

University of Alberta

Dimerization of the DEAD-Box Cyanobacterial RNA Helicase Redox,  
CrhR

by

Reem M. Skeik

A thesis submitted to the Faculty of Graduate Studies and Research  
in partial fulfillment of the requirements for the degree of

Master of Science

in

Molecular Biology and Genetics

Department of Biological Sciences

©Reem M. Skeik

Fall 2012

Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

## Abstract

The DEAD-box cyanobacterial RNA helicase redox, or CrhR, in *Synechocystis* sp. PCC 6803 is capable of unwinding dsRNA and in annealing ssRNA in a bidirectional ATP-dependent manner. This is a feature shared by only four other DEAD-box RNA helicases. Two of which, the eukaryotic p68 and p72 proteins, were also shown to self-dimerize. Self-dimerization is a characteristic rarely possessed by an RNA helicase. In this study, CrhR was found to exhibit self-interaction using the yeast two-hybrid system and differentially tagged-CrhR protein exchange (swap) analysis, with the dimerization domain localized to the N-terminus, and some assistance, or partial dimerization occurring through the C-terminus. FPLC analysis also revealed CrhR dimerization to occur in an RNA-independent manner. In addition to FPLC analysis, mass spectrometry also suggests CrhR interaction with protein complexes *in vivo*. These findings suggest physiological functions for CrhR association with ribosomes in multi-subunit complexes upon acclimatization of *Synechocystis* cells to low temperature.

## Acknowledgments

My grateful appreciation and thanks go first off to my supervisor, Dr. George W. Owtrim, for his vital encouragement, guidance and support throughout the course of my degree. I am thankful to have had a great mentor who not only pushed me to excel in my research but also made me realize the joy of teaching.

I would also like to thank my supervisory committee, Dr. Brenda K. Leskiw and Dr. Richard P. Fahlman for their valuable suggestions and advice. I would like to extend my thanks to Dr. Fahlman and David Kramer for help with Mass Spectrometry analysis.

Great thanks to Troy Locke in MBSU, for his genuine assistance and patience during FPLC analysis. He remained courteous after several failed promises of “this is my last sample” and he kept a smile on his face even after having to order new columns twice because of my samples. I would also like to thank Dr. Dana Chamot for her assistance with the many new techniques I was able to learn in the lab, Dr. Claudio De Virgilio for the yeast two-hybrid strains, Jeremy Wideman for help with BN-PAGE analysis and Matthew Schellenburg for help with Crystallization and trypsin digests and just allowing me to pick his brains when I didn’t have an answer for my results.

This journey would not have been as enjoyable without the support and good sense of humor of both my lab mates, Oxana Tarassova and Albert Rosana. Thank you for the great long conversations and the pleasant memories.

And finally, none of this would have been possible without the great deal of support from my family and friends, most especially my parents, who kept me sane and satisfied all my life. I love you. And foremost, to God, who made all things possible.

## Table of Contents

I. INTRODUCTION	1
1.1 Helicases: Structural Characterization and Superfamilies	1
1.1.1 DNA and RNA helicase families	2
1.1.2 The identification of structural themes and superfamilies	4
1.1.3 DEAD-Box RNA helicases and the catalytic core	11
1.1.4 Cellular roles of DEAD-box helicases	16
1.2 Accessory domains and SF2 helicase functionality	21
1.2.1 Specificity and catalytic functionality	22
1.2.2 RNA helicase dimerization	25
1.3 Helicase Mechanisms	27
1.3.1 Active and passive states of helicases	27
1.3.2 Helicase translocation models	32
1.4 Cyanobacterial DEAD-box RNA helicases	37
1.4.1 <i>Synechocystis</i> sp. PCC 6803	39
1.4.2 DEAD-box cyanobacterial RNA helicase redox, CrhR	40
1.5 Objectives	42
II. MATERIALS AND METHODS	45
2.1 Strains, Plasmids and Growth Conditions	45
2.1.1 Yeast strains and plasmids	45
2.1.2 Growth and maintenance of yeast strains	45
2.1.3 <i>E. coli</i> strains and plasmids	47
2.1.4 Growth and maintenance of <i>E. coli</i> strains	47
2.1.5 Growth conditions of <i>Synechocystis</i> sp. PCC 6803	48
2.2 Yeast Two-Hybrid Analysis	48
2.2.1 Cloning into pEG202 and pJG4-5	48
2.2.2 Chemical transformation of EGY48 using LiOAc	50
2.2.3 Yeast two-hybrid $\beta$ -Galactosidase screening assay	52
2.2.4 Yeast two-hybrid leucine selection assay	53

2.2.5	Overexpression protein analysis of EGR-3' and JGR-3'	53
2.3	Tagged-CrhR Protein Exchange or Swap Analysis	54
2.3.1	Gateway® fusion protein constructs: GST-R and MBP-R	54
2.3.2	Construction of the 6xHis-tagged CrhR fusion protein: His-R	55
2.3.3	Construction of the mutant fusion proteins: His-C2R and His-N2R	56
2.3.4	PCR amplification	57
2.3.5	<i>E. coli</i> BL21 transformation	58
2.3.6	DNA Sequencing	58
2.3.7	Tagged protein expression	59
2.3.8	Affinity purification	60
2.3.9	Tagged-CrhR Exchange Assay	61
2.4	FPLC – Gel Filtration Analysis	61
2.4.1	Affinity purification of His-R (+/-) RNase A	61
2.4.2	Protein isolation from <i>Synechocystis</i> sp. PCC 6803 (+/-) RNase A	64
2.4.3	Gel filtration	64
2.4.4	TCA and acetone protein precipitation	65
2.5	Ribosomal Preparation	66
2.5.1	Ribosome isolation <i>Synechocystis</i> sp. PCC 6803	66
2.5.2	Ultracentrifugation of ribosomal proteins	66
2.6	Protein Electrophoresis	67
2.6.1	SDS-polyacrylamide gel electrophoresis (SDS-PAGE)	67
2.6.2	Coomassie staining SDS-PAGE	68
2.6.3	Blue native-polyacrylamide gel electrophoresis (BN-PAGE)	69
2.6.4	Western blot analysis	70
2.6.5	Far-western analysis	71
2.7	Mass Spectrometry Analysis	72
2.7.1	Sample preparation from <i>Synechocystis</i> sp. PCC 6803	72
2.7.2	Trypsin digest of samples	72
2.7.3	Liquid chromatography mass spectrometry analysis (LC-MS)	74

III. RESULTS	76
3.1 CrhR Self-Dimerization through the N-terminal Domain	76
3.1.1 Yeast two-hybrid interaction in $\beta$ -Galactosidase assay	76
3.1.2 Yeast two-hybrid interaction on plates lacking leucine	78
3.1.3 Self-interaction analysis using tagged-CrhR (swap analysis)	79
3.1.4 Swap analysis using non-denaturing BN-PAGE	85
3.1.5 Far-western analysis using purified His-R as a probe	85
3.1.6 Eliminating non-specific binding of GST-R, MBP-R and His-R	87
3.1.7 FPLC-gel filtration analysis <i>in vitro</i> using His-R	93
3.2 Localization of the self-dimerization domain of CrhR	100
3.2.1 Yeast two-hybrid analysis to identify the dimerization domain	100
3.2.2 Yeast two-hybrid analysis of the C-terminal domain and C-terminal extension	102
3.2.3 Truncation of the extended C-terminal domain, His-C2R	106
3.2.4 Partial N-terminal truncation, His-N2R	109
3.3 CrhR Interaction with Protein Complex	112
3.3.1 RNA-independent multimerization of CrhR <i>in vivo</i> and <i>in vitro</i>	112
3.3.2 Temperature-dependent multimerization of CrhR <i>in vivo</i>	120
3.3.3 Potential CrhR and ribosomal protein interactions	127
IV. DISCUSSION	132
4.1 CrhR RNA-Independent Dimerization <i>In Vitro</i> and <i>In Vivo</i>	133
4.2 Potential Functions for CrhR DEAD-Box RNA Helicase	141
V. REFERENCES	161

## List of Tables

### II. Materials and Methods

Table 2.1	List of Yeast Two-Hybrid Strains	46
Table 2.2	Primers used for Yeast Two-Hybrid Constructs	51

### III. Results

Table 3.1	Yeast two-hybrid interactions on S-Gal and (-) Leucine plates	77
-----------	---	----

## List of Figures

### I. Introduction

Figure 1.1	Conserved helicase cores of superfamilies SF1-SF6	5
Figure 1.2	Identification of SF2 families/groups	9
Figure 1.3	DEAD-box RNA helicase core	12
Figure 1.4	Models of helicase translocation	34

### II. Materials and Methods

Figure 2.1	Structure of Cyanobacterial <i>Synechocystis</i> sp. PCC 6803 DEAD-box RNA helicase redox, CrhR	49
Figure 2.2	Tagged-CrhR Protein Exchange - Swap Analysis	62
Figure 2.3	Mass spectrometry analyses of native cold shocked <i>Synechocystis</i> sp. PCC 6803 proteins in a 96 well plate	73

### III. Results

Figure 3.1	Dimerization of three differently tagged-CrhR in protein exchange analysis	80
Figure 3.2	In-solution interaction of affinity purified His-R and MBP-R	83
Figure 3.3	Self-dimerization of His-R and MBP-R on BN-PA	86
Figure 3.4	Far-western self-dimerization analysis of MBP-R using a purified His-R probe	88
Figure 3.5	Controls to exclude non-specific binding of CrhR in protein exchange analysis	91
Figure 3.6	Non-specific controls to rule out binding of the three differently tagged-CrhR to non-respective affinity columns	94

Figure 3.7	FPLC gel filtration analysis of standard molecular weight markers using the Superdex 200BKL column	96
Figure 3.8	FPLC gel filtration analysis of overexpressed His-R using the Superdex 200BKL column	98
Figure 3.9	Protein overexpression analysis of Yeast two-hybrid constructs containing the C-terminal truncation	104
Figure 3.10	Protein overexpression analysis of Yeast two-hybrid constructs containing the extended C-terminal truncation (END)	105
Figure 3.11	Overexpression analysis of His-tagged and truncated CrhR lacking the C-terminal extension	107
Figure 3.12	Self-dimerization of the His-tagged and truncated CrhR lacking the C-terminal extension	108
Figure 3.13	Overexpression analysis of His-tagged and truncated CrhR lacking part of the N-terminal domain	110
Figure 3.14	Self-dimerization of His-tagged and truncated CrhR lacking part of the N-terminal domain	111
Figure 3.15	FPLC gel filtration analysis of the purified His-R treated with RNase A	113
Figure 3.16	FPLC gel filtration analysis of native <i>Synechocystis</i> sp. PCC 6803 proteins with/ or without RNase A treatment	116
Figure 3.17	Coomassie and western blot analysis of native <i>Synechocystis</i> sp. PCC 6803 proteins from FPLC gel filtration analysis	118
Figure 3.18	Coomassie and western blot analysis of native <i>Synechocystis</i> sp. PCC 6803 proteins from FPLC gel filtration analysis treated with RNase A	119
Figure 3.19	FPLC gel filtration analysis of standard molecular weight markers using the Sephacryl S300 column	121

Figure 2.20	FPLC gel filtration analysis of native cold-shocked <i>Synechocystis</i> sp. PCC 6803 proteins using a Sephacryl S300 column	123
Figure 3.21	FPLC gel filtration analysis of native non cold-shocked <i>Synechocystis</i> sp. PCC 6803 proteins using a Sephacryl S300 column	125
Figure 3.22	Coomassie stained and western blot analysis of ultracentrifuged ribosomal proteins extracted from native <i>Synechocystis</i> sp. PCC 6803	130

#### IV. Discussion

Figure 4.1	Phylogenetic analysis of DEAD-box RNA helicases CrhR, CsdA and RhIE.	144
Figure 4.2	A model for CrhR dimerization	156

## List of Abbreviations

AAA <sup>+</sup>	ATPase Associated with various cellular Activities
$\beta$	Beta
$\beta$ -Gal	$\beta$ -Galactosidase
$\beta$ -ME	$\beta$ -Mercaptoethanol or 2-Mercaptoethanol
$\mu$ g	Microgram(s)
$\mu$ L	Microliter(s)
aa	Amino acid(s)
Ab	Antibody
AD	Activation domain
Amp <sub>100</sub>	Ampicillin at 100 $\mu$ g/ml
Amp <sup>R</sup>	Ampicillin resistant
APS	Ammonium persulfate
bp	basepair(s)
Cm <sub>30</sub>	Chloramphenicol at 30 $\mu$ g/ml
Cm <sup>R</sup>	Chloramphenicol resistant
CrhC	Cyanobacterial RNA helix cold
CrhR	Cyanobacterial RNA helix redox
DBD	DNA binding domain
DNA	Deoxyribonucleic acid
DTT/CO <sub>3</sub>	Dithiothreitol and sodium carbonate
dsRNA	Double stranded ribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
FWD	Forward primer
g/L	Grams per liter
GST	Glutathione S-transferase
His	Histidine

HRP	Horse radish peroxidase
IPTG	Isopropyl $\beta$ -D-thiogalactopyranoside
kbp	Kilobasepair(s)
kDa	KiloDalton(s)
LB	Luria Bertani medium
LEU	Leucine
LiOAc	Lithium acetate
M	Molar
mA	Milliamp(s)
MBP	Maltose binding protein
MCS	Multiple cloning site
mg	Milligram(s)
min	Minute(s)
mL	Milliliter(s)
mM	Millimolar
mRNA	Messenger RNA
MW	Molecular weight
ng	Nanogram(s)
nm	Nanometer(s)
OD	Optical density
O/N	Overnight
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
Pre-mRNA	Precursor mRNA
R	CrhR
REV	Reverse primer
RNA	Ribonucleic acid
RNP	Ribonucleoprotein

RNase A	Ribonuclease A
RNPase	Ribonucleoprotein displacement activity
rpm	Revolutions per minute
RT	Room temperature
SDS	Sodium dodecyl sulfate
SF2	Superfamily 2
S-Gal	Galactose induction
ssRNA	Single stranded ribonucleic acid
TAE	Tris-acetate EDTA buffer
TBE	Tris-borate EDTA buffer
TBS	Tris-buffered saline
TBST	Tris-buffered saline and Tween 20
TCA	Trichloroacetic acid
TE	Tris-EDTA buffer
TEMED	Tetramethyl ethylene diamine
T <sub>m</sub>	Melting temperature
Tris	Tris (hydroxymethyl) aminomethane
tRNA	Transfer RNA
U	Unit(s)
URA	Uracil
UTR	Untranslated region
V	Volt(s)
vol	Volume(s)
v/v	Volume per volume
W	Watts
w/o	Without
w/v	Weight per volume
X-Gal	5-bromo-4-chloro-3-indoyl- β-D-galactopyranoside
YNB	Yeast nitrogen base
YPD	Yeast extract Peptone Dextrose growth medium
YTH	Yeast Tryptophan Histidine media

## I. INTRODUCTION

### 1.1 Helicases: Structural Characterization and Superfamilies

Helicases are a class of enzymes that alter the secondary structure and functionality of double-stranded DNA (dsDNA), double-stranded RNA (dsRNA), RNA-DNA duplexes, and most recently been shown to remodel RNA/protein complexes known as ribonucleoproteins (RNPs). Using the energy of ATP-hydrolysis, helicases bind to the double-stranded nucleic-acids (dsNA) and catalyze unwinding of the double strand or alter the structure and composition of RNA-protein complexes (Yang and Jankowsky, 2005; Yodh *et al.*, 2010). All forms of cellular life, including many viruses, encode helicases (Fairman-Williams *et al.*, 2010), whose activity is frequently downregulated until required. When upregulated they may become incorporated into complexes with other protein partners and loaded onto DNA/RNA to perform their function (Singleton *et al.*, 2007).

Helicases are very closely related based on the highly conserved sequence motifs and modular accessory domains. Structure/function analyses of helicases has revealed their implication in many cellular processes; these processes range from DNA replication, recombination, repair and overall genome stability (Hilbert *et al.*, 2009), to regulating gene expression through an association with RNA transcription, editing, splicing, export, translation and ribosome biogenesis (De la Cruz *et al.*, 1999; Cordin *et al.*, 2006). By mixing and matching core motifs and accessory domains of various types, nature has created a diverse range of nucleic

acid processing machinery (Singleton *et al.*, 2007). For that reason, helicases have been classified as DNA or RNA helicase based on their specificity, and further divided into superfamilies, based on their amino acid sequence motifs and structural similarities.

### 1.1.1 DNA and RNA helicase families

Helicases are ubiquitous molecular motors that are employed in virtually all aspects of DNA and/or RNA metabolism (Yodh *et al.*, 2010; Rabhi *et al.*, 2010). They are subdivided into families on the basis of mechanistic differences, which stem from their specificity to DNA, RNA or RNA-DNA hybrids (Singleton *et al.*, 2007). DNA helicases catalyze the unwinding of duplex DNA, the most stable form of DNA *in vivo*. The dsDNA must be unwound transiently in order to allow for the separation of the two ss-DNA intermediates that will thus become substrates for crucial cellular processes such as DNA replication, repair, recombination, as well as DNA transfer during conjugation (Lohman and Bjornson, 1996; Hilbert *et al.*, 2009).

In contrast, dsRNA or RNA-DNA hybrid separation is catalyzed by RNA helicases. The majority of RNA in the cell is synthesized into single-stranded transcripts (ssRNA). During transcription, the ssRNA is known to fold and basepair with itself, creating complex higher order secondary and tertiary structures. The thermodynamic strength of the Watson-Crick base pairing that is formed due to the folding of the ssRNA often makes these structures quite stable (Wong and Pan, 2009). However, this dsRNA needs to undergo extensive

conformational changes, and thus be unwound to allow for dynamic interactions with protein complexes, termed ribonucleoprotein (RNP) complexes in order to function. This remodeling is known to regulate several aspects of gene expression (Chen *et al.*, 2008).

Helicases can either have specificity for RNA or DNA, but some of these enzymes can have specificity to both, and are capable of unwinding RNA-DNA duplexes. This includes proteins from the viral NS3/NPH-II group of helicases (Taylor *et al.*, 2010), eukaryotic RNA helicase A (DHX9) (Zhang *et al.*, 2004), and eukaryotic Upf1-like helicases (Guenther *et al.*, 2009). However, the lack of clear correlation between the helicase families and specificity for RNA or DNA suggests that discrimination between RNA and DNA might not have been a predominant evolutionary force for the differentiation of the helicase families (Jankowsky, 2011).

Recently, there has been extensive work done on helicase substrate specificity, kinetics and mechanisms of unwinding which embody just how conserved different types of helicases are. This conservation is primarily based on the sequence motifs and structural similarities shared by DNA and RNA helicases that results in their specificity and activity, and is the reason why they are considered to be very closely related (Yodh *et al.*, 2010; Fairman-Williams *et al.*, 2010). Therefore, DNA and RNA helicases have been classified into six superfamilies, as a result of their sequence and structural similarities.

### 1.1.2 Identification of structural themes and superfamilies

All proteins with demonstrated helicase activity contain a set of highly conserved motifs which make up the catalytic core responsible for helicase function. It has also been shown that a subset of helicases also possesses additional conserved motifs within the core and outside of it that not only differentiates all helicases structurally but mechanistically as well (Singleton *et al.*, 2007). These motifs are used as reliable identifiers of the respective groups of helicases using database searches.

Pioneering sequence analysis showed that helicases can be classified into several superfamilies based on these sequence-specific structural similarities defined as the “signature” motifs (Gorbalenya and Koonin, 1993). There are 5 superfamily classifications of helicases, SF1 to SF5, and recently the AAA<sup>+</sup> proteins (ATPases Associated with various cellular Activities) has been classified as superfamily six, or SF6 (Singleton *et al.*, 2007). Each superfamily can include a mixture of DNA and RNA helicases, and it is the specific changes in the primary structure of the helicase core and modular accessory domains in the N- and C-terminal domains that are responsible for substrate specificity and biochemical characteristics (Patel and Donmez, 2006).

In the case of all helicases within SF1 to SF6, the catalytic core of a single polypeptide chain is comprised of either two RecA-like domains, in SF1 and SF2 proteins, or one RecA-like domain, in SF3 to SF6 proteins (Figure 1.1). RecA is an *E. coli* DNA helicase involved in rearrangement of dsDNA and strand exchange during repair. The RNA helicase core domains are connected by an

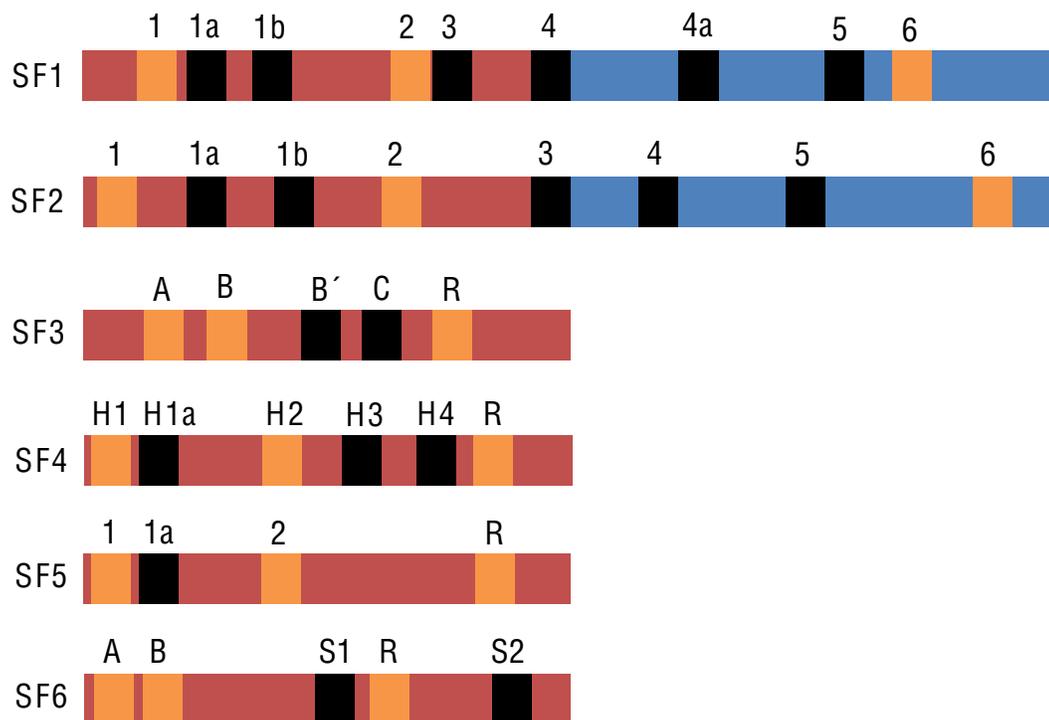


Figure 1.1: Conserved helicase cores of superfamilies SF1-SF6. A schematic representation of the six helicase superfamilies adapted from Singleton *et al.*, 2007. Red and blue bars indicate RecA-like domains or folds of the helicase core; SF1 and SF2 consist of two RecA-like domains, in most cases functioning as a monomeric core, with some functioning as dimers. SF3-SF6 consists of six individual RecA-like domains, and members of these families form hexameric rings. Yellow and black boxes within each RecA-like domain represent the structural motifs within each superfamily; yellow boxes represent conserved structural motifs across all helicase superfamilies involved in ATP-binding and hydrolysis and black boxes represent the RNA binding motifs.

interdomain cleft which allows them to flexibly fold, and these folds resemble the structure of the RecA DNA helicase.

The N-terminal RecA-like domain of the core carries the classical Walker A motif, which is the phosphate-binding motif known as the P-loop and the Walker B motif which is the  $Mg^{2+}$ -binding aspartic acid motif (Walker *et al.*, 1982; Caruthers and McKay, 2002). These motifs are the two well characterized ATP-binding signatures of all helicases which X-ray crystallography shows are located at the cleft of the RecA-like domain interface, and the universal features of the core domain, but can also be found in non-helicase proteins (Bleichert and Baserga, 2007; Rabhi *et al.*, 2010). While they are highly conserved, the sequence of the Walker B motif shows considerable variability among SF1 to SF6 proteins compared to the Walker A motif (Walker *et al.*, 1982; Geourjon *et al.*, 2001; Singleton *et al.*, 2007).

These motifs of the core domain convert chemical to mechanical energy by NTP-binding and hydrolysis to catalyze protein conformational changes (Ye *et al.*, 2004). This affects the activity of the helicase as structural motions result from the altered relative positions of the RecA-like domain and the NA-binding motifs. This leads to a disruptive and/or driving force on the NA substrate (Rabhi *et al.*, 2010). It is believed that the spatial arrangement of the RecA-like domains in helicases gives rise to their specific polarity, as opposed to it being an intrinsic property of the protein (Bird *et al.*, 1998).

While the helicase core is a highly conserved region, there are some minor differences that differentiate helicases from each other structurally; however two

functional organizations also exist based on whether a helicase can oligomerize, which in many cases is required for motor activity (Rabhi *et al.*, 2010).

Superfamilies 3 to 5 show limited sequence homology either with each other or with SF1 and SF2, but make up the toroidal or ring-forming hexameric helicases (Fairman-Williams *et al.*, 2010). SF3 comprises helicases from small DNA and RNA viruses which possess a 3' to 5' polarity and contain a small putative core domain of approximately 100 amino acid residues and three conserved motifs with either a RecA-like core or an AAA<sup>+</sup> core (Caruthers and McKay, 2002; Hickman and Dyda, 2005). SF4 and SF5 contain the 5' to 3' hexameric helicases with five motifs related either to the *E. coli* replicative helicase DnaB, the main member of superfamily 4, or the bacterial transcription termination factor, Rho of superfamily 5 (Brennan *et al.*, 1987; Bird *et al.*, 1998; Singleton *et al.*, 2007).

A number of hexameric motor proteins that contain the AAA<sup>+</sup> core do not fall into SF3, but have been recently classified as SF6 (Singleton *et al.*, 2007). The domain architecture of the AAA<sup>+</sup> proteins, consist of a non-ATPase, N-terminal domain (the N-domain), considered to be the primary RNA or DNA substrate recognition site, followed by either one or two AAA<sup>+</sup> ATPase domains named D1 and D2 (Lupas and Martin, 2002). SF6 are active only as oligomeric assemblies and like SF1 and SF2 possess a 3' to 5' and 5' to 3' polarity.

On the other hand, the two largest helicase superfamilies, SF1 and SF2 comprise the non-ring-forming DNA helicases and the majority of the RNA helicases known (Ogilvie *et al.*, 2003; Klostermeier and Rudolph, 2008; Pyle, 2008; Lehnik-Habrink *et al.*, 2010). They are composed of two RecA-like

domains and are generally monomeric, although some dimeric helicases have been described. SF1 and SF2 members share similar patterns of seven to nine conserved sequence motifs that make up the catalytic core (Figure 1.1), including long poorly conserved spacers found within some of these helicases (Singleton *et al.*, 2007).

The SF2 proteins are divided into 10 families or groups based on the presence or absence of distinct sequence features and characteristic domain organization (Fairman-Williams *et al.*, 2010) (Figure 1.2). Four out of the 10 families were named after a single prototype protein, in which the name of the protein is followed with –like. These four families are the eukaryotic and prokaryotic RecQ-like proteins (Chu and Hickson, 2009), the Ski2-like proteins (Banroques *et al.*, 2008) and the prokaryotic RecG-like proteins (Gorbalenya and Koonin, 1993; Singleton *et al.*, 2001) and RIG-I-like proteins (Nakhaei *et al.*, 2009). The remaining families were named either according to terms already in use, like the eukaryotic and prokaryotic Rad3/XPD (Fan *et al.*, 2008), Swi/Snf (Dürr *et al.*, 2005), the prokaryotic T1R (type 1 restriction enzymes) and the viral NS3/NPH-II proteins (Pang *et al.*, 2002) or according to the single letter amino acid sequence in the corresponding signature motif II, like the eukaryotic and prokaryotic DEAD (asp-glu-ala-asp) and the DEAH (asp-glu-ala-his)-box proteins (Dürr *et al.*, 2006; Fairman-Williams *et al.*, 2010).

Of the ten families within the SF2 RNA helicases, the DEAD-box proteins comprise the largest family, and include hundreds of proteins that range from viral to prokaryotic to eukaryotic organisms (Jankowsky and Putnam, 2010). While

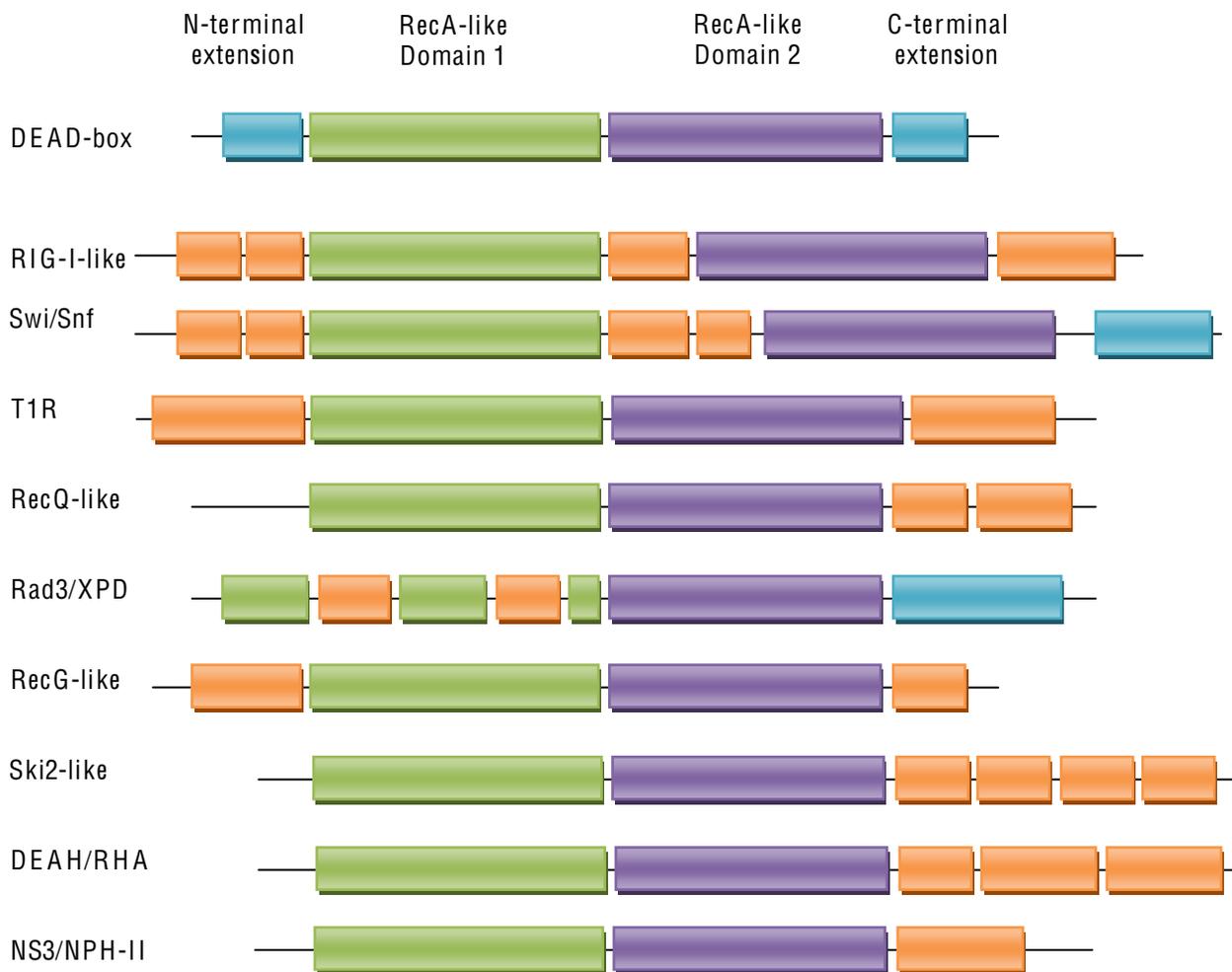


Figure 1.2: Identification of SF2 families/groups. Structural domain and motif organizations of the nine families and one group classified within SF2 helicases (Adapted from Fairman-Williams *et al.*, 2010). The green and purple boxes represent the two RecA-like domains within the highly conserved helicase core or SF2 proteins, and the blue boxes represent the non-conserved N- and C-terminal extensions. Orange boxes represent highly conserved family-typical motifs either inserted within the helicase core or present in the accessory domains of the families/group.

most of the genetic and biochemical analysis obtained is mainly from yeast DEAD-box helicases, it is clear that these helicases play important roles in a variety of crucial cellular processes, altering the secondary structure and functionality of target RNAs, and also remodeling of RNP complexes (Linder, 2006).

### 1.1.3 DEAD-box RNA helicases and the catalytic core

In the late 1980's, several research groups isolated and characterized genes in prokaryotic and eukaryotic organisms ranging from *E. coli* to humans that shared a similar structural homology to the eukaryotic initiation factor, eIF-4A (Linder *et al.*, 1989). This led to the identification of a new family of motor proteins which were termed DEAD-box RNA helicases, due to the highly conserved amino acid sequence asp-glu-ala-asp or D-E-A-D. DEAD-box proteins constitute the largest classification within superfamily 2 (SF2) RNA helicases.

Members of this family share an approximately 400 amino acid core region comprising nine highly conserved signature motifs within two RecA-like domains or folds that affect nucleotide binding, ATP hydrolysis, RNA binding, or RNA unwinding (Figure 1.3). The two RecA-like domains possess two conformations: an open conformation in the absence of ATP and an RNA substrate, in which the two RecA-like domains in the helicase core do not interact. The domains transition from the open conformation to the more rigid closed conformation, when both ATP and RNA are present, which is stabilized upon contacting the RNA (Theissen *et al.*, 2008; Hilbert *et al.*, 2009). Several DEAD-

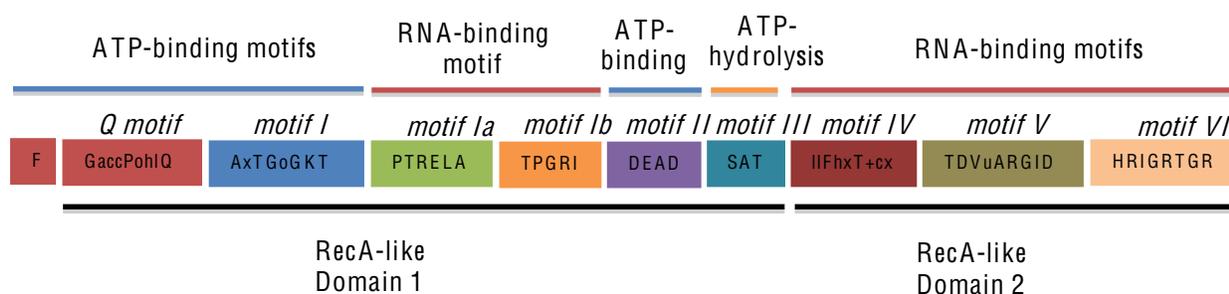


Figure 1.3: DEAD-box RNA helicase core. A schematic representation of the DEAD-box helicase core adapted from Cordin *et al.*, 2006. The DEAD-box family of SF2 RNA helicases is named after the highly conserved motif II. The core consists of nine conserved motifs, and folds into two RecA-like domains. First domain is made up of the Q motif, motif I, Ia, Ib, II and III, while the second domain consists of motifs IV-VI. The Q motif consists of an aromatic group located 17 aa upstream of it (F). The Q motif, along with motif I and II or the DEAD motif, are sites of ATP binding. Motif Ia, Ib, IV-VI are RNA binding sites. While motif III or the SAT motif is required for ATP-hydrolysis.

box helicases have been observed in either the open conformation or the closed conformation, but only structural analysis of the Ddx19 and eIF4A-III DEAD-box helicases has been determined in both open and closed conformation (Andersen *et al.*, 2006; Collins *et al.*, 2009; Von Moeller *et al.*, 2009).

The N-terminal RecA-like domain contains the Q motif, which is unique to DEAD-box proteins, and motifs I-III, while the C-terminal RecA-like domain contains motifs IV-VI. The DEAD-box specific Q motif, upstream of the motif I or the Walker A motif and motif II or the Walker B motif was discovered in 2003 by Tanner and colleagues. The Q motif consists of a nine amino acid sequence with an invariant glutamine, and an isolated aromatic group that is generally located 17 amino acids upstream of the Q motif (Figure 1.3). This motif was originally considered not part of the helicase core; however, after the discovery that the invariant glutamine and the aromatic group interacts with adenine in ATP, it has become one of the conserved motifs that comprises the DEAD-box helicase core (Tanner, 2003; Tanner *et al.*, 2003).

As mentioned previously, the Walker A (motif I) and Walker B motif (motif II), which comprises the DEAD-motif, are the most highly conserved across helicase families. Residues from motifs I and II contact the triphosphate portion of the nucleotide directly, through water and  $Mg^{2+}$  (Hilbert *et al.*, 2009). Motif VI often referred to as the HRIGRxxR box, of the C-terminal RecA-like domain is also expected to interact with the triphosphate in the closed conformation (Pause *et al.*, 1993), and has been thought to be involved in RNA binding as well.

Both RecA-like domains in the closed conformation form the RNA binding site of the helicase core. It involves contacts between the RNA and motif Ia and Ib in the N-terminal domain with the 3'-region of the ssRNA substrate, and to motifs IV, V and VI, in the C-terminal domain with the 5'-region of the ssRNA substrate. Motif Ia and motif IV contact a specific phosphate in ATP, an interaction thought to contribute to the stabilization of the closed conformation (Hilbert *et al.*, 2009).

DEAD-box RNA helicases have a distinct unwinding mechanism; rather than using the energy of ATP-hydrolysis to unwind the RNA duplex, they are directly loaded onto the RNA strand and that leads to the double-strand being pried open with little or no ATP-hydrolysis. This is also referred to as thermal breathing of the duplex, and has been estimated to occur at a rate of about 1,000 times per sec. The protein binds to the duplex, which becomes unstable, and results either in complete thermal melting or in dissociation of the protein from the RNA without unwinding (Rocak and Linder, 2004). This indicates the DEAD-box proteins do not require ATP-hydrolysis to unwind duplexes. As a matter of fact, ATP-hydrolysis is used to release the protein from the substrate to enable recycling, and thus a new round of binding (Rocak and Linder, 2004; Liu *et al.*, 2008). Thus, the dsRNA is unwound without a defined polarity, which explains why many DEAD-box helicases are observed to bidirectionally unwind duplexes, as their mechanism is not based on translocation on one of the nucleic acid strands (Jankowsky, 2011).

This was demonstrated through the use of nonhydrolyzable ATP analogs ADP-beryllium fluoride (ADP-BeF<sub>x</sub>), which mimics the ATP-prehydrolysis state, and the ADP-aluminum fluoride (ADP-AlF<sub>4</sub>), which resembles the transition or the posthydrolysis state, which have been used in mechanistic and structural studies in other ATPase motors including myosin, F1-ATPases and AAA<sup>+</sup> ATPases (Fisher *et al.*, 1995; Kagawa *et al.*, 2004; Chen *et al.*, 2007). DEAD-box proteins unwind the RNA duplexes with the ADP-BeF<sub>x</sub> analog, but not with ADP-AlF<sub>4</sub>, which indicates that ATP binding is necessary for duplex unwinding, while the energy from ATP hydrolysis is dispensable.

However, ATP hydrolysis does appear to be necessary for efficient release of the DEAD-box helicase from the RNA, and thus for multiple substrate turnovers (Liu *et al.*, 2008). The helicase core is mainly in the open conformation in the presence of ADP and RNA, and this suggests that phosphate release from the ADP-Pi state is in fact what triggers the reopening of the interdomain cleft, and thus the release of RNA, as observed by kinetic evidence (Henn *et al.*, 2008). Therefore, it seems that ATP-hydrolysis is required to reset the enzyme to a low affinity for the RNA substrate, which in turn allows for multiple rounds of unwinding, initiated and sometimes even completed by distorting the bound RNA in the ATP-bound state (Hilbert *et al.*, 2009).

In terms of processivity, the distinct unwinding mode of DEAD-box proteins appears uniquely suited for the localized separation of short duplexes in the cell. The efficiency of unwinding depends on the length and stability of the duplex, unwinding greatly decreasing as with increasing duplex length. Most

DEAD-box proteins only unwind duplexes containing less than 10-12 basepairs with appreciable activity, as observed with the three *E. coli* DEAD-box proteins CsdA, RhlE and SrmB (Bizebard *et al.*, 2004), cyanobacterial CrhC (Chamot *et al.*, 1999) and many eukaryotic DEAD-box helicases, as RNAs in eukaryotic cells form few, if any, uninterrupted duplexes that exceed this length (Jankowsky, 2011). This indicates DEAD-box helicase recruitment is a local action to unwind a limited dsRNA or dissociate a protein from the RNA, to allow subsequent steps in a process to occur (Linder, 2006).

Despite the progress that has been made in the recent years on DEAD-box helicases, and RNA helicases in general, many questions remain unanswered regarding the physiological substrates targeted by these helicases. It is important to understand the connections between structural and mechanistic aspects, in order to elucidate the role of helicases in a physiological context. Recent studies have started to reveal the interaction of RNA helicases with their physiological substrates on a physical basis (Fairman-Williams *et al.*, 2010).

#### 1.1.4 Cellular roles of DEAD-box helicases

RNA is considered to be the most structurally and functionally diverse molecule within living cells. Functionality however is dependent on correct folding into the proper tertiary structure, and association with the correct set of proteins in ribonucleoprotein complexes (RNPs). For that, RNAs require “helper” proteins to disrupt these RNA-RNA or RNA-DNA interactions (Bleichert and Baserga, 2007). RNA helicases have received significant attention, ever since

their identification in the 1980s. Many RNA helicases are essential for viability, and a growing number of these enzymes are known to play major regulatory roles in cells (Jankowsky, 2011).

RNA helicases remodel ribonucleoprotein complexes (RNPs), which is why it is no surprise that many of the characterized helicases are involved in essentially every aspect of RNA metabolism. Functions include numerous crucial housekeeping functions such as translation initiation, RNA degradation, pre-mRNA splicing and ribosome biogenesis. In eukaryotic cells, it has been determined that 19 of the 38 DEAD-box helicases in humans and 19 of the 26 in yeast perform roles in pre-mRNA splicing and/or ribosome biogenesis (Fairman-Williams *et al.*, 2010; Jankowsky and Putnam, 2010).

It has been speculated that the dissociation of dsRNA structures is functionally coupled to the formation of new structures in the splicing of pre-mRNA (Rössler *et al.*, 2001). The splicing apparatus is composed of an RNP machine termed the spliceosome, consisting of five RNA components (U1, U2, U4, U5 AND U6) and more than 100 proteins (Murray and Jarrell, 1999). Each round of splicing requires the ordered binding and release of RNPs. Early assembly steps appear to be energy independent; nevertheless, all subsequent steps require ATP hydrolysis (Staley and Guthrie, 1998). This process can take place through temporal modification reactions like phosphorylation, but a more likely explanation is through active remodeling, which in all cases requires one if not more than one highly specific helicase at each step (Linder, 2006).

Splicing of pre-mRNAs has become a paradigm for the analysis of the function of DEAD-box helicases. A total of eight RNA helicases are associated with the spliceosome and required to catalyze specific reactions in eukaryotic mRNA splicing (Linder, 2006). DEAD-box proteins are involved in the establishment of a functional spliceosome, which is essential for splicing to occur (Linder, 2006). In yeast, the well-studied DEAD-box proteins Sub2, Prp28, and Prp5 have been shown to be required for *in vivo* splicing, participating in multiple steps of the spliceosome assembly process (Staley and Guthrie, 1998), which requires the five small nucleolar RNPs (snRNPs) U1, U2, U4, U5 and U6 (Linder, 2006).

In higher eukaryotes, both the p68 and p72 RNA helicases have been implicated in alternative splicing, playing a role in the transition from pre-spliceosome to spliceosome during assembly through pre-mRNA splicing (Liu, 2002; Guil *et al.*, 2003; Honig *et al.*, 2002). Furthermore, p68 and p72 have also been reported as essential players in pre-rRNA maturation, in which p68 was shown to be capable of promoting an RNA structural rearrangement within the pre-60S ribosomal subunit (Jalal *et al.*, 2007). Eukaryotic ribosomes are composed of 4 rRNAs, three of which are transcribed as a large pre-60S rRNA, processed to the three mature 18S, 5.8S and 25S rRNAs during the assembly reaction. In yeast, eight DEAD-box proteins are required for the synthesis of the large ribosomal subunit (Linder, 2006). DEAD-box proteins play a variety of roles in ribosome biogenesis, from reorganizing the pre-ribosomal complexes to

allow new and mutually exclusive RNA-RNA interactions, to remodeling RNPs by altering RNA-protein folding (Linder, 2006).

Ribosome biogenesis engages by far the largest number of helicases in prokaryotes, many of which are members of the DEAD-box family (Bleichert and Baserga, 2007). Four of the five prokaryotic DEAD-box helicases have been implicated in ribosome biogenesis. The *E. coli* DEAD-box RNA helicase CsdA is involved in the biogenesis of the large 50S ribosomal subunit, second to SrmB, which is considered to be involved during an earlier step of 50S assembly (Charollais *et al.*, 2004). Sucrose gradient fractionation indicated that CsdA associated with a pre-50S particle while CsdA deletion resulted in a deficit of 50S subunits. However, its activity appears to be required only at temperatures below 30°C. This is consistent with the proposed role for RNA helicase activity at low temperature, where the RNA structures will be more thermodynamically stable (Charollais *et al.*, 2004).

In addition to pre-mRNA splicing and ribosome biogenesis, many DEAD-box helicases have also been proposed to have additional functions in RNA metabolism, based on their association with certain large RNPs. These functions include and are not limited to nuclear RNA export, turnover and quality control, translation initiation, termination and inhibition, small RNA processing and mitochondrial RNA metabolism. The Ded1 RNA helicase, along with Prp43, Dbp2, Mtr4, and Sub2 are all examples of DEAD-box proteins implicated in multiple cellular functions in yeast (Bleichert and Baserga, 2007).

But an even more surprising example of a multifunctional RNA helicase is the eukaryotic translation initiation factor eIF-4AIII (Linder, 2003). eIF-4AIII is a member of the eIF-4A family. The eIF-4A protein is the prototypical RNA helicase consisting of only the RNA helicase core. It is required for translation initiation of all mRNAs in eukaryotic cells, and is the first and most extensively studied DEAD-box protein. While eIF-4A is only involved in a single cellular role, eIF-4AIII has been implicated in several functions. It is an important component of the exon-junction complex (EJC) in pre-mRNA splicing (Linder, 2006), and also has a role in developmental cell signaling (Li and Li, 2006). This new function is not only independent of its function in cap-dependent translation initiation, but also reveals the role of an RNA helicase in a pathway that does not involve RNA (Li *et al.*, 2005; Li and Li, 2006). This may indicate a non-helicase or enzymatic role for eIF-4A.

It is evident that while many DEAD-box helicases perform several functions in the cell, they carry out these functions in a very specific manner and cannot substitute for each other. This specificity can be explained by interactions between these helicases and accessory proteins, or cofactors (Bleichert and Baserga, 2007). Cofactors can confer substrate specificity and/or increase the affinity of the helicase for its substrate, as well as stimulate or inhibit ATPase and helicase activity (Cordin *et al.*, 2006). A well-known example of cofactor is eIF-4B which interacts with the translation initiation factor eIF-4A in the EJC. The eukaryotic p68 and p72 proteins also play important roles in transcription, often as coactivators or corepressors through their interaction with key components of the

transcriptional machinery (Endoh *et al.*, 1999; Wilson *et al.*, 2004; Fuller-Pace, 2006).

While the helicase motifs that are organized in the core domain are thought to provide the catalytic function of the proteins, it has been proposed that N- and C-terminal sequences flanking the conserved helicase core may be able to confer substrate specificity and cofactor interactions and modulate the RNA helicase activity (Bleichert and Baserga, 2007). These extensions or accessory domains are characterized by a high degree of sequence and length variability (Lüking *et al.*, 1998). A few DEAD-box RNA helicases have been shown to have additional biochemical activities, and some are even capable of forming oligomers. This can possibly be explained by the optional inserts and amino- and carboxy-terminal sequences of DEAD-box RNA helicases (Gorbalenya and Koonin, 1993) that may comprise distinct domains that lead to these interactions and diverse roles.

## 1.2 Accessory Domains and SF2 Helicase Functionality

Within a helicase family, the architecture might differ due to the presence of extended domains which can be found within the N- or C-terminals of the proteins or even as inserts within the core regions. These N- or C-terminal domains can often be larger than the helicase core itself (Singleton *et al.*, 2007; Fairman-Williams *et al.*, 2010). Unlike some accessory domains of the Ski2-like and DEAH families, in which recent work show some structural conservation

between the extended C-terminal between the members (He *et al.*, 2010; Büttner *et al.*, 2007), the accessory domains within the DEAD-box family are not conserved. Many of these accessory domains have been shown to aid in the helicase being targeted to specific cellular locations (Singleton *et al.*, 2001; Linder and Lasko, 2006), or may add a catalytic functionality to the activity of the helicase, such as annealing or strand exchange (Flores-Rozas and Hurwitz, 1993; Valdez *et al.*, 1997; Rössler *et al.*, 2001; Yang and Jankowsky, 2005), or may even promote oligomerization (Ogilvie *et al.*, 2003; Klostermeier and Rudolph, 2008; Lehnik-Habrink, 2010).

### 1.2.1 Specificity and catalytic functionality

N- and C-terminal domains are thought to be critical for the cellular specificity of helicases by facilitating recruitment to specific complexes, either through interactions with other proteins, or by facilitating recognition of specific nucleic acid regions as seen in the bacterial DEAD box protein DbpA (Jankowsky, 2011). The nonprocessive DbpA ATPase and duplex RNA unwinding activities are specific to a 93 base pair region of the 23S rRNA domain V of the ribosomal peptidyltransferase centre (PTC), which was found to be necessary and sufficient for complete activation of ATPase activity (Nicol and Fuller-Pace, 1995). It is to date the only DEAD-box protein to which a specific RNA substrate has been identified, with these specific interactions occurring in the C-terminal region of the protein.

While RNA-binding and unwinding by other DEAD-box proteins shows no substrate specificity, in some cases, some factors exist that direct the roles of these enzymes. In other words, some helicases may be restricted to defined cellular compartments by specific interaction with anchoring proteins (Rössler *et al.*, 2001). The eukaryotic p68 and p72 DEAD-box proteins interact with the nucleolar protein fibrillarin through their C-terminal extensions suggesting a role for them in the activation of ribosomal DNA transcription/preribosomal RNA processing (Nicol *et al.*, 2000). Additionally, many recent reports imply a wide range of functions for p68 and p72, as their interaction with each other (Ogilvie *et al.*, 2003), forming heterodimers, through their C-terminals allows them to interact with different RNA substrates in the cell.

Both p68 and p72 are two of only four eukaryotic DEAD-box helicases that are not only capable of unwinding dsRNA, but also have been shown to anneal complementary ssRNA substrates (Rössler *et al.*, 2001). While the mechanistic manner of p68 and p72 functionality remains to be elucidated, they have been demonstrated to form branch migration structures in an RNA duplex in an ATP-dependent manner utilizing both the annealing and unwinding activities exhibited by the heterodimer (Ogilvie *et al.*, 2003). This additional enzymatic activity suggests a role for these proteins in post-transcriptional regulation, as coupling of both their unwinding and annealing activities to alter secondary structures of pre-mRNA through strand exchange may assist with splicing (Rössler *et al.*, 2001); p68 for instance has been previously detected in spliceosomes (Neubauer *et al.*, 1998).

The RNA annealing activity that results in the formation of RNA duplexes catalyzed by p68 and p72 proteins is unique and is only shared by two other DEAD-box RNA helicases, the eukaryotic Ded1 (Yang and Jankowsky, 2005) and RNA helicase II/Gu (RH-II/Gu). In terms of RNA helicase II/Gu (RH-II/Gu), a domain in the non-conserved C-terminal appears to contain this additional activity (Valdez *et al.*, 1997). Unlike p68 and p72 however, Ded1 and RH-II/Gu do not demonstrate branch migration, which suggests unwinding and annealing activities are distinct. In the case of Ded1, it establishes an ATP-dependent steady state that enables the enzyme to modulate the balance between the two opposing activities through cellular ATP and ADP concentrations (Yang and Jankowsky, 2005).

The human Rh-II/Gu also requires different conditions for its RNA annealing activity. Low concentrations of ATP and  $Mg^{2+}$ , as well as amino acid substitutions in the DEAD motif II appear to enhance the RNA annealing activity, while these conditions are determined to be inhibitory for the RNA unwinding. On the other hand, truncation of 76 amino acids of the C-terminal of the human RNA helicase II/Gu was still able to unwind dsRNA, but was not able to catalyze the annealing of ssRNA (Valdez *et al.*, 1997).

To date, there is no mechanistic understanding of how these DEAD-box helicases coordinate RNA unwinding and annealing activities, but findings suggest an importance of both activities in regards to remodeling RNA and RNP complex structures (Yang and Jankowsky, 2005). Although the physiological functions of these proteins are not yet known, the annealing activity, and in the

case of p68 and p72 RNA strand exchange, could be a means in the cell to compensate for the low processivity observed in these helicases, as they demand resolution of longer RNA duplexes. As identified with respect to the RecA helicase, the protein-induced annealing seems to require more than one nucleic acid binding site within a given enzyme (Kurumizaka *et al.*, 1999), but while hexamer formation is common to several DNA helicases (Rössler *et al.*, 2001), it has not been observed in the SF2 RNA helicases. However, there have been reports of a rare few RNA helicases that are capable of self-interaction to form dimers.

### 1.2.2 RNA helicase dimerization

While the helicase core possibly does not contain any conserved protein-protein interaction motifs, it is most likely the accessory domains in the N- and C-terminal regions that are responsible for protein interaction, as well as self-interaction. While there has not been one case of DEAD-box RNA helicases functioning as hexamers, like DNA helicases, a few DEAD-box helicases have been shown to function as dimers. These helicases also exhibit other biochemical and/or physiological functions not observed in other helicases.

The eukaryotic DEAD-box helicases p68 and p72 interact with each other through the C-terminal domain, and also have been shown recently to self-associate and form dimers through the N-terminal domain (Ogilvie *et al.*, 2003). This suggests that possibly the formation of the homodimers could occur in a somewhat different way than that of the heterodimer, and while it is still unknown

what role the dimer has in cellular processes, it reinforces the various reports implying a wide range of functions for both p68 and p72 in the cell.

Dimerization has also been observed at the prokaryotic level, with the DEAD-box proteins Hera and CshA. Crystal structures of the *Thermus thermophilus* helicase Hera have revealed a C-terminal domain with a novel dimerization motif as well as an RNA binding module outside of the Hera helicase core. They are however connected to the second RecA-like domain by a hinge region that confers added flexibility. The 23S rRNA has been identified as a specific substrate for Hera, and thus it has been suggested that dimerization maybe important in directing unwinding or remodeling of larger RNA molecules that display a higher degree of conformational freedom (Klostermeier and Rudolph, 2008). Similarly, analysis of the *Bacillus subtilis* helicase CshA revealed the C-terminal domain to be crucial for the dimerization of CshA as well as the interaction with components of the RNA degradosome that include RNase Y which appears to be a functional equivalent of the RhlB helicase of the *E. coli* RNA degradosome (Lehnik-Habrink *et al.*, 2010).

While each of these DEAD-box RNA helicases are associated with different functions including transcription regulation, RNA processing and RNA degradation, there has not been a link demonstrated between dimerization and biochemical activity. Therefore it is difficult at this point to connect helicase activity, biochemical activity and oligomerization, but the ability of these proteins to exist as monomers and dimers adds further complexity, and may allow for

alterations in function that can be relevant to other DEAD-box RNA helicases (Ogilvie *et al.*, 2003).

The currently available X-ray structures of some helicases has provided an important advance in terms of structure and function analysis of these motor proteins, emphasizing the close relationship of both the conserved motifs and the three-dimensional structures of the enzymatic core (Tanner and Linder, 2001). Many of these structures do not fully represent the entire length of the protein, and inferences cannot be made until full-length structures, including extended N- or C-terminals, are available (Fairman-Williams *et al.*, 2010). However, all recent data seems to imply that the catalytic cores of the proteins are functionally and mechanistically equivalent even if they have different biological activities and substrates.

### 1.3 Helicase Mechanisms

Most helicase superfamilies are further divided into families or groups based on their family/group-typical mechanistic features. This includes NTP usage versus specificity for adenosine triphosphate, the ability to unwind duplexes, and the polarity of their unwinding activity.

#### 1.3.1 Active and passive states of helicase activity

Helicases are considered a subgroup of motor proteins referred to as translocases, since they couple ATP-hydrolysis to directional movement along single or double-stranded nucleic-acids (Singleton *et al.*, 2007). They have been

proposed to carry out their unwinding activity in two ways: either an active or a passive mechanism (Amaratunga and Lohman, 1993; Geiselmann *et al.*, 1993). Helicases that participate in unwinding of a duplex are referred to as “active”, meaning they can actively destabilize the duplex and stimulate the rate constant for unwinding (Amaratunga and Lohman, 1993; Geiselmann *et al.*, 1993; Pyle, 2008). The physical mechanism of these active helicases is not yet very well defined, but many suggestions have been proposed regarding how destabilization of the duplex occurs. The protein can either spread out over the duplex region and electrostatically destabilizes it, or it can form interactions that lead to the destabilization of the bases. Both cases can exert force or torque on the duplex to pry it apart (Pyle, 2008).

DNA helicases and certain viral RNA helicases have been shown to possess three essential biochemical activities characteristic of a helicase: nucleic acid-dependent (NA-dependent) NTP-hydrolysis, an NTPase-powered unidirectional translocation along the single-stranded nucleic acid (ssNA) and an NTPase-dependent unwinding of the duplex NA (Yodh *et al.*, 2010). One example is the well-studied UvrD DNA helicase, originally known as DNA helicase II, in *E. coli*, which plays a very crucial role in DNA replication, recombination and repair of mismatched bases as a result of ultraviolet (UV) damage (Arthur and Lloyd, 1980; Lee and Yang, 2006). Crystal structures of UvrD complexed with DNA and ATP-hydrolysis intermediates reveal that ATP-hydrolysis leads to unwinding by directional rotation and translation of the DNA

duplex, to which the release of ADP and Pi result in the translocation of the developing strand (Lee and Yang, 2006).

However, the SF2 DEAD-box family proteins (section 1.1.3) are an exception to this mode of action. DEAD-box proteins have been shown to unwind RNA duplexes in an NTP-dependent manner, induced upon binding of the helicase to the RNA, with limited NTP-hydrolysis (Chen *et al.*, 2008; Liu *et al.*, 2008; Hilbert *et al.*, 2009; Yodh *et al.*, 2010). In studies using the eukaryotic DEAD-box RNA helicase Ded1 in *Saccharomyces cerevisiae*, ATP hydrolysis was found to be required for fast enzyme release from the RNA and multiple substrate turnovers, and thus enzyme recycling; whereas ATP binding alone appears to be the necessary step for promoting unwinding of the duplex RNA, achieving complete strand separation by using only a single ATP (Chen *et al.*, 2008; Liu *et al.*, 2008). This is currently the preferred model for these helicases.

In contrast, another mode of helicase activity that was proposed was termed a passive or opportunistic mechanism (Amaratunga and Lohman, 1993; Geiselman *et al.*, 1993). In this mechanism, the helicase simply waits at the junction for transient opening or thermodynamic breathing of the base pairs, which occurs rapidly and spontaneously at duplex termini (Pyle, 2008), meaning that the helicase does not physically unwind the duplex NA, but does couple NTP-hydrolysis to directional motion along the NA. These opportunistic helicases wait for the duplex to separate, usually by thermal fraying, and then they work by locking the strands in the unwound state (Singleton *et al.*, 2007). This type of mechanism involves helicases that can translocate and occupy one base pair at a

time, since the terminal base pairs at a junction open and close at a very fast rate (Patel and Donmez, 2006).

Many SF6 members of the AAA<sup>+</sup> (ATPases Associated with various cellular Activities) protein family (section 1.1.2) fall into this category of helicases (Erzberger and Berger, 2006). The mini chromosome maintenance (MCM) protein complex is the most well-known AAA<sup>+</sup> protein member, postulated to be an archaeal and eukaryotic helicase, and also thought to be the main eukaryotic replicative helicase. The MCM complex is involved in the initiation of replication, as well as the elongation of the DNA and is composed of six proteins known as MCM2-7, of which MCM4, 6 and 7 have been shown to comprise the “putative helicase” activity (Tye, 1999; Erzberger and Berger, 2006).

While the mechanism for strand separation is not-well known for helicases, they are likely influenced by the physical processes involved in both physical translocation and engaging with the duplex. However, some combine aspects of the active and passive mechanisms, assuming that the proteins translocate along a single loading strand in a biased directional manner (Pyle, 2008). Directionality is determined by the direction (5'-3' or 3'-5') of the nucleic acid strand onto which the helicase loads (Singleton *et al.*, 2007). Most helicases unwind in a unidirectional 3'-5' direction, but some have been shown to unwind in the 5'-3' direction, and a few have been shown to bind to either strand and therefore can be considered as a bidirectional helicase, like the Vasa DEAD-box helicase (Linder and Lasko, 2006).

Many helicases display their activities in isolation, but most work efficiently as part of a larger protein complex (Patel and Donmez, 2006). In fact, the processivity of helicases is thought to be determined by their interaction with other proteins in a complex (Singleton *et al.*, 2007). This increase in unwinding efficiency can mean an association of two helicases, like the case of the RecBCD helicase/nuclease complex; RecB was known to possess a 3' to 5' helicase activity, but recent studies also showed that RecD possessed a helicase activity, however in the opposite 5' to 3' direction, thus making RecBCD a bipolar DNA helicase (Dillingham *et al.*, 2003). Association of a helicase with a polymerase can also increase efficiency of unwinding, like the T4 helicase and T7 DNA polymerase; it was shown that the T4 replication helicase and the T7 polymerase coupled system can carry out ATP-dependent strand displacement DNA synthesis at physiological rates of 400 to 500 base pairs per sec (Dong *et al.*, 1996). Even an association of the helicase with a single stranded binding protein can increase its efficiency to unwind a duplex NA (Patel and Donmez, 2006).

While there may be several mechanisms for helicase biochemical activity, it is required for efficiently catalyzing most DNA and RNA metabolic processes in order for helicases to perform their diverse functions in the cell (Patel and Donmez, 2006). How these enzymes physically perform their cellular tasks has been a longstanding question, but what is known to date is that any defects or deregulations of these proteins leads to a number of developmental diseases, including cancer (Fairman-Williams *et al.*, 2010; Jankowsky, 2011). It is important to note that while helicases may differ in their mode of action this is not

the only determining factor for efficiency. The structure of a helicase and its oligomeric state are also determining factors in how the helicase functions. In recent years however, there has been an increase in intriguing models that have started to link the mechanism and structure of helicases to their cellular biological function (Jankowsky, 2011).

### 1.3.2 Helicase translocation models

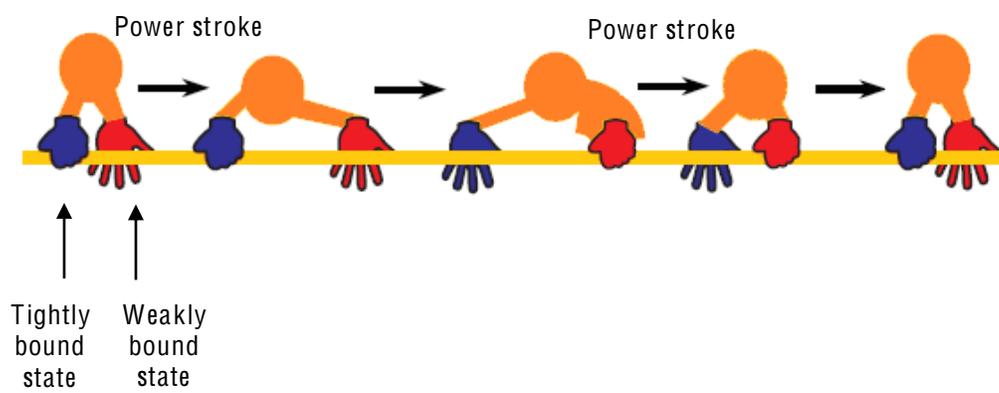
Helicase processivity is critical when it comes to carrying out such cellular functions as genome replication, recombination and repair. Most DNA helicases within SF3-SF6 display ring-shaped architectures and function as hexamers. There are two unique RNA helicases that have also been reported to function as ring-shaped helicases; the bacterial Rho helicase of SF5, which unwinds RNA-DNA duplexes, and the  $\Phi$ 29 P4 viral RNA helicase of SF4 (Brennan *et al.*, 1987; Kainov *et al.*, 2006; Rabhi *et al.*, 2010; Jankowsky, 2011). However, most RNA helicases belong to SF1 and SF2, in which the majority have been shown to function as monomers, although a few can dimerize; therefore, motor processivity may not be required in the case of monomeric helicases, especially if they are not responsible for unwinding more than two helical turns of RNA duplex per enzymatic round (Jankowsky and Fairman, 2007).

There have also been a few DNA helicases that do not appear to require the formation of oligomers to carry out their functions. Structural and functional analysis show RecQ helicase functions as a monomer, as both DNA unwinding activity and ssDNA-stimulated ATPase specific activity were independent of

RecQ concentration (Xu *et al.*, 2003). Some DNA helicases require dimerization for function; Rep DNA helicase in *E. coli* is a stable monomer in the absence of DNA, however it dimerizes when bound to ssDNA or duplex DNA which is its functionally active state, retaining both its DNA-dependent ATPase and helicase activity (Yarranton and Gefter, 1979; Chao and Lohman, 1991; Wong *et al.*, 1992). The UvrD DNA helicase is another example of a DNA helicase that functions as a dimer. In the absence of DNA, UvrD can still self-assemble to form a dimer, however, that state is stabilized further upon binding short oligodeoxynucleotides (Runyon *et al.*, 1993).

As cooperation among multiple monomers may be required for enzymatic efficiency, there have been a number of models suggested for the different oligomeric states of helicases and their translocation mechanisms. One proposed model is the inchworm stepping mechanism, in which two nucleic acid binding sites independently bind and release nucleic acids in response to signals received from the NTPase site (Figure 1.4a). This works for both monomeric and dimeric helicases, which would involve either two-NA binding sites on one helicase, or one-NA binding site from each of the two monomers making up the dimeric helicase, respectively. Thus, coordinated NTPase activity leads to a cycle of NA binding, release and translocation (Patel and Donmez, 2006). UvrD translocates in a strained-inchworm mechanism in which a conformational change that bends and tenses the ssDNA is required to activate the dimer (Sun *et al.*, 2008). PcrA is a monomeric DNA helicase that has also been shown to translocate in an inchworm mechanism (Velankar *et al.*, 1999).

A



B

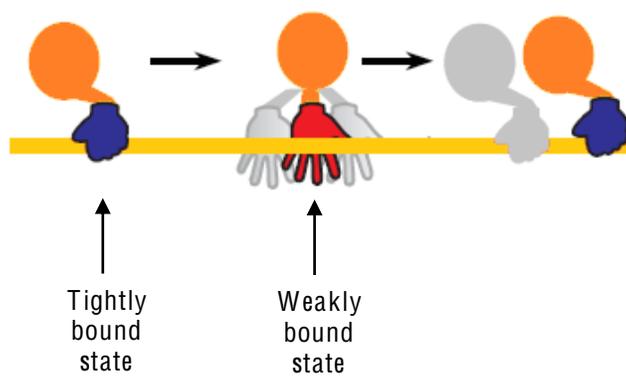


Figure 1.4: Models of helicase translocation. Mechanisms of translocation along nucleic acids by monomeric or oligomeric helicases (adapted from Patel and Donmez, 2006). (A) An inchworm mechanism which involves either monomeric or dimeric helicases with two nucleic acid binding sites (SF1 and SF2). In this mechanism, each binding site domain of the monomer (represented by the blue and red hands) takes on the tightly bound conformation or the weakly bound conformation in order to translocate in a power stroke without dissociation from the nucleic acid. The weakly bound state is therefore capable of sliding across the nucleic acid strand. In the case of a dimer, the two states would be exhibited by each monomer of the dimer complex. (B) In the brownian mechanism, the only binding site would be able to exhibit each of the tightly and weakly conformations. This mechanism is mostly seen in hexameric helicases that form ring structures (SF3-SF6), since the weakly bound state can move in either direction and can dissociate from the nucleic acid strand to allow the other monomer to load and exhibit the tightly bound state.

A second model of translocation proposed is the Brownian motor mechanism, based on two conformational states of the helicase and one binding site, in which one state is weakly bound and the other is tightly bound to the NA strand (Figure 1.4b). These modes result from different NTP ligation states (NTP-hydrolysis, or product release). When tightly bound, movement is restricted, therefore in order to translocate, the helicase changes its NTP ligation state, and thus loosens and can either move in a Brownian motion, or completely dissociates from the nucleic acid strand until it resumes the tight state and makes a step forward or a power stroke (Patel and Donmez, 2006). This model has two basic requirements: first, an oligomeric protein structure with subunits that work cooperatively in pairs and second the ability of each subunit to bind either ds- or ss-DNA, but not both simultaneously (Bird *et al.*, 1998).

In the Brownian motion model, the unidirectional translation of the helicase along one strand of DNA, while excluding the other strand, results in separation. Most, perhaps all, hexameric helicases operate by this model based upon their ability to pass over a bulky substituent when it is present upon one strand but not the other (Enemark and Joshua-Tor, 2008). Hepatitis C virus helicase is an example of a helicase that translocates in a Brownian motor with a simple two stroke cycle. The energy of binding to ssDNA fuels the directional movement, while ATP binding is responsible for the brief period of random movement that prepares the helicase for the next cycle (Levin *et al.*, 2005). The HCV NS3 (non-structural protein 3) helicase unwinds dsDNA and RNA, and can function as a monomer, however, oligomerization does occur and while it is

unsure whether both subunits function independently or cooperatively, this state depends on the relative concentration of DNA (Tackett *et al.*, 2005; Sikora *et al.*, 2008; Jennings *et al.*, 2009). The bacteriophage T4 Dda helicase can catalyze streptavidin displacement from a DNA strand as a monomer, but exhibits increased activity when multiple helicase molecules are bound, and functions in a Brownian motion mechanism (Byrd and Raney, 2004).

Helicases are ubiquitous and central players in DNA and RNA metabolism of organisms ranging from prokaryotes to eukaryotes and including many viruses (Jankowsky, 2011). While much still needs to be uncovered in regards to helicase activity, significant progress has been made linking structure and unwinding mechanisms of many helicases and further progress will lead to better understanding of the functions performed by helicases in several cellular processes.

#### 1.4 Cyanobacterial DEAD-box RNA Helicases

Cyanobacteria were recognized as a biological group early in the 19<sup>th</sup> century. Subject to the rules of the Botanical Code (Statlen, 1972), they were treated as a subdivision of algae. Prior to the 1970's it was most commonly known as blue-green algae or cyanophytes. However, the nomenclature came under the rules of the Bacteriological Code (Lapage *et al.*, 1975) due to two taxonomically significant facts that became evident: first of all, their structural properties, which were typical of prokaryotes as they were unlike those of any other algal groups; and secondly, the prokaryotic structure of the cell constitutes

the only set of properties common to all bacteria, as they are extraordinarily diverse (Stanier and Cohen-Bazire, 1977; Stanier *et al.*, 1978). They are Gram-negative bacteria, as they contain an outer membrane, a peptidoglycan layer and a plasma membrane. They are also obligate photoautotrophs, performing oxygenic photosynthesis similar to that observed in chloroplasts.

They are divided into an array of subgroups or subclasses, orders, families and genera, with over 200 representative strains. Based on their structural properties, these Gram-negative bacteria can be unicellular or filamentous, motile or immotile, rod or spherical-shaped (Stanier and Cohen-Bazire, 1977). In addition, cyanobacteria contain a distinct intracellular membrane system known as thylakoids, which are energy-transducing membranes and the sites for both photosynthesis and respiration (Huang *et al.*, 2002). They constitute the largest, most diverse and most widely distributed group of photosynthetic prokaryotes. Most are grown photoautotrophically, since they use light-harvesting pigments, like chlorophyll *a* and phycobiliproteins, to perform oxygenic photosynthesis; however, some can also be grown photoheterotrophically or chemoheterotrophically, by metabolizing organic molecules like glucose for energy in dark conditions, but growth is at a much slower rate (Stanier and Cohen-Bazire, 1977).

Cyanobacteria are major global players in nitrogen and carbon cycles. All cyanobacteria, heterocystous (nitrogen-fixing) or non-heterocystous (non-nitrogen fixing), synthesize glycogen, which is their principle non-nitrogenous organic reserve material, and cyanophycin, their nitrogenous organic reserve material

(Geisy, 1964). Like several groups of prokaryotes, they can also fix CO<sub>2</sub> using an organelle called the carboxysome, which is filled with closely packed molecules of ribulose 1,5 biphosphate carboxylase oxygenase (RuBisCO), the enzyme responsible for photosynthetic CO<sub>2</sub> fixation (Shively *et al.*, 1973).

Cyanobacteria can be found all over the world, inhabiting aquatic and terrestrial environments. The majority of cyanobacteria live in freshwater, however, they can also be found in saltwater, hot springs, and even on soil and rocks (Paerl *et al.*, 2002). Therefore, they are required to be able to adapt to a vast majority of different environmental changes. Thus, cyanobacteria are considered good model organisms for studying adaptability to environmental changes. In fact, one of the two most extensively studies cyanobacterial species, in that respect, is the unicellular freshwater cyanobacterium *Synechocystis* sp. PCC 6803.

#### 1.4.1 *Synechocystis* sp. strain PCC 6803

In the Chroococcacean subgroup of cyanobacteria, which are the simplest in structural respects, the *Synechocystis* is substantially studied. It is a fresh water cyanobacterium that is covered with pilus-like appendages, and reproduces by binary fission (Bhaya *et al.*, 2000). This unicellular, naturally transformable genus has been widely used in genetic and biochemical studies aimed at understanding photosynthesis and other metabolic processes (Pakrasi, 1995).

*Synechocystis* sp. PCC 6803 was the first photoautotrophic organism to have its genome sequenced, at 3,573,470 bp (Kaneko *et al.*, 1996b; Kaneko and Tabata, 1997), comprising 3168 ORFs, approximately 60% of which are

annotated (Huang *et al.*, 2002; Ikeuchi and Tabata, 2006). To date, only three DEAD-box proteins have been studied in cyanobacteria; the two divergent RNA helicases CrhB and CrhC, in the *Anabaena* sp. PCC 7120 (Chamot *et al.*, 1999), and the redox regulated CrhR in *Synechocystis* sp. PCC 6803 (Kujat and Owtrim, 2000). CrhB appears to be expressed under a broad range of conditions, like CrhR, and belongs to the DEAD-box family. However, CrhC and CrhR have been shown to encode novel RNA helicases, whose expression appears to be elevated under cold shock conditions, with a basal level of expression at higher temperature observed only for CrhR.

#### 1.4.2 DEAD-box cyanobacterial RNA helicase redox, CrhR

The Cyanobacterial RNA Helicase Redox, or CrhR, is a stress-induced helicase, whose expression is linked to light-driven changes in the redox states of the electron transport chain (Kujat and Owtrim, 2000), as well as cold- and salt-dependent gene expression (Vinnemeier and Hagemann, 1999). CrhR transcript accumulation is observed in response to the presence of light. Protein levels, however, appear to be present regardless of the redox state (Kujat and Owtrim, 2000). CrhR transcript levels also accumulate after salt-shock, in which a 5 fold increase was observed after only 3 hrs of induction, and after cold-shock, increasing approximately 25-fold after only 1 hr (Vinnemeier and Hagemann, 1999). CrhR protein levels also show a significant increase with a downshift in temperature (Rosana *et al.*, 2012).

Taking into consideration cyanobacterial habitats, temperature fluctuation is potentially the stress that is most frequently encountered by cyanobacteria. It is intriguing also to note that the expression of other DEAD-box proteins identified in cyanobacteria, CrhB and CrhC, are also cold-regulated (Chamot *et al.*, 1999). This suggests that the RNA secondary structure rearrangement activity of these helicases is required for normal cellular function in response to a downshift from the organisms' optimal growth temperature (Owtrim, 2006). At low temperature, RNA secondary structure will be thermodynamically stabilized and thus its rearrangement may require an active (ATP-dependent) mechanism.

The *Synechocystis* CrhR is a unique DEAD-box RNA helicase capable of bidirectionally unwinding dsRNA, annealing ssRNA in an ATP-dependent manner and RNA strand exchange (Chamot *et al.*, 2005). These functions are also shared with the p68 and p72 DEAD-box helicases (Rössler *et al.*, 2001). CrhR has not yet been linked to any specific cellular functions within cyanobacteria; however, its regulation implies roles in RNA metabolism that may be connected to *Synechocystis* survival at low temperature, processes associated with the maintenance of optimal photosynthetic capacity, either in terms of light harvesting or carbon acquisition and fixation (Owtrim, 2006; Rosana *et al.*, 2012).

Like many DEAD-box RNA helicases, CrhR contains a unique 157 aa C-terminal extension that is not conserved in other prokaryotic and eukaryotic organisms, and not very highly conserved in other cyanobacterial species. It has not yet been elucidated whether the CrhR C-terminal extension contains accessory domains that may be required for its biochemical activities. However, other studies

on DEAD-box proteins with N- or C-terminal extensions, have proposed possible roles for these extensions in helicase activity, by providing extra RNA or protein binding sites.

Protein interaction studies have been initiated in order to deduce functional aspects of CrhR. From these observations, it became apparent that CrhR may be able to interact with itself in the yeast two-hybrid screen. If CrhR is capable of dimerization, this would make it the fourth DEAD-box RNA helicase capable of self-interaction, and like those proteins, it may actually indicate an even more critical and broader range of functions for CrhR in *Synechocystis*.

### 1.5 Objectives

The objective of this thesis was to characterize CrhR structurally and functionally. The main question posed was whether CrhR was capable of self-interaction and the formation of functional dimers. There are three primary objectives of this thesis that included determining and confirming CrhR self-interaction, the identification of the dimerization domain either within the N- or C-terminal, and determining the conditions under which CrhR dimerizes.

These questions were investigated using a range of techniques primarily yeast two-hybrid analysis, tagged-CrhR protein exchange or swap analysis and FPLC or gel filtration analysis. CrhR self-interaction was observed in the yeast two-hybrid analysis using  $\beta$ -Gal screening and leucine selection. This was also confirmed in the protein exchange analysis, in which three fusion protein

constructs tagged to CrhR were affinity purified and used to test for CrhR self-interaction using SDS-PAGE and western analysis.

CrhR contains a unique extended C-terminal domain, that is not very highly conserved in other DEAD-box RNA helicases. Mutant constructs lacking the C-terminal and or extension were designed to investigate whether the C-terminal extension is required for dimerization using both the yeast two-hybrid system, as well as in the fusion tagged-CrhR constructs. The analysis suggests that the C-terminal domain is not crucial for CrhR dimerization. Therefore constructs were designed to investigate the requirement of the N-terminal domain for dimerization. Using both the yeast two-hybrid system and tagged-CrhR protein exchange or swap analysis, showed that the primary dimerization domain is located within the N-terminal of CrhR. However, the C-terminal extension may still contribute to dimerization for stability possibly through an unidentified RNA binding domain in this region.

CrhR self-interaction was also confirmed using FPLC or gel filtration analysis, resulting in the detection of not only dimer formation, but also CrhR localization in multi-subunit complexes *in vivo*. Dimer and protein complex association was tested in the presence and absence of RNA, and shown to be RNA-independent. Association with a multi-subunit complex was also tested in warm and cold-stressed native *Synechocystis* proteins, which was found to be temperature-regulated when cold-shocked. These observations lead to the conclusion that CrhR was located in a higher protein complex *in vivo*. Mass spectrometry ID of proteins which co-elute with CrhR during FPLC analysis

indicates the prominent presence of ribosomal proteins, predominantly 50S subunit polypeptides.

The results presented in this thesis confirm CrhR dimerization, and suggest physiological functions of CrhR may be associated with the assembly or activity of ribosomes in multi-subunit complexes. This alludes to CrhR association with RNP complexes in *Synechocystis* upon acclimatization of cells to low temperatures.

## II. MATERIALS AND METHODS

### 2.1 Strains, Plasmids and Growth Conditions

#### 2.1.1 Yeast strains and plasmids

All transformations were done using the EGY48 parental yeast strain (*his3 trp1 ura3* LEU2::pLexAop6-LEU2), which contains the reporter plasmid pSH18-34 (Ura<sup>+</sup>) (Zervos *et al.*, 1993). The list of yeast two-hybrid constructs made in this study can be found in Table 2.1. Inserts were cloned into either pEG202 (His<sup>+</sup>) bait containing the activation domain (AD) and/or pJG4-5 (Trp<sup>+</sup>) prey containing the DNA binding domain (DBD).

#### 2.1.2 Growth and maintenance of yeast strains

The EGY48 parental strain (from Dr. Claudio DeVirgilio, University of Fribourg, Switzerland; Zervos *et al.*, 1993) was maintained in YTH dropout media (1.7 g/L YNB w/o amino acids, 1.4 g/L [his, trp, ura, leu] dropout mix and 5 g/L Ammonium sulfate) lacking uracil with 2% [w/v] glucose as liquid cultures, or on Yeast amino acid dropout medium plates lacking uracil with 2% [w/v] bacto agar as solid cultures. Yeast two-hybrid constructs were grown at 28°C O/N as liquid cultures in YTH dropout medium (1.7 g/L YNB w/o amino acids, 1.4 g/L [his, trp, ura, leu] dropout mix and 5 g/L ammonium sulfate) lacking histidine, tryptophan and uracil with 2% [w/v] glucose, and were grown at 28°C for 3-5 days as solid cultures on the same media containing 2% [w/v] agar. Transformants

Table 2.1: List of Yeast Two-Hybrid Strains

Strain Name AD/DBD	Plasmid		Description
	pEG202	pJG4-5	
<b>EGR/JGR</b>	CrhR ORF	CrhR ORF	Interaction between ORF of CrhR and itself
<b>EGR/JG5'</b>	CrhR ORF	5' CrhR-1 5' CrhR-2 5' CrhR-6	Interaction between ORF of CrhR and the N-terminal of CrhR
<b>EGR/JG3'</b>	CrhR ORF	3' CrhR-2 3' CrhR-12	Interaction between ORF of CrhR and the full C-terminal of CrhR
<b>EG5'/JG5'</b>	5' CrhR-13	5' CrhR-14	Interaction between the N-terminal of CrhR and itself
<b>EG5'/JG3'</b>	5' CrhR-13	3' CrhR-13	Interaction between the N-terminal and the C-terminal of CrhR
<b>EGR/JG-END</b>	CrhR END CrhR-3	ORF	Interaction between ORF of CrhR and the extended C-terminal only of CrhR
<b>EG-END/JGR</b>	END CrhR ORF	CrhR	Interaction between ORF of CrhR and the extended C-terminal only of CrhR
<b>EG-END/JG-END</b>	END END CrhR-3	CrhR	Interaction between the extended C-terminal only of CrhR and itself
<b>EGR/JGC</b>	CrhR CrhC ORF	ORF	Interaction between ORF of CrhR and ORF of CrhC
<b>EGC/JGR</b>	CrhC CrhR ORF	ORF	Interaction between ORF of CrhR and ORF of CrhC
<b>EGR/JGP</b>	CrhR ORF	7-2B5 Rps1a 7-2B6 Rps1a 9-1 Rps1a	Interaction between ORF of CrhR and ORF of Rps1a
<b>EGR/JGQ</b>	CrhR ORF	10-1 RecQ 10-3 RecQ	Interaction between ORF of CrhR and ORF of RecQ
<b>EGR/JGD</b>	CrhR ORF	DED1-2	Interaction between ORF of CrhR and ORF of DED1
<b>EGC/JGD</b>	CrhC ORF	DED1-2	Interaction between ORF of CrhC and ORF of DED1
<b>EG-END/JGD</b>	END CrhR	DED1-2	Interaction between the extended C-terminal of CrhR and ORF of DED1
<b>EGD/JGD</b>	DED1-3	DED1-2	Interaction between ORF of DED1 and itself
<b>EGR/pJG4-5</b>	CrhR ORF	pJG4-5	Interaction between ORF of CrhR and an empty vector (- ve control)
<b>pEG202/JGR</b>	pEG202	CrhR ORF	Interaction between ORF of CrhR and an empty vector (- ve control)
<b>EG-END/pJG4-5</b>	END CrhR	pJG4-5	Interaction between the extended C-terminal of CrhR and an empty vector (- ve control)
<b>pEG202/JG-END</b>	pEG202	END CrhR-3	Interaction between the extended C-terminal of CrhR and an empty vector (- ve control)
<b>YKR/YBR</b>	YKR	YBR	Yeast positive control (from Dr. Claudio DeVirgilio, University of Fribourg, Switzerland )

were selected on the dropout medium lacking leucine, and screened on yeast  $\beta$ -Galactosidase ( $\beta$ -Gal) induction media, consisting of Z media containing 2% [w/v] galactose and 1% [w/v] raffinose to test for protein interactions (more in section 2.2.4 and 2.2.5). Strains were maintained as frozen stocks at  $-86^{\circ}\text{C}$  in 20% glycerol.

### 2.1.3 *E. coli* strains and plasmids

*E. coli* overexpression constructs (MBP- and GST-fusion constructs) were made using the Gateway® Recombination Cloning Technology (Invitrogen) (described in 2.3.1). Fusion protein constructs cloned into pDEST containing the MBP- or GST-tags were transformed into *E. coli* BL21 ( $\text{Amp}^{\text{R}}$ ); fusion protein constructs cloned into pRSET-A containing the 6xHis-tag were cloned into *E. coli* BL21 pLysS ( $\text{Amp}^{\text{R}}$  and  $\text{Cm}^{\text{R}}$ ). Fusion constructs were also transformed into *E. coli* DH5 $\alpha$  host cells for propagation.

### 2.1.4 Growth and maintenance of *E. coli* strains

Solid *E. coli* cultures were grown overnight (O/N) at  $37^{\circ}\text{C}$  on LB $\text{Amp}_{100}$  (Luria broth media containing 5 g/L bacto tryptone, 5 g/L yeast extract, 5 g/L NaCl, buffered with 1 mL of 1N NaOH, 1.2% [w/v] bacto agar and supplemented with ampicillin at 100  $\mu\text{g}/\text{ml}$ ) for strains resistant to ampicillin and on LB $\text{Amp}_{100}\text{Cm}_{30}$  (supplemented with both ampicillin, as described above, and Chloramphenicol at 30  $\mu\text{g}/\text{ml}$ ) for strains resistant to both Amp and Cm. Liquid

LBamp<sub>100</sub> and LBamp<sub>100</sub>Cm<sub>30</sub> were inoculated with one colony O/N at 37°C.

Strains were maintained as frozen stocks at -86°C in 20% glycerol.

### 2.1.5 Growth conditions of *Synechocystis* sp. strain PCC 6803

*Synechocystis* sp. PCC 6803 cells were maintained on BG-11 solid media (1.5 g/L NaNO<sub>3</sub>, 0.04 g/L K<sub>2</sub>HPO<sub>4</sub>, 0.075 g/L MgSO<sub>4</sub>·7H<sub>2</sub>O, 0.036 g/L CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.006 g/L citric acid H<sub>2</sub>O, 0.006 g/L ferric ammonium citrate, 0.001 g/L DiNaEDTA, 0.002 g/L NaCO<sub>3</sub> and 1 mL A5 microelements (2.86 g/L H<sub>3</sub>BO<sub>3</sub>, 1.81 g/L MnCl<sub>2</sub>·4H<sub>2</sub>O, 0.222 g/L ZnSO<sub>4</sub>·7H<sub>2</sub>O, 0.079 g/L CuSO<sub>4</sub>·5H<sub>2</sub>O and 0.040 g/L CoCl<sub>2</sub>·6H<sub>2</sub>O)) containing 1% [w/v] bacto agar and grown at 30°C under constant illumination (Castenholz, 1988). Liquid cultures were inoculated into 50 mL BG-11 media and grown for 3 days at 30°C under constant illumination and shaking at 200 rpm. Liquid cultures were then transferred to 300 mL BG-11 media and grown for another 3-5 days at 30°C in the light, shaking at 200 rpm and aerated by bubbling with humidified air. For higher protein expression, cultures were then transferred to 20°C cold shock O/N in the light, shaking at 200 rpm and bubbling with air.

## 2.2 Yeast Two-Hybrid Analysis

### 2.2.1 Cloning into pEG202 and pJG4-5

Two plasmids were used to carry out the Yeast two-hybrid interaction assays: pEG202 bait which is 10166 bp and contains the HIS3 marker and the AD; as well as pJG4-5 prey which is 6449 bp and contains the TRP1 marker and the

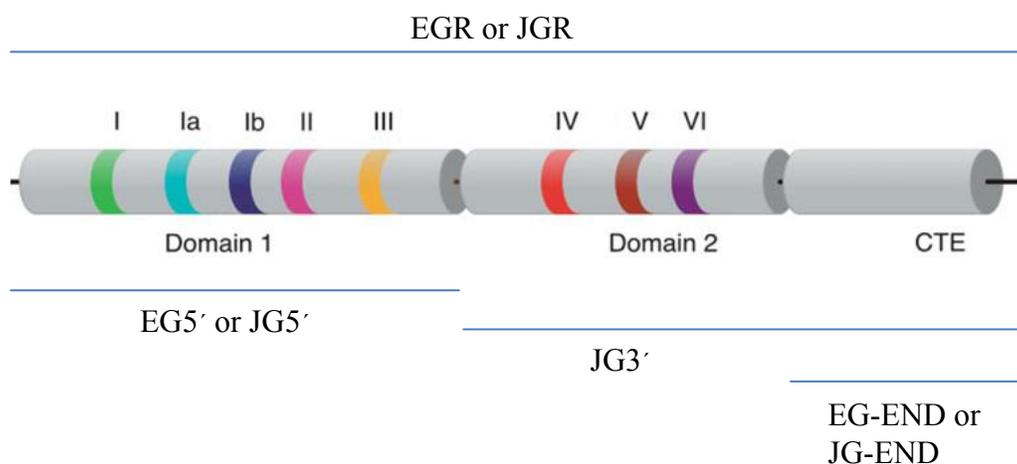


Figure 2.1: Structure of *Synechocystis* sp. PCC 6803 DEAD-box RNA helicase redox, CrhR. Original image from Jarmoskaite and Russell (2010). Domain 1 and 2 are the RecQ-like domains which are conserved across all DEAD-box RNA helicases. CTE represents the extended C-terminal domain unique to CrhR. EGR and JGR refers to the full length CrhR ORF insert cloned into pEG202 and pJG405; EG5' and JG5' contain only Domain 1 of the N-terminal cloned into pEG202 and pJG4-5, respectively. JG3' contains the entire C-terminal, including Domain 2 and the C-terminal extension cloned into pJG4-5, while EG-END and JG-END contain only the extended C-terminal domain cloned into pEG202 and pJG4-5, respectively.

DBD. The inserts that were cloned into pEG202 and pJG4-5 are shown in Figure 2.1. This is a diagram representation of the CrhR ORF and the N- and C-terminal deletions made within the CrhR ORF. The inserts were cloned into the *EcoRI* and *XhoI* sites using primers that had both restriction sites incorporated into them:

5' end of each FWD primer with *EcoRI* site:      CAT GAA TTC

3' end of each REV primer with *XhoI* site:      G GTC GTC TCG AG

Specific primers for each insert were then generated using the sequence above for the forward primer followed by ~15-20 nt of the *crhR* ORF starting at the ATG start codon, to generate an in-frame fusion peptide. For the reverse primer, sequence included 6 nt downstream of the TAA stop codon and ~12-15 nt of the ORF, which was not required to be in-frame. Primers were also made for several other cyanobacterial proteins, in order to test for interactions. Inserts were PCR amplified as described in section 2.3.4. A list of primers used for the Yeast two-hybrid can be found in Table 2.2. The plasmid constructs used in the initial determination of CrhR-CrhR interaction were generated by George Owtrim.

### 2.2.2 Chemical transformation of EGY48 using LiOAc

The EGY48 parental strain was inoculated into 5 mL YTH dropout media (section 2.1.2) lacking uracil with 2% [w/v] glucose as a liquid culture O/N. Cultures were diluted to an OD<sub>600</sub> of 0.15 and grown for ~4-5 hours at 28°C to an OD<sub>600</sub> of 0.6. Cultures were harvested at 3000 rpm using an Eppendorf 5415D centrifuge (Eppendorf), and resuspended in 50 µL LiOAc/TE (1 vol 10X TE (100 mM Tris-HCl and 10 mM EDTA), 1 vol 10 X LiOAc (1 M LiOAc, pH 7.5) and 8

Table 2.2: Primers used for Yeast Two-Hybrid Constructs

<b>Primer</b>	<b>Sequence (5'-3')</b>	<b>Origin of Sequence</b>
<b>GWO-1F</b>	CAT <u>GAA</u> TTCATGACTAATACTTTGACTAGTACC	Forward primer for CrhR ORF insert (1485 bp)
<b>GWO-1R</b>	GGTCGTCTC <u>GAG</u> GGAATCAATGTTTACTGTTGG	Reverse primer for CrhR ORF insert (1485 bp)
<b>GWO-M-R</b>	GGTCGTCTC <u>GAG</u> GCGGATTCGGGATCTTCC	Reverse primer for 5' CrhR construct used with GWO-1F; 5' of CrhR to 3' of Motif III (743 bp)
<b>GWO-M-F</b>	CAT <u>GAA</u> TTCATGGAAGATCCCGAATCCGC	Forward primer for 3' CrhR construct used with GWO-1R; 3' of Motif III to 3' of CrhR (776 bp)
<b>GWO-3'-F</b>	CAT <u>GAA</u> TTCGGGGCTGGTAAGACTGGTAAGG	Forward primer for END CrhR construct used with GWO-1R; 3' of Motif VI to 3' of CrhR (478 bp)
<b>DED1-F</b>	CAT <u>GAA</u> TTCATGGCTGAACTGAGCGAACAAGTGC	Forward primer for DED1 ORF insert (604 bp)
<b>DED1-R</b>	GGTCGTCTC <u>GAG</u> CTGAAATCACCACCAAGAAGAG	Reverse primer for DED1 ORF insert (604 bp)
<b>GWO-9F</b>	CAT <u>GAA</u> TTCATGGTCAGTCAAACCTTCTACAGC	Forward primer for Rps1a ORF insert (992 bp)
<b>GWO-9R</b>	GGTCGTCTC <u>GAG</u> TTGTTTATTATTCATCCACAGC	Reverse primer for Rps1a ORF insert (992 bp)
<b>GWO-10F</b>	CAT <u>GAA</u> TTCATGGCTGATCGCCAATCGTTGG	Forward primer for RecQ ORF insert (1436 bp)
<b>GWO-10R</b>	GGTCGTCTC <u>GAG</u> CCTGAAAATTAGCAGCGTCC	Reverse primer for RecQ ORF insert (1436 bp)

\* underlined bp indicate the restriction sites found within each primer: GAA TTC (EcoRI) and CTC GAG (XhoI)

vol H<sub>2</sub>O) for each single transformation. EGY48 was transformed with ~2 µg of plasmid DNA, with the aid of ~20-50 µg of boiled ss-carrier DNA (Sigma) in LiOAc/TE/PEG (1 vol 10X TE, 1 vol 10X LiOAc and 8 vol 50% PEG 4000). Transformants were incubated for 30 min at 28°C and heat shocked for 20 min at 43°C, then selected on (-his -trp -ura)YTH medium at 28°C for 3-5 days.

### 2.2.3 Yeast two-hybrid $\beta$ -Galactosidase screening assay

Determination of protein interactions in the yeast two-hybrid system were performed using the  $\beta$ -Gal screening assay. Single colony transformants were inoculated into 100 µL mQH<sub>2</sub>O, and quadruplicate 10 µL aliquots of each transformant were deposited on S-Gal induction media (YTH dropout media containing 2% [w/v] galactose, 1% [w/v] raffinose, 0.5M KH<sub>2</sub>PO<sub>4</sub> and 40 mg/mL X-gal). Plates were incubated for 3-5 days at 28°C, to allow color development on the YKR/YBR positive control (from Dr. Claudio DeVirgilio, University of Fribourg, Switzerland; Zervos *et al.*, 1993).

Protein interactions were also observed using the  $\beta$ -Gal filter assay (Herskowitz lab protocol). Transformants were inoculated into 100 µL mQH<sub>2</sub>O, and quadruplicate 10 µL aliquots of each transformant were deposited on YTH dropout media (section 2.1.2) and incubated for 3-5 days at 28°C. Yeast colonies were replica plated onto a Whatmann 50 filter placed on YTH dropout media plates, using sterile velvets, and incubated for 3-5 days at 28°C. The Whatmann filter paper with yeast colonies was placed into liquid nitrogen for 5 to 10 sec, thawed for 30 to 60 sec and placed onto a Whatmann #3 filter in a petri dish

containing 3 mL Z buffer (8.5 g/L anhydrous  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ , 5.5 g/L  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ , 0.75 g/L KCl and 0.246 g/L  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ ), and incubated at 37 °C until the color substrate was visible on the positive control.

#### 2.2.4 Yeast two-hybrid leucine selection assay

Protein interaction was also tested by determining growth in the absence of leucine on YTH dropout medium. EGY48 parental strain (*his3 trp1 ura3* LEU2::pLexAop6-LEU2) has the LexA promoter. If proteins attached to the AD and the DBD interact, LEU2 is expressed which complements the leu auxotrophy.

Similar to the  $\beta$ -Gal assay, single colony transformants were inoculated into 100  $\mu\text{L}$  mQH<sub>2</sub>O, and quadruplicate 10  $\mu\text{L}$  aliquots of each transformant were deposited on YTH dropout media lacking leucine. Plates were incubated for 3-5 days at 28°C, to allow for colony growth, which indicates protein interaction.

#### 2.2.5 Overexpression protein analysis of EGR-3' and JGR-3'

The END-construct was found to give anomalous results for self-interaction; therefore constructs containing either EG-END or JG-END (Figure 2.1) were tested for protein overexpression. Yeast protein extraction was performed using the alkaline lysis method. Cultures were grown O/N in 5 mL YPD media (1% [w/v] yeast extract, 2% [w/v] bacto peptone and 2% [w/v] glucose/dextrose) at 28°C. Cultures were harvested at 14,000 rpm, resuspended in 150  $\mu\text{L}$  of yeast protein extract buffer (1M NaOH and 1M  $\beta$ -ME) and incubated on ice for 10 min. Cultures were harvested at 14,000 rpm and resuspended in 150

$\mu$ L of 50% TCA and incubated on ice 10 min. Cultures were harvested as above, washed 3X with ice cold acetone and precipitated proteins were collected by centrifugation at 14,000 rpm for 2 min. Pellets were air dried for 10 min at RT and resuspended in 30  $\mu$ L Laemmli (0.1M DDT/CO<sub>3</sub>) buffer containing one-third volume SDS loading dye (125mM Tris, pH 6.8, 4% [w/v] SDS, 20% [v/v] glycerol, 10% [v/v]  $\beta$ -ME and 0.02% bromophenol blue). Proteins were further denatured by boiling for 5 min, and proteins were size fractionated by SDS-PAGE (section 2.6.1).

## 2.3 Tagged-CrhR Protein Exchange or Swap Analysis

### 2.3.1 Gateway® fusion protein constructs: GST-R and MBP-R

Fusion protein constructs overexpressed in *E. coli* BL21 were produced using the Gateway® Cloning System (Invitrogen) which involves site specific recombination. Primers were designed to clone CrhR into the Gateway® donor plasmid pDONR221, to generate an entry clone. The primers contained *attB* bacterial attachment sites (in bold below) consisting of four homologous core sequences of about 25 nt followed either by 24 nt of CrhR ORF containing the ATG start codon in the forward primer, or preceded by 24 nt of CrhR ORF containing the TAA stop codon in the reverse primer:

5' end of FWD primer with *attB*1 site:

GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGACTAATA  
CTTTGACTAGTACC

3' end of REV primer with *attB2* site:

GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACTGTTGGCG  
ATCACTATAGGC

The *attB*-flanked CrhR product (~150 ng) was combined with the Amp<sup>R</sup> pDONR221 donor vector containing the *attB*-flanked *ccdB* gene and Cm<sup>R</sup> marker (~150 ng/μL), in TE, pH 8.0 containing the BP Clonase<sup>TM</sup> II enzyme (Invitrogen) for 1 hour reaction at RT to allow for recombination. The reaction was terminated at 37°C for 10 min by the addition of Proteinase K.

To generate the MBP- and the GST-tagged CrhR constructs, the newly synthesized entry vector was shuttled into the Gateway® destination vector pDEST containing either the MBP-tag or the GST-tag (Invitrogen). The C-terminal of the MBP- and GST-tags in pDEST is attached to an *attR*-flanked *ccdB* gene and Cm<sup>R</sup> marker. Recombination between the entry vector and the destination vector generate an MBP-R and GST-R fusion, producing plasmid pDEST-MBP and pDEST-GST, which were transformed in *E.coli* BL21 (section 2.3.4).

### 2.3.2 Construction of the 6xHis-tagged CrhR fusion protein: His-R

A His-R fusion protein construct was also generated, using pRSET-A containing the 6xHis-tag (Invitrogen) through restriction enzyme (RE) digestion and ligation. Primers for CrhR were designed to incorporate a *Bam*HI and *Eco*RI site at each end of the CrhR ORF.

5' end of FWD primer with *Bam*HI site:

TTTAACGGATCCATGACTAATACTTTG

3' end of REV primer with *EcoRI* site:

GCGGGATCC**GAATTC**GGTG

The bold letters outline the bases that have been mutated to give the two restriction sites. The CrhR insert along with pRSET-A were digested using *Bam*HI and *Eco*RI FastDigest enzymes (Fermentas) for 30 min at 37°C. Vector and insert were ligated using T4 ligase and 1X ligase buffer O/N at 16°C. The ligated pRSET-A His-R was transformed into *E. coli* BL21 (section 2.3.4).

### 2.3.3 Construction of the mutant fusion proteins: His-C2R and His-N2R

Using pRSET-A, a C-terminal and an N-terminal CrhR truncation were produced. The 492 aa CrhR protein contains 334 aa of the RNA helicase core plus a 158 aa C-terminal extension unique to CrhR (Figure 2.1). The C-terminal truncation was constructed by removing 104 amino acids from His-R through RE digestion and religation. The His-R insert in pRSET-A was digested with *Pst*I, which is 320 bp upstream of the CrhR ORF TAA translation stop codon, and *Eco*RI, which is ~ 1260 bp downstream of the translation stop codon. The truncated plasmid was blunt ended using 0.1 mM dNTPs, which is a mixture of all 4 dNTPs, and 2 U of Klenow DNA polymerase at RT for 30 min, which cut back and filled in the overhangs generated by the restriction digest. The reaction products were purified using the QIAquick PCR Purification kit (Qiagen), self-ligated and transformed into *E. coli* BL21 (section 2.3.4) to generate His-C2R, which contains the RNA helicase core lacking 104 aa's of the C-terminal domain

or extension. His-C2R is 388 aa long, and encodes a 48 kDa polypeptide possessing 53 aa downstream of the HRIGR-box, which is characteristic of DEAD-box RNA helicases. This mutant was confirmed through sequencing (section 2.3.5) and western blot analysis (section 2.6.4).

The N-terminal CrhR truncation was generated using an alternate methodology than that used for His-C2R. Primers were designed to incorporate a *Bam*HI site 327 bp downstream of the ATG translation start codon, and an *Eco*RI site 92 bp downstream of the TAA translation stop codon of the CrhR ORF:

5' end of FWD primer with *Bam*HI site:

GGTGGT**G GATC**CATCGAGCGGC

3' end of REV primer with *Eco*RI site:

CCGAGC**G AATTC**CCATCAAAGTAA

The bold letters indicate the mutated bases to generate the RE sites, with the forward primer required to be in-frame. The digested CrhR insert, lacking part of the N-terminal including motif I and Ia, but containing the DEAD-box motif, was cloned into pRSET-A to generate His-N2R, and was transformed into *E. coli* BL21 (section 2.3.4). The 422 aa His-N2R encoding a 49 kDa polypeptide was confirmed through sequencing (section 2.3.5) and western blot analysis (section 2.6.4).

#### 2.3.4 PCR amplification

PCR amplification of CrhR was performed using the primer pairs listed in sections 2.3.1, 2.3.2 and 2.3.3. In a total volume of 50  $\mu$ L, approximately 200 ng

of *Synechocystis* sp. PCC 6803 genomic DNA was combined with 0.2 mM dNTP mixture, 0.5  $\mu$ M of forward and reverse primers, 1X HF buffer (New England BioLabs, NEB) and 1U Phusion DNA polymerase (New England BioLabs, NEB). The reaction was amplified in a Techgene Thermal Cycler (Life Sciences) for a total of 30 cycles [consisted of: 98°C for 10 sec denaturation, 60-68°C for 30 sec annealing ( $T_m$  was determined empirically for each primer pair) and 72°C for 30 sec elongation] with an initial denaturation at 98°C for 30 sec and a final elongation/extension at 72°C for 10 min. the 30 cycles.

### 2.3.5 *E. coli* BL21 transformation

Transformation into competent *E. coli* BL21 was done chemically. Total ligation products, were combined with 200  $\mu$ L of chemically competent *E. coli* BL21, and total ligation reaction of His-R, His-C2R and His-N2R, were combined with 200  $\mu$ L of chemically competent *E. coli* BL21 containing pLysS, and incubated on ice for 30 min. The cells were heat shocked at 42°C for 2 min and incubated at 37°C for an hour for growth by the addition of 800  $\mu$ L of LB. *E. coli* BL21 containing MBP-R and GST-R were plated on LBamp<sub>100</sub>, and *E. coli* BL21 pLysS containing His-R, His-C2R and His-N2R were selected on LBamp<sub>100</sub>Cm<sub>30</sub>, and incubated O/N at 37°C.

### 2.3.6 DNA sequencing

DNA sequencing was performed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Plasmid DNA (~500 ng) was combined

with 2  $\mu$ L BigDye-premix (Thermo sequenase II DNA polymerase, ddNTPs, dNTPs, 200 mM Tris, pH9.0, 5 mM MgCl<sub>2</sub>), 0.75X dilution buffer (200 mM Tris, pH 9.0, 5 mM MgCl<sub>2</sub>) and 5 pmol of the appropriate sequencing primer (Table 2.2). The reaction was amplified in the Techgene Thermal Cycler (Life Sciences) for a total of 25 cycles: 95°C denaturation for 30 sec, 50°C annealing for 15 sec and 60°C elongation for 1 min. The reaction was precipitated by the addition of 2  $\mu$ L NaOAc/EDTA (150 mM NaOAc, pH 8.0 and 225 mM EDTA) and 80  $\mu$ L of ice cold 95% [v/v] ethanol for 15 min at 4°C. Precipitated samples were centrifuged for 15 min at 4°C, washed with 0.4 mL ice cold 70% [v/v] ethanol and centrifuged for 5 min at 14,000 rpm. The pellet was air dried and provided to the Molecular Biology Services Unit (MBSU) for automated sequencing using the Applied Bioscience 377 analyzer.

### 2.3.7 Tagged protein expression

For protein expression, overnight cultures were diluted 100-fold, grown to an OD<sub>600</sub> of 0.4-0.6 at 37°C and induced with 0.1 mM IPTG (MBP-R and GST-R) and 1 mM IPTG (His-R, His-C2R and His-N2R) and grown for a further 3 h at 37°C (O/N for the His-N2R construct). Cells were harvested at 37°C by centrifugation at 6000 xg in a Janetzki T5 table centrifuge (Jena Instruments Ltd., Surrey, British Columbia, Canada) for 10 min and stored at -86°C.

Cells were lysed by sonication using a Braun-Sonic 2000 (Braun). MBP-R and GST-R cells were lysed by 4 cycles of 30 sec sonicated followed by 30 sec in an ice water bath. The His-tagged CrhR constructs were sonicated for less,

because the *E. coli* BL21 strain they were transformed into also contains the pLysS plasmid; 15 sec sonication and an ice water bath for 30 sec for 4 consecutive cycles. Lysates were clarified by centrifugation for 15 min at 14,000 rpm and the supernatant used for affinity purification (section 2.3.7 and section 2.3.8).

### 2.3.8 Affinity purification

Fusion proteins were affinity purified on tag-specific columns; MBP-R on amylose resin (NEB), GST-R on glutathione agarose (Invitrogen) and His-R, His-C2R and His-N2R on Ni-NTA (Qiagen). The amylose column was equilibrated with a wash buffer containing 1 M Tris-HCl pH 7.4, 1 M KCl and 0.5 M EDTA; the glutathione column was equilibrated with 1X PBS buffer pH 7.2 (Invitrogen); and the Ni-NTA column was equilibrated with a wash buffer containing 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl and 50 mM imidazole. As shown in a schematic diagram in Figure 2.2B, clarified cell lysate of each construct (section 2.3.6) was incubated for 1 h with the respective beads at 4°C with gentle shaking. The bound beads were washed 5X with the appropriate equilibration buffer with a pulse centrifugation to pellet the beads after each wash using an Eppendorf 5415D centrifuge (Eppendorf). Proteins were eluted with the respective wash buffers containing 10 mM β-ME and 10 mM maltose for MBP-R; 5 mM reduced glutathione for GST-R; and 250 mM imidazole for His-R, His-C2R and His-N2R. Proteins were quantified using the Bradford Assay (Bio-Rad Protein Assay Dye Reagent, 200 μL) using the appropriate elution buffer as a control. BSA was used

to generate a standard curve. The purity of eluted proteins was determined by SDS-PAGE (section 2.6.1) and western blot analysis (section 2.6.4).

### 2.3.9 Tagged-CrhR Exchange Assays

The theory of tagged-CrhR exchange assays, or swap analysis, is outlined in Figure 2.2B. A fusion protein was affinity purified as described in section 2.3.7. A second differently-tagged CrhR was passed over the bound beads and incubated for another hour at 4°C. Columns were washed 5X with the appropriate wash buffer and eluted with the appropriate elution buffer to the first bound construct on the column. This was repeated for all possible tagged-CrhR protein combinations, as well as with the N- and C-terminal truncated constructs, in order to test for interaction between two differently-tagged CrhR constructs. Proteins were quantified using the Bradford Assay (as described in section 2.3.8), and analyzed using SDS-PAGE (section 2.6.1) and western blot analysis (section 2.6.4).

## 2.4 FPLC – Gel Filtration Analysis

### 2.4.1 Affinity purification of His-R (+/-) RNase A

His-R was overexpressed and affinity purified as described in sections 2.3.6 and 2.3.7. Prior to FPLC gel filtration analysis, RNase A was added to one sample of purified His-R at 10 µg/mL and incubated for 10-30 min at 37°C. Gel filtration analysis is described in section 2.4.3.

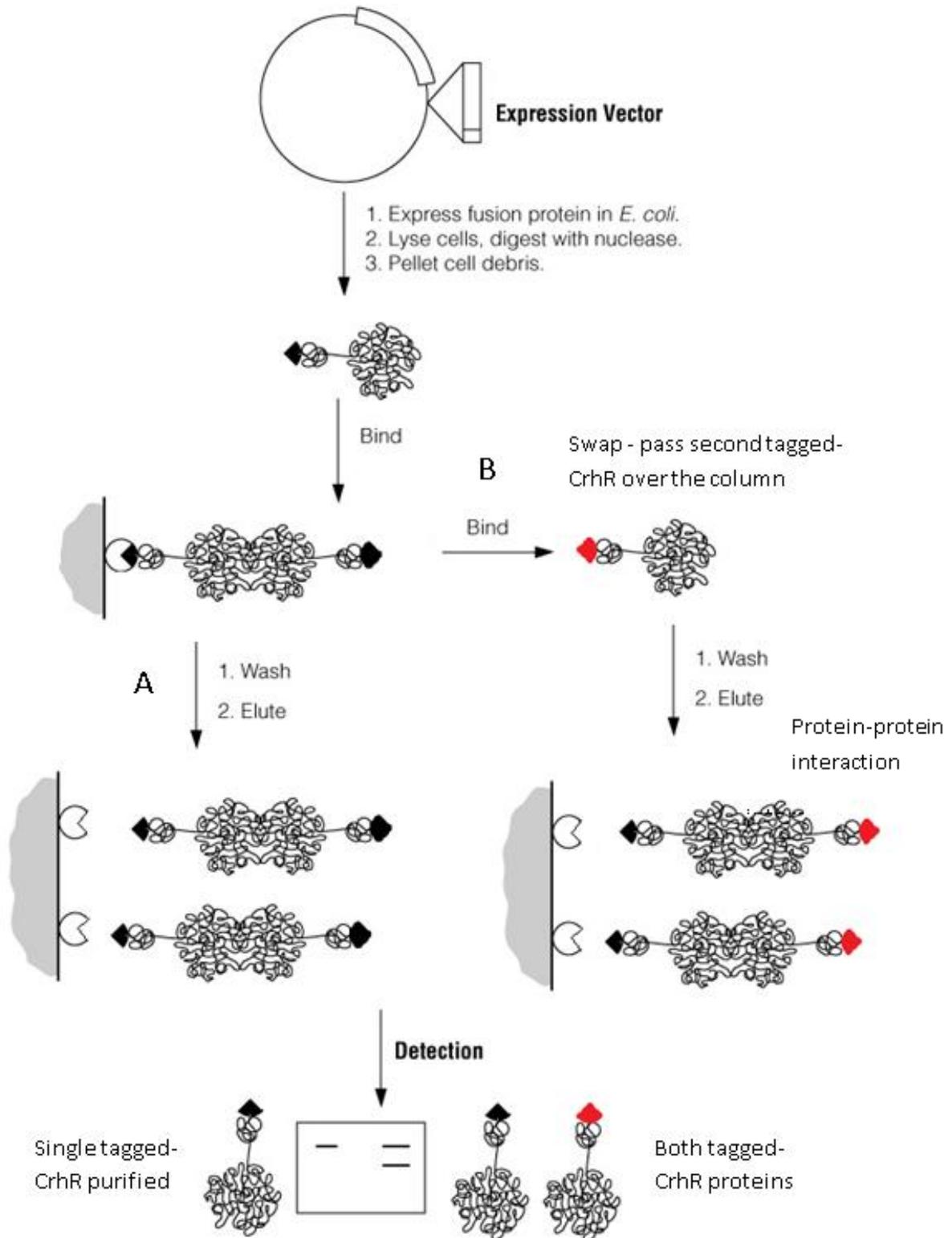


Figure 2.2: Tagged-CrhR Protein Exchange - Swap Analysis. A schematic representation of a tagged-CrhR construct affinity purified on its respective column, in this example MBP-R (black) on the amylose resin column (white). In (A), the bound construct was washed several times and eluted off the column, which results in detection of MBP-R followed by Western analysis. In (B), before eluting the bound construct, a second differently tagged-CrhR construct was passed over the column, in this example His-R (red). If there is interaction between the two-differently tagged-CrhR proteins, then SDS-PAGE and western analysis should detect both the MBP-R and His-R constructs. This is referred to as a swap analysis.

## 2.4.2 Protein isolation from *Synechocystis* sp. PCC 6803 (+/-)

### RNase A

Warm and cold shocked *Synechocystis* sp. strain PCC 6803 cells were harvested by centrifugation using a Beckman Model J2-21 centrifuge (Beckman Coulter) at 6000 rpm for 15 minutes and resuspended in 1 mL extraction buffer consisting of 20 mM Tris-HCl, pH 8.0, 100 mM KCl and 1mM MgCl<sub>2</sub>, and containing a protease inhibitor cocktail (Complete Mini-Roche). One-third volume of Dyno-Mill Lead free 0.2 - 0.3 mm glass beads (Impandex Inc.) and DTT to a final concentration of 3mM were added and cells lysed by seven cycles of 30 sec vortexing followed by 30 sec in an ice water bath. Lysates were clarified by centrifugation for 10 minutes at 4°C. Protein samples that needed to be RNA-free were incubated with 10 µg/µL RNase A at 37°C for 30 min prior to gel filtration analysis (section 2.4.3).

## 2.4.3 Gel filtration

Gel filtration analysis was performed in the Molecular Biology Services Unit (MBSU) using the ÄKTA FPLC<sup>TM</sup> (Amersham Biosciences) for size separation of affinity purified His-R and native *Synechocystis* sp. PCC 6803 proteins. Purified His-R samples were separated on a Superdex 200BKL column (Amersham Biosciences) with a fractionation range of 10,000 to 600,000 Daltons. Samples were loaded manually by injection onto columns and mixed with a running buffer consisting of 50 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>, pH 7.4, 300 mM NaCl and 10 mM imidazole. Protein elution was monitored at 280 nm and displayed on

the UNICORN software interface. Total run time of each sample was 100 min, using a total of 54 mL of running buffer at a flow rate of 0.5 mL/min and 250  $\mu$ L fractions were collected.

Crude protein cell lysate of *Synechocystis* sp. PCC 6803 was separated on a Sephacryl S300 column (Amersham Biosciences) with a fractionation range of 10,000 to 1,500,000 Daltons. Running buffer used for native samples consisted of 20mM Tris-HCl pH 8.0, 100mM KCl and 1mM MgCl<sub>2</sub>. Protein elution was also monitored at 280 nm and displayed on the software interface. Total run time of each sample was 100 min, using a total of 54 mL of running buffer with a 0.5 mL/min flow rate and 500  $\mu$ L fractions were collected.

#### 2.4.4 TCA and acetone protein precipitation

When required, protein fractions were concentrated by either TCA precipitation or acetone precipitation. For TCA precipitation, single or multiple (3-5) fractions, were mixed with one-tenth volume of TCA and one-tenth volume of deoxycholate at 10 mg/mL and incubated on ice for 30 min. Samples were centrifuged at 14,000 rpm for 15 min at RT and washed 2X with ice cold acetone and pelleted at 14,000 rpm for 5 min at RT. Protein pellets were resuspended in 30  $\mu$ L of 0.1M DTT/CO<sub>3</sub> and mixed with one-third volume SDS loading buffer (125mM Tris, pH 6.8, 4% [w/v] SDS, 20% [v/v] glycerol, 10% [v/v]  $\beta$ -ME and 0.02% bromophenol blue). Samples were boiled for 4 min then size fractionated by SDS-PAGE (section 2.6.1).

For acetone precipitation, single or combined FPLC fractions were mixed with 3 volumes of ice cold 100% acetone, and incubated at 20°C for 30 min. Samples were pelleted at 14,000 rpm for 15 min at 4°C and air dried for 10 min. Protein pellets were resuspended in 30  $\mu$ L of 0.1M DTT/CO<sub>3</sub>, and one-third volume SDS loading buffer, boiled and size fractionated by SDS-PAGE (section 2.6.1).

## 2.5 Ribosomal Preparation

### 2.5.1 Ribosome isolation from *Synechocystis* sp. PCC 6803

Native cold shocked *Synechocystis* sp. PCC 6803 cells were grown as described in section 2.4.2, and harvested by centrifugation using a Beckman Model J2-21 centrifuge (Beckman Coulter) at 6,000 rpm for 15 minutes. Cells were resuspended in 25 mL ribosome extraction buffer (20 mM Tris-HCl, pH 7.5, 300 mM NH<sub>4</sub>Cl, 10 mM MgCl<sub>2</sub>, 0.5 mM EDTA and 1 mM DTT) and lysed by the addition of one-third volume of glass beads by vortexing for 30 sec, and incubating in an ice/water bath for 30sec, for 7 min. Cells were centrifuged for 10 minutes at 4°C and immediately used for ultracentrifugation analysis.

### 2.5.2 Ultracentrifugation of ribosomal proteins

The 25 mL aliquot of total cell lysate from section 2.5.1 was centrifuged O/N at 100,000 x g in a Beckman Coulter Optima L-90K using the SW 70 Ti rotor to produce a crude ribosome pellet. The supernatant was stored at 4°C on ice, and the reddish-brown pellet was resuspended in 12.5 mL ribosome extraction

buffer and layered onto a 12.5 mL 10% sucrose cushion buffer (20 mM Tris-HCl, pH 7.5, 300 mM NH<sub>4</sub>Cl, 10 mM MgCl<sub>2</sub>, 0.5 mM EDTA, 1 mM DTT and 10% [w/v] sucrose). The mixture was centrifuged O/N at 100,000 x g using the SW 70 Ti rotor. The supernatant was stored at 4°C on ice, and the yellowish/brown pellet was suspended in 12.5 mL ribosome extraction buffer and layered onto a 12.5 mL 30% sucrose cushion, prepared in ribosome extraction buffer and centrifuged O/N at 100,000 x g in the SW 70 Ti rotor. The supernatant was stored at 4°C on ice and the greenish pellet was resuspended in 500 µL of Ribosomal extraction buffer and stored at 4°C on ice.

A 1 mL aliquot of the resuspended pellet from each centrifugation was also collected and stored at 4°C on ice. Approximately the same protein concentration was used for western analysis; 1:10 volume of original 1 mL lysate and pellet, and 1:250 volume of the 25 mL supernatants were concentrated and combined with one-third volume SDS loading buffer, and size fractionated by SDS-PAGE (section 2.6.1). Supernatants were TCA precipitated (section 2.4.4) prior to loading on SDS-PAGE, while pellets were not as they were already concentrated samples.

## 2.6 Protein Electrophoresis

### 2.6.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were denatured by boiling in a water bath for 4 min and separated by 10% SDS-PAGE. The 3.75 mL resolving gel consisted of: 937.5 µL of 40% [w/v] acrylamide:bis (37.5:1) (Bio-Rad), 468.75 µL of 3M Tris, pH 8.8,

37.5  $\mu\text{L}$  of 10% [w/v] SDS, 2.12 mL of mQH<sub>2</sub>O, 187.5  $\mu\text{L}$  of 1.5% [w/v] APS, and 1.875  $\mu\text{L}$  of TEMED. The 1.25 mL stacking gel consisted of: 125  $\mu\text{L}$  of 40% [w/v] acrylamide:bis (37.5:1) (Bio-Rad), 315  $\mu\text{L}$  of 0.5M Tris-HCL, pH 6.8, 12.5  $\mu\text{L}$  of 10% [w/v] SDS, 741.25  $\mu\text{L}$  of mQH<sub>2</sub>O, 62.5  $\mu\text{L}$  of 1.5% [w/v] APS, and 1.25  $\mu\text{L}$  of TEMED. For limited proteolysis, samples were run on 15% SDS-PAGE, differing only in the volumes of the 40% acrylamide:bis (37.5:1) and water (1.406 mL of 40% [w/v] acrylamide:bis (37.5:1) (Bio-Rad), 1.6515 mL mQH<sub>2</sub>O for the resolving gel and 187.5  $\mu\text{L}$  of 40% [w/v] acrylamide:bis (37.5:1) (Bio-Rad), 678.75  $\mu\text{L}$  mQH<sub>2</sub>O for the stacking gel).

The wells were flushed out with SDS Reservoir running buffer (25 mM Tris, 0.192 M glycine, 0.1% [w/v] SDS), and approximately 30  $\mu\text{L}$  of samples and 2.5  $\mu\text{L}$  of Fermentas Prestained Protein ladder were loaded, and electrophoresed at 200 Volts for 45 min to 1 hour in 1X SDS Reservoir running buffer.

## 2.6.2 Coomassