



Provided by the author(s) and University of Galway in accordance with publisher policies. Please cite the published version when available.

Title	Development of drugs targeting the human tRNA ligase RtcB in breast cancer
Author(s)	Nandy, Argha
Publication Date	2019-02-15
Publisher	NUI Galway
Item record	<a href="http://hdl.handle.net/10379/15324">http://hdl.handle.net/10379/15324</a>

Downloaded 2024-03-13T10:01:03Z

Some rights reserved. For more information, please see the item record link above.



# DEVELOPMENT OF DRUGS TARGETING THE HUMAN tRNA LIGASE RtcB IN BREAST CANCER

A thesis submitted in partial fulfilment of the requirements for the  
degree of Doctor of Philosophy



NATIONAL UNIVERSITY OF IRELAND, GALWAY

By

**ARGHA NANDY MSc. B.Tech.**

SUPERVISION BY  
PROF. AFSHIN SAMALI  
&  
DR. ADRIENNE GORMAN

APOPTOSIS RESEARCH CENTRE  
DISCIPLINE OF BIOCHEMISTRY

NATIONAL UNIVERSITY OF IRELAND, GALWAY, IRELAND

**FEBRUARY 2019**

---

## Dedication

To my parents, Samir Kumar Nandy and Chitra Nandy whose guidance, wisdom, kindness, blessing, persistent love and infinite moral support have brought me this far. Also, to my late grandparents Chitta Ranjan Nandy and Shanti Nandy, I owe it all to you. Many Thanks!

---

## Epigraph

*“Inspiration follows aspiration.”*

--**Rabindranath Tagore**  
(1861-1941)

---

## Declaration

The work presented in this thesis is to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that apart from the due acknowledgements, it is entirely my own work and have not submitted this material, either in whole or in part, for a degree at this or any other university.

Signed: *Argha Nandy*

Date: 15.02.2019

**Argha Nandy**

---

## Acknowledgements

Sailing through troubled waters and finally stepping on to the shore. A feeling that can't be described unless experienced. Time to thank everyone who made this journey worthwhile.

First and foremost, I would like to thank my principal supervisors Prof. Afshin Samali and Dr. Adrienne Gorman for the opportunity to be a part of their research laboratory and their guidance and support over the last 4.5 years. I would also like to extend my sincerest appreciation to Prof. Leif Eriksson and Dr. Patricia Saenz-Méndez for their unending patience, guidance and very positive approach to work during my 3-month secondment in Sweden. Coming from an entirely wet-lab background, I had no other choice but to annoy you every day by asking so many questions on the amazing computational work we published in *Proteins*. My gratitude extends to my everyday wet-lab mentors Drs. Susan Logue and Shane Deegan for all the invaluable suggestions, advice and great tips. I would like to thank Dr. Muriel Voisin for her valuable inputs and guidance during the microscopy experiments. Microscopy was a whole new world for me. Luckily the immunostaining clicked right from the first experiment, and since then we were on a roll. I would also like to thank my graduate review committee members, Prof. Kevin Sullivan, Prof. Uri Frank, and Prof. Heinz-Peter Nasheuer for serving as my committee members even at hardship and helping me reach my goal of finishing Ph.D. successfully. A very special gratitude goes out to the Hardiman Research Scholarships for helping and providing the funding for the work.

Special thanks to the lab members Izabella, Eoghan, and Aaron with whom I spent my initial years in the lab. Iza, you helped me book an apartment for me even before I landed in Ireland. See, I kept my promise and you are at the top of the list. Eoghan, I will never forget the times when we pipetted tiny droplets into the qPCR assay plates for hours. It was a hell of a job. Chetan, Aitor, and Brian, you guys made the lab livelier. The most difficult days of my final years were spent with you guys. Thanks for those evenings in the college bar that helped me to go through complaining about life and discussing the science. Mari, Emma, and Sangeetha, thanks for your support as well. Thanks to Alexis from Chile, all other lab members, and the ARC over the years that helped me grow both scientifically and personally.

Finally, my eternal gratitude goes to my dearest family, my beloved cousin Rupali and to my friend cum sisters Priyanka, Pamela from back home whose eternal support over the years has made a world of difference and to whom I owe a debt of gratitude that I can never fully repay.

---

# Table of contents

<b>ABSTRACT .....</b>	<b>IX</b>
<b>LIST OF FIGURES .....</b>	<b>XI</b>
<b>LIST OF TABLES .....</b>	<b>XIII</b>
<b>LIST OF ABBREVIATIONS .....</b>	<b>XIV</b>
<b>CHAPTER 1: INTRODUCTION.....</b>	<b>1</b>
1.1 RtcB .....	2
1.1.1 RtcB complex.....	3
1.1.2 RNA ligation mechanism.....	6
1.1.3 RtcB subcellular localization .....	6
1.1.4 RtcB facilitates tRNA ligation .....	7
1.1.5 RtcB facilitates XBP1 mRNA ligation .....	7
1.1.6 RtcB facilitates other biological processes .....	8
1.2 IRE1 .....	9
1.2.1 IRE1 $\alpha$ – XBP1 .....	13
1.2.2 IRE1 $\alpha$ – RIDD .....	15
1.3 IRE1 $\alpha$ – XBP1/RIDD in cancer .....	18
1.4 Breast cancer.....	21
1.4.1 Histological classification.....	21
1.4.2 Molecular classification.....	21
1.4.3 Triple negative breast cancer.....	21
1.5 Research hypothesis and objectives .....	24
1.5.1 Hypothesis .....	24
1.5.2 Objectives.....	24
<b>CHAPTER 2: MATERIALS AND METHODS .....</b>	<b>26</b>
2.1 Cell culture of adherent cell lines .....	26

---

2.2	Drug treatment .....	26
2.2.1	<i>Tunicamycin</i> .....	26
2.2.2	<i>Thapsigargin</i> .....	26
2.2.3	<i>MKC8866</i> .....	27
2.2.4	<i>Paclitaxel</i> .....	27
2.3	Immunofluorescence and microscopy.....	27
2.4	siRNA preparation.....	28
2.4.1	<i>siRNA transfections</i> .....	28
2.5	RNA extractions .....	29
2.6	Reverse transcription polymerase chain reaction (cDNA synthesis).....	30
2.7	Conventional polymerase chain reaction (PCR).....	31
2.8	Real time PCR .....	32
2.9	Protein expression analysis.....	34
2.9.1	<i>Protein sample preparation</i> .....	34
2.9.2	<i>SDS-PAGE</i> .....	34
2.9.3	<i>Western blotting</i> .....	35
2.10	Computational methods.....	36
2.10.1	<i>Homology structure modelling</i> .....	37
2.10.2	<i>Quality assessment of predicted structures</i> .....	38
2.10.3	<i>Molecular dynamics simulations</i> .....	39
2.10.4	<i>Umbrella sampling simulations</i> .....	39
2.10.5	<i>Virtual high throughput screening</i> .....	40
2.11	Antibiotic kill curve.....	41
2.12	Plasmid DNA preparation .....	42
2.13	Plasmid transfections.....	44
2.14	Screening hit compounds by cellular assay.....	45
2.15	Luciferase assay .....	46



---

2.16	BCA assay.....	47
2.17	Cell viability assay – Hoechst/PI staining .....	48
2.18	Statistical Analysis.....	49
<b>CHAPTER 3: HOMOLOGY MODEL OF THE HUMAN RTCB .....</b>		<b>50</b>
3.1	Introduction and aims .....	50
3.2	Results.....	52
3.2.1	<i>Human RtcB homology model</i> .....	52
3.2.2	<i>In silico model quality assessment</i> .....	55
3.2.2.1	QMEAN quality assessment.....	55
3.2.2.2	Estimating stereochemical quality .....	56
3.2.2.3	Distinguishing between correct and incorrect folding.....	56
3.2.2.4	Structure alignment of the human and bacterial homologs .....	58
3.2.2.5	Binding pocket and druggability .....	60
3.2.2.6	Metal coordination microenvironment .....	63
3.2.3	<i>Mn binding in wild type and C122A mutant hRtcB</i> .....	66
3.3	Discussion.....	68
	Supplementary Material .....	73
<b>CHAPTER 4: LIGAND DISCOVERY FROM A HOMOLOGY MODEL OF THE HUMAN RTCB.....</b>		<b>82</b>
4.1	Introduction and aims .....	82
4.2	Results.....	84
4.2.1	<i>Hit discovery through the VHTS campaign</i> .....	84
4.2.2	<i>Structural interaction fingerprint</i> .....	85
4.2.3	<i>Protein binding site similarity</i> .....	87
4.2.4	<i>Analysis of the top eight hit compounds</i> .....	91
4.2.5	<i>Establishment of an HTS assay for screening inhibitors against XBP1 splicing</i> .....	93
4.2.6	<i>Depletion of hRtcB significantly reduces XBP1s in MDA-MB-231 /XBP1-luciferase and MDA-MB-231 cells</i> .....	95

---

4.2.7	<i>Hit compound-001 reduces ER stress-induced XBP1 splicing in MDA-MB-231/XBP1-luciferase cells.....</i>	97
4.2.8	<i>Compounds-001 and -007 are not toxic at 7 hours of treatment under ER stress .....</i>	99
4.2.9	<i>Compound-001 does not inhibit hRtcB in the IRE1-XBP1 axis in MDA-MB-231/XBP1-luciferase cells.....</i>	100
4.2.10	<i>Compounds-001 and -007 do not inhibit luciferase.....</i>	101
4.2.11	<i>Compound-001 does not reduce XBP1s in MDA-MB-231 cells ....</i>	102
4.2.12	<i>Compound-007 does not reduce XBP1s in MDA-MB-231 cells ....</i>	103
4.2.13	<i>Depletion of hRtcB reduces paclitaxel-induced XBP1 splicing in MDA-MB-231/XBP1-luciferase cells.....</i>	104
4.2.14	<i>Inducibility of hRtcB in MDA-MB-231 cells .....</i>	106
4.2.15	<i>Localization of hRtcB in MDA-MB-231 cells.....</i>	107
4.3	Discussion.....	108
<b>CHAPTER 5: GENERAL DISCUSSION AND FUTURE PERSPECTIVES .</b>		<b>113</b>
5.1	Understanding the druggability of hRtcB protein.....	113
5.2	Understanding the role of hRtcB in generation of XBP1s.....	117
5.3	Understanding the need of targeting hRtcB in breast cancer.....	118
5.4	Concluding remarks.....	121
<b>REFERENCES.....</b>		<b>122</b>

---

## Abstract

The RNA ligase RtcB has been shown to regulate RNA function by direct exon ligation of spliced *XBPI* messenger RNA cleaved by the endoplasmic reticulum (ER) stress sensor IRE1 $\alpha$ . Ligation by RtcB leads to the expression of the transcription factor XBP1s that controls ER homeostasis and is surprisingly also utilized for oncogenic transformation, survival, and progress. Through elevated levels of XBP1s, tumor cells acquire a pro-survival advantage and promote tumor aggressiveness by an increase in tumor immune infiltration, angiogenesis, and cell invasive properties. Using RNA interference (RNAi), depleting human RtcB (*hRtcB*) in MDA-MB-231 triple-negative breast cancer (TNBC) cells significantly reduced basal XBP1s expression both at the mRNA and protein level. As the crystal structure of *hRtcB* is yet to be resolved, *in silico* techniques were therefore employed to develop a 3-dimensional predicted model of the *hRtcB* protein for the first time. Built from the *Pyrococcus horikoshii* RtcB (bacterial RtcB) crystal structure, PDB ID 4DWQA, the *hRtcB* homology model is in complex with manganese and covalently bound ligand GMP. The developed model clearly showed that the GTP binding site of the enzyme is a well-defined pocket that can be utilized as a target site for *in silico* drug discovery. A virtual high throughput screening (VHTS) campaign was thereafter executed to develop drugs that would halt or modulate the *hRtcB*-mediated XBP1s formation in TNBC. ~Eighteen million small molecule compounds were screened against the ligand binding site/active site of the *hRtcB* homology model. From the 3129 hits generated, the top 200 highest scoring binders were subjected to the structural hit fingerprint analysis. The top 100 highest scoring binders that exhibited similar interactions to the natural ligand GMP in the active site of the *hRtcB* homology model were further narrowed down to the top 20 and finally to the top 8 hit candidates based on criteria including high, medium, and poor scoring binders and purchasability. Through establishing an *in vitro* screening platform (MDA-MB-231/XBP1-Luciferase cell line), the top 8 hit candidates were tested in the luciferase assay from which compound-001 yielded a significant reduction (~50%) in XBP1s-luciferase under ER stress. Additionally, compound-007 showed a trend in the reduction of XBP1s-luciferase under stress in the assay. Testing these 2 compounds further in MDA-MB-231 TNBC cells didn't yield any reduction in XBP1s under ER stress. Based

---

on the luciferase assay results, these two hit compounds can become potential leads once tested in a biochemical assay setting and if successful, further optimized towards structure-activity relationships.

---

## List of figures

Figure 1.1 RtcB-mediated <i>XBPI</i> mRNA splicing, RIDD and their involvement in tumorigenesis .....	2
Figure 1.2 RtcB ligase facilitates re-ligation of tRNA and <i>XBPI</i> mRNA ends.....	5
Figure 1.3 Domain structure of IRE1 .....	10
Figure 1.4 ER stress-mediated <i>XBPI</i> mRNA splicing.....	14
Figure 2.1 Antibiotic kill curve .....	42
Figure 2.2 The luciferase reaction.....	46
Figure 3.1 Sequence alignment and predicted structure .....	53
Figure 3.2 <i>In silico</i> model quality assessment .....	57
Figure 3.3 Structure alignment assessment.....	59
Figure 3.4 Ligand binding site .....	62
Figure 3.5 Metal-binding microenvironment.....	65
Figure 3.6 Umbrella sampling histograms .....	67
Figure 3.7 Potential of mean force (PMF) curves.....	68
Figure 3.8 Ligand binding/active site of <i>bRtcB</i> .....	71
Figure S1 Residue-by-residue mapping of 3D structural alignment.....	80
Figure 4.1 The VHTS workflow .....	84
Figure 4.2 Structural interaction fingerprint of top 200 small molecule compounds	86
Figure 4.3 Binding site similarity .....	89
Figure 4.4 Predicted binding patterns of GMP and the eight hit compounds in the <i>hRtcB</i> active site .....	92
Figure 4.5 HTS platform for screening inhibitors against <i>XBPI</i> splicing .....	94
Figure 4.6 Depletion of <i>hRtcB</i> in MDA-MB-231/ <i>XBPI</i> -luciferase and MDA-MB-231 cells .....	96
Figure 4.7 Hit compound-001 reduces ER stress-induced <i>XBPI</i> s-luciferase activity .....	98
Figure 4.8 Toxicity of hit compounds-001 and -007 .....	100
Figure 4.9 Compound-001 does not inhibit <i>hRtcB</i> .....	101
Figure 4.10 Compounds-001 and -007 do not inhibit luciferase .....	102

---

Figure 4.11 Compound-001 does not reduce XBP1s in TNBC cell model .....	103
Figure 4.12 Compound-007 does not reduce XBP1s in TNBC cell model .....	104
Figure 4.13 Depletion of <i>hRtcB</i> reduces paclitaxel-induced XBP1s-luciferase activity .....	105
Figure 4.14 Inducibility of <i>hRtcB</i> in non-tumorigenic and tumorigenic cell models .....	106
Figure 4.15 Subcellular localization of <i>hRtcB</i> in non-tumorigenic and tumorigenic cell models .....	108

---

## List of tables

Table 2.1 PCR primers and cycle conditions .....	32
Table 2.2 Real time PCR primers and probes .....	33
Table 2.3 SDS PAGE lysis buffer (2X) .....	34
Table 2.4 SDS-PAGE .....	35
Table 2.5 Antibodies for western blotting.....	36
Table 2.6 Plasmid.....	43
Table 2.7 Cell culture lysis reagent buffer composition .....	47
Table 3.1 Model Rankings .....	54
Table 3.2 Scores for structure reliability, absolute quality, Ramachandran plot, and VERIFY-3D .....	58
Table 3.3 Scores for 3D structure alignment .....	60
Table 3.4 Scores for ligand binding site and their 3D structure alignments .....	63
Table 3.5 Scores for metal-binding site microenvironment.....	66
Table 4.1 Top Hit compounds from the VHTS .....	87
Table 4.2 Proteins with similar sites from PDB.....	89

---

## List of abbreviations

3D	3-dimension
3'-P	3'-phosphate
5'-OH	5'-hydroxyl
Arg	Arginine
ATF6	Activating transcription factor 6
BCA	Bicinchoninic acid assay
BCSC	Breast cancer stem cell
BiP	Binding immunoglobulin protein
BLAST	Basic local alignment search tool
<i>bRtcB</i>	Bacterial RtcB
BSA	Bovine serum albumin
bZIP	Basic leucine zipper
CBCS	Chemical Biology Consortium Sweden
cDNA	Complementary DNA
COM	Center of mass
CSD	Cambridge Structural Database
CXCL1	Chemokine (C-X-C) ligand 1
DAPI	4',6-diamidino-2-phenylindole
DEPC	Diethyl pyrocarbonate
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DoG	Difference of Gaussian
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EMT	Epithelial-to-mesenchymal transition
ER	Endoplasmic reticulum
ER <sup>+</sup>	Estrogen receptor positive
ERAD	Endoplasmic reticulum associated degradation
ERLIN2	ER lipid raft-associated 2



---

ERSE	ER stress response element
FBS	Fetal bovine serum
FITC	Fluorescein isothiocyanate
GBM	Glioblastoma multiforme
GFP	Green fluorescent protein
GM-CSF	Granulocyte-macrophage-colony-stimulating-factor
GRP	Glucose-regulated protein
GSTO1	Glutathione S-transferase omega 1
HER2	Human epidermal growth factor receptor 2
HIF	Hypoxia-inducible factor
<i>hRtcB</i>	Human RtcB
HTS	High-throughput screening
IL	Interleukin
Ile	Isoleucine
IRE1	Inositol-requiring enzyme 1
I-TASSER	Iterative Threading ASSEmbly Refinement
LB	Luria broth
MD	Molecular dynamics
MHC	Major histocompatibility complex
miRNA	Microribonucleic acid
MOE	Molecular Operating Environment
mRNA	Messenger ribonucleic acid
ORF	Open reading frame
PARP	Poly ADP-ribose polymerase
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PERK	Protein kinase RNA-like ER kinase
Phyre2	Protein Homology/analogY Recognition Engine V 2.0
PI	Propidium iodide
PME	Particle-mesh Ewald
PMF	Potential of mean force
PPi	Pyrophosphate
pre-tRNA	Precursor transfer ribonucleic acid

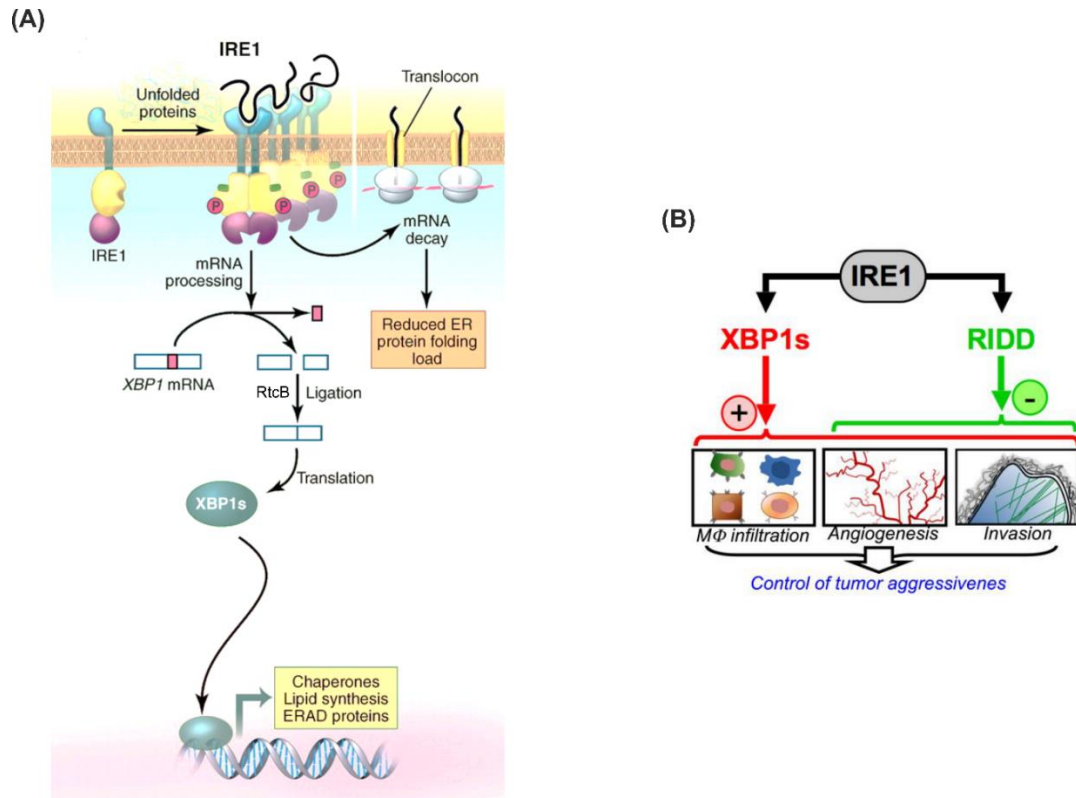
---

QMEAN	Qualitative Model Energy ANalysis
RIDD	Regulated IRE1-dependent decay
RMSD	Root-mean-square deviations
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Endoribonuclease
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
Ser	Serine
SERCA	Sarco-endoplasmic reticulum calcium-ATPase
SIFt	Structural interaction fingerprint
SL	Stem loop
tDC	Tumor-associated dendritic cell
TE	Trypsin EDTA
TGF	Transforming growth factor
Tg	Thapsigargin
Thr	Threonine
Tm	Tunicamycin
TNBC	Triple-negative breast cancer
tRNA	Transfer RNA
TXNIP	Thioredoxin-interacting protein
Tyr	Tyrosine
UPR	Unfolded protein response
UPRE	Unfolded protein response element
VECSUM	Vector sum of bond valence
VHTS	Virtual high-throughput screening
V	Volt
WHAM	Weighted histogram analysis method
XBP1s	X-box-binding protein-1 spliced
XBP1u	X-box-binding protein-1 unspliced
YASARA	Yet Another Scientific Artificial Reality Application

---

# CHAPTER 1: Introduction

RNA ligation by the tRNA ligase RtcB regulates the expression of spliced X-box-binding protein-1 (XBP1s), a transcription factor that drives expression of genes important for protein homeostasis in the endoplasmic reticulum (ER) (Figure 1.1A). Cleavage of *XBPI* messenger RNA (mRNA) is mediated by the endoribonuclease activity of the ER stress sensor inositol requiring protein 1 $\alpha$  (IRE1 $\alpha$ ) and is followed by ligation of the cleaved ends by RtcB, generating XBP1s. RtcB-mediated XBP1s generation acts as a cellular pro-survival mechanism which is surprisingly also utilized by the tumor cells to promote tumor aggressiveness by an increase in tumor immune infiltration, angiogenesis, and cell invasive properties (Figure 1.1B). Therefore, to decrease tumor progression through the RtcB-XBP1s-endoribonuclease signaling, both the endoribonuclease activity of IRE1 and RtcB serve as prime targets for cancer therapy. Although pharmacological inhibition of the endoribonuclease has achieved success in reducing the pro-survival effects of XBP1s in tumors, it is thought to also attenuate a recently discovered anti-tumorigenic function of its second downstream signaling arm called regulated IRE1-dependent decay (RIDD) (Figure 1.1B). Hence, selective inhibition of the *XBPI* mRNA ligase RtcB that directly joins/re-ligates the spliced *XBPIu* mRNA ends might lead to better treatment outcomes when used as an adjuvant/neo-adjuvant therapy in combination with frontline chemo- and radiotherapies as it is expected to leave the RIDD active. This thesis, therefore, aims to develop drugs that would selectively target the *XBPI* mRNA ligase RtcB in triple-negative breast cancer (TNBC) that exhibits constitutive expression of XBP1s, keeping the RIDD active.



**Figure 1.1| RtcB-mediated *XBP1* mRNA splicing, RIDD and their involvement in tumorigenesis.** (A) RtcB re-ligates IRE1α cleaved *XBP1* mRNA that encodes the transcription factor XBP1 leading to the expression of active transcription factor XBP1s. XBP1s upregulates UPR target genes related to protein quality control, folding and ERAD. IRE1α also cleaves ER-resident mRNAs/microRNAs through a process called RIDD. “Adapted from Volume 334, Issue 6059, (Walter and Ron, 2011) The unfolded protein response: from stress pathway to homeostatic regulation, Page 1082, Copyright (2011). Reprinted with permission from American Association for the Advancement of Science (AAAS).” (B) Biological impact of the antagonistic XBP1s and RIDD signals in glioblastoma multiforme aggressiveness. “Adapted from Volume 10, Issue 3, (Lhomond *et al.*, 2018) Dual IRE1 RNase functions dictate glioblastoma development, Page 12, Copyright (2018), with permission from Wiley.”

## 1.1 RtcB

RtcB is an important tRNA ligase that catalyzes the formation of a phosphodiester bond between cleaved 2',3'-cyclic phosphate and 5'-hydroxyl (5'-OH) RNA ends generated by specific endonucleases during tRNA maturation and *XBP1* mRNA splicing (discussed in detail in section 1.2) of the unfolded response (UPR) (Abelson *et al.*, 1998, Popow *et al.*, 2011, Lu *et al.*, 2014b, Jurkin *et al.*, 2014b, Kosmaczewski *et al.*, 2014). 2',3'-cyclic phosphate and 5'-OH termini, therefore, serve as potential substrates for RtcB. Highly

---

conserved from eukaryotes, bacteria to archaea, the RtcB ligase is also known as HSPC117, C22orf28, FAAP, and D10Wsu52e. RtcB, hence the prime focus of the thesis is an RNA ligase without which, protein stress results in the accumulation of unligated *XBPI* mRNA fragments, and defects in the UPR (Kosmaczewski *et al.*, 2014). In B-cell-specific RtcB knockout mouse, depleting *RtcB* in B cells blocked XBP1s formation during differentiation of B cells into antibody-secreting plasma cells and significantly reduced the capability of plasma cells to secrete antibodies (Jurkin *et al.*, 2014b). Knocking out *RtcB* in *Caenorhabditis elegans* (*C. elegans*) reduced lifespan under conditions of constant ER stress (Kosmaczewski *et al.*, 2014) suggesting that loss of XBP1 functioning in *RtcB* null animals increased the sensitivity to ER stress. Knocking out *RtcB* in mouse ESC cells grew significantly slower compared to control cells (Lu *et al.*, 2014b). There is yet no evidence of *RtcB* knockout in human cells in the literature. However, knocking down *RtcB* in normal fertilized mouse embryos, post-implantation survival of embryos at stages E9 (prior to placenta formation), E12 (after the placenta was fully functional) and E19 (post-natal) resulted in substantial fetal death, accompanied by placental blood clotting, at E12. Deficiency of *RtcB* led to failure in embryo development and fetal formation (Wang *et al.*, 2010). Through XBP1 signaling RtcB was shown to induce neuroprotection in worms. Pan-neuronal-specific knockdown of *RtcB* blocked the neuroprotective phenotype of XBP1 in worms expressing  $\alpha$ -synuclein. Compared to  $\alpha$ -synuclein control worms, RtcB mutants failed to protect dopamine neurons significantly (Ray *et al.*, 2014).

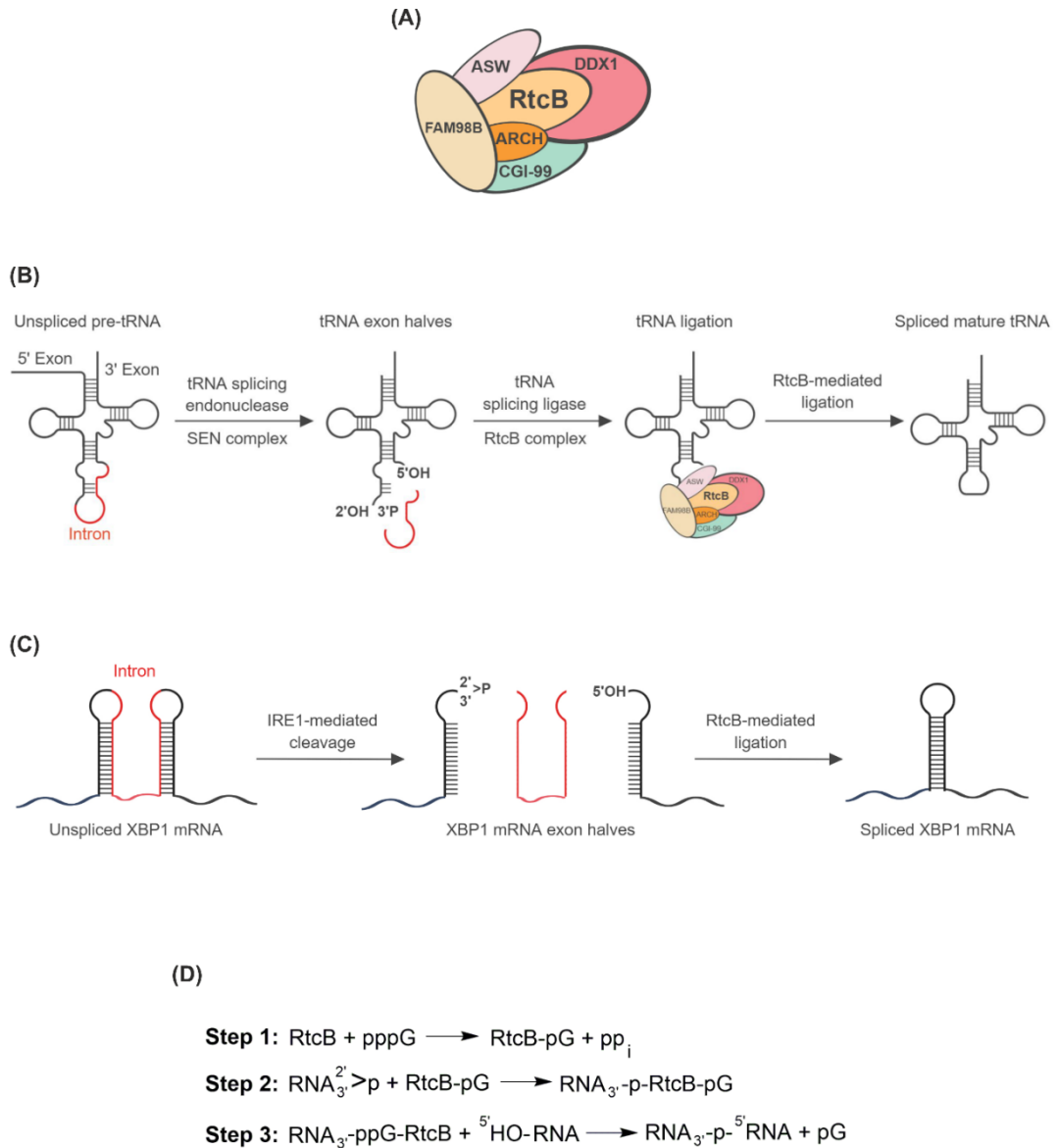
### **1.1.1 RtcB complex**

Human RtcB (*hRtcB*) forms a multimeric protein complex with ASW (also termed as C2orf49), CGI-99 (also termed as C14orf166), FAM98B, Archease (ARCH; also termed as ZBTB8OS) and the DEAD-box helicase DDX1 (Popow *et al.*, 2011, Popow *et al.*, 2014) (Figure 1.2A). Serving as the essential and catalytic subunit of the tRNA ligase complex (Popow *et al.*, 2011, Popow *et al.*, 2014), RtcB re-ligates cleaved tRNA ends (Figure 1.2B), and *XBPIu* mRNA fragments forming *XBP1s* (Figure 1.2C) through a reading frameshift leading to translation of the *XBP1s* mRNA (Lu *et al.*, 2014b, Jurkin *et al.*, 2014b, Kosmaczewski *et al.*, 2014). Two other subunits of the complex, namely, Archease and DDX1 also participate in the re-ligation process. Along with *hRtcB*, Archease and DDX1 catalyze the formation of *hRtcB*-guanylate (first step of ligation)

---

(Popow *et al.*, 2014). The remaining three subunits of the complex namely, ASW, CGI-99, and FAM98B have functions that are yet unknown.

Importantly, the involvement of Archease in catalysis of *hRtcB*-guanylate could be explained by its role as a stimulatory co-factor required for full activity of the human tRNA ligase complex (Popow *et al.*, 2014). RNA ligase activity of *hRtcB* is only activated upon Archease co-expression (Poothong *et al.*, 2017) and Archease accelerates the RNA ligation efficiency of *Pyrococcus horikoshii* RtcB (*bRtcB*) (Desai *et al.*, 2014). Found in the form of a small 16 kDa acidic protein conserved in all three domains of life, (Canaves, 2004) Archease has no RNA ligase activity (Poothong *et al.*, 2017, Jurkin *et al.*, 2014a, Desai *et al.*, 2014) and functionally interacts with *bRtcB* (Desai *et al.*, 2014). Effect of Archease on RNA ligation kinetics revealed that titrating Archease into *bRtcB*-containing ligation reactions yielded ~3.5 times more ligation product as compared to control (without Archease) (Desai *et al.*, 2014). In contrast, in spite of having 5-fold excess *bRtcB* over RNA substrates *in vitro*, the ligation reaction with RtcB alone allowed only ~10% of *bRtcB* molecules to execute 43% complete ligation reaction with the ligation completion rising to 49% upon Archease addition (Desai *et al.*, 2014). This limited ligation efficiency (43 – 49%) of *bRtcB* with or without Archease prompts the involvement of other subunits of the ligase complex or other unknown factors that catalyze the ligation efficiency of RtcB. DDX1 helicase was also shown to be necessary for full human RNA ligase activity to mediate pre-tRNA ligation (Popow *et al.*, 2014). In association with Archease, DDX1 facilitates the formation of *hRtcB*-guanylate intermediate central to mammalian RNA ligation (Popow *et al.*, 2014). In sharp contrast, the *Escherichia coli* (*E. coli*) RtcB alone catalyzes RNA ligation without any co-factors (Tanaka *et al.*, 2011b, Poothong *et al.*, 2017). Hence, a dual co-factor mediated stimulation is essential for the *hRtcB* to carry out RNA ligation.



**Figure 1.2| RtcB ligase facilitates re-ligation of tRNA and XBP1 mRNA ends.** (A) tRNA ligase complex. The complex consists of six subunits including RtcB (catalytic subunit), DDX1, FAM98B, ASW, CGI-99, ASW, and Archease. (B) Enzymatic splicing of pre-tRNA. Splicing of the pre-tRNA intron (in color red) is accomplished by an endonucleolytic cleavage followed by re-ligation by the tRNA ligase RtcB. (C) Splicing of XBP1 mRNA intron (in color red) is accomplished by separate cleavage and ligation reactions catalyzed by the endoribonuclease activity of IRE1 and the tRNA ligase RtcB, respectively. (D) RtcB-mediated RNA ligation reaction mechanism. RtcB interacts with GTP to form RtcB-guanylate intermediate central to mammalian RNA ligation in step 1. In step 2, the GMP moiety interacts with the RNA 3'-P, and in step 3, the 5'-OH from the other RNA strand attacks the activated 3'-P to form a phosphodiester bond and releases the GMP (Desai *et al.*, 2013, Englert *et al.*, 2012).

---

### **1.1.2 RNA ligation mechanism**

Ligation by RtcB proceeds through a three-step nucleotidyl transfer mechanism (Figure 1.2D) with 2',3'-cyclic phosphate termini being hydrolyzed to 3'-phosphate (3'-P) termini in a step that precedes 3'-P activation with GMP (Desai *et al.*, 2013). In the first step, RtcB reacts with GTP to form a covalent RtcB-histidine-GMP intermediate and release PP<sub>i</sub>; in the second step, the GMP moiety is transferred to the RNA 3'-P; in the third step, the 5'-OH from the opposite RNA strand attacks the activated 3'-P to form a 3',5'-phosphodiester bond and release GMP (Desai *et al.*, 2013). Compared to classical RNA ligases which rely on ATP and Mg(II) for catalysis, ligation by RtcB is dependent on GTP and Mn<sup>2+</sup> (Tanaka *et al.*, 2011a, Chakravarty *et al.*, 2012, Chakravarty and Shuman, 2012, Desai *et al.*, 2013).

### **1.1.3 RtcB subcellular localization**

RtcB localizes both in the cytoplasm and in the nuclear/perinuclear regions (Jurkin *et al.*, 2014b, Lu *et al.*, 2014b). While mammalian tRNA splicing occurs in the nucleus (De Robertis *et al.*, 1981, Lund and Dahlberg, 1998), *XBPI* mRNA splicing takes place in the cytoplasm (Yanagitani *et al.*, 2009, Sidrauski *et al.*, 1996). A study by Pérez-González *et al.* (2014), demonstrated the human *hCLE/C14orf166* protein, a constituent of nuclear and cytoplasmic protein complexes involved in different steps of nuclear and cytoplasmic RNA metabolism such as transcription, maturation, and translation, forms a complex with *hRtcB* and its interacting subunits DDX1 and FAM98B of the tRNA ligase complex. This *hCLE-hRtcB-DDX1-FAM98B* complex was found to be present both in the nucleus and cytoplasm of HEK293T cells. *hCLE* was shown to modulate the expression of its interacting partners as RNAi-mediated depletion of *hCLE* provoked the destabilization of the complex and consequently the degradation of all the associated proteins both in the nuclear and cytoplasmic compartments. By means of *hCLE* tagged with a photoactivatable GFP and live cell imaging, nuclear import and export of *hCLE* were reported and that *hCLE* probably in complex with its interacting partners shuttles between the nucleus and cytoplasm transporting RNAs. Shuttling between the cytoplasm and nucleus, the RtcB ligase therefore fulfills the common goal of joining broken RNA ends, controlling both nuclear and cytoplasmic RNA fate (Perez-Gonzalez *et al.*, 2014, Jurkin *et al.*, 2014a).



---

#### **1.1.4 *RtcB* facilitates tRNA ligation**

RtcB promotes post-transcriptional maturation of endogenous intron-containing precursor tRNAs (pre-tRNAs) in the nucleus (Laski *et al.*, 1983, De Robertis *et al.*, 1981, Lund and Dahlberg, 1998, Popow *et al.*, 2011, Kosmaczewski *et al.*, 2014). Depletion of *hRtcB* *in vitro* abolishes interstrand ligation of intron-containing pre-tRNAs, impairs processing of pre-tRNA transcripts into mature tRNAs and leads to accumulation of tRNA exon halves (Popow *et al.*, 2011). With a 73% identity to *hRtcB*, loss-of-function of *C. elegans* RtcB *in vivo* impairs ligation of endogenous intron-containing tRNAs and leads to the accumulation of unligated tRNA exon halves (Kosmaczewski *et al.*, 2014).

In 2011 Javier Martinez's group identified the *hRtcB* complex (Popow *et al.*, 2011), and showed that efficient depletion of the *hRtcB*-interacting subunits of the complex including DDX1, CGI-99, FAM98B, and ASW did not impair interstrand ligation and maturation of intron-containing pre-tRNAs. However, depleting *Archease* impaired maturation of intron-containing human pre-tRNAs *in vitro* to a comparable extent as observed upon *hRtcB* depletion (Popow *et al.*, 2014); suggesting further an elusive role of the other four subunits of the complex towards RtcB's ligase activity. The authors also noted that depleting *hRtcB* simultaneously depleted *DDX1* and *FAM98B* (Popow *et al.*, 2011). Importantly, in humans, only 32 (Ile, Tyr and Arg families) of 506 tRNA genes contain intron sequences (Kosmaczewski *et al.*, 2014) and therefore reduction in splicing-dependent mature tRNAs did not lead to a global defect in protein synthesis in RtcB- and Archease-depleted cells (Jurkin *et al.*, 2014a).

#### **1.1.5 *RtcB* facilitates XBP1 mRNA ligation**

RtcB re-ligates cleaved *XBP1u* mRNA ends forming *XBP1s* during the UPR (Lu *et al.*, 2014b, Jurkin *et al.*, 2014b, Kosmaczewski *et al.*, 2014). Depleting *hRtcB* impaired *in vitro* ligation of the *XBP1* mRNA exon halves with results slightly improving upon depletion of both *hRtcB* and *Archease* (Jurkin *et al.*, 2014a). Surprisingly, RNAi-mediated depletion of *hRtcB* alone led to *XBP1s* mRNA expression almost to wild-type levels endogenously. This suggests that the stimulatory activity of Archease sustains ligation activity in the presence of reduced amounts of RtcB (Jurkin *et al.*, 2014a). Depletion of neither *hRtcB* nor *Archease* alone could fully abolish XBP1s induction endogenously (Jurkin *et al.*, 2014a). However, on depleting both *hRtcB* and *Archease*,

---

XBP1s was abolished at the protein level and greatly reduced at the mRNA level (Jurkin *et al.*, 2014b). This further reduced the levels of total *XBPI* and *XBPIu* mRNA and also the mRNA levels of XBP1s target genes, *EDEMI*, and *DNAJB9* (Jurkin *et al.*, 2014b). Importantly, complete knockout of *RtcB* alone sufficiently blocked XBP1s protein in mouse plasma cells (Jurkin *et al.*, 2014a) as well as in mouse ES cells (Lu *et al.*, 2014b). However, in the mouse *RtcB* knockout phenotype, low levels of *XBPIs* mRNA were still observed (Lu *et al.*, 2014b), suggesting ligation of the cleaved *XBPIu* substrates by yet unknown compensatory RNA ligases. Interestingly in *RtcB* knockout *C. elegans* model, *XBPIs* was not detectable under ER stress (Kosmaczewski *et al.*, 2014). More importantly, mRNA levels of known RIDD substrates, namely *BLOS1* and *PDGFRB* remained unchanged after *hRtcB* and *Archease* depletion (Jurkin *et al.*, 2014a). This demonstrated that depletion of *hRtcB* and *Archease* does not interfere with IRE1's endonucleolytic activity but specifically interrupts *XBPI* mRNA splicing and the induction of its downstream target genes; providing promise for the development of *RtcB* targeted drugs in the thesis. More details about IRE1, RIDD and XBP1 signaling will be provided in the next section 1.2.

### **1.1.6 *RtcB* facilitates other biological processes**

*RtcB* in other organisms has been reported to be involved in other biological processes. The *RtcB* gene in *E. coli* resides within a  $\sigma 54$ -regulated operon that contains two additional open reading frames (ORFs) named *RtcA* and *RtcR* (Genschik *et al.*, 1998, Engl *et al.*, 2016). *RtcA* is positioned immediately downstream of *RtcB* whereas *RtcR* is positioned upstream of the *RtcA/RtcB* (*RtcAB*) unit and is transcribed in the opposite direction. Conserved across all three domains of life (Genschik *et al.*, 1997), *RtcA* functions as an RNA 3'-P terminal cyclase enzyme that catalyzes the enzymatic modification of a 3'-P RNA end to a 2',3'-cyclic phosphodiester (Genschik *et al.*, 1998). Expression of *RtcAB* genes is transcriptionally activated by the enhancer binding protein *RtcR* in a  $\sigma 54$ -RNA polymerase-dependent manner. *RtcR* transduces an unknown signal that drives the elevated expression of *RtcAB* (Engl *et al.*, 2016). In fact, the whole conglomerate of three ORFs termed as the *Rtc* system, under normal condition, functions in maintaining ribosome homeostasis and chemotaxis in bacterial cells. Cells that lack *RtcR* and/or *RtcAB* show *Rtc*-dependent changes in the transcriptome with a substantial proportion of genes differentially regulated that correspond to the translation apparatus.

---

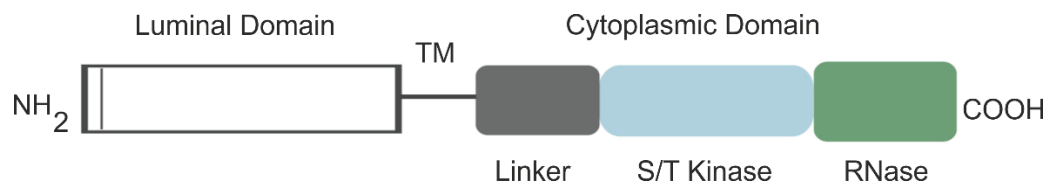
Moreover, translation inhibition by exogenous stress upregulated *RtcAB* promoter activity in a RtcR-dependent manner. Some other roles of RtcAB include the inhibition of post-injury axon regeneration in the *Drosophila* and *C. elegans* central nervous system (Song *et al.*, 2015, Kosmaczewski *et al.*, 2015). Further, loss of *RtcB* causes defects in development and fertility in *C. elegans* (Kosmaczewski *et al.*, 2014), that is independent of tRNA maturation and the UPR; a role outside its essential functioning in tRNA biogenesis and the UPR. Engineered RtcB mutants expressing pre-spliced tRNAs still resulted in a defective-developmental phenotype and other mutations that abrogate the UPR including that of *IRE1* and *XBPI* did not result in any developmental defects (Kosmaczewski *et al.*, 2014). In a broader perspective, the Rtc system hence plays a role far beyond the classical role of RtcB in ligating broken RNA ends.

## 1.2 IRE1

IRE1 (also known as ERN1), a type I ER transmembrane protein is a core element of the UPR signaling network. It is one of three ER stress sensors (other two being PERK and ATF6), and upon activation it produces a transcription factor (XBP1s) that activates genes to increase the protein-folding capacity in the ER (Hetz *et al.*, 2013). From yeast to humans, it's a highly conserved enzyme with two conserved paralogs found in mammals, IRE1 $\alpha$  and IRE1 $\beta$ . While IRE1 $\alpha$  is expressed ubiquitously, expression of IRE1 $\beta$  is majorly limited to gut epithelial cells (Human Protein Atlas) (Boot-Handford and Briggs, 2010, Bertolotti *et al.*, 2001). IRE1 $\beta$  regulates microsomal triglyceride transfer protein (MTP) and chylomicron production that is needed to assemble chylomicrons in the ER of enterocytes. Mice deficient in IRE1 $\beta$  developed more pronounced hyperlipidemia as a result of elevated chylomicron secretion and intestinal MTP expression. In cell culture, IRE1 $\beta$ , but not IRE1 $\alpha$ , decreased MTP mRNA through increased posttranscriptional degradation (Iqbal *et al.*, 2008). Other tissues where IRE1 $\beta$  is expressed include brain, endocrine, bone marrow, muscle, lung, liver and gallbladder, pancreas, kidney and urinary bladder, tissues of the reproductive system, adipose and soft tissues, and skin (Human Protein Atlas).

Structural analysis revealed that IRE1 $\alpha$  consists of an N-terminal luminal domain, a single pass transmembrane segment and a cytoplasmic domain. The luminal domain is tethered via the transmembrane segment to the cytoplasmic domain which is further

subdivided into a linker region, an auto-phosphorylating Ser/Thr protein kinase domain and a C-terminal endoribonuclease (RNase) domain (Figure 1.3) (Tirasophon *et al.*, 1998, Ali *et al.*, 2011). Catalytic activities of IRE1 $\alpha$  are housed in these cytoplasmic domains. The linker region of IRE1, however, is thought to play a role in minimizing the chances of other mRNA editing by error (van Anken *et al.*, 2014).



**Figure 1.3| Domain structure of IRE1.** Domain organization of IRE1. TM: transmembrane segment; Linker: a region having no homology to known proteins; S/T kinase: catalytic domain of Ser/Thr protein kinase; RNase: endoribonuclease domain.

The mechanism of IRE1 $\alpha$  activation includes the protein being held in a monomeric and inactive state by the ER chaperone binding immunoglobulin protein (BiP) (Bertolotti *et al.*, 2000, Kimata *et al.*, 2003) which binds to the luminal domain of IRE1 $\alpha$ . Upon accumulation of unfolded proteins in the ER, BiP is released, permitting dimerization of the luminal domains (Credle *et al.*, 2005, Zhou *et al.*, 2006) which further promotes higher-order oligomers and transphosphorylation (phosphotransfer) of the cytoplasmic kinase domain (Shamu and Walter, 1996, Welihinda and Kaufman, 1996, Han *et al.*, 2009). The kinase domains of human IRE1 $\alpha$  form a face-to-face dimer in a way that the activation segment of one monomer can be phosphorylated by the active site of the other (Ali *et al.*, 2011). This leads to the activation of the splicing activity of the cytoplasmic ribonuclease domain (Tirasophon *et al.*, 2000, Ali *et al.*, 2011) which is in sharp contrast to the yeast IRE1 RNase activation that proceeds from IRE1 monomers forming back-to-back dimers (Lee *et al.*, 2008, Korennykh *et al.*, 2009, Wiseman *et al.*, 2010). Further, the RNase activity has been reported to be dependent on the kinase activity as kinase-dead IRE1 mutants and mutants in which the autophosphorylation site is mutated exhibit a defective RNase and UPR signaling (Shamu and Walter, 1996, Welihinda and Kaufman, 1996, Tirasophon *et al.*, 2000, Ali *et al.*, 2011).

Importantly, it is the RNase activity by which IRE1 is able to initiate *XBPI* mRNA splicing (Sidrauski and Walter, 1997, Tirasophon *et al.*, 2000) and therefore is of prime

---

importance to the thesis. Besides splicing *XBP1*, IRE1 $\alpha$  RNase is also involved in degrading a subset of mRNA/miRNAs through a process termed as RIDD (Hollien and Weissman, 2006, Han *et al.*, 2009, Hollien *et al.*, 2009, Tam *et al.*, 2014) as mentioned at the beginning. While XBP1s confers pro-survival advantages, RIDD promotes cell death (Maurel *et al.*, 2014). Essentially, as the IRE1 $\alpha$  RNase activity cleaves target RNAs through its XBP1s and RIDD outputs, RtcB, therefore, finds the novel opportunity to re-ligate these broken RNA ends. As stated before, IRE1 $\alpha$  RNase activity cleaves *XBP1* mRNA generating 2',3'-cyclic phosphate and 5'-OH RNA ends. Importantly, RIDD has also been recently suggested to generate similar RNA ends, hence making them ideal substrates to be re-ligated by RtcB (Filipowicz, 2014), although experimental evidence of such an event is not yet available.

IRE1 is involved in multiple physiological events including Ca<sup>2+</sup> regulation (Son *et al.*, 2014), glucose homeostasis (Shao *et al.*, 2014), ER membrane lipid composition (Volmer *et al.*, 2013), adipocyte differentiation (Sha *et al.*, 2009), regulating the cell cycle *cyclin A1* gene (Thorpe and Schwarze, 2010), and secretory immune functions (Reimold *et al.*, 2001). IRE1 $\alpha$  through its downstream target XBP1 regulates expression and activities of key enzymes in phospholipid biosynthesis and contributes to ER membrane expansion (Sriburi *et al.*, 2007). Mutant IRE1 $\alpha$  lacking luminal stress-sensing domain still responds to increased lipid saturation through the transmembrane domain (Volmer *et al.*, 2013). The non-muscle myosin IIB interacts with IRE1 during ER stress, suggesting that IRE1's interaction with cytoskeleton is required for optimal IRE1 activity (He *et al.*, 2012). IRE1 was shown to regulate pancreatic  $\beta$ -cell homeostasis by controlling insulin levels synthesized in the ER. Treating  $\beta$  cells with high glucose hyperactivated IRE1, which decreased insulin mRNA expression (Lipson *et al.*, 2008). *In vitro* studies also revealed that IRE1 cleaves Insulin 1 and Insulin 2 mRNAs in a site-specific manner (Han *et al.*, 2009).

The broad functioning of IRE1 $\alpha$  is possibly due to its ability to assemble with many regulatory proteins forming a complex dynamic scaffold. These modulatory proteins directly interact with IRE1 $\alpha$  forming a macromolecular complex often termed as the UPRosome (Sitia and Braakman, 2003, Hetz and Glimcher, 2008). Different UPRosome elicits different downstream signals leading to different cell fate. One such modulatory protein is the heat shock protein (HSP). Known for its cell survival functions, HSP70

---

(also known as HSP72) and HSP90 have been shown to directly interact with IRE1 $\alpha$ . HSP70 imparts a strong cytoprotective effect inhibiting apoptosis. By binding to IRE1 $\alpha$ , HSP70 enhances IRE1 $\alpha$ 's RNase activity, promoting adaptation to ER stress and facilitating cell survival (Gupta *et al.*, 2010). Silencing HSP70 abrogated ER stress-induced IRE1 $\alpha$  upregulation and promoted cell death (Stornuolo *et al.*, 2015). HSP90, on the other hand, was shown to regulate IRE1 $\alpha$  oligomerization and autophosphorylation in a Cdc37 cochaperone-dependent manner (Ota and Wang, 2012). Another protein called RING finger protein 13 (RNF13) was shown to enhance IRE1 $\alpha$  stability through direct interaction with IRE1 $\alpha$  (Arshad *et al.*, 2013). The pro-apoptotic Bcl-2 family Bax and Bak form a protein complex with IRE1 $\alpha$ 's cytosolic domain regulating IRE1 $\alpha$  activation and is independent of Bax/Bak's proapoptotic function at the mitochondria (Hetz *et al.*, 2006). Interestingly, the Bax inhibitor-1 (BI-1) protein that suppresses cell death was reported to form a stable protein complex with IRE1 $\alpha$ , decreasing its RNase activity; suggesting BI-1's role in early adaptive responses against ER stress that contrasts with its known downstream function in apoptosis (Lisbona *et al.*, 2009). Poly (ADP-ribose) polymerase 16 (PARP16), one of the members of the PARP family that regulate DNA damage repair and cytoplasmic stress response was shown to interact with IRE1 $\alpha$  both in the presence and absence of ER stress. Both the kinase and RNase activities of IRE1 $\alpha$  was elevated by PARP16, and further facilitated cell survival by dissociating BiP from IRE1 $\alpha$  under ER stress (Jwa and Chang, 2012). Some other IRE1 $\alpha$  interactors that were reported in the early years include Nck1 (Nguyen *et al.*, 2004), protein-tyrosine phosphatase 1B (Gu *et al.*, 2004), AIP1 (Luo *et al.*, 2008), BIM/PUMA (Ren *et al.*, 2010) and RACK1 (Qiu *et al.*, 2010).

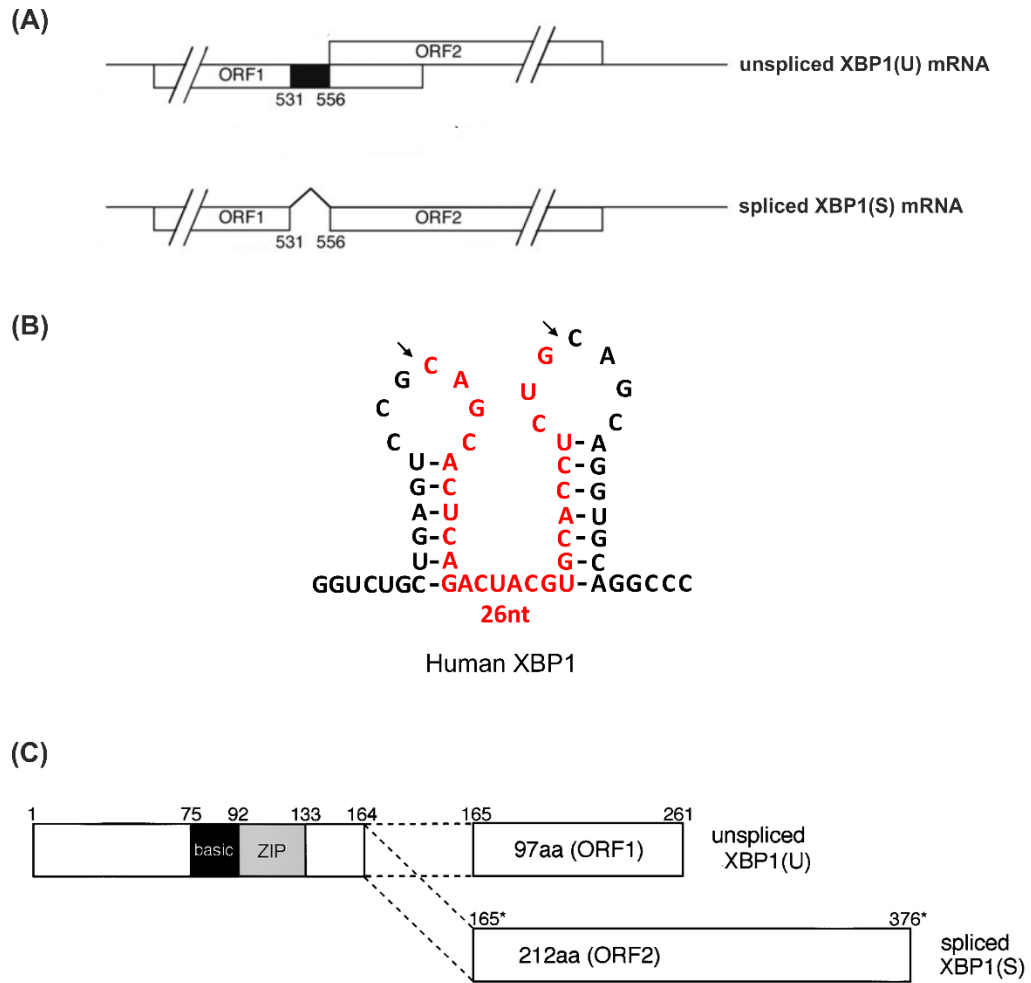
IRE1 $\alpha$  is predominantly activated in the placenta during mouse embryogenesis. Loss of IRE1 $\alpha$  affects the expression of VEGF-A and leads to placental dysplasia. Furthermore, loss of IRE1 $\alpha$  was shown to affect proliferation, ER development, and ER stress level in trophoblasts of the placenta. Studies in IRE1 $\alpha$  knockout mice revealed reduced cell proliferation of trophoblasts, poor ER development in trophoblasts, and higher BiP expression levels compared to the wild-type placenta (Iwawaki *et al.*, 2009). IRE1 is also activated in lipid-laden macrophages that infiltrate the atherosclerotic lesions. Expression of many proatherogenic genes is regulated by IRE1, including several important cytokines and chemokines. Inhibiting IRE1 *in vivo*, significantly decreased hyperlipidemia-induced IL-1 $\beta$  and IL-18 production, lowered T-helper type-1 immune

---

responses, and reduced atherosclerotic plaque size in mice, thereby counteracting metaflammation and alleviating atherosclerosis (Tufanli *et al.*, 2017).

### **1.2.1 IRE1 $\alpha$ – XBP1**

Activated IRE1 $\alpha$  RNase excises a 26-nucleotide (531 to 556) intron from its substrate; an mRNA that encodes the transcription factor *XBP1* in mammals (Figure 1.4A) or *HAC1* in yeast (Cox and Walter, 1996, Yoshida *et al.*, 2001). *XBP1* mRNA encodes two ORFs. ORF1 and ORF2. ORF1 encodes basic leucine zipper (bZIP) protein of 261 amino acid. ORF2 partially overlaps but is not in frame with ORF1 encoding protein of 222 amino acid. Excision occurs within the cytoplasm at a dual stem loop (SL) structure (Figure 1.4B), that leaves a 2',3'-cyclic phosphate end and a 5'-OH end which is then subsequently re-ligated by RtcB (Yoshida *et al.*, 2001, Lu *et al.*, 2014b, Jurkin *et al.*, 2014b, Kosmaczewski *et al.*, 2014). Re-ligation by RtcB causes a reading frame-shift (Figure 1.4C) of the excised substrate that leads to translation of a more active and stable isoform termed XBP1s (Yoshida *et al.*, 2001, Liu *et al.*, 2000). XBP1s is a basic leucine zipper protein which consists of an acidic domain, a serine/threonine, glutamine, and proline rich regions. Excision causes the removal of the C-terminal 97 amino acid portion from ORF1 and re-ligation causes addition of 212 amino acid portion of ORF2 to the N-terminal 164 amino acid of ORF1 containing the bZIP domain, resulting in a frame shift at amino acid 165 (Yoshida *et al.*, 2001) (Figure 1.4C). Hence, *XBP1u* (for unspliced form) mRNA encodes a protein of 261 amino acids while *XBP1s* mRNA encodes a protein of 376 amino acids. The C-terminal replacement, however, does not affect protein stability of both XBP1 isoforms (Yoshida *et al.*, 2001). This C-terminal region of XBP1s is the key to its transactivation activity and importantly harbors the nuclear localization signal (Clauss *et al.*, 1996). However, in unstressed condition, *XBP1* is constitutively transcribed and translated but due to the presence of a premature stop codon that prevents the translation of the transactivation domain, the XBP1 protein remains non-functional (Tirosh *et al.*, 2006).



**Figure 1.4| ER stress-mediated *XBP1* mRNA splicing.** (A) Locations of the ORFs 1 and 2 within *XBP1u* and *XBP1s* mRNA are indicated by closed boxes. Nucleotides 531-556 represent the position of the excised nucleotides. “Adapted from Volume 172, Issue 4, (Yoshida *et al.*, 2006) pXBP1(U) encoded in XBP1 pre-mRNA negatively regulates unfolded protein response activator pXBP1(S) in mammalian ER stress response, Page 567, Copyright (2006), with permission from Rockefeller University Press.” (B) The sequence of human XBP1 mRNA stem loop cleaved by IRE1 $\alpha$ . Cleavage sites are indicated by arrows and the 26 nucleotides that are spliced out indicated in color red. (C) Reading frame-shift during *XBP1* splicing. Numbers denote amino acid positions with initiation methionine set at 1. Numbers with asterisks denote the amino acid encoded by ORF2. Basic and leucine zipper (ZIP) domains denoted in shaded boxes. “Adapted from Volume 107, Issue 7, (Yoshida *et al.*, 2001) XBP1 mRNA Is Induced by ATF6 and Spliced by IRE1 in Response to ER Stress to Produce a Highly Active Transcription Factor, Page 884, Copyright (2001), with permission from Elsevier.”

XBP1s functions both in UPR and non-UPR-related pathways. Following excision, XBP1s enters the nucleus and binds to the unfolded protein response element (UPRE) present in the promoter regions of UPR-target genes that drives protein folding, protein translocation to the ER, maturation, secretion, and degradation (ERAD) to correct ER



---

homeostasis (Lee *et al.*, 2003b, Acosta-Alvear *et al.*, 2007, Sriburi *et al.*, 2004). UPRE is the CCACG site of ER stress response element (ERSE); an evolutionary conserved tripartite structure CCAAT(N<sub>9</sub>)CCACG (N being a GC-rich region of 9 bp) conserved amongst glucose-regulated protein (GRP) promoters (Roy and Lee, 1999). XBP1s can also dimerize with ATF6 $\alpha$  for transcriptional upregulation of major ERAD components (Yamamoto *et al.*, 2007). Some non-UPR-related pathways upregulated by XBP1s include the production of pro-inflammatory cytokines, lipid and hexosamine biosynthesis and hypoxia response (Bettigole and Glimcher, 2015). In contrast, nascent XBP1u directs the mRNA-ribosome nascent chain complex to the ER membrane for splicing to occur (Yanagitani *et al.*, 2009). Interestingly, both these isoforms contain a DNA binding region. XBP1s can bind to DNA as a monomer but also forms a heterodimer with c-FOS during transcription of a subset of MHC class II genes (Ono *et al.*, 1991).

### **1.2.2 IRE1 $\alpha$ – RIDD**

In 2006, for the first time, IRE1 RNase activity was shown to be involved in degrading ER resident mRNAs in *Drosophila melanogaster* (*D. melanogaster*) through a process called RIDD (Hollien and Weissman, 2006) that is conserved in mammals (Han *et al.*, 2009, Hollien *et al.*, 2009), and fission yeast (Kimmig *et al.*, 2012). While splicing of XBP1 favors cell survival by increasing cellular adaptability to stress, RIDD reduces cell growth in a cell-specific manner by degrading mRNAs that encode growth-promoting proteins leading to increased cell death (Maurel *et al.*, 2013b, Dejeans *et al.*, 2012, Pluquet *et al.*, 2013). Although *XBP1* splicing is an outcome of IRE1's oligomerization and phosphotransfer activation modes, RIDD was thought to be solely induced by the phosphotransfer activation, transmitting proapoptotic signals independent of *XBP1* splicing (Han *et al.*, 2009). However, this proapoptotic signaling failed to induce apoptosis in kinase-active/RNase-dead mutants suggesting further that an active kinase/RNase is indeed required to induce cell death and that the proapoptotic outputs do not proceed through the phosphotransfer activity alone (Han *et al.*, 2009).

RIDD is enhanced by stress. Time-course analyses from several studies revealed that under ER stress when *XBP1* splicing is maximal, expression levels of RIDD targets have already started to go down and continue to decay with time before cell death occurs (Han *et al.*, 2009, Upton *et al.*, 2012, Lerner *et al.*, 2012). If the stress is not alleviated RIDD

---

leads to cell death as its activity keeps rising proportionally with the intensity of stress (Maurel *et al.*, 2014). In fact, numerous RIDD targets have been found to be upregulated basally under conditions of impaired IRE1 activity or expression (Gaddam *et al.*, 2013, Dejeans *et al.*, 2012, Maurel *et al.*, 2013b, Pluquet *et al.*, 2013, So *et al.*, 2012a), which implies that RIDD can happen both in the presence and absence of ER stress (Gaddam *et al.*, 2013, Cho *et al.*, 2013) and that its activity increases gradually with the intensity/duration of stress (Hollien *et al.*, 2009, Dejeans *et al.*, 2012, Pluquet *et al.*, 2013). This further implies that RIDD does not require ER stress but only gets enhanced by the stress. What's more intriguing is that while XBP1 splicing activity is only transient during adaptive ER stress, RIDD activity is constitutive that gradually increases with the intensity/duration of the stress and subsequently becomes cytotoxic under prolonged or unmitigated stress. However, the exact mechanism by which RIDD switches from cytoprotective to cytotoxic remains elusive (Maurel *et al.*, 2014).

RIDD relies on mRNA structure, sequence, and translation status of target mRNAs. Unlike in *D. melanogaster* where RIDD was originally discovered (Hollien and Weissman, 2006), the presence of XBP1-like SLs is more prevalent in mammalian RIDD targets. IRE1 directly cleaves RIDD targets in their XBP1-like SLs. Although RIDD appears to be a sequence-specific event that occurs at XBP1-like SLs, the presence of such SLs within mRNAs is always not sufficient to predict RIDD targets (Moore and Hollien, 2015). In fact, a common consensus sequence 5'-CUGCAG-3' has been reported in IRE1 $\alpha$  cleavage sites both for RIDD mRNA/miRNAs and XBP1 mRNAs (Oikawa *et al.*, 2010, So *et al.*, 2012a, Oikawa *et al.*, 2007, Hur *et al.*, 2012, Coelho *et al.*, 2013, Maurel *et al.*, 2013a, Upton *et al.*, 2012, Moore *et al.*, 2013, Maurel *et al.*, 2014). Importantly, translation attenuation of target mRNAs itself is important for degradation by RIDD. Unlike the wild-type counterpart, introducing translation-halting sequences in RIDD reporters, degradation of the translationally halted reporter mRNAs remained unaffected by PERK (translation attenuator) depletion (Moore and Hollien, 2015). This implies that mammalian RIDD, therefore, relies on the mRNA sequence, structure, and translational status of target mRNAs.

The physiological relevance of RIDD has been demonstrated in several studies. RIDD plays a role in controlling the expression of lipogenic enzymes and modulating lipid levels in the serum. In studies where mice were fed with high fat or high cholesterol diet,

---

IRE1 $\beta$  played a protective role against hyperlipidemia by decreasing lipid absorption in the intestine (Iqbal *et al.*, 2008). In the liver, upon *XBPI* deficiency, IRE1 $\alpha$  gets hyperactivated leading to a hypolipogenic phenotype in mice thereby reducing plasma cholesterol and triglycerides (So *et al.*, 2012b). Silencing *IRE1 $\alpha$*  in *XBPI* deficient mice induced genes encoding Angiopoietin-like protein 3 (Angptl3) and *ces1* genes that are involved in lipid metabolism which were further validated as RIDD substrates (So *et al.*, 2012b). Inefficient maturation of insulin in  $\beta$ -cells deficient for *XBPI* occurs partly due to RIDD. Silencing *XBPI* hyperactivated IRE1 $\alpha$  in Min6 insulinoma cells allowing cleavage of a subset of mRNAs encoding insulin secretory pathway genes, namely *PC1*, *PC2*, and *CPE* (Lee *et al.*, 2011). RIDD also participates in maintaining homeostasis of the immune system. RIDD is required in CD8 $\alpha^+$  dendritic cells for cross-presentation of cell-derived antigens via major histocompatibility complex (MHC)-class I to CD8 $^+$  T cells (Osorio *et al.*, 2014). In CD8 $\alpha^+$  conventional dendritic cells, RIDD plays a critical role in regulating the expression of integrins and components of the MHC-class I antigen-presentation machinery (Osorio *et al.*, 2014). Finally, RIDD was also active in B cells, where it cleaved the mRNA of secretory  $\mu$  chains (Benhamron *et al.*, 2014).

RIDD regulates proapoptotic signaling. Upton and coworkers (Upton *et al.*, 2008) have shown that the ER stress-mediated BAX/BAK-dependent pro-apoptotic pathway operates through RIDD of miRNA precursors. They showed that by degrading respective miRNA precursors, RIDD reduces expression of miRNA-17, -34a, -96 and -125b which normally represses translation of proapoptotic caspase-2. Furthermore, by using anti-miRNA oligonucleotides against these miRNAs, the authors showed that caspase-2 levels increased in an IRE1 $\alpha$ -dependent manner. This clearly explains that IRE1 $\alpha$  facilitates RIDD of miRNA-mediated cell death via downregulating specific miRNAs that in turn de-repress caspase-2 translation which then leads to activation of this caspase and subsequent BAX/BAK-mediated cell death. However, this novel finding is controversial as Sandow and coworkers (Sandow *et al.*, 2013) did not detect any increase in caspase-2 protein expression under similar experimental conditions although they never challenged the fact that IRE1 cleaves several precursor miRNAs. Another study by Lerner *et al.* (2012) also reported RIDD of miRNA-17-dependent cell death. It was shown that IRE1-mediated reduction of miRNA-17 upregulates expression of the pro-oxidant thioredoxin-interacting protein (TXNIP), another target of miR-17. Elevated TXNIP protein activates the NLRP3 inflammasome leading to pro-caspase-1 cleavage

---

and interleukin (IL) 1 $\beta$  (IL-1 $\beta$ ) secretion thereby promoting inflammation and cell death.

The assessment by Maurel and colleagues (Maurel *et al.*, 2014) that RIDD, in general, tends to be pro-death is controversial. RIDD was shown to counteract apoptosis by mediating death receptor 5 (DR5) mRNA decay (Lu *et al.*, 2014a). While PERK-CHOP activity induced *DR5* mRNA up-regulation, RIDD suppressed *DR5* transcripts, allowing cellular adaptation to apoptosis (Lu *et al.*, 2014a). In *XBPI* deficient mice, RIDD inhibition did not rescue lung conventional dendritic cells (cDC1s), arguing against RIDD's pro-death role (Tavernier *et al.*, 2017). In fact, RIDD was shown to mediate a protective role in *XBPI* deficient cDC1 survival. Loss of *XBPI* in cDC1s activated both IRE1 and PERK resulting in inhibition of cell death in the intestine in a RIDD-dependent manner (Tavernier *et al.*, 2017). Hence, IRE1-mediated RIDD activity alleviated the deleterious effects of *XBPI* deficiency in cDC1s and protected against cell death, promoting cell survival. Furthermore, RIDD was shown to promote cell survival by regulating protein translation in *XBPI* deficient hepatocytes by degrading Ppp1r15b mRNA. RIDD-mediated decrease of Ppp1r15b mRNA increased eIF2 $\alpha$  phosphorylation, suppressing protein synthesis thereby decreasing cellular stress (So *et al.*, 2015). Also, it is interesting to note that RIDD has been implicated in protecting the liver from drug-induced toxicity. Several drug-metabolizing enzymes in the liver, including CYP1A2, CYP2E1, and CYP3A4, contribute to liver toxicity. RIDD degrades the mRNAs that encode *CYP1A2* and *CYP2E1* genes, reducing JNK phosphorylation and protecting hepatocytes from drug-induced toxicity (Hur *et al.*, 2012).

### **1.3 IRE1 $\alpha$ – XBPI/RIDD in cancer**

Since the late 1990s, the UPR has been shown to be involved in tumorigenesis (Ma and Hendershot, 2004). Conditions including hypoxia, nutrient deprivation, and oxidative stress have been shown to activate the UPR in various tumors (Vandewynckel *et al.*, 2013). Utilizing the IRE1-XBPI axis, tumors sustain their growth driving further oncogenic transformation and progress. Sequencing cancer genomes across different human cancers identified somatic mutations in IRE1 as likely drivers in tumor development (Greenman *et al.*, 2007). Supporting these findings were other independent studies that reported somatic alterations in IRE1 in hepatocellular carcinoma (Guichard *et al.*, 2012) and glioblastoma multiforme (GBM) (Parsons *et al.*, 2008). High levels of

---

XBPs have been reported in TNBC (Chen *et al.*, 2014, Logue *et al.*, 2018) and are associated with poor patient survival (Davies *et al.*, 2008, Chen *et al.*, 2014). XBP1 has been shown to drive TNBC tumorigenicity by forming a complex with hypoxia-inducible factor (HIF)1 $\alpha$ , that maintains the stem cell CD44<sup>high</sup>CD24<sup>low</sup> populations (Chen *et al.*, 2014). Depleting *XBPI* inhibited tumor growth and relapse and reduced the CD44<sup>high</sup>CD24<sup>low</sup> populations (Chen *et al.*, 2014). Constitutive splicing of *XBPI* modulates the tumor secretome through the production of pro-tumorigenic factors including IL-6, IL-8, chemokine (C-X-C) ligand 1 (CXCL1), granulocyte-macrophage-colony-stimulating-factor (GM-CSF), and transforming growth factor  $\beta$  2 (TGF $\beta$ 2) in TNBC cells (Logue *et al.*, 2018). Inhibiting the IRE1 $\alpha$  RNase activity modulated the tumor cell secretome and improved response to chemotherapeutic drugs in TNBC cells (Logue *et al.*, 2018). IRE1 $\alpha$ -XBP1 has also been shown to be involved in the development of multiple myeloma (Carrasco *et al.*, 2007, Chapman *et al.*, 2011), a type of cancer that arises due to the uncontrolled proliferation of malignant plasma cells. XBP1 is required for the differentiation of B cells, into antibody-secreting plasma cells (Reimold *et al.*, 2001) and importantly, *RtcB* is required for the induction of XBPs during this differentiation process (Jurkin *et al.*, 2014a). Depleting *RtcB* *in vivo* in B cells abolished XBPs at the protein level and significantly reduced the *XBPIs* and total *XBPI* mRNAs. Activation of XBPs has also been reported in melanoma (Chen and Zhang, 2017) and squamous cell carcinoma (Sun *et al.*, 2018). Overexpression of XBPs has also been linked to poor prognosis in pulmonary adenocarcinoma (Kwon *et al.*, 2018). These findings indicate that XBPs confers pro-survival advantages that lead to further growth and progress of different tumor types.

In contrast, compared to XBPs, RIDD has very recently been implicated to play an antagonistic role on tumor outcomes. In GBM, while XBPs was shown to promote tumor aggressiveness, RIDD was shown to play an anti-tumoral role (Lhomond *et al.*, 2018). By studying RIDD signatures, from a group of 1141 mRNAs susceptible to be cleaved by IRE1, 37 GBM-specific potential RIDD targets were identified which were then clustered into RIDD-positive and RIDD-negative tumor populations. The role of miRNA-17, one of the RIDD targets was evaluated towards its involvement in tumor progression with respect to expression of monocyte (IBA1), angiogenesis (CD31), and migration/invasion (RHOA). Compared to XBPs which promoted tumor aggressiveness through tumor immune infiltration and angiogenesis, RIDD imparted

---

anti-tumor effects by selectively reducing tumor angiogenesis and invasion. Low miRNA-17 levels in tumors correlated with better survival compared to patients with high miRNA-17 tumors. Hierarchical clustering using TCGA dataset (miRNAs and RNAseq) led to four major groups such as  $\text{XBP1s}^{\text{high}}/\text{RIDD}^{\text{low}}$ ,  $\text{XBP1s}^{\text{low}}/\text{RIDD}^{\text{high}}$ ,  $\text{XBP1s}^{\text{low}}/\text{RIDD}^{\text{low}}$ , and  $\text{XBP1s}^{\text{high}}/\text{RIDD}^{\text{high}}$ . Patient survival analysis revealed a statistically longer survival in the  $\text{XBP1s}^{\text{low}}/\text{RIDD}^{\text{high}}$  population compared to the  $\text{XBP1s}^{\text{high}}/\text{RIDD}^{\text{low}}$  population.  $\text{XBP1s}^{\text{low}}/\text{RIDD}^{\text{high}}$  exhibited low levels of markers for monocyte (IBA1), angiogenesis (CD31) and migration/invasion (RHOA) compared to the  $\text{XBP1s}^{\text{high}}/\text{RIDD}^{\text{low}}$  population. Furthermore,  $\text{XBP1s}^{\text{low}}/\text{RIDD}^{\text{high}}$  tumors exhibited a pro-neural and neural subtype compared to  $\text{XBP1s}^{\text{high}}/\text{RIDD}^{\text{low}}$  tumors which exhibited mesenchymal features. A similar clustering method applied to 12 primary GBM lines clustered them into the similar four groups as discussed above. Upon injection in mice, these 12 lines resulted in tumors where the  $\text{XBP1s}^{\text{low}}/\text{RIDD}^{\text{high}}$  yielded very small tumors compared to the  $\text{XBP1s}^{\text{high}}/\text{RIDD}^{\text{low}}$  that yielded the most aggressive tumors, thereby confirming the results observed from the TCGA cohort.

miRNA-34a, another IRE1 miRNA target has been shown to suppress pro-neural subtype GBM (Genovese *et al.*, 2012). GBM patients with low miR-34a expression exhibited an improved overall survival compared to miR-34a high-expressing pro-neural patients. Similarly, in breast cancer, high miR-34a expression correlated with poor prognosis factors including positive nodal status, high tumor grade, estrogen receptor negativity, HER2 positivity, and high proliferation rate (Peurala *et al.*, 2011). Under pathologic conditions, in adult neovascularization such as tumor neovascularization, neovascular diseases of the retina, VEGF has been reported to be the best-characterized pro-angiogenic factor. While transcription factors from all three arms of the UPR bind to VEGF driving its transcription (Ghosh *et al.*, 2010), RIDD has been shown to impede revascularization of the ischemic retina by degrading the pro-angiogenic neuroimmune guidance cue, netrin-1 that modulates the expression of VEGF and other proangiogenic factors such as Angiopoietin 2 in macrophage/microglial cells (Binet *et al.*, 2013). Taken together, inhibition of the IRE1 $\alpha$  RNase (Logue *et al.*, 2018) or selective inhibition of the *XBP1* mRNA ligase RtcB may have significant therapeutic benefits in treating different cancers. More importantly, uncoupling IRE1's dual RNase signaling, by an RtcB targeted drug serves as an attractive target for drug discovery for the treatment of cancers.

---

## 1.4 Breast cancer

Global breast cancer incidence and mortality are on the rise (Bray *et al.*, 2018). Due to its diverse and heterogeneous nature with respect to clinical, morphological and molecular attributes, standard classifications are used to guide patient management. These classifications are continuously updated to include new aspects that reflect present-day knowledge, thus allowing greater stratification for improved clinical outcomes.

### 1.4.1 Histological classification

Clinically, treatment is primarily governed based on histological classifications which broadly classifies breast cancer into *in-situ* carcinoma and invasive (infiltrating) carcinoma (Malhotra *et al.*, 2010). While *in-situ* carcinoma is further classified into ductal or lobular subtypes, invasive carcinoma is differentiated into infiltrating ductal, invasive lobular, ductal/lobular, mucinous, tubular, medullary and papillary subtypes (Malhotra *et al.*, 2010). However, whether these classifications include discrete molecular entities, remains to be determined.

### 1.4.2 Molecular classification

In the year 2000, Perou and colleagues (Perou *et al.*, 2000) first coined the term ‘molecular classification’ of breast cancers based on gene expression which categorized tumors into luminal-like (reflecting estrogen receptor positive (ER<sup>+</sup>), the human epidermal growth factor receptor 2 (HER2; also called Erb-B2 or Neu) positive (HER2/Erb-B2/Neu<sup>+</sup>), basal-like and the normal breast-like (reflecting more normal tissue components and low tumor cell content) subtypes. Additionally, the luminal subtype is further classified into luminal A, B, and C subtypes (Sorlie *et al.*, 2001, Cancer Genome Atlas, 2012, Dvorkin-Gheva and Hassell, 2014). Current therapy decision-making is increasingly governed based on these molecular classifications eventually evolving into three basic therapeutic groups; ER<sup>+</sup>, HER2<sup>+</sup> and the triple-negative (ER<sup>-</sup>, PR<sup>-</sup> and HER2 negative) (Cancer Genome Atlas, 2012).

### 1.4.3 Triple negative breast cancer

TNBC, unlike the ER<sup>+</sup> and HER2<sup>+</sup> tumors, has very limited treatment options due to the lack of expression of estrogen, progesterone and HER2 receptors that mostly fuel breast cancer growth (Foulkes *et al.*, 2010). TNBC patients have a relatively poor outcome and

---

cannot be treated with hormone therapy (e.g. tamoxifen, fulvestrant, aromatase inhibitors) or targeted therapies such as trastuzumab, pertuzumab, lapatinib, which target the HER2<sup>+</sup> tumors. Chemotherapy and radiotherapy, therefore, continue to be the mainstream therapy for treating TNBC. Well known for its aggressive behavior, TNBC encompasses other molecular subtypes such as claudin-low (enriched with properties similar to stem cells and epithelial-to-mesenchymal transitions (EMTs)), and the normal-breast-like subgroup (Foulkes *et al.*, 2010). Importantly, EMTs that are often associated with increased aggressiveness along with invasive, and metastatic characteristics, have been shown to generate tumor cells with stem cell properties (Devarajan *et al.*, 2012). Closely related to TNBC is the basal-like breast cancer subtype (Foulkes *et al.*, 2010). An estimated 15% of all invasive breast cancers comprise of TNBC and the basal-like subtype (Foulkes *et al.*, 2010). TNBC has also been associated with the mutation status of the *BRCA1* tumor suppressor gene. It has been reported that more than 75% of breast tumors carrying *BRCA1* mutation exhibit either a TNBC or basal-like phenotype or both (Rakha *et al.*, 2008, Reis-Filho and Tutt, 2008). These TNBC and basal-like tumors exhibit a profile of cell surface markers that are similar to the breast cancer stem cells (BCSCs), characterized by the CD44<sup>high</sup>/CD24<sup>low</sup> phenotype (Honeth *et al.*, 2008). Gene-expression patterns of the basal-like subtype are often consistent with those exhibiting properties of EMT (Sarrio *et al.*, 2008).

Strikingly, chemotherapy has been shown to enrich BCSCs (Yu *et al.*, 2007, Fillmore and Kuperwasser, 2008). Chemotherapy induces HIF-dependent expression of the glutathione S-transferase omega 1 (GSTO1) in TNBCs. GSTO1 then interacts with the ryanodine receptor embedded in the ER, resulting in increased Ca<sup>2+</sup> release from the ER. This, in turn, triggers a signaling pathway that increases the expression of pluripotency factors and enrichment of the BCSCs (Lu *et al.*, 2017). XBP1 has also been shown to directly interact with HIF1 $\alpha$  that maintains the BCSC population CD44<sup>high</sup>CD24<sup>low</sup> (Chen *et al.*, 2014). Depleting *XBPI* inhibited tumor growth, relapse and reduced the BCSC population CD44<sup>high</sup>CD24<sup>low</sup> in TNBC tumors (Chen *et al.*, 2014). Recently, the poly ADP-ribose polymerase (PARP) inhibitor olaparib was approved by the U.S. Food and Drug Administration for the treatment of breast tumors carrying BRCA mutations (Robson *et al.*, 2017, Stark, 2018, Robson *et al.*, 2019). Although not exclusively meant for TNBC treatment, this drug provides a new treatment option for those ~15% of patients with TNBC and a BRCA mutation. Olaparib, an alternative to chemotherapy,



---

inhibits an enzyme involved in cell recovery from DNA damage. Inhibiting the enzyme leads to damaged BRCA genes less likely to be repaired, leading to cell death and reduced tumor growth (Stark, 2018).

With no prominent targeted therapy yet available exclusively for TNBCs, this thesis, therefore, makes an effort towards developing a targeted approach via the *hRtcB*-XBP1s signaling of the UPR for improved treatment of TNBCs.

---

## 1.5 Research hypothesis and objectives

TNBCs are inherently aggressive in nature. Therefore, targeting RtcB in TNBC was investigated as high levels of XBP1 activity have been reported in primary TNBC patient samples compared to ER<sup>+</sup>/PR<sup>+</sup> patients (Chen *et al.*, 2014). Survival analysis from a cohort of 193 TNBC patient data set demonstrated that tumors with an elevated XBP1 signature displayed shorter relapse-free survival (Chen *et al.*, 2014). Silencing *XBPI* in a patient-derived TNBC xenograft mouse model significantly reduced tumor incidence (Chen *et al.*, 2014). Isolating CD44<sup>high</sup>CD24<sup>low</sup> cells from TNBC patient samples displayed high levels of XBP1s and that silencing *XBPI* impaired their mammosphere-forming ability (Chen *et al.*, 2014). Additionally, highest levels of XBP1s expression were reported in MDA-MB-231 TNBC cells both at the transcript and protein level compared to cells from other breast cancer subtypes (Logue *et al.*, 2018). Therefore, this thesis utilizes MDA-MB-231 cells as an experimental model for developing drugs against RtcB.

### 1.5.1 Hypothesis

Inhibition of *hRtcB* by small molecules modulates expression of XBP1s of the IRE1 arm of the UPR in TNBC.

### 1.5.2 Objectives

#### Objective 1

To establish a 3-dimension (3D) predicted structure of the *hRtcB* protein for lead discovery and optimization. Obtaining experimental structures from X-ray crystallography or NMR spectroscopy are often difficult and time-consuming often requiring a very high level of expertise in those techniques. Also, through email communication with Javier Martinez group who identified the *hRtcB* complex, purification of recombinant *hRtcB* protein in their laboratory resulted in an inactive enzyme suggesting the presence of the other five subunits of the complex might serve as a prerequisite for RtcB to remain active. Therefore, modeling 3D predicted structures can provide useful structural information for generating hypotheses about a protein's function and its druggability, directing further experimental work. As the

---

crystal structure of *hRtcB* is yet to be resolved, this 3D predicted structure would assist in analyzing the probable ligand binding site/active site of the *hRtcB* protein for *in silico* drug discovery purposes.

## **Objective 2**

To identify small molecules that target *hRtcB* in TNBC. Targeting *hRtcB* with small molecules should modulate the expression of XBP1s that could reduce XBP1s-mediated tumor aggressiveness in TNBCs.

---

## CHAPTER 2: Materials and Methods

### 2.1 Cell culture of adherent cell lines

MCF10A (ATCC) normal breast epithelial cells were maintained in culture with Dulbecco's modified Eagle's medium (DMEM)/Nutrient Mixture F-12 Ham from Sigma (Catalog No: D6421) supplemented with 5% horse serum, 20 (ng)/ml epidermal growth factor, 0.5 (mg)/ml hydrocortisone, 100 ng/ml cholera toxin and 10 (µg)/ml insulin. MDA-MB-231 (ATCC) triple-negative breast cancer cells were maintained in culture with DMEM high glucose (4500 mg/L) from Sigma (Catalog No: D6429) supplemented with 10% heat inactivated fetal bovine serum (FBS) and 2 (mM) L-glutamine. All cells were cultured at 37 °C in a static humidified 5% CO<sub>2</sub> incubator.

### 2.2 Drug treatment

The pharmacological ER stress inducers Tunicamycin (Tm) and Thapsigargin (Tg) were used over the course of the study.

#### 2.2.1 Tunicamycin

Tm from *Streptomyces* sp. was used to induce ER stress in MCF10A and MDA-MB-231 cells at the indicated concentrations and up till time points mentioned in relevant chapters of the thesis. It is a mixture of homologous antibiotics that inhibits N-acetylglucosamine transferases, hence preventing the first step of N-linked glycosylation of newly synthesized proteins in the ER (Bassik and Kampmann, 2011). This leads to accumulation of misfolded proteins in the ER, activating the UPR. There are at least 10 homologs of this drug, combinations of which are used in commercially available drugs. Tm from Sigma-Aldrich (Catalog No: T7765) was re-suspended to a stock concentration of 2 mg/ml in DMSO.

#### 2.2.2 Thapsigargin

Tg from *Thapsia garganica* was also used to induce ER stress in MCF10A and MDA-MB-231 cells at the indicated concentrations and up till time points mentioned in relevant chapters of the thesis. It is an irreversible non-competitive inhibitor of the

---

ubiquitous sarco-endoplasmic reticulum calcium-ATPase (SERCA) family of calcium pumps in mammalian cells which inhibits  $\text{Ca}^{2+}$  uptake into the ER resulting in depletion of  $\text{Ca}^{2+}$  pool within the ER (Lytton *et al.*, 1991). Tg from Sigma-Aldrich (Catalog No: T9033) was re-suspended to a stock concentration of 4 mM in DMSO.

### **2.2.3 MKC8866**

MKC8866 was a kind gift from Dr. John Patterson, MannKind Corp., Valencia, CA, was used at a dose of 20 ( $\mu\text{M}$ ) and up till time points mentioned in relevant chapters of the thesis. The inhibitor specifically inhibits the IRE1 RNase domain and blocks *XBPI* mRNA splicing (Volkman *et al.*, 2011).

### **2.2.4 Paclitaxel**

Paclitaxel from *Taxus brevifolia*, a standard chemotherapy agent in the treatment of breast cancers was used in the study at a dose of 10 nM and up till time points mentioned in ‘Results’ section of the thesis. This drug binds to the N-terminus of  $\beta$ -tubulin and stabilizes microtubules arresting the cell cycle at the G2/M phase. This microtubule damage induces apoptosis. Paclitaxel from Sigma-Aldrich (Catalog No: T7402) was re-suspended to a stock concentration of 1 mM in DMSO.

## **2.3 Immunofluorescence and microscopy**

Developed by Albert Coons (Coons *et al.*, 1941, Coons, 1942, Coons and Kaplan, 1950, Coons, 1961) immunofluorescence staining remains a cornerstone of cell biology. MCF10A and MDA-MB-231 cells were seeded at a density of 70,000 cells/well and grown on glass cover slips in 500  $\mu\text{l}$  of DMEM/Nutrient Mixture F-12 Ham and DMEM high glucose respectively in Greiner CELLSTAR® 24 Well Cell Culture Plate for Adherent Cultures, 1.9  $\text{cm}^2$ , TC, Sterile, with Lid (Catalog No: 662 160). After treatment, cells were rinsed off twice in 1 ml 1X phosphate buffered saline (PBS) pH 7.4 and fixed in 4% formaldehyde from Sigma-Aldrich (F8775) for 20 minutes at room temperature (RT). Fixed cells were washed thrice in 1 ml 1% BSA in PBS for 5 minutes each and permeabilized with 0.1% Triton X-100 for 10 minutes at 4 °C. Cells were then gently rinsed thrice in 1 ml PBS. Coverslips were blocked in 1 ml 5% BSA in PBS for 1 hour at RT and incubated in primary antibody diluted in 1

---

ml 1% BSA in PBS overnight at 4 °C. Next day, cells were washed thrice in 1ml PBS for 5 minutes each. Cells were then incubated in secondary antibody diluted in 1 ml 1% BSA in PBS for 1 hour at RT in the dark. Cells were again washed thrice in 1ml PBS for 5 minutes each in the dark. Coverslips were then mounted in Fluoroshield™ with DAPI mounting medium from Sigma-Aldrich (Catalog No: F6057). Coverslips were sealed with nail polish to prevent drying and movement under microscope and stored in dark at 4 °C. Primary and secondary antibodies used were rabbit RtcB-specific (1:200; Proteintech Europe) and FITC-conjugated anti-rabbit (1:200) respectively. Fixed samples were imaged on a DeltaVision microscope. Images were edited for brightness and contrast using ImageJ (Abràmoff, 2004).

## **2.4 siRNA preparation**

Posttranscriptional gene silencing was first discovered in plants by David Baulcombe's group (Hamilton and Baulcombe, 1999) and later shown to induce RNAi in mammalian cells (Elbashir *et al.*, 2001). Dharmacon siRNAs, ON-TARGETplus Human RtcB (*hRtcB*) (51493) siRNA – SMARTpool (Catalog No: L-017647-00-0005), 5 (nM) and ON-TARGETplus Non-targeting control siRNA (Catalog No: D-001810-01-20; Lot No: 1628597), 20 nM were obtained from Fisher Scientific Ireland. 51493 is the NCBI gene ID. The RtcB siRNA was re-suspended in 1X siRNA buffer prepared in the laboratory according to “Basic siRNA Resuspension” protocol provided by Dharmacon to a stock concentration of 20 µM and stored as smaller volume aliquots at -80 °C. However, following resuspension, the final RtcB siRNA concentration measured by NanoDrop spectrophotometer was 18 µM. This minor difference was probably due to the reason that salts present in 1X buffer (or other resuspension solution) are known to cause a decrease in the absorbance reading of RNA as stated in the protocol.

### **2.4.1 siRNA transfections**

MDA-MB-231 cells were transfected with both *hRtcB* and Non-targeting control siRNAs using chemical-based transfection method (Felgner *et al.*, 1987, Felgner *et al.*, 1994). At 24 hours prior transfection, cells at a density of 70,000 cells/well were seeded in 1 (ml) of DMEM in Greiner CELLSTAR® 12 Well Cell Culture Plate for Adherent Cultures, 3.9 cm<sup>2</sup>, TC, Sterile, with Lid (Catalog No: 665 180). Similarly,

---

cells at a density of 168,000 cells/well were seeded in 2 ml of DMEM in Greiner CELLSTAR® 6 Well Cell Culture Plate for Adherent Cultures, 9.6 cm<sup>2</sup>, TC, Sterile, with Lid (Catalog No: 657 160). The cells were then incubated overnight at 37 °C. After 24 hours, cells on reaching 60-70% confluency, the medium in the wells was replaced with fresh DMEM before starting with experiments. Desired siRNA concentrations were used for respective individual experiments which as mentioned in relevant chapters of the thesis. Both *hRtcB* and control siRNAs were diluted in serum-free medium to reach the desired concentrations for the different experiments and incubated for 5 minutes at RT. Transfections were carried out using DharmaFECT 4 transfection reagent (Catalog No: T-2004-02; Lot No: 140801T) obtained from Fisher Scientific Ireland, which was again diluted in empty medium to reach the desired concentrations for the experiments and incubated for 5 minutes at RT. Each of the *hRtcB* and control siRNAs and the DharmaFECT 4 reagent were diluted separately in empty medium to make up a final volume of 100 µl. After 5 minutes of incubation, siRNAs for *hRtcB* and control were added to the diluted DharmaFECT and incubated for 20 minutes at RT. The siRNA-DharmaFECT complex was then added dropwise to the wells of the transfection plate. The final volume in each of the wells was made up to 1000 µl for 12 well and 2000 µl for 6 well plate. The plate was gently shaken before leaving it for overnight incubation at 37 °C. For all experiments, transfection medium was replaced after 48 hours post-transfection. While transfection is likely to be completed within the 6 hours timeframe, gene silencing detection was not carried out prior to 48 hours post-transfection. Moreover, as per the protocol, use of antibiotics were avoided, as during transfection cells are very sensitive in terms of permeability and the presence of antibiotics could greatly increase cell death due to uptake of antibiotic along with the siRNA-transfection reagent complex. In agreement to the manufacturer recommendation, all the transfections were carried out completely in antibiotic-free medium.

## 2.5 RNA extractions

A method of total RNA isolation by a single extraction with an acid guanidinium thiocyanate-phenol-chloroform mixture was described by Chomczynski and colleagues (Chomczynski and Sacchi, 1987). MDA-MB-231 cells at a density of 168,000 cells/well were seeded in 6 well plates. After experimental treatments

---

mentioned in the 'Results' section of the thesis, plates were removed from incubator, kept on ice and cells were scraped into the media and transferred to 2 ml Eppendorf tubes. Samples were centrifuged at 1699 g for 4 minutes at 4 °C to pellet the cells. Cells were resuspended and lysed in 1 ml TRI reagent from Invitrogen (Catalog No: #9738) and incubated for 5 minutes at RT to allow full dissociation of any nucleoprotein complexes. Next, 200 µl of chloroform was added to each Eppendorf tubes and shaken vigorously for 15 secs. Samples were incubated for 15 minutes at RT before centrifugation at top speed (20,817 g) for 15 minutes at 4 °C. Following centrifugation, the sample separates into 3 phases, an upper colorless aqueous phase, a white interphase and a lower red phenol-chloroform organic phase. The aqueous phase contains RNA whereas the organic and interphase contain a mixture of proteins and DNA. Next, 400 µl of the upper aqueous phase was carefully pipetted out and transferred into a sterile Eppendorf tube. 1 volume (400 µl) of cold isopropanol was then added dropwise on top of the aqueous phase and quickly inverted 6 times followed by incubation for 10 minutes at RT or overnight at -20 °C. Next, samples were centrifuged at top speed (20,817 g) for 15 minutes at 4 °C. Supernatant was removed and the RNA pellet was washed in 1 ml of 85% ethanol plus 15% Diethyl pyrocarbonate (DEPC) treated water. Samples were centrifuged at top speed (20,817 g) for 15 minutes at 4 °C. Supernatant was removed and RNA pellet was left to air dry for approximately 25 minutes at RT. RNA pellets were resuspended in 15 µl of DEPC treated water and heated for 15 minutes at 65 °C followed by pulse centrifuge twice at an interval of few seconds at 4000 g. RNA concentration was determined by measuring absorbance at 260 nm (A<sub>260</sub> nm) using NanoDrop spectrophotometer. RNA purity was assessed using A<sub>260</sub> nm/A<sub>280</sub> nm ratio, ensuring the ratio of nucleic acid/protein was between 1.7 and 2.1. RNA was stored in the -80 °C freezer.

## **2.6 Reverse transcription polymerase chain reaction (cDNA synthesis)**

The independent and simultaneous discovery of reverse transcriptase in RNA tumor viruses (retroviruses) by Howard Temin and David Baltimore in 1970 revolutionized molecular biology (Temin and Mizutani, 1970, Baltimore, 1970). To monitor gene expression, the isolated RNA was reverse transcribed into cDNA by reverse transcription. For this, 1 µg of RNA was transferred to a PCR tube and made up to 10



---

$\mu$ l with sterile water from Sigma-Aldrich (Catalog No: W4502). For real time polymerase chain reactions, amount of RNA needed to make cDNA depends on the abundance of the gene. Next, 1  $\mu$ l DNase-I and 1  $\mu$ l of 10X DNase buffer was added to allow degradation of any DNA contamination and pulse centrifuged at 4000 g twice at an interval of few seconds. The DNase was inactivated by adding 1  $\mu$ l of 25 mM EDTA and incubating for 8 minutes at 65 °C. 1  $\mu$ l of Oligo dT (cDNA primer) was added to the reaction mixture and incubated for 2 minutes (denaturation) at 65 °C followed by another incubation for 2 minutes at 42 °C to allow the primer to anneal to the poly(A) tail of the mRNA. To the reaction mixture, 10  $\mu$ l of master mix was added which contained 2.6  $\mu$ l sterile water, 4  $\mu$ l of 5X 1st strand buffer, 2  $\mu$ l of 0.1 M DTT (dithiothreitol), 1  $\mu$ l of 10 mM dNTPs and 0.4  $\mu$ l superscript II (reverse transcriptase enzyme). Mixture was incubated at 42 °C for 50 minutes for cDNA synthesis followed by 75 °C for 10 min to inactivate superscript reverse transcriptase. Reagents were purchased from Invitrogen.

## **2.7 Conventional polymerase chain reaction (PCR)**

For conventional-PCR (Saiki *et al.*, 1985), a reaction mixture was made, which contained 2  $\mu$ l of cDNA product, 3  $\mu$ l of a forward and 3  $\mu$ l of a reverse primer (Table 2.1), 4.5  $\mu$ l sterile water (Sigma-Aldrich) and 12.5  $\mu$ l PCR master mix (GoTaq Polymerase enzyme). The reactions were placed in a thermocycler and cDNA was amplified under the following conditions (Table 2.1). The reaction mixtures underwent 25-40 cycles, depending on the gene abundance. Primers were designed and purchased from Integrated DNA Technologies (IDT).

**Table 2.1|** PCR primers and cycle conditions

Species	Gene	Forward primer	Reverse primer
Human	XBP1	5'-GGAACAGCAAGTGGTAGA-3'	5'-CTGGAGGGGTGACAACTG-3'
Human	GAPDH	5'-ACCACAGTCCATGCCATC-3'	5'-TCCACCACCCTGTTGCTG-3'

**PCR cycle conditions for XBP1**

Steps	Temperature (°C)	Time	Number of cycles
Initial Denaturation	94	5 min	40
Denaturation	94	1 min	
Primer Annealing	55	30 sec	
Extension	72	30 sec	
Final Extension	72	5 min	
Finish	4	∞	

**PCR cycle conditions for GAPDH**

Steps	Temperature (°C)	Time	Number of cycles
Initial Denaturation	94	3 min	25
Denaturation	94	30 sec	
Primer Annealing	55	30 sec	
Extension	72	30 sec	
Final Extension	72	7 min	
Finish	4	∞	

## 2.8 Real time PCR

Real time PCR assays for RtcB, XBP1s, EDEM1, GRP78 and MRPL19 were purchased from IDT (Table 2.2). Probe based detection chemistry (Taqman®) was utilized as it avoids non-specific amplification during PCR (Heid *et al.*, 1996, Bustin, 2000). These probes are fluorescently labelled DNA oligonucleotides. The probe has a fluorescent reporter dye attached to its 5' end and a quencher dye at its 3' end. The quencher dye effectively quenches the output from the reporter dye. Therefore, when

the reporter and quencher are physically close to one another the overall level of fluorescent output is low. As DNA polymerase amplifies the target it begins to cleave the probe separating the fluorescent dye from the quencher allowing release of a fluorescent signal detected by the real time PCR machine. cDNA was diluted 1:5 ratio in sterile water (Sigma-Aldrich). The dilution mixture was vortexed for 10 seconds and pulse centrifuged at 1.5 g twice at an interval of few seconds. Next, the assay probes (0.5 µl/well) were diluted in sterile water (2.5 µl/well) and the supplied 5X master mix enzyme (2 µl/well; Takyon™ Rox Probe MasterMix dTTP (Eurogentec)). 5 µl of cDNA was pipetted into each well of the MicroAmp™ Fast Optical 96-Well Reaction Plate with Barcode, 0.1 ml from Applied Biosystems (Catalog No: 4346906). To this, 5 µl of the master mix containing the probe of interest was added to each well. The plate was sealed and centrifuged at 1000 g for 1 minute twice at an interval of few seconds to mix the reaction. The plate was placed into the thermocycler and subjected to 40 cycles of PCR (Hold at 95 °C for 3 mins, then 40 cycles at 95 °C for 12 secs, and 60 °C for 30 secs). Relative gene expression was calculated using the  $\Delta\Delta CT$  method (Livak and Schmittgen, 2001) and MRPL19 was used as the house keeping gene to normalize gene expression.

**Table 2.2|** Real time PCR primers and probes

Species	Gene	Primer 1	Primer 2	Probe
Human	RtcB	5'- GGCTGGACATC A CTTTCATCT-3'	5'-TGAATGACCC TGAAGCAGTAG -3'	5'-/56- FAM/TGTCAAAC C/ZEN/CGACACCAC CTGG/ 3IABkFQ/-3'
Human	XBP1s	5'- GGAATGAAGTG AGGCCAGT-3'	5'- AGAGTCAATA CCGCCAGAATC- 3'	5'-/56- FAM/TGAGTCCG C/ZEN/AGCAGGTGC A /3IABkFQ/-3'
Human	MRPL19	5'- CTTAGGAATGT TATCGAAGGAC A AG-3'	5'-GCTATATTCA GGAAGGGCATC T-3'	5'-/56- FAM/CTCGGGTC C/ZEN/AGGAGATTC AG GTG/3IABkFQ/-3'

---

## 2.9 Protein expression analysis

### 2.9.1 Protein sample preparation

Cells were seeded in 12 or 6 well plates for protein sample preparation. Following treatment, plates were kept on ice. Cells were scraped and collected in appropriately sized Eppendorf tubes on ice. Eppendorf tubes were centrifuged at 20,000 g for 20 seconds at 4 °C and the supernatant was removed from cell pellet. 150 µl of ice cold 1 X PBS was added to the pellet, resuspended and centrifuged again at 20,000 g for 20 seconds at 4 °C. Supernatant was discarded. All samples were lysed together in appropriate volume of 1 X Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) lysis buffer (Table 2.3) at  $2.5 \times 10^6$  cells/ml. Samples were left on ice for 5 minutes. Next, protein samples were boiled for 5 minutes at 100 °C on heat block. Samples were then placed on ice for 2 minutes, pulse centrifuged at top speed and either stored at -20 °C or loaded directly onto an SDS-PAGE gel.

**Table 2.3|** SDS PAGE lysis buffer (2X). Prior to use, buffer diluted 1:2 and β-mercaptoethanol added. 1 ml distilled H<sub>2</sub>O plus 950 µl 2X SDS PAGE Buffer plus 50 µl β-mercaptoethanol.

Reagents	2X SDS-PAGE lysis buffer
10% SDS	20 ml
1 M Tris-HCL pH 6.8	5 ml
0.1% Bromophenol Blue	0.05g
100% Glycerol	10 ml
dH <sub>2</sub> O	15 ml
Total Volume	50 ml

### 2.9.2 SDS-PAGE

Protein separation was performed by loading samples on to an SDS-PAGE (Shapiro *et al.*, 1967, Pitt-Rivers and Impiombato, 1968, Weber and Osborn, 1969) alongside a protein molecular weight marker. The constituents and volumes used to make these gels are described (Table 2.4). The SDS-gels were placed in a tank containing running buffer and electrophoresis was carried out at 60 volt (V) for 1 hour followed by 90 V until the samples reach the bottom of the gel.

**Table 2.4| SDS-PAGE**

<b>Reagents</b>	<b>8% Resolving Gel (ml)</b>	<b>10% Resolving Gel (ml)</b>	<b>2% Stacking Gel (ml)</b>
dH <sub>2</sub> O	4.6	4	1.36
30% Acrylamide mix	2.6	3.4	0.34
1.5 M Tris pH 8.8	2.6	2.6	0.26
10% SDS	0.1	0.1	0.02
10% Ammonium Persulfate (APS)	0.1	0.1	0.02
TEMED	0.006	0.004	0.002
Total Volume	10	10	2

### **2.9.3 Western blotting**

Following protein separation, the SDS-PAGE gel was placed in a sandwich with a sponge, filter paper, nitrocellulose membrane, gel, filter paper, and a sponge (Towbin *et al.*, 1979, Burnette, 1981). The proteins were electrophoretically transferred onto the nitrocellulose membrane for 90 minutes at 110 V in transfer buffer (10 X CAPs at pH 11, 20% methanol and water). Following transfer, the nitrocellulose membrane was then blocked with 5% non-fat milk in PBS containing 0.1% Tween for 1 hour and probed with an appropriate primary antibody at optimized concentrations (Table 2.5). Next, the membrane was washed 3 times for 10 minutes each with PBS Tween 0.1% and further incubated in an appropriate horseradish peroxidase-conjugated secondary antibody for 1 hour at RT (Table 2.5). Following incubation, the membrane was washed 3 times for 10 minutes each with PBS Tween 0.1%. Antibodies were visualized using horseradish peroxidase substrate chemiluminescent Western Lightning ECL substrates (PerkinElmer).

**Table 2.5|** Antibodies for western blotting

Antibody	Company	Cat. No.	Mono/ Poly	Species reactivity	Raised in	Primary dilution	Primary incubation
RtcB-Specific	Proteintech Europe	19809-1-AP	Poly	H	Rabbit	1:1000 5% Milk	Overnight at 4 °C
XBP1s	BioLegend	647502	Mono	H	Mouse	1:1000 5% Milk	Overnight at 4 °C
IRE1 $\alpha$	Cell Signaling Technology	3294	Mono	H, M	Rabbit	1:1000 5% Milk	Overnight at 4 °C
PERK	Cell Signaling Technology	3192	Mono	H, M, R	Rabbit	1:1000 5% Milk	Overnight at 4 °C
Anti-rabbit HRP-2Ab	Cell Signaling Technology	7074		R	Goat	1:10,000 5% Milk	1-hour RT
Anti-mouse HRP-2Ab	Cell Signaling Technology	7076		M	Horse	1:10,000 5% Milk	1-hour RT

## 2.10 Computational methods

Homology modeling was performed using the simulation programs Yet Another Scientific Artificial Reality Application (YASARA, v15.3.8) (Krieger *et al.*, 2004), Molecular Operating Environment (MOE 2014.09; Chemical Computing Group) (2014), Iterative Threading ASSEmbly Refinement (I-TASSER) (Roy *et al.*, 2010), Protein Homology/analogY Recognition Engine V 2.0 (Phyre2) (Kelley *et al.*, 2015), and SWISS-MODEL (Biasini *et al.*, 2014). Model quality was assessed through the QMEAN Z scores, RAMPAGE, VERIFY-3D, PDBeFOLD, DoGSiteScorer, MatchMaker (UCSF CHIMERA) (Pettersen *et al.*, 2004), and CheckMyMetal servers. Molecular dynamics (MD) simulations were performed using the YASARA program, and umbrella sampling of Mn<sup>2+</sup> movement using Gromacs (Berendsen *et al.*, 1995, Abraham *et al.*, 2015).

---

### 2.10.1 Homology structure modelling

The crystal structures of *bRtcB* with its bound natural ligand GMP were retrieved from the Protein Data Bank (PDB). A common strategy for all homology modeling techniques involves identifying known structures (templates) from the PDB database. Best-matching template is chosen by comparing the sequence alignment between the query protein (one to be modeled), and the template proteins. This is followed by mapping the query sequence to the template structure, referred to as threading or fold recognition (Bowie *et al.*, 1991, Jones and Taylor, 1992). This steers the process toward 3D model construction and quality estimation (Sali and Blundell, 1993, Schwede *et al.*, 2008). Homology modeling was performed using the MOE, I-TASSER, Phyre2, SWISS-MODEL, and YASARA simulation programs, although the differences in accuracy between these tools are minor (Moult *et al.*, 2014). All of these tools build different models whose quality then needs to be evaluated using separate *in silico* structure quality estimation tools described in section 2.10.2.

MOE utilizes the segment-matching procedure (Levitt, 1992) and the approach to the modeling of insertion/deletion regions (Fechteler *et al.*, 1995). By default, MOE creates 10 models using a knowledge-based loop searching method and sidechain rotamer selection method followed by the creation of an average model.

I-TASSER is a server-based modeling approach where sequence alignments are submitted to the I-TASSER server after selecting the option “Specify template with alignment.” This option is used to specify both the template structure and the target–template sequence alignment which is in contrast to the default mode used in some of the other modeling techniques where one submits the target sequence only and allows the program to provide templates and sequence alignments.

Phyre2 is again a server-based approach and the key differentiating factor is that it provides a user-friendly interface which enables biologists inexperienced in bioinformatics to use state-of-the-art techniques without the very steep learning curve typical of many online modeling tools (Kelley *et al.*, 2015). Phyre2 combines a large number of software components whose advanced facilities include Backphyre, to search a structure against a range of genomes, batch submission of a large number of protein

---

sequences for modeling, one-to-one threading of a user sequence onto a user structure, Phyrealarm, for automatic weekly scans for proteins that are difficult to model, and Phyre Investigator for in-depth analysis of model quality (Kelley *et al.*, 2015).

SWISS-MODEL is a fully automated protein modeling server. It uses an aligned query–template sequence as input and uses the knowledge-based ProModeII program (Peitsch, 1996) for building models. Compared to other modeling techniques where energy refinement is carried out separately at the end of the modeling process, SWISS-MODEL attempts to produce a complete, minimized model using the Gromos96 force field (Van Gunsteren, 1996).

Finally the YASARA simulation program features a complete homology modeling module that fully automatically takes all the steps from an amino acid sequence to a refined high-resolution model using a critical assessment of techniques for protein structure prediction (CASP) approved protocol (Krieger *et al.*, 2009). Briefly, the target sequence was PSI-BLASTed (Altschul *et al.*, 1997) against Uniprot which was then used to search the PDB for potential modeling templates. Based on the alignment score and the structural quality (Hooft *et al.*, 1996b, Hooft *et al.*, 1996a), templates were ranked. Through the PSI-BLAST iterations, templates structures were identified, which was followed by searching the PDB for a match (i.e. hits with an E-value below the homology modeling cutoff 0.5). 3 hits (template structures: 4DWR (Englert *et al.*, 2012), 4DWQ (Englert *et al.*, 2012), and 2EPG) were obtained. Altogether, these steps ensure that good template structures are used even if the alignment score is lower. To aid alignment correction and loop modeling, a secondary structure prediction for the target sequence was obtained. For each of the three templates, models were built. Either a single model if the alignment was certain, or a number of alternative models if the alignment was ambiguous. 13 predicted models were initially generated and sorted by their overall quality Z-scores by the program. From these models, YASARA combined the best parts of the 13 models to obtain a hybrid model, to increase the accuracy.

### **2.10.2 Quality assessment of predicted structures**

The generated homology models were assessed carefully for (i) absolute structure



---

quality through QMEAN6 Z scores; (ii) stereochemical quality of the constituting amino acids using the RAMPAGE server; (iii) folding reliability through VERIFY-3D investigations; (iv) extent of secondary structure identity/similarity between predicted and template structures using PDBeFOLD; (v) active site pocket detection plus estimation of druggability through DoGSiteScorer; (vi) extent of secondary structure identity/similarity between pockets from predicted and template structures through MatchMaker (Pettersen *et al.*, 2004), and (vii) validating metal binding sites through the CheckMyMetal server.

### **2.10.3 Molecular dynamics simulations**

To obtain relaxed geometries, molecular dynamics (MD) simulations were performed using YASARA. The AM1-BCC (Jakalian *et al.*, 2002) model already implemented in YASARA was utilized to compute partial atomic charges for the ligands. Of the seven predicted structures that were built, the best fit structure for lead discovery and optimization was solvated in a periodic box of water with dimensions  $90 \times 90 \times 90$  Å<sup>3</sup>, at constant temperature (298 K). This was followed by simulated annealing minimization. Water molecules were randomly replaced by Na<sup>+</sup> or Cl<sup>-</sup> ions to provide an overall charge neutral system and to achieve a physiological salt concentration of 0.9%, followed by a final energy minimization step. The final system contained ~75,000 atoms including 21,500 water molecules. The generated structure was simulated for 10 ns in an NVE ensemble to allow for it to relax and optimize its interactions. Particle-mesh Ewald (PME) summation (Essmann *et al.*, 1995) with a cut-off of 7.86 Å was used for long-range Coulomb interactions. Multiple time steps including 1.25 fs for intramolecular forces, 2.5 fs for intermolecular forces and data collection every 100 ps marked the entire simulation process.

### **2.10.4 Umbrella sampling simulations**

In order to shed further light on the noted inactivity of the C122A mutant of hRtcB, (Popow *et al.*, 2011) the binding interaction between the enzyme and metal ion was investigated for the wild type system and its C122A mutant. Thus, the binding energy of Mn<sup>2+</sup> was derived from the potential of mean force (PMF) (Kirkwood, 1935) extracted from umbrella sampling simulations (Torrìe and Valleau, 1977, Lemkul and Bevan, 2010, Luo and Karplus, 2011, Allner *et al.*, 2012, Plazinski and Knys-

---

Dzieciuch, 2013, Yu *et al.*, 2016). Umbrella sampling is based on the idea that by running several restrained MD simulations in which each window explores a particular region along the reaction coordinate between two interacting species ( $\xi$ ), the entire conformational space is sampled. The free energy profile (Wong and York, 2012) can thereafter be obtained using the weighted histogram analysis method (WHAM) (Kumar *et al.*, 1992). First, the  $\text{Mn}^{2+}$  of *hRtcB* (both wild and mutant variants) was included in the force field parameter files, with a full +2 charge. For the van der Waals parameters of the metal, a well-depth ( $\epsilon$ ) of  $5.85760 \times 10^{-2} \text{ kJ mol}^{-1}$ , and a  $\sigma$  of  $3.01124 \times 10^{-1} \text{ \AA}$  were employed. These parameters were obtained from the work of Bradbrook (Bradbrook *et al.*, 1998) and adapted to fit the units of the Amber ff99sb-ildn force field in Gromacs. This force field contains updated backbone  $\phi$  and  $\psi$  torsions and improved side-chain torsion angles (Lindorff-Larsen *et al.*, 2010). The protein was placed in a box of sufficient dimensions to satisfy the minimum image convention and provide space to pull the metal ion along the z axis. Water molecules were added, and the counter-ion concentration modified to have a charge neutral system with a final NaCl concentration equal to a physiological solution (0.1 mol/L), using the genion algorithm of Gromacs (Berendsen *et al.*, 1995, Abraham *et al.*, 2015). In order to generate equilibrated structures as starting points for the pulling simulations, each protein (wild type and mutant) was energy minimized and equilibrated at  $T = 300 \text{ K}$  and  $P = 1 \text{ bar}$ . These structures were used as initial configurations for pulling simulations. Restraints were removed only for  $\text{Mn}^{2+}$  which was pulled away from the protein along the z axis  $>1000 \text{ ps}$ , using a spring constant of  $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  and a pull rate of  $0.1 \text{ nm ps}^{-1}$ . Snapshots were saved every 1 ps. The final center of mass (COM) distance between the protein and  $\text{Mn}^{2+}$  was ca. 10 nm. From this pulling procedure, several snapshots (spaced from 0.1 to 0.2 nm) were considered to get the initial configurations for the umbrella sampling windows. This resulted in 68 and 64 windows for wild-type and mutated systems, respectively. Each window was independently equilibrated, and 10 ns MD simulation was performed. Analysis of the results was carried out using WHAM (Kumar *et al.*, 1992).

### **2.10.5 Virtual high throughput screening**

Virtual High Throughput Screening (VHTS) of compound libraries was performed using Schrödinger Software Release (2015-4). Briefly, the VHTS is designed for

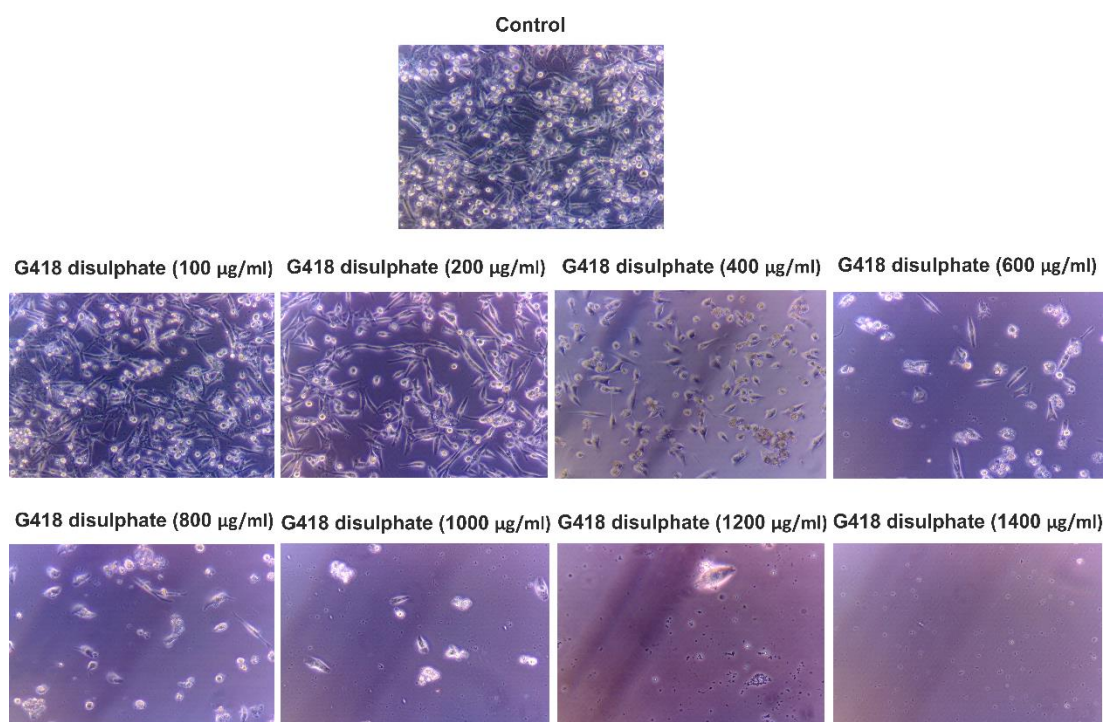
---

screening large collections of compounds against one or more targets. The workflow for the VHTS campaign included protein preparation, ligand preparation using LigPrep, receptor grid generation, and Glide docking at the three accuracy levels, HTVS, SP, and XP, in each stage retaining the top 10% binders. The ligand files for the workflow consisted of 3D structures. The three docking stages include HTVS docking at which the ligands that were retained were then passed to the next stage, which performed the SP docking and the survivors of this stage were then passed on to the third and final stage which performed the XP docking. These stages involved flexibility of ligand and target area and number of docked poses. The VHTS campaign was carried out using the national Swedish supercomputing facilities at C3SE (Gothenburg) to screen the ZINC database (Irwin *et al.*, 2012) of potential lead compounds (over 13 million compounds) and in-house databases of approximately 140,000 compounds provided by the Chemical Biology Consortium Sweden (CBCS). On completion of the VHTS campaign, through the generation of structural interaction fingerprints, similarities and diversity of the binding modes of the top 200 highest scoring binders (based on binding affinity scores/binding free energy values ( $\Delta G$ ), docking scores) were assessed. Thereafter, the top 100 highest scoring binders that exhibited similar interactions to the natural ligand GMP in the active site of the *hRtcB* homology model were selected and narrowed down to the top 20 and finally to the top 8 hit compounds, as identified based on their binding energy scores, docking scores, their chemical nature (e.g. based on visual inspection such as similar backbones but different sidechains or vice versa), synthesizability or availability. These top 8 hit compounds were considered to be further tested through cell-based assays. All hit compounds were either purchased or synthesized from MolPort ([www.molport.com](http://www.molport.com)).

### **2.11 Antibiotic kill curve**

Determining the optimal antibiotic concentration is the first step for generation of a stable cell line (MacGowan *et al.*, 1996). MDA-MB-231 cells at a density of 30,000 cells/well were seeded in 2 ml of DMEM in Greiner CELLSTAR® 6 Well Cell Culture Plate for Adherent Cultures, 9.6 cm<sup>2</sup>, TC, Sterile, with Lid (Catalog No: 657 160). After 24 hours, cells were subjected to increasing doses (100, 200, 400, 600, 800, 1200 and 1400) µg/ml of G418 disulfate to determine the minimum antibiotic concentration needed to kill all the cells over a course of 10 days (Figure 2.1). Every 2 days, old

medium was replaced with fresh DMEM containing the G418 doses and incubated at 37 °C. The optimal antibiotic concentration is cell type dependent. Performing a kill curve is recommended with each new cell type or when a new selection antibiotic or different lot of selection antibiotic is used.



**Figure 2.1| Antibiotic kill curve.** MDA-MB-231 cells were treated with increasing doses of G418 disulphate over a course of 10 days.

## 2.12 Plasmid DNA preparation

The term plasmid was first coined in 1952 by the American molecular biologist Joshua Lederberg to describe 'any extrachromosomal hereditary element' (Lederberg, 1952). The XBP1-Luciferase/pcDNA3 plasmid was received as a kind gift from Prof. Etsu Tashiro, Department of Biosciences and Informatics, Keio University, Japan (Tashiro *et al.*, 2007). The plasmid which was received on a filter paper was stored at 4 °C. Next, 70% ethanol was sprayed on forceps and scissor and was held against the Bunsen burner flame (flame sterilization). A circular mark was already made to indicate the plasmid location on the filter paper. Using the sterile scissor and forceps, the filter paper was cut around the marked area and placed at the bottom of a 2 ml Eppendorf tube. 200 µl of sterile 1x trypsin EDTA (TE) was added to cover the filter paper in the

Eppendorf tube. All steps were performed under flame sterilization. The Eppendorf tube was kept at 4 °C overnight. Next, the Eppendorf tube was spun in a centrifuge at 20,817 g for a minute and the plasmid solution at the bottom was transferred into a fresh 0.5 ml sterile Eppendorf tube and stored at 4 °C. The plasmid was then transformed (Griffith, 1928) into competent *Escherichia coli* strain DH5α cells. Transformation is a technique used to introduce foreign DNA in the form of a plasmid into competent bacterial cells. First, 50 µl of DH5α cells were thawed on ice. Next, in an Eppendorf tube, 2 µl of TE/plasmid was added into 50 µl of DH5α cells, mixed gently and incubated on ice for 30 minutes. This allowed the plasmid to coat bacteria. The Eppendorf tube was then placed on a heat block at 42 °C and the cells were heat-shocked for 40 seconds to allow plasmid uptake. Cells were placed directly back onto the ice and left to recover for 2 minutes. 450 µl of Luria broth (LB) (without antibiotic) was added to the cells and incubated with shaking for 1 h at 37 °C at 200 rpm. This allowed expression of the antibiotic resistance gene. Following incubation, 50 and 100 µl of the transformation mixture was spread onto agar plates containing 20 ml agar/plate and appropriate antibiotic, 100 µg/ml of ampicillin. The plates were placed upside down and incubated overnight at 37 °C to allow for colony formation. The following day, a single colony was picked and inoculated in 5 ml of LB broth containing the appropriate antibiotic (starter culture) and incubated with shaking for 6 hours at 37 °C at 200 rpm. 5 starter cultures were made. Following incubation, 1 ml of starter culture was added to 100 ml of LB broth containing the appropriate antibiotic and incubated with shaking overnight at 37 °C at 200 rpm. Next day, the overnight bacterial culture was harvested/centrifuged at 6000 g for 15 minutes at 4 °C to pellet the bacterial cells containing the plasmid of interest. Plasmid extraction was carried out using Qiagen maxi-prep kit (Catalog No. 12162). Following maxi-prep, all preps were run on an agarose gel (0.8% in 100 ml sodium borate) at 120 V for plasmid detection. Normally, detection of supercoiled DNA is expected from a plasmid prep. The plasmid from the preps was then quantified using a NanoDrop spectrophotometer.

**Table 2.6|** Plasmid

Vector	Insert	Selection
pcDNA3	Human XBP1-luciferase	Neomycin

---

## 2.13 Plasmid transfections

Uptake of DNA into cells have been reported earlier (Wong and Neumann, 1982, Chu *et al.*, 1987). MDA-MB-231 cells were transfected with human XBP1-Luciferase/pcDNA3 expression plasmid, pcDNA3 (empty vector control plasmid) and pEGFP plasmid (for determining transfection efficiency). The XBP1-Luciferase/pcDNA3 plasmid consists of human XBP1 cDNA fused upstream of luciferase cDNA (Tashiro *et al.*, 2007). At 24 hours prior transfection, 100,000 cells/well were seeded in quadruplicate per sample plasmid in 6 well plates. The cells were then kept for overnight incubation at 37 °C. Transfections were carried out using Turbofect transfection reagent (Catalog No: R0531; Lot No: 00455346). Next day, 1 µg of plasmid/well and 2 µl of Turbofect/well, were directly added to empty DMEM medium and incubated at RT for 20 minutes. Transfection of these cells with Turbofect requires a 1:2 ratio of DNA to lipid. During the incubation period the medium from the cells was replaced with 1 ml of fresh DMEM. Following incubation of DNA and lipids, the DNA-lipid complexes were added drop wise to the cells. The drops were allowed to fall into different areas of the well and the plate was swirled gently to ensure an even spread of the reagent. Cells were incubated for 24 hours at 37 °C. Following day, transfection efficiency was determined by examining GFP transfected cells under a fluorescent microscope. The media on cells was replaced with fresh DMEM containing 1400 µg/ml G418 and allowed to grow for a further 48 hours. Next, cells were trypsinized and pooled from quadruplicate wells and seeded at a density of 80,000 cells/6 well in sextuplicate in fresh DMEM. The following day, cells were treated with 1400 µg/ml G418. From here on, over the next few weeks, every 2 days, the medium on the cells was replaced with medium containing 1400 µg/ml G418 until all un-transfected cells were dead and colonies were only present in the XBP1-Luciferase/pcDNA3 transfected cells. Sometimes when the cells looked too stressed, the G418 concentration was reduced to 600 µg/ml for a few days to allow the DNA transfected colonies to expand. Once the colonies started expanding and cells (MDA-MB-231/XBP1-Luciferase) looked healthy the G418 concentration was again raised to 1400 µg/ml. This ensured growth of only DNA transfected cells in the colonies. Finally, when the colonies grew large enough, they were trypsinized and transferred to a 12 well plate and expanded thereafter to be tested for luciferase activity through treatments as discussed in the relevant section of the thesis.

---

## 2.14 Screening hit compounds by cellular assay

Screening compounds have become an important part of the academic drug discovery process (Kaiser, 2008, Fox, 2005, Coma *et al.*, 2009). The top 8 hit compounds from the VHTS campaign were screened in Corning® 96 well white polystyrene microplates (Catalog No: 3610) in duplicate wells towards their ability to inhibit the enzymatic activity of the ligase. Each plate had a negative control (100% inhibition; cells alone without Tg and compound) and a positive control (0% inhibition; cells with 0.5  $\mu$ M Tg). MDA-MB-231/XBP1-Luciferase cells at a density of 10,000 cells/well were seeded in 100  $\mu$ l of DMEM in 96 well mentioned above. After 24 hours of incubation, cells were first pre-treated with appropriate concentrations of the compounds from a 100 mM DMSO compound stock in 95  $\mu$ l of fresh DMEM and incubated at 37 °C with 5% CO<sub>2</sub> for 1 hour. Next, 5  $\mu$ l of 10  $\mu$ M Tg was added to each reaction well having the compounds using fresh tips. The final concentration of Tg was 0.5  $\mu$ M, for the assay. Wells were incubated with 0.5  $\mu$ M Tg together with or without screening compounds. The final volume for each well was 100  $\mu$ l. The final concentration of DMSO was <0.3% (v/v). The plates were incubated at 37 °C with 5% CO<sub>2</sub> for an additional 6 hours. Cells were then harvested, and XBP1-luciferase activity was measured as stated in section 2.21. Plates were analyzed for Z' factor (formula below) (Zhang *et al.*, 1999, Birmingham *et al.*, 2009) (0.79 in this case) based on controls, the most common quality metrics reported for small-molecule screens.

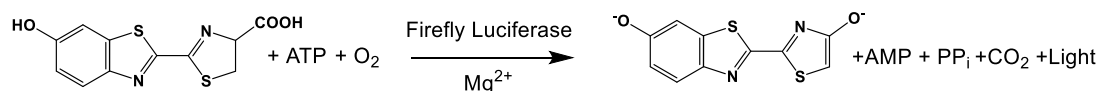
$$Z' \text{ factor} = 1 - (3\sigma_{\text{hc}} + 3\sigma_{\text{lc}}) / |\mu_{\text{hc}} - \mu_{\text{lc}}|$$

where  $\mu$  indicates mean,  $\sigma$  indicates standard deviation (SD),  $3\sigma$  indicates data variation defined by 3 SDs: both sample and control data is assumed to obey a normal distribution profile where ~68% of the data values are within one SD of the mean ( $\mu \pm \sigma$ ), ~95% of the data values are within two SDs of the mean ( $\mu \pm 2\sigma$ ), and ~99.7% lie within three SDs ( $\mu \pm 3\sigma$ ), “hc” indicates the high-value control, “lc” indicates the low-value control. The range of measurement is negative infinity to 1, with > 0.5 as a very good assay, > 0 an acceptable assay and < 0 an unacceptable assay. The cut-off for Hit compounds to be considered as primary hits was considered at ~50% inhibition of the positive control.

---

## 2.15 Luciferase assay

Transcriptional regulation coupled to the expression of a reporter gene is routinely used to study a wide range of physiological events. Luciferase is widely used as a reporter (Gould and Subramani, 1988) as a functional enzyme is created immediately upon translation and the assay is reliable, rapid and easy to perform. Analysis using luciferase as the genetic reporter is well suited to laboratory automation and high-throughput applications. The ONE-Glo™ Luciferase Assay System from Promega Corporation (Catalog No: # E6110) allows the measurement of luciferase activity as per the different treatments mentioned in chapter 5 of the thesis. This highly sensitive assay is based on a simple chemical reaction (Figure 2.2) where mono-oxygenation of the substrate (luciferin) is catalyzed by luciferase in the presence of magnesium, ATP and molecular oxygen. Beetle luciferin is a relatively stable molecule found only in luminous beetles (including fireflies). The enzyme uses ATP as a co-factor although most of the energy for photon production comes from molecular oxygen. The reaction produces light proportional to the activity of luciferase. The temperature of the ONE-Glo™ Reagent was ensured to be at room temperature before beginning measurements.



**Figure 2.2| The luciferase reaction.** Mono-oxygenation of luciferin is catalyzed by luciferase in the presence of magnesium, ATP and molecular oxygen.

For harvesting procedures, cell culture media was removed from the experimental 96-multiwell plate and gently rinsed with 100 µl 1X PBS buffer twice without dislodging any cells. Cells were lysed with 50 µl 1X cell culture lysis reagent buffer made as per composition (Table 2.7) for 30 minutes on ice. 2 µl of cell lysate from each of the sample wells was transferred to a 96-well plate to quantify the protein content through bicinchoninic acid assay (BCA) assay. The remaining 48 µl of cell lysates were transferred to a white 96-multiwell plate. Assays were run on duplicate wells. White plates provide maximum reflection and minimum auto-luminescence. 48 µl Luciferase Assay reagent was added to the wells and luminescence was measured 30 times with a delay of 30 seconds between each read in a Perkin Elmer Janus Automated



Workstation Elmer Plate Reader using the ‘Luciferase 96 well protocol’. For data analysis, background luminescence was subtracted from all samples. Data were normalized by dividing the reporter activity (luminescence) of each sample by the total protein content of that sample. Normalization to total protein content removes sample-to-sample variability due to cell number. Luminescence measurements of each plate were normalized to the control sample contained within the same plate. This allowed for the correction of small variations in luminescence that can occur over time or due to other variables such as temperature. Equation 1 was used to determine the normalized fold change in activity between test samples. This normalization gave a relative difference in activity between test samples within a single experiment. The normalized fold changes in activity from each experiment were averaged together and statistical significance was determined.

**Table 2.7|** Cell culture lysis reagent buffer composition

Stock Solution Concentration	Working Concentration	Volume/50 ml
50 mM Tris-HCL (pH 7.8)	25 mM	25 ml
1 M DTT	2 mM	100 $\mu$ l
1 M EDTA	2 mM	100 $\mu$ l
Glycerol	10 %	5 ml
Titron-X-100	1%	500 $\mu$ l
dH <sub>2</sub> O		Make up to 50 ml

$$\Delta \text{ Fold Activity} = \frac{\text{Average (Luciferase/Protein) from Test samples}}{\text{Average (Luciferase/Protein) from Control sample}}$$

**Equation 1|** Normalized fold change in activity between test samples.

## 2.16 BCA assay

BCA assay was first described by Smith and colleagues (Smith *et al.*, 1985). 1 mg/ml bovine serum albumin (BSA) in distilled water was prepared and frozen in 10  $\mu$ l aliquots. Briefly, 2  $\mu$ l of protein lysate was added in triplicate to a clear 96 well plate. A standard curve was generated using the 1 mg/ml BSA stock. 200  $\mu$ l of BCA assay

---

reagent from Thermo Fisher Scientific (Catalog No: # 23225) was added to each well and the plate was incubated at 37 °C for 30 minutes. Absorbance was measured at 595 nM.

## **2.17 Cell viability assay – Hoechst/PI staining**

Loss of cell viability was measured through double staining with Hoechst 33342/Propidium Iodide (PI) staining. Originally developed by Hoechst AG, Hoechst 33342 (Latt *et al.*, 1975, Latt and Stetten, 1976), the 33,342nd Bis-benzimide compound, a cell-permeant nuclear counterstain that emits blue fluorescence when bound to dsDNA, was used to stain cells. Hoechst 33342 from Invitrogen™ (Catalog No: H3570) was re-suspended to a stock concentration of 200 µg/ml in deionized water.

PI, a popular red-fluorescent nuclear and chromosome counterstain was used to detect dead cells in a population. Due to the loss of membrane integrity, PI easily enters dead cells. PI is not permeant to viable cells due to an intact cell membrane. This allows easy differentiation between viable and dead cells within a population. PI functions through binding to DNA by intercalating between the bases with little or no sequence preference. PI from Sigma-Aldrich (Catalog No: P4170) was re-suspended to a stock concentration of 50 µg/ml in deionized water.

MDA-MB-231/XBP1-Luciferase cells were seeded and treated with various treatments/drugs. Post-treatment, cells were stained with 25 µl 1X (1 µg/ml) Hoechst 33342/PI reagent-mix/well made from a 5X (5 µg/ml) stock at 37 °C with 5% CO<sub>2</sub> for 10 minutes. Final concentration of Hoechst 33342/PI/well was 1 µg/ml each. Cell viability was measured in the PerkinElmer Operetta® High Content Imaging System using the ‘Hoechst 33342/PI staining protocol’. For data analysis, an average of ‘% dead cells (((Number of dead cells/Number of all cells) × 100)/well)’ from duplicate treatment wells were calculated.

---

## 2.18 Statistical Analysis

Cell death, densitometry, Real-Time PCR and luciferase data were expressed as mean  $\pm$  SD/SEM for three independent experiments. GraphPad Prism's One-way ANOVA and paired two-tailed *t*-test or unpaired two-tailed *t*-test with Welch's correction were used to analyze the statistical significance of differences between the control and treatment groups. The p-values with  $*p < 0.05$  were considered statistically significant and values with  $***p < 0.001$  were considered very statistically significant.

---

## CHAPTER 3: Homology model of the human RtcB

**Copyright Information:** Copyright © 2017 Wiley Periodicals, Inc. All rights reserved.

### 3.1 Introduction and aims

RNA ligases catalyze the formation of a phosphodiester bond between the RNA strands generated by specific endonucleases during tRNA maturation, antiphage response, and in the UPR (Lu *et al.*, 2014b, Jurkin *et al.*, 2014a, Kosmaczewski *et al.*, 2014, Abelson *et al.*, 1998, Levitz *et al.*, 1990, Amitsur *et al.*, 1987, Popow *et al.*, 2012). For the most part these are classical ATP-dependent RNA ligases which, due to their inability to directly join the endonuclease generated RNA cleavage strands ending with 2',3'-cyclic phosphate and 5'-OH (Calvin and Li, 2008, Cuchillo *et al.*, 2011), utilize a multienzymatic pathway to repair the RNAs (Amitsur *et al.*, 1987, Pascal, 2008, Schwer *et al.*, 2004, Englert *et al.*, 2010, Englert *et al.*, 2012). In a sharp contrast to classical ligases, the RNA ligase RtcB facilitates a noncanonical direct joining/ligation of the 2',3'-cyclic phosphate and 5'-OH ends (Englert *et al.*, 2012, Tanaka *et al.*, 2011a, Chakravarty *et al.*, 2012, Perkins *et al.*, 1985, Chakravarty and Shuman, 2012, Tanaka and Shuman, 2011, Desai and Raines, 2012, Englert *et al.*, 2011, Desai *et al.*, 2013, Popow *et al.*, 2011). The RtcB mediated ligation proceeds through a three-step nucleotidyl transfer mechanism relying on GTP and  $Mn^{2+}$  for catalysis, in contrast to classical ligases which depend on ATP and  $Mg(II)$  (Tanaka *et al.*, 2011a, Chakravarty *et al.*, 2012, Chakravarty and Shuman, 2012, Desai *et al.*, 2013).

Crystal structures of the holo-RtcB complexes solved from the hyperthermophilic archaeobacterium *Pyrococcus horikoshii*, reveal a divalent metal ion specificity in the ligand binding site along with a covalently bound GMP (Englert *et al.*, 2012, Desai *et al.*, 2013). The structures provide important insight into how the active site, at its core, cycles through various states of structural and mechanistic rearrangements along the

---

entire guanylylation pathway. These geometrical rearrangements are required to accommodate the multiple conformational and stereo-chemical reactions in the single catalytic site of RtcB, remodeling the active site contacts and propelling the ligation reactions forward. Moreover, phylogenetic distributions of RtcB/HSPC117 reveal a high degree of conservation in a wide range of organisms from all three domains of life (bacteria, archaea, and mammals), with a deep and wide pocket lined by several conserved histidine residues and a cysteine (Chakravarty *et al.*, 2012, Popow *et al.*, 2011, Tanaka *et al.*, 2011b, Okada *et al.*, 2006).

From a clinical perspective, in order to arrest the *hRtcB* operated, XBP1 ligation-mediated tumor survival and progress, a promising strategy to search for new drug candidates would therefore initially involve the construction of a structural model of the ligase. With the determination of recent crystal structures for *P. horikoshii* RtcB (*bRtcB*), interest in structure-based ligand discovery for *hRtcB* has increased. Unfortunately, in spite of significant progress in X-ray crystallography and high-field NMR spectroscopy, structures of many therapeutically relevant target receptors including *hRtcB* remain absent. Such structures are of particular interest during early phase of drug discovery.

Therefore, in chapter 3 of the thesis, the primary aim was to construct an *in silico* 3D structure of the *hRtcB* protein using *bRtcB* as the main template, which could then eventually be used as an early phase drug discovery platform to discover inhibitors of the *hRtcB* RNA ligase. The accuracy of modeling and the inherent quality of the model constructed relies heavily on the sequence identity/similarity toward templates, performance of the modeling technique itself (including correctness of the alignment, loop folding and metal binding), and the choice of template structures used. To address the accuracy and reliability of *hRtcB* protein structure modeling that would eventually determine its possible application in lead discovery and optimization, seven different homology models were initially developed thereof, using five different modeling programs, based on the *bRtcB* crystal structures available from the Protein Data Bank. A thorough *in silico* quality assessment/evaluation of the models were executed using different computational tools. The best model for lead discovery and optimization was then distilled out and analyzed further. In addition, simulation techniques of wild type

---

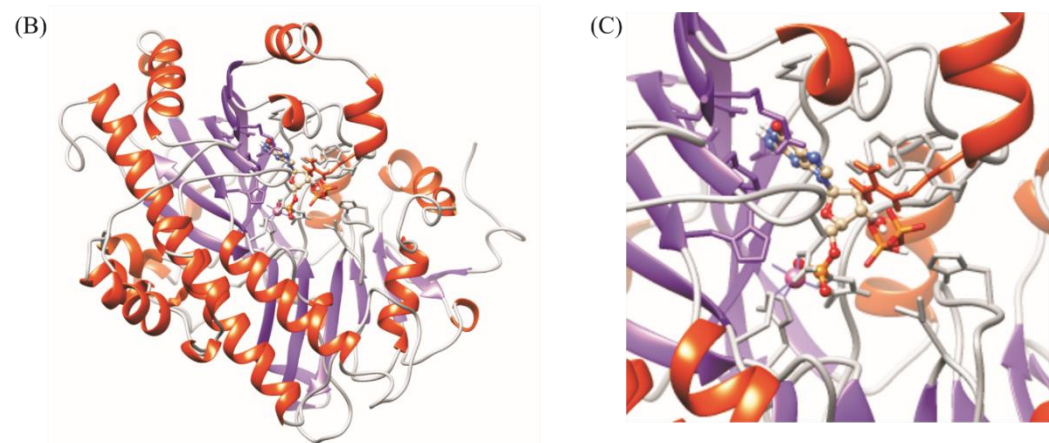
and C122A mutant models were employed in order to provide a rationale for the observed inability of the mutant to perform its catalytic activity (Popow *et al.*, 2011).

## 3.2 Results

### 3.2.1 Human *RtcB* homology model

Through a BLAST (Altschul *et al.*, 1997) search with solved PDB structures, a protein amino acid sequence identity between the human and *P. horikoshii* RtcB of 48% was determined. This is well above the minimal acceptable margin of 25% sequence identity required for creating homology models with an accuracy comparable to those of medium to low resolution X-ray structures (Xiang, 2016). To display the degree of conserved residues of the ligase between bacteria, archaea, and mammals, multiple sequence alignment was performed (Figure 3.1A). Importantly, the homology in the ligand-binding site/active site between human and *P. horikoshii* RtcB was 52% as determined by sequence identity. The sequence identities as stated above formed the rationale for developing a homology model of the *hRtcB* protein.

Five simulation programs, YASARA, MOE, I-TASSER, Phyre2, and SWISS-MODEL were employed in predicting a total of seven distinct structures of the human RtcB protein. Next, considering the vital role metals play toward catalysis and architecture of biological macromolecules, the predicted structures were examined for metal existence. Surprisingly, only the structures from YASARA and I-TASSER contained metal ions, whereas the remaining programs did not consider metal inclusion while constructing the 3D structures. Henceforth, the structures predicted by MOE, Phyre2, and SWISS-MODEL were not included in further assessments of their suitability in lead discovery and optimization. The *bRtcB* crystal structures 4DWR (Englert *et al.*, 2012), 4DWQ (Englert *et al.*, 2012), and 2EPG served as templates for modeling the human protein. 4DWR corresponds to a complex with two  $Mn^{2+}$  ions at the active site, surrounded by sulfate ions, while 4DWQ displays a complex with  $Mn^{2+}$  and the covalently bound ligand GMP. Upon detailed *in silico* quality assessments, described in the methods segment, the YASARA-built predicted structure (Figure 3.1B) was ranked as the structure of choice for future lead discovery and optimization. A possible reason for the overall high quality could be traced back to YASARA's exclusive ability to generate a hybrid model by assembling the best parts from (in this

[illegible]

53

585 of the unaligned intern domain were removed from the intein-containing precursor during splicing. Residues in red colored background represent identical residues. Residues in black colored background represents coordination with  $Mn^{2+}$  in the active site of the protein. Residues indicated with an overhead arrow interact with GMP in the active site of the protein structures. “\*” denotes fully conserved residues; “:” denotes conserved constitution and “.” denotes semiconserved. **(B)** Human RtcB homology model from YASARA, after MD simulation. The color orange-red denotes helix, purple denotes beta-sheet and gray denotes loop. The pink ball at the center depicts the  $Mn^{2+}$  metal cofactor. The natural ligand GMP is shown in ball and stick model, and PPi and some of the key interacting amino acid residues in stick model. **(C)** Zooming in on the active site. Color scheme same as in (B). “Reprinted from Volume 85, Issue 11, (Nandy *et al.*, 2017) Homology model of the human tRNA splicing ligase RtcB, Page 1986, Copyright (2017), with permission from John Wiley and Sons.”

**Table 3.1| Model Rankings.** The hybrid model and the 13 models generated in YASARA, listed by their overall quality Z-scores. Alignment variants have been designated by numbers 01 to 05 in the template column.

Rank	Model	Template	Z-score
1	The Hybrid Model	<u>4DWQ-A02</u> ; 2EPG-B03; 2EPG-B05; 2EPG-B01	-0.758
2	1	4DWQ-A01	-0.864
3	2	4DWQ-A03	-0.865
4	3	4DWQ-A02	-0.895
5	4	4DWQ-A04	-1.032
6	5	4DWR-B01	-1.040
7	6	4DWR-B02	-1.087
8	7	4DWR-B03	-1.158
9	8	4DWR-B04	-1.185
10	9	2EPG-B04	-1.315
11	10	2EPG-B03	-1.367
12	11	2EPG-B01	-1.372
13	12	2EPG-B02	-1.533
14	13	2EPG-B05	-1.564



---

### 3.2.2 *In silico* model quality assessment

Estimating the quality of predicted structures is necessary to select the best candidate from the ensemble of alternative structures generated by protein structure prediction methods. To this end, several *in silico* quality assessments/evaluations were executed as outlined in section 2.13, to determine the accuracy of the final two metal-containing structures. The YASARA-built predicted structure was found to have a better overall quality over the I-TASSER-built one, in particular in the Mn binding region, and was therefore subjected to a 10 ns MD simulation for further improvement (CREDIT to Leif A. Eriksson for performing the MD simulation).

#### 3.2.2.1 QMEAN quality assessment

To estimate the absolute quality of the predicted structures, the Qualitative Model Energy ANalysis (QMEAN), a composite scoring function that analyses the geometry of protein structures through six different structural descriptors was utilized (Benkert *et al.*, 2009). By means of QMEAN6 Z scores, the absolute quality was estimated and compared to scores from a nonredundant set of high-resolution experimental structures of similar size ( $\pm 10\%$ ) solved through X-ray crystallography. The QMEAN Z score can be defined as an estimate of the "degree of nativeness" of the structural features observed in a model by describing the likelihood that a model is of comparable quality to high-resolution experimental structures. A high positive score (Z score  $> 1$ ) corresponds to better structural quality. QMEAN Z-scores around zero indicate good agreement between the model structure and experimental structures of similar size. Scores of -4.0 or below are an indication of models with low quality. Therefore, the possible range of scores will be in general between -4.0 to 1. The area covered by the circles in different shades of grey in the plot (Figure 3.2A) represent the QMEAN scores of the reference structures from the PDB. The model's QMEAN score is compared to the scores obtained for experimental structures of similar size (model size  $\pm 10\%$ ) and a Z score is calculated. Low-quality models are expected to have strongly negative Z-scores for QMEAN. The QMEAN6 scores, a sign of structure reliability as a whole, showed the YASARA-built predicted structure to be reliable. In addition, despite the close similarity between the pre- and post-simulated structures, the MD simulations improved the absolute quality from -0.04 (pre-MD) to 0.35 (post-MD) (Figure 3.2A), confirming the simulated structure to be of highest prediction quality

---

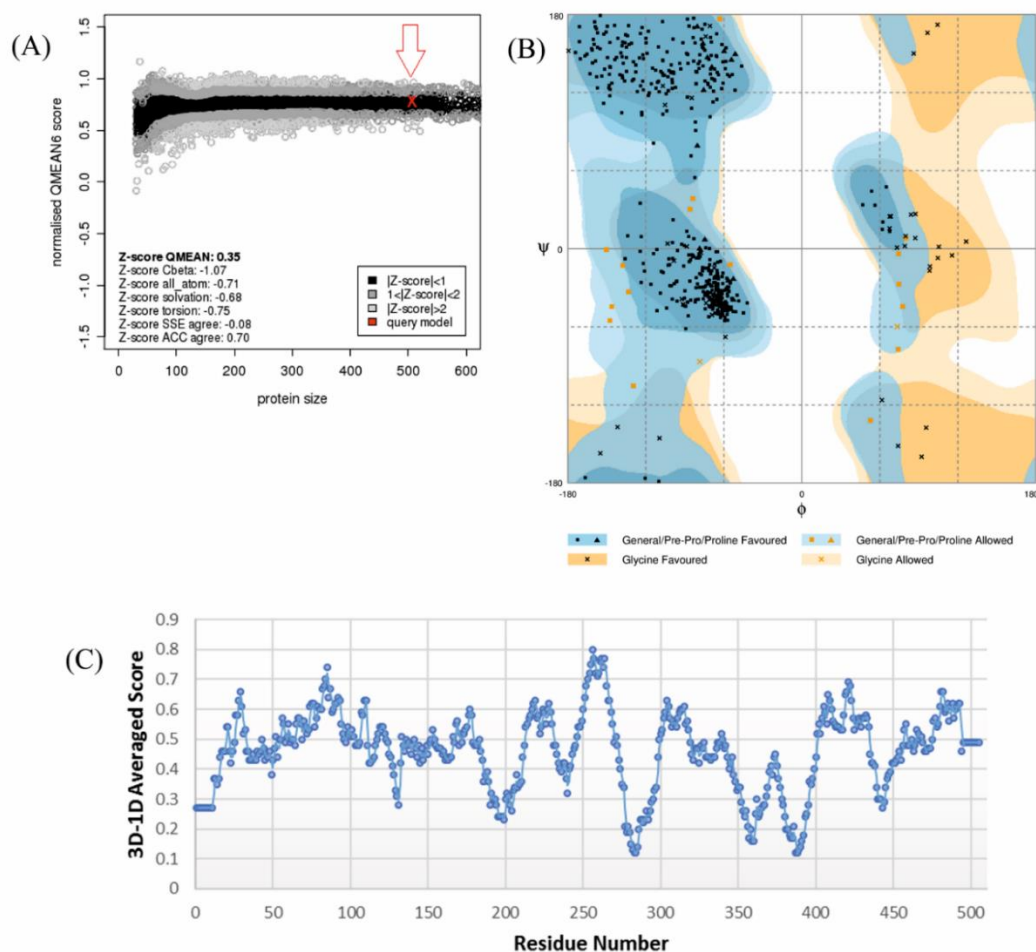
(Table 3.2).

### 3.2.2.2 Estimating stereochemical quality

A 2D Ramachandran plot of the backbone torsional angles  $\phi$  against  $\psi$  of residues in a protein structure signify the main chain N-C $\alpha$  and C $\alpha$ -C rotations. Using the RAMPAGE server to compute and display the different angles relative to energetically allowed regions for the ( $\phi, \psi$ ) pairs, enables an assessment of the stereochemical quality of the predicted structure. Already the initial homology model was found to be of high stereochemical quality, as the majority of the ( $\phi, \psi$ ) pairs fell in favored regions and only a few in the lower ranked “allowed” regions of the plot (Table 3.2). Two residues, a histidine and a leucine, were initially found in the outlier region in the structure. Although these residues are not part of the binding site, the MD simulation significantly improved the overall stereochemical quality from 95.2% favored and 0.4% outliers (pre-MD) to 96.4% favored and no outliers (post-MD) (Figure 3.2B).

### 3.2.2.3 Distinguishing between correct and incorrect folding

Using the VERIFY-3D server, the compatibility of the 3D amino acid environment of the predicted structures to their own 1D amino acid sequence were tested by a 3D profile score. In short, this differentiates between correct and incorrect folding through a residue-by-residue moving-window scan. The 3D profile of a protein structure is computed from the atomic coordinates of the structure that can be used to score the compatibility of the 3D structure model with any amino acid sequence. 3D profiles computed from correctly folded protein structures match their own sequences with high scores in contrast to 3D profiles computed from protein models known to be incorrect score poorly. The test of folding correctness is the compatibility of the model with its own amino acid sequence (Eisenberg *et al.*, 1997). Scores below zero indicate serious folding problems. The homology models both before and after MD refinement ‘passed’ the test (Table 3.2) as at least 80% of residues scored  $\geq 0.2$  in the 3D-1D profile. The MD simulation further corrected the folding quality from 93.5% residues scoring  $\geq 0.2$  (pre-MD) to 96.2% (post-MD) (Figure 3.2C). The low scoring regions in the plots (around residues 280, 360, and 390) are found in flexible loop regions facing the solvent.



**Figure 3.2| *In silico* model quality assessment.** (A) Normalized QMEAN6 Z scores. Analysis of absolute quality of the YASARA-built (post-MD) *hRtcB* structure indicated by red 'X'. SSE agree: secondary structure element agreement. ACC agree: solvent accessibility agreement. (B) Ramachandran plot. Analysis of stereochemical accuracy of *hRtcB* post-MD. (C) VERIFY-3D plot. The x axis indicates the sequence position while the y axis indicates the average 3D-1D score for residues in a moving-window of residues. Scores for the first and last 11 sequence positions have no meaning. “Reprinted from Volume 85, Issue 11, (Nandy *et al.*, 2017) Homology model of the human tRNA splicing ligase RtcB, Page 1987, Copyright (2017), with permission from John Wiley and Sons.”

**Table 3.2| Scores for structure reliability, absolute quality, Ramachandran plot, and VERIFY-3D.** Scores for YASRARA-built *hRtcB* homology model. QMEAN6 depict scores for structure reliability ranging from 0 to 1. QMEAN6 Z scores depict scores for structure absolute quality. Ramachandran scores indicate stereochemical quality. VERIFY-3D scores are an assessment of correct and incorrect folding.

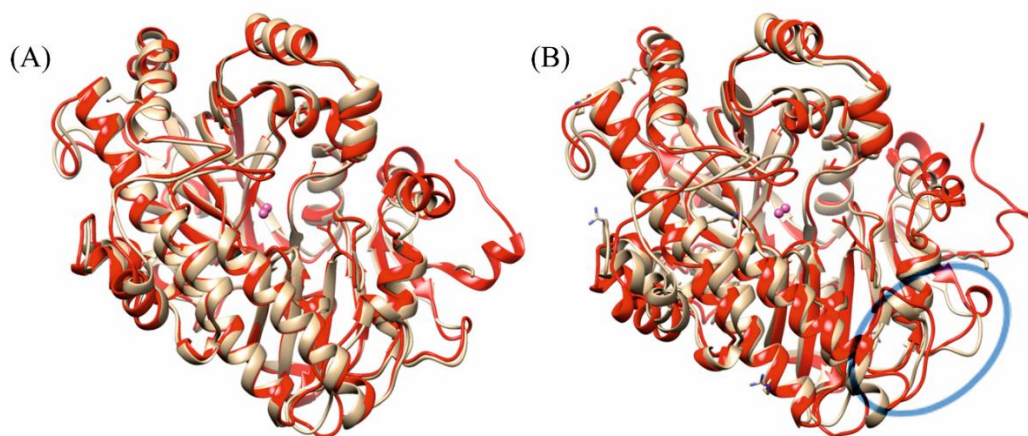
Prediction models	QMEAN6 scores	QMEAN6 Z-scores	Ramachandran scores		VERIFY-3D scores
			Favoured	Outliers	
<i>hRtcB</i> (pre-MD)	0.76	-0.04	95.2%	0.4%	93.4% Lowest score 0.06
<i>hRtcB</i> (post-MD)	0.79	0.35	96.4%	0.0%	96.2% Lowest score 0.12

### 3.2.2.4 Structure alignment of the human and bacterial homologs

The 3D fold of a protein greatly shapes its stability, ability to bind to other proteins or to ligands and drug candidates, and also the mechanical aspects of its behavior. Since structures are more highly conserved than sequences, 3D alignment-based protein structure comparisons, such as the PDBeFOLD server, can be used to evaluate the evolutionary distances between predicted and template structures. 3D coordinates of both structure types were aligned, and secondary structure identity/similarity computed by identifying residues that occupy equivalent geometric positions. The root-mean-square deviations (RMSDs) (Table 3.3) demonstrate high similarity between the aligned homology model and the *hRtcB* structures. Although three different templates were involved in the generation of the final model of *hRtcB*, the structure of the main template (PDB ID 4DWQ-A) was used in the determination of evolutionary distances to the predicted structure. Unlike previous quality assessments, where the structure after MD simulation displayed improved quality, the RMSD calculated against the main template increased somewhat (from 0.90 Å pre-MD to 1.5 Å post-MD; Figure 3.3A, B). The main source behind the reduced RMSD values from the MD simulations are fluctuations in loop structures, as indicated in the bottom right of Figure 3.3B. Interestingly, MD simulations did increase the %sse value (Table 3.3), indicating that the MD simulations generated a better agreement in terms of secondary

---

structure elements, to the template. A 3D residual alignment (Supplementary Material Figure S1) between the post-MD and template structures (residue-by-residue mapping) displays the extent of superposition between similar secondary structure folds and verifies the close structural similarity.



**Figure 3.3| Structure alignment assessment.** 3D structure alignment of *hRtcB* homology model (red) and *bRtcB* template 4DWQ-A (brown). (A) Before and (B) after MD simulation of *hRtcB*. Region of main fluctuation encircled in (B). “Reprinted from Volume 85, Issue 11, (Nandy *et al.*, 2017) Homology model of the human tRNA splicing ligase RtcB, Page 1988, Copyright (2017), with permission from John Wiley and Sons.”

**Table 3.3| Scores for 3D structure alignment.** Scores depict evolutionary distances between *hRtcB* homology model and *bRtcB* crystal structures. Q-score: quality of alignment, with 1 being the highest score; also known as the C $\alpha$  alignment. P-score: considers RMSD, number of aligned residues, number of gaps, number of matched secondary structure elements and the SSE match score; and Z-score is based on Gaussian Statistics. %seq: a score for amino acid sequence identity. %sse: score for secondary structure identity (sse: secondary structure element) between the two structures. RMSD, %sse and Q-score serve as optimal parameters toward assessing model identity/similarity with templates.

Prediction models	Template models	Scoring			RMSD (Å)	N <sub>align</sub>	%seq	%sse
		Q	P	Z				
<i>hRtcB</i> (pre-MD)	4DWQ-A	0.81	50.9	21.9	0.90	464	51	78
<i>hRtcB</i> (post-MD)	4DWQ-A	0.68	38.1	19.3	1.50	459	52	83

### 3.2.2.5 Binding pocket and druggability

The DoGSite algorithm-based pocket detection server (DoGSiteScorer) was used to predict (Volkamer *et al.*, 2010, Volkamer *et al.*, 2012) the ability of the *hRtcB* ligand binding site/active site to bind small molecules. Compared to other active site prediction methods which include geometry-based methods (Peters *et al.*, 1996, Liang *et al.*, 1998, Nayal and Honig, 2006), energy-based methods (Goodford, 1985, Laurie and Jackson, 2005), and evolutionary-based methods (Casari *et al.*, 1995, Pupko *et al.*, 2002, Lichtarge and Sowa, 2002), DoGSite maps proteins onto a grid and a difference of Gaussian (DoG) filter approach (Marr and Hildreth, 1980) is applied to detect potential pockets on the protein surface. The grid points forming the pocket and residues lining this, are used to calculate the different pocket properties including size, shape and hydrophobicity. By this approach, the best-fit active binding pocket/cavity was scored and its druggability predicted. The prediction is based on analyzing the geometric and physical-chemical properties of the binding pocket by means of a support vector mechanism (incorporated within the server), trained and tested on a druggability data set of 1069 target structures with a prediction accuracy of 88%. DoGSite correctly predicts binding pockets for over 92% of the PDBBind (Wang *et al.*, 2004) and the scPDB (Kellenberger *et al.*, 2006) data set that contains 828 and

---

6754 protein structures with druglike ligands, respectively. This method introduces a more precise prediction performance measure by taking into account the pairwise ligand and pocket coverage properties. In 90% of the cases, DoGSite predicts a pocket that contains at least half of the ligand. To address the heterogeneous nature of active sites, DoGSite introduces the concept of sub-pockets allowing a refined structural description of the topology of the entire active site. Consideration of sub-pockets produces an increase in coverage, yielding a success rate of 83% (Volkamer *et al.*, 2010).

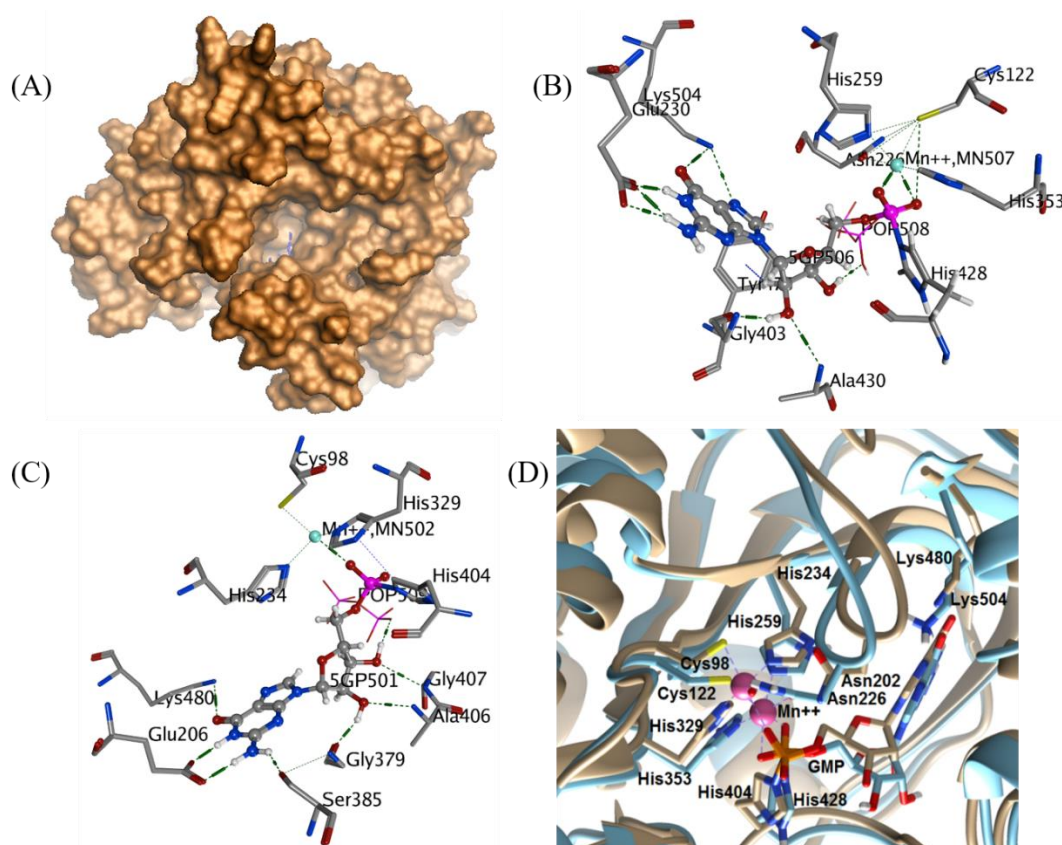
A predicted drug score of 0.80 shows the active binding pocket to be highly druggable (Table 3.4). In fact, a simple view through the surface topology (Figure 3.4A), especially the large and deep binding pocket of *hRtcB* with the natural ligand GMP buried deep inside, pinpoints the presence of protein folds that could favor interactions with small molecule compounds; an indication of a highly druggable target. The MD simulation improved the structure quality by increasing the druggability score to 0.83 and reducing the pocket volume from 1210.9 to 821.0 Å<sup>3</sup>, pocket area from 989.2 to 702.6 Å<sup>2</sup>, and pocket depth from 30.1 to 22.9 Å; crafting a geometrically more ordered and intact pocket suitable for drug targeting.

A detailed ligand interaction diagram as shown in Figure 3.4B reveals several residues interacting through hydrogen bonds with GMP; namely, Glu230, Gly403, Ala430, and Lys504. GMP is also forming a covalent bond with the Nε of His428, which is produced after the nucleophilic attack of His428 to GTP, with the release of pyrophosphate. The pyrophosphate remains present in the active site, where it interacts through a hydrogen bond with the C2' hydroxyl group of the ribose moiety. GMP is coordinating the divalent Mn, which is further coordinated by Cys122, Asn226, His259, and His353. The binding pattern is comparable to that observed for the template structure (Figure 3.4C), displaying hydrogen bond interactions of GMP with Glu206, Gly379, Ser685, Ala406, and Lys480. GMP is there forming a covalent bond with His404, while pyrophosphate is interacting with the ribose hydroxyl group. In *bRtcB*, Mn<sup>2+</sup> is coordinated to GMP, Cys98, His234, and His329. Moreover, a point mutation of the strictly conserved Cys122 residue (Cys98 in *bRtcB*) has been shown to render the ligase inactive (Popow *et al.*, 2011). The cysteine coordinates the Mn ion, essential for GTP binding. The homology model places Cys122 in perfect coordination



to Mn507, on the opposing side of the manganese to the bound GMP phosphate group.

Next, the structural similarities between the active sites was studied. Prior studies on *b*RtcB have already shown that upon GTP binding, the formation of a covalent RtcB-histidine-GMP intermediate with release of a pyrophosphate (PPi) induces conformational changes that extends to the metal ion binding site (Englert *et al.*, 2012, Desai *et al.*, 2013). A 3D structure alignment between the residues in the active site pocket of the predicted and template structure (Figure 3.4D) reveals a considerable overlap with its bacterial counterpart (RMSD of selected residues is 0.797 Å). Again, we note a slight increase in RMSD compared to the situation prior to MD simulation (Table 3.4).



**Figure 3.4| Ligand binding site.** (A) Surface topology of *h*RtcB, depicting the large, wide and deep binding pocket/cavity with the GTP ligand (blue) inside. (B) Binding pattern of GMP in *h*RtcB model (ball and stick representation). Key interacting residues of *h*RtcB are displayed in stick drawing, and pyrophosphate is displayed in line drawing. Hydrogen atoms of the protein are omitted for clarity. Green dashed lines represent hydrogen bonds, and blue dashed lines H- $\pi$  interactions. (C) Binding pattern of GMP (ball and stick representation) in *b*RtcB template. Key interacting residues of *b*RtcB are displayed



in stick drawing, and pyrophosphate is displayed in line drawing. Hydrogen atoms of the protein are omitted for clarity. Green dashed lines represent hydrogen bonds. **(D)** Superposition of binding cavity of human (blue) and bacterial RtcB (4DWQ-A) template structure (brown). “Reprinted from Volume 85, Issue 11, (Nandy *et al.*, 2017) Homology model of the human tRNA splicing ligase RtcB, Page 1989, Copyright (2017), with permission from John Wiley and Sons.”

**Table 3.4| Scores for ligand binding site and their 3D structure alignments.** Scores depict the druggability and different pocket properties from the *hRtcB* 3D prediction and *bRtcB* template structures. For each queried input structure, a druggability score between zero and one is returned. The higher the score the more druggable the pocket is estimated to be. RMSDs of the two *hRtcB* models are relative to the *bRtcB* template.

Models	Volume [Å <sup>3</sup> ]	Surface [Å <sup>2</sup> ]	Depth [Å]	Drug Scores	RMSD [Å]
<i>bRtcB</i> (4DWQ-A)	1039.0	941.3	28.0	0.80	0.000
<i>hRtcB</i> (pre-MD)	1210.9	989.2	30.1	0.80	0.555
<i>hRtcB</i> (post-MD)	821.0	702.6	22.9	0.83	0.797

### 3.2.2.6 Metal coordination microenvironment

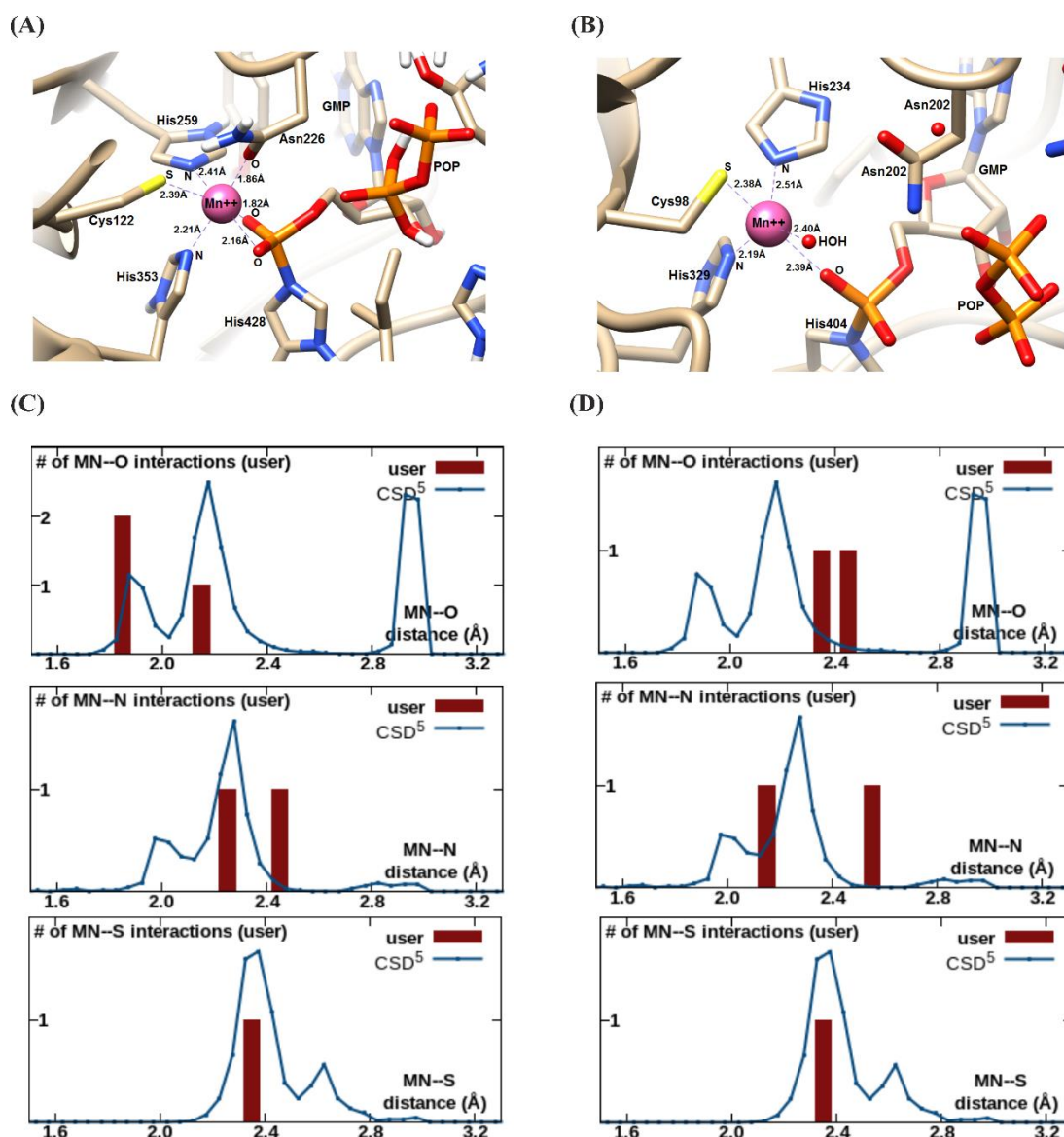
Metal cations play essential roles in catalysis and architecture of biological macromolecules. To further analyze the metal-binding microenvironment and its underlying contribution to the catalytic mechanism of *hRtcB*, we utilized the CheckMyMetal (CMM) Metal Binding Site Validation Server. Several well-established concepts of metal binding sites, coordination geometries, bond valence, vector sum of bond valence (VECSUM), and metal binding environment were considered for structure validations. The system under study was then compared to the corresponding data obtained from >7000 metal-binding sites in a benchmark dataset of >2000 high-resolution metal-containing X-ray structures from the PDB, employed by the server. Incorrect metal assignments and their geometries can lead to misinterpretation of molecular mechanisms and distort future lead discovery and optimization studies.

---

In the current system,  $\text{Mn}^{2+}$  is the central metal ion forming a coordination complex with a surrounding array of amino acid ligands in both the predicted and template RtcB structures.  $\text{Mn}^{2+}$  coordinates Cys122, Asn226, His259, His353, and GMP in the predicted structure (Figure 3.5A) while in the template (Figure 3.5B), it coordinates with Cys98, His234, His329, GMP, and a water molecule. Compared to the template, loss of a manganese-water coordination and movement of Asn226 (Asn202 in the template) to form a metal-ligand coordination, demonstrates the uniqueness of the human structure model. A rotation by 32.5 degrees of the His-bound phosphate group of GMP resulted in an additional Mn-O coordination for GMP in the simulated structure, relative to the template.

Evaluation of the binding environment through six parameters (ligands, valence, nVESCUM, geometry, gRMSD, vacancy) confirmed a complete and well modelled metal-binding microenvironment in the predicted structure (Table 3.5). Explanation of all six parameters are provided in the table legend.

Furthermore, all nonhydrogen atoms serving as metal ligands for the post-MD structure and its template were plotted in form of metal-ligand distance distribution graphs. These graphs show the distribution distances of the query structure in comparison to the distribution distances observed in high-resolution X-ray structures of small molecules deposited in the Cambridge Structural Database (CSD). Interestingly, compared to the template crystal structure (Figure 3.5C), the Mn-O, Mn-N, and Mn-S distributions observed in the simulated structure showed much higher similarity with the CSD structures (Figure 3.5D); a sign of a properly modeled metal binding site. Outliers indicate the presence of potential problematic metal-binding sites that might need further inspections. This analysis is in accordance with all the above described quality evaluations and confirms the MD simulated *hRtcB* homology model to be of high accuracy to be used in future lead discovery and optimization.



**Figure 3.5| Metal-binding microenvironment.** (A) Metal-binding microenvironment of the *hRtcB* homology model after MD simulation. The metal (pink ball) forms a coordination complex with four amino acid ligands (Cys122, Asn226, His259, His353) and the natural ligand GMP. (B) Metal-binding microenvironment of *bRtcB* template 4DWQ-A. The metal (pink ball) forms a coordination complex with three amino acid ligands (Cys98, His234, His329), a water molecule and the natural ligand GMP. Metal-ligand distance distribution graphs for (C) *hRtcB* and (D) *bRtcB*. The Mn-O, Mn-N, Mn-S metal-ligand interactions have been examined. Each graph signifies a normalized distribution of metal-ligand interaction distances found in the Cambridge Structure Database (CSD; blue lines) and distance distributions found in the post-MD structure (red boxes). The x axis shows the metal-ligand distance and the y axis the number of interactions in each bin. “Reprinted from Volume 85, Issue 11, (Nandy *et al.*, 2017) Homology model of the human tRNA splicing ligase RtcB, Page 1991, Copyright (2017), with permission from John Wiley and Sons.”

**Table 3.5| Scores for metal-binding site microenvironment.** Scores for the six different parameters used for evaluating the inherent consistency and geometrical arrangement of each metal-binding site. Each table cell colored in red, yellow or green depicting outlier, borderline and acceptable scores, respectively. **Ligands:** elemental composition of the coordination sphere; **Valence:** coordination based on metal-ligand distances; **nVECSUM:** sum over all vectors drawn from the central atom to its ligands within a coordination sphere; should be approximately zero. **Geometry:** completeness of the coordination sphere. **gRMSD:** evaluates the overall RMSD of all observed ligand-metal-ligand geometry angles compared to an ideal coordination geometry. Deviations of  $>20^\circ$  usually point towards a problematic metal-binding environment. **Vacancy:** the percentage of unoccupied sites within the coordination sphere for a particular detected geometry. **Bidentate:** evaluates all possible bidentate interactions. <sup>a</sup> Data relating to Figure 3.5A. <sup>b</sup> Data relating to Figure 3.5B.

Parameter Model	Metal	Ligands	Valence	nVECSUM	Geometry	gRMSD	Vacancy	Bidentate
<b><i>hRtcB</i> (pre-MD)</b>	Mn	O <sub>3</sub> N <sub>2</sub> S <sub>1</sub>	4	0.32	Octahedral	14.7°	0	0
<b><i>hRtcB</i><sup>a</sup> (post-MD)</b>	Mn	O <sub>3</sub> N <sub>2</sub> S <sub>1</sub>	4.5	0.35	Octahedral	16.9°	0	0
<b>Template<sup>b</sup> (4DWQ:A)</b>	Mn	O <sub>2</sub> N <sub>2</sub> S <sub>1</sub>	1.6	0.22	Trigonal Bipyramidal	11.3°	0	0

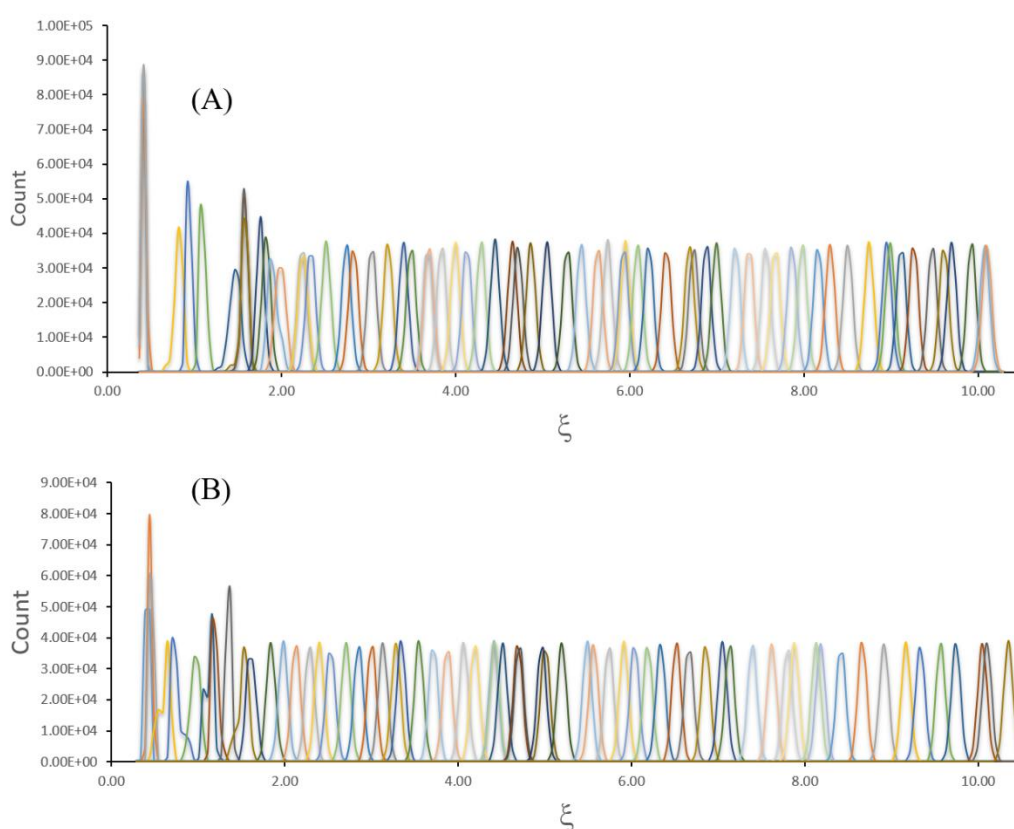
### 3.2.3 Mn binding in wild type and C122A mutant *hRtcB*

Umbrella sampling was employed to calculate the  $\Delta G$  of binding of  $Mn^{2+}$  to the protein structure (both wild and mutant variants), employing the distance between the metal ion and the center of mass (COM) of the protein as the reaction coordinate ( $\xi$ ). The histograms of the configurations within the umbrella sampling windows for both the wild-type *hRtcB* protein (Figure 3.6A) and the C122A mutant (Figure 3.6B), show that there is sufficient overlap between adjacent windows.

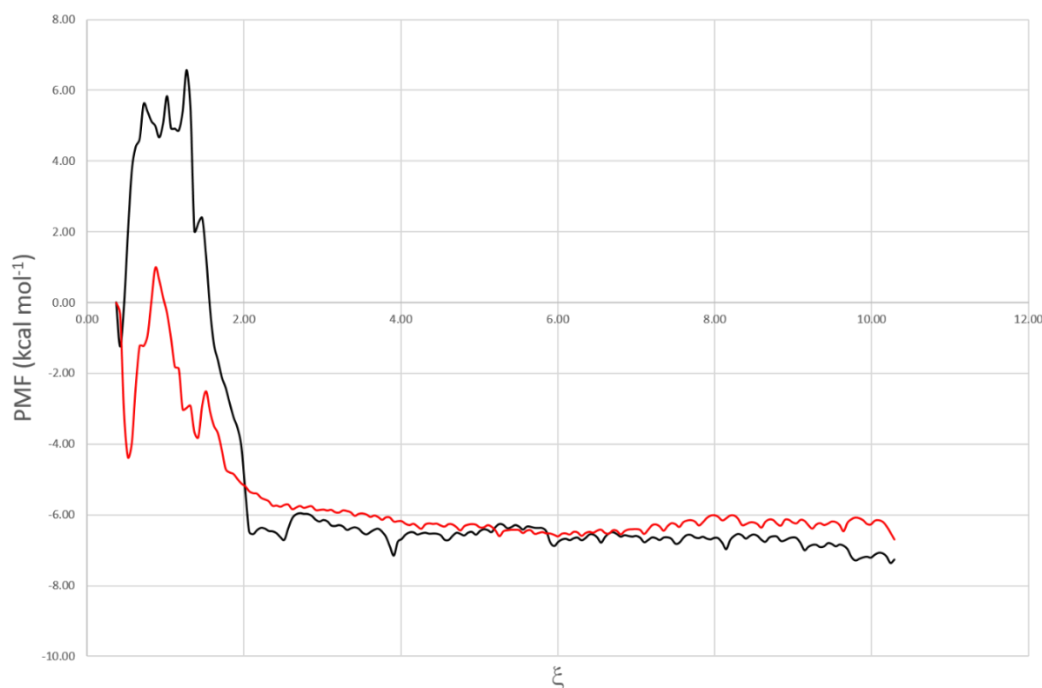
Using 68 and 64 sampling windows for wild-type and mutated systems, respectively, potential of mean force (PMF) curves were derived for  $Mn^{2+}$  (Figure 3.7). (CREDIT to Patricia Saenz-Méndez (P.S.-M.) for results obtained in figures 3.6A, B, and 3.7)

From these graphs, the values of  $\Delta G$  were calculated as the differences between the energy minimum and maximum of the PMF curve. For the wild type protein, the  $\Delta G$

of binding for  $\text{Mn}^{2+}$  was  $-13.93 \text{ kcal mol}^{-1}$ , while for the C122A mutant was  $-7.69 \text{ kcal mol}^{-1}$ . This difference is large enough to be statistically relevant. The metal ion is thus bound much stronger to the protein in the wild-type system, where Cys122 is participating in the binding. For the C122A mutant, this interaction is missing, dramatically diminishing the strength and number of  $\text{Mn}^{2+}$ -protein interactions. A plausible explanation for the lack of catalytic activity (Popow *et al.*, 2011) is thus that the markedly reduced coordination bond strength in C122A leads to impaired metal binding. Furthermore, changed coordination of the ligands in the mutant system can also play an important role, should the metal ion be able to locate to the active site.



**Figure 3.6| Umbrella sampling histograms.** (A) Wild-type protein and (B) C122A mutant (CREDIT to P.S.-M.). “Reprinted from Volume 85, Issue 11, (Nandy *et al.*, 2017) Homology model of the human tRNA splicing ligase RtcB, Supplementary information Page 11, Copyright (2017), with permission from John Wiley and Sons.”



**Figure 3.7| Potential of mean force (PMF) curves.** Wild-type *hRtcB* in black and C122A in red (CREDIT to P.S.-M.). “Reprinted from Volume 85, Issue 11, (Nandy *et al.*, 2017) Homology model of the human tRNA splicing ligase RtcB, Supplementary information Page 12, Copyright (2017), with permission from John Wiley and Sons.”

### 3.3 Discussion

Recent years have witnessed the identification of the elusive RNA ligase RtcB capable of joining 2',3'-cyclic phosphate and 5'-OH ends of RNA (Lu *et al.*, 2014b, Jurkin *et al.*, 2014a, Kosmaczewski *et al.*, 2014, Englert *et al.*, 2012, Chakravarty *et al.*, 2012, Desai *et al.*, 2013, Popow *et al.*, 2011), a phenomenon well explained during *XBPI* mRNA splicing. RNA repair of the 2',3'-cyclic phosphate and 5'-OH spliced ends broadly occurs through either the healing and sealing pathway or the direct ligation pathway (Chakravarty *et al.*, 2012). In contrast to classical ligases, RtcB follows the noncanonical direct ligation pathway. In the present study, we modeled a reliable 3D predicted protein model of the human RtcB using homology modeling. Although the determined structure is only a predicted model that has not been validated experimentally, such structures are of particular interest especially during early phase drug discovery, prior to the availability of experimental structures. Homology modeling based on the *bRtcB* templates, followed by MD simulation, provided a reliable predicted structure as verified by several independent *in silico* quality checks.

---

A vital aspect of drug discovery is the druggability of the ligand binding pocket, the control center that drives enzyme catalysis to carry out important biological processes. Druggability is described as the presence of protein folds that favor interactions with drug-like chemical compounds (Hopkins and Groom, 2002). Proteins that lack these structural features, although they might have interesting biological properties, are unlikely readily amenable to pharmacological manipulation. A good number of drug discovery projects fail due to the target of interest is found to be undruggable, i.e., the lack of a druggable domain. Based on a 2014 report, there are at least 18,097 different proteins (Wilhelm *et al.*, 2014) while the number of druggable targets in humans is estimated to be just above 3000 (Hopkins and Groom, 2002). Hence an automatic structure-based target assessment tool was employed to rapidly and reliably predict, characterize and estimate the druggability of the potential binding pockets and sub-pockets of the *hRtcB* homology model. The ligand binding pocket yielded an estimated high drug score of 0.83. Further analysis of the surface topology revealed a large and deep binding pocket with the natural ligand GMP buried deep inside (Figure 3.4A). The presence of protein folds in the ligand binding pocket showed promise to favor interactions with small molecule compounds. Collectively, these findings elucidate our predicted model to be a highly druggable target.

In our modeled *hRtcB*-GMP-Mn<sup>2+</sup> structure, Mn<sup>2+</sup> displays an octahedral coordination with four conserved ligands (Cys122, Asn226, His259, and His353) and GMP in its active binding site. The coordination differs slightly from that of the template structure (Englert *et al.*, 2012) in that a coordinating water molecule is replaced by Asn226, loss of a Mn<sup>2+</sup>-H<sub>2</sub>O coordination, and finally rotation of the His-bound phosphate group of GMP by 32.5 degrees leading to an additional Mn-O coordination for GMP; all of which define the unique features of our modeled structure. These important differences noted in our modeled structure should be important subjects of observation once the X-ray structure of *hRtcB* is solved. The *hRtcB*-GTP interaction forming a 5'-GMP-His intermediate through a covalent link with Nε2 of His428 in our structure is similar to the *bRtcB*-GTP interaction mechanism in the template structure (Englert *et al.*, 2012), and represents the end product of the guanylylation pathway. A probable interpretation behind the PPi group leaving GTP in the human structure can be deduced from the template (Englert *et al.*, 2012) as well as from another *bRtcB* structure (Desai *et al.*, 2013) where it is apically oriented to Nε of His404 (His428 in human), which is the

---

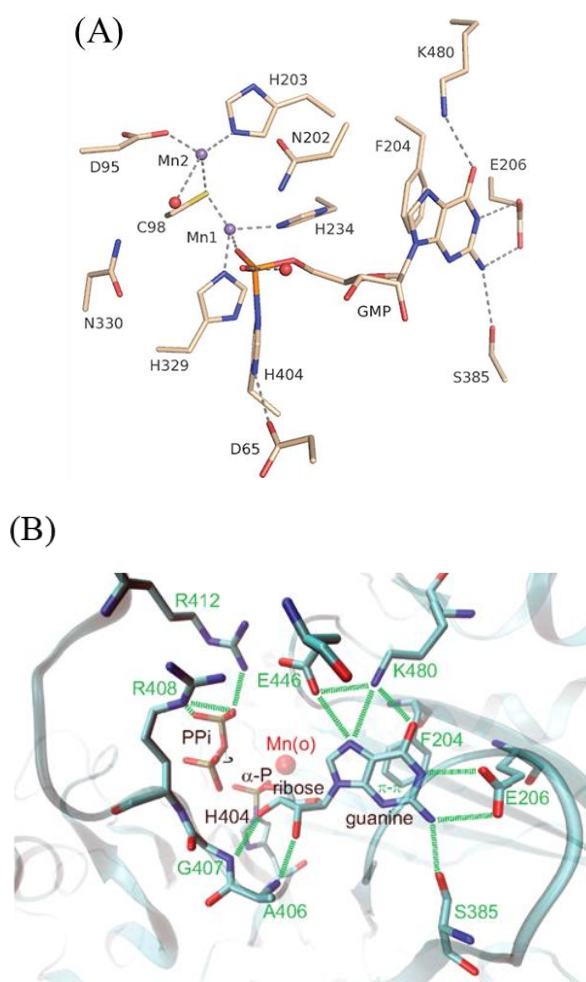
atom attacking the  $\alpha$ -phosphorus atom of GTP. This histidine guanylylation, as per published data from the bacterial homolog structures, should drive the RNA end sealing to a second step where the *hRtcB*-GMP-His moiety becomes transferred to, and activates, the RNA 3'-P end, followed by the final nucleotidyl transfer step where the 5'-OH from the opposite RNA strand attacks the activated 3'-P end forming *XBP1s* through a phosphodiester bond formation and GMP release.

The steps immediately preceding *bRtcB* guanylylation have been very well described in X-ray diffraction studies where the authors captured snapshots of the entire guanylylation pathway (Desai *et al.*, 2013). Structures from this study revealed that a single  $Mn^{2+}$  ion (Mn1) is bound to *bRtcB* prior to binding GTP. Secondly, in the step immediately preceding *bRtcB* guanylylation, *bRtcB* is in complex with  $Mn^{2+}$  and the unreactive GTP analog guanosine 5'-( $\alpha$ -thio)triphosphate (GTP $\alpha$ S). This *bRtcB*-GTP $\alpha$ S- $Mn^{2+}$  complex indicated the presence of two  $Mn^{2+}$  ions (Mn1 and Mn2) and the released pyrophosphate group of GTP $\alpha$ S is oriented apically to His404. Binding of GTP $\alpha$ S imparted significant conformational changes in the *bRtcB* ligand binding/active site. In the final step, *bRtcB* guanylylation marks the end of the guanylylation pathway where a *bRtcB*-His404-GMP intermediate is formed. This structure also indicated the presence of two  $Mn^{2+}$  ions (Mn1 and Mn2) in the ligand binding site (Figure 3.8A). These findings did suggest that *bRtcB* does share a similar two-metal mechanism utilized by the classical ATP-dependent nucleic acid ligases (El Omari *et al.*, 2006) in spite of using a different metal ion.

However, X-ray structures of *bRtcB* published in 2012 in the *PNAS*, although revealed the presence of two  $Mn^{2+}$  ions in the ligand binding site, they did report the presence of only one  $Mn^{2+}$  in the *bRtcB*-GMP- $Mn^{2+}$  structure with the loss of the second  $Mn^{2+}$  ion (Figure 3.8B) (Englert *et al.*, 2012). In line with the *bRtcB* guanylylated structures published by both the groups, our predicted *hRtcB* structure did represent a covalent interaction between GMP and His428 (His404 in *bRtcB*) resulting in a *hRtcB*-His428-GMP guanylylated form with the release of pyrophosphate but with one  $Mn^{2+}$  in the ligand binding site. The presence of only one  $Mn^{2+}$  in our predicted model was probably due to the fact that the template that was utilized to build our model traces back to the structures previously published by Jimin Wang's group (Englert *et al.*, 2012). Therefore, once the X-ray structure of *hRtcB* is available, an important subject



of observation would be whether the guanylation is a one- or two-metal mechanism and what sort of coordination do they form with the surrounding residues in the ligand binding site to carry out the enzymatic catalysis.



**Figure 3.8| Ligand binding/active site of *bRtcB*.** (A) *bRtcB*-His404-GMP covalent intermediate and the presence of two Mn<sup>2+</sup> ions (Mn1 and Mn2). “Reprinted with permission from (Desai, K. K., Bingman, C. A., Phillips, G. N., Jr. & Raines, R. T. Structures of the noncanonical RNA ligase RtcB reveal the mechanism of histidine guanylation. *Biochemistry* 52, 2518-2525, doi:10.1021/bi4002375 (2013)). Copyright (2013) American Chemical Society.” (B) *bRtcB*-His404-GMP covalent intermediate and the presence of one Mn<sup>2+</sup> ion. “Adapted from Volume 109, Issue 38, (Englert *et al.*, 2012) Structural and mechanistic insights into guanylation of RNA-splicing ligase RtcB joining RNA between 3'-terminal phosphate and 5'-OH, Page 15237, Copyright (2012), with permission from the National Academy of Sciences.”

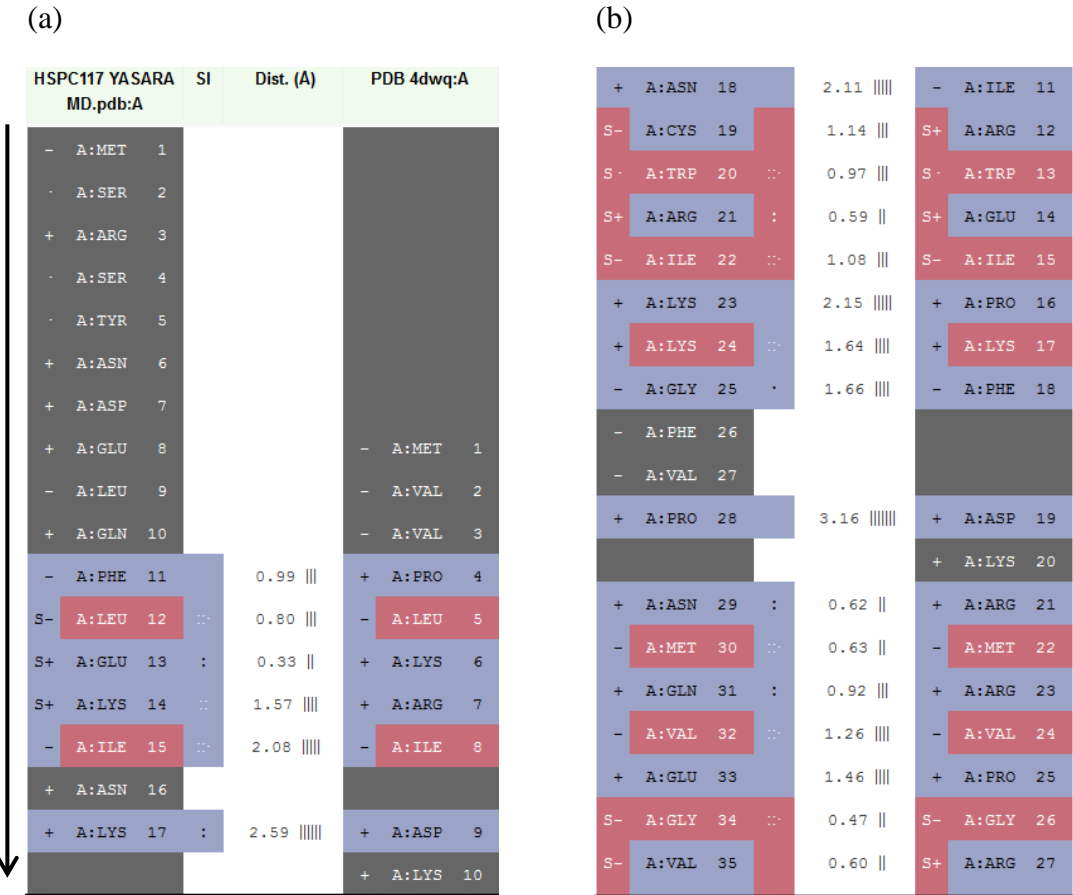
Finally, as the predicted model of *hRtcB* is now available, future investigations should aim to exploit the highly druggable ligand binding domain of the modeled structure as

---

a target site to screen small molecule libraries to discover inhibitors of *hRtcB*. This would preferentially inhibit the formation of *hRtcB*-histidine-GMP intermediate; the first step in the three-step re-ligation of broken RNA ends. Results presented in this chapter hence demonstrate the potential application of the predicted model in early phase drug discovery programs in search of potential inhibitors of the *hRtcB* ligase.

Supplementary Material

Figure S1: Residue-by-residue mapping of 3D structural alignment.



(c)

S-	A:PHE	36		0.83	
S-	A:TYR	37		1.57	
-	A:VAL	38	:	2.16	
+	A:ASN	39			
+	A:ASP	40			
H-	A:ALA	41			
H-	A:LEU	42			
H+	A:GLU	43			
H+	A:LYS	44			
H-	A:LEU	45		3.23	
H-	A:MET	46		3.05	
H-	A:PHE	47		2.51	
H+	A:GLU	48		2.60	
H+	A:GLU	49	:	3.07	
H-	A:LEU	50		2.70	
H+	A:ARG	51		2.60	
H+	A:ASN	52		3.34	
H-	A:ALA	53		3.20	
-	A:CYS	54			

(d)

+	A:ARG	55			
-	A:GLY	56			
-	A:GLY	57			
-	A:GLY	58			
-	A:VAL	59			
-	A:GLY	60			
-	A:GLY	61			
-	A:PHE	62			
-	A:LEU	63			
H+	A:PRO	64		0.45	
H-	A:ALA	65	.	0.12	
H-	A:MET	66		0.55	
H+	A:LYS	67	:	0.57	
H+	A:GLN	68		0.63	
H-	A:ILE	69	:	0.67	
H-	A:GLY	70		0.68	
H+	A:ASN	71		0.41	
H-	A:VAL	72		0.48	
H-	A:ALA	73		0.47	
H+	A:ARG	42			
H-	A:THR	43			
H-	A:LEU	44			
H+	A:GLU	45			
H+	A:GLN	46			
H-	A:ALA	47			
H-	A:THR	48			
H+	A:ASN	49			
H-	A:VAL	50			
H-	A:ALA	51			

(e)

-	A:ALA	74	:	0.70	
-	A:LEU	75		1.20	
+	A:PRO	76		2.13	
-	A:GLY	77		2.40	
S-	A:ILE	78		1.96	
S-	A:VAL	79		2.10	
S+	A:HIS	80	:	2.27	
S+	A:ARG	81		1.41	
S-	A:SER	82		0.40	
S-	A:ILE	83		0.82	
S-	A:GLY	84	.	0.88	
S-	A:LEU	85		0.99	
S+	A:PRO	86		1.00	
S+	A:ASP	87		0.90	
S-	A:VAL	88	.	0.72	
S+	A:HIS	89		0.74	
-	A:SER	90		0.73	
-	A:GLY	91		1.08	
-	A:TYR	92		0.99	
-	A:MET	52			
-	A:LEU	53			
+	A:PRO	54			
-	A:GLY	55			
-	A:ILE	56			
-	A:TYR	57			
+	A:LYS	58			
-	A:TYR	59			
S-	A:SER	60			
S-	A:ILE	61			
S-	A:VAL	62			
S-	A:MET	63			
S+	A:PRO	64			
S+	A:ASP	65			
S-	A:GLY	66			
S+	A:HIS	67			
+	A:GLN	68			
-	A:GLY	69			
-	A:TYR	70			

(f)

-	A:GLY	93		0.93	
-	A:PHE	94		0.73	
-	A:ALA	95	:	0.58	
S-	A:ILE	96		0.74	
S-	A:GLY	97		0.44	
S+	A:ASN	98		0.73	
S-	A:MET	99		0.41	
S-	A:ALA	100		0.36	
S-	A:ALA	101		1.07	
S-	A:PHE	102		1.08	
S+	A:ASP	103		1.19	
-	A:MET	104		1.34	
+	A:ASN	105	:	1.79	
+	A:ASP	106		1.88	
+	A:PRO	107			
+	A:GLU	108			
-	A:ALA	109	:	2.68	
-	A:VAL	110		1.77	
-	A:VAL	111		0.85	
-	A:GLY	71			
-	A:PHE	72			
+	A:PRO	73			
S-	A:ILE	74			
S-	A:GLY	75			
S-	A:GLY	76			
S-	A:VAL	77			
S-	A:ALA	78			
S-	A:ALA	79			
S-	A:PHE	80			
S+	A:ASP	81			
-	A:VAL	82			
+	A:LYS	83			
+	A:GLU	84			
-	A:GLY	85			
-	A:VAL	86			
-	A:ILE	87			

(g)

H-	A:SER 112	0.51	
H+	A:PRO 113	0.23	
H-	A:GLY 114	0.99	
H-	A:GLY 115	0.73	
H-	A:VAL 116	0.58	
-	A:GLY 117	0.48	
-	A:PHE 118	0.64	
+	A:ASP 119	0.41	
-	A:ILE 120	0.23	
+	A:ASN 121	0.46	
-	A:CYS 122	0.90	
S-	A:GLY 123	0.96	
S-	A:VAL 124	0.70	
S+	A:ARG 125	0.31	
S-	A:LEU 126	0.53	
S-	A:LEU 127	0.51	
S+	A:ARG 128	0.88	
S-	A:THR 129	0.93	
S+	A:ASN 130	1.42	

(h)

S-	A:LEU 131	1.16	
S+	A:ASP 132	1.30	
H+	A:GLU 133	0.99	
H-	A:SER 134	1.01	
H+	A:ASP 135	1.08	
H-	A:VAL 136	1.07	
H+	A:GLN 137	2.10	
+	A:PRO 138	1.71	
H-	A:VAL 139	0.84	
H+	A:LYS 140	0.74	
H+	A:GLU 141	0.53	
H+	A:GLN 142	0.70	
H-	A:LEU 143	0.71	
H-	A:ALA 144	1.19	
H+	A:GLN 145	0.86	
H-	A:ALA 146	0.59	
H-	A:MET 147	1.09	
H-	A:PHE 148	1.41	
H+	A:ASP 149	1.17	

(i)

H+	A:HIS 150	1.10	
H-	A:ILE 151	1.40	
+	A:PRO 152	1.07	
-	A:VAL 153	1.88	
-	A:GLY 154	2.29	
-	A:VAL 155	2.74	
-	A:GLY 156	3.99	
-	A:SER 157	2.95	
+	A:LYS 158	1.67	
-	A:GLY 159	0.68	
-	A:VAL 160	2.72	
-	A:ILE 161	3.65	
+	A:PRO 162		
-	A:MET 163		
H+	A:ASN 164		
H-	A:ALA 165	4.98	
H+	A:LYS 166	4.60	
H+	A:ASP 167		

(j)

H-	A:LEU 168		
H+	A:GLU 169		
H+	A:GLU 170	4.79	
H-	A:ALA 171	4.43	
H-	A:LEU 172		
H+	A:GLU 173		
H-	A:MET 174		
H-	A:GLY 175		
H-	A:VAL 176		
H+	A:ASP 177		
H-	A:TRP 178		
H-	A:SER 179	2.78	
H-	A:LEU 180	2.96	
H+	A:ARG 181	2.88	
H+	A:GLU 182	2.89	
-	A:GLY 183	3.03	

(k)

-	A:TYR 184	1.82	
-	A:ALA 185	2.74	
H-	A:TRP 186	1.37	
H-	A:ALA 187	2.43	
H+	A:GLU 188	2.46	
H+	A:ASP 189	0.93	
H+	A:LYS 190	0.66	
H+	A:GLU 191	0.86	
+	A:HIS 192	0.28	
-	A:CYS 193	0.39	
H+	A:GLU 194	0.65	
H+	A:GLU 195	0.87	
H-	A:TYR 196	2.26	
H-	A:GLY 197	2.36	
H+	A:ARG 198	1.65	
-	A:MET 199	0.44	
-	A:LEU 200	1.28	
+	A:GLN 201	2.82	
-	A:ALA 202	0.84	

(l)

H+	A:ASP 203	1.14	
H+	A:PRO 204	0.96	
H+	A:ASN 205	1.26	
H+	A:LYS 206	0.92	
H-	A:VAL 207	1.14	
H-	A:SER 208	1.37	
H-	A:ALA 209	1.62	
H+	A:ARG 210	2.24	
H-	A:ALA 211	1.87	
H+	A:LYS 212	2.15	
H+	A:LYS 213	2.79	
H+	A:ARG 214	3.51	
H-	A:GLY 215	3.92	
H-	A:LEU 216	3.39	
H+	A:PRO 217	1.55	
H+	A:GLN 218	1.06	
H-	A:LEU 219	1.56	
-	A:GLY 220	1.80	
-	A:THR 221	1.23	

(m)

-	A:LEU 222	0.86	
-	A:GLY 223	1.04	
-	A:ALA 224	2.14	
-	A:GLY 225	2.56	
+	A:ASN 226	2.30	
+	A:HIS 227	1.75	
S-	A:TYR 228	1.49	
S-	A:ALA 229	1.22	
S+	A:GLU 230	1.03	
S-	A:ILE 231	0.90	
S+	A:GLN 232	0.49	
S-	A:VAL 233	0.40	
S-	A:VAL 234	0.87	
S+	A:ASP 235	0.46	
S+	A:GLU 236	1.14	
S-	A:ILE 237	1.06	
-	A:PHE 238	1.63	
H+	A:ASN 239	1.58	
H+	A:GLU 240	1.67	

(n)

H-	A:TYR 241	1.87	
H-	A:ALA 242	1.96	
H-	A:ALA 243	2.07	
H+	A:LYS 244	1.97	
H+	A:LYS 245	2.14	
H-	A:MET 246	2.22	
-	A:GLY 247	2.36	
-	A:ILE 248	2.60	
+	A:ASP 249		
+	A:HIS 250	3.26	
+	A:LYS 251	1.05	
-	A:GLY 252	0.72	
S+	A:GLN 253	1.03	
S-	A:VAL 254	0.25	
S-	A:CYS 255	0.34	
S-	A:VAL 256	0.61	
S-	A:MET 257	0.53	
S-	A:ILE 258	0.27	
S+	A:HIS 259	0.73	

(o)

-	A:SER 260	0.37		S-	A:THR 235
-	A:GLY 261	0.52		-	A:GLY 236
H-	A:SER 262	0.45		H-	A:SER 237
H+	A:ARG 263	0.47		H+	A:ARG 238
H-	A:GLY 264	0.96		H-	A:GLY 239
H-	A:LEU 265	0.96		H-	A:LEU 240
H-	A:GLY 266	0.66		H-	A:GLY 241
H+	A:HIS 267	0.80		H+	A:HIS 242
H+	A:GLN 268	1.00		H+	A:GLN 243
H-	A:VAL 269	0.73		H-	A:VAL 244
H-	A:ALA 270	0.61		H-	A:ALA 245
H-	A:THR 271	1.04		H-	A:SER 246
H+	A:ASP 272	0.66		H+	A:ASP 247
H-	A:ALA 273	0.27		H-	A:TYR 248
H-	A:LEU 274	0.89		H-	A:LEU 249
H-	A:VAL 275	1.52		H+	A:ARG 250
H-	A:ALA 276	1.53		H-	A:ILE 251
H-	A:MET 277	1.43		H-	A:MET 252
H+	A:GLU 278	1.94		H+	A:GLU 253
H+	A:LYS 279	2.74		H+	A:ARG 254

(p)

-	A:ALA 280	2.91		H-	A:ALA 255
-	A:MET 281	2.07		H-	A:ILE 256
+	A:LYS 282				
+	A:ARG 283	1.87		+	A:ARG 257
+	A:ASP 284	2.50		+	A:LYS 258
					A:TYR 259
+	A:LYS 285	4.00		+	A:ARG 260
-	A:ILE 286	3.50		-	A:ILE 261
-	A:ILE 287	3.47		+	A:PRO 262
-	A:VAL 288	2.80			A:TRP 263
+	A:ASN 289	2.65		+	A:PRO 264
+	A:ASP 290	1.65		H+	A:ASP 265
+	A:ARG 291	1.59		H+	A:ARG 266
+	A:GLN 292	1.22		H+	A:GLU 267
-	A:LEU 293	1.04		H-	A:LEU 268
-	A:ALA 294	1.64		H-	A:VAL 269
-	A:CYS 295	1.65			A:SER 270
-	A:ALA 296	1.66		-	A:VAL 271
+	A:ARG 297	1.67		+	A:PRO 272

(q)

-	A:ILE 298	1.24		-	A:PHE 273
-	A:ALA 299	1.87		+	A:GLN 274
H-	A:SER 300	1.02		H-	A:SER 275
H+	A:PRO 301	0.52		H+	A:GLU 276
H+	A:GLU 302	0.82		H+	A:GLU 277
H-	A:GLY 303	0.57		H-	A:GLY 278
H+	A:GLN 304	0.64		H+	A:GLN 279
H+	A:ASP 305	0.80		H+	A:ARG 280
H-	A:TYR 306	0.29		H-	A:TYR 281
H-	A:LEU 307	0.39		H-	A:PHE 282
H+	A:LYS 308	0.51		H-	A:SER 283
H-	A:GLY 309	0.40		H-	A:ALA 284
H-	A:MET 310	0.24		H-	A:MET 285
H-	A:ALA 311	0.50		H+	A:LYS 286
H-	A:ALA 312	0.80		H-	A:ALA 287
H-	A:ALA 313	0.63		H-	A:ALA 288
H-	A:GLY 314	0.73		H-	A:ALA 289
H+	A:ASN 315	0.86		H+	A:ASN 290
H-	A:TYR 316	0.79		H-	A:PHE 291

(r)

H-	A:ALA 317	0.63		H-	A:ALA 292
H-	A:TRP 318	0.71		H-	A:TRP 293
H-	A:VAL 319	0.67		H-	A:ALA 294
H+	A:ASN 320	0.56		H+	A:ASN 295
H+	A:ARG 321	0.55		H+	A:ARG 296
H-	A:SER 322	0.53		H+	A:GLN 297
H-	A:SER 323	1.15		H-	A:MET 298
H-	A:MET 324	1.13		H-	A:ILE 299
H-	A:THR 325	0.66		H-	A:THR 300
H-	A:PHE 326	1.19		H+	A:HIS 301
H-	A:LEU 327	1.35		H-	A:TRP 302
H-	A:THR 328	0.80		H-	A:VAL 303
H+	A:ARG 329	1.02		H+	A:ARG 304
H+	A:GLN 330	1.40		H+	A:GLU 305
H-	A:ALA 331	1.06		H-	A:SER 306
H-	A:PHE 332	1.01		H-	A:PHE 307
H-	A:ALA 333	1.06		H+	A:GLN 308
H+	A:LYS 334	0.86		H+	A:GLU 309
H-	A:VAL 335	0.52		H-	A:VAL 310

(s)

H-	A:PHE 336	⊖	1.01		H-	A:PHE 311
H+	A:ASN 337	:	1.67		H+	A:LYS 312
-	A:THR 338		2.02		+	A:GLN 313
-	A:THR 339		1.78		+	A:ASP 314
+	A:PRO 340	⊖	1.44		+	A:PRO 315
+	A:ASP 341	:	1.05		+	A:GLU 316
					-	A:GLY 317
+	A:ASP 342	⊖	1.65		+	A:ASP 318
-	A:LEU 343	⊖	1.88		-	A:LEU 319
+	A:ASP 344		1.64		-	A:GLY 320
-	A:LEU 345	⊖	1.34		-	A:MET 321
+	A:HIS 346		1.08		+	A:ASP 322
S-	A:VAL 347	⊖	0.44		S-	A:ILE 323
S-	A:ILE 348	⊖	0.30		S-	A:VAL 324
S-	A:TYR 349	⊖	0.39		S-	A:TYR 325
S+	A:ASP 350	⊖	0.58		S+	A:ASP 326
S-	A:VAL 351	⊖	0.71		S-	A:VAL 327
S-	A:SER 352	*	0.93		S-	A:ALA 328
+	A:HIS 353	⊖	1.04		+	A:HIS 329

(t)

S+	A:ASN 354	⊖	0.77		S+	A:ASN 330
S-	A:ILE 355	⊖	0.55		S-	A:ILE 331
S-	A:ALA 356	:	0.32		S-	A:GLY 332
S+	A:LYS 357	⊖	1.46		S+	A:LYS 333
S-	A:VAL 358	⊖	1.16		S-	A:VAL 334
S+	A:GLU 359	⊖	0.70		S+	A:GLU 335
S+	A:GLN 360	:	0.61		S+	A:GLU 336
S+	A:HIS 361	⊖	0.45		S+	A:HIS 337
S-	A:VAL 362		0.57		S+	A:GLU 338
S-	A:VAL 363	⊖	1.71		S-	A:VAL 339
+	A:ASP 364	⊖	3.84		+	A:ASP 340
-	A:GLY 365	⊖	2.52		-	A:GLY 341
S+	A:LYS 366	⊖	0.74		S+	A:LYS 342
S+	A:GLU 367	:	0.28		S+	A:ARG 343
S+	A:ARG 368		1.49		S-	A:VAL 344
S-	A:THR 369		0.66		S+	A:LYS 345
S-	A:LEU 370	⊖	1.03		S-	A:VAL 346
S-	A:LEU 371	⊖	0.70		S-	A:ILE 347
S-	A:VAL 372	⊖	0.62		S-	A:VAL 348

(u)

S+	A:HIS 373	⊖	0.46		S+	A:HIS 349
S+	A:ARG 374	⊖	0.45		S+	A:ARG 350
S+	A:LYS 375	⊖	0.32		S+	A:LYS 351
-	A:GLY 376		0.48		-	A:GLY 352
-	A:SER 377	*	0.58		S-	A:ALA 353
-	A:THR 378	⊖	1.05		S-	A:THR 354
+	A:ARG 379	⊖	0.66		S+	A:ARG 355
-	A:ALA 380	⊖	1.64		-	A:ALA 356
-	A:PHE 381	⊖	1.47		-	A:PHE 357
+	A:PRO 382	⊖	1.45		+	A:PRO 358
+	A:PRO 383	⊖	1.46		+	A:PRO 359
+	A:HIS 384		1.79		-	A:GLY 360
+	A:HIS 385	⊖	1.65		+	A:HIS 361
+	A:PRO 386		2.05		+	A:GLU 362
-	A:LEU 387	:	2.40		-	A:ALA 363
-	A:ILE 388	⊖	3.25		-	A:VAL 364
H-	A:ALA 389	:	3.45		H+	A:PRO 365
H-	A:VAL 390		2.69		H+	A:ARG 366
H+	A:ASP 391		3.53		H-	A:LEU 367

(v)

H-	A:TYR 392	⊖	2.83		H-	A:TYR 368
H+	A:GLN 393	:	1.79		H+	A:ARG 369
-	A:LEU 394		1.03		+	A:ASP 370
-	A:THR 395		1.94		-	A:VAL 371
-	A:GLY 396	⊖	1.62		-	A:GLY 372
+	A:GLN 397	⊖	0.86		+	A:GLN 373
S+	A:PRO 398	⊖	0.73		S+	A:PRO 374
S-	A:VAL 399	⊖	0.69		S-	A:VAL 375
S-	A:LEU 400	⊖	0.49		S-	A:LEU 376
S-	A:ILE 401	⊖	0.75		S-	A:ILE 377
S-	A:GLY 402	:	0.69		+	A:PRO 378
-	A:GLY 403	⊖	1.24		-	A:GLY 379
-	A:THR 404	:	1.22		-	A:SER 380
-	A:MET 405	⊖	1.16		-	A:MET 381
-	A:GLY 406	⊖	1.19		-	A:GLY 382
-	A:THR 407	⊖	0.76		-	A:THR 383
-	A:CYS 408		0.78		-	A:ALA 384
S-	A:SER 409	⊖	0.76		S-	A:SER 385
S-	A:TYR 410	⊖	0.84		S-	A:TYR 386



(w)

S-	A:VAL 411	∅	0.75		S-	A:ILE 387
S-	A:LEU 412	∅	0.70		S-	A:LEU 388
S-	A:THR 413	*	0.39		S-	A:ALA 389
S-	A:GLY 414	∅	0.74		-	A:GLY 390
H-	A:THR 415	∅	0.85		H-	A:THR 391
H+	A:GLU 416	∅	0.95		H+	A:GLU 392
H+	A:GLN 417		1.27		H-	A:GLY 393
H-	A:GLY 418	:	0.77		H-	A:ALA 394
H-	A:MET 419	∅	1.01		H-	A:MET 395
H-	A:THR 420		1.53		H+	A:LYS 396
H+	A:GLU 421	∅	0.77		H+	A:GLU 397
H-	A:THR 422	∅	0.69		H-	A:THR 398
-	A:PHE 423	∅	1.16		-	A:PHE 399
-	A:GLY 424	∅	0.76		-	A:GLY 400
-	A:THR 425	*	0.22		S-	A:SER 401
-	A:THR 426	∅	0.65		S-	A:THR 402
-	A:CYS 427	∅	1.82		S-	A:CYS 403
+	A:HIS 428	∅	1.43		+	A:HIS 404
-	A:GLY 429	∅	1.40		-	A:GLY 405

(x)

-	A:ALA 430	∅	1.67		-	A:ALA 406
-	A:GLY 431	∅	1.44		-	A:GLY 407
+	A:ARG 432	∅	1.41		+	A:ARG 408
-	A:ALA 433	:	1.64		-	A:VAL 409
-	A:LEU 434	∅	1.26		-	A:LEU 410
H-	A:SER 435	∅	1.27		H-	A:SER 411
H+	A:ARG 436	∅	1.02		H+	A:ARG 412
H-	A:ALA 437		1.12		H+	A:LYS 413
H+	A:LYS 438		1.61		H-	A:ALA 414
H-	A:SER 439	*	1.22		H-	A:ALA 415
H+	A:ARG 440		1.67		H-	A:THR 416
H+	A:ARG 441	∅	3.02		H+	A:ARG 417
H+	A:ASN 442	:	3.20		H+	A:GLN 418
H-	A:LEU 443	∅	2.06		H-	A:TYR 419
+	A:ASP 444	:	1.61		H+	A:ARG 420
-	A:PHE 445	*	2.45		H-	A:GLY 421
H+	A:GLN 446	*	1.82		H+	A:ASP 422
H+	A:ASP 447	:	1.26		H+	A:ARG 423
H-	A:VAL 448	∅	1.87		H-	A:ILE 424

(y)

H-	A:LEU 449		1.29		H+	A:ARG 425
H+	A:ASP 450	*	1.23		H+	A:GLN 426
H+	A:LYS 451	:	1.73		H+	A:GLU 427
H-	A:LEU 452	∅	1.94		H-	A:LEU 428
H-	A:ALA 453	:	1.54		H-	A:LEU 429
H+	A:ASP 454	*	2.49		H+	A:ASN 430
H-	A:MET 455		3.10		H+	A:ARG 431
-	A:GLY 456	∅	2.38		-	A:GLY 432
-	A:ILE 457	∅	1.12		-	A:ILE 433
S-	A:ALA 458	:	0.33		S-	A:TYR 434
S-	A:ILE 459	∅	0.52		S-	A:VAL 435
S+	A:ARG 460	∅	0.85		S+	A:ARG 436
S-	A:VAL 461	:	0.82		S-	A:ALA 437
-	A:ALA 462	∅	1.04		-	A:ALA 438
H-	A:SER 463	∅	1.83		H-	A:SER 439
H+	A:PRO 464		2.02		H-	A:MET 440
H+	A:LYS 465	∅	2.12		H+	A:ARG 441
H-	A:LEU 466	∅	2.33		H-	A:VAL 442
H-	A:VAL 467	∅	2.27		H-	A:VAL 443

(z)

H-	A:MET 468	:	1.70		H-	A:ALA 444
+	A:GLU 469	∅	1.30		H+	A:GLU 445
+	A:GLU 470	∅	1.03		H+	A:GLU 446
-	A:ALA 471	∅	0.99		-	A:ALA 447
+	A:PRO 472	∅	0.90		+	A:PRO 448
+	A:GLU 473		0.59		-	A:GLY 449
-	A:SER 474	*	0.97		-	A:ALA 450
-	A:TYR 475	∅	0.71		-	A:TYR 451
+	A:LYS 476	∅	1.03		+	A:LYS 452
+	A:ASN 477	∅	0.47		H+	A:ASN 453
H-	A:VAL 478	∅	1.53		H-	A:VAL 454
H-	A:THR 479		1.50		H+	A:ASP 455
H+	A:ASP 480	*	2.19		H+	A:ASN 456
H-	A:VAL 481	∅	1.69		H-	A:VAL 457
H-	A:VAL 482	∅	0.97		H-	A:VAL 458
H+	A:ASN 483	:	1.36		H+	A:LYS 459
H-	A:THR 484		1.28		H-	A:VAL 460
H-	A:CYS 485		0.70		H-	A:VAL 461
H+	A:HIS 486	*	1.42		H-	A:SER 462

(z1)

H+	A:ASP 487	·	1.05		H+	A:GLU 463
H-	A:ALA 488		0.92		H-	A:ALA 464
-	A:GLY 489		0.90		H-	A:GLY 465
-	A:ILE 490		0.81		-	A:ILE 466
S·	A:SER 491	·	0.94		-	A:ALA 467
S+	A:LYS 492		1.06		S+	A:LYS 468
S+	A:LYS 493		1.50		S-	A:LEU 469
S-	A:ALA 494	:	1.05		S-	A:VAL 470
S-	A:ILE 495	:	0.73		S-	A:ALA 471
S+	A:LYS 496		1.03		S+	A:ARG 472
S-	A:LEU 497		1.01		S-	A:MET 473
S+	A:ARG 498		0.93		S+	A:ARG 474
S+	A:PRO 499		0.77		S+	A:PRO 475
S-	A:ILE 500		0.59		S-	A:ILE 476
S-	A:ALA 501	:	0.92		S-	A:GLY 477
S-	A:VAL 502		0.63		S-	A:VAL 478
S-	A:ILE 503	:	1.09		S-	A:ALA 479
S+	A:LYS 504		1.32		S+	A:LYS 480
S-	A:GLY 505		1.19		-	A:GLY 481

**Figure S1| Residue-by-residue mapping of 3D structural alignment.** Human RtcB YASARA-built prediction structure (10 ns MD simulated) aligned over *b*RtcB 4DWQ-A template disclosing the extent of superposition between similar secondary structure folds. ***h*RtcB YASARA MD.pdb:A column:** Residues related to this structure running sequentially along the structure chain. The column has two sub-columns. Left sub-column, on a colored background, denotes identification of secondary structure and hydropathy of the residue. Secondary structure is identified by H (Helix) and S ( $\beta$ -strand) residues; white space means that the residue does not belong to either helix or strand. Hydropathy of residues is denoted either by hydrophilic (+), hydrophobic (-) or a weak hydropathy (dot). Background color red denotes matched residues that have identical secondary structure, cyan denotes other matched residues and black denotes unmatched residues. Right sub-column on a colored background denotes the residue-chain ID and residue name. Red background denotes matched residues of similar names, cyan denotes other matched residues and black denotes unmatched residues. **Similarity index (SI) column:** Degree of structural and chemical similarity between matched residues from zero to five dots. Five dots on red background denote identical or highly similar residues; two-three dots marked in cyan indicate somewhat similarity; no dots marked on grey reveal dissimilarity between residues. **Distance column:** Distance ( $\text{\AA}$ ) between matched residues at the best 3D alignment of *h*RtcB and 4dwq:A. The indicator one to ten vertical strips emphasize better and worse regions of alignment. The SI and distance columns for residues that are not matched by 3D alignment are denoted by an empty space. **PDB 4dwq:A column:** Residues related to this structure corresponding to residues of *h*RtcB YASARA MD.pdb:A by 3D matching. The structure of this column is similar to the *h*RtcB YASARA MD.pdb:A column and hence all the descriptors reveal similar connotations. “Reprinted from Volume 85, Issue 11, (Nandy *et*

---

*al.*, 2017) Homology model of the human tRNA splicing ligase RtcB, Supplementary information Page 2, Copyright (2017), with permission from John Wiley and Sons.”

---

## CHAPTER 4: Ligand discovery from a homology model of the human RtcB

### 4.1 Introduction and aims

Expression of XBP1 is significantly associated with treatment outcomes in breast cancer. High levels of XBP1s are associated with shorter patient survival (Davies *et al.*, 2008, Chen *et al.*, 2014). Neo-adjuvant chemotherapy, an element of the standard care for treating TNBCs has been shown to increase the levels of XBP1s expression (Logue *et al.*, 2018). Besides standard neo-adjuvant regimens which include anthracyclines, taxanes, and cyclophosphamide (Senkus *et al.*, 2015, Huober *et al.*, 2010), different systemic treatment options have also been investigated in the recent years (Gerber *et al.*, 2013, Yardley *et al.*, 2015, Tutt *et al.*, 2010, Traina *et al.*, 2015, Nabholz *et al.*, 2014). Due to lack of targeted therapies for TNBC in the neo-adjuvant setting, different candidate drugs are currently being tested. Although a range of small molecule inhibitors targeting the formation of XBP1s (Thamsen *et al.*, 2018, Volkmann *et al.*, 2011) are available, they also inhibit the IRE1 $\alpha$ -mediated RIDD activity that elicits anti-tumorigenic properties at least in glioblastoma (Lhomond *et al.*, 2018). Hence, a more selective intervention such as inhibiting the RtcB ligase that re-ligates cleaved *XBPIu* mRNA forming *XBPIs* seems promising as its depletion by RNAi has been shown to keep the RIDD unaffected in a HeLa cell model (Jurkin *et al.*, 2014a). Currently, there are no known drugs that target the RtcB ligase. A drug discovery program was therefore initiated to selectively target *hRtcB*-mediated production of *XBPIs*. A range of drug screening paradigms exists which include HTS (Fox *et al.*, 2006), virtual screening (McInnes, 2007, Lavecchia and Di Giovanni, 2013), fragment screening (Law *et al.*, 2009), NMR screening (Fernandez and Jahnke, 2004, Kim and Wyss, 2015) or a more physiologically relevant, phenotypic screening (Zheng *et al.*, 2013). With the advancement of *in silico* techniques, efforts towards screening compound libraries have become much focused and straight-forward. Millions of compounds can now be easily screened via virtual high-throughput screening (VHTS) techniques at the early drug discovery phase without the need of undertaking an extensive HTS program. An example of a successful VHTS was the identification of a first-in-class small molecule activin antagonist for treating activin-

---

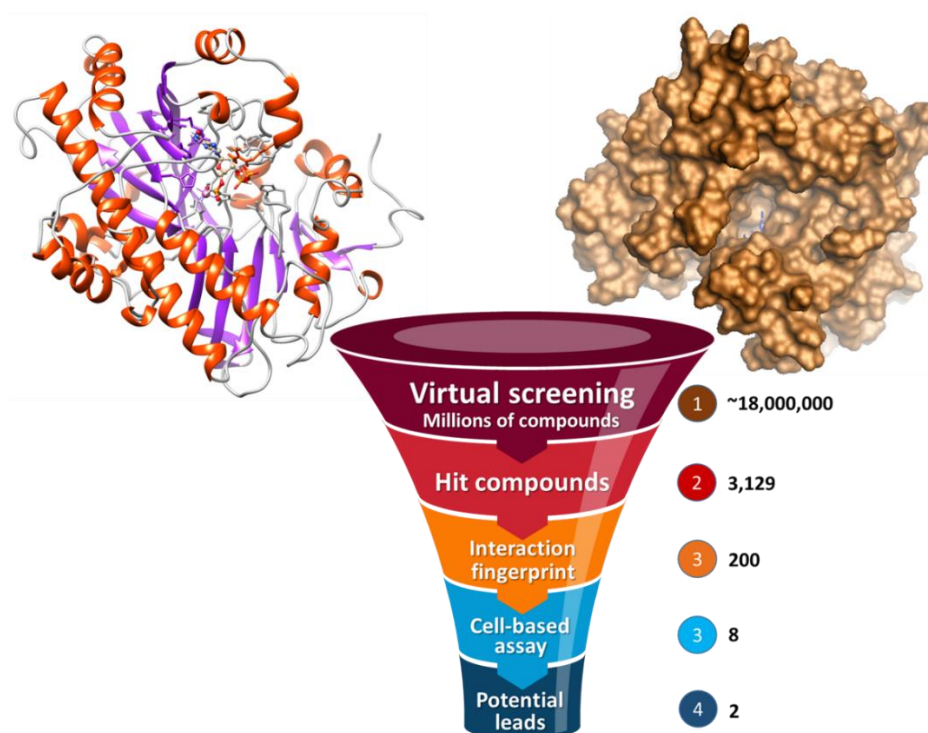
mediated diseases (Zhu *et al.*, 2015). In continuation from chapter 3 of the thesis where a 3D predicted structure of the *hRtcB* was developed, here in chapter 4, an attempt to discover inhibitors of *hRtcB* via VHTS has been made. The primary aim was to virtually screen small molecule compound libraries against the ligand binding site/active site (described in chapter 3 of the thesis) of the published *hRtcB* homology model (Nandy *et al.*, 2017) to identify ‘hit’ molecules. In the thesis, a ‘hit’ molecule is defined as being a compound which showed binding affinity to the ligand binding site of the *hRtcB* homology model in the VHTS campaign. Two libraries, namely, ZINC (Irwin *et al.*, 2012) and the CBCS database were used for the virtual screening purpose. ZINC represents a free database of over twenty million commercially available chemical compounds prepared especially for virtual screening purposes in ready-to-dock formats. Widely used by biologists and chemists in academic institutes, pharma, and biotechnology companies, the database hosts purchasable compounds from over one hundred vendors in 3D molecular representations which is easy to search and download for rapid testing of docking hypothesis. Other available databases, for instance, include PubChem (Li *et al.*, 2010), ChEMBL (Gaulton *et al.*, 2012), DrugBank (Knox *et al.*, 2011), and BindingDB (Liu *et al.*, 2007) but all of these lack ZINC’s unique focus on docking and purchasability. Another database called the ChemSpider (Ayers, 2012), although providing purchasing information, again lacks ZINC’s important focus on biologically relevant representation for docking (i.e., 3D molecular representations that actually bind to proteins). The CBCS database, on the other hand, contains six in-house databases that contain ~500,000 compounds of different types. Further, by docking these compound libraries onto the ligand binding site of the *hRtcB* homology model, the aim was to displace the natural ligand GTP/GMP (shown in Chapter 3) from the ligand binding site by the hit compounds to inhibit the catalytic activity of the ligase.

Importantly, from a clinical perspective, the ability of *hRtcB* to reduce chemotherapy-induced XBP1s upregulation (Logue *et al.*, 2018) was additionally explored. This was coupled with *hRtcB* localization studies to confirm its presence both in the nuclear and cytoplasmic compartments (Perez-Gonzalez *et al.*, 2014, Jurkin *et al.*, 2014a, Popow *et al.*, 2011).

## 4.2 Results

### 4.2.1 Hit discovery through the VHTS campaign

To screen potential inhibitors of *hRtcB*, a VHTS campaign was conducted utilizing the Schrödinger Software Release (2015-4). Molecular docking was performed using the *hRtcB* homology model developed in Chapter 3, as crystal structure for the human ligase is yet unavailable. Small molecule compound libraries from ZINC and the CBCS database were screened against the active site of the *hRtcB* homology model to identify potential inhibitors of the human ligase. Three levels of docking accuracy included the HTVS docking, SP docking, and the XP docking; each stage retaining the top 10% of binders. Filtering through each stage increased the flexibility of ligand and target area and the number of docked poses. Approximately eighteen million small molecule compounds were screened initially from which a range of top hit compounds were selected (selection method described in the next section 4.2.2) to be tested in cell-based assays to identify potential lead molecules. A methodical screening approach employed has been shown in Figure 4.1.



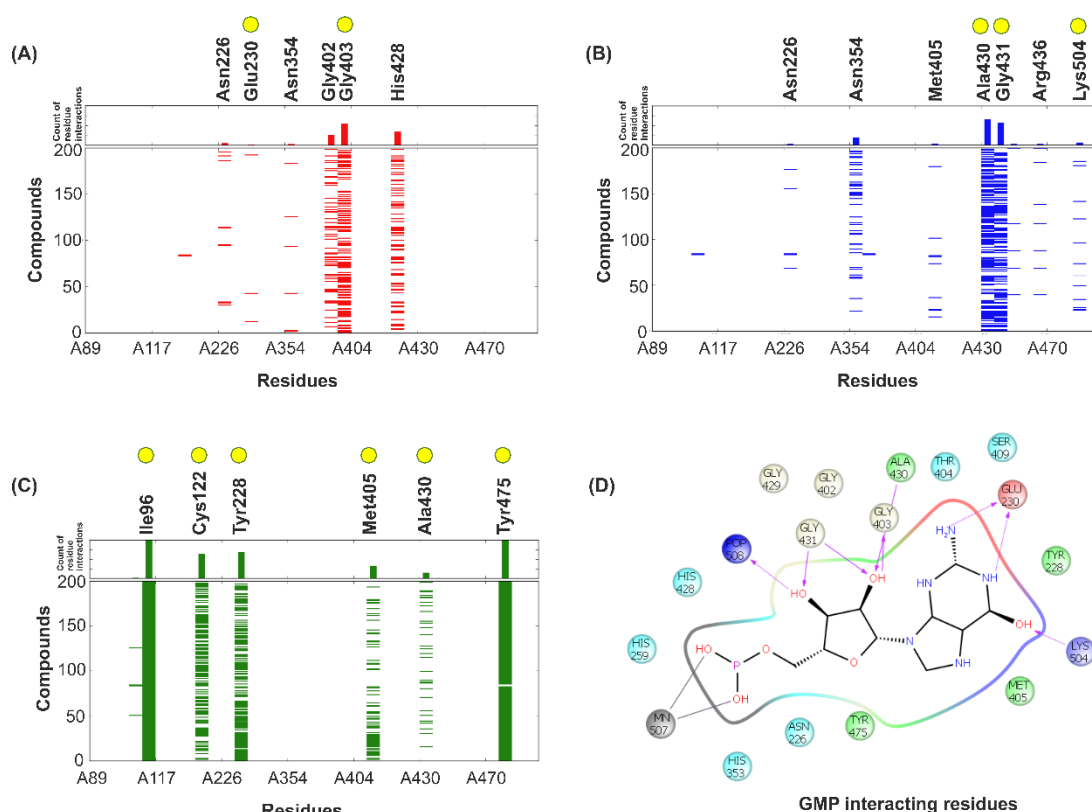
**Figure 4.1| The VHTS workflow.** Homology model of *hRtcB* was employed to screen approximately 18 million of small molecule compounds through the VHTS campaign to generate hits which were then further subjected to cell-based assays to identify potential lead candidates. The number of compounds

---

at each stage of the screening workflow (~18,000,000 compounds to 2 potential leads) have been indicated.

#### **4.2.2 Structural interaction fingerprint**

To understand the 3D structural protein-small molecule binding interactions, structural interaction fingerprints (SIFt) were generated (Figure 4.2). Through this method, interaction fingerprints are generated that translates 3D structural binding information from a protein-ligand complex into a one-dimensional binary string (Deng *et al.*, 2004, Singh *et al.*, 2006). Commonly used in VHTS campaigns, this method is used to filter, cluster, and analyze large datasets of the screening results. Using SIFt, the similarities and diversity of the binding modes of the small molecule compounds were identified, clustered, and compared to the binding interaction of the natural ligand GMP in the active site of the *hRtcB* homology model. During the VHTS process, this method served as an effective molecular filter to predict small molecule compounds with desirable binding mode(s) and/or desirable interaction patterns with the active site residues of the *hRtcB* homology model that was comparable to the binding pattern of GMP. From the ~eighteen million compounds screened, 3129 hits initially showed a binding affinity towards the active site of the *hRtcB* homology model. From these 3129 hits (ranging from a very high binding affinity to a low affinity), the top 200 highest scoring binders (based on binding affinity scores/binding free energy values ( $\Delta G$ ), docking scores), were subjected to the SIFt analysis. Important protein-small molecule intermolecular interactions that were assessed through SIFt include hydrogen bond acceptor (Figure 4.2A), hydrogen bond donor (Figure 4.2B) and hydrophobic interactions (Figure 4.2C). Additionally, the interaction of GMP with the active site residues in the *hRtcB* homology model was assessed (Figure 4.2D). As seen in the SIFts (Figure 4.2A, B, C), the majority of the top 200 small molecule compounds showed interactions with the active site residues with which GMP interacts in the *hRtcB* homology model.



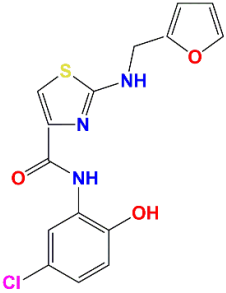
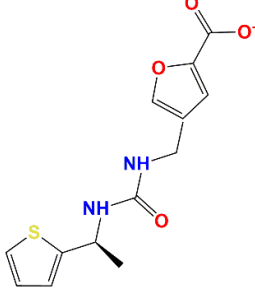
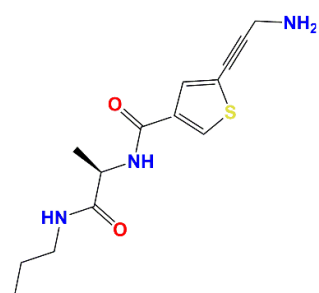
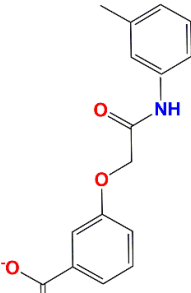
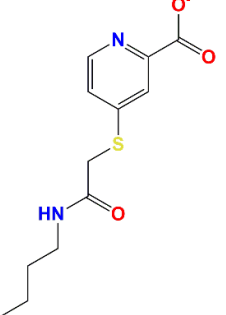
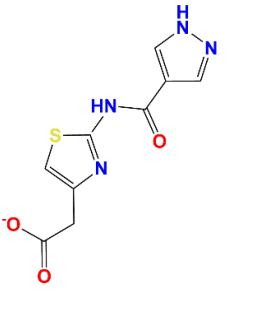
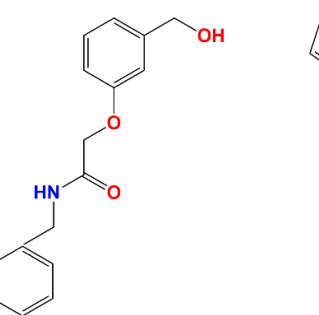
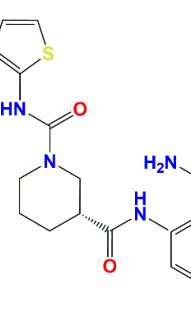
**Figure 4.2| Structural interaction fingerprint of top 200 small molecule compounds.** Predicted interaction fingerprint of top 200 compounds displaying interactions in form of (A) hydrogen bond acceptor, (B) hydrogen bond donor, and (C) hydrophobic interactions with residues in the active site of the *hRtcB* homology model. Yellow circles on top of each fingerprint denote the active site residues with which the natural ligand GMP interacts in the *hRtcB* homology model. (D) Ligand interaction diagram depicts the hydrogen bond interactions and metal coordination of the natural ligand GMP with the active site residues in the *hRtcB* homology model.

Following SIFT analysis, the top 100 highest scoring binders that show similar interactions to the GMP in the active site of the *hRtcB* homology model were selected. Thereafter, based on criteria including high, medium, and poor scoring binders along with structures with similar backbones but slightly different side chains or vice versa, the top 20 candidates were selected. Finally, a similar procedure was applied to narrow down the top 20 candidates to the top 8 hits (Table 4.1), with the inclusion of purchasability as an additional criterion. These top 8 hits were finally chosen to be evaluated in cell-based assays to narrow down to one or two candidate molecules to be considered as potential leads.



**Table 4.1| Top Hit compounds from the VHTS.** The large negative  $\Delta G$  scores represent the compounds with the highest binding affinity.

### Top eight predicted hit compounds

			
Compound-001 $\Delta G$ : -72.237 kcal/mol Cat. No.: STK931298 ZINC ID: 44891857 Mw: 349.799 g/mol	Compound-002 $\Delta G$ : -76.432 kcal/mol Cat. No.: BBV-43342639 ZINC ID: 94555779 Mw: 293.324 g/mol	Compound-003 $\Delta G$ : -76.217 kcal/mol Cat. No.: BBV-42682379 ZINC ID: 93976050 Mw: 293.392 g/mol	Compound-004 $\Delta G$ : -52.313 kcal/mol Cat. No.: BBV-32824904 ZINC ID: 13418713 Mw: 284.291 g/mol
			
Compound-005 $\Delta G$ : -68.066 kcal/mol Cat. No.: BBV-42229870 ZINC ID: 93518111 Mw: 267.33 g/mol	Compound-006 $\Delta G$ : -36.482 kcal/mol Cat. No.: BBV-25193082 ZINC ID: 36778224 Mw: 251.247 g/mol	Compound-007 $\Delta G$ : -71.032 kcal/mol Cat. No.: BBV-212872 ZINC ID: 20368557 Mw: 271.316 g/mol	Compound-008 $\Delta G$ : -75.686 kcal/mol Cat. No.: 1700614-Q093 ZINC ID: 49428955 Mw: 372.45 g/mol

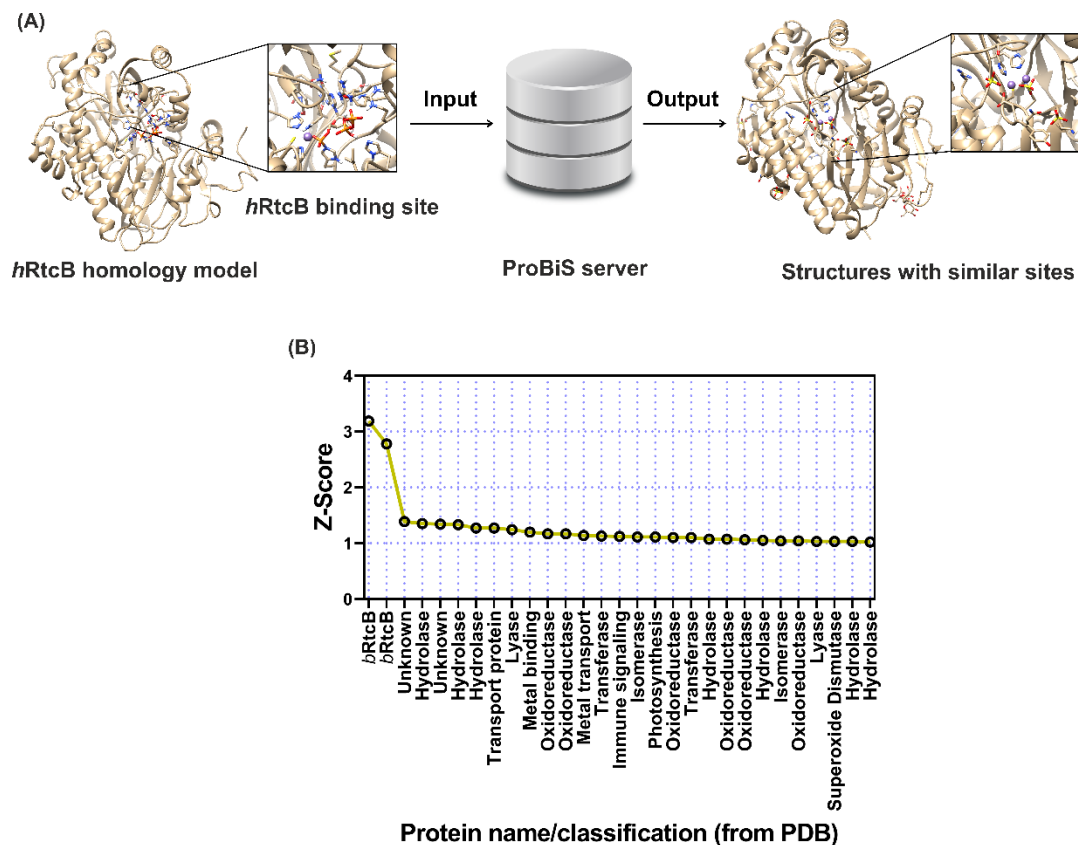
$\Delta G$ : Binding free energy; **Cat. No.:** Catalogue number; **Mw:** Molecular weight

### 4.2.3 Protein binding site similarity

Binding sites in proteins play a major role towards protein function. Identification of binding sites is essential for understanding their interactions with ligands, including other proteins. Binding sites can also contribute towards conservation of structure (Keskin *et al.*, 2005). Even in absence of sequence similarity, structural similarity between two proteins, especially in the binding site/active site can suggest similar functions. From a drug discovery perspective these can in turn suggest similar interactions with small molecule compounds. Detection of structurally similar proteins is therefore essential in

---

predicting side-effects/off-target effects of known drugs that bind to our protein target of interest. The ProBiS server was therefore employed to identify sites in other proteins that are structurally similar to that of the GTP binding site in the *hRtcB* homology model (Figure 4.3A). ProBiS is an algorithm that detects structurally similar sites on protein surfaces by local surface structure alignment (Konc and Janezic, 2010a). By comparing a query protein to a database of non-redundant protein 3D structures (nr-PDB), the server detects structurally similar sites as patterns of physicochemical properties on the protein surface (Konc and Janezic, 2010b, Konc *et al.*, 2012). Local structural alignments are then ranked by a Z-score, providing a statistical and structural significance of protein similarity. Z-scores >2 represent protein structures that are significantly similar to the query protein. Only Z-scores >1 are shown in the database user interface. As seen in Figure 4.3B, the GTP-binding site of the *hRtcB* homology model showed a high similarity Z-score with the first two proteins which were the *bRtcB* structures from the PDB, one from *P. horikoshii* and the other from *T. thermophilus* and then there was a drop in Z-score for proteins from other families that harbor similar sites to the GTP-binding site of the *hRtcB* homology model (Figure 4.3B). Only 29 proteins (including 3 from humans) from the PDB (Table 4.2) showed a structurally similar site to the GTP-binding site of the *hRtcB* homology model which depicts the uniqueness of the *hRtcB* homology model as a query protein.



**Figure 4.3| Binding site similarity.** (A) ProBiS workflow. (B) Structural similarity profile for *hRtcB* query protein homology model. Proteins that harbor structurally similar sites to the GTP-binding site of the *hRtcB* homology model are ranked by Z-scores.

**Table 4.2|** Proteins with similar sites from PDB. Z-scores >2 represent protein structures that are significantly similar to the query protein.

PDB ID	Protein name	Classification	Organism	Z-score
4DWR	RtcB	Ligase	Bacteria	3.19
2EPG	RtcB-like	Ligase	Bacteria	2.78
2AZ4	Hypothetical protein EF2904	Unknown	Bacteria	1.39
2XR1	Cleavage and polyadenylation specificity factor	Hydrolase	Archaea	1.35
2V9T	Roundabout homolog 1	Structural protein	Human	1.34
4UWP	Metallo-beta-lactamase vim-26	Hydrolase	Bacteria	1.33
4OJX	3',5'-cyclic-nucleotide phosphodiesterase 1	Hydrolase	Yeast	1.27

---

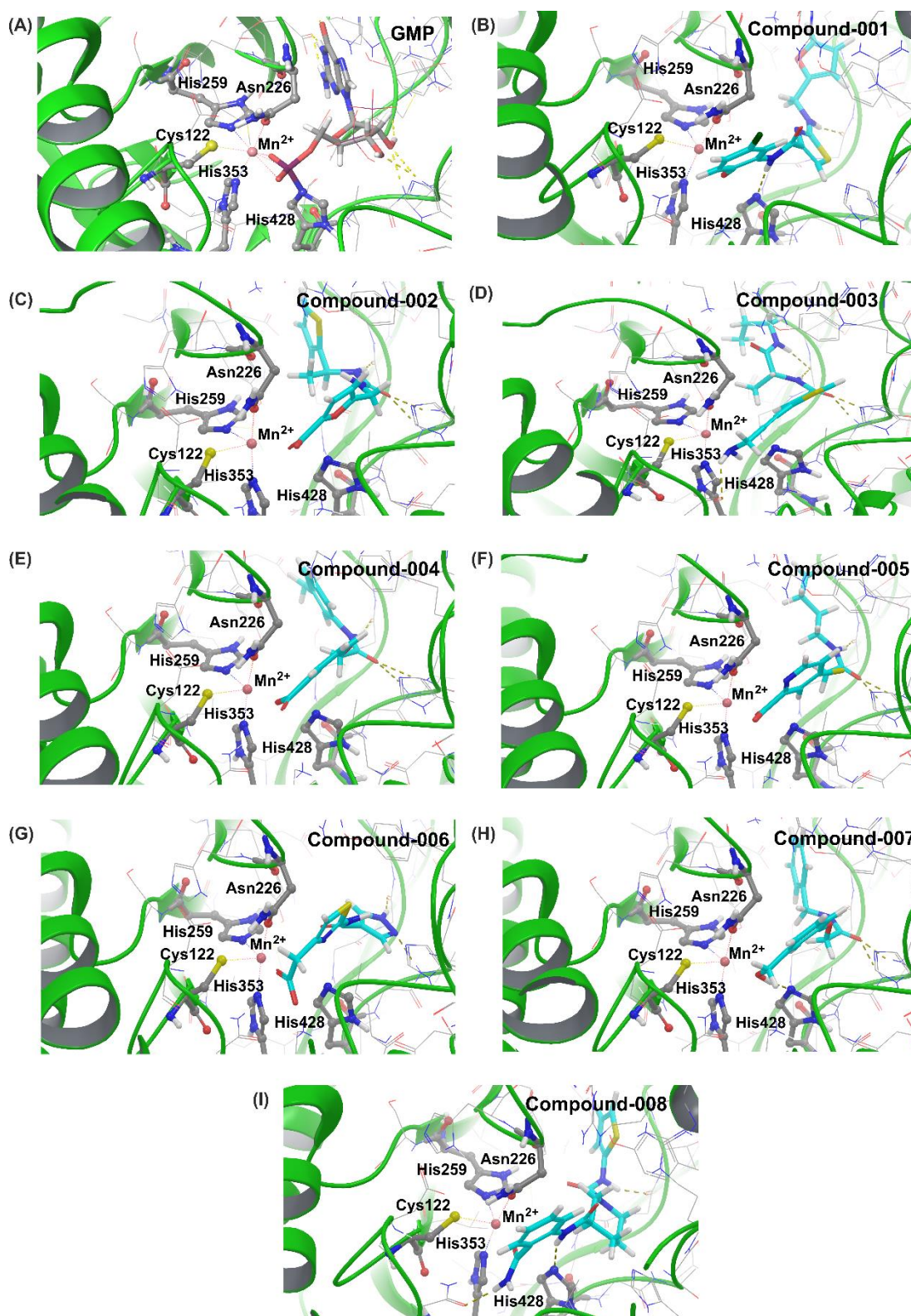
3T6O	Sulfate transporter	Transport protein	Bacteria	1.27
4IZO	Phosphoribosyl-aminoimidazole carboxylase	Lyase	Bacteria	1.24
4WWF	Nickel and cobalt resistance protein CNRR	Metal binding protein	Bacteria	1.20
3RF7	Iron-containing alcohol dehydrogenase	Oxidoreductase	Bacteria	1.17
4M3H	Copper oxidase	Oxidoreductase	Bacteria	1.17
2K70	Uncharacterized protein TTHA1943	Metal transport	Bacteria	1.14
3T6S	CERJ	Transferase	Bacteria	1.13
3RZF	MGC80376 protein	Immune system signaling protein	Frog	1.12
4MSP	Peptidyl-prolyl cis-trans isomerase FKBP14	Isomerase	Human	1.11
2LNJ	Uncharacterized protein SLL1418	Photosynthesis	Bacteria	1.11
3KW8	Copper oxidase	Oxidoreductase	Bacteria	1.10
4DU5	PFKB	Transferase	Bacteria	1.10
5AHO	5' exonuclease apollo	Hydrolase	Human	1.07
4HD4	Tyrosinase	Oxidoreductase	Bacteria	1.07
3PQA	Lactaldehyde dehydrogenase	Oxidoreductase	Archaea	1.06
3ZQ4	Ribonuclease J1	Hydrolase	Bacteria	1.05
3NVL	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	Isomerase	Protist	1.04
3KBE	Superoxide dismutase	Oxidoreductase	Worm	1.04
3O3M	2-hydroxyisocaproyl-coa dehydratase	Lyase	Bacteria	1.03
2APS	Cu, Zn superoxide dismutase	Superoxide dismutase	Bacteria	1.03
4J6O	Metallophosphoesterase	Hydrolase	Bacteria	1.03
1Y44	Ribonuclease Z	Hydrolase	Bacteria	1.02

---

---

#### ***4.2.4 Analysis of the top eight hit compounds***

The binding patterns of GMP (Chapter 3) and the top eight candidate compounds bound to the active site of the *hRtcB* homology model as predicted by MM-GBSA have been shown in Figure 4.4. Binding pose analysis revealed GMP sharing two hydrogen bonds with Glu230, one hydrogen bond with Gly403, one hydrogen bond with Ala430, two hydrogen bonds with Gly431, and one hydrogen bond with Lys504 (Figure 4.4a). Docking pose analysis revealed two hydrogen bonds between compound-001 and His428 and Gly403 with a  $\pi$ - $\pi$  interaction with His428; three hydrogen bonds between compound-002 and Gly403, Ala430, and Gly431 with two  $\pi$ - $\pi$  interactions with His89, and Tyr475; three hydrogen bonds between compound-003 and Asn354, Ala430, and Gly431 and also forming two hydrogen bonds with Gly403; three hydrogen bonds between compound-004 and Gly403, Ala430, and Gly431 with a  $\pi$ - $\pi$  interaction with His89; three hydrogen bonds between compound-005 and Gly403, Ala430, and Gly431 with a  $\pi$ - $\pi$  interaction with His89; two hydrogen bonds between compound-006 and Gly403, and Ala430 with two  $\pi$ - $\pi$  interactions with His89, and His428; four hydrogen bonds between compound-007 and Gly403, His428, Ala430, and Gly431 with two  $\pi$ - $\pi$  interactions with His89, and Tyr475; three hydrogen bonds between compound-008 and Asn354, Gly403, and His428.



**Figure 4.4| Predicted binding patterns of GMP and the eight hit compounds in the *hRtcB* active site.** Predicted binding interactions formed by (A) GMP and (B-I) compounds-001 to -008 in the active site. GMP is represented in the grey stick. The compounds are represented in cyan sticks. The green ribbon diagram represents the *hRtcB* predicted model to which the GMP is bound, and the compounds were docked. Yellow dashed lines represent hydrogen bonds. Key interacting residues in the *hRtcB* ligand binding site are displayed in the ball and stick representation. The pink ball at the center depicts

---

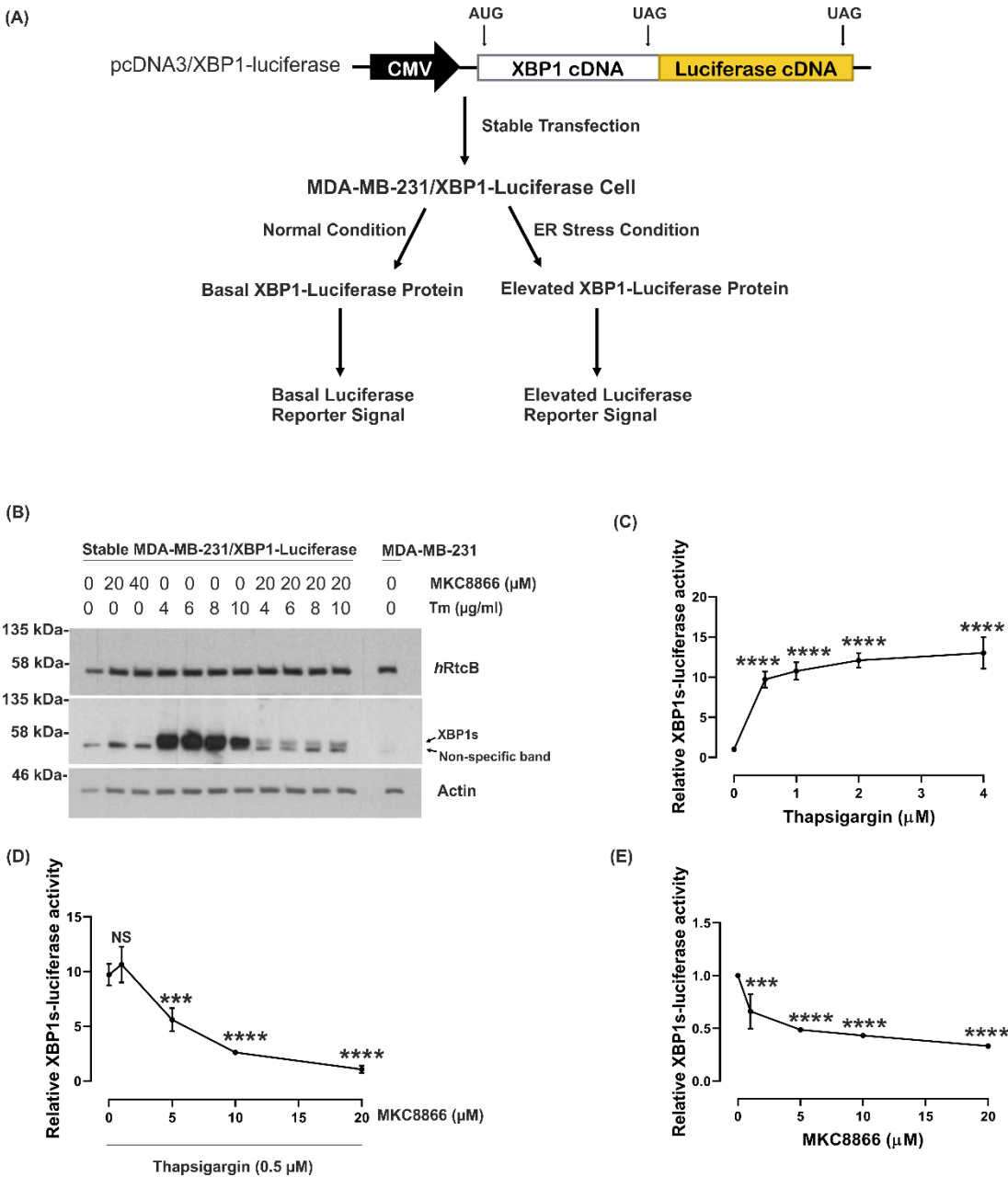
the Mn<sup>2+</sup> metal cofactor. The binding patterns of the compounds to hRtcB were obtained by Prime MM-GBSA of the Schrödinger Software Release (2015-4). Prime MM-GBSA is a physics-based approach that calculates the force field energies of the bound and unbound states of the protein-ligand complex (Sirin *et al.*, 2014).

#### **4.2.5 Establishment of an HTS assay for screening inhibitors against XBP1 splicing**

To establish an HTS assay for screening inhibitors against XBP1s, XBP1-Luciferase/pcDNA3 plasmid was stably transfected into MDA-MB-231 cells, and MDA-MB-231/XBP1-luciferase cells were isolated. The XBP1-Luciferase/pcDNA3 plasmid consists of full length human XBP1 cDNA fused upstream of luciferase cDNA and was a kind gift from Etsu Tashiro (Tashiro *et al.*, 2007). As per information provided by the Etsu laboratory, in this plasmid, the stop codon of XBP1 was eliminated and placed immediately upstream from the luciferase coding sequence. Under normal conditions (Figure 4.5A), luciferase is not expressed, as the XBP1-luciferase mRNA translation is terminated at the stop codon situated upstream of the luciferase mRNA. However, under conditions of ER stress, removal of the 26-nucleotide intron leads to a frame shift as a result of which the XBP1-luciferase mRNA translation is terminated at the stop codon of the luciferase mRNA that leads to the expression of the full XBP1-luciferase fused protein. As shown in Figure 4.5A, C, D, and E, under normal conditions, basal XBP1s-luciferase activity was detected in MDA-MB-231/XBP1-luciferase cells due to the basal UPR activity. To optimize treatment conditions, MDA-MB-231/XBP1-luciferase cells were treated with the ER stress-inducing agent Tm (Figure 4.5B) along with the IRE1 $\alpha$  RNase inhibitor MKC8866 for 6 hours. As seen in the blot, elevated levels of endogenous XBP1s was observed with the Tm which was reduced upon addition of the MKC8866. The size of the luciferase coding sequence within the plasmid is about 1,650 bp (as per information provided) and therefore the luciferase protein should be ~61 kDa. This would further mean that the XBP1s-luciferase fusion protein should be ~120 kDa. However, in our experiments the XBP1s monoclonal antibody used did not detect any XBP1s-luciferase fusion protein at the predicted molecular weight in the western blots. In Figure 4.5C, cells were treated with increasing doses of another ER stress-inducing agent Tg which significantly elevated (~10 to 14-fold) the XBP1s-luciferase activity. It was evident that 0.5  $\mu$ M Tg was enough to significantly induce the XBP1s-luciferase



activity. Hence in Figure 4.5D, cells were subjected to 0.5  $\mu$ M Tg treatment to which the MKC8866 was added in a dose-dependent manner. As expected, when cells were subjected to Tg treatment, XBP1s-luciferase activity was significantly elevated ( $\sim$ 10-fold) which was reduced to almost basal levels upon addition of the MKC8866. Indeed, the addition of MKC8866 alone significantly reduced the basal IRE1 $\alpha$  RNase activity again in a dose-dependent manner (Figure 4.5E); suggesting that the HTS platform was well suited for screening compounds.



**Figure 4.5| HTS platform for screening inhibitors against XBP1 splicing.** (A) Illustration of the HTS platform for screening inhibitors. MDA-MB-231/XBP1-luciferase cells were treated for 6 hours

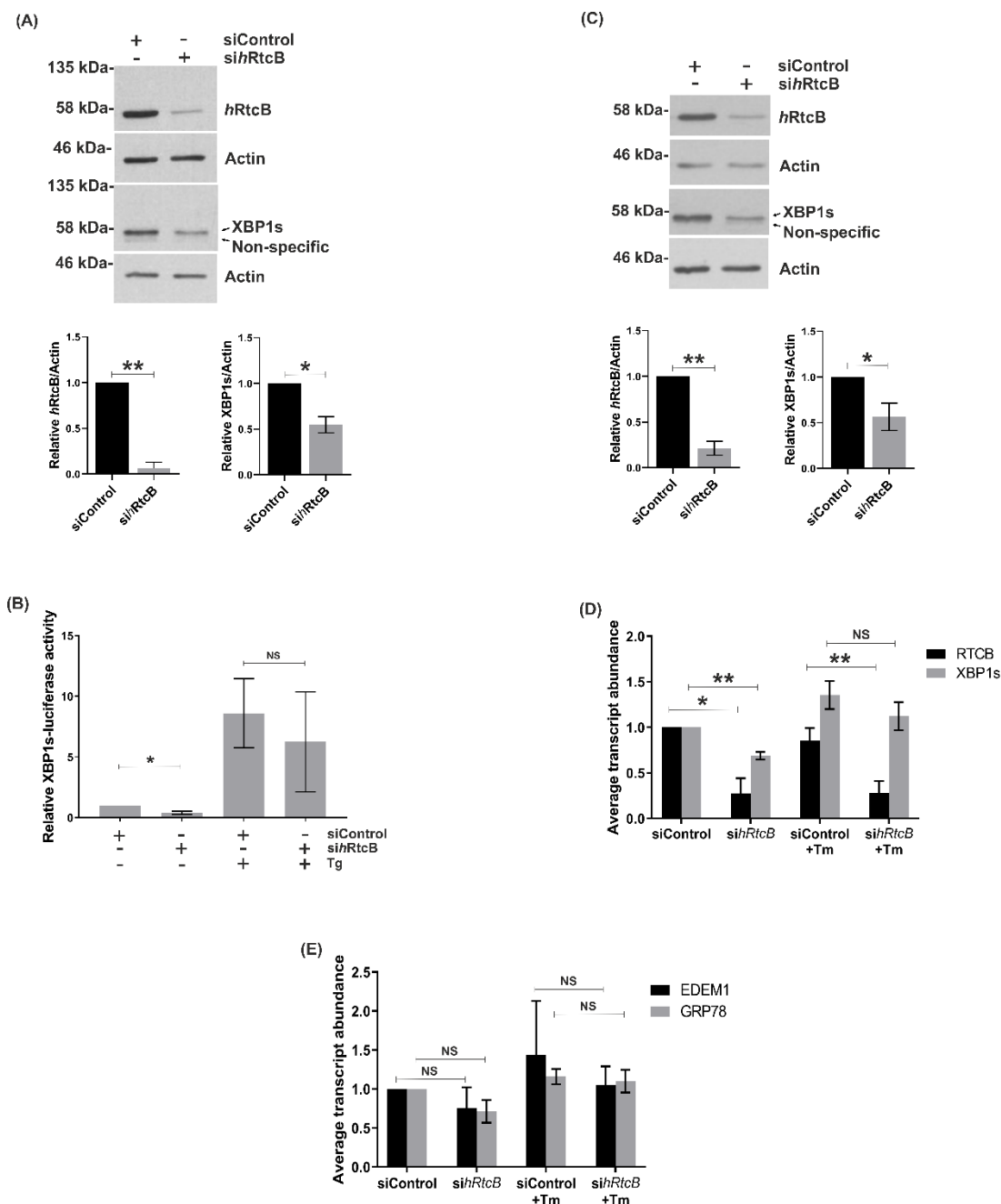


---

with **(B)** increasing doses of MKC8866, Tm, or Tm in combination with 20  $\mu$ M MKC8866, **(C)** increasing doses of Tg alone, **(D)** 0.5  $\mu$ M Tg in combination with increasing doses of MKC8866, and **(E)** increasing doses of MKC8866 alone ( $n = 3$ , mean XBP1s-luciferase activities and SD are displayed). XBP1s-luciferase reporter activity was normalized to the total protein content of each sample and to the DMSO control sample. One-way ANOVA multiple comparisons were used to analyze the statistical significance of differences in reporter activity between control and treated cells (\* $p < 0.05$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , NS is non-significant).

#### **4.2.6 Depletion of *hRtcB* significantly reduces XBP1s in MDA-MB-231 /XBP1-luciferase and MDA-MB-231 cells**

To further confirm that *hRtcB* catalyzes re-ligation of cleaved *XBP1u* mRNA ends to generate *XBP1s*, *hRtcB* was depleted both in MDA-MB-231/XBP1-luciferase and in parental MDA-MB-231 cells by treating with control siRNA and siRNA against *hRtcB* for 96 hours. Compared to control cells, while depletion of *hRtcB* significantly reduced basal expression of endogenous XBP1s (Figure 4.6A) and basal XBP1s-luciferase activity (Figure 4.6B), it did not yield any reduction of XBP1s-luciferase activity under conditions of ER stress (Figure 4.6B). Similarly, in the MDA-MB-231 cells, *hRtcB* depletion compared to control cells, significantly reduced basal XBP1s expression both at the protein and mRNA levels (Figure 4.6C, D) and further did not yield any reduction in *XBP1s* and its target genes *EDEM1* and *GRP78* (Lee *et al.*, 2003a) at the mRNA level under ER stress (Figure 4.6D, E). Although both the western blots from the two cell lines (Figure 4.6A, C) represent endogenous XBP1s reduction, similar levels of XBP1s reduction as observed in both cell lines confirmed that the engineered cell line works similar to the parental MDA-MB-231 cells with respect to their regulation of XBP1s. Also, a direct comparison between the results obtained in Figure 4.6 B, and D was not possible as two different ER stress-inducing agents were utilized. Tg was used to induce ER stress in MDA-MB-231/XBP1-luciferase cells (Figure 4.6B) and Tm was used to induce ER stress in MDA-MB-231 cells (Figure 4.6D).



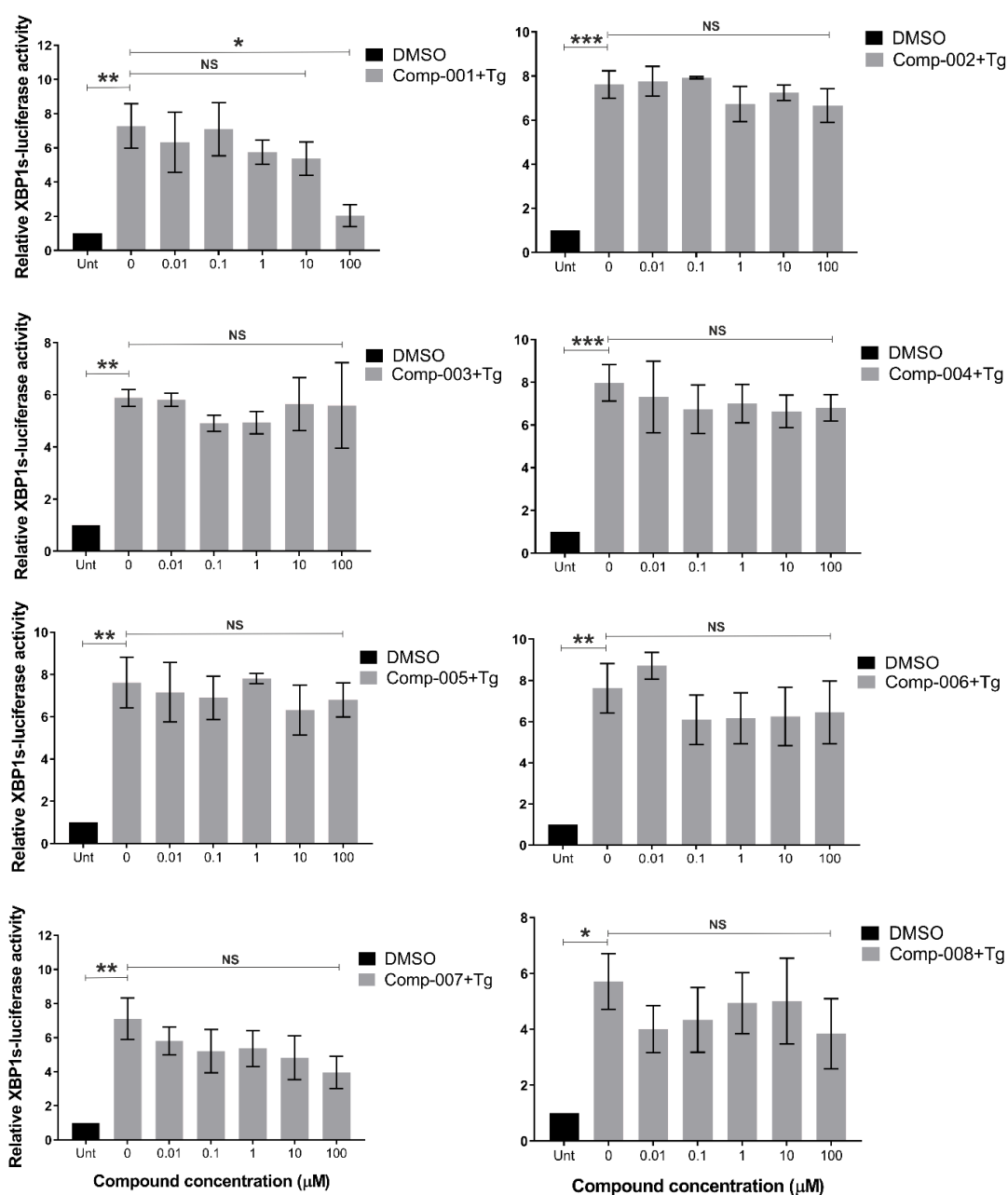
**Figure 4.6| Depletion of *hRtcB* in MDA-MB-231/XBP1-luciferase and MDA-MB-231 cells.** (A, B) MDA-MB-231/XBP1-luciferase cells were treated with 10 nM of both control siRNA and siRNA against *hRtcB* for 96 hours followed by 0.5  $\mu$ M Tg treatment for an additional 6 hours ( $n = 3$ , mean XBP1s-luciferase activity and SD are displayed). XBP1s-luciferase reporter activity was normalized as stated in earlier section 4.2.5. Blots are representative of 3 independent experiments. Levels of *hRtcB* and XBP1s protein were quantified by densitometry and normalized to the housekeeping protein (Actin) ( $n = 3$ , mean expression levels and SD are displayed). Paired two-tailed  $t$ -test was used to analyze the statistical significance of differences between control and treated cells (\* $p < 0.05$ , \*\* $p < 0.01$ , NS is non-significant). (C, D) MDA-MB-231 cells were treated with 10 nM of both control siRNA and siRNA against *hRtcB* for 96 hours followed by 0.5  $\mu$ g/ml Tm treatment for an additional 6 hours. Blots are representative of 3 independent experiments. Levels of *hRtcB* and XBP1s protein were quantified

---

and normalized to the housekeeping protein and analyzed by densitometry ( $n = 3$ , mean expression levels and SD are displayed; Paired two-tailed  $t$ -test was used to analyze the statistical significance of differences between control and treated cells ( $*p < 0.05$ ,  $**p < 0.01$ , NS is non-significant). (E) Samples are same as in (D). Relative mRNA levels of *hRtcB*, *XPB1s*, *EDEM1*, and *GRP78* were analyzed by RT-qPCR. Expression levels were normalized to the housekeeping gene *MRPL19* mRNA levels and to the untreated control sample ( $n = 3$ , mean expression levels and SD are displayed; Paired two-tailed  $t$ -test was used to analyze the statistical significance of differences in mRNA levels between control and *hRtcB*-depleted cells ( $*p < 0.05$ ,  $**p < 0.01$ , NS is non-significant).

#### **4.2.7 Hit compound-001 reduces ER stress-induced XBP1 splicing in MDA-MB-231/XBP1-luciferase cells**

The top 8 hit compounds identified from the VHTS campaign as potential inhibitors of *hRtcB* were screened through the XBP1s-luciferase reporter assay towards their potential to reduce the XBP1s-luciferase activity. MDA-MB-231/XBP1-luciferase cells were pre-treated with compounds-001 to -008 for 1 hour at 37 °C followed by 0.5  $\mu$ M Tg treatment for an additional 6 hours. The screening was performed in a 96-well plate format at 6 doses – 0.01, 0.1, 1, 10 and 100  $\mu$ M (Figure 4.7). Each plate had a negative control (100% inhibition; cells alone without Tg and compound) and a positive control (0% inhibition; cells with 0.5  $\mu$ M Tg).  $Z'$  factor was calculated to be 0.79, indicating that the assay was robust for 96-well plate HTS. Hits that yielded ~50% inhibition activities of positive control were considered as potential lead molecules. Hit compound-001 yielded a significant (~5.2 fold) reduction of the XBP1s-luciferase. The other seven compounds that were tested in the assay did not show any reduction in the XBP1s-luciferase activity. Therefore, compound-001 was chosen to be re-tested in further assays in the thesis. Additionally, compound-007 was also selected to be re-tested.

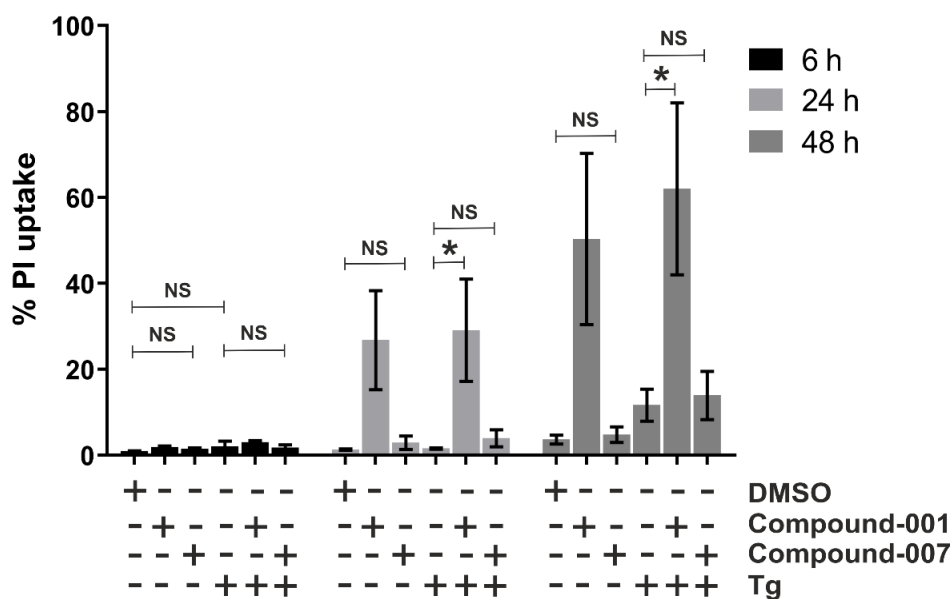


**Figure 4.7| Hit compound-001 reduces ER stress-induced XBP1s-luciferase activity.** MDA-MB-231/XBP1-luciferase cells were pre-treated with 0.01, 0.1, 1, 10 and 100 μM of compounds-001 to -008 for 1 hour followed by 0.5 μM Tg treatment for an additional 6 hours (for compounds-001 and -007,  $n = 4$ , and compounds-002, -003, -004, -005, -006, -008,  $n = 3$ , mean XBP1s-luciferase activity and SEM are displayed). XBP1s-luciferase reporter activity was normalized as stated in earlier section 4.2.5. One-way ANOVA multiple comparisons were used to analyze the statistical significance of differences in reporter activity between control and treated cells (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , NS is non-significant).

---

#### ***4.2.8 Compounds-001 and -007 are not toxic at 7 hours of treatment under ER stress***

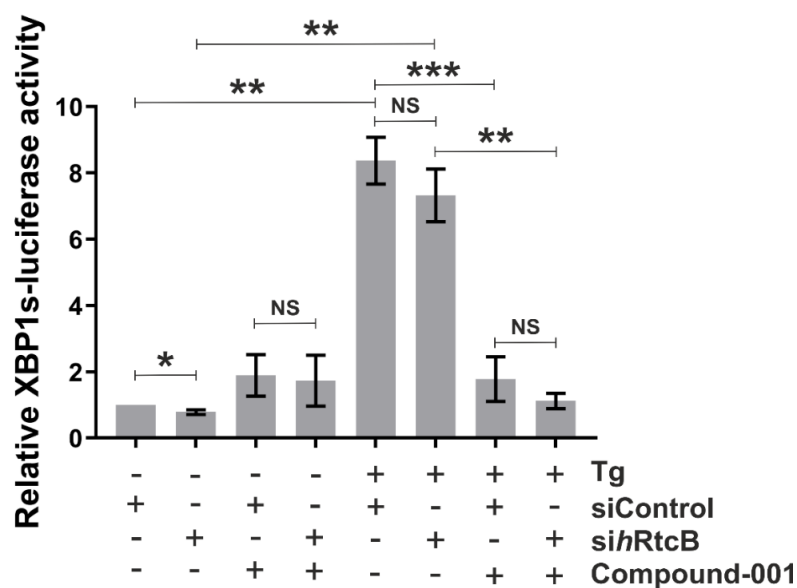
To determine any toxicity associated with higher doses of the compounds, both the potential lead candidates were tested towards their toxic effects on the cells using Hoechst/PI staining at the treatment time of 6 hours and beyond. As shown in Figure 4.8, MDA-MB-231/XBP1-luciferase cells when treated with compounds-001 and -007 alone at 100  $\mu$ M each for 6 hours, proved to be not toxic (1.5 to 2% increase in dead cells) when compared to DMSO control. In a similar way, cells treated with these two compounds at 100  $\mu$ M each for 6 hours under ER stress were also not toxic (1.5 to 3% increase in dead cells) compared to control cells (Tg alone); suggesting that the inhibition of Tg-induced XBP1s-luciferase activity as seen in the previous sections was not due to any toxic effect of the compounds. However, extending the toxicity evaluation up till 48 hours of treatment time, while compound-007 did not impart any toxic effects, compound-001 proved to be toxic to the cells at later times (24 and 48 hours). Also, the level of cell death at 48 hours in cells treated with Tg alone was low. This could be due to the higher resistance of these engineered cells towards Tg. Although re-analysis of the data to include small debris/apoptotic bodies could have resulted in increased cell death, this was however not possible since the experiment was performed using the operetta high-Content imaging system and not the flow cytometry technique. Therefore, the aforesaid statement of higher resistance discussed is presented as a fact rather than a possibility and is not substantiated by any evidence. In the thesis, any cell death above 10% is considered as toxic.



**Figure 4.8| Toxicity of hit compounds-001 and -007.** MDA-MB-231/XBP1-luciferase cells were pre-treated with 100  $\mu$ M of compounds-001 and -007 for 1 hour followed by 0.5  $\mu$ M Tg treatment for an additional 6 hours. Cells were harvested at 7, 25, and 49 hours respectively. Toxicity assessments were performed by Hoechst/PI staining assay. ( $n = 3$ , mean PI uptake levels and SEM are displayed). One-way ANOVA multiple comparisons were used to analyze the statistical significance of differences in PI uptake between control and treated cells (\* $p < 0.05$ , NS is non-significant).

#### 4.2.9 Compound-001 does not inhibit *hRtcB* in the IRE1-XBP1 axis in MDA-MB-231/XBP1-luciferase cells

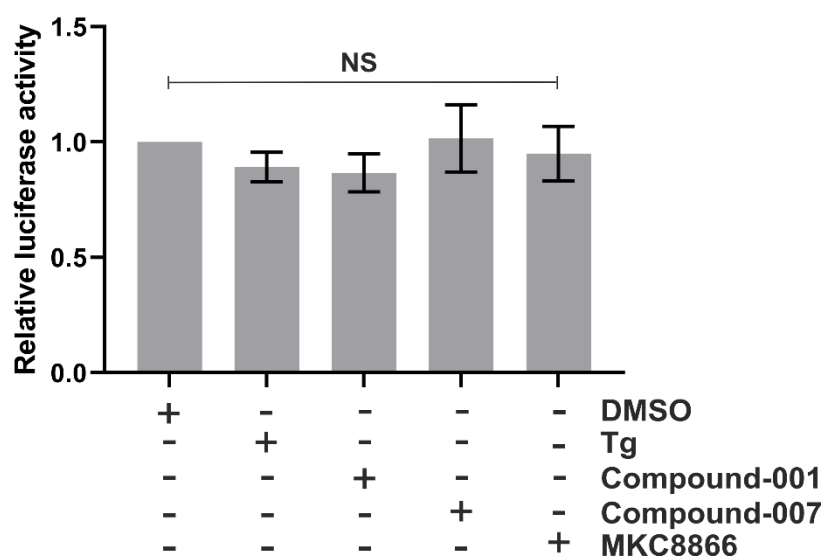
To understand the probable course of action of compound-001 at least at the molecular level, *hRtcB* was first depleted in MDA-MB-231/XBP1-luciferase cells until 96 hours. Following depletion, cells were pre-treated with 100  $\mu$ M compound-001 for 1 hour and further treated with 0.5  $\mu$ M Tg for an additional 6 hours. *hRtcB* depletion (Figure 4.9) yielded a statistically significant reduction (22%) of the basal XBP1s-luciferase and the addition of compound-001 in depleted cells under stress significantly reduced (~7-fold) the XBP1s-luciferase activity compared to depleted cells under Tg treatment alone. Importantly, the addition of compound-001 in *hRtcB* depleted cells both with and without stress did not reduce the XBP1s-luciferase activity when compared to control cells treated with the compound with and without ER stress.



**Figure 4.9| Compound-001 does not inhibit *hRtcB*.** MDA-MB-231/XBP1-luciferase cells were treated with 10 nM of both control siRNA and siRNA against *hRtcB* for 96 hours followed by pre-treatment with 100  $\mu$ M compound-001 for 1 hour plus 0.5  $\mu$ M Tg treatment for an additional 6 hours ( $n = 3$ , mean XBP1s-luciferase activity and SD are displayed). XBP1s-luciferase reporter activity was normalized as stated in earlier section 4.2.5 Unpaired two-tailed *t*-test with Welch's correction was used to analyze the statistical significance of differences in reporter activity between control and treated cells as well as within the different treatments as indicated (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , NS is non-significant).

#### 4.2.10 Compounds-001 and -007 do not inhibit luciferase

To further rule out the possibility that reduction in XBP1s-luciferase activity by compounds-001 and -007 is not due to the compounds' inhibitory effect on the luciferase enzyme itself, luciferase/pcDNA3.1 plasmid was transiently transfected in MDA-MB-231 cells. Following 24 hours of transfection, transfected cells were treated with 0.5  $\mu$ M Tg, 100  $\mu$ M of compounds-001, -007 and 20  $\mu$ M MKC8866 for 7 hours. As expected, compounds-001 and -007 did not show any inhibition of the enzymatic activity of the luciferase itself (Figure 4.10). Therefore, it was concluded that the reduction in XBP1s-luciferase activity by compound-001 and the trend in reduction of XBP1s-luciferase activity by compound-007 (as seen in earlier sections) was not due to the inhibition of luciferase activity by the compounds.

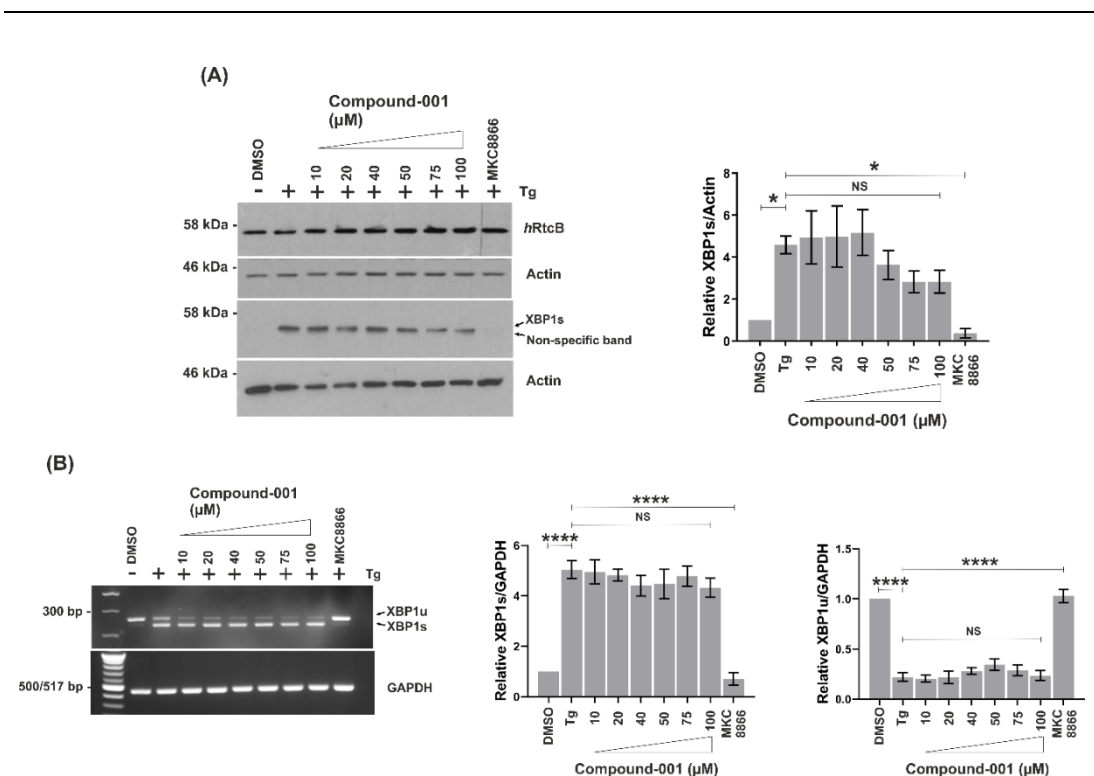


**Figure 4.10| Compounds-001 and -007 do not inhibit luciferase.** MDA-MB-231 cells were transiently transfected with 50 ng luciferase/pcDNA3.1 plus 50 ng PEGFP plasmid in independent wells of a 96 well plate. 24 hours post-transfection, all cells in the plate were treated with DMSO, 0.5  $\mu$ M Tg, 100  $\mu$ M compound-001, and 100  $\mu$ M compound-007 for 7 hours. 20  $\mu$ M MKC8866 served as an additional control ( $n = 3$ , mean luciferase activity and SD are displayed). Luciferase reporter activity was normalized as stated in earlier section 4.2.5. One-way ANOVA multiple comparisons were used to analyze the statistical significance of differences in reporter activity between control and treated cells (NS is non-significant).

#### 4.2.11 Compound-001 does not reduce XBP1s in MDA-MB-231 cells

Finally, to test whether compound-001 reduces XBP1s expression levels in the human TNBC cell model, MDA-MB-231 cells were pre-treated with compound-001 in a dose-dependent manner for 1 hour following which they were subjected to ER stress for an additional 6 hours using 0.5  $\mu$ M Tg. Changes in XBP1s expression levels were monitored both at the protein and mRNA level. Compared to the positive control (Tg alone), there was no reduction of XBP1s by compound-001 at the protein level (Figure 4.11A). No reduction in relative XBP1s/GAPDH or XBP1u/GAPDH was observed at the mRNA level with the compound compared to Tg alone (Figure 4.11B). A good inhibitor of *hRtcB* would be expected to reduce both the spliced and unspliced transcripts; reduced levels of the latter could be accounted for the subsequent degradation of the cleaved *XBP1u* mRNA due to the absence of re-ligation.

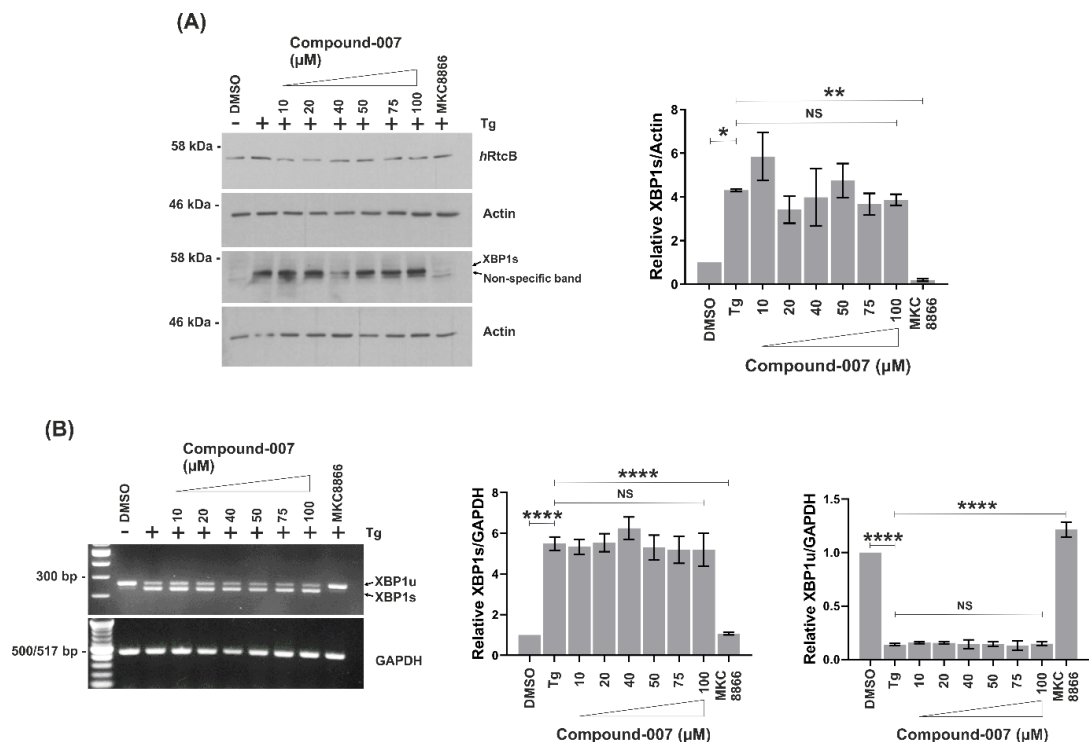




**Figure 4.11| Compound-001 does not reduce XBP1s in TNBC cell model.** (A, B) MDA-MB-231 cells were pre-treated with 10, 20, 40, 50, 75 and 100  $\mu$ M of compound-001 for 1 hour, followed by 0.5  $\mu$ M Tg treatment for an additional 6 hours. 20  $\mu$ M MKC8866 served as an additional control. Levels of XBP1s for the protein were quantified and normalized to the housekeeping protein/gene and analyzed by densitometry. Relative XBP1s/GAPDH and XBP1u/GAPDH for the mRNA was quantified and analyzed by densitometry ( $n = 3$ , mean expression levels and SEM are displayed). One-way ANOVA multiple comparisons were used to analyze the statistical significance of differences in expression levels between control and treated cells (\* $p < 0.05$ , \*\*\* $p < 0.001$ , NS is non-significant).

#### 4.2.12 Compound-007 does not reduce XBP1s in MDA-MB-231 cells

Since compound-007 showed a trend in the reduction of XBP1s-luciferase activity in figure 4.7, this compound was also tested in the MDA-MB-231 cells in a similar fashion to the compound-001. Cells were pre-treated with compound-007 in a dose-dependent manner for 1 hour following which they were subjected to ER stress for an additional 6 hours using 0.5  $\mu$ M Tg. Changes in XBP1s expression levels were monitored both at the protein and mRNA level. Compound-007 yielded no reduction in XBP1s and relative XBP1s/GAPDH or XBP1u/GAPDH at the protein and mRNA levels respectively compared to the positive control (Figure 4.12A, B).

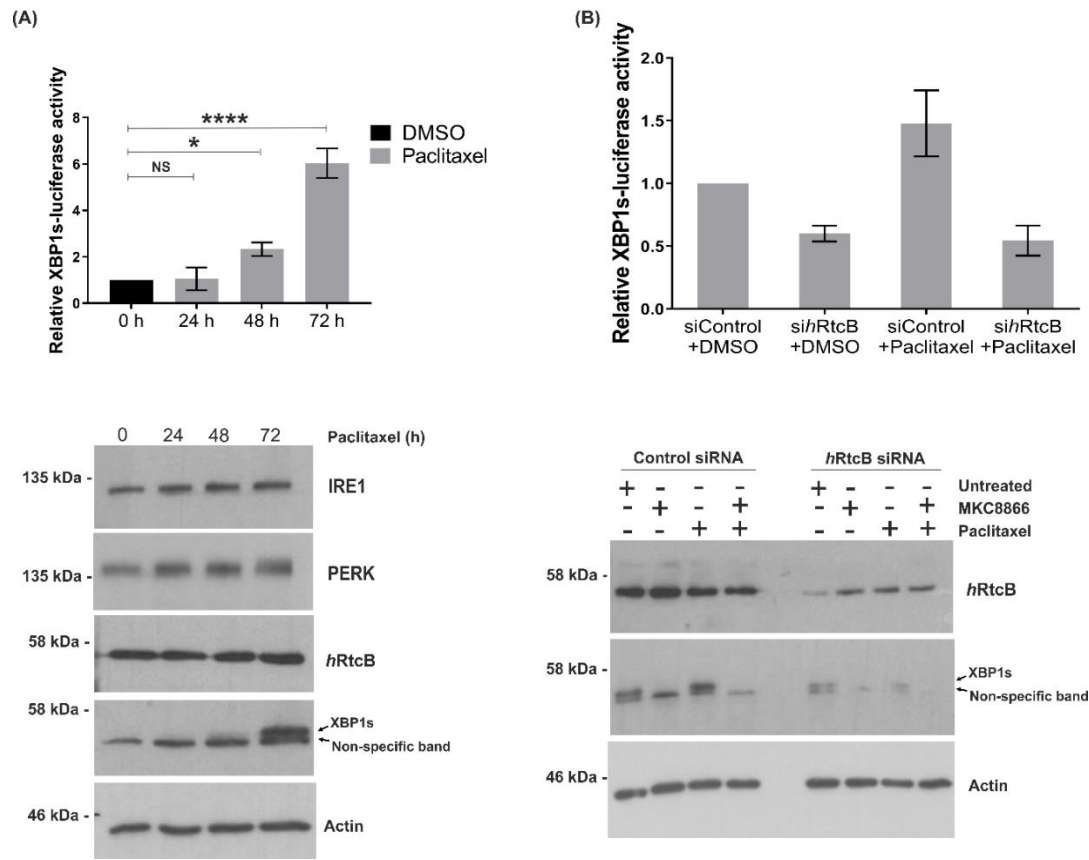


**Figure 4.12| Compound-007 does not reduce XBP1s in TNBC cell model.** (A, B) MDA-MB-231 cells were pre-treated with 10, 20, 40, 50, 75 and 100  $\mu$ M of compound-007 for 1 hour, followed by 0.5  $\mu$ M Tg treatment for an additional 6 hours. 20  $\mu$ M MKC8866 served as an additional control. Levels of XBP1s for the protein were quantified and normalized to the housekeeping protein/gene and analyzed by densitometry. Relative XBP1s/GAPDH and XBP1u/GAPDH for the mRNA was quantified and analyzed by densitometry ( $n = 3$ , mean expression levels and SEM are displayed). One-way ANOVA multiple comparisons were used to analyze the statistical significance of differences in expression levels between control and treated cells (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$ , NS is non-significant).

#### 4.2.13 Depletion of *hRtcB* reduces paclitaxel-induced XBP1 splicing in MDA-MB-231/XBP1-luciferase cells

Lastly, the ultimate goal of developing an *hRtcB* inhibitor is for eventual use in adjuvant/neo-adjuvant cancer therapy. A recent study from our group has shown that paclitaxel, a commonly used chemotherapy for the treatment of TNBC produces pro-tumorigenic factors in a manner partly dependent on IRE1 $\alpha$  RNase activity by elevating levels of XBP1s (Logue *et al.*, 2018). Since data presented in the thesis has revealed that depleting *hRtcB* modulates expression of XBP1s, hence, to investigate whether depleting *hRtcB* could also reduce paclitaxel-induced XBP1s expression, *hRtcB* was depleted in MDA-MB-231/XBP1-luciferase cells. MDA-MB-231/XBP1-luciferase cells were treated with 10 nM paclitaxel for 3 days, and elevated levels of XBP1s-luciferase activity was reported (Figure 4.13A). As seen in the western blot

analysis in Figure 4.13A, paclitaxel treatment at 72 hours significantly elevated levels of endogenous XBP1s. While *hRtcB* levels remained unaffected at 72 hours, PERK and IRE1 seemed to get elevated at 72 hours, suggesting that paclitaxel treatment also affects these two arms of the UPR. Following 72 hours of *hRtcB* depletion, cells were treated with 10 nM paclitaxel for 3 days. As expected, depletion of *hRtcB* reduced (~3-fold) the paclitaxel-induced XBP1s-luciferase activity compared to cells treated with control siRNA plus paclitaxel (Figure 4.13B). Similarly, western blot analysis in Figure 4.13B revealed that depletion of *hRtcB* besides reducing basal XBP1s also reduces paclitaxel-induced XBP1s compared to control cells endogenously.

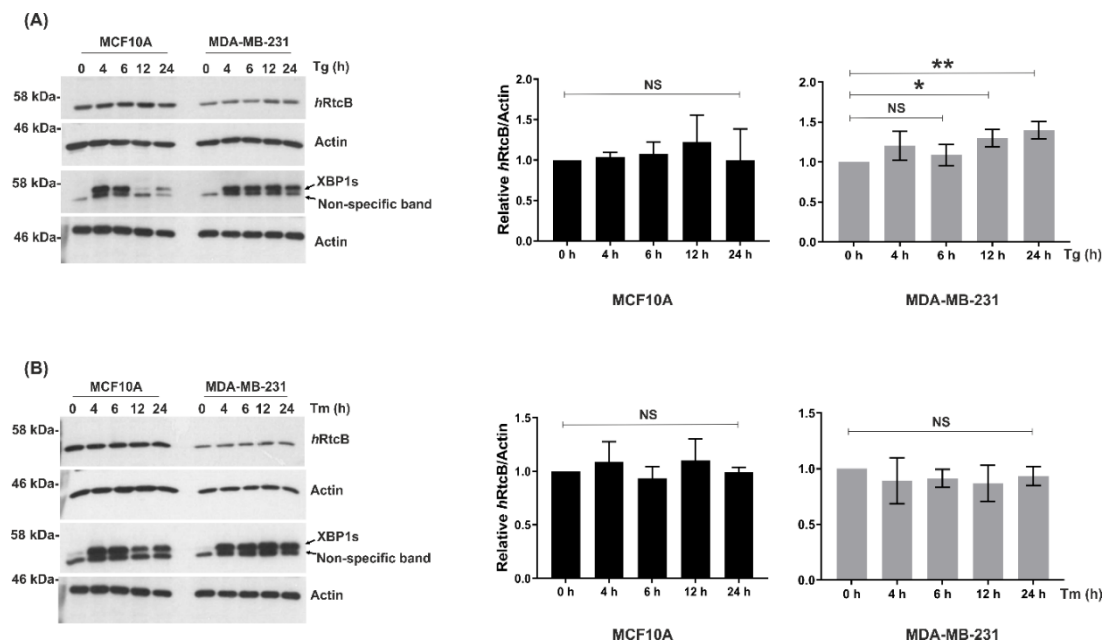


**Figure 4.13| Depletion of *hRtcB* reduces paclitaxel-induced XBP1s-luciferase activity.** MDA-MB-231/XBP1-luciferase cells were treated with (A) 10 nM paclitaxel for up to 72 hours ( $n = 3$ , mean XBP1s-luciferase activity and SD are displayed; One-way ANOVA multiple comparisons were used to analyze the statistical significance of differences in reporter activity between control and treated cells;  $*p < 0.05$ ,  $***p < 0.001$ ,  $****p < 0.0001$ ) and (B) 10 nM of both control siRNA and siRNA against *hRtcB* for 6 days which included 3 days of 10 nM paclitaxel treatment post-72 hours of *hRtcB* depletion ( $n = 2$ , mean XBP1s-luciferase activity  $\pm$  range are displayed) Cells were also treated with 20  $\mu$ M MKC8866 for 3 days in combination with paclitaxel which served as an additional control. XBP1s-

luciferase reporter activity was normalized as stated in earlier section 4.2.5. Blots are representative of 2 independent experiments.

#### 4.2.14 Inducibility of *hRtcB* in MDA-MB-231 cells

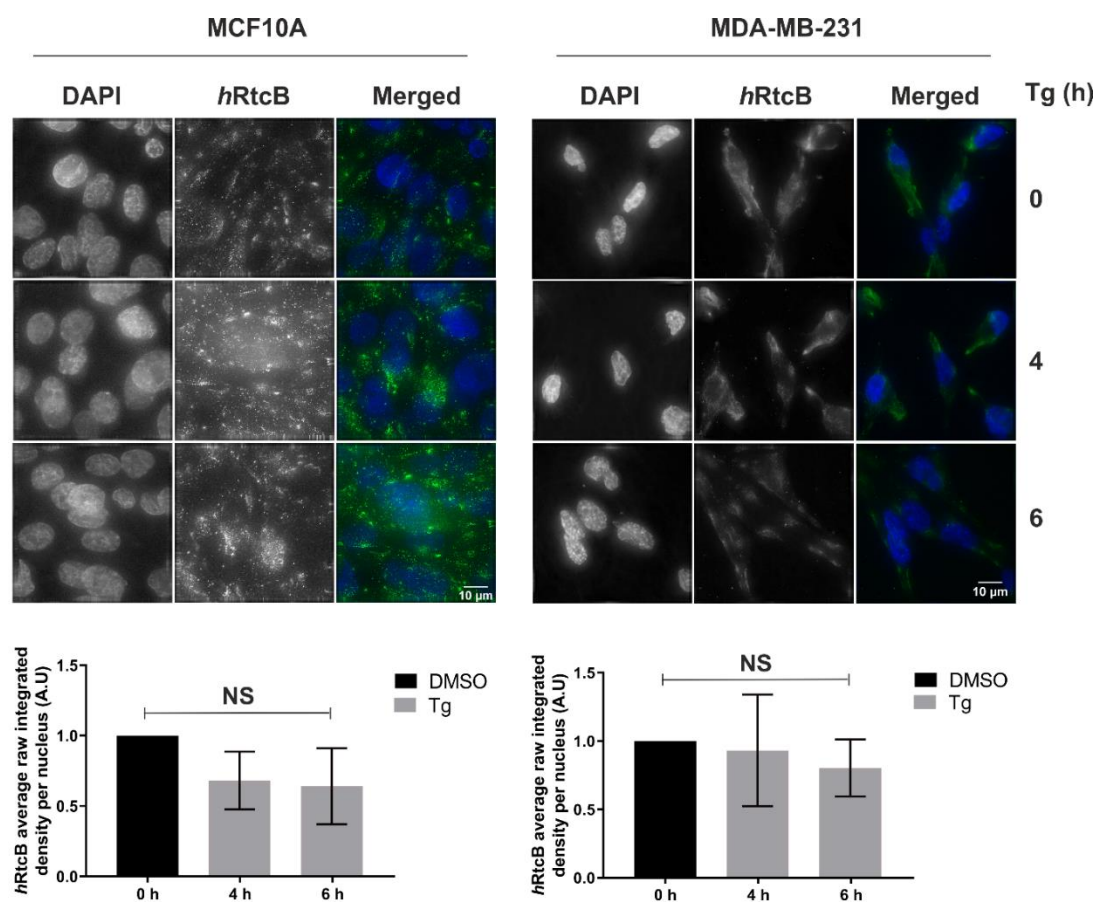
Additionally, the inducibility of *hRtcB* upon UPR induction was monitored. Inducibility was examined by monitoring the expression of the *hRtcB* protein in MDA-MB-231 cells treated with UPR inducing agents Tg and Tm compared to the non-tumorigenic MCF10A cells treated under the same conditions (Figure 4.14A, B). Consistent with a previous report (Jurkin *et al.*, 2014a), expression of the *hRtcB* protein was not induced during the UPR in the MCF10A cells compared to control cells. However, in contrast to previous report, there was a time-dependent increase in *hRtcB* protein expression in MDA-MB-231 cells, specifically with Tg treatment reaching 1.5-fold increase at 24 hours compared to the non-tumorigenic cells.



**Figure 4.14| Inducibility of *hRtcB* in non-tumorigenic and tumorigenic cell models.** MCF10A and MDA-MB-231 cells treated with (A) 0.5  $\mu$ M Tg and (B) 0.5  $\mu$ g/ml Tm for the indicated time periods. Blots are representative of 3 independent experiments. Levels of *hRtcB* protein were quantified and normalized to the housekeeping protein/gene and analyzed by densitometry ( $n = 3$ , mean expression levels and SD are displayed). One-way ANOVA multiple comparisons were used to analyze the statistical significance of differences in expression levels between control and treated cells (\* $p < 0.05$ , \*\* $p < 0.01$ , NS is non-significant).

#### 4.2.15 Localization of *hRtcB* in MDA-MB-231 cells

Mammalian tRNA splicing has been reported to occur in the nucleus (De Robertis *et al.*, 1981, Lund and Dahlberg, 1998), while *XBPI* mRNA splicing occurs in the cytoplasm (Yanagitani *et al.*, 2009, Sidrauski *et al.*, 1996). Previously Jurkin *et al.* (2014) have shown that during the UPR, in HeLa cells, there is no active recruitment of *hRtcB* to the site of *XBPI* mRNA splicing and that the subcellular distribution of the protein, both in the nucleus and in the cytoplasmic compartment, does not change after Tg treatment. To determine the subcellular localization of *hRtcB* in MDA-MB-231 cells, immunofluorescence staining was performed in both control and Tg treated cells (Figure 4.15). *hRtcB* was detected both in the nucleus and in the cytoplasm in MDA-MB-231 TNBC cells and also in non-tumorigenic MCF10A cells. Localization of *hRtcB* in MCF10A cells during UPR induction was compared with subcellular distribution of the ligase in MDA-MB-231 TNBC cells. Further quantification of the nuclear *hRtcB* in both the non-tumorigenic and tumorigenic cell models yielded no change in *hRtcB* levels during the UPR.



---

**Figure 4.15| Subcellular localization of *hRtcB* in non-tumorigenic and tumorigenic cell models.** Subcellular localization of *hRtcB* visualized by immunofluorescence staining of MCF10A and MDA-MB-231 cells treated with 0.5  $\mu$ M Tg for the indicated time periods. DeltaVision immunofluorescence images show *hRtcB* localization both in the nucleus and in the cytoplasm. The nucleus was visualized by DAPI staining. Images are representative of 3 independent experiments. 180-260 cells were counted across 36 different fields per cell line, per experiment. Levels of nuclear *hRtcB* protein were quantified and normalized to the control cells ( $n = 3$ , mean expression levels and SD are displayed). One-way ANOVA multiple comparisons were used to analyze the statistical significance of differences in expression levels between control and treated cells (NS is non-significant).

### 4.3 Discussion

A drug discovery initiative was undertaken to selectively inhibit the *hRtcB*-mediated XBP1s formation, to keep the RIDD unaffected. In an effort to discover an inhibitor of *hRtcB*, a VHTS campaign was conducted to screen through small molecule libraries provided by ZINC and the CBCS database against the ligand binding/active site of the *hRtcB* homology model. Such initiatives hold promise towards determining the druggability of the target of interest especially during early phase drug discovery when the experimental structures are yet to be resolved. 3,129 hits were generated in the campaign and were filtered down to top 200 molecules based on binding affinity scores/binding free energy values ( $\Delta G$ ) and docking scores. These 200 molecules were analyzed using the SIFt analysis to identify similarities and diversity of the binding modes of the molecules compared to GMP's interaction in the active site of the *hRtcB* homology model. Based on SIFt analysis, the top 100 molecules that displayed similar interactions to the GMP in the active site of the *hRtcB* homology model were selected. This helped to further filter down to the top 20 and finally to top 8 potential candidates that were further screened through a cell-based reporter assay such as the luciferase assay to identify potential lead molecules. Advancing through the hit-to-lead discovery phase, one compound, named compound-001, significantly reduced XBP1s in the luciferase assay. Another compound, named compound-007, however only showed a trend in the reduction of XBP1s in the luciferase assay. These two compounds were therefore considered as potential lead molecules to be further tested in successive assays.

As witnessed in the VHTS, compounds-001, -002, -003, -007, and -008 yielded a high binding affinity (between -71.03 and -76.43 kcal/mol) (Table 4.1) against the ligand

---

binding/active site of the *hRtcB* homology model. The virtual binding affinity scores correlate to some extent with the *in vitro* potency observed, as one potential candidate (compounds-001) yielded a statistically significant reduction of XBP1s in the luciferase assay. Similarly, compounds-004, -005, and -006 that exhibited medium to very poor virtual binding affinity (-68.07 to -36.48 kcal/mol) (Table 4.1), were completely unsuccessful in reducing XBP1s levels in the assay. An additional aspect resulting from the assay was its inherent variability. Luciferase assays are attractive as assay reporters due to their high signal to background bioluminescence, as no excitation light is required to generate a signal, resulting in very low background signal. However, the variability encountered across the experiments presented in the thesis was mainly due to the fact that following stable transfection of the XBP1-Luciferase/pcDNA3 plasmid, the MDA-MB-231/XBP1-luciferase cells that were isolated were in a mixed population and not from a single cell. Therefore, cell-to-cell variability in terms of gene integration into the host cells, may contribute to the variability of the assay. Importantly, while compound-001 proved to be toxic to the cells beyond 6 hours of treatment time, compound-007 did not impart any toxicity to the cells. If the compounds are further optimized to increase the potency of binding towards *hRtcB*, it might yield low or even negligible levels of toxicity. This is due to the fact, with increased potency, a much lower dose of the compounds should effectively reduce XBP1s expression levels. On the contrary, it cannot be excluded that following optimization of the compound towards improved binding to the *hRtcB* might also increase the toxicity levels and should be a subject of future study.

Testing the compounds-001 and -007 further in MDA-MB-231 TNBC cells, neither of them yielded any reduction of XBP1s both at the protein and mRNA levels. This may be due to lower sensitivity of the RT-PCR assay compared with the luciferase assay in detecting inhibition of XBP1s production, where the generation of light signal can easily pick up even minor changes in activity or expression of the target of interest. Based on the luciferase assay results, these two compounds might be having a minor effect on the ligase in the MDA-MB-231 TNBC cells. But this minor effect cannot be detected in the RT-PCR assay employed here.

These findings were further coupled with RNAi-mediated experiments in the luciferase assay. RNAi-mediated depletion of *hRtcB* significantly reduced both basal

---

endogenous XBP1s protein and basal XBP1s-luciferase activity and yielded ~3-fold reduction in paclitaxel-induced *XBPI* mRNA splicing. Further, as the exact mechanism of action of the potential lead molecules is yet to be validated, one of the lead molecules (compound-001) was added to *hRtcB* depleted cells to have an idea of the course of action of the compound. Addition of compound-001 to *hRtcB* depleted cells did not further reduce XBP1s-luciferase activity.

Treating cells with Tg and Tm did not induce *hRtcB* levels in the non-tumorigenic cell model. However, in the tumorigenic cell model, while Tg caused a minor increase in *hRtcB* levels, Tm did not induce *hRtcB* during the UPR. Moreover, immunofluorescence data presented in Figure 4.15 is consistent with previous findings in HeLa cells (Jurkin *et al.*, 2014a) and mouse ES cells (Lu *et al.*, 2014b) that RtcB localizes both in the nucleus and in the cytoplasmic compartments of MDA-MB-231 cells. Importantly, as reported by Jurkin *et al.* (2014) there is neither any active recruitment of *hRtcB* to the site of *XBPI* mRNA splicing nor in the nucleus where tRNA splicing occurs during the UPR. Quantification of the immunofluorescence data representing nuclear *hRtcB* in figure 4.15 (bottom panel) yielded a similar trend in both the non-tumorigenic and tumorigenic cell models.

Considering that both compounds-001 and -007 didn't yield any promising reduction of XBP1s under stress conditions in MDA-MB-231 TNBC cells, two strategies could be pursued from hereon. First, these two compounds could be tested in a biochemical assay such as microscale thermophoresis to detect or confirm their binding to the *hRtcB* protein. If successful, this should be followed by lead optimization to improve the potency and selectivity of the candidate molecules. Second, if unsuccessful, other compounds in the top 20 hit list could be purchased or synthesized and tested towards their activity against XBP1s reduction and should be tested in the biochemical assay in search of an inhibitor of *hRtcB*. Only when a binder of *hRtcB* is confirmed, following lead optimization steps, an *in vitro* RNA ligation assay should be performed, e.g., as employed by Javier Martinez group (Jurkin *et al.*, 2014b) to show inhibition of mRNA re-ligation by the compounds. As far as the mechanism of RNA re-ligation is concerned (discussed in section 1.1.2), potential inhibitors of RtcB should inhibit the first step of ligation (Figure 2D), i.e., the binding of the natural ligand GTP to RtcB and inhibit the formation of the RtcB-histidine-GMP intermediate.



---

Considering the transition from ligand discovery phase to the establishment of an inhibitor of *hRtcB* through the lead optimization steps, generating analog structures and testing/re-testing them will involve several years of work, important functional effects that need to be tested in the distant future should involve the selective reduction of XBP1s while unaffected the RIDD levels. All binders of *hRtcB* should be tested across a range of different cancer cell lines where there is reported evidence for a role for XBP1s in the cancer – e.g. multiple myeloma, GBM, TNBC etc. to verify this important functional effect. Simple assays such as the real-time PCR could be employed to verify that treating cancer cells with an *hRtcB* inhibitor would selectively reduce XBP1s levels while not affecting levels of RIDD targets, such as CD59 in MDA-MB-231 cells. Additionally, based on the ProBiS binding site similarity analysis, treating cells with *hRtcB* inhibitors may affect Roundabout homolog 1, Peptidyl-prolyl cis-trans isomerase FKBP14 (PPIase FKBP14), and 5' exonuclease Apollo proteins in humans. Roundabout homolog 1 interacts with Slit proteins that mediate guidance during projection of axons to different regions during neuronal development (Brose *et al.*, 1999) and is also involved in lung tumor suppression (Kong *et al.*, 2015). The ER resident PPIase FKBP22 (encoded by the FKBP14 gene) is involved in collagen biosynthesis in the rough ER. Lack of FKBP22 leads to myopathy, hearing loss, aortic rupture, and Ehlers-Danlos syndrome (Ishikawa and Bachinger, 2014). The 5' exonuclease Apollo (also called SNM1 homolog B (*hSNM1B*)) is a DNA exonuclease and functions as a caretaker protein (Lenain *et al.*, 2006). *hSNM1B* is a component of the human telomeric complex and is required for the protection of telomeres during or after their replication and to protect chromosome termini from being recognized and processed as DNA damage (Lenain *et al.*, 2006). Loss of *hSNM1B* led to the activation of DNA-damage signal at telomeres (van Overbeek and de Lange, 2006). The remaining other proteins from other organisms that came up in the ProBiS analysis displaying similar sites to the GTP-binding site of the *hRtcB* homology model, if they have human homologs may also get affected by *hRtcB* inhibitors. None of these proteins are ligases but mainly belong to the lyase, hydrolase, and oxidoreductase families. Unlike the ligases which are involved in catalyzing bond formation reactions, lyases are involved in catalyzing bond breaking reactions. Hydrolases on the other hand use water to break chemical bonds and commonly include lipases, phosphatases, glycosidases, peptidases, and nucleosidases. Finally, the oxidoreductases catalyze the transfer of electrons from electron donor

---

(reductant) to electron acceptor (oxidant). Another important functionality of *hRtcB* which includes re-ligation of a sub-set of pre-tRNAs into mature tRNAs is expected to be affected by all *hRtcB* inhibitors. However, this should not impart any global defects in protein translation as it has been already shown by Jurkin *et al.* (2014) that reduction of splicing-dependent mature tRNAs in *hRtcB*-depleted cells does not lead to a global defect in protein synthesis. Other important functions of *RtcB* already discussed in the introductory chapter of the thesis including development and fertility (Kosmaczewski *et al.*, 2014), ribosome homeostasis and chemotaxis in bacterial cells (Engl *et al.*, 2016) should also get affected by inhibitors of *RtcB*.

---

## CHAPTER 5: General discussion and future perspectives

Involvement of a tRNA ligase in regulating the UPR mRNA splicing was first evidenced by Peter Walter's group in 1996 (Sidrauski *et al.*, 1996). During the same time and in the following years while the identification of the *XBPI* mRNA splicing event went on to gain recognition (Cox and Walter, 1996, Yoshida *et al.*, 2001), it was finally after eighteen years in 2014 three independent groups reported that RtcB is the elusive tRNA ligase that regulates the *XBPI* mRNA splicing (Lu *et al.*, 2014b, Jurkin *et al.*, 2014b, Kosmaczewski *et al.*, 2014). Over the recent years, our understanding of the tRNA ligase RtcB continues to evolve as the structural and mechanistic insights (Englert *et al.*, 2012, Desai *et al.*, 2013, Nandy *et al.*, 2017) allowed us to explore the RtcB guanylylation and RNA ligation, with a future emphasis being placed on its pathological implications. Henceforth, as an advancement to the field, this thesis builds on the potential of *hRtcB* as a druggable target towards therapeutic treatment of TNBCs.

### 5.1 Understanding the druggability of *hRtcB* protein

In this thesis our understanding of the druggability of *hRtcB* has been developed for the first time through building a 3D protein prediction model from the crystal structure of its bacterial homolog as a template. With a 51% amino acid sequence identity between the human and *bRtcB*, the *hRtcB* model developed in chapter 3 is in complex with manganese and covalently bound GMP and shows the presence of protein folds that is suitable for drug targeting. Although X-ray structures for *bRtcB* have been developed (Desai *et al.*, 2013, Englert *et al.*, 2012) structures for the human homolog are yet to be resolved. Therefore, our predicted structure provides the first insights about the probable 3D protein structure of the *hRtcB* ligase that could be used for *in silico* drug discovery purposes. Furthermore, MD simulations, a routinely used technique to refine 3D protein structures based on experimental constraints from X-ray crystallography or NMR spectroscopy was employed. Such simulations allow atoms and molecules to interact for a fixed period of time, leading to a dynamic

---

evolution of the system. A 10 ns MD simulation did improve the quality of our predicted structure as compared to the pre-MD structure (Figure 3.3) when passed through several *in silico* quality assessments of which the most important was the increase in druggability of the ligand binding pocket.

Considering the importance of metal ions in contributing to the architecture of proteins and to enzyme catalytic reactions, the *hRtcB* model developed in the thesis harbors  $Mn^{2+}$  in an octahedral coordination with four conserved ligands (Cys122, Asn226, His259, and His353) and GMP in its ligand binding site. This was comparable to the *bRtcB* template structure where GMP is coordinated to  $Mn^{2+}$ , Cys98, His234, and His329. Additionally, the GMP in our predicted model forms a covalent bond with the N $\epsilon$  of His428, which is produced after the nucleophilic attack of His428 to GTP, with the release of pyrophosphate. The pyrophosphate interacts with the C2' hydroxyl group of the ribose moiety within the pocket. In contrast, the GMP in the *bRtcB* template structure forms a covalent bond with His404, while the residual pyrophosphate interacts with the ribose hydroxyl group. Such similar interactions between the residues in our model and the template structure produced considerable overlap between the key residues within the ligand binding site/active site when subjected to 3D structural alignment evaluation, holding promise as a realistic model that could be utilized for drug discovery.

One key aspect to note is the formation of the RtcB-guanylate intermediate; the first step in the three-step RNA ligation process that will be pharmacologically inhibited by any RtcB drug. While studies from the *bRtcB* template structure, showed that addition of  $\alpha[^{32}P]GTP$  to *bRtcB* directly resulted in the formation of *bRtcB*- $[^{32}P]GMP$  guanylate (Englert *et al.*, 2012), in contrast, addition of  $\alpha[^{32}P]GTP$  to *hRtcB* did not yield the formation of *hRtcB*-guanylate (Popow *et al.*, 2014). Formation of the *hRtcB*-guanylate could only occur upon addition of wild-type Archease, and inactive Archease point mutants did not support the guanylation (Popow *et al.*, 2014). Archease stimulated the activity of the *hRtcB* ligase complex only in the presence of GTP. Moreover, *hRtcB* guanylation also depends on the ATP-binding DDX1 subunit of the *hRtcB* ligase complex that is required for maximal stimulation of *hRtcB* by Archease (Popow *et al.*, 2014), providing evidence for the requirement of a dual co-factor for *hRtcB* guanylation. This points to the notion that pharmacological

---

inhibition of the *hRtcB*-guanylate may not be very straightforward unlike its bacterial homolog, given its complex nature where at least three out of six subunits of the *hRtcB* ligase complex collectively participate in the guanylation process along with GTP. This thesis demonstrates an approach to deter the binding of GTP to *hRtcB* by small molecule compounds which is supposed to inhibit the catalytic rearrangements that take place inside the ligand binding site upon GTP binding, however, does not take into consideration the involvement of Archease and DDX1 in the guanylation process. Further knowledge about the structural and mechanistic insights by which both Archease and DDX1 catalyze the formation of *hRtcB*-guanylate would be required if we are to consider targeting *hRtcB* therapeutically. Based on Javier Martinez's finding that interaction between GTP and *hRtcB* do not lead to *hRtcB*-guanylate formation (Popow *et al.*, 2014), it is hence possible that the inhibition of the *hRtcB*-GTP interaction by the compounds screened in the thesis are unable to significantly inhibit the *hRtcB*-guanylate formation. As a matter of fact, the inability of the compounds to significantly reduce XBP1s in MDA-MB-231 TNBC cells could be well inferred from the aforesaid finding reported by Martinez group.

The Cys122 residue (Cys98 in *bRtcB*) is a strictly conserved residue involved in metal ion coordination in the ligand binding site of RtcB proteins (Popow *et al.*, 2011). In an extension to Martinez's finding that mutating the Cys122 would render the ligase inactive (Popow *et al.*, 2011), virtual mutation of the Cys122 in our predicted model yielded a dramatically weaker binding of  $Mn^{2+}$  to the protein structure ( $-7.69 \text{ kcal mol}^{-1}$ ) when compared to the wild-type model ( $-13.93 \text{ kcal mol}^{-1}$ ). Such reduced coordination bond strength together with changed coordination of the ligands probably leads to an impaired catalytic activity in the mutant. Further, Lu and colleagues (Lu *et al.*, 2014b) have suggested that ligase-dead RtcB (Cys122 mutant) might behave as a dominant-negative protein. Compared to the *RtcB* knockout phenotype, the ligase-dead protein yielded better outcomes with a nearly complete loss of XBP1s both at the protein and mRNA levels. Hence if it is possible to interfere with this  $Mn^{2+}$ -Cys122 coordination with a compound that would turn the ligase into a dominant-negative protein, this might in future provide better inhibition of *hRtcB*-guanylate and abrogate cleaved *XBPIu* mRNA ends from being re-ligated by yet unknown compensatory RNA ligase(s). This notion is consistent with the study from the *bRtcB* template structure where interaction between  $\alpha[^{32}P]GTP$  and mutated *bRtcB* (Cys98Ala) failed

---

to form *bRtcB*-[<sup>32</sup>P]GMP guanylate (Englert *et al.*, 2012). Collectively, these findings indicate that interfering with the metal-ligand coordination might pave a better way for pharmacological modulation of the *hRtcB*-mediated catalysis.

Based on the ProBiS binding site similarity analysis, it is interesting to note that only 29 proteins (including 3 from humans) from the PDB displayed a structurally similar site to the GTP-binding site of the *hRtcB* homology model. None of these proteins represent any ligase and mainly belong to the lyase, hydrolase, and oxidoreductase families. This elucidates the uniqueness of the GTP-binding site of the *hRtcB* homology model and further suggests the possibility of fewer off-target effects. However, due to the small proportion of RNA ligases that have been characterized so far, one may speculate that the other, yet unknown RNA ligases may mediate the closest off-target effects for *hRtcB* targeted drugs. Alternatively, considering the canonical RNA ligases reported in other organisms so far, *hRtcB* targeted drugs may/may not interfere/inhibit the enzymatic activity of their human homologs if they exist at all. Currently five structurally characterized families of ATP-dependent RNA ligases exist: (i) Rnl1 family – bacteriophage T4 RNA ligase 1 (El Omari *et al.*, 2006), (ii) Rnl2 family – bacteriophage T4 RNA ligase 2 and RNA editing ligases (RELs) of the kinetoplastid protozoa *Trypanosoma* and *Leishmania* (Nandakumar *et al.*, 2006, Ho and Shuman, 2002) (iii) Rnl3 family – *Pyrococcus abyssi* RNA ligase (Brooks *et al.*, 2008), (iv) Rnl4 family – *Clostridium thermocellum* RNA ligase (Smith *et al.*, 2012), and (v) Rnl5 family – *Naegleria gruberi* RNA ligase (NgrRnl) (Unciuleac *et al.*, 2015). Additionally, a structurally uncharacterized *Deinococcus radiodurans* RNA ligase (DraRnl), the bacterial homolog of the eukaryal RNA ligase NgrRnl is also a member of the Rnl5 family (Martins and Shuman, 2004, Unciuleac and Shuman, 2015). While members of Rnl1 and Rnl4 heal single-strand breaks in RNA stem-loops (Wang *et al.*, 2007, Zhang *et al.*, 2012), Rnl2 and DraRnl are involved in healing 3'-OH/5'-P ends in duplex RNAs and RNA:DNA hybrids (Nandakumar *et al.*, 2004, Martins and Shuman, 2004). NgrRnl proved to be feeble at ligating 3'-OH/5'-P single-strand RNA breaks but efficiently ligated RNA:DNA hybrids (Unciuleac and Shuman, 2015). The Rnl3 family, on the other hand, besides sealing single-stranded circular RNAs, also seals single-stranded DNAs (Becker *et al.*, 2017).

---

## 5.2 Understanding the role of *hRtcB* in generation of XBP1s

Since the first report in *Science* in 2011 identifying RtcB as the catalytic subunit of a tRNA ligase complex (Popow *et al.*, 2011), it was soon discovered that it is similar to the tRNA ligase that regulates *XBPI* mRNA splicing during the UPR (Lu *et al.*, 2014b, Jurkin *et al.*, 2014b, Kosmaczewski *et al.*, 2014). Three years later, a sixth member of the tRNA ligase complex termed Archease was further reported in *Nature* (Popow *et al.*, 2014). This paper highlighted an essential requirement for Archease as a stimulatory co-factor for full activity of the tRNA ligase complex. Thus, RtcB, along with 5 other proteins forms a tRNA ligase complex that regulates both nuclear (tRNA splicing) and cytoplasmic (*XBPI* mRNA splicing) RNA fate (Perez-Gonzalez *et al.*, 2014, Jurkin *et al.*, 2014a, Popow *et al.*, 2011). Although Archease does not possess any RNA ligation activity (Jurkin *et al.*, 2014b), without Archease *hRtcB* is unable to catalyze RNA ligation (Poonthong *et al.*, 2017). Compared to control cells, RNAi-mediated depletion of *hRtcB* alone led to XBP1s expression almost to wild-type levels endogenously (Jurkin *et al.*, 2014a). Similarly, in MDA-MB-231 TNBC cells in chapter 4, RNAi-mediated depletion of *hRtcB* demonstrated a limited capacity to abolish *XBPIs* and *XBPIs* target genes which is consistent with two other previously published reports (Iwawaki and Tokuda, 2011, Jurkin *et al.*, 2014a). In conjunction with results presented in chapter 4 and other published reports, it is likely that even in the presence of reduced amounts of RtcB, stimulatory activity of Archease sustains ligation activity to re-ligate cleaved *XBPIu* mRNA ends. Additionally, it has been reported that simultaneous depletion of *hRtcB* and *Archease* in HeLa cells abolished XBP1s at the protein level and greatly reduced *XBPIs* and its target genes *EDEMI* and *DNAJB9* at the transcript level (Jurkin *et al.*, 2014a). Depletion of neither *hRtcB* nor *Archease* alone could completely abrogate XBP1s induction (Jurkin *et al.*, 2014a).

As stated earlier, while only a minor reduction in XBP1s was observed in shRNA-mediated *hRtcB* depleted HeLa cells, a complete knockout of *RtcB* alone sufficiently blocked XBP1s protein in plasma cells (Jurkin *et al.*, 2014a) as well as in mouse ES cells and did not affect the other UPR branches (Lu *et al.*, 2014b). However, in the mouse *RtcB* knockout phenotype, while levels of total *XBPI* mRNA were reduced, probably due to the cleavage and subsequent degradation of the cleaved *XBPIu* mRNA exon halves, low levels of *XBPIs* mRNA were still observed (Lu *et al.*, 2014b).

---

Importantly, the levels of RIDD targets *BLOS1* and *PDGFRB* mRNAs in HeLa cells remained unchanged after RNAi-mediated simultaneous depletion of both *hRtcB* and *Archease* and were equally reduced after UPR induction (Jurkin *et al.*, 2014a).

### 5.3 Understanding the need of targeting *hRtcB* in breast cancer

The involvement of the mammalian UPR mRNA splicing event represented by the splicing of transcription factor XBP1 followed by re-ligation of the cleaved *XBPIu* mRNA ends by *hRtcB* to generate *XBPIs* in promoting tumorigenesis has gained attention. *XBPI* knockout cells could not form tumors *in vivo* (Romero-Ramirez *et al.*, 2004). Overexpression of *XBPIs* has been reported in both cell lines and primary patient samples across a range of breast cancer subtypes (Fujimoto *et al.*, 2003, Zhu *et al.*, 2006, Li *et al.*, 2015), the highest being reported in TNBCs (Li *et al.*, 2015, Logue *et al.*, 2018, Chen *et al.*, 2014), a highly aggressive malignancy with limited treatment options including surgery, radio- and chemotherapy (Carey *et al.*, 2010, Foulkes *et al.*, 2010, Narod *et al.*, 2015). High levels of *XBPIs* corresponds to shorter overall and relapse-free survival of TNBC patients (Chen *et al.*, 2014, Li *et al.*, 2015, Logue *et al.*, 2018).

*XBPI* has also been reported to act as a regulator of EMT; a process crucial for invasion of epithelial tumors (Li *et al.*, 2015). Findings by Wang *et al.* (2012) report that *XBPI* modulates the ER lipid raft-associated 2 (ERLIN2) oncoprotein, promoting breast cancer survival. Recently, *XBPI* was shown to drive TNBC tumors through direct interaction with HIF1 $\alpha$ ; a master transcriptional regulator of tumor cell response to hypoxia (Chen *et al.*, 2014). Depleting *XBPI* inhibited tumor growth, relapse and reduced the breast cancer stem cell population CD44<sup>high</sup>/CD24<sup>low</sup> in TNBC tumors (Chen *et al.*, 2014). Furthermore, our group has recently shown that constitutive expression of *XBPIs* in TNBC induces pro-tumorigenic secretory factors including IL6, IL8, GM-CSF, CXCL1, and TGF $\beta$ 2 (Logue *et al.*, 2018). Through constitutive activation in tumor-associated dendritic cells (tDCs), *XBPI* directly promotes intrinsic tumor growth by disrupting normal DC mediated T cell-dependent anti-tumor immunity (Cubillos-Ruiz *et al.*, 2015). Depleting *XBPI* in tDCs decreased tumor burden and improved survival by evoking protective type 1 anti-tumor response in



---

preclinical ovarian cancer models; an important step towards cancer immunotherapy (Cubillos-Ruiz *et al.*, 2015).

Importantly, up-regulation of XBP1s was not observed in MCF10A cells, a spontaneously immortalized, non-transformed, non-tumorigenic breast epithelial cell line (Logue *et al.*, 2018). Addition of MKC8866, a selective IRE1 $\alpha$  RNase inhibitor, reduced breast cancer cell proliferation by reducing the number of cells entering S phase of the cell cycle (Logue *et al.*, 2018). In contrast, inhibiting the IRE1 $\alpha$  RNase in MCF10A cells, did not alter cell proliferation (Logue *et al.*, 2018). Targeting the IRE1 $\alpha$  - XBP1s branch hence seems attractive as the functioning of XBP1s is linked to the breast cancer phenotype only and not to the non-tumorigenic breast epithelial cells and therefore could be valuable as an alternative treatment strategy especially for treating TNBCs (Romero-Ramirez *et al.*, 2004, Chen *et al.*, 2014).

More importantly, recently a dual role played by the RNase of IRE1 $\alpha$  was established in GBM (Lhomond *et al.*, 2018). While the IRE1 $\alpha$  - XBP1s branch was shown to be pro-tumorigenic, the IRE1 $\alpha$  - RIDD branch elicited anti-tumorigenic features (Lhomond *et al.*, 2018). Using primary cell lines and human cancer tissue samples of GBM, it was shown that the IRE1 $\alpha$  - XBP1s branch promotes tumor aggressiveness through an increase in tumor immune infiltration, angiogenesis and cell invasive properties whereas the IRE1 $\alpha$  - RIDD branch elicits anti-tumorigenic properties by selectively reducing tumor invasion and angiogenesis. Although this dual mechanistic role shown in GBM would be an important subject of further investigation in other tumor types, uncoupling the antagonistic outputs, XBP1 splicing and RIDD of the RNase of IRE1 $\alpha$  seems promising especially in treating TNBC tumors that harbor constitutive expression of XBP1s (Logue *et al.*, 2018). A range of small molecule inhibitors targeting the RNase domain of IRE1 $\alpha$  have been developed, all of which block XBP1s formation (Papandreou *et al.*, 2011, Cross *et al.*, 2012, Sanches *et al.*, 2014, Volkmann *et al.*, 2011, Thamsen *et al.*, 2018). However, besides reducing XBP1s formation, these inhibitors although limiting tumor growth in myeloma (Papandreou *et al.*, 2011), TNBC (Logue *et al.*, 2018), GBM (Lhomond *et al.*, 2018), and leukemia (Vieri, 2017) would also inhibit RIDD.

---

Direct pharmacological inhibition of XBP1 is challenging due to its multifaceted nature as a transcription factor being involved in diverse cascades including development of secretory tissues (Lee *et al.*, 2005), osteoblast differentiation (Zambelli *et al.*, 2005), hepatogenesis (Reimold *et al.*, 2000), adipogenesis (Sha *et al.*, 2009), myogenic regulatory program (Blais *et al.*, 2005), innate immunity (dendritic cell development (Iwakoshi *et al.*, 2007), production of pro-inflammatory cytokines in macrophages (Martinon *et al.*, 2010)), and adaptive immunity (plasma cell differentiation (Reimold *et al.*, 2001), immunoglobulin M synthesis (Tirosh *et al.*, 2005)). Therefore, to uncouple the *XBP1* mRNA splicing event and RIDD activity, selective inhibition of the *XBP1* mRNA ligase RtcB may lead to better tumor treatment outcome. Through targeting of the *XBP1* mRNA ligase *hRtcB*, this thesis presents the possibility of reducing the expression of XBP1s in TNBC cells. Importantly, since depletion of *hRtcB* in MDA-MB-231 TNBC cells demonstrated a limited capacity to reduce basal *XBP1s* levels, therefore it might prove difficult to show that targeting *hRtcB* by RNAi does not change the RIDD levels. To present a proof-of-principle or proof-of-concept study, generation of a *hRtcB* knockout TNBC cell model therefore might need to be considered in future as this would mean Archease will not have access to *hRtcB* anymore. This would establish a therapeutic model where selective knockout of *hRtcB* might lead to significant impairment of tumor growth by uncoupling the antagonistic XBP1 splicing and RIDD branches at least theoretically.

Mainstream cancer therapies that include chemo- and radiotherapy induce DNA damage. Unlike cleaved RNA, bearing 3'-P and 5'-OH ends discussed so far, DNA damage often generates 3'-OH and 5'-P termini. However, there are instances when DNA damage leads to the generation of incorrect or “dirty” ends representing a 3'-P and a 5'-OH end which cannot be ligated by classical DNA ligases (DNA ligase I, II, and III in mammals). Reported first in 2013 by Stewart Shuman's group, RtcB offers such lesions a direct ligation of DNA 3'-P end to a DNA 5'-OH end (Das *et al.*, 2013). DNA ligation by RtcB operates by activation of the 3'-P end, a reverse strategy employed by classical DNA/RNA ligases which operate by activating a 5'-P DNA or RNA end (Das *et al.*, 2013). Importantly, it was reported that the *E. coli* RtcB was unable to join DNA 3'-P and 5'-OH ends at a nick in duplex DNA but was only able to guanylate the 3'-P end suggesting the possibility of further DNA repairing steps. Radiomimetic drugs such as the neocarzinostatin and calicheamicin induce DNA

---

double-strand breaks with 3'-P ends (Kappen and Goldberg, 1983, Smith *et al.*, 1994, Povirk, 2012) and RtcB thus offers a novel alternative to rectify such lesions.

## 5.4 Concluding remarks

To conclude, compared to the classical RNA ligases, RtcB, therefore, offers a novel family of “reverse ligases”. Unlike *bRtcB*, where addition of GTP directly led to the formation of *bRtcB*-GMP guanylate (step 1 of RNA ligation) (Englert *et al.*, 2012), significant inhibition of the *hRtcB*-guanylate pharmacologically may not be easily achievable as addition of GTP to *hRtcB* did not yield the formation of *hRtcB*-guanylate (Popow *et al.*, 2014). Formation of the *hRtcB*-guanylate could only occur upon the addition of Archease under stimulation by DDX1 subunit of the *hRtcB* ligase complex (Popow *et al.*, 2014). Owing to the moderate differences observed in the active site of the *hRtcB* homology model compared to the *bRtcB* template structure (Nandy *et al.*, 2017) and that the interaction between *hRtcB*, Archease, and DDX1 seems to be an essential event for *hRtcB* guanylylation, the idea of developing *hRtcB* targeted drugs could be strengthened by solving the crystal structures of *hRtcB* in complex with GTP, and if possible, in complex with Archease and DDX1. These structures will bring valuable insights towards the formation of *hRtcB*-guanylate. Importantly, besides ligating cleaved *XBPIu* mRNA ends, RIDD-generated mRNA fragments also possess 2',3'-cyclic-P and 5'-OH ends which makes them ideal candidates for RtcB-mediated ligation (Filipowicz, 2014). Solving these important structures will undoubtedly pave the way for effective development of RtcB drugs and their targeting in the therapeutic setting.

---

## References

- CheckMyMetal (CMM): Metal Binding Site Validation Server.: [http://csgid.org/csgid/metal\\_sites/](http://csgid.org/csgid/metal_sites/).
- The Human Protein Atlas. <https://www.proteinatlas.org/ENSG00000134398-ERN2/tissue>.
- PDB Database.: <http://www.rcsb.org>.
- PDBeFold. Structure Similarity.: <http://www.ebi.ac.uk/msd-srv/ssm/>.
- ProteinsPlus - Structure-Based Modeling Support Server.: <http://proteinsplus.zbh.uni-hamburg.de/>.
- QMEAN Server.: <http://swissmodel.expasy.org/qmean/cgi/index.cgi>.
- RAMPAGE server.: <http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>.
- VERIFY 3D Structure Evaluation Server.: [http://services.mbi.ucla.edu/Verify\\_3D/](http://services.mbi.ucla.edu/Verify_3D/).
2014. Molecular Operating Environment. *Chemical Computing Group*. version 2014.09 ed. Montreal, QC.
- 2015-4. Schrödinger Release 2015-4: Maestro, Schrödinger, LLC, New York, NY, 2018.
- ABELSON, J., TROTTA, C. R. & LI, H. 1998. tRNA Splicing. *J. Biol. Chem.*, 273, 12685-12688.
- ABRAHAM, M. J., MURTOLA, T., SCHULZ, R., PÁLL, S., SMITH, J. C., HESS, B. & LINDAHL, E. 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2, 19-25.
- ABRÀMOFF, M. D., MAGALHÃES, P. J. & RAM, S. J. 2004. Image processing with ImageJ. *Biophotonics Intern.*, 11, 36–42.
- ACOSTA-ALVEAR, D., ZHOU, Y., BLAIS, A., TSIKITIS, M., LENTS, N. H., ARIAS, C., LENNON, C. J., KLUGER, Y. & DYNLACHT, B. D. 2007. XBP1 controls diverse cell type- and condition-specific transcriptional regulatory networks. *Mol Cell*, 27, 53-66.
- ALI, M. M., BAGRATUNI, T., DAVENPORT, E. L., NOWAK, P. R., SILVA-SANTISTEBAN, M. C., HARDCASTLE, A., MCANDREWS, C., ROWLANDS, M. G., MORGAN, G. J., AHERNE, W., COLLINS, I., DAVIES, F. E. & PEARL, L. H. 2011. Structure of the Ire1 autophosphorylation complex and implications for the unfolded protein response. *EMBO J*, 30, 894-905.
- ALLNER, O., NILSSON, L. & VILLA, A. 2012. Magnesium Ion-Water Coordination and Exchange in Biomolecular Simulations. *J. Chem. Theory Comput.*, 8, 1493-502.
- ALTSCHUL, S. F., MADDEN, T. L., SCHÄFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a New Generation of Protein Database Search Programs. *Nucleic Acids Res.*, 25, 3389-3402.
- AMITSUR, M., LEVITZ, R. & KAUFMANN, G. 1987. Bacteriophage T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine tRNA. *EMBO J.*, 6, 2499-2503.
- ARSHAD, M., YE, Z., GU, X., WONG, C. K., LIU, Y., LI, D., ZHOU, L., ZHANG, Y., BAY, W. P., YU, V. C. & LI, P. 2013. RNF13, a RING finger protein, mediates endoplasmic reticulum stress-induced apoptosis through the inositol-requiring enzyme (IRE1 $\alpha$ )/c-Jun NH2-terminal kinase pathway. *J Biol Chem*, 288, 8726-36.
- AYERS, M. 2012. ChemSpider: The Free Chemical Database.
- BALTIMORE, D. 1970. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature*, 226, 1209-11.
- BASSIK, M. C. & KAMPMANN, M. 2011. Knocking out the door to tunicamycin entry. *Proc Natl Acad Sci U S A*, 108, 11731-2.

- 
- BECKER, H. F., HELIOU, A., DJAOUT, K., LESTINI, R., REGNIER, M. & MYLLYKALLIO, H. 2017. High-throughput sequencing reveals circular substrates for an archaeal RNA ligase. *RNA Biol*, 14, 1075-1085.
- BENHAMRON, S., HADAR, R., IWAWAKY, T., SO, J. S., LEE, A. H. & TIROSH, B. 2014. Regulated IRE1-dependent decay participates in curtailing immunoglobulin secretion from plasma cells. *Eur J Immunol*, 44, 867-76.
- BENKERT, P., KÜNZLI, M. & SCHWEDE, T. 2009. QMEAN Server for Protein Model Quality Estimation. *Nucleic Acids Res.*, 37, W510-W514.
- BERENDSEN, H. J. C., VAN DER SPOEL, D. & VAN DRUNEN, R. 1995. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comp. Phys. Comm.*, 91, 43-56.
- BERTOLOTTI, A., WANG, X., NOVOA, I., JUNGREIS, R., SCHLESSINGER, K., CHO, J. H., WEST, A. B. & RON, D. 2001. Increased sensitivity to dextran sodium sulfate colitis in IRE1beta-deficient mice. *J Clin Invest*, 107, 585-93.
- BERTOLOTTI, A., ZHANG, Y., HENDERSHOT, L. M., HARDING, H. P. & RON, D. 2000. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nat Cell Biol*, 2, 326-32.
- BETTIGOLE, S. E. & GLIMCHER, L. H. 2015. Endoplasmic Reticulum Stress in Immunity. *Annual Review of Immunology*, 33, 107-138.
- BIASINI, M., BIENERT, S., WATERHOUSE, A., ARNOLD, K., STUDER, G., SCHMIDT, T., KIEFER, F., GALLO CASSARINO, T., BERTONI, M., BORDOLI, L. & SCHWEDE, T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.*, 42, W252-W258.
- BINET, F., MAWAMBO, G., SITARAS, N., TETREAU, N., LAPALME, E., FAVRET, S., CERANI, A., LEBOEUF, D., TREMBLAY, S., REZENDE, F., JUAN, A. M., STAHL, A., JOYAL, J. S., MILOT, E., KAUFMAN, R. J., GUIMOND, M., KENNEDY, T. E. & SAPIEHA, P. 2013. Neuronal ER stress impedes myeloid-cell-induced vascular regeneration through IRE1alpha degradation of netrin-1. *Cell Metab*, 17, 353-71.
- BIRMINGHAM, A., SELFORS, L. M., FORSTER, T., WROBEL, D., KENNEDY, C. J., SHANKS, E., SANTOYO-LOPEZ, J., DUNICAN, D. J., LONG, A., KELLEHER, D., SMITH, Q., BEIJERSBERGEN, R. L., GHAZAL, P. & SHAMU, C. E. 2009. Statistical methods for analysis of high-throughput RNA interference screens. *Nat Methods*, 6, 569-75.
- BLAIS, A., TSIKITIS, M., ACOSTA-ALVEAR, D., SHARAN, R., KLUGER, Y. & DYNLACHT, B. D. 2005. An initial blueprint for myogenic differentiation. *Genes Dev*, 19, 553-69.
- BOOT-HANDFORD, R. P. & BRIGGS, M. D. 2010. The unfolded protein response and its relevance to connective tissue diseases. *Cell Tissue Res*, 339, 197-211.
- BOWIE, J. U., LUTHY, R. & EISENBERG, D. 1991. A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 253, 164-170.
- BRADBROOK, G. M., GLEICHMANN, T., HARROP, S. J., HABASH, J., RAFTERY, J., KALB, J., YARIV, J., HILLIER, I. H. & HELLIWELL, J. R. 1998. X-Ray and molecular dynamics studies of concanavalin-A glucoside and mannoside complexes. *J. Chem. Soc., Faraday Trans.*, 94, 1603-1611.
- BRAY, F., FERLAY, J., SOERJOMATARAM, I., SIEGEL, R. L., TORRE, L. A. & JEMAL, A. 2018. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*.
- BROOKS, M. A., MESLET-CLADIÈRE, L., GRAILLE, M., KUHN, J., BLONDEAU, K., MYLLYKALLIO, H. & VAN TILBEURGH, H. 2008. The structure of an archaeal homodimeric ligase which has RNA circularization activity. *Protein Sci*, 17, 1336-45.
- BROSE, K., BLAND, K. S., WANG, K. H., ARNOTT, D., HENZEL, W., GOODMAN, C. S., TESSIER-LAVIGNE, M. & KIDD, T. 1999. Slit Proteins Bind Robo Receptors and Have an Evolutionarily Conserved Role in Repulsive Axon Guidance. *Cell*, 96, 795-806.

- 
- BURNETTE, W. N. 1981. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem*, 112, 195-203.
- BUSTIN, S. A. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol*, 25, 169-93.
- CALVIN, K. & LI, H. 2008. RNA-splicing endonuclease structure and function. *Cell. Mol. Life Sci.*, 65, 1176-1185.
- CANAVES, J. M. 2004. Predicted role for the archease protein family based on structural and sequence analysis of TM1083 and MTH1598, two proteins structurally characterized through structural genomics efforts. *Proteins: Structure, Function, and Bioinformatics*, 56, 19-27.
- CANCER GENOME ATLAS, N. 2012. Comprehensive molecular portraits of human breast tumours. *Nature*, 490, 61-70.
- CAREY, L., WINER, E., VIALE, G., CAMERON, D. & GIANNI, L. 2010. Triple-negative breast cancer: disease entity or title of convenience? *Nat Rev Clin Oncol*, 7, 683-92.
- CARRASCO, D. R., SUKHDEO, K., PROTOPOPOVA, M., SINHA, R., ENOS, M., CARRASCO, D. E., ZHENG, M., MANI, M., HENDERSON, J., PINKUS, G. S., MUNSHI, N., HORNER, J., IVANOVA, E. V., PROTOPOPOV, A., ANDERSON, K. C., TONON, G. & DEPINHO, R. A. 2007. The differentiation and stress response factor XBP-1 drives multiple myeloma pathogenesis. *Cancer Cell*, 11, 349-60.
- CASARI, G., SANDER, C. & VALENCIA, A. 1995. A method to predict functional residues in proteins. *Nat Struct Biol*, 2, 171-8.
- CHAKRAVARTY, A. K. & SHUMAN, S. 2012. The sequential 2',3'-cyclic phosphodiesterase and 3'-phosphate/5'-OH ligation steps of the RtcB RNA splicing pathway are GTP-dependent. *Nucleic Acids Res.*, 40, 8558-8567.
- CHAKRAVARTY, A. K., SUBBOTIN, R., CHAIT, B. T. & SHUMAN, S. 2012. RNA ligase RtcB splices 3'-phosphate and 5'-OH ends via covalent RtcB-(histidiny)-GMP and polynucleotide-(3')pp(5')G intermediates. *Proc. Natl. Acad. Sci.*, 109, 6072-6077.
- CHAPMAN, M. A., LAWRENCE, M. S., KEATS, J. J., CIBULSKIS, K., SOUGNEZ, C., SCHINZEL, A. C., HARVIEW, C. L., BRUNET, J. P., AHMANN, G. J., ADLI, M., ANDERSON, K. C., ARDLIE, K. G., AUCLAIR, D., BAKER, A., BERGSAGEL, P. L., BERNSTEIN, B. E., DRIER, Y., FONSECA, R., GABRIEL, S. B., HOFMEISTER, C. C., JAGANNATH, S., JAKUBOWIAK, A. J., KRISHNAN, A., LEVY, J., LIEFELD, T., LONIAL, S., MAHAN, S., MFUKO, B., MONTI, S., PERKINS, L. M., ONOFRIO, R., PUGH, T. J., RAJKUMAR, S. V., RAMOS, A. H., SIEGEL, D. S., SIVACHENKO, A., STEWART, A. K., TRUDEL, S., VIJ, R., VOET, D., WINCKLER, W., ZIMMERMAN, T., CARPTEN, J., TRENT, J., HAHN, W. C., GARRAWAY, L. A., MEYERSON, M., LANDER, E. S., GETZ, G. & GOLUB, T. R. 2011. Initial genome sequencing and analysis of multiple myeloma. *Nature*, 471, 467-72.
- CHEN, C. & ZHANG, X. 2017. IRE1alpha-XBP1 pathway promotes melanoma progression by regulating IL-6/STAT3 signaling. *J Transl Med*, 15, 42.
- CHEN, X., ILIOPOULOS, D., ZHANG, Q., TANG, Q., GREENBLATT, M. B., HATZIAPOSTOLOU, M., LIM, E., TAM, W. L., NI, M., CHEN, Y., MAI, J., SHEN, H., HU, D. Z., ADORO, S., HU, B., SONG, M., TAN, C., LANDIS, M. D., FERRARI, M., SHIN, S. J., BROWN, M., CHANG, J. C., LIU, X. S. & GLIMCHER, L. H. 2014. XBP1 promotes triple-negative breast cancer by controlling the HIF1alpha pathway. *Nature*, 508, 103-107.
- CHO, J. A., LEE, A. H., PLATZER, B., CROSS, B. C. S., GARDNER, B. M., DE LUCA, H., LUONG, P., HARDING, H. P., GLIMCHER, L. H., WALTER, P., FIEBIGER, E., RON, D., KAGAN, J. C. & LENCER, W. I. 2013. The unfolded protein response element IRE1alpha senses bacterial proteins invading the ER to activate RIG-I and innate immune signaling. *Cell Host Microbe*, 13, 558-569.

- 
- CHOMCZYNSKI, P. & SACCHI, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*, 162, 156-9.
- CHU, G., HAYAKAWA, H. & BERG, P. 1987. Electroporation for the efficient transfection of mammalian cells with DNA. *Nucleic Acids Res*, 15, 1311-26.
- CLAUSS, I. M., CHU, M., ZHAO, J. L. & GLIMCHER, L. H. 1996. The basic domain/leucine zipper protein hXBP-1 preferentially binds to and transactivates CRE-like sequences containing an ACGT core. *Nucleic Acids Res*, 24, 1855-64.
- COELHO, D. S., CAIRRÃO, F., ZENG, X., PIRES, E., COELHO, A. V., RON, D., RYOO, H. D. & DOMINGOS, P. M. 2013. Xbp1-independent Ire1 signaling is required for photoreceptor differentiation and rhabdomere morphogenesis in *Drosophila*. *Cell reports*, 5, 10.1016/j.celrep.2013.09.046.
- COMA, I., CLARK, L., DIEZ, E., HARPER, G., HERRANZ, J., HOFMANN, G., LENNON, M., RICHMOND, N., VALMASEDA, M. & MACARRON, R. 2009. Process validation and screen reproducibility in high-throughput screening. *J Biomol Screen*, 14, 66-76.
- COONS, A. H. 1961. The Beginnings of Immunofluorescence. *The Journal of Immunology*, 87, 499.
- COONS, A. H., CREECH, H. J. & JONES, R. N. 1941. Immunological Properties of an Antibody Containing a Fluorescent Group. *Proceedings of the Society for Experimental Biology and Medicine*, 47, 200-202.
- COONS, A. H., CREECH, HUGH J., JONES, R. NORMAN, BERLINER, E. 1942. The Demonstration of Pneumococcal Antigen in Tissues by the Use of Fluorescent Antibody. *The Journal of Immunology*, 45, 159-170.
- COONS, A. H. & KAPLAN, M. H. 1950. Localization of antigen in tissue cells; improvements in a method for the detection of antigen by means of fluorescent antibody. *J Exp Med*, 91, 1-13.
- COX, J. S. & WALTER, P. 1996. A Novel Mechanism for Regulating Activity of a Transcription Factor That Controls the Unfolded Protein Response. *Cell*, 87, 391-404.
- CREDLE, J. J., FINER-MOORE, J. S., PAPA, F. R., STROUD, R. M. & WALTER, P. 2005. On the mechanism of sensing unfolded protein in the endoplasmic reticulum. *Proc Natl Acad Sci U S A*, 102, 18773-84.
- CROSS, B. C., BOND, P. J., SADOWSKI, P. G., JHA, B. K., ZAK, J., GOODMAN, J. M., SILVERMAN, R. H., NEUBERT, T. A., BAXENDALE, I. R., RON, D. & HARDING, H. P. 2012. The molecular basis for selective inhibition of unconventional mRNA splicing by an IRE1-binding small molecule. *Proc Natl Acad Sci U S A*, 109, E869-78.
- CUBILLOS-RUIZ, J. R., SILBERMAN, P. C., RUTKOWSKI, M. R., CHOPRA, S., PERALES-PUCHALT, A., SONG, M., ZHANG, S., BETTIGOLE, S. E., GUPTA, D., HOLCOMB, K., ELLENSON, L. H., CAPUTO, T., LEE, A. H., CONEJO-GARCIA, J. R. & GLIMCHER, L. H. 2015. ER Stress Sensor XBP1 Controls Anti-tumor Immunity by Disrupting Dendritic Cell Homeostasis. *Cell*, 161, 1527-38.
- CUCHILLO, C. M., NOGUES, M. V. & RAINES, R. T. 2011. Bovine pancreatic ribonuclease: fifty years of the first enzymatic reaction mechanism. *Biochemistry*, 50, 7835-7841.
- DAS, U., CHAKRAVARTY, A. K., REMUS, B. S. & SHUMAN, S. 2013. Rewriting the rules for end joining via enzymatic splicing of DNA 3'-PO4 and 5'-OH ends. *Proc Natl Acad Sci U S A*, 110, 20437-42.
- DAVIES, M. P., BARRACLOUGH, D. L., STEWART, C., JOYCE, K. A., ECCLES, R. M., BARRACLOUGH, R., RUDLAND, P. S. & SIBSON, D. R. 2008. Expression and splicing of the unfolded protein response gene XBP-1 are significantly associated with clinical outcome of endocrine-treated breast cancer. *Int J Cancer*, 123, 85-8.
- DE ROBERTIS, E. M., BLACK, P. & NISHIKURA, K. 1981. Intranuclear location of the tRNA splicing enzymes. *Cell*, 23, 89-93.

- 
- DEJEANS, N., PLUQUET, O., LHOMOND, S., GRISE, F., BOUCHECAREILH, M., JUIN, A., MEYNARD-CADARS, M., BIDAUD-MEYNARD, A., GENTIL, C., MOREAU, V., SALTEL, F. & CHEVET, E. 2012. Autocrine control of glioma cells adhesion and migration through IRE1 $\alpha$ -mediated cleavage of SPARC mRNA. *Journal of Cell Science*, 125, 4278.
- DENG, Z., CHUAQUI, C. & SINGH, J. 2004. Structural interaction fingerprint (SIFT): a novel method for analyzing three-dimensional protein-ligand binding interactions. *J Med Chem*, 47, 337-44.
- DESAI, K. K., BINGMAN, C. A., PHILLIPS, G. N., JR. & RAINES, R. T. 2013. Structures of the noncanonical RNA ligase RtcB reveal the mechanism of histidine guanylation. *Biochemistry*, 52, 2518-2525.
- DESAI, K. K., CHENG, C. L., BINGMAN, C. A., PHILLIPS, G. N., JR. & RAINES, R. T. 2014. A tRNA splicing operon: Archease endows RtcB with dual GTP/ATP cofactor specificity and accelerates RNA ligation. *Nucleic Acids Res*, 42, 3931-42.
- DESAI, K. K. & RAINES, R. T. 2012. tRNA ligase catalyzes the GTP-dependent ligation of RNA with 3'-phosphate and 5'-hydroxyl termini. *Biochemistry*, 51, 1333-1335.
- DEVARAJAN, E., SONG, Y. H., KRISHNAPPA, S. & ALT, E. 2012. Epithelial-mesenchymal transition in breast cancer lines is mediated through PDGF-D released by tissue-resident stem cells. *Int J Cancer*, 131, 1023-31.
- DVORKIN-GHEVA, A. & HASSELL, J. A. 2014. Identification of a novel luminal molecular subtype of breast cancer. *PLoS One*, 9, e103514.
- EISENBERG, D., LUTHY, R. & BOWIE, J. U. 1997. VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol*, 277, 396-404.
- EL OMARI, K., REN, J., BIRD, L. E., BONA, M. K., KLARMANN, G., LEGRICE, S. F. & STAMMERS, D. K. 2006. Molecular architecture and ligand recognition determinants for T4 RNA ligase. *J Biol Chem*, 281, 1573-9.
- ELBASHIR, S. M., HARBORTH, J., LENDECKEL, W., YALCIN, A., WEBER, K. & TUSCHL, T. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, 411, 494-8.
- ENGL, C., SCHAEFER, J., KOTTA-LOIZOU, I. & BUCK, M. 2016. Cellular and molecular phenotypes depending upon the RNA repair system RtcAB of Escherichia coli. *Nucleic Acids Res*, 44, 9933-9941.
- ENGLERT, M., SHEPPARD, K., ASLANIAN, A., YATES, J. R., III & SÖLL, D. 2011. Archaeal 3'-phosphate RNA splicing ligase characterization identifies the missing component in tRNA maturation. *Proc. Natl. Acad. Sci.*, 108, 1290-1295.
- ENGLERT, M., SHEPPARD, K., GUNDLLAPALLI, S., BEIER, H. & SÖLL, D. 2010. Branchiostoma floridae has separate healing and sealing enzymes for 5'-phosphate RNA ligation. *Proc. Natl. Acad. Sci.*, 107, 16834-16839.
- ENGLERT, M., XIA, S., OKADA, C., NAKAMURA, A., TANAVDE, V., YAO, M., EOM, S. H., KONIGSBERG, W. H., SOLL, D. & WANG, J. 2012. Structural and mechanistic insights into guanylation of RNA-splicing ligase RtcB joining RNA between 3'-terminal phosphate and 5'-OH. *Proc Natl Acad Sci U S A*, 109, 15235-15240.
- ESSMANN, U., PERERA, L., BERKOWITZ, M. L., DARDEN, T., LEE, H. & PEDERSEN, L. G. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103, 8577-8593.
- FECHTELER, T., DENGLER, U. & SCHOMBURG, D. 1995. Prediction of protein three-dimensional structures in insertion and deletion regions: a procedure for searching data bases of representative protein fragments using geometric scoring criteria. *J Mol Biol*, 253, 114-31.
- FELGNER, J. H., KUMAR, R., SRIDHAR, C. N., WHEELER, C. J., TSAI, Y. J., BORDER, R., RAMSEY, P., MARTIN, M. & FELGNER, P. L. 1994. Enhanced gene delivery and mechanism studies with a novel series of cationic lipid formulations. *J Biol Chem*, 269, 2550-61.



- 
- FELGNER, P. L., GADEK, T. R., HOLM, M., ROMAN, R., CHAN, H. W., WENZ, M., NORTHROP, J. P., RINGOLD, G. M. & DANIELSEN, M. 1987. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc Natl Acad Sci U S A*, 84, 7413-7.
- FERNANDEZ, C. & JAHNKE, W. 2004. New approaches for NMR screening in drug discovery. *Drug Discov Today Technol*, 1, 277-83.
- FILIPOWICZ, W. 2014. Making ends meet: a role of RNA ligase RTCB in unfolded protein response. *EMBO J*, 33, 2887-9.
- FILLMORE, C. M. & KUPERWASSER, C. 2008. Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Res*, 10, R25.
- FOULKES, W. D., SMITH, I. E. & REIS-FILHO, J. S. 2010. Triple-negative breast cancer. *N Engl J Med*, 363, 1938-48.
- FOX, S., BIROS, M., FARR-JONES, S., SOPCHAK, L., BOGGS, A., NICELY, H., KHOURY, R., CHRISTENSEN, D., SATIN, A., FOX, E., EDWARDS, S. 2005. *High Throughput Screening: New Users, More Cell-Based Assays, and a Host of New Tools*. Moraga, CA: HighTech Business Decisions.
- FOX, S., FARR-JONES, S., SOPCHAK, L., BOGGS, A., NICELY, H. W., KHOURY, R. & BIROS, M. 2006. High-throughput screening: update on practices and success. *J Biomol Screen*, 11, 864-9.
- FUJIMOTO, T., ONDA, M., NAGAI, H., NAGAHATA, T., OGAWA, K. & EMI, M. 2003. Upregulation and overexpression of human X-box binding protein 1 (hXBP-1) gene in primary breast cancers. *Breast Cancer*, 10, 301-6.
- GADDAM, D., STEVENS, N. & HOLLIEN, J. 2013. Comparison of mRNA localization and regulation during endoplasmic reticulum stress in *Drosophila* cells. *Mol Biol Cell*, 24, 14-20.
- GAULTON, A., BELLIS, L. J., BENTO, A. P., CHAMBERS, J., DAVIES, M., HERSEY, A., LIGHT, Y., MCGLINCHEY, S., MICHALOVICH, D., AL-LAZIKANI, B. & OVERINGTON, J. P. 2012. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res*, 40, D1100-7.
- GENOVESE, G., ERGUN, A., SHUKLA, S. A., CAMPOS, B., HANNA, J., GHOSH, P., QUAYLE, S. N., RAI, K., COLLA, S., YING, H., WU, C. J., SARKAR, S., XIAO, Y., ZHANG, J., ZHANG, H., KWONG, L., DUNN, K., WIEDEMAYER, W. R., BRENNAN, C., ZHENG, H., RIMM, D. L., COLLINS, J. J. & CHIN, L. 2012. microRNA regulatory network inference identifies miR-34a as a novel regulator of TGF-beta signaling in glioblastoma. *Cancer Discov*, 2, 736-49.
- GENSCHIK, P., BILLY, E., SWIANIEWICZ, M. & FILIPOWICZ, W. 1997. The human RNA 3'-terminal phosphate cyclase is a member of a new family of proteins conserved in Eucarya, Bacteria and Archaea. *EMBO J*, 16, 2955-67.
- GENSCHIK, P., DRABIKOWSKI, K. & FILIPOWICZ, W. 1998. Characterization of the *Escherichia coli* RNA 3'-terminal phosphate cyclase and its sigma54-regulated operon. *J Biol Chem*, 273, 25516-26.
- GERBER, B., LOIBL, S., EIDTMANN, H., REZAI, M., FASCHING, P. A., TESCH, H., EGGEMANN, H., SCHRADER, I., KITTEL, K., HANUSCH, C., KREIENBERG, R., SOLBACH, C., JACKISCH, C., KUNZ, G., BLOHMER, J. U., HUOBER, J., HAUSCHILD, M., NEKLJUDOVA, V., UNTCH, M., VON MINCKWITZ, G. & GERMAN BREAST GROUP, I. 2013. Neoadjuvant bevacizumab and anthracycline-taxane-based chemotherapy in 678 triple-negative primary breast cancers; results from the geparquinto study (GBG 44). *Ann Oncol*, 24, 2978-84.
- GHOSH, R., LIPSON, K. L., SARGENT, K. E., MERCURIO, A. M., HUNT, J. S., RON, D. & URANO, F. 2010. Transcriptional regulation of VEGF-A by the unfolded protein response pathway. *PLoS One*, 5, e9575.

- 
- GOODFORD, P. J. 1985. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J Med Chem*, 28, 849-57.
- GOULD, S. J. & SUBRAMANI, S. 1988. Firefly luciferase as a tool in molecular and cell biology. *Anal Biochem*, 175, 5-13.
- GREENMAN, C., STEPHENS, P., SMITH, R., DALGLIESH, G. L., HUNTER, C., BIGNELL, G., DAVIES, H., TEAGUE, J., BUTLER, A., STEVENS, C., EDKINS, S., O'MEARA, S., VASTRIK, I., SCHMIDT, E. E., AVIS, T., BARTHORPE, S., BHAMRA, G., BUCK, G., CHOUDHURY, B., CLEMENTS, J., COLE, J., DICKS, E., FORBES, S., GRAY, K., HALLIDAY, K., HARRISON, R., HILLS, K., HINTON, J., JENKINSON, A., JONES, D., MENZIES, A., MIRONENKO, T., PERRY, J., RAINE, K., RICHARDSON, D., SHEPHERD, R., SMALL, A., TOFTS, C., VARIAN, J., WEBB, T., WEST, S., WIDAA, S., YATES, A., CAHILL, D. P., LOUIS, D. N., GOLDSTRAW, P., NICHOLSON, A. G., BRASSEUR, F., LOOIJENGA, L., WEBER, B. L., CHIEW, Y. E., DEFAZIO, A., GREAVES, M. F., GREEN, A. R., CAMPBELL, P., BIRNEY, E., EASTON, D. F., CHENEVIX-TRENCH, G., TAN, M. H., KHOO, S. K., TEH, B. T., YUEN, S. T., LEUNG, S. Y., WOOSTER, R., FUTREAL, P. A. & STRATTON, M. R. 2007. Patterns of somatic mutation in human cancer genomes. *Nature*, 446, 153-8.
- GRIFFITH, F. 1928. The Significance of Pneumococcal Types. *J Hyg (Lond)*, 27, 113-59.
- GU, F., NGUYEN, D. T., STUIBLE, M., DUBE, N., TREMBLAY, M. L. & CHEVET, E. 2004. Protein-tyrosine phosphatase 1B potentiates IRE1 signaling during endoplasmic reticulum stress. *J Biol Chem*, 279, 49689-93.
- GUICHARD, C., AMADDEO, G., IMBEAUD, S., LADEIRO, Y., PELLETIER, L., MAAD, I. B., CALDERARO, J., BIOULAC-SAGE, P., LETEXIER, M., DEGOS, F., CLEMENT, B., BALABAUD, C., CHEVET, E., LAURENT, A., COUCHY, G., LETOUZE, E., CALVO, F. & ZUCMAN-ROSSI, J. 2012. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet*, 44, 694-8.
- GUPTA, S., DEEPTI, A., DEEGAN, S., LISBONA, F., HETZ, C. & SAMALI, A. 2010. HSP72 protects cells from ER stress-induced apoptosis via enhancement of IRE1alpha-XBP1 signaling through a physical interaction. *PLoS Biol*, 8, e1000410.
- HAMILTON, A. J. & BAULCOMBE, D. C. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, 286, 950-2.
- HAN, D., LERNER, A. G., WALLE, L. V., UPTON, J.-P., XU, W., HAGEN, A., BACKES, B. J., OAKES, S. A. & PAPA, F. R. 2009. IRE1 $\alpha$  Kinase Activation Modes Control Alternate Endoribonuclease Outputs to Determine Divergent Cell Fates. *Cell*, 138, 562-575.
- HE, Y., BEATTY, A., HAN, X., JI, Y., MA, X., ADELSTEIN, R. S., YATES, J. R., 3RD, KEMPHUES, K. & QI, L. 2012. Nonmuscle myosin IIB links cytoskeleton to IRE1alpha signaling during ER stress. *Dev Cell*, 23, 1141-52.
- HEID, C. A., STEVENS, J., LIVAK, K. J. & WILLIAMS, P. M. 1996. Real time quantitative PCR. *Genome Res*, 6, 986-94.
- HETZ, C., BERNASCONI, P., FISHER, J., LEE, A. H., BASSIK, M. C., ANTONSSON, B., BRANDT, G. S., IWAKOSHI, N. N., SCHINZEL, A., GLIMCHER, L. H. & KORSMEYER, S. J. 2006. Proapoptotic BAX and BAK modulate the unfolded protein response by a direct interaction with IRE1alpha. *Science*, 312, 572-6.
- HETZ, C., CHEVET, E. & HARDING, H. P. 2013. Targeting the unfolded protein response in disease. *Nat Rev Drug Discov*, 12, 703-19.
- HETZ, C. & GLIMCHER, L. 2008. The daily job of night killers: alternative roles of the BCL-2 family in organelle physiology. *Trends Cell Biol*, 18, 38-44.
- HO, C. K. & SHUMAN, S. 2002. Bacteriophage T4 RNA ligase 2 (gp24.1) exemplifies a family of RNA ligases found in all phylogenetic domains. *Proc Natl Acad Sci U S A*, 99, 12709-14.

- 
- HOLLIEN, J., LIN, J. H., LI, H., STEVENS, N., WALTER, P. & WEISSMAN, J. S. 2009. Regulated Ire1-dependent decay of messenger RNAs in mammalian cells. *J Cell Biol*, 186, 323-31.
- HOLLIEN, J. & WEISSMAN, J. S. 2006. Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. *Science*, 313, 104-7.
- HONETH, G., BENDAHL, P. O., RINGNER, M., SAAL, L. H., GRUVBERGER-SAAL, S. K., LOVGREN, K., GRABAU, D., FERNO, M., BORG, A. & HEGARDT, C. 2008. The CD44+/CD24- phenotype is enriched in basal-like breast tumors. *Breast Cancer Res*, 10, R53.
- HOOFT, R. W., SANDER, C., SCHARF, M. & VRIEND, G. 1996a. The PDBFINDER database: a summary of PDB, DSSP and HSSP information with added value. *Comput Appl Biosci*, 12, 525-9.
- HOOFT, R. W., VRIEND, G., SANDER, C. & ABOLA, E. E. 1996b. Errors in protein structures. *Nature*, 381, 272.
- HOPKINS, A. L. & GROOM, C. R. 2002. The druggable genome. *Nat Rev Drug Discov*, 1, 727-30.
- HUOBER, J., VON MINCKWITZ, G., DENKERT, C., TESCH, H., WEISS, E., ZAHM, D. M., BELAU, A., KHANDAN, F., HAUSCHILD, M., THOMSEN, C., HÖGEL, B., DARB-ESFAHANI, S., MEHTA, K. & LOIBL, S. 2010. Effect of neoadjuvant anthracycline-taxane-based chemotherapy in different biological breast cancer phenotypes: overall results from the GeparTrio study. *Breast Cancer Research and Treatment*, 124, 133-140.
- HUR, K. Y., SO, J. S., RUDA, V., FRANK-KAMENETSKY, M., FITZGERALD, K., KOTELIANSKY, V., IWAWAKI, T., GLIMCHER, L. H. & LEE, A. H. 2012. IRE1alpha activation protects mice against acetaminophen-induced hepatotoxicity. *J Exp Med*, 209, 307-18.
- IQBAL, J., DAI, K., SEIMON, T., JUNGREIS, R., OYADOMARI, M., KURIAKOSE, G., RON, D., TABAS, I. & HUSSAIN, M. M. 2008. IRE1beta inhibits chylomicron production by selectively degrading MTP mRNA. *Cell Metab*, 7, 445-55.
- IRWIN, J. J., STERLING, T., MYSINGER, M. M., BOLSTAD, E. S. & COLEMAN, R. G. 2012. ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model*, 52, 1757-68.
- ISHIKAWA, Y. & BACHINGER, H. P. 2014. A substrate preference for the rough endoplasmic reticulum resident protein FKBP22 during collagen biosynthesis. *J Biol Chem*, 289, 18189-201.
- IWAKOSHI, N. N., PYPAERT, M. & GLIMCHER, L. H. 2007. The transcription factor XBP-1 is essential for the development and survival of dendritic cells. *J Exp Med*, 204, 2267-75.
- IWAWAKI, T., AKAI, R., YAMANAKA, S. & KOHNO, K. 2009. Function of IRE1 alpha in the placenta is essential for placental development and embryonic viability. *Proc Natl Acad Sci U S A*, 106, 16657-62.
- IWAWAKI, T. & TOKUDA, M. 2011. Function of yeast and amphioxus tRNA ligase in IRE1alpha-dependent XBP1 mRNA splicing. *Biochem Biophys Res Commun*, 413, 527-31.
- JAKALIAN, A., JACK, D. B. & BAYLY, C. I. 2002. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem*, 23, 1623-41.
- JONES, D. T. & TAYLOR, W. R. 1992. A new approach to protein fold recognition. *Nature*, 358, 86-89.
- JURKIN, J., HENKEL, T., NIELSEN, A. F., MINNICH, M., POPOW, J., KAUFMANN, T., HEINDL, K., HOFFMANN, T., BUSSLINGER, M. & MARTINEZ, J. 2014a. The mammalian tRNA ligase complex mediates splicing of XBP1 mRNA and controls antibody secretion in plasma cells. *EMBO J*, 33, 2922-2936.
- JURKIN, J., HENKEL, T., NIELSEN, A. F., MINNICH, M., POPOW, J., KAUFMANN, T., HEINDL, K., HOFFMANN, T., BUSSLINGER, M. & MARTINEZ, J. 2014b. The mammalian tRNA ligase

- 
- complex mediates splicing of XBP1 mRNA and controls antibody secretion in plasma cells. *EMBO J*, 33, 2922-36.
- JWA, M. & CHANG, P. 2012. PARP16 is a tail-anchored endoplasmic reticulum protein required for the PERK- and IRE1alpha-mediated unfolded protein response. *Nat Cell Biol*, 14, 1223-30.
- KAISER, J. 2008. Molecular biology. Industrial-style screening meets academic biology. *Science*, 321, 764-6.
- KAPPEN, L. S. & GOLDBERG, I. H. 1983. Deoxyribonucleic acid damage by neocarzinostatin chromophore: strand breaks generated by selective oxidation of C-5' of deoxyribose. *Biochemistry*, 22, 4872-8.
- KELLENBERGER, E., MULLER, P., SCHALON, C., BRET, G., FOATA, N. & ROGNAN, D. 2006. sc-PDB: an annotated database of druggable binding sites from the Protein Data Bank. *J Chem Inf Model*, 46, 717-27.
- KELLEY, L. A., MEZULIS, S., YATES, C. M., WASS, M. N. & STERNBERG, M. J. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, 10, 845-858.
- KESKIN, O., MA, B. & NUSSINOV, R. 2005. Hot regions in protein--protein interactions: the organization and contribution of structurally conserved hot spot residues. *J Mol Biol*, 345, 1281-94.
- KIM, H. Y. & WYSS, D. F. 2015. NMR screening in fragment-based drug design: a practical guide. *Methods Mol Biol*, 1263, 197-208.
- KIMATA, Y., KIMATA, Y. I., SHIMIZU, Y., ABE, H., FARCASANU, I. C., TAKEUCHI, M., ROSE, M. D. & KOHNO, K. 2003. Genetic evidence for a role of BiP/Kar2 that regulates Ire1 in response to accumulation of unfolded proteins. *Molecular biology of the cell*, 14, 2559-2569.
- KIMMIG, P., DIAZ, M., ZHENG, J., WILLIAMS, C. C., LANG, A., ARAGON, T., LI, H. & WALTER, P. 2012. The unfolded protein response in fission yeast modulates stability of select mRNAs to maintain protein homeostasis. *Elife*, 1, e00048.
- KIRKWOOD, J. G. 1935. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.*, 3, 300-313.
- KNOX, C., LAW, V., JEWISON, T., LIU, P., LY, S., FROLKIS, A., PON, A., BANCO, K., MAK, C., NEVEU, V., DJOUMBOU, Y., EISNER, R., GUO, A. C. & WISHART, D. S. 2011. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res*, 39, D1035-41.
- KONC, J., CESNIK, T., KONC, J. T., PENCA, M. & JANEZIC, D. 2012. ProBiS-database: precalculated binding site similarities and local pairwise alignments of PDB structures. *J Chem Inf Model*, 52, 604-12.
- KONC, J. & JANEZIC, D. 2010a. ProBiS algorithm for detection of structurally similar protein binding sites by local structural alignment. *Bioinformatics*, 26, 1160-8.
- KONC, J. & JANEZIC, D. 2010b. ProBiS: a web server for detection of structurally similar protein binding sites. *Nucleic Acids Res*, 38, W436-40.
- KONG, R., YI, F., WEN, P., LIU, J., CHEN, X., REN, J., LI, X., SHANG, Y., NIE, Y., WU, K., FAN, D., ZHU, L., FENG, W. & WU, J. Y. 2015. Myo9b is a key player in SLIT/ROBO-mediated lung tumor suppression. *The Journal of Clinical Investigation*, 125, 4407-4420.
- KORENNYKH, A. V., EGEA, P. F., KOROSTELEV, A. A., FINER-MOORE, J., ZHANG, C., SHOKAT, K. M., STROUD, R. M. & WALTER, P. 2009. The unfolded protein response signals through high-order assembly of Ire1. *Nature*, 457, 687-693.
- KOSMACZEWSKI, S. G., EDWARDS, T. J., HAN, S. M., ECKWAHL, M. J., MEYER, B. I., PEACH, S., HESSELBERTH, J. R., WOLIN, S. L. & HAMMARLUND, M. 2014. The RtcB RNA ligase is an essential component of the metazoan unfolded protein response. *EMBO Rep*, 15, 1278-85.

- 
- KOSMACZEWSKI, S. G., HAN, S. M., HAN, B., IRVING MEYER, B., BAIG, H. S., ATHAR, W., LIN-MOORE, A. T., KOELLE, M. R. & HAMMARLUND, M. 2015. RNA ligation in neurons by RtcB inhibits axon regeneration. *Proc Natl Acad Sci U S A*, 112, 8451-6.
- KRIEGER, E., DARDEN, T., NABUURS, S. B., FINKELSTEIN, A. & VRIEND, G. 2004. Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins*, 57, 678-683.
- KRIEGER, E., JOO, K., LEE, J., LEE, J., RAMAN, S., THOMPSON, J., TYKA, M., BAKER, D. & KARPLUS, K. 2009. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins*, 77 Suppl 9, 114-22.
- KUMAR, S., BOUZIDA, D., SWENDSEN, R. H., KOLLMAN, P. A. & ROSENBERG, J. M. 1992. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method *J. Comput. Chem.*, 13, 1011-1021.
- KWON, D., KOH, J., KIM, S., GO, H., MIN, H. S., KIM, Y. A., KIM, D. K., JEON, Y. K. & CHUNG, D. H. 2018. Overexpression of endoplasmic reticulum stress-related proteins, XBP1s and GRP78, predicts poor prognosis in pulmonary adenocarcinoma. *Lung Cancer*, 122, 131-137.
- LASKI, F. A., FIRE, A. Z., RAJBHANDARY, U. L. & SHARP, P. A. 1983. Characterization of tRNA precursor splicing in mammalian extracts. *J Biol Chem*, 258, 11974-80.
- LATT, S. A. & STETTEN, G. 1976. Spectral studies on 33258 Hoechst and related bisbenzimidazole dyes useful for fluorescent detection of deoxyribonucleic acid synthesis. *J Histochem Cytochem*, 24, 24-33.
- LATT, S. A., STETTEN, G., JUERGENSEN, L. A., WILLARD, H. F. & SCHER, C. D. 1975. Recent developments in the detection of deoxyribonucleic acid synthesis by 33258 Hoechst fluorescence. *J Histochem Cytochem*, 23, 493-505.
- LAURIE, A. T. & JACKSON, R. M. 2005. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics*, 21, 1908-16.
- LAVECCHIA, A. & DI GIOVANNI, C. 2013. Virtual screening strategies in drug discovery: a critical review. *Curr Med Chem*, 20, 2839-60.
- LAW, R., BARKER, O., BARKER, J. J., HESTERKAMP, T., GODEMANN, R., ANDERSEN, O., FRYATT, T., COURTNEY, S., HALLETT, D. & WHITTAKER, M. 2009. The multiple roles of computational chemistry in fragment-based drug design. *J Comput Aided Mol Des*, 23, 459-73.
- LEDERBERG, J. 1952. Cell genetics and hereditary symbiosis. *Physiol Rev*, 32, 403-30.
- LEE, A.-H., HEIDTMAN, K., HOTAMISLIGIL, G. S. & GLIMCHER, L. H. 2011. Dual and opposing roles of the unfolded protein response regulated by IRE1 $\alpha$  and XBP1 in proinsulin processing and insulin secretion. *Proceedings of the National Academy of Sciences*, 108, 8885.
- LEE, A.-H., IWAKOSHI, N. N. & GLIMCHER, L. H. 2003a. XBP-1 Regulates a Subset of Endoplasmic Reticulum Resident Chaperone Genes in the Unfolded Protein Response. *Mol. Cell. Biol.*, 23, 7448-7459.
- LEE, A. H., CHU, G. C., IWAKOSHI, N. N. & GLIMCHER, L. H. 2005. XBP-1 is required for biogenesis of cellular secretory machinery of exocrine glands. *EMBO J*, 24, 4368-80.
- LEE, A. H., IWAKOSHI, N. N. & GLIMCHER, L. H. 2003b. XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response. *Mol Cell Biol*, 23, 7448-59.
- LEE, K. P. K., DEY, M., NECULAI, D., CAO, C., DEVER, T. E. & SICHERI, F. 2008. Structure of the dual enzyme Ire1 reveals the basis for catalysis and regulation in non-conventional RNA splicing. *Cell*, 132, 89-100.
- LEMKUL, J. A. & BEVAN, D. R. 2010. Assessing the Stability of Alzheimer's Amyloid Protofibrils Using Molecular Dynamics. *J. Phys. Chem.*, 114, 1652-1660.

- 
- LENAIN, C., BAUWENS, S., AMIARD, S., BRUNORI, M., GIRAUD-PANIS, M. J. & GILSON, E. 2006. The Apollo 5' exonuclease functions together with TRF2 to protect telomeres from DNA repair. *Curr Biol*, 16, 1303-10.
- LERNER, ALANA G., UPTON, J.-P., PRAVEEN, P. V. K., GHOSH, R., NAKAGAWA, Y., IGBARIA, A., SHEN, S., NGUYEN, V., BACKES, BRADLEY J., HEIMAN, M., HEINTZ, N., GREENGARD, P., HUI, S., TANG, Q., TRUSINA, A., OAKES, SCOTT A. & PAPA, FEROS R. 2012. IRE1 $\alpha$  Induces Thioredoxin-Interacting Protein to Activate the NLRP3 Inflammasome and Promote Programmed Cell Death under Irremediable ER Stress. *Cell Metabolism*, 16, 250-264.
- LEVITT, M. 1992. Accurate modeling of protein conformation by automatic segment matching. *J Mol Biol*, 226, 507-33.
- LEVITZ, R., CHAPMAN, D., AMITSUR, M., GREEN, R., SNYDER, L. & KAUFMANN, G. 1990. The optional E. coli prr locus encodes a latent form of phage T4-induced anticodon nuclease. *EMBO J.*, 9, 1383-1389.
- LHOMOND, S., AVRIL, T., DEJEANS, N., VOUTETAKIS, K., DOULTSINOS, D., MCMAHON, M., PINEAU, R., OBACZ, J., PAPADODIMA, O., JOUAN, F., BOURIEN, H., LOGOTHETI, M., JEGOU, G., PALLARES-LUPON, N., SCHMIT, K., LE RESTE, P. J., ETCHEVERRY, A., MOSSER, J., BARROSO, K., VAULEON, E., MAUREL, M., SAMALI, A., PATTERSON, J. B., PLUQUET, O., HETZ, C., QUILLIEN, V., CHATZIOANNOU, A. & CHEVET, E. 2018. Dual IRE1 RNase functions dictate glioblastoma development. *EMBO Mol Med*, 10.
- LI, H., CHEN, X., GAO, Y., WU, J., ZENG, F. & SONG, F. 2015. XBP1 induces snail expression to promote epithelial- to-mesenchymal transition and invasion of breast cancer cells. *Cell Signal*, 27, 82-9.
- LI, Q., CHENG, T., WANG, Y. & BRYANT, S. H. 2010. PubChem as a public resource for drug discovery. *Drug Discov Today*, 15, 1052-7.
- LIANG, J., EDELSBRUNNER, H. & WOODWARD, C. 1998. Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein Sci*, 7, 1884-97.
- LICHTARGE, O. & SOWA, M. E. 2002. Evolutionary predictions of binding surfaces and interactions. *Curr Opin Struct Biol*, 12, 21-7.
- LINDORFF-LARSEN, K., PIANA, S., PALMO, K., MARAGAKIS, P., KLEPEIS, J. L., DROR, R. O. & SHAW, D. E. 2010. Improved Side-Chain Torsion Potentials for the Amber ff99sb Protein Force Field. *Proteins*, 78, 1950-8.
- LIPSON, K. L., GHOSH, R. & URANO, F. 2008. The role of IRE1 $\alpha$  in the degradation of insulin mRNA in pancreatic beta-cells. *PLoS One*, 3, e1648.
- LISBONA, F., ROJAS-RIVERA, D., THIELEN, P., ZAMORANO, S., TODD, D., MARTINON, F., GLAVIC, A., KRESS, C., LIN, J. H., WALTER, P., REED, J. C., GLIMCHER, L. H. & HETZ, C. 2009. BAX inhibitor-1 is a negative regulator of the ER stress sensor IRE1 $\alpha$ . *Mol Cell*, 33, 679-91.
- LIU, C. Y., SCHRODER, M. & KAUFMAN, R. J. 2000. Ligand-independent dimerization activates the stress response kinases IRE1 and PERK in the lumen of the endoplasmic reticulum. *J Biol Chem*, 275, 24881-5.
- LIU, T., LIN, Y., WEN, X., JORISSEN, R. N. & GILSON, M. K. 2007. BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res*, 35, D198-201.
- LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, 25, 402-8.
- LOGUE, S. E., MCGRATH, E. P., CLEARY, P., GREENE, S., MNICH, K., ALMANZA, A., CHEVET, E., DWYER, R. M., OOMMEN, A., LEGEMBRE, P., GODEY, F., MADDEN, E. C., LEUZZI, B., OBACZ, J., ZENG, Q., PATTERSON, J. B., JAGER, R., GORMAN, A. M. & SAMALI, A. 2018.

- 
- Inhibition of IRE1 RNase activity modulates the tumor cell secretome and enhances response to chemotherapy. *Nat Commun*, 9, 3267.
- LU, H., CHEN, I., SHIMODA, L. A., PARK, Y., ZHANG, C., TRAN, L., ZHANG, H. & SEMENZA, G. L. 2017. Chemotherapy-Induced Ca(2+) Release Stimulates Breast Cancer Stem Cell Enrichment. *Cell Rep*, 18, 1946-1957.
- LU, M., LAWRENCE, D. A., MARSTERS, S., ACOSTA-ALVEAR, D., KIMMIG, P., MENDEZ, A. S., PATON, A. W., PATON, J. C., WALTER, P. & ASHKENAZI, A. 2014a. Opposing unfolded-protein-response signals converge on death receptor 5 to control apoptosis. *Science*, 345, 98-101.
- LU, Y., LIANG, F.-X. & WANG, X. 2014b. A synthetic biology approach identifies the mammalian UPR RNA ligase RtcB. *Mol. Cell*, 55, 758-770.
- LUND, E. & DAHLBERG, J. E. 1998. Proofreading and Aminoacylation of tRNAs Before Export from the Nucleus. *Science*, 282, 2082.
- LUO, D., HE, Y., ZHANG, H., YU, L., CHEN, H., XU, Z., TANG, S., URANO, F. & MIN, W. 2008. AIP1 is critical in transducing IRE1-mediated endoplasmic reticulum stress response. *J Biol Chem*, 283, 11905-12.
- LUO, G. & KARPLUS, M. 2011. Determining the conformational change that accompanies donor-acceptor distance fluctuations: an umbrella sampling analysis. *J. Phys. Chem. B*, 115, 7991-5.
- LYTTON, J., WESTLIN, M. & HANLEY, M. R. 1991. Thapsigargin inhibits the sarcoplasmic or endoplasmic reticulum Ca-ATPase family of calcium pumps. *J Biol Chem*, 266, 17067-71.
- MA, Y. & HENDERSHOT, L. M. 2004. The role of the unfolded protein response in tumour development: friend or foe? *Nat Rev Cancer*, 4, 966-77.
- MACGOWAN, A. P., WOOTTON, M., HEDGES, A. J., BOWKER, K. E., HOLT, H. A. & REEVES, D. S. 1996. A new time-kill method of assessing the relative efficacy of antimicrobial agents alone and in combination developed using a representative beta-lactam, aminoglycoside and fluoroquinolone. *J Antimicrob Chemother*, 38, 193-203.
- MALHOTRA, G. K., ZHAO, X., BAND, H. & BAND, V. 2010. Histological, molecular and functional subtypes of breast cancers. *Cancer Biol Ther*, 10, 955-60.
- MARR, D. & HILDRETH, E. 1980. Theory of edge detection. *Proc R Soc Lond B Biol Sci*, 207, 187-217.
- MARTINON, F., CHEN, X., LEE, A. H. & GLIMCHER, L. H. 2010. TLR activation of the transcription factor XBP1 regulates innate immune responses in macrophages. *Nat Immunol*, 11, 411-8.
- MARTINS, A. & SHUMAN, S. 2004. An RNA ligase from *Deinococcus radiodurans*. *J Biol Chem*, 279, 50654-61.
- MAUREL, M., CHEVET, E., TAVERNIER, J. & GERLO, S. 2014. Getting RIDD of RNA: IRE1 in cell fate regulation. *Trends Biochem Sci*, 39, 245-54.
- MAUREL, M., DEJEANS, N., TAOUI, S., CHEVET, E. & GROSSET, C. F. 2013a. MicroRNA-1291-mediated silencing of IRE1 $\alpha$  enhances Glypican-3 expression. *RNA*, 19, 778-88.
- MAUREL, M., DEJEANS, N., TAOUI, S., CHEVET, E. & GROSSET, C. F. 2013b. MicroRNA-1291-mediated silencing of IRE1 $\alpha$  enhances Glypican-3 expression. *RNA*, 19, 778-788.
- MCINNES, C. 2007. Virtual screening strategies in drug discovery. *Curr Opin Chem Biol*, 11, 494-502.
- MOORE, K. & HOLLIEN, J. 2015. Ire1-mediated decay in mammalian cells relies on mRNA sequence, structure, and translational status. *Mol Biol Cell*, 26, 2873-84.
- MOORE, K. A., PLANT, J. J., GADDAM, D., CRAFT, J. & HOLLIEN, J. 2013. Regulation of sumo mRNA during endoplasmic reticulum stress. *PLoS One*, 8, e75723.

- 
- MOULT, J., FIDELIS, K., KRYSHTAFOVYCH, A., SCHWEDE, T. & TRAMONTANO, A. 2014. Critical assessment of methods of protein structure prediction (CASP)--round x. *Proteins*, 82 Suppl 2, 1-6.
- NABHOLTZ, J. M., ABRIAL, C., MOURET-REYNIER, M. A., DAUPLAT, M. M., WEBER, B., GLIGOROV, J., FOREST, A. M., TREDAN, O., VANLEMMENS, L., PETIT, T., GUIU, S., VAN PRAAGH, I., JOUANNAUD, C., DUBRAY-LONGERAS, P., TUBIANA-MATHIEU, N., BENMAMMAR, K. E., KULLAB, S., BAHADOOR, M. R., RADOSEVIC-ROBIN, N., KWIATKOWSKI, F., DESRICHARD, A., CAYRE, A., UHRHAMMER, N., CHALABI, N., CHOLLET, P. & PENAULT-LLORCA, F. 2014. Multicentric neoadjuvant phase II study of panitumumab combined with an anthracycline/taxane-based chemotherapy in operable triple-negative breast cancer: identification of biologically defined signatures predicting treatment impact. *Ann Oncol*, 25, 1570-7.
- NANDAKUMAR, J., HO, C. K., LIMA, C. D. & SHUMAN, S. 2004. RNA substrate specificity and structure-guided mutational analysis of bacteriophage T4 RNA ligase 2. *J Biol Chem*, 279, 31337-47.
- NANDAKUMAR, J., SHUMAN, S. & LIMA, C. D. 2006. RNA ligase structures reveal the basis for RNA specificity and conformational changes that drive ligation forward. *Cell*, 127, 71-84.
- NANDY, A., SAENZ-MENDEZ, P., GORMAN, A. M., SAMALI, A. & ERIKSSON, L. A. 2017. Homology model of the human tRNA splicing ligase RtcB. *Proteins*, 85, 1983-1993.
- NAROD, S. A., DENT, R. A. & FOULKES, W. D. 2015. CCR 20th Anniversary Commentary: Triple-Negative Breast Cancer in 2015-Still in the Ballpark. *Clin Cancer Res*, 21, 3813-4.
- NAYAL, M. & HONIG, B. 2006. On the nature of cavities on protein surfaces: application to the identification of drug-binding sites. *Proteins*, 63, 892-906.
- NGUYEN, D. T., KEBACHE, S., FAZEL, A., WONG, H. N., JENNA, S., EMADALI, A., LEE, E. H., BERGERON, J. J., KAUFMAN, R. J., LAROSE, L. & CHEVET, E. 2004. Nck-dependent activation of extracellular signal-regulated kinase-1 and regulation of cell survival during endoplasmic reticulum stress. *Mol Biol Cell*, 15, 4248-60.
- OIKAWA, D., TOKUDA, M., HOSODA, A. & IWAWAKI, T. 2010. Identification of a consensus element recognized and cleaved by IRE1 alpha. *Nucleic Acids Res*, 38, 6265-73.
- OIKAWA, D., TOKUDA, M. & IWAWAKI, T. 2007. Site-specific cleavage of CD59 mRNA by endoplasmic reticulum-localized ribonuclease, IRE1. *Biochem Biophys Res Commun*, 360, 122-7.
- OKADA, C., MAEGAWA, Y., YAO, M. & TANAKA, I. 2006. Crystal structure of an RtcB homolog protein (PH1602-extein protein) from *Pyrococcus horikoshii* reveals a novel fold. *Proteins*, 63, 1119-1122.
- ONO, S. J., LIOU, H. C., DAVIDON, R., STROMINGER, J. L. & GLIMCHER, L. H. 1991. Human X-box-binding protein 1 is required for the transcription of a subset of human class II major histocompatibility genes and forms a heterodimer with c-fos. *Proceedings of the National Academy of Sciences of the United States of America*, 88, 4309-4312.
- OSORIO, F., TAVERNIER, S. J., HOFFMANN, E., SAEYS, Y., MARTENS, L., VETTERS, J., DELRUE, I., DE RYCKE, R., PARTHOENS, E., POULIOT, P., IWAWAKI, T., JANSSENS, S. & LAMBRECHT, B. N. 2014. The unfolded-protein-response sensor IRE-1alpha regulates the function of CD8alpha+ dendritic cells. *Nat Immunol*, 15, 248-57.
- OTA, A. & WANG, Y. 2012. Cdc37/Hsp90 protein-mediated regulation of IRE1alpha protein activity in endoplasmic reticulum stress response and insulin synthesis in INS-1 cells. *J Biol Chem*, 287, 6266-74.
- PAPANDREOU, I., DENKO, N. C., OLSON, M., VAN MELCKEBEKE, H., LUST, S., TAM, A., SOLOW-CORDERO, D. E., BOULEY, D. M., OFFNER, F., NIWA, M. & KOONG, A. C. 2011. Identification of an Ire1alpha endonuclease specific inhibitor with cytotoxic activity against human multiple myeloma. *Blood*, 117, 1311-4.



- 
- PARSONS, D. W., JONES, S., ZHANG, X., LIN, J. C., LEARY, R. J., ANGENENDT, P., MANKOO, P., CARTER, H., SIU, I. M., GALLIA, G. L., OLIVI, A., MCLENDON, R., RASHEED, B. A., KEIR, S., NIKOLSKAYA, T., NIKOLSKY, Y., BUSAM, D. A., TEKLEAB, H., DIAZ, L. A., JR., HARTIGAN, J., SMITH, D. R., STRAUSBERG, R. L., MARIE, S. K., SHINJO, S. M., YAN, H., RIGGINS, G. J., BIGNER, D. D., KARCHIN, R., PAPADOPOULOS, N., PARMIGIANI, G., VOGELSTEIN, B., VELCULESCU, V. E. & KINZLER, K. W. 2008. An integrated genomic analysis of human glioblastoma multiforme. *Science*, 321, 1807-12.
- PASCAL, J. M. 2008. DNA and RNA ligases: structural variations and shared mechanisms. *Curr. Opin. Struct. Biol.*, 18, 96-105.
- PEITSCH, M. C. 1996. ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. *Biochem Soc Trans*, 24, 274-9.
- PEREZ-GONZALEZ, A., PAZO, A., NAVAJAS, R., CIORDIA, S., RODRIGUEZ-FRANSEN, A. & NIETO, A. 2014. hCLE/C14orf166 associates with DDX1-HSPC117-FAM98B in a novel transcription-dependent shuttling RNA-transporting complex. *PLoS One*, 9, e90957.
- PERKINS, K. K., FURNEAUX, H. & HURWITZ, J. 1985. Isolation and Characterization of an RNA Ligase from HeLa Cells. *Proc. Natl. Acad. Sci.*, 82, 684-688.
- PEROU, C. M., SORLIE, T., EISEN, M. B., VAN DE RIJN, M., JEFFREY, S. S., REES, C. A., POLLACK, J. R., ROSS, D. T., JOHNSEN, H., AKSLIN, L. A., FLUGE, O., PERGAMENSHIKOV, A., WILLIAMS, C., ZHU, S. X., LONNING, P. E., BORRESEN-DALE, A. L., BROWN, P. O. & BOTSTEIN, D. 2000. Molecular portraits of human breast tumours. *Nature*, 406, 747-52.
- PETERS, K. P., FAUCK, J. & FROMMEL, C. 1996. The automatic search for ligand binding sites in proteins of known three-dimensional structure using only geometric criteria. *J Mol Biol*, 256, 201-13.
- PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.*, 25, 1605-1612.
- PEURALA, H., GRECO, D., HEIKKINEN, T., KAUR, S., BARTKOVA, J., JAMSHIDI, M., AITTOMAKI, K., HEIKKILA, P., BARTEK, J., BLOMQVIST, C., BUTZOW, R. & NEVANLINNA, H. 2011. MiR-34a expression has an effect for lower risk of metastasis and associates with expression patterns predicting clinical outcome in breast cancer. *PLoS One*, 6, e26122.
- PITT-RIVERS, R. & IMPIOMBATO, F. S. 1968. The binding of sodium dodecyl sulphate to various proteins. *Biochem J*, 109, 825-30.
- PLAZINSKI, W. & KNYS-DZIECIUCH, A. 2013. The 'order-to-disorder' conformational transition in CD44 protein: an umbrella sampling analysis. *J. Mol. Graph. Model.*, 45, 122-7.
- PLUQUET, O., DEJEANS, N., BOUCHECAREILH, M., LHOMOND, S., PINEAU, R., HIGA, A., DELUGIN, M., COMBE, C., LORIOT, S., CUBEL, G., DUGOT-SENANT, N., VITAL, A., LOISEAU, H., GOSLINE, S. J. C., TAOUI, S., HALLETT, M., SARKARIA, J. N., ANDERSON, K., WU, W., RODRIGUEZ, F. J., ROSENBAUM, J., SALTEL, F., FERNANDEZ-ZAPICO, M. E. & CHEVET, E. 2013. Posttranscriptional Regulation of PER1 Underlies the Oncogenic Function of IREα. *Cancer research*, 73, 4732-4743.
- POOTHONG, J., TIRASOPHON, W. & KAUFMAN, R. J. 2017. Functional analysis of the mammalian RNA ligase for IRE1 in the unfolded protein response. *Biosci Rep*, 37.
- POPOW, J., ENGLERT, M., WEITZER, S., SCHLEIFFER, A., MIERZWA, B., MECHTLER, K., TROWITZSCH, S., WILL, C. L., LÜHRMANN, R., SÖLL, D. & MARTINEZ, J. 2011. HSPC117 is the essential subunit of a human tRNA splicing ligase complex. *Science*, 331, 760-764.
- POPOW, J., JURKIN, J., SCHLEIFFER, A. & MARTINEZ, J. 2014. Analysis of orthologous groups reveals archease and DDX1 as tRNA splicing factors. *Nature*, 511, 104-7.

- 
- POPOW, J., SCHLEIFFER, A. & MARTINEZ, J. 2012. Diversity and roles of (t)RNA ligases. *Cell. Mol. Life Sci.*, 69, 2657-2670.
- POVIRK, L. F. 2012. Processing of damaged DNA ends for double-strand break repair in mammalian cells. *ISRN Mol Biol*, 2012.
- PUPKO, T., BELL, R. E., MAYROSE, I., GLASER, F. & BEN-TAL, N. 2002. Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. *Bioinformatics*, 18 Suppl 1, S71-7.
- QIU, Y., MAO, T., ZHANG, Y., SHAO, M., YOU, J., DING, Q., CHEN, Y., WU, D., XIE, D., LIN, X., GAO, X., KAUFMAN, R. J., LI, W. & LIU, Y. 2010. A crucial role for RACK1 in the regulation of glucose-stimulated IRE1alpha activation in pancreatic beta cells. *Sci Signal*, 3, ra7.
- RAKHA, E. A., REIS-FILHO, J. S. & ELLIS, I. O. 2008. Basal-like breast cancer: a critical review. *J Clin Oncol*, 26, 2568-81.
- RAY, A., ZHANG, S., RENTAS, C., CALDWELL, K. A. & CALDWELL, G. A. 2014. RTCB-1 mediates neuroprotection via XBP-1 mRNA splicing in the unfolded protein response pathway. *J Neurosci*, 34, 16076-85.
- REIMOLD, A. M., ETKIN, A., CLAUSS, I., PERKINS, A., FRIEND, D. S., ZHANG, J., HORTON, H. F., SCOTT, A., ORKIN, S. H., BYRNE, M. C., GRUSBY, M. J. & GLIMCHER, L. H. 2000. An essential role in liver development for transcription factor XBP-1. *Genes Dev*, 14, 152-7.
- REIMOLD, A. M., IWAKOSHI, N. N., MANIS, J., VALLABHAJOSYULA, P., SZOMOLANYI-TSUDA, E., GRAVALLESE, E. M., FRIEND, D., GRUSBY, M. J., ALT, F. & GLIMCHER, L. H. 2001. Plasma cell differentiation requires the transcription factor XBP-1. *Nature*, 412, 300-7.
- REIS-FILHO, J. S. & TUTT, A. N. 2008. Triple negative tumours: a critical review. *Histopathology*, 52, 108-18.
- REN, D., TU, H. C., KIM, H., WANG, G. X., BEAN, G. R., TAKEUCHI, O., JEFFERS, J. R., ZAMBETTI, G. P., HSIEH, J. J. & CHENG, E. H. 2010. BID, BIM, and PUMA are essential for activation of the BAX- and BAK-dependent cell death program. *Science*, 330, 1390-3.
- ROBSON, M., IM, S. A., SENKUS, E., XU, B., DOMCHEK, S. M., MASUDA, N., DELALOGUE, S., LI, W., TUNG, N., ARMSTRONG, A., WU, W., GOESSL, C., RUNSWICK, S. & CONTE, P. 2017. Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N Engl J Med*, 377, 523-533.
- ROBSON, M. E., TUNG, N., CONTE, P., IM, S. A., SENKUS, E., XU, B., MASUDA, N., DELALOGUE, S., LI, W., ARMSTRONG, A., WU, W., GOESSL, C., RUNSWICK, S. & DOMCHEK, S. M. 2019. OlympiAD final overall survival and tolerability results: Olaparib versus chemotherapy treatment of physician's choice in patients with a germline BRCA mutation and HER2-negative metastatic breast cancer. *Ann Oncol*.
- ROMERO-RAMIREZ, L., CAO, H., NELSON, D., HAMMOND, E., LEE, A. H., YOSHIDA, H., MORI, K., GLIMCHER, L. H., DENKO, N. C., GIACCIA, A. J., LE, Q. T. & KOONG, A. C. 2004. XBP1 is essential for survival under hypoxic conditions and is required for tumor growth. *Cancer Res*, 64, 5943-7.
- ROY, A., KUCKURAL, A. & ZHANG, Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protoc.*, 5, 725-738.
- ROY, B. & LEE, A. S. 1999. The mammalian endoplasmic reticulum stress response element consists of an evolutionarily conserved tripartite structure and interacts with a novel stress-inducible complex. *Nucleic Acids Res*, 27, 1437-43.
- SAIKI, R. K., SCHARF, S., FALOONA, F., MULLIS, K. B., HORN, G. T., ERLICH, H. A. & ARNHEIM, N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230, 1350-4.

- 
- SALI, A. & BLUNDELL, T. L. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Biol. Chem.*, 234, 779-815.
- SANCHES, M., DUFFY, N. M., TALUKDAR, M., THEVAKUMARAN, N., CHIOVITTI, D., CANNY, M. D., LEE, K., KURINOV, I., UEHLING, D., AL-AWAR, R., PODA, G., PRAKESCH, M., WILSON, B., TAM, V., SCHWEITZER, C., TORO, A., LUCAS, J. L., VUGA, D., LEHMANN, L., DUROCHER, D., ZENG, Q., PATTERSON, J. B. & SICHERI, F. 2014. Structure and mechanism of action of the hydroxy-aryl-aldehyde class of IRE1 endoribonuclease inhibitors. *Nat Commun*, 5, 4202.
- SANDOW, J. J., DORSTYN, L., O'REILLY, L. A., TAILLER, M., KUMAR, S., STRASSER, A. & EKERT, P. G. 2013. ER stress does not cause upregulation and activation of caspase-2 to initiate apoptosis. *Cell Death And Differentiation*, 21, 475.
- SARRIO, D., RODRIGUEZ-PINILLA, S. M., HARDISSON, D., CANO, A., MORENO-BUENO, G. & PALACIOS, J. 2008. Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. *Cancer Res*, 68, 989-97.
- SCHWEDE, T., SALI, A., ESWAR, N. & PEITSCH, M. C. 2008. Protein Structure Modeling. In: SCHWEDE, T. & PEITSCH, M. C. (eds.) *Computational Structural Biology. Methods and Applications*. Singapore: World Scientific Publishing Co. Pte. Ltd.
- SCHWER, B., SAWAYA, R., HO, C. K. & SHUMAN, S. 2004. Portability and fidelity of RNA-repair systems. *Proc. Natl. Acad. Sci.*, 101, 2788-2793.
- SENKUS, E., KYRIAKIDES, S., OHNO, S., PENAULT-LLORCA, F., POORTMANS, P., RUTGERS, E., ZACKRISSON, S., CARDOSO, F. & COMMITTEE, E. G. 2015. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 26 Suppl 5, v8-30.
- SHA, H., HE, Y., CHEN, H., WANG, C., ZENNO, A., SHI, H., YANG, X., ZHANG, X. & QI, L. 2009. The IRE1alpha-XBP1 pathway of the unfolded protein response is required for adipogenesis. *Cell Metab*, 9, 556-64.
- SHAMU, C. E. & WALTER, P. 1996. Oligomerization and phosphorylation of the Ire1p kinase during intracellular signaling from the endoplasmic reticulum to the nucleus. *EMBO J*, 15, 3028-39.
- SHAO, M., SHAN, B., LIU, Y., DENG, Y., YAN, C., WU, Y., MAO, T., QIU, Y., ZHOU, Y., JIANG, S., JIA, W., LI, J., LI, J., RUI, L., YANG, L. & LIU, Y. 2014. Hepatic IRE1alpha regulates fasting-induced metabolic adaptive programs through the XBP1s-PPARalpha axis signalling. *Nat Commun*, 5, 3528.
- SHAPIRO, A. L., VINUELA, E. & MAIZEL, J. V., JR. 1967. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochem Biophys Res Commun*, 28, 815-20.
- SIDRAUSKI, C., COX, J. S. & WALTER, P. 1996. tRNA Ligase Is Required for Regulated mRNA Splicing in the Unfolded Protein Response. *Cell*, 87, 405-413.
- SIDRAUSKI, C. & WALTER, P. 1997. The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. *Cell*, 90, 1031-9.
- SINGH, J., DENG, Z., NARALE, G. & CHUAQUI, C. 2006. Structural interaction fingerprints: a new approach to organizing, mining, analyzing, and designing protein-small molecule complexes. *Chem Biol Drug Des*, 67, 5-12.
- SIRIN, S., KUMAR, R., MARTINEZ, C., KARMILOWICZ, M. J., GHOSH, P., ABRAMOV, Y. A., MARTIN, V. & SHERMAN, W. 2014. A computational approach to enzyme design: predicting omega-aminotransferase catalytic activity using docking and MM-GBSA scoring. *J Chem Inf Model*, 54, 2334-46.
- SITIA, R. & BRAAKMAN, I. 2003. Quality control in the endoplasmic reticulum protein factory. *Nature*, 426, 891-4.

- 
- SMITH, B. L., BAUER, G. B. & POVIRK, L. F. 1994. DNA damage induced by bleomycin, neocarzinostatin, and melphalan in a precisely positioned nucleosome. Asymmetry in protection at the periphery of nucleosome-bound DNA. *J Biol Chem*, 269, 30587-94.
- SMITH, P., WANG, L. K., NAIR, P. A. & SHUMAN, S. 2012. The adenylyltransferase domain of bacterial Pnkp defines a unique RNA ligase family. *Proc Natl Acad Sci U S A*, 109, 2296-301.
- SMITH, P. K., KROHN, R. I., HERMANSON, G. T., MALLIA, A. K., GARTNER, F. H., PROVENZANO, M. D., FUJIMOTO, E. K., GOEKE, N. M., OLSON, B. J. & KLENK, D. C. 1985. Measurement of protein using bicinchoninic acid. *Anal Biochem*, 150, 76-85.
- SO, J.-S., CHO, S., MIN, S.-H., KIMBALL, S. R. & LEE, A.-H. 2015. IRE1 $\alpha$ -Dependent Decay of CREP/Ppp1r15b mRNA Increases Eukaryotic Initiation Factor 2 $\alpha$  Phosphorylation and Suppresses Protein Synthesis. *Molecular and Cellular Biology*, 35, 2761.
- SO, J.-S., HUR, K. Y., TARRIO, M., RUDA, V., FRANK-KAMENETSKY, M., FITZGERALD, K., KOTELIANSKY, V., LICHTMAN, A. H., IWAWAKI, T., GLIMCHER, L. H. & LEE, A.-H. 2012a. Silencing of lipid metabolism genes through IRE1 $\alpha$ -mediated mRNA decay lowers plasma lipids in mice. *Cell metabolism*, 16, 487-499.
- SO, J. S., HUR, K. Y., TARRIO, M., RUDA, V., FRANK-KAMENETSKY, M., FITZGERALD, K., KOTELIANSKY, V., LICHTMAN, A. H., IWAWAKI, T., GLIMCHER, L. H. & LEE, A. H. 2012b. Silencing of lipid metabolism genes through IRE1 $\alpha$ -mediated mRNA decay lowers plasma lipids in mice. *Cell Metab*, 16, 487-99.
- SON, S. M., BYUN, J., ROH, S. E., KIM, S. J. & MOOK-JUNG, I. 2014. Reduced IRE1 $\alpha$  mediates apoptotic cell death by disrupting calcium homeostasis via the InsP3 receptor. *Cell Death Dis*, 5, e1188.
- SONG, Y., SRETAVAN, D., SALEGIO, E. A., BERG, J., HUANG, X., CHENG, T., XIONG, X., MELTZER, S., HAN, C., NGUYEN, T. T., BRESNAHAN, J. C., BEATTIE, M. S., JAN, L. Y. & JAN, Y. N. 2015. Regulation of axon regeneration by the RNA repair and splicing pathway. *Nat Neurosci*, 18, 817-25.
- SORLIE, T., PEROU, C. M., TIBSHIRANI, R., AAS, T., GEISLER, S., JOHNSEN, H., HASTIE, T., EISEN, M. B., VAN DE RIJN, M., JEFFREY, S. S., THORSEN, T., QUIST, H., MATESE, J. C., BROWN, P. O., BOTSTEIN, D., LONNING, P. E. & BORRESEN-DALE, A. L. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*, 98, 10869-74.
- SRIBURI, R., BOMMIASAMY, H., BULDAK, G. L., ROBBINS, G. R., FRANK, M., JACKOWSKI, S. & BREWER, J. W. 2007. Coordinate regulation of phospholipid biosynthesis and secretory pathway gene expression in XBP-1(S)-induced endoplasmic reticulum biogenesis. *J Biol Chem*, 282, 7024-34.
- SRIBURI, R., JACKOWSKI, S., MORI, K. & BREWER, J. W. 2004. XBP1: a link between the unfolded protein response, lipid biosynthesis, and biogenesis of the endoplasmic reticulum. *The Journal of Cell Biology*, 167, 35-41.
- STARK, A. 2018. FDA approves first treatment for breast cancer with a certain inherited genetic mutation. U.S. Food and Drug Administration.
- STORNILOLO, A., RACITI, M., CUCINA, A., BIZZARRI, M. & DI RENZO, L. 2015. Quercetin affects Hsp70/IRE1 $\alpha$  mediated protection from death induced by endoplasmic reticulum stress. *Oxid Med Cell Longev*, 2015, 645157.
- SUN, Y., JIANG, F., PAN, Y., CHEN, X., CHEN, J., WANG, Y., ZHENG, X. & ZHANG, J. 2018. XBP1 promotes tumor invasion and is associated with poor prognosis in oral squamous cell carcinoma. *Oncol Rep*, 40, 988-998.
- TAM, A. B., KOONG, A. C. & NIWA, M. 2014. Ire1 Has Distinct Catalytic Mechanisms for XBP1/HAC1 Splicing and RIDD. *Cell reports*, 9, 850-858.

- 
- TANAKA, N., CHAKRAVARTY, A. K., MAUGHAN, B. & SHUMAN, S. 2011a. Novel mechanism of RNA repair by RtcB via sequential 2',3'-cyclic phosphodiesterase and 3'-Phosphate/5'-hydroxyl ligation reactions. *J. Biol. Chem.*, 286, 43134-43143.
- TANAKA, N., MEINEKE, B. & SHUMAN, S. 2011b. RtcB, a novel RNA ligase, can catalyze tRNA splicing and HAC1 mRNA splicing in vivo. *J. Biol. Chem.*, 286, 30253-30257.
- TANAKA, N. & SHUMAN, S. 2011. RtcB is the RNA ligase component of an Escherichia coli RNA repair operon. *J. Biol. Chem.*, 286, 7727-7731.
- TASHIRO, E., HIRONIWA, N., KITAGAWA, M., FUTAMURA, Y., SUZUKI, S., NISHIO, M. & IMOTO, M. 2007. Trierixin, a novel Inhibitor of ER stress-induced XBP1 activation from Streptomyces sp. 1. Taxonomy, fermentation, isolation and biological activities. *J Antibiot (Tokyo)*, 60, 547-53.
- TAVERNIER, S. J., OSORIO, F., VANDERSARREN, L., VETTERS, J., VANLANGENAKKER, N., VAN ISTERDAEL, G., VERGOTE, K., DE RYCKE, R., PARTHOENS, E., VAN DE LAAR, L., IWAWAKI, T., DEL VALLE, J. R., HU, C. C., LAMBRECHT, B. N. & JANSSENS, S. 2017. Regulated IRE1-dependent mRNA decay sets the threshold for dendritic cell survival. *Nat Cell Biol*, 19, 698-710.
- TEMIN, H. M. & MIZUTANI, S. 1970. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature*, 226, 1211-3.
- THAMSEN, M., GHOSH, R., AUYEUNG, V. C., BRUMWELL, A., CHAPMAN, H. A., BACKES, B. J., PERARA, G., MALY, D. J., SHEPPARD, D. & PAPA, F. R. 2018. Small molecule inhibition of IRE1 $\alpha$  kinase/RNase has anti-fibrotic effects in the lung. *bioRxiv*.
- THORPE, J. A. & SCHWARZE, S. R. 2010. IRE1 $\alpha$  controls cyclin A1 expression and promotes cell proliferation through XBP-1. *Cell Stress Chaperones*, 15, 497-508.
- TIRASOPHON, W., LEE, K., CALLAGHAN, B., WELIHINDA, A. & KAUFMAN, R. J. 2000. The endoribonuclease activity of mammalian IRE1 autoregulates its mRNA and is required for the unfolded protein response. *Genes Dev*, 14, 2725-36.
- TIRASOPHON, W., WELIHINDA, A. A. & KAUFMAN, R. J. 1998. A stress response pathway from the endoplasmic reticulum to the nucleus requires a novel bifunctional protein kinase/endoribonuclease (Ire1p) in mammalian cells. *Genes Dev*, 12, 1812-24.
- TIROSH, B., IWAKOSHI, N. N., GLIMCHER, L. H. & PLOEGH, H. L. 2005. XBP-1 specifically promotes IgM synthesis and secretion, but is dispensable for degradation of glycoproteins in primary B cells. *The Journal of Experimental Medicine*, 202, 505.
- TIROSH, B., IWAKOSHI, N. N., GLIMCHER, L. H. & PLOEGH, H. L. 2006. Rapid turnover of unspliced Xbp-1 as a factor that modulates the unfolded protein response. *J Biol Chem*, 281, 5852-60.
- TORRIE, G. M. & VALLEAU, J. P. 1977. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.*, 23, 187-199.
- TOWBIN, H., STAHELIN, T. & GORDON, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A*, 76, 4350-4.
- TRAINA, T. A., MILLER, K., YARDLEY, D. A., O'SHAUGHNESSY, J., CORTES, J., AWADA, A., KELLY, C. M., TRUDEAU, M. E., SCHMID, P., GIANNI, L., GARCÍA-ESTEVEZ, L., NANDA, R., ADEMUYIWA, F. O., CHAN, S., STEINBERG, J. L., BLANEY, M. E., TUDOR, I. C., UPPAL, H., PETERSON, A. C. & HUDIS, C. A. 2015. Results from a phase 2 study of enzalutamide (ENZA), an androgen receptor (AR) inhibitor, in advanced AR+ triple-negative breast cancer (TNBC). *Journal of Clinical Oncology*, 33, 1003-1003.
- TUFANLI, O., TELKOPARAN AKILLILAR, P., ACOSTA-ALVEAR, D., KOCATURK, B., ONAT, U. I., HAMID, S. M., CIMEN, I., WALTER, P., WEBER, C. & ERBAY, E. 2017. Targeting IRE1 with small molecules counteracts progression of atherosclerosis. *Proc Natl Acad Sci U S A*, 114, E1395-E1404.

- 
- TUTT, A., ROBSON, M., GARBER, J. E., DOMCHEK, S. M., AUDEH, M. W., WEITZEL, J. N., FRIEDLANDER, M., ARUN, B., LOMAN, N., SCHMUTZLER, R. K., WARDLEY, A., MITCHELL, G., EARL, H., WICKENS, M. & CARMICHAEL, J. 2010. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. *Lancet*, 376, 235-44.
- UNCIULEAC, M. C., GOLDGUR, Y. & SHUMAN, S. 2015. Structure and two-metal mechanism of a eukaryal nick-sealing RNA ligase. *Proc Natl Acad Sci U S A*, 112, 13868-73.
- UNCIULEAC, M. C. & SHUMAN, S. 2015. Characterization of a novel eukaryal nick-sealing RNA ligase from *Naegleria gruberi*. *RNA*, 21, 824-32.
- UPTON, J. P., AUSTGEN, K., NISHINO, M., COAKLEY, K. M., HAGEN, A., HAN, D., PAPA, F. R. & OAKES, S. A. 2008. Caspase-2 cleavage of BID is a critical apoptotic signal downstream of endoplasmic reticulum stress. *Mol Cell Biol*, 28, 3943-51.
- UPTON, J. P., WANG, L., HAN, D., WANG, E. S., HUSKEY, N. E., LIM, L., TRUITT, M., MCMANUS, M. T., RUGGERO, D., GOGA, A., PAPA, F. R. & OAKES, S. A. 2012. IRE1alpha cleaves select microRNAs during ER stress to derepress translation of proapoptotic Caspase-2. *Science*, 338, 818-22.
- VAN ANKEN, E., PINCUS, D., COYLE, S., ARAGON, T., OSMAN, C., LARI, F., GOMEZ PUERTA, S., KORENNYKH, A. V. & WALTER, P. 2014. Specificity in endoplasmic reticulum-stress signaling in yeast entails a step-wise engagement of HAC1 mRNA to clusters of the stress sensor Ire1. *Elife*, 3, e05031.
- VAN GUNSTEREN, W. F., BILLETER, S.R., EISING, A.A., HÜNENBERGER, P.H., KRÜGER, P., MARK, A.E., SCOTT, W.R.P., TIRONI, I.G. 1996. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*.
- VAN OVERBEEK, M. & DE LANGE, T. 2006. Apollo, an Artemis-related nuclease, interacts with TRF2 and protects human telomeres in S phase. *Curr Biol*, 16, 1295-302.
- VANDEWYNCKEL, Y. P., LAUKENS, D., GEERTS, A., BOGAERTS, E., PARIDAENS, A., VERHELST, X., JANSSENS, S., HEINDRYCKX, F. & VAN VLIERBERGHE, H. 2013. The paradox of the unfolded protein response in cancer. *Anticancer Res*, 33, 4683-94.
- VIERI, M., SALIMI, A., PATTERSON, J. B., SAMALI, A., CHEVET, E., BRÜMMENDORF, T. H., APPELMANN, I., KHARABI, B. 2017. Preclinical combination of a novel IRE1 RNase inhibitor MKC-8866 and tyrosine kinase inhibition acts synergistic in acute lymphoblastic leukemia. *European Hematology Association*.
- VOLKAMER, A., GRIEWEL, A., GROMBACHER, T. & RAREY, M. 2010. Analyzing the topology of active sites: on the prediction of pockets and subpockets. *J Chem Inf Model*, 50, 2041-52.
- VOLKAMER, A., KUHN, D., RIPPMANN, F. & RAREY, M. 2012. DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics*, 28, 2074-5.
- VOLKMANN, K., LUCAS, J. L., VUGA, D., WANG, X., BRUMM, D., STILES, C., KRIEBEL, D., DER-SARKISSIAN, A., KRISHNAN, K., SCHWEITZER, C., LIU, Z., MALYANKAR, U. M., CHIOVITTI, D., CANNY, M., DUROCHER, D., SICHERI, F. & PATTERSON, J. B. 2011. Potent and selective inhibitors of the inositol-requiring enzyme 1 endoribonuclease. *J Biol Chem*, 286, 12743-55.
- VOLMER, R., VAN DER PLOEG, K. & RON, D. 2013. Membrane lipid saturation activates endoplasmic reticulum unfolded protein response transducers through their transmembrane domains. *Proc Natl Acad Sci U S A*, 110, 4628-33.
- WALTER, P. & RON, D. 2011. The unfolded protein response: from stress pathway to homeostatic regulation. *Science*, 334, 1081-6.
- WANG, L. K., NANDAKUMAR, J., SCHWER, B. & SHUMAN, S. 2007. The C-terminal domain of T4 RNA ligase 1 confers specificity for tRNA repair. *RNA*, 13, 1235-44.

- 
- WANG, R., FANG, X., LU, Y. & WANG, S. 2004. The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J Med Chem*, 47, 2977-80.
- WANG, Y., HAI, T., LIU, Z., ZHOU, S., LV, Z., DING, C., LIU, L., NIU, Y., ZHAO, X., TONG, M., WANG, L., JOUNEAU, A., ZHANG, X., JI, W. & ZHOU, Q. 2010. HSPC117 deficiency in cloned embryos causes placental abnormality and fetal death. *Biochem Biophys Res Commun*, 397, 407-12.
- WEBER, K. & OSBORN, M. 1969. The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. *J Biol Chem*, 244, 4406-12.
- WELIHINDA, A. A. & KAUFMAN, R. J. 1996. The unfolded protein response pathway in *Saccharomyces cerevisiae*. Oligomerization and trans-phosphorylation of Ire1p (Ern1p) are required for kinase activation. *J Biol Chem*, 271, 18181-7.
- WILHELM, M., SCHLEGL, J., HAHNE, H., GHOLAMI, A. M., LIEBERENZ, M., SAVITSKI, M. M., ZIEGLER, E., BUTZMANN, L., GESSULAT, S., MARX, H., MATHIESON, T., LEMEER, S., SCHNATBAUM, K., REIMER, U., WENSCHUH, H., MOLLENHAUER, M., SLOTTA-HUSPENINA, J., BOESE, J. H., BANTSCHIEFF, M., GERSTMAYER, A., FAERBER, F. & KUSTER, B. 2014. Mass-spectrometry-based draft of the human proteome. *Nature*, 509, 582-7.
- WISEMAN, R. L., ZHANG, Y., LEE, K. P., HARDING, H. P., HAYNES, C. M., PRICE, J., SICHERI, F. & RON, D. 2010. Flavonol activation defines an unanticipated ligand-binding site in the kinase-RNase domain of IRE1. *Mol Cell*, 38, 291-304.
- WONG, K. Y. & YORK, D. M. 2012. Exact Relation between Potential of Mean Force and Free-Energy Profile. *J. Chem. Theory Comput.*, 8, 3998-4003.
- WONG, T. K. & NEUMANN, E. 1982. Electric field mediated gene transfer. *Biochem Biophys Res Commun*, 107, 584-7.
- XIANG, Z. 2016. Advances in Homology Protein Structure Modeling. *Curr. Protein Pept. Sci.*, 7, 217-227.
- YAMAMOTO, K., SATO, T., MATSUI, T., SATO, M., OKADA, T., YOSHIDA, H., HARADA, A. & MORI, K. 2007. Transcriptional Induction of Mammalian ER Quality Control Proteins Is Mediated by Single or Combined Action of ATF6 $\alpha$  and XBP1. *Developmental Cell*, 13, 365-376.
- YANAGITANI, K., IMAGAWA, Y., IWAWAKI, T., HOSODA, A., SAITO, M., KIMATA, Y. & KOHNO, K. 2009. Cotranslational targeting of XBP1 protein to the membrane promotes cytoplasmic splicing of its own mRNA. *Mol Cell*, 34, 191-200.
- YARDLEY, D. A., SHIPLEY, D. L., PEACOCK, N. W., SHASTRY, M., MIDHA, R., PRIEGO, V. M. & HAINSWORTH, J. D. 2015. Phase I/II trial of neoadjuvant sunitinib administered with weekly paclitaxel/carboplatin in patients with locally advanced triple-negative breast cancer. *Breast Cancer Res Treat*, 152, 557-67.
- YOSHIDA, H., MATSUI, T., YAMAMOTO, A., OKADA, T. & MORI, K. 2001. XBP1 mRNA Is Induced by ATF6 and Spliced by IRE1 in Response to ER Stress to Produce a Highly Active Transcription Factor. *Cell*, 107, 881-891.
- YOSHIDA, H., OKU, M., SUZUKI, M. & MORI, K. 2006. pXBP1(U) encoded in XBP1 pre-mRNA negatively regulates unfolded protein response activator pXBP1(S) in mammalian ER stress response. *J Cell Biol*, 172, 565-75.
- YU, F., YAO, H., ZHU, P., ZHANG, X., PAN, Q., GONG, C., HUANG, Y., HU, X., SU, F., LIEBERMAN, J. & SONG, E. 2007. let-7 regulates self renewal and tumorigenicity of breast cancer cells. *Cell*, 131, 1109-23.
- YU, R., TABASSUM, N. & JIANG, T. 2016. Investigation of alpha-conotoxin unbinding using umbrella sampling. *Bioorg. Med. Chem. Lett.*, 26, 1296-300.
- ZAMBELLI, A., MONGIARDINI, E., VILLEGAS, S. N., CARRI, N. G., BOOT-HANDFORD, R. P. & WALLIS, G. A. 2005. Transcription factor XBP-1 is expressed during osteoblast

- 
- differentiation and is transcriptionally regulated by parathyroid hormone (PTH). *Cell Biol Int*, 29, 647-53.
- ZHANG, C., CHAN, C. M., WANG, P. & HUANG, R. H. 2012. Probing the substrate specificity of the bacterial Pnkp/Hen1 RNA repair system using synthetic RNAs. *RNA*, 18, 335-44.
- ZHANG, J. H., CHUNG, T. D. & OLDENBURG, K. R. 1999. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen*, 4, 67-73.
- ZHENG, W., THORNE, N. & MCKEW, J. C. 2013. Phenotypic screens as a renewed approach for drug discovery. *Drug Discov Today*, 18, 1067-73.
- ZHOU, J., LIU, C. Y., BACK, S. H., CLARK, R. L., PEISACH, D., XU, Z. & KAUFMAN, R. J. 2006. The crystal structure of human IRE1 luminal domain reveals a conserved dimerization interface required for activation of the unfolded protein response. *Proc Natl Acad Sci U S A*, 103, 14343-8.
- ZHU, J., MISHRA, R. K., SCHILTZ, G. E., MAKANJI, Y., SCHEIDT, K. A., MAZAR, A. P. & WOODRUFF, T. K. 2015. Virtual High-Throughput Screening To Identify Novel Activin Antagonists. *J Med Chem*, 58, 5637-48.
- ZHU, Y., SINGH, B., HEWITT, S., LIU, A., GOMEZ, B., WANG, A. & CLARKE, R. 2006. Expression patterns among interferon regulatory factor-1, human X-box binding protein-1, nuclear factor kappa B, nucleophosmin, estrogen receptor-alpha and progesterone receptor proteins in breast cancer tissue microarrays. *Int J Oncol*, 28, 67-76.