtained by superposing the BsuTrz/tRNA complex (PDB: 2FK6) onto Trz1. (tRNA is in orange and BsuTrz is not shown). The position of the flexible arm and acceptor stem are indicated by arrows. The N- / C- terminal halves and linker of Trz1 are shown in green/cyan and purple, respectively. The active center of Trz1 is shown in pink stick, and zinc ions shown in yellow sphere. The α17 helix and “PαKαRN” loop are shown in yellow. (B) Zoom in on the active site indicated by a dashed line square in (A). The bound phosphate ion is shown in thick sticks. (C) Electrostatic potential surface of Trz1 as generated by Pymol (positively charged regions are in blue, negatively charged regions in red, neutral regions in white); and (D) Zoom in of dashed line square in (C). (E-F) Sequence conservation projected onto the surface, as generated with default parameters by the webserver (<http://consurf.tau.ac.il>). Conservation based on the multiple sequence alignment can be found in the annex. Trz1 in figure (E) has the same orientation as in figure (A) and is turned by 180° in figure (F). (G) Potential reaction mechanism of Trz1 (adapted from (Redko et al. 2007), the numbering of the residues and nucleotide were adapted to Trz1 and tRNA^{Ser} from *S. cerevisiae* used in our the activity assay).



FigureS 3 (continue) Model of Trz1 in complex with tRNA.

1.2 Catalytic mechanism of Trz1

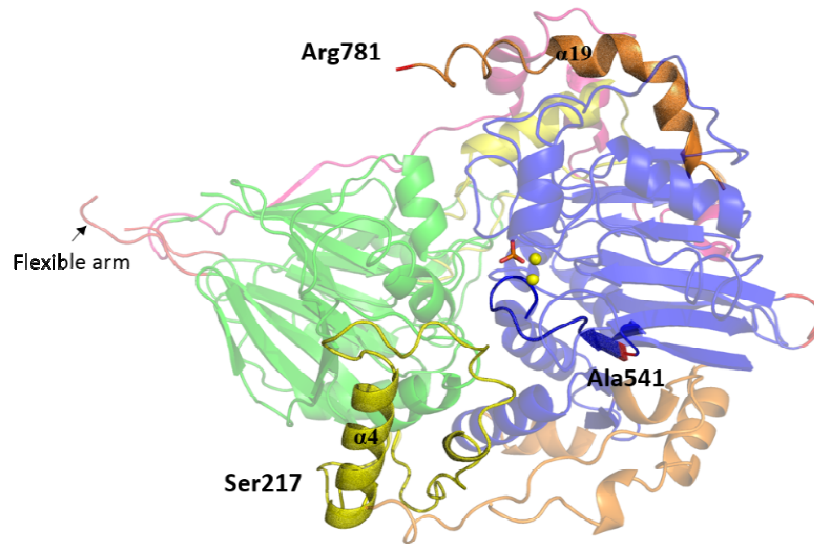
In order to describe the catalytic machinery of Trz1, we superposed its structure with those of both the free form and the tRNA complex form of BusTrz. The complete catalytic machinery used by BsuTrz is perfectly superposable onto that of the C-terminal domain of Trz1. The histidines from motifs II-V, the aspartate from motif I, II, a phosphate ion and a

water molecule provide a pseudo-octahedral coordination of the two zinc ions, constituting the catalytic center. In analogy with what was proposed for BsuTrz, the negatively charged Asp544 (Asp67 in BsuTrz) acts as a general base and deprotonates a neighboring water (Redko et al. 2007). The nucleophilic hydroxyl group then attacks the cleavable phosphate linkage of the tRNA substrate that is polarized by the two zinc ions, resulting in the breakage of the phosphor diester bond between nucleotide 82 and 83. Then the 3' position of tRNA restores its hydroxy group by protonation through a water molecule (Figure S2 G).

In the free form of BsuTrz (PDB code: 1Y44) only one subunit carries the catalytic Zn ions, and the other is therefore inactive. In the presence of tRNA, the inactive subunit is activated by the other subunit, behaves as an allosteric enzyme, and cleave two tRNAs eventually. The structure of Trz1 demonstrates that it is capable of processing only one tRNA at a time rather than two. The long form of RNase Z from *D. melanogaster* behaves as a Michaelian enzyme rather than an allosteric (Mohan et al. 1999) proves our model.

1.3 Mutations related to prostate cancer susceptibility

A number of mutations found in the ELAC2 gene, the human RNase Z homolog, were connected with various diseases. Trz1 and ELAC2 share 21% of sequence identity and 49% sequence similarity. With the help of the structure of Trz1 we tried to explore the structural and functional consequences of these mutations (Figure S3). ELAC2 was marked as a prostate cancer susceptible gene (Tavtigian et al. 2001). The related mutations are S217L, A541T, R781H, and the last one is a guanine insertion after base 1641, causing premature termination of the ELAC2 transcript (Takaku et al. 2003). The three mutated residues of ELAC2 are not conserved in the gene coding for Trz1. Two of the three mutated residues, Ser217 and Arg781, locate in the mobile loop regions implicated in the interactions between the two domains of Trz1: Ser217 lies in the fifty-residue-long extension region (Extension-I, residues 136 to 190) between the strands $\beta 8$ and $\beta 9$ which contains the helix $\alpha 4$; and Arg781 is in the disordered tail region following the last long helix $\alpha 19$. The third mutated residue Ala541 corresponds to Met535 in Trz1, and although it precedes by a few amino acids the Histidine motif in the sequence it is distant from the active centre in the 3-D structure. The positions of these mutations in the 3-D structure of Trz1 support the finding that none of these mutations affect the tRNA 3'-processing activity of ELAC2 (Figure 21 of introduction, Table S1).



FigureS 4 Localization of prostate cancer-related mutations in the Trz1 3-D structure.

Mutated residues are labeled with ELAC2 numbering. The coloring is the same as previously described: The N-terminal domain is shown in green containing the extensions in yellow, the C-terminal domain is in blue, the extensions and tails are in orange and the linker region is in purple.

Table S 1 ELAC2 disease-related mutations and the corresponding residues in Trz1

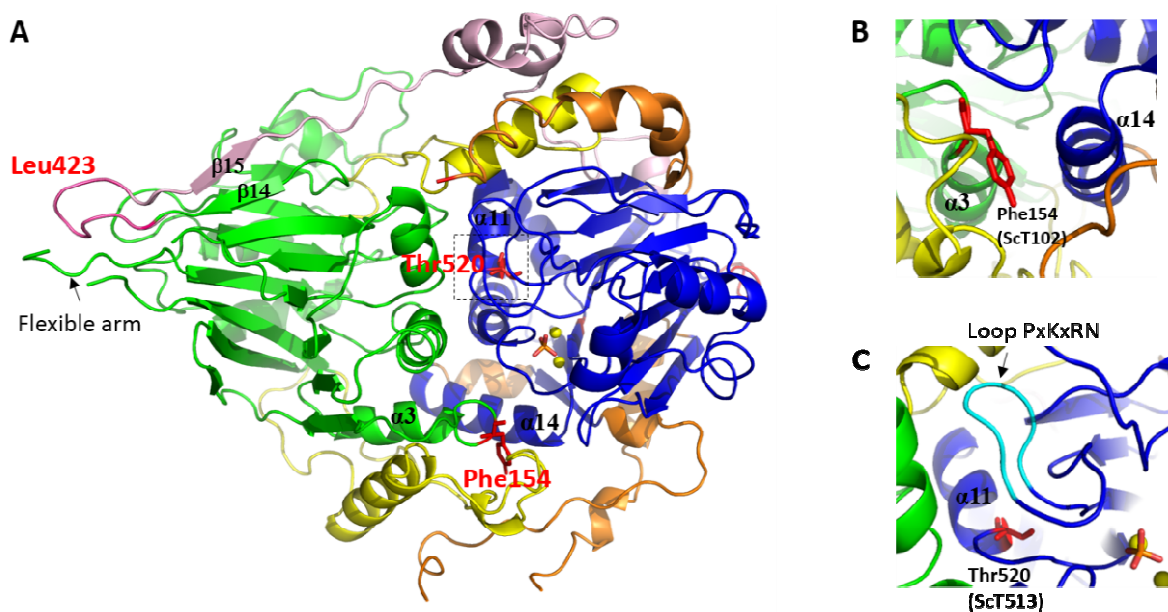
Disease	Mutation in Human	Residue in <i>S. cerevisiae</i>	Location in Trz1 structure	tRNA-processing defect
Hypertrophic cardiomyopathy	F154L	Tyr102	$\alpha 3$ end (Nt-Ct interface)	Yes ⁽¹⁾
	L423F	Pro404	linker	Yes ⁽¹⁾
	T520I	Thr513	Motif I (Ct)	yes ⁽¹⁾⁽³⁾
	Arg211	--	--	STOP ⁽¹⁾
Prostate cancer	S217L	136-147	Extension-I	No ⁽²⁾
	A541T	M535	Before His-motif	No ⁽²⁾
	R781H	Val819	Tail region	No ⁽²⁾
	1641G (base)	--	--	STOP ⁽²⁾

(1) (Haack et al. 2013); (3) (Zareen 2006). (2) (Takaku et al. 2003).

I.4 Mutations related to hypertrophic cardiomyopathy

ELAC2 mutations were also found to be associated with hypertrophic cardiomyopathy, which is caused by a mitochondrial RNA processing defect (Haack et al. 2013). The four related mutations are F154L, L423F, and T520I, and the last one is a STOP mutation at Arg211. Two mutations, Phe154 and Thr520, locate in the N-terminal- and C-terminal β -lactamase domains respectively, and are highly conserved residues among RNase Z^L in eukaryotes (Figure 22 in Introduction). The residues in Trz1 equivalent with the mutated

residues in ELAC2 are shown in the Trz1 structure (Figure S4). Phe154, corresponding to Tyr102 in Trz1, is at end of the helix α_3 at the interface of the N-terminal- and C-terminal domain. Thr520 is equivalent to Thr513 in Trz1, which locates in the conserved Motif I of the C-terminal domain. The *S. cerevisiae* strain harbouring the T513I mutation showed reduced growth rate on a non-fermentable medium, which is a typical phenotype indicative of impaired respiration; a higher *petite* frequency is also observed with this mutant, indicating quantitatively reduced mitochondria translation (Haack et al. 2013). Mutation of Thr471 in the DmeTrz^L Δ 23 (equivalent with human Thr520) reduced the activity 10 fold (Zareen 2006). It is interesting to note that Thr520 is part of helix α_1 , which is also at the interface of N-terminal- and C-terminal domains. Leu423 locates in the beginning of the linker region between the strands β_{14} and β_{15} , and is not conserved in Trz1.



FigureS 5 Localization of cardiomyopathy-related mutation in Trz1 3-D structure.
Residues are labeled in red with ELAC2 numbering. The coloring is the same as previously described.

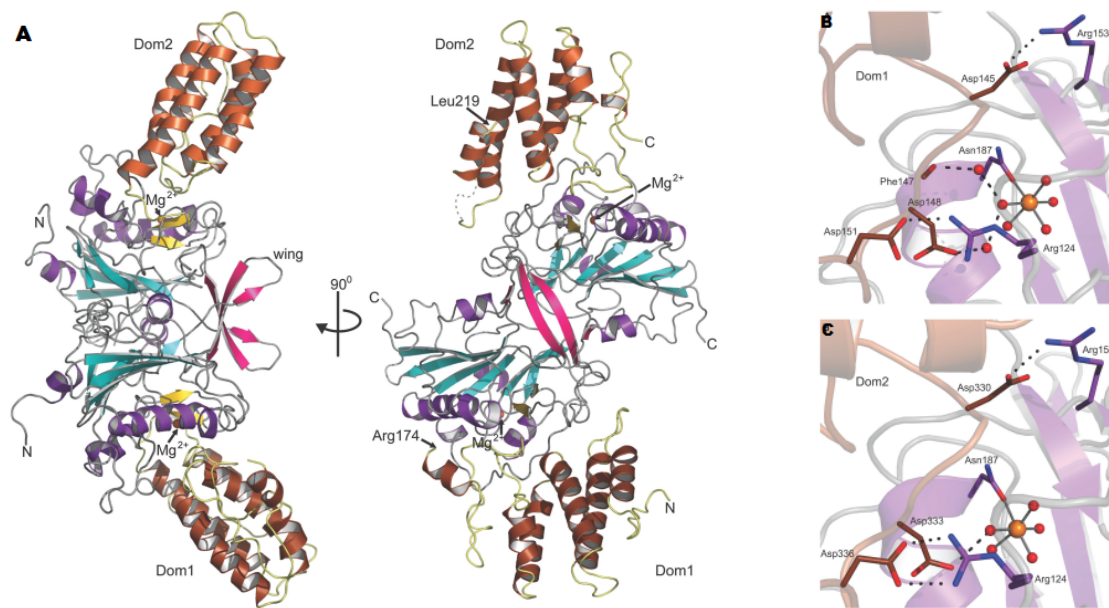
As the first resolved structure of the long form of tRNase Z, thanks to the high similarity of Trz1 and ELAC2 in sequence and the conservation among tRNase Z^L proteins, the 3-D structure of Trz1 certainly provides some information on understanding the mechanisms of these diseases and could be helpful in clinical drug development.

II. Biological significance of Trz1/ Nuc1/mutarotase ternary complex model

We proved the existence of two binary complexes Trz1/Nuc1, Nuc1/mutarotase by gel filtration and SEC-MALLS, and one ternary Trz1/Nuc1/mutarotase complex. These complexes all proved to be very stable and their low-resolution structures could therefore be studied by SAXS. Trz1, mutarotase are monomeric and Nuc1 forms a homodimer. Our experimental data suggest that the Nuc1 dimer forms the kernel of the binary and ternary complexes, which form either hetero-tetramers (ZN, NM complexes) or hetero-hexamers (ZNM complex).

One of the long-term objectives of this project is to understand why these three proteins are forming such a stable complex. We started to investigate this question by verifying whether the enzymatic properties of the three protein partners are affected by the formation of the ternary complex. Our tRNA 3'-end processing assays showed that the activities of the binary and ternary complexes are similar with that of Trz1 alone. The glucose-6-phosphate mutarotase activity of the complexes also was the same as mutarotase alone. In contrast, the DNA-nuclease activity of Nuc1 is remarkably inhibited by its interaction with mutarotase in the context of either the binary NM or the ternary ZNM complex. This inhibition was confirmed by two totally different assays using as substrate either plasmid dsDNA or a hairpin structured artificial short DNA. This inhibiting effect on Nuc1 by mutarotase was totally unexpected and may provide a new and exciting direction to start looking for a potential biological role of the ZNM complex. In *Drosophila*, the EndoG nuclease was shown to be translocated to the nucleus during apoptosis and participated to DNA degradation (Zhao et al. 2007). Different organisms have developed different mechanisms to regulate the deleterious activity of the EndoG homologs in order to protect cellular nucleic acids from sorting failures. *Drosophila* possesses a EndoG inhibitor protein (EndoGI) mostly located into the nucleus (Temme et al. 2009). EndoGI binds to EndoG with subpicomolar affinity and the interaction is dependent on the active site of the nuclease. EndoGI consists of two very similar domains that are linked by a long peptide stretch (Figure S5). EndoGI binds to EndoG with a one to two stoichiometry (one EndoGI per EndoG dimer). Each domain, when tested on its own is capable of inhibiting EndoG activity. In the crystal structure of EndoGI in complex with EndoG, the two domains of EndoGI block the active sites of the two EndoG subunits in the complex (Loll et al. 2009).

There exists no obvious homolog of EndoGI in yeast. The 3D structure of EndoGI, consisting of two very similar helical bundles connected by a disordered linker, is totally different from that of mutarotase. One of the ridges of the EndoGI helical bundles contain a highly acidic surface patch that binds to the positively charged DNA binding groove of EndoG. Interestingly, EndoGI is composed of HEAT repeats that are present in many transport proteins and is related to the structure of importin β which was shown to be involved in Nuc1 translocation in *S. cerevisiae* (Büttner, Eisenberg, et al. 2007).



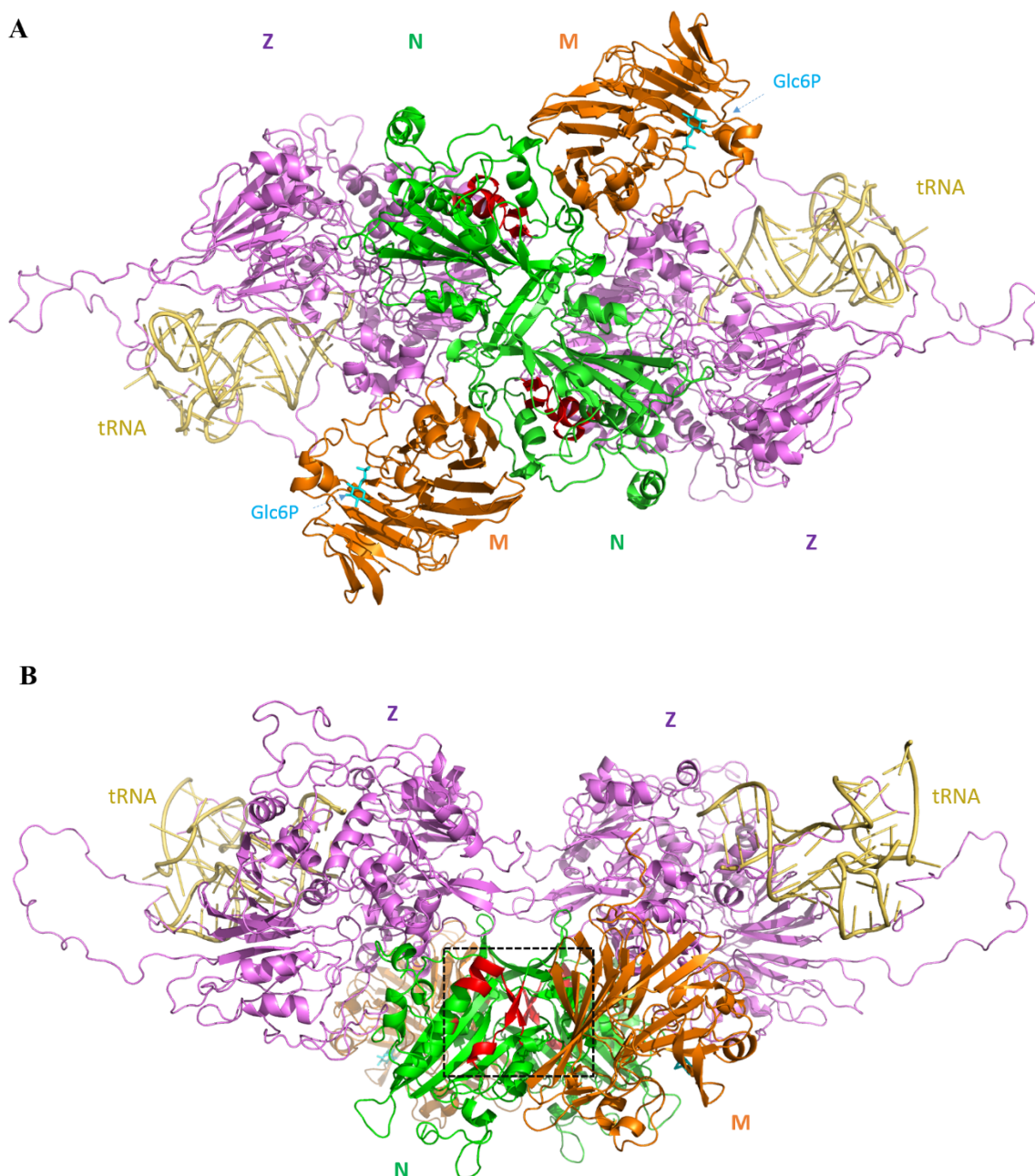
FigureS 6 Structure of dEndoG/dEndoGI complex (Loll et al. 2009).

(A) Overall structure. Helices of dEndoG are illustrated as purple ribbons and strands of the central β -sheet in cyan. The small two-stranded β -sheet (β -strands D and E) involved in metal ion binding is highlighted in yellow, and the bound metal ion is shown as an orange sphere. The two wings forming an intermolecular β -sheet are shown in red (β -strands G and H). Dom1 and Dom2 of Endo GI, which sandwich the dEndoG dimer, are shown as brown ribbons. The missing linker between the two domains would start at Arg174 in Dom1 and connect the polypeptide chain with Leu219 in Dom2. (B) Inhibition of the active site of dEndoG by dEndoGI. The interface between Dom1 of dEndoGI (brown) and the active site of one dEndoG monomer (violet) is shown in (B) and the equivalent interface between Dom2 and the second dEndoG molecule in (C). Residues involved in binding are shown as sticks and for simplicity only salt bridges of important arginine residues (dotted black lines) and hydrogen bonds to water molecules (dashed black lines) are presented.

In order to establish a link between our enzymatic activity data and the SAXS model of the ternary ZNM complex, we superimposed it on our model of the Trz1/tRNA complex (itself based on the BsuTrz/tRNA structure; PDB code 2FK6). The superposition is presented in Figure S6. As can be seen, neither mutarotase nor Nuc1 seem to block the access of the tRNA substrate to Trz1 in the ZNM 3D-model. This observation is compatible with the absence of inhibitory effect of mutarotase and Nuc1 on the tRNAse activity of Trz1. On the

Conclusion & Discussion

other hand, our SAXS model places mutarotase near the conserved $\beta\beta\alpha$ -Me-finger motif of Nuc1, which is essential in substrate catalysis/binding and metal coordination (Wu et al. 2009; Lin et al. 2012; Schäfer et al. 2004)(Büttner, Eisenberg, et al. 2007). In this position mutarotase would block the entry of DNA/RNA to the active site of Nuc1 (Figure S6 B) causing the inhibitory effect we observed. However, the active site of mutarotase (indicated by the stick model of the Glc6P substrate in Figure S6 B) is neither in contact with Nuc1 nor with Trz1, explaining why the mutarotase activity is not affected in the ternary ZNM complex. We conclude that our enzymatic activity data are compatible with (but do not prove) the SAXS model of the ZNM complex.



FigureS 7 Representation of ZNM complex in complexed with substrates.

Z (Trz1), N (Nuc1), and M (mutarotase) are colored in purple, green, and orange, respectively. tRNA (yellow) and Glc6P (cyan) was positioned by superimposition with corresponding enzyme/substrate complex (PDB code 2FK6, 2CIR, respectively). The conserved $\beta\beta$ -Me-finger in Nuc1 involved in DNA catalysis/binding is colored in red (residues 165-144, 160-172). (A) Front view. (B) Top view.

The nuclease activity of Nuc1 is indispensable for the Nuc1-mediated apoptosis: the yeast strain Nuc1H138A (His138 from the DRGH motif) barely responds to apoptosis inducer H_2O_2 (Büttner, Eisenberg, et al. 2007). Similarly, the reduced DNase activity of CPS-6 (the *C. elegans* homologue of Nuc1) mutants was positively correlated with its diminished cell killing activity in *C. elegans* (Lin et al. 2012). Therefore, we speculate that the ZNM complex might be a regulator of the apoptosis signal mediated by Nuc1. This is positively correlated with the fact that Nuc1 is pro-live rather than pro-death in the glucose medium, whose pathway is unknown (Burhans & Weinberger 2007) (Figure 34 in introduction). Since glucose-6-phosphate is the substrate of mutarotase, this might provide some clue for understanding the unknown mechanism of Nuc1 pro-live effect in glucose medium. Many metabolic enzymes were discovered to be involved in moonlighting functions, and mutarotase might be another example of this phenomenon. One of the best-known examples of moonlighting is cytochrome c which besides its well-known electron transport function is also an activator of apoptosis. The knowledge of the precise 3D structure of the ZNM complex would be a very stimulating asset to investigate the hypothesis of an implication of this complex in the regulation of apoptosis.

MATERIALS & METHODS

I. Nuc1 cloning

NUC1 (YJL208C) was cloned from yeast chromosomal DNA using primers including His-tag and Strep-tag sequence (Table M1, M2). Two truncated forms of Nuc1, Nuc1 Δ 24, and Nuc1 Δ 43 (corresponding to N-terminal truncations of 24 and 43 residues respectively) fused with the two tags, separately, were amplified by PCR. The amplified segment was inserted into the vector pET29a between enzyme restriction sites NdeI and XhoI. The plasmids were amplified in the *E. coli* strain XL1 Blue and sequenced.

TableM 1 PCR system for amplifying Nuc1

Buffer 5x	10 μ l	98 °C	5 min	
DMSO	5 μ l	98 °C	30 sec	
Chromosome DNA (50 ng)	1 μ l	<u>55 °C</u>	30 sec	34 cycles
dNTPs mix (2.5 mM each)	1 μ l	72 °C	<u>30 sec</u>	
Oligos (10 pmoles/ μ l)		72 °C	10 min	
Upper	1 μ l			
Lower	1 μ l			
Polymerase Phusion	1 μ l			
H ₂ O	30 μ l			
Final volume	50 μ l			

TableM 2 Oligos used in Nuc1 cloning

Oligo name	Sequence (5' to 3')
Nuc1-D24-Nde-His-5'	GGGAA CATATG CAC CAT CAC CAT CAC CAT aacaacattctccaacgc
Nuc1-D24-Nde-His-5'	GGGAA CATATG TGG AGC CAC CCG CAG TTC GAA AAA aacaacattctccaacgc
Nuc1-D43-Nde-His-5'	GGGAA CATATG CAC CAT CAC CAT CAC CAT aatagtaattcaatctc
Nuc1-D43-Nde-Strep-5'	GGGAA CATATG TGG AGC CAC CCG CAG TTC GAA AAA aatagtaattcaatctc
Nuc1-stop-Xho-3'	GAC CTCGAG TTA attccttttttttgaggaggta

II. Protein expression and purification

II.1.1 Constructs of Trz1, Nuc1 and mutarotase

The constructs, ORFs of Trz1, Nuc1 and mutarotase, and their protein parameters are listed in the TableM 3.

TableM 3 Protein parameters.

Protein	Molecular weight	Ext. coefficient	Theoretical pI	Tag	vector
Z (Trz1)	101 kDa	90.44	5.55	N-terminal His	pET-45 (Amp+)
N (Nuc1Δ43)	33.8 kDa	33.015	8.65	N-terminal Strep	pET-24a(+) (Kan+)
M (mutarotase)	34.8 kDa	61.9	6.10	C-terminal His	pET-9 (Kan+)
ZN	134.8 kDa	123.45	5.86	-	-
NM	68.6 kDa	95	6.75	-	-
ZNM	169.5 kDa	185.4	5.92	-	-
ZNM_poly ⁽¹⁾	165.4 kDa	178.4	5.99	-	pET-21a(+) (Amp+)

(1) ZNM complex expressed from the polycistronic plasmid.

II.1.2 Native protein expression in *E. coli*

The plasmids harboring the ORF of the target protein were transfected into *E. coli* competent cells, and inoculated to 2YT medium as preculture and incubated at 37°C at 200 rpm overnight. 8ml preculture were inoculated to 800 ml 2YT medium supplemented with antibiotics. The cell cultures were incubated at 37 °C at 200 rpm until the O.D.₆₀₀ reached 0.6-0.8, protein expression was induced by adding 0.5 mM IPTG and the cultures were further incubated at 37°C for 3-4 h or at 15 °C 16-20 h at 200 rpm. Cells were collected by centrifuging at 4000 rpm for 20 min; the pellets were resuspended by lysis buffer (about 30 ml for pellets of 800 ml culture) are frozen at -80 °C.

On the day of purification, the pellet suspension were thawed, sonicated, and centrifuged at 20,000 rpm for 30 min to remove the insoluble fraction. The supernatants were loaded to affinity purification columns.

Medium of Culture: 2YT: 5 g/L NaCl, 16 g/L Bacto-tryptone, and 10g/L yeast extract.

II.1.3 Selenium-labelled protein expression in *E. coli*

Selenium-labelled Trz1 was produced with the auto-inducible synthetic medium (“M-AI”) containing SeMet. Two precultures were needed: normal 2TY medium was used for the 1st pre-culture; the synthetic medium not auto-inducible (“M-NAI”) was used for the 2nd preculture. Finally, for the production culture in large volume the auto-inducible medium containing SeMet was used. The protocol is summarized in TableM 4.

In the case of Se-methionine labeling of Trz1, the same induction temperature at 15 °C was used as that for the native protein.

TableM 4 Protocol for Se-labelled protein expression

Start time	Events	Lasting time	Volume	Medium
Day 1 morning	Transfection, 1 st Preculture	1 day	10 ml	2 YT selective
Day 1 evening	2 nd preculture	Overnight	10 ml	M-NAI
Day 2 morning	Production culture	65-72 h at 15 °C*	1000 ml	M-AI

TableM 5 Compositions of non- auto-inducible medium (M-NAI)

	STOCKS		Volume	Pour 100 mL	Pour 200 mL	[C]finale	Molarité
			STOCKS			g/L	
NPS	g/L	x10	100 mL	10 mL	20 mL	g/L	
(NH ₄) ₂ SO ₄	66 g					6,6	
KH ₂ PO ₄	136 g (1 M)					13,6	0,1 M
Na ₂ HPO ₄	142 g (1 M)					14,2	0,1 M
H ₂ O				85 mL	170		
Substrats Carbonés						g/L	
Glycérol	50% V/V	x62,5		1,6 mL	3,2 mL	10	
glucose	50% W/V	x100		1 mL	2 mL	5	
Antibiotiques (ou)							
Kanamycine	25mg/mL	x1000		0,1mL	0,2mL	25µg/mL	
Chloramphénicol	30mg/mL	x1000		0,1mL	0,2mL	30µg/mL	
Ampicilline	50mg/mL	x500		0,2 ml	0,4 ml	100µg/mL	
OLIGO ET VITAMINES						mg/L	
Thiamine	2 g/L	x500		0,2 mL	0,4 mL	4	
CaCl ₂ , 2H ₂ O	10 g/L	x100		1 mL	2 mL	100	
MgSO ₄ . 7H ₂ O	80 g/L	x100		1 mL	2 mL	800	
FeSO ₄ ,7H ₂ O	10 g/L	x2000		0,05 mL	0,1 mL	5	
OLIGO-ELEMENTS (mg/L)		x500	1 L	0,2 mL	0,4 mL	mg/L	
ZnSO ₄ ,7H ₂ O	400					0,8	
MnSO ₄ ,1H ₂ O	400					0,8	
CuSO ₄ ,5H ₂ O	40					0,08	
Na ₂ MoO ₄	200					0,4	
H ₃ Bo ₃	500					1,0	
KI	100					0,2	

Table M 6 Compositions of auto-inducible medium (M-AI) containing SeMet

	STOCKS		Volume STOCKS	Pour 1L	Pour 2 L	[C]finale g/L	Molarité
NPS	g	x10	100 mL	100mL	200 mL	g/L	
(NH ₄) ₂ SO ₄	66 g					6,6	0,1 M
KH ₂ PO ₄	136 g (1 M)					13,6	0,1 M
Na ₂ HPO ₄	142 g (1 M)					14,2	
H ₂ O				750 mL	1500		
Substrats Carbonés						g/L	
Glycérol	50% V/V	x62,5		16 ml	32 mL	10	
glucose	50% W/V	x500		2 ml	4mL	1	
lactose	40 g/L	x10	1L	100 mL	200 mL	4	
Antibiotiques (ou)							
Kanamycine	25mg/mL	x1000		1mL	2 mL	25µg/mL	
Chloramphénicol	30mg/mL	x1000		1mL	2 mL	30µg/mL	
Ampicilline	50mg/mL	x1000		1mL	2 mL	100µg/mL	
OLIGO ET VITAMINES						mg/L	
Thiamine	2 g/L	x500		2 mL	4 mL	4	
CaCl ₂ , 2H ₂ O	10 g/L	x100		10 mL	20 mL	100	
MgSO ₄ . 7H ₂ O	80 g/L	x100		10 mL	20 mL	800	
FeSO ₄ ,7H ₂ O	10 g/L	x2000		0,5 mL	1 mL	5	
OLIGO-ELEMENTS (mg/L)		x500	1 L	2 mL	4 mL	mg/L	
ZnSO ₄ ,7H ₂ O	400					0,8	
MnSO ₄ ,1H ₂ O	400					0,8	
CuSO ₄ ,5H ₂ O	40					0,08	
Na ₂ MoO ₄	200					0,4	
H ₃ Bo ₃	500					1,0	
KI	100					0,2	
Solution AA	g/L	du Milieu Studier				mg/L	
leucine	10		1 L	20 mL	40 mL	200	
isoleucine	10					200	
valine	10					200	
Tryptophane	10	x50				200	
phenylalanine	10					200	
lysine	10					200	
alanine	10					200	
arginine	10					200	
aspartic	10					200	
glycine	10					200	
asparagine	10					200	
proline	10					200	
serine	10					200	
threonine	10					200	
histidine	10					200	
glutamine	10					200	
glutamic	10					200	
selenomethionine	2,5	x20		50	100	125	

III. Enzyme activity assays

III.1 RNase Z activity assay

The templates for *in vitro* transcription of the tRNA precursor substrates were amplified from yeast chromosomal DNA (for yeast tRNA^{Ser}) or plasmid pHMT1 DNA (for *B. subtilis* tRNA^{Thr}) by PCR (TableM 7). The upstream primers contain a T7 RNA polymerase promoter; and the downstream primers were designed to cover the 3'-trailer sequence of the pre-tRNA (Table M8). The PCR product were purified by a clean up column (Nucleospin™ PCR clean up kit) and examined by 2% agarose gel electrophoresis.

TableM 7 PCR system for amplifying templates of tRNA transcription.

Chromosome DNA (1 µg)	4 µl (0.25 µg/µl)	94 °C	5 min	35 cycles
dNTPs mix (2.5 mM each)	4 µl	94 °C	2 min	
Buffer 10X	5 µl	<u>55 °C</u>	30 sec	
Oligos (10 pmoles/µl)		72 °C	<u>30 sec</u>	
Upper	2.5 µl	72 °C	5 min	
Lower	2.5 µl			
Enzyme (Taq)	0.8 µl			
H ₂ O	31.2 µl			
Final volume	50 µl			

TableM 8 PCR oligos used for amplifying templates of tRNA transcription.

Target tRNA		No.	Sequences (5'-3')
Sc tRNA ^{Ser}	Upstream	CC1610	GCTCTAATACGACTCACTATA <u>G</u> GCACTATGGCCGAGTGGTTAA GGC
	Downstream	CC1611	GAAGATATCTGATAATCAGTAAGCG
Bs tRNA ^{Thr}	Upstream	HP560	ATTAATACGACTCACTATA <u>G</u> CTTCATAGCTCAGCAGGTAG
	Downstream	HP612	AATACATTTAAGCTTCCAAG

Legends: AAA T7 RNA promoter; G last nucleotide of T7 promoter and meanwhile the first nucleotide of tRNA.

In vitro transcription was carried out using an Ambion MegaShortScript kit according to the manufacturer's instructions. The components are described in the results chapter in which UTP- $\alpha^{32}\text{P}$ was included, and mix were incubated 2 hours and 15 minutes at 37°C. Then 2 µl RQ DNase (Promega™) was added to the transcription product and incubated 15 minutes at 37°C. The template DNA was removed by addition of RQ DNase. The product was adjusted to 30 µl and purified by G50 spin column (GE Healthcare™). It is important to centrifuge the

column always in the same direction. Unincorporated nucleotides were removed by running the mixture over a G50 column. The labelling by UTP- $\alpha^{32}\text{P}$ was verified by a radioactive detector.

The pre-tRNAs were prepared by Dr. Olivier Pellegrini from Institute de Biologie Physico-Chimique (CNRS UPR 9073).

The reaction mixtures were prepared in two parts separately:

- Substrate-mix* (TableM 9): The reaction mix containing radioactive transcripts and 5X reaction buffer were complemented with H₂O to 2 μl and multiplied. Reaction buffer for BsuTrz and Trz1 (SceTrz) are listed in TableM 10.

- Enzyme (TableM 9): Stock solutions of Trz1, Nuc1, mutarotase were diluted at equal molar concentrations (TableM 11), so that one protein in the combination corresponds to 1 μl in the reaction since Z, N, M are present in a 1:1:1 stoichiometry in the complexes.

The experiments were carried out by adding the substrate-mix to the enzymes followed by immediate incubation at 37°C for 20min. The reactions were stopped by addition of 5 μl loading buffer (95% formamide, 20mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) and this mixture were applied directly to 5% polyacrylamide/7 M urea gels.

TableM 9 Experiment set up for RNase Z activity

Substrate Mix* (2 μl)	Enzyme combination(s) (3 μl)					Final volume (5 μl)
0.2 μl transcript*	1 μl individual protein (Z or N or M)	or	2 μl binary complex (ZN or NM)	or	3 μl ternary complex (ZNM)	
1 μl buffer 5X						
0.8 μl H ₂ O						
	Qsp to 3 μl by protein buffer A					

*radioactive (“hot”).

TableM 10 Reaction buffer for RNase Z activity

BsuTrz		Trz1 (SceTrz)	
Buffer 10X		Buffer 10X	
200 mM Tris-HCl, pH 8.4	2.0 ml (1 M)	500 mM Tris-HCl, pH 7.1	5.0 ml (1 M)
10 mM MgCl ₂	0.1 ml (1 M)	50 mM MgCl ₂	0.5 ml (1 M)
10 mM KCl	0.1 ml (1 M)	50 mM KCl	0.5 ml (1 M)
10 mM DDT	0.1 ml (1 M)	20 mM DDT	0.2 ml (1 M)
	qsp 10 ml		qsp 10 ml

BsuTrz Buffer 1X: 20 mM Tris-HCl, pH 8.4, 1 mM MgCl₂, 1 mM KCl, 1 mM DDT.

Trz1 buffer 1X: 50 mM Tris-HCl, pH 7.1, 5 mM Mg Cl₂, 5 mM KCl, 2 mM DDT.

Protein buffer A: 20mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM DTT.

TableM 11 Manipulation of protein dilution.

	Mr (kDa)	1 mg/ml equals	Dilute to 6 pmol/μl		Dilute to 5 pmol/μl	
			Pr. (μl)	Buffer (μl)	Pr. (μl)	Buffer (μl)
Z	101 kDa	10 pmol/μl	6	4	5	5
N	33.8 kDa	30 pmol/μl	2	8	2	10
M	34.8 kDa	30 pmol/μl (28.7 pmol/μl)	2	8	2	10
ZN	134.8 kDa	7.4 pmol/μl	6	1.4	5	2.4
NM	68.6 kDa	14.6 pmol/μl	6	8.6	5	9.6
ZNM_poly	165.4 kDa	6 pmol/μl	No dilution		5	1

Dilution buffer: buffer A.

III.2 *Nuc1 activity assay using dsDNA as substrate*

Nuc1 or Nuc1-complexes were mixed with 200ng plasmid DNA (pUC18) in reaction buffer (50mM Tris pH7.5, 5mM MgCl₂, 1mM DTT) in a 10-μl-reaction volume, and incubated at 37 °C for 30 min. The 10-μl-reaction mixture were stopped by addition of 2 μl of 6X DNA loading dye (10 mM Tris-HCl, pH 7.6, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol 60 mM EDTA) and electrophoresis on 1% agarose gel, stained with SYBR and visualized by UV. The reaction mixture were prepared as shown in TableM 12.

TableM 12 Experiment set up for Nuc1 activity assay using dsDNA as substrate

Substrate Mix (7 μl)		Enzyme combination(s) (3 μl)				Final volume (10 μl)
Plasmid	1 μl (200 ng/μl)	1 μl individual protein (Z or N or M)	or	2 μl binary complex (ZN or NM)	or	
buffer 10X	1 μl	Qsp to 3 μl by protein buffer A				
H ₂ O	5 μl					

III.3 *Nuc1 activity assay using fluorescent molecular beacon as substrate*

A molecular beacon (5'-6-FAM-CAAGATATCCAGCTGCCCCCCCCCAGCTGGATATCTTG -BHQ-1-3') labeled at its 5' end with 6-FAM (6-Carboxyfluoresceine) and at its 3' end with BHQ-1 (Black Hole Quencher-1) were commercially synthesized (Eurofin™). The beacon was dissolved by applied buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) to 10 μM (200x). The assays were carried out in a 200 μl volume: 50 nM beacon was mixed with 5 μM of enzyme in the reaction buffer (50 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 1 mM DTT). The change in fluorescence over time was recorded at 25°C at an emission wavelength of 517 nm using a fluorometer.

The reaction mixture were prepared as shown in TableM 13. A preliminary experiment with various concentrations of Nuc1 was performed, in order to obtain a measurable fluorescence signal during the 20 min to 30 min recording time. Nuc1 concentrations at 50nM, 500nM, 1 μ M and 5 μ M were tested. And finally 5 μ M was also used for the Nuc1-relating complexes. A same reaction mixture without enzyme was used as control. An emission scan was performed with the control before and after the experiments (excitation at 495 nm) to verify the fluorescent peak was at 517 nm, proving the fluorescence dye was not degraded during experiments.

The fluorescence signal was recorded by the software supplied by the manufacture. The data was converted into the csv format and fed into the SigmaPLOT software for analysis.

TableM 13 Experiment set up for Nuc1 activity assay using fluorescent beacon as substrate

Compositions	Volume for 200 μ l reaction	Final concentration
Reaction buffer 10X	20 μ l	50mM Tris-HCl, pH 7.5, 5 mM MgCl ₂ , 1 mM DTT
Substrate 200X (10 μ M)	1 μ l	50 nM
Enzymes	Vpr*	5 μ M
H ₂ O	qsp 200 μ l	-

*Vpr: Volume of the proteins (Z, N, M) were calculated according to their stock concentration to reach a final concentration of 5 μ M in a 200 μ l reaction volume.

III.4 Mutarotase activity

An ammonium sulfate suspension of Glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides* (Sigma™) was centrifuged and the white pellet was dissolved in adjusted volume of reaction buffer (50 mM imidazole pH 7.5, 50 mM KCl, 8 mM MgSO₄), to a concentration of 9 mg/ml. According to the manufacture instructions this corresponds to an activity of about 7.5 units/ μ l (9 mg/ml * 822 units/mg). The substrates glucose-6-phosphate and β -NADP (Sigma™) were dissolved in reaction buffer at concentrations of 60 mM and 200 mM (20 mg in 1 ml), respectively.

Experiments were performed at 25 °C in the buffer 50mM imidazole, pH 7.5, 50 mM KCl, 8 mM MgSO₄ using an Applied Photophysics SW18-MV stopped-flow spectrometer. Reactions were initiated by mixing equal volumes of 30 μ M equilibrated Glc6P and 1 mM NADP⁺ with 75 units glucose-6-phosphate dehydrogenase, 1 mM NADP⁺, and various concentrations of mutarotase or related complexes (TableM 14). Stopped-flow data were

fitted to a double exponential using a non-linear least square curve fitting using software supplied by Applied Photophysics. Experiments were repeated at least five times.

A preliminary experiment with various concentration of mutarotase were performed, in order to find a proper concentration range where the two phases of the reaction curve could be separated. A starting concentration at 108 nM was too high that the second phase is too fast so that merge in to the first phase. Then then concentration was reduced to 54 nM and 27 nM. A control without any enzyme in the reaction mixture was performed at the starting of the experiments.

TableM 14 Experiment set up for mutarotase activity assay using Stopped-flow

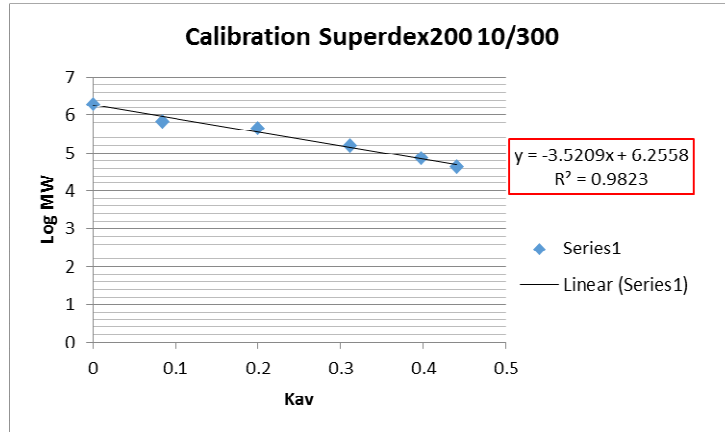
2ml	Chamber 1 (Substrates)	Chamber 2 (Enzymes)	2ml
1 μ l (60 mM)	α/β - Glc6P 30 μ M	Glc6P dehydrogenase (75 units/ml - excess)	20 μ l (7.5 units/ μ l)
10 μ l (200 mM)	NADP ⁺ 1 mM	NADP ⁺ 1 mM	10 μ l (200 mM)
-	--	Mutarotase or the complexes*	Vpr.
Qsp 2 ml	Reaction buffer	Reaction buffer	Qsp 2 ml

* Variable factor. *Vpr: Volume of the proteins (Z, N, M) were calculated according to their stock concentration to reach a final working concentration.

IV. Mr Estimation of proteins by gel filtration

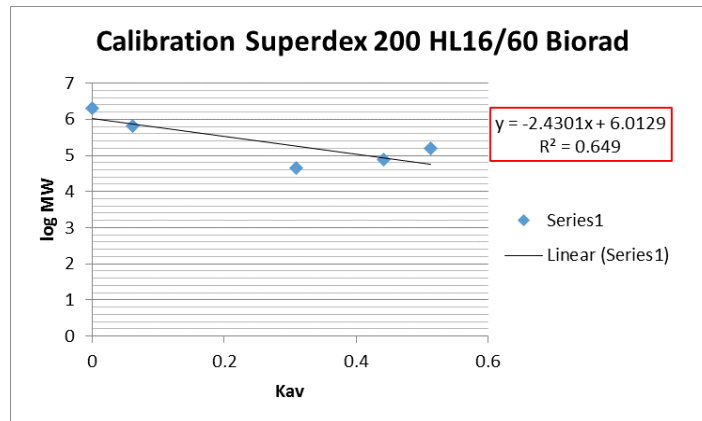
The experimental Mr of proteins obtained by gel-filtration is calculated from the protein elution volume obtained by gel filtration chromatography. The gel filtration columns were calibrated by several standard proteins. An linear fit equation established the correlation between molecular weights and elution volumes. Using the equation, we estimated the molecular weight of target protein by its elution volume.

Two columns were calibrated, which are Superdex 200 Hiload 16/60 (column volume 120 ml) and Superdex 200 10/300 GL (column volume 24 ml). The standard proteins are Aldolase, Conalbumin, Ovalbumin, Ferritin and Thyroglobulin, their molecular weights and elution volumes on column are listed below. Dextran blue was used to measure to the void volume (Vo) of the columns. All the runs were performed in the classic buffer A (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM DTT).



Calibration	Mr (Da)	Ve (ml)	Ve-Vo	Kav*	Log MW
Dextran Blue	2000 000	8.87	0	0	6.301029996
Aldolase	158 000	13.58	4.71	0.311302	5.198657087
Conalbumin	75 000	14.88	6.01	0.3972241	4.875061263
Ovalbumin	44 000	15.54	6.67	0.440846	4.643452676
Ferritin	440 000	11.9	3.03	0.2002644	5.643452676
Thyroglobulin	669 000	10.14	1.27	0.0839392	5.825426118

*Kav = (Ve - Vo) / (Vc - Vo). Vo: void volumn. Vc: column volumn: 24 ml. Ve: elution volumn.



Calibration	MW (Da)	Ve (ml)	Ve-Vo	Kav*	Log MW
Dextran Blue	2000 000	46.63	0	0	6.301029996
Aldolase	158 000	84.27	37.64	0.5130162	5.198657087
Conalbumin	75 000	79.04	32.41	0.4417337	4.875061263
Ovalbumin	44 000	69.29	22.66	0.3088456	4.643452676

*Kav = (Ve - Vo) / (Vc - Vo). Vo: void volumn. Vc: column volumn: 120 ml. Ve: elution volumn.

V. SEC-MALLS

The size exclusion chromatography (SEC) coupled to multi-angle laser light scattering (MALLS) uses multi-angle light scattering to characterize molar mass, molecular weight and

size of the proteins. The instrument does a triple detection of concentration, viscosity and light scattering. RI (deflection refractometry) measures accurately the concentration; the LALS (low angle light scattering) gives accurate molecular weight; and the differential viscometer (DV) provides all important structural data.

The SEC column was pre-equilibrated by the buffer 20 mM Tris-HCl, pH 7.5, 200 mM NaCl and 5 mM β ME. The protein samples were freshly purified with same buffer and injected to the SEC column Superdex 200 10/300 GL (GE healthcare™). The data was analyzed by the software Omniseq applied by the manufacture.

VI. Initial Data processing of Trz1 structure determination

Diffraction tests of the crystals were performed using the X-ray beam from the 3rd generation SUN (Source Optimized Light Energy Intermediate LURE Saint-Aubin, Proxima-1 lines or Proxima-2) or the ESRF (European Synchrotron Radiation Facility Grenoble, line ID23). The diffraction of the crystals appear as spots on the PILATUS detector and collected.

The collected images were processed with the XDS packages (Kabsch 2010). The initial data processing includes several steps:

- Determination of cell parameters and Bravais lattice characterizing the crystal lattice (program XDS)
- Measurement and integration of the intensities of the diffracted reflections (program XDS)
- Scaling (program XSCALE)
- Determination of space group from systematic extinctions, if possible.
- Estimated number of molecules in the asymmetric unit

VI.1.1 Determination of lattice parameters and Bravais lattice

The first step consists of image indexing. This adjustment is usually performed by measuring the position of the diffraction spots and determining the three lattice vectors a^* , b^* and c^* compatible with their positions. These vectors are used to calculate the dimensions (a, b, c) and angles (α , β , γ) that they form, which define the crystal lattice parameters. The XDS software (Kabsch 2010) provides a list of different possible Bravais lattices for

measured reflection. Each solution is associated with a proportional penalty of the difference between the estimated lattice parameters and geometric constraints imposed by the proposed Bravais lattice.

A Bravais lattice is considered correct when it combines the highest symmetry and the lowest penalty score (generally less than 20). For each Bravais lattice, there may be several Laue groups, or space groups, depending on the internal symmetry of the cell.

VI.1.2 Integration

The collected images are indexed one by one, by refining the estimated cell parameters of the selected space group, the crystal-detector distance, the crystal orientation and mosaicity. The positions of the spots are predicted and adjusted to the actually observed spots. Each spot is indexed.

The intensity of the spots is measured by applying a shell around the spot: the center of the shell corresponds to the integrated area, the outer shells corresponds to background noise. Thus, each spot intensity can be corrected against the background noise.

VI.1.3 Scaling

Each image collected contain a fraction of all possible reflections, obtained by the complete rotation of the crystal in the beam. The images are then integrated to form a complete dataset. However, variations in intensity may occur during collection. They are caused by the change in beam intensity, and most importantly by the decrease of the intensity of the spots caused by radiation damage of the crystal during the collection. A correction factor is applied by comparing the symmetry related reflections.

The quality of the dataset is evaluated:

- By comparing the intensity of each symmetry equivalent reflection to the average value of this reflection quantified by the values R_{meas} or R_{sym} (Karplus and Diederichs, 2012), defined by

$$R_{sym}(I) = \frac{\sum_{hkl} \sum_{j=1}^{j=n_{hkl}} | \langle I_{hkl} \rangle - I_{hkl,j} |}{\sum_{hkl} \sum_{j=1}^{j=n_{hkl}} I_{hkl,j}} \quad ; \quad \langle I_{hkl} \rangle = \frac{1}{n_{hkl}} \sum_{j=1}^{j=n_{hkl}} I_{hkl,j}$$

Where $I_{hkl,j}$ measure the reflection (h, k, l) and $\langle I_{hkl} \rangle$ represent the average density of all equivalent reflections (h, k, l) measured.

- And by the signal to noise ratio, $I/\sigma(I)$, where $\sigma(I)$ is the standard deviation of the intensity and represents background noise. The value of $I/\sigma(I)$ decreases as the resolution

increases, reflecting a decrease in the diffraction intensity. R_{sym} values and $I/\sigma(I)$ are used to decide the maximum resolution at which data may be used.

VII. SAXS:

VII.1 Principle

The small angle X-ray scattering (SAXS) is a contrast method where the scattering signal is derived from the difference in the electron density, $\Delta\rho(\mathbf{r})$, of solute particles of interest, $\rho(\mathbf{r})$, and the solvent (ρ_0), that is

$$\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_0$$

Where ρ_0 is $\sim 0.334 \text{ e}^-/\text{\AA}^3$ for pure water; and $\rho(\mathbf{r})$ of the proteins is typically $\sim 0.44 \text{ e}^-/\text{\AA}^3$ (Putnam et al. 2007). In practice, data are collected on a sample and a buffer as blank. The scattering signal due to proteins is obtained by subtracting the buffer scattering from the protein solution scattering. The resulting scattering pattern, $I(q)$, is radially symmetric (isotropic) because particles in solution have random orientation. The modulus q of the momentum transfer is given by:

$$q = 4\pi \sin\theta / \lambda$$

where 2θ is the scattering angle, and λ is the wavelength of X-ray beam (Figure MSAX-1). The ideality and monodispersity of the sample has to be checked. The monodispersity means the molecules of the sample are identical, which could be checked separately during the purification protocol by size-exclusion chromatography; and the ideality of the sample means there are no intermolecular interactions. In these conditions $I(q)$ can be expressed as:

$$I(q) = N \left| \int_V \Delta\rho(\vec{r}) e^{-i\vec{q}\cdot\vec{r}} d\vec{r} \right|^2$$

where N is the number of particles and V the particle volume.

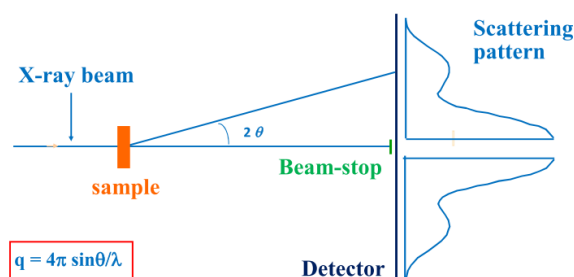


Figure MM-1. Experimental set up of SAXS (P. Vachette).

VII.2 *Sample preparation and data collection*

The protein samples Trz1, Nuc1, and mutarotase were individually purified as described before, and were solubilized in the buffer 20 mM Tris-HCl, pH 7.5, 200 mM NaCl. Trz1 and mutarotase stored at -80 °C were thawed on ice right before the experiment. Nuc1 was freshly purified and 1mM MgCl₂ was added after purification. The binary complexes ZN and NM were prepared by mixing the two partners at equal molar. The ternary complex ZNM was freshly purified separately after recombinant expression using a polycistronic plasmid (thus Z was tag-free in the ZNM complex). All the samples were centrifuged at 12,000 rpm at 4 °C for 30 min before injection into the SAXS flow-through capillary cell mounted on the SAXS beamline SWING (Synchrotron SOLEIL, Saint-Aubin, France).

SAXS experiments were carried out on the SWING beamline at the SOLEIL synchrotron radiation facility (Saint-Aubin, France). The sample to detector (Aviex CCD) distance was set to 1820 mm and the wavelength λ to 1.0 Å, allowing useful data collection over the momentum transfer range $0.006 \text{ \AA}^{-1} < q < 0.5 \text{ \AA}^{-1}$. The solution was injected into the SAXS flow-through capillary cell (1.5 mm in diameter) under vacuum. In order to use mono-disperse solutions devoid of aggregates, SAXS data were collected directly after elution through an on-line size-exclusion high-performance liquid chromatography (SEC-HPLC) column available on SWING (Saint-Aubin, France). Flow rate was 200 $\mu\text{l}/\text{min}$, frame duration was 1 s and the dead time between frames was 0.5 s. For each frame, the protein concentration (below than 1 mg/ml at the top of elution peak) was estimated from UV absorption at 280 nm using a spectrometer located immediately upstream of the SAXS measuring cell.

Data were first analyzed using Foxtrot, the Swing in-house software, and the program package Primus (<https://www.embl-hamburg.de/biosaxs/primus.html>). Scattering of the elution buffer before void volume was recorded and subtracted from all protein scattering curves. The scattered intensities were displayed on an absolute scale using the scattering by water. In order to obtain the scattering curve of the main species fully devoid from the contribution of aggregates, the elution profiles were fitted by a set of Gaussian curves using the program US-SOMO (Brookes et al. 2016).

VII.3 Preliminary data processing (Putnam et al., 2007; P. Vachette)

The SAXS data were processed by Dr. Dominique Durand from our team.

VII.3.1 Radius of gyration (R_g) obtained by the Guinier approximation

The overall size of the particles, described by the Radius of gyration (R_g), is estimated by the Guinier approximation.

$$I(q) \cong I(0)e^{-\frac{q^2 R_g^2}{3}}; \quad \ln[I(q)] = \ln[I(0)] - \frac{q^2 R_g^2}{3}$$

Where $I(0)$ is the intensity of the scattering profile extrapolated to $q = 0$. This is the most common method of estimating R_g . It is measured via the slope of the plot $\ln[I(q)]$ vs. q^2 ("Guinier plot" in the program <Primus>). For globular molecules $qR_g < 1.3$.

VII.3.2 Distance-distribution Function $P(r)$ and Maximum dimension (D_{max})

The Distance-distribution Function $P(r)$ is the histogram of the distances between any pair of scattering elements within the particles. $P(r)$ can be directly calculated from $I(q)$ with

$$P(r) = \frac{r}{2\pi^2} \int_0^\infty I(q)q \sin(qr) dq$$

Indirect Fourier transform methods have been developed for calculating $P(r)$ (Program GNOM), based on the hypothesis that the particle has a finite size characterized by D_{max} .

D_{max} is the maximum dimension of the particle, and it is the value of r at $P(r) = 0$ for large r . Assignment of D_{max} may be complicated by flexibility or multimerization.

The Radius of gyration (R_g) could also be obtained from the $P(r)$ function using the equation

$$R_G^2 = \frac{\int_0^{D_{max}} r^2 P(r) dr}{\int_0^{D_{max}} P(r) dr}$$

This alternative estimation of R_g is a good consistency check for R_g , D_{max} , and $P(r)$.

VII.3.3 Molar Mass (MM) and Volume (V)

The molar mass (MM) of the scattering objects is usually calculated from the zero-angle scattered intensity $I(0)$ which is proportional to MM and to the protein concentration, following the expression:

$$I(0) = \frac{c \cdot MM \cdot r_0^2}{N_A} \cdot \left[v_p (\rho_{\text{prot}} - \rho_{\text{buf}}) \right]^2$$

where c is the mass concentration, r_0 the classical electron radius, v_p the protein specific volume and $\rho_{\text{prot}} - \rho_{\text{buf}}$ the electronic density contrast. In fact, this approach is only suitable if c is precisely known.

An alternative estimate of the molar mass can be obtained using the SAXSMoW2 program (available at URL <http://saxs.ifsc.usp.br/>) that is based on the determination of the Prod volume given by:

$$V = 2\pi^2 I_{\text{exp}}^2(0) / \left(\int_0^\infty I(q) q^2 dq \right)$$

Interestingly, V is independent of the sample concentration.

A third method uses the program ScÅtter (available at URL <https://bl1231.als.lbl.gov/scatter/>) based on the so-called correlation volume V_c also independent of the concentration:

$$V_c = \frac{I(0)}{\int q \cdot I(q) dq}$$

VII.3.4 Construction of the model from the SAXS data

For a molecule whose structure is totally unknown, the *ab initio* program DAMMIN (Svergun 1999) can be used. This program describes the particle as a compact ensemble of identical spheres the scattering of which is as close as possible to the experimental curve. Their assembly yields a uniform volume that accounts for the shape of the scattering object. We obtained the envelope of the NM and ZNM complexes using this program. The program GASBOR, which describes the protein as a chain of N dummy residues where N is the number of amino acid residues, could also be used to construct an envelope but it is not suitable in the case of this work because of the large number of atoms.

To construct the Trz1 model from X-ray scattering data, whose crystal structure was provided but several flexible parts were missing, the program AllosMod-FoXS (Weinkam et al. 2012; Schneidman-Duhovny et al. 2010) was used.

For constructing an atomic model of NM and ZNM complex from SAXS data, whose atomic model of subunits was already known, we used the program SASREF (Petoukhov & Svergun 2005). SASREF performs rigid body modeling of a complex formed by several subunits with known atomic structure against solution scattering data, which means positioning of the subunits in the complexes.

VII.3.5 Final validation of the model

The program CRY SOL (Svergun et al. 1995) is used for calculating the scattering curve from an atomic model and adjusting it to experimental scattering curve. Fitting is done by varying three parameters: (i) the radius of the atomic group (ii) the excluded volume (iii) the contrast of the hydration shell. The goodness of fit is generally characterized by χ^2 with the following function:

$$\chi^2 = \frac{1}{N-1} \sum_j \left[\frac{I_{exp}(q_j) - cI_{calc}(q_j)}{\sigma(q_j)} \right]^2$$

Where N is the number of experimental points, c is a scaling factor, and $I_{calc}(q_i)$ and $\sigma(q_i)$ are the calculated intensity and the experimental error at the scattering vector q_j , respectively.

ANNEX

❖ Protein sequences in the constructs

▪ Trz1-His

```

MAHHHHHHVHG TGSNDDDDKS PDPNWELVYT ARLQEF

 2  sktvnfrat knfylqfvsv ssrdtscipc ihlfffskry vfgsvgegcq railsqqrlr
61  skikdvflmq gssisspdy dssssstts vsdmlqlddr dkvivserns mcstvnyptw
121 wdscggfpgf llslndisep getgeaspfv lhgpsvehqf lssmrhftyh tnvnlvtqgy
181 tsaepvfvd enicvtpvvv slvknsfkkk kheninrgtn arplkedran tsphwshvs
241 ndtsfvvena myntpaplep dkpelfisyi vqshptpgkf daakakslgi tkgldcgrla
301 rgepvtleng ktvypkevig psipgssffi ihcpnelvid lvienhkwn apkpvsvihs
361 vtpevyknpr yqswissfpe evshliaste vnevinyprs avaiatlntl dskvplgfn
421 cyevknvqn nriafakpkl rfafgkktgi ddsevgsie elkdilkek pdyksfveea
481 qkyvskpka psfagsdiqi ctltgtsamp slyrnvssty vripvdkkcm edsaismkn
541 lldcgegtlg rlsrqygdnl kyeiaslrwi yishmhadhh agvigvlkaw tkysdgrskl
601 fitappqfef wlleysridy lplsnivfis nsalrtdrpk salesrlss lfkefdlvf
661 rtvpaihcpy sycmeitnss gwkiaysgdt rpsedfania kdstllieha tledsmheia
721 ikkqhstyse alevakkagt knvilthfsq rypklpdidi stedlhiala fdgmtlkisd
781 islfryfgkp laylfneenl keesdplkf

```

MAHHHHHHVHG TGSNDDDDKS PDPNWELVYT ARLQEF: HIS-tag, enterokinase cleavage site (underline) and linker

▪ Strep -Nuc1Δ43

```

MWSHPQFEK
24 NSNIQSH SFNVDPGSGFF
61 KYGFFPGPIHD LQNREEFISC YNRQTQNPYW VLEHITPESL AARNADRKNS FFKEDEVIPE
121 KFRGKLRDYF RSGYDRGHA PAADAKFSQQ AMDDTFYLSN MCPQVGEFNF RYWAHLEYF
181 CRGLTKKYKS VRIVTGPLYL PKKDPIDNKF RVNYEVIGNP PSIAVPTFFF KLIVAEAPTA
241 NPAREDIAVA AFVLPNEPIS NETKLTDFEV PIDALERSTG LELLQKVPPS KKKALCKEVN
301 CQIVVRDFSN AAIKQSKDVK LLPPPCKRN

```

MWSHPQFEK: Strep-tag.

▪ Mutarotase-His

```

1 MPIKETDKEV VLTHPADETT SVHILKYGAT VYSWKLKSEE QLWLSTAAKL DGSKPVRGGI
61 PLVFPVFGKN STDEHLSKLP QHGLARNSTW EFLGQTKENP PTVQFGLKPE IANPELTKLW
121 PMDYLLILTV ELGSDYLKTA IEVENTSSSK ELKFNWLFHT YFRIEDIEGT MVSNLAGMKL
181 YDQLLKESYV DKHPVVTFNQ ETDVIYQNVS AERAIQIVDK GVQIHTLKRY NLPDVTWVNP
241 WIEKSQGMAD FEPKTYGQM ICIEPGHVHD FISLAPGKKW NAYQLLCKEE LKYQAIQHHHHHH

```

HHHHHH: His-tag.

▪ Sequence of the transcription region in the ZNM polycistronic plasmid

```

CATATGTTACATTTATACCCATCACCCATCCTACATCGGATACAAAGCACCCATTGCTGCTAGTCCAGTCTGCACATGGGGA
AAAGTATTTCTTCGGTAAAATTGGTGAAGGATCCCAAAGGAGTCTGACTGAAAATAAGATCAGGATATCCAAATTGAAGGATA
TTTTCCTTACTGGTGAATTAACCTGGTCAGATATAGGTGGATTACCTGGAATGATTTTGACTATTGCTGATCAAGGGAAAAGT
AATCTTGTTTTGCATTACGGCAATGACATTTTGAATTACATAGTTTCCACTTGGAGATACTTCGTCTTTAGATTCCGGAATAGA
CTTGAACGATCACATTATGAAAGACAAGGAAGTATATAAAGATAAGATAATAGCTGTTAAATCCTTTAATGTTCTGAAAAATG
GGGGGAAGACAGGTTAGGCGTCTTCGATAGTTTTCAAAAAGGTGTATTACGTTCCATAGTAGCAAAAATGTTCCCAAACAT

```

GCACCCACCGATAGGTACGATCCTTCTAGTGATCCGCACTTGAATGTAGAGTTGCCTGACTTAGACGCCAAAGTGAAGTTTC
TACGAATTACGAGATTTCAATCAGTCCAGTGAGGGTAAATTTAAAGTGAGGAAGCTATTAACTAGGTGTTCCGAAGGGTC
CCTTATTTGCAAAGTTAACCAAGGGCCAAACAATTACTTTGGATAACGGTATTGTTGTAACCTCCGAACAGGTATTGGAGAAT
GAACGTCATTTTGCCAAAGTATTGATCCTGGATATCCCAGATGACCTATATTTGAACGCTTTCGTAGAAAAATTCAAGGATTA
TGATTGTGCTGAGCTTGGCATGGTGTATTATTTTCTTGGTGATGAGGTTACCATTAATGATAATCTATTCGCGTTCATTGACA
TATTTGAGAAAAACAATTATGGTAAAGTAAAT CACATGATATCCCACAATAAAATTTCTCCAAACACGATATCATTTTTTCGGT
TCTGCATTGACCACATTGAAATTAAGGCACTACAAGTAAATAATTACAATTTACCAAAAACAGATCGTGTGTTTTCCAAGGA
CTTCTACGACAGATTTCGATACACCCTCAGCAGAGGTACATCTATGTGTAAATCCCAGGAAGAGCCTTTGAATACAATAATAG
AGAAGGATAACATTCATATTTTTTCCAAAAACAAGACAGTAACTTTGAAACCATTTCCGGATGAACGAAGAACCAGTGAATGC
AACATCAACGGTGAAGTGGCGGATTTCTCGTGGCAAGAAATTTTCGAAGAACATGTAAAACCATTAGAATTTCCCTTAGCTGA
TGTCGATACAGTTATCAATAATCAACTACAGTGGATAACTTTAACAATTCAGCAGAAAAAGAAACACGTTGAAATTTATCA
CCTTAGGAACCGGTAGTGCATTGCCTTCTAAATATAGAAACGTTGTCTCCACACTTGTAAAGTTCCCTTTACTGACGCCGAT
GGAAATACCATAAATAGAAACATTTATGCTAGATGCTGGTGAATAACTTTAGGTACCATACACAGGATGTTTTCTCAGCTAGC
AGTCAAGTCAATATTTAGGATTTGAAATGATATATCTGAGTCACTTGCATGCAGACCACCATTTGGGAATAATCAGCGTGC
TAAATGAATGGTACAAATATAACAAGGATGATGAAACGAGTTATATATATGTGGTTACT CCGTGGCAATATCACAAATTTGTT
AATGAATGGTTAGTTCTAGAAAAATAAGAGATTTTAAAGAGAATCAAATACATAAGTTGTGAGCATTTCATCAATGATTCGTT
TGTAAGAATGCAGACACAATCTGTTCTTTGGCAGAGTTCAATGAAATATTGAAAGAAAAATAGCAATCAAGAATCAAACGAA
AACTGGAACCTGGATAGAGATTTCTCATATAGGGATGTTGACTTGATCAGACAAATGTATGAGGATTTATCGATAGAATATTTT
CAAACCTGCAGAGCTATACATTGTGACTGGGCATATTCGAACTCAATTACCTTCCGAATGGACGAAAACAATGAGCATAAATAC
ATTCAGGTTTCATATTCAGGCGATACAAGACCTAACATCGAGAAATTTTCCCTCGAAATAGGCTATAAATTCAGATCTATTA
TTCACGAAGCTACACTAGAAAAATCAGCTACTGGAGGATGCCGTGAAGAAAAAACACTGCACTATTAATGAAGCAATCGGTGTT
TCGAACAAAATGAATGCTAGGAAGTTGATCTTAACACACTTTTCCCAGAGATATCCCAAATGCCCCAATTAGACAATAATAT
TGATGTGATGGCGAGAGAATTTTGTGTTGCTTTTCGACAGTATGATCGTTGATTATGAGAAAAATGGTGAACAGCAGCGTATTT
TTCCACTGCTGAATAAGGCATTTGTTGAAGAAAAGGAAGAAGAAGATGTTGATGACGTTGAAAGCGTACAAGATTTGGAA
GTCAAACCTAAGAAACACAAGAAAAAT TAActcta aagctta aat tttgt ttaacttta ag aaggag atata CCatg GGC TGA
GCCACCCG CAGTTCGAAAAA AATAGTAATATTCAATCTCACTCTTCAACGTCGATCCTTCCGGGTTCTTCAAGTATGGTTTT
CCTGGTCCGATTCATGATCTGCAGAACCGGAAGAGTTTATCTCATGTTACAACAGACAAACTCAAATCCTTATTGGGTCTCT
CGAACATATAACGCCAGAATCATTTGGCTGCAAGGAATGCTGACAGAAAAAATTCCTTTTTCAAGGAAGATGAAGTAATTCAG
AAAAGTTTAGAGGTAAACTAAGAGACTATTTTAGGTGGGCTATGATCGAGGCCATCAAGCCCAGCTGCAGACGCAAAATTT
TCTCAACAGG C GATGGATGATACATCTACTTATCCAAATATGTGTCTCAAGTAGGAGAAGGTTTTAATAGAGACTATTGGGC
GCATTTGGAGTACTTTTGTAGGGGATGACTAAGAAATATAAGAGTGAAGAATCGTGACTGGTCCATTGTATCTACCAAAA
AGGATCCCATAGATAATAAATTTAGGGTTAATTATGAAGTTATTGGCAATCCACCCAGTATTGCTGTTCCAACGCACCTTTTTT
AAATTGATTGTTGCAGAACCAACAGCCAATCCTGCTAGAGAGGATATTGCTGTGCGGCATTTGTATTGCCAAACGAACC
GATATCAAATGAGACGAAATGACTGACTTTGAAGTTCTATAGATGCTTTAGAGAGAAGTACTGGGCTAGAACTTCTGCAAA
AAGTACCACCTTCAAAGAAGAAGGCATTTATGCAAAGAGGTAAATTTGCAAAATGTAGTGAGAGATTTCTCTAACGCGGCATC
AAACAATCGAAAGATGTGAAATGTTACCTCCTCAAAAAAAGGAAT TGA ctAGa GAATTC at tttgt ttaacttta ag aag
gag atata CATATGCCTATCAAAGAACTGATAAAGAGTTGTTTTGACTCATCCAGCTGATGAGACCACCAGCGTTCATATT
CTAAAGTACGGTGCTACAGTTTATTCTTGGAAATGAAATCTGAAGAACAGTTGTGGTTGTCTACTGCTGCTAAATTTGGATGG
TAGCAAACTGTGAGAGGTGGTATACCTTTGGTCTTTCCTGTATTCCGGAAAAATAGCACCGATGAACATTTGAGTAAATTA
CTCAACATGGTCTTGAAGAAATCTACTTGGGAGTTTTTGGGTCAAACCTAAGGAAAACCCACCGACCGTACAATTTGGCTTG
AAACCAGAAATGCTAACCCAGAATTGACCAATTTGTGGCAATGGATTATCTTTTGATTTTGACTGTTGAATTTAGGCTCCGA
TTATTTGAAACTGCCATAGAAGTAGAAAACACATCTAGTTCCAAGGAATTAAGTTCAACTGGTTGTTCCATACATACTTCC
GTATCGAAGATATTGAAGGAACAATGGTCTCTAATTTAGCTGGCATGAACTTTATGACCAACTGTTGAAGGAATCCTACGTC
GACAAGCACCCAGTCGTTACCTTCAATCAAGAAACCGATGTAATTTATCAGAATGTGAGCGCCGAACGGGCCATTCAAATAGT
TGACAAGGGCGTTCAAATTCACACTCTAAAAAGATACAACCTTGCCCGACACTGTTGTTTGAAT CCGTGGATTGAGAAGTCTC

```
AAGGGATGGCCGATTTCGAACCAAAAACCTGGTTACCAACAAATGATATGTATTGAACCTGGTCATGTTTCATGATTTTATTCC
TTGGCTCCTGGTAAAAAATGGAATGCTTATCAATTACTTTGCAAAGAAGAATTGAAATATCAAGCTATTCAACACCATCACCA
TCACCATTAACtcGAg
```

*Legends:

ORF	-	bases in grey background
RBS sequence	enhancer + SD	TTAACTTTAAGAAGGAG
Tags	STREP-Tag	TGG AGC CAC CCG CAG TTC GAA AAA
	His-tag	CAC CAT CAC CAT CAC CAT
Enzyme restriction sites	NdeI	CATATG
	HindIII	A AGCTT
	EcoRI	G AATTC
	XhoI	C TCGAG
	NcoI	CCATGG
Mutations	TRZ1	T945C (CAT=CAC=His)
Mutation to remove NcoI sites (CCATGG)	TRZ1	CCATGG to CCCTGG, A1719G (CCA = CCG = Pro)
	NUC1	GCCATGG to GCCATGG, C453G (GCC = GCG = Ala)
	YMR099C	CCATGG to CCGTGG, A720G (CCA = CCG = Pro)

❖ **IDs of the sequences used for generating conservation surface to Trz1 structure**

Input_pdb_ATOM_A
UniRef90_A0A0P1KXD7_1_804
UniRef90_G0V5Z9_1_814
UniRef90_G8ZP37_1_800
UniRef90_J8TXF2_1_818
UniRef90_J8Q5H1_1_809
UniRef90_P36159_1_809
UniRef90_A0A0L8RG30_1_818
UniRef90_C5DNU3_1_799
UniRef90_W0W1Y6_1_810
UniRef90_C5DJM8_1_808
UniRef90_H2AYP1_1_802
UniRef90_A7TQ27_1_807
UniRef90_A0A0C7MLM9_1_812
UniRef90_A0A0W0E6U4_1_813
UniRef90_G0W3T6_1_847
UniRef90_G8JSX7_1_823
UniRef90_W0T9S2_1_802
UniRef90_G8BX39_1_825
UniRef90_I2GX64_1_818
UniRef90_Q759R1_1_818
UniRef90_A0A0A8L0R9_1_810
UniRef90_Q6CJR0_1_810
UniRef90_J7S6F8_1_797
UniRef90_I2H450_1_806
UniRef90_A0A0X8HTN5_1_809
UniRef90_K0KNQ0_1_776
UniRef90_A0A061AR67_1_818
UniRef90_A0A0H5BZC1_1_798

❖ **Interface analysis of N-terminal domain and C-terminal domains of Trz1** (by PDBePISA)

- Structure 1: Trz1 C-terminal domain (residues 466-818)

- Structure 2: Trz1 N-terminal domain (residues 1-405)

Interface Summary

	<u>Structure 1</u>		<u>Structure 2</u>	
<u>Selection range</u>	B		A	
class	Protein		Protein	
symmetry operation	x,y,z		x,y,z	
symmetry ID	1_555		0_555	
<u>Number of atoms</u>				
interface	284	10.1%	299	10.1%
surface	1492	53.0%	1681	56.8%
total	2814	100.0%	2961	100.0%
<u>Number of residues</u>				
interface	70	20.5%	73	20.0%
surface	300	88.0%	335	91.8%
total	341	100.0%	365	100.0%
<u>Solvent-accessible area, Å</u>				
interface	2822.6	16.6%	2782.3	14.9%
total	17048.4	100.0%	18623.8	100.0%
<u>Solvation energy, kcal/mol</u>				
isolated structure	-305.1	100.0%	-354.4	100.0%
gain on complex formation	-17.2	5.6%	-8.3	2.3%
average gain	-4.8	1.6%	-4.8	1.4%
P-value	0.010		0.268	

Interface details

Table X-1: Hydrogen bonds

##		CtD			Dist. [Å]	NtD			
1		B:ASN	512	[ND2]	3.32	A:ASP	13	[O]	
2		B:ARG	519	[NE]	2.74	A:ASP	13	[OD2]	
3		B:ARG	519	[NH1]	2.58	A:ASN	325	[OD1]	
4		B:ARG	519	[NH1]	3.35	A:ASN	325	[O]	
5		B:ARG	519	[NH1]	3.37	A:LYS	351	[O]	
6		B:ARG	519	[NH2]	2.92	A:THR	326	[O]	
7		B:HIS	542	[NE2]	3.68	A:ASP	172	[OD1]	
8	α 13	B:SER	551	[OG]	2.85	A:GLU	36	[OE2]	α 1
9	α 15	B:GLN	575	[NE2]	3.60	A:SER	174	[O]	Ex-I
10	Ex-III	B:LYS	589	[NZ]	3.53	A:LEU	186	[O]	Ex-I
11		B:ARG	610	[NH2]	3.32	A:SER	175	[OG]	
12		B:SER	615	[OG]	3.32	A:ASN	180	[OD1]	
13		B:VAL	616	[N]	3.11	A:ASN	180	[O]	
14		B:ASP	672	[N]	2.80	A:TYR	171	[OH]	
15		B:TRP	673	[NE1]	3.86	A:SER	175	[OG]	
16		B:SER	787	[OG]	3.71	A:ASN	348	[O]	
17		B:ILE	789	[N]	2.65	A:ASN	346	[O]	
18		B:SER	479	[O]	2.76	A:ARG	40	[NE]	α 1
19		B:LYS	480	[O]	3.25	A:ARG	40	[NH2]	α 1
20		B:LYS	480	[O]	3.45	A:SER	332	[N]	
21	α 11	B:ARG	519	[O]	2.86	A:LYS	351	[N]	
22		B:MSE	520	[O]	2.99	A:ASN	348	[ND2]	
23		B:LEU	547	[O]	3.02	A:TRP	63	[NE1]	
24	α 14	B:GLU	582	[OE2]	2.36	A:ARG	101	[NH2]	α 3
25	α 14	B:GLU	582	[O]	3.04	A:TYR	95	[OH]	α 3
26	Ex-III	B:THR	613	[OG1]	3.00	A:ASN	180	[ND2]	Ex-I
27		B:THR	613	[O]	2.86	A:ASN	180	[ND2]	
28		B:VAL	616	[O]	3.48	A:GLU	182	[N]	
29		B:ASP	672	[OD2]	2.85	A:SER	175	[OG]	
30		B:ASP	672	[O]	3.73	A:SER	174	[OG]	
31	Tail	B:SER	787	[OG]	3.82	A:ASN	348	[N]	Ex-II
32		B:GLU	797	[OE2]	3.85	A:ASN	346	[ND2]	
33		B:GLN	798	[OE1]	3.54	A:GLN	343	[NE2]	
34		B:GLU	812	[OE1]	2.14	A:LYS	338	[NZ]	

Table X-2: Salt bridges

##		CtD			Dist. [Å]	NtD			
1		B:ARG	519	[NE]	3.57	A:ASP	13	[OD1]	
2		B:ARG	519	[NE]	2.74	A:ASP	13	[OD2]	
3		B:ARG	519	[NH2]	2.82	A:ASP	13	[OD2]	
4		B:HIS	542	[NE2]	3.68	A:ASP	172	[OD1]	
5	α 14	B:GLU	582	[OE1]	2.52	A:ARG	101	[NH2]	α 3
6		B:GLU	582	[OE2]	3.91	A:ARG	101	[NH1]	
7		B:GLU	582	[OE2]	2.36	A:ARG	101	[NH2]	
8	α 1	B:GLU	812	[OE1]	2.14	A:LYS	338	[NZ]	Ex-III
9		B:GLU	812	[OE2]	3.90	A:LYS	338	[NZ]	

REFERENCE

Aravind, L., 1999. An evolutionary classification of the metallo-beta-lactamase fold proteins. *In Silico Biol*, 1(2), pp.69–91. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11471246> <http://iospress.metapress.com/content/7wu0chexlj69nwmt/?genre=article&issn=1386-6338&volume=1&issue=2&spage=69>.

Arnoult, D. et al., 2003. Mitochondrial release of AIF and EndoG requires caspase activation downstream of Bax/Bak-mediated permeabilization. *The EMBO journal*, 22(17), pp.4385–99. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12941691> [Accessed September 26, 2016].

Asha, P.K. et al., 1983. Ribonuclease BN: identification and partial characterization of a new tRNA processing enzyme. *Proceedings of the National Academy of Sciences of the United States of America*, 80(11), pp.3301–4. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/6344080> [Accessed December 19, 2016].

Bahi, N. et al., 2006. Switch from Caspase-dependent to Caspase-independent Death during Heart Development: ESSENTIAL ROLE OF ENDONUCLEASE G IN ISCHEMIA-INDUCED DNA PROCESSING OF DIFFERENTIATED CARDIOMYOCYTES. *Journal of Biological Chemistry*, 281(32), pp.22943–22952. Available at: <http://www.jbc.org/cgi/doi/10.1074/jbc.M601025200> [Accessed July 30, 2016].

Bebrone, C., 2007. Metallo-β-lactamases (classification, activity, genetic organization, structure, zinc coordination) and their superfamily. *Biochemical Pharmacology*, 74(12), pp.1686–1701.

Benschop, J.J. et al., 2010. A consensus of core protein complex compositions for *Saccharomyces cerevisiae*. *Molecular cell*, 38(6), pp.916–28. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20620961> [Accessed December 13, 2014].

Brookes, E. et al., 2016. US-SOMO HPLC-SAXS module: dealing with capillary fouling and extraction of pure component patterns from poorly resolved SEC-SAXS data. *Journal of applied crystallography*, 49(Pt 5), pp.1827–1841. Available at: <http://scripts.iucr.org/cgi-bin/paper?S1600576716011201> [Accessed October 26, 2016].

Brzezniak, L.K. et al., 2011. Involvement of human ELAC2 gene product in 3' end processing of mitochondrial tRNAs. *RNA Biology*, 8(4), pp.616–626.

Burhans, W.C. & Weinberger, M., 2007. Yeast Endonuclease G: Complex Matters of Death, and of Life. *Molecular Cell*, 25(3), pp.323–325. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17289580>.

Büttner, S., Carmona-Gutierrez, D., et al., 2007. Depletion of endonuclease G selectively kills polyploid cells. *Cell Cycle*, 6(9), pp.1072–1076.

Büttner, S., Eisenberg, T., et al., 2007. Endonuclease G regulates budding yeast life and death. *Molecular cell*, 25(2), pp.233–46. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17244531>.

Canino, G. et al., 2009. Arabidopsis encodes four tRNase Z enzymes. *Plant physiology*, 150(3), pp.1494–502. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2705019&tool=pmcentrez&rendertype=abstract> [Accessed October 14, 2013].

Carfi, A. et al., 1995. The 3-D structure of a zinc metallo-beta-lactamase from *Bacillus cereus* reveals a new type of protein fold. *The EMBO journal*, 14(20), pp.4914–21. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=394593&tool=pmcentrez&rendertype=abstract>.

Ceballos, M. & Vioque, A., 2007. tRNase Z. *Protein and peptide letters*, 14(2), pp.137–45. Available

at:

<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=17305600&retmode=ref&cmd=prlinks%5Cnpapers3://publication/uuid/AB0AC4A0-338B-4FD2-9172-D04ECAA44D64>.

Chen, Y. et al., 2005. Characterization of TRZ1, a yeast homolog of the human candidate prostate cancer susceptibility gene ELAC2 encoding tRNase Z. *BMC molecular biology*, 6, p.12. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1156898&tool=pmcentrez&rendertype=abstract> [Accessed October 14, 2013].

Cheung, W.L. et al., 2003. Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase. *Cell*, 113(4), pp.507–17. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12757711>.

Condon, C. & Putzer, H., 2002. The phylogenetic distribution of bacterial ribonucleases. *Nucleic acids research*, 30(24), pp.5339–46. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=140075&tool=pmcentrez&rendertype=abstract>.

Côté, J., Renaud, J. & Ruiz-Carrillo, A., 1989. Recognition of (dG)_n(dC)_n sequences by endonuclease G. Characterization of the calf thymus nuclease. *The Journal of biological chemistry*, 264(6), pp.3301–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2914952>.

Cowan, K., 2006. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D Biological Crystallography*, 62(9), pp.1002–1011. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0907444906022116>.

Daiyasu, H. et al., 2001. Expansion of the zinc metallo-hydrolase family of the β -lactamase fold. *FEBS Letters*, 503(1), pp.1–6.

Dake, E. et al., 1988. Purification and properties of the major nuclease from mitochondria of *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 263(16), pp.7691–7702. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3286639>.

David, K.K. et al., 2006. EndoG is dispensable in embryogenesis and apoptosis. *Cell Death and Differentiation*, 13(7), pp.1147–1155. Available at: <http://www.nature.com/doi/10.1038/sj.cdd.4401787> [Accessed July 30, 2016].

Deutscher, M.P., 1993. Promiscuous exoribonucleases of *Escherichia coli*. *Journal of bacteriology*, 175(15), pp.4577–83. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8335617> [Accessed December 20, 2016].

Dutta, T. & Deutscher, M.P., 2009. Catalytic Properties of RNase BN/RNase Z from *Escherichia coli*: RNase BN IS BOTH AN EXO- AND ENDORIBONUCLEASE. *Journal of Biological Chemistry*, 284(23), pp.15425–15431. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19366704> [Accessed December 20, 2016].

Dutta, T. & Deutscher, M.P., 2010. Mode of action of RNase BN/RNase Z on tRNA precursors: RNase BN does not remove the CCA sequence from tRNA. *The Journal of biological chemistry*, 285(30), pp.22874–81. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20489203> [Accessed December 20, 2016].

Dutta, T., Malhotra, A. & Deutscher, M.P., 2012. Exoribonuclease and endoribonuclease activities of RNase BN/RNase Z both function in vivo. *The Journal of biological chemistry*, 287(42), pp.35747–55. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22893707> [Accessed December 20, 2016].

Dutta, T., Malhotra, A. & Deutscher, M.P., 2013. How a CCA Sequence Protects Mature tRNAs and tRNA Precursors from Action of the Processing Enzyme RNase BN/RNase Z. *The Journal of biological chemistry*, 288, pp.30636–44. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24022488>.

Ekert, P.G. & Vaux, D.L., 2005. The mitochondrial death squad: Hardened killers or innocent bystanders? *Current Opinion in Cell Biology*, 17(6), pp.626–630.

Emsley, P. & Cowtan, K., 2004. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography*, 60(12), pp.2126–2132. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0907444904019158>.

Ezraty, B., Dahlgren, B. & Deutscher, M.P., 2005. The RNase Z Homologue Encoded by Escherichia coli elaC Gene Is RNase BN. *Journal of Biological Chemistry*, 280(17), pp.16542–16545. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15764599> [Accessed December 19, 2016].

Fan, L. et al., 2011. A survey of green plant tRNA 3'-end processing enzyme tRNase Zs, homologs of the candidate prostate cancer susceptibility protein ELAC2. *BMC evolutionary biology*, 11(1), p.219. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3161902&tool=pmcentrez&rendertype=abstract> [Accessed August 18, 2014].

Garber, R.L. & Gage, L.P., 1979. Transcription of a cloned Bombyx mori tRNA^{2Ala} gene: nucleotide sequence of the tRNA precursor and its processing in vitro. *Cell*, 18(3), pp.817–28. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/260697>.

Graille, M. et al., 2006. Structure-based Functional Annotation: YEAST ymr099c CODES FOR A D-HEXOSE-6-PHOSPHATE MUTAROTASE. *Journal of Biological Chemistry*, 281(40), pp.30175–30185. Available at: <http://www.jbc.org/lookup/doi/10.1074/jbc.M604443200>.

Haack, T.B. et al., 2013. ELAC2 mutations cause a mitochondrial RNA processing defect associated with hypertrophic cardiomyopathy. *American journal of human genetics*, 93(2), pp.211–23. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3738821&tool=pmcentrez&rendertype=abstract> [Accessed March 24, 2014].

Hagenbüchle, O. et al., 1979. The primary transcription product of a silkworm alanine tRNA gene: identification of in vitro sites of initiation, termination and processing. *Cell*, 18(4), pp.1217–29. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/519766>.

Hartmann, R.K. et al., 2009. The making of tRNAs and more - RNase P and tRNase Z. *Progress in molecular biology and translational science*, 85(8), pp.319–68. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19215776> [Accessed September 24, 2013].

Hazbun, T.R. et al., 2003. Assigning Function to Yeast Proteins by Integration of Technologies. , 12, pp.1353–1365.

Hopper, A.K., Pai, D. a & Engelke, D.R., 2010. Cellular dynamics of tRNAs and their genes. *FEBS letters*, 584(2), pp.310–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19931532>.

Howard, M.J. et al., 2013. RNase P enzymes: divergent scaffolds for a conserved biological reaction. *RNA biology*, 10(6), pp.909–14. Available at: <http://www.tandfonline.com/doi/full/10.4161/rna.24513>.

Huang, K.-J., Ku, C.-C. & Lehman, I.R., 2006. Endonuclease G: a role for the enzyme in recombination and cellular proliferation. *Proceedings of the National Academy of Sciences of the United States of America*, 103(24), pp.8995–9000. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1482554&tool=pmcentrez&rendertype=abstract>.

Irvine, R.A. et al., 2005. Generation and characterization of endonuclease G null mice. *Molecular and cellular biology*, 25(1), pp.294–302. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15601850> [Accessed September 25, 2016].

Ishii, R. et al., 2005. Crystal structure of the tRNA 3' processing endoribonuclease tRNase Z from *Thermotoga maritima*. *The Journal of biological chemistry*, 280(14), pp.14138–44. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15701599> [Accessed October 14, 2013].

Ishii, R. et al., 2007. The structure of the flexible arm of *Thermotoga maritima* tRNase Z differs from those of homologous enzymes. *Acta crystallographica. Section F, Structural biology and crystallization*

communications, 63(Pt 8), pp.637–41. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2335171&tool=pmcentrez&rendertype=abstract> [Accessed March 27, 2014].

Kabsch, W., 2010. XDS. *Acta Crystallographica Section D Biological Crystallography*, 66(2), pp.125–132. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0907444909047337> [Accessed October 24, 2016].

Karkashon, S., Hopkinson, A. & Levinger, L., 2007. tRNase Z catalysis and conserved residues on the carboxy side of the His cluster. *Biochemistry*, 46(33), pp.9380–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17655328>.

Kelley, L.A. et al., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6), pp.845–858. Available at: <http://www.nature.com/doi/10.1038/nprot.2015.053>.

Kieper, J. et al., 2010. Production and characterization of recombinant protein preparations of Endonuclease G-homologs from yeast, *C. elegans* and humans. *Protein Expression and Purification*, 73(1), pp.99–106.

Klemm, B.P. et al., 2016. The Diversity of Ribonuclease P: Protein and RNA Catalysts with Analogous Biological Functions. *Biomolecules*, 6(2), p.27. Available at: <http://www.mdpi.com/2218-273X/6/2/27>.

Korver, W. et al., 2003. The product of the candidate prostate cancer susceptibility gene ELAC2 interacts with the gamma-tubulin complex. *International journal of cancer*, 104(3), pp.283–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12569551>.

Kostelecny, B., Pohl, E. & Vogel, A., 2006. The crystal structure of the zinc phosphodiesterase from *Escherichia coli* provides insight into function and cooperativity of tRNase Z-family proteins. *Journal of ...*, 188(4), pp.1607–1614. Available at: <http://jb.asm.org/content/188/4/1607.short> [Accessed August 27, 2014].

Kroemer, G. & Martin, S.J., 2005. Caspase-independent cell death. *Nature medicine*, 11(7), pp.725–30. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16015365>.

Kufel, J. & Tollervey, D., 2003. 3'-processing of yeast tRNA^{Trp} precedes 5'-processing. *RNA (New York, N.Y.)*, 9(2), pp.202–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12554863>.

Lee, M.-W. et al., 2007. Global protein expression profiling of budding yeast in response to DNA damage. *Yeast (Chichester, England)*, 24(3), pp.145–54. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17351896>.

Li, L.Y., Luo, X. & Wang, X., 2001. Endonuclease G is an apoptotic DNase when released from mitochondria. *Nature*, 412(6842), pp.95–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11452314> [Accessed September 26, 2016].

Li de la Sierra-Gallay, I. et al., 2006. Structure of the ubiquitous 3' processing enzyme RNase Z bound to transfer RNA. *Nature Structural & Molecular Biology*, 13(4), pp.376–377. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16518398> [Accessed October 14, 2013].

Li de la Sierra-Gallay, I., Pellegrini, O. & Condon, C., 2005. Structural basis for substrate binding, cleavage and allostery in the tRNA maturase RNase Z. *Nature*, 433(7026), pp.657–61. Available at: <http://www.nature.com/nature/journal/v433/n7026/abs/nature03284.html> [Accessed March 19, 2014].

Lin, J.L.J. et al., 2012. Structural insights into apoptotic DNA degradation by CED-3 protease suppressor-6 (CPS-6) from *Caenorhabditis elegans*. *Journal of Biological Chemistry*, 287(10), pp.7110–7120.

Loll, B. et al., 2009. Crystal structure of the EndoG/EndoGI complex: mechanism of EndoG inhibition. *Nucleic acids research*, 37(21), pp.7312–20. Available at:

<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2790893&tool=pmcentrez&rendertype=abstract> [Accessed June 5, 2014].

van Loo, G. et al., 2001. Endonuclease G: a mitochondrial protein released in apoptosis and involved in caspase-independent DNA degradation. *Cell death and differentiation*, 8(12), pp.1136–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11753562> [Accessed September 26, 2016].

Low, R.L., 2003. Mitochondrial Endonuclease G function in apoptosis and mtDNA metabolism: A historical perspective. *Mitochondrion*, 2(4), pp.225–236. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16120323> [Accessed July 30, 2016].

Maraia, R.J. & Lamichhane, T.N., 2011. 3' processing of eukaryotic precursor tRNAs. *Wiley Interdisciplinary Reviews: RNA*, 2(3), pp.362–375. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21572561>.

McCoy, A.J. et al., 2007. Phaser crystallographic software. *Journal of Applied Crystallography*, 40(4), pp.658–674. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0021889807021206> [Accessed August 26, 2016].

Minagawa, A. et al., 2004. A novel endonucleolytic mechanism to generate the CCA 3' termini of tRNA molecules in *Thermotoga maritima*. *The Journal of biological chemistry*, 279(15), pp.15688–97. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14749326> [Accessed August 27, 2014].

Minagawa, A. et al., 2006. Identification by Mn²⁺ rescue of two residues essential for the proton transfer of tRNase Z catalysis. *Nucleic Acids Research*, 34(13), pp.3811–3818.

Minagawa, A. et al., 2008. The flexible arm of tRNase Z is not essential for pre-tRNA binding but affects cleavage site selection. *Journal of molecular biology*, 381(2), pp.289–99. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18602113> [Accessed January 3, 2014].

Minagawa, A. et al., 2005. The missense mutations in the candidate prostate cancer gene ELAC2 do not alter enzymatic properties of its product. *Cancer letters*, 222(2), pp.211–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15863270>.

Mohan, A. et al., 1999. The 3' end CCA of mature tRNA is an antideterminant for eukaryotic 3'-tRNase. *Rna*, 5(2), pp.245–256. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10024176.

Nashimoto, M., Tamura, M. & Kaspar, R.L., 1999. Minimum requirements for substrates of mammalian tRNA 3' processing endoribonuclease. *Biochemistry*, 38(37), pp.12089–12096.

Ninomiya, S. et al., 2015. Potential small guide RNAs for tRNase ZL from human plasma, peripheral blood mononuclear cells, and cultured cell lines. *PloS one*, 10(3), p.e0118631. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4346264&tool=pmcentrez&rendertype=abstract>.

Ohsato, T. et al., 2002. Mammalian mitochondrial endonuclease G: Digestion of R-loops and localization in intermembrane space. *European Journal of Biochemistry*, 269(23), pp.5765–5770. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12444964>.

Palermo, V. et al., 2013. p53 death signal is mainly mediated by Nucl(EndoG) in the yeast *Saccharomyces cerevisiae*. *FEMS yeast research*, 13(7), pp.682–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23875998>.

Parrish, J. et al., 2001. Mitochondrial endonuclease G is important for apoptosis in *C. elegans*. *Nature*, 412(6842), pp.90–94. Available at: <http://www.nature.com/doifinder/10.1038/35083608> [Accessed July 30, 2016].

Pellegrini, O. et al., 2012. Activation of tRNA Maturation by Downstream Uracil Residues in B.

subtilis. *Structure*, 20(10), pp.1769–1777. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22940585> [Accessed March 31, 2014].

Pellegrini, O. et al., 2003. Endonucleolytic processing of CCA-less tRNA precursors by RNase Z in *Bacillus subtilis*. *The EMBO journal*, 22(17), pp.4534–43. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=202377&tool=pmcentrez&rendertype=abstract>.

Perwez, T. & Kushner, S.R., 2006. RNase Z in *Escherichia coli* plays a significant role in mRNA decay. *Molecular Microbiology*, 60(3), pp.723–737. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16629673>.

Petoukhov, M. V & Svergun, D.I., 2005. Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophysical journal*, 89(2), pp.1237–50. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15923225> [Accessed August 29, 2016].

Phizicky, E.M. & Hopper, A.K., 2015. tRNA processing, modification, and subcellular dynamics: past, present, and future. *RNA*, 21(4), pp.483–485. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25780105>.

Pierce, B.G. et al., 2014. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics (Oxford, England)*, 30(12), pp.1771–3. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24532726> [Accessed August 29, 2016].

Putnam, C.D. et al., 2007. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly reviews of biophysics*, 40(3), pp.191–285.

Redko, Y., Li de la Sierra-Gallay, I. & Condon, C., 2007. When all's zed and done: the structure and function of RNase Z in prokaryotes. *Nature reviews. Microbiology*, 5(4), pp.278–86. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17363966> [Accessed March 27, 2014].

Robert, X. & Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research*, 42(Web Server issue), pp.W320–4. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24753421> [Accessed September 24, 2016].

Rossmannith, W., 2011. Localization of Human RNase Z Isoforms: Dual Nuclear/Mitochondrial Targeting of the ELAC2 Gene Product by Alternative Translation Initiation B. Lightowers, ed. *PLoS ONE*, 6(4), p.e19152. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3084753&tool=pmcentrez&rendertype=abstract> [Accessed January 3, 2014].

Rossmannith, W., 2012. Of P and Z: Mitochondrial tRNA processing enzymes. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1819(9–10), pp.1017–1026. Available at: <http://dx.doi.org/10.1016/j.bbagr.2011.11.003>.

Salinas-Giegé, T., Giegé, R. & Giegé, P., 2015. tRNA Biology in Mitochondria. *International Journal of Molecular Sciences*, 16(3), pp.4518–4559. Available at: <http://www.mdpi.com/1422-0067/16/3/4518/>.

Schäfer, P. et al., 2004. Structural and functional characterization of mitochondrial EndoG, a sugar non-specific nuclease which plays an important role during apoptosis. *Journal of Molecular Biology*, 338(2), pp.217–228.

Schiffer, S., 2002. Assigning a function to a conserved group of proteins: the tRNA 3'-processing enzymes. *The EMBO Journal*, 21(11), pp.2769–2777. Available at: <http://emboj.embopress.org/cgi/doi/10.1093/emboj/21.11.2769>.

Schiffer, S. et al., 2001. The plant tRNA 3' processing enzyme has a broad substrate spectrum. *Biochemistry*, 40(28), pp.8264–8272.

Schilling, O. et al., 2004. Characterization of an Escherichia coli elaC deletion mutant. *Biochemical and biophysical research communications*, 320(4), pp.1365–73. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15303284> [Accessed December 20, 2016].

Schilling, O. et al., 2005. Exosite modules guide substrate recognition in the ZiPD/ElaC protein family. *Journal of Biological Chemistry*, 280(18), pp.17857–17862.

Schneidman-Duhovny, D., Hammel, M. & Sali, A., 2010. FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic acids research*, 38(Web Server issue), pp.W540–4. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20507903> [Accessed August 31, 2016].

Scholz, S.R. et al., 2003. Experimental evidence for a beta beta alpha-Me-finger nuclease motif to represent the active site of the caspase-activated DNase. *Biochemistry*, 42(31), pp.9288–94. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12899615> [Accessed August 1, 2016].

Seidman, J.G. et al., 1975. A mutant of escherichia coli defective in removing 3' terminal nucleotides from some transfer RNA precursor molecules. *Cell*, 5(4), pp.389–400. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1098779> [Accessed December 20, 2016].

Sheldrick, G.M., 2008. A short history of SHELX. *Acta Crystallographica Section A Foundations of Crystallography*, 64(1), pp.112–122. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0108767307043930> [Accessed December 19, 2016].

Sheldrick, G.M., 2010. Experimental phasing with SHELXC/D/E: Combining chain tracing with density modification. *Acta Crystallographica Section D: Biological Crystallography*, 66(4), pp.479–485.

Shibata, H.S. et al., 2005. The T loop structure is dispensable for substrate recognition by tRNase ZL. *The Journal of biological chemistry*, 280(23), pp.22326–34. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15824113>.

Skowronek, E. et al., 2014. tRNA 3' processing in yeast involves tRNase Z, Rex1, and Rrp6. *RNA*, 20(1), pp.115–130. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24249226>.

Smith, M.M. & Levitan, D.J., 2004. The Caenorhabditis elegans homolog of the putative prostate cancer susceptibility gene ELAC2, hoe-1, plays a role in germline proliferation. *Developmental biology*, 266(1), pp.151–60. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14729485>.

Späth, B. et al., 2005. Analysis of the functional modules of the tRNA 3' endonuclease (tRNase Z). *The Journal of biological chemistry*, 280(42), pp.35440–7. Available at: <http://www.jbc.org/lookup/doi/10.1074/jbc.M506418200> [Accessed September 11, 2016].

Späth, B., Canino, G. & Marchfelder, A., 2007. tRNase Z: the end is not in sight. *Cellular and molecular life sciences: CMLS*, 64(18), pp.2404–12. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17599240> [Accessed September 24, 2013].

Svergun, D. et al., 1995. CRY SOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography*, 28(6), pp.768–773. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0021889895007047> [Accessed August 29, 2016].

Svergun, D.I., 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical journal*, 76(6), pp.2879–86. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10354416> [Accessed August 29, 2016].

Takaku, H. et al., 2003. A candidate prostate cancer susceptibility gene encodes tRNA 3' processing endoribonuclease. *Nucleic acids research*, 31(9), pp.2272–8. Available at: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkg337> [Accessed October 14, 2013].

Takaku, H., 2004. The N-terminal half-domain of the long form of tRNase Z is required for the RNase 65 activity. *Nucleic Acids Research*, 32(15), pp.4429–4438. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=516050&tool=pmcentrez&rendertype=abstract>

[Accessed October 14, 2013].

Tavtigian, S. V et al., 2001. A candidate prostate cancer susceptibility gene at chromosome 17p. *Nature genetics*, 27(2), pp.172–80. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11175785>.

Temme, C. et al., 2009. The *Drosophila melanogaster* gene cg4930 encodes a high affinity inhibitor for endonuclease G. *Journal of Biological Chemistry*, 284(13), pp.8337–8348. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19129189>.

Vogel, A. et al., 2002. ElaC encodes a novel binuclear zinc phosphodiesterase. *Journal of Biological Chemistry*, 277(32), pp.29078–29085.

Vogel, A. et al., 2005. The tRNase Z family of proteins: physiological functions, substrate specificity and structural properties. *Biological chemistry*, 386(12), pp.1253–64. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16336119> [Accessed October 14, 2013].

Vogel, A., Schilling, O. & Meyer-Klaucke, W., 2004. Identification of metal binding residues for the binuclear zinc phosphodiesterase reveals identical coordination as glyoxalase II. *Biochemistry*, 43(32), pp.10379–86. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15301536> [Accessed September 11, 2016].

Walker, D.C. et al., 2002. Mutagenic scan of the H-N-H motif of colicin E9: implications for the mechanistic enzymology of colicins, homing enzymes and apoptotic endonucleases. *Nucleic acids research*, 30(14), pp.3225–34. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12136104> [Accessed August 1, 2016].

Wang, Z. et al., 2012. Identification and sequence analysis of metazoan tRNA 3'-end processing enzymes tRNase Zs. N. Salamin, ed. *PloS one*, 7(9), p.e44264. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22962606>.

Weinkam, P., Pons, J. & Sali, A., 2012. Structure-based model of allostery predicts coupling between distant sites. *Proceedings of the National Academy of Sciences of the United States of America*, 109(13), pp.4875–80. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22403063> [Accessed August 31, 2016].

Wu, J. et al., 2016. Histone H2B gene cloning, with implication for its function during nuclear shaping in the Chinese mitten crab, *Eriocheir sinensis*. *Gene*, 575(2 Pt 1), pp.276–84. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26343795>.

Wu, S.-L. et al., 2009. Mutagenesis identifies the critical amino acid residues of human endonuclease G involved in catalysis, magnesium coordination, and substrate specificity. *Journal of biomedical science*, 16, p.6.

Xie, X. & Dubrovsky, E.B., 2015. Knockout of *Drosophila* RNase ZL impairs mitochondrial transcript processing, respiration and cell cycle progression. *Nucleic Acids Research*, pp.1–12. Available at: <http://nar.oxfordjournals.org/content/early/2015/11/06/nar.gkv1149.long>.

Zamzami, N., Larochette, N. & Kroemer, G., 2005. Mitochondrial permeability transition in apoptosis and necrosis. *Cell Death and Differentiation*, 12, pp.1478–1480. Available at: <http://www.nature.com/doi/10.1038/sj.cdd.4401682> [Accessed September 5, 2016].

Zareen, N. et al., 2005. Residues in the conserved His domain of fruit fly tRNase Z that function in catalysis are not involved in substrate recognition or binding. *Journal of molecular biology*, 350(2), pp.189–99. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15935379> [Accessed January 3, 2014].

Zareen, N., 2006. Residues in two homology blocks on the amino side of the tRNase Z His domain contribute unexpectedly to pre-tRNA 3' end processing. *RNA*, 12(6), pp.1104–1115. Available at: <http://www.rnajournal.org/cgi/doi/10.1261/rna.4206>.

Zassenhaus, H.P. & Denniger, G., 1994. Analysis of the role of the NUC1 endo/exonuclease in yeast mitochondrial DNA recombination. *Current genetics*, 25(2), pp.142–9. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/8087883>.

Zhang, X., Bian, X. & Kong, J., 2014. The Proapoptotic Protein BNIP3 Interacts with VDAC to Induce Mitochondrial Release of Endonuclease G D. Tang, ed. *PLoS ONE*, 9(12), p.e113642. Available at: <http://dx.plos.org/10.1371/journal.pone.0113642>.

Zhao, T. et al., 2007. Granzyme K directly processes bid to release cytochrome c and endonuclease G leading to mitochondria-dependent cell death. *The Journal of biological chemistry*, 282(16), pp.12104–11. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17308307> [Accessed October 4, 2016].

Zhao, Z. et al., 2009. Functional conservation of tRNase ZL among *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and humans. *The Biochemical journal*, 422(3), pp.483–92. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19555350> [Accessed October 14, 2013].

'Protein interfaces, surfaces and assemblies' service PISA at the European Bioinformatics Institute. (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html). E. Krissinel and K. Henrick (2007). 'Inference of macromolecular assemblies from crystalline state.'. *J. Mol. Biol.* 372, 774--797.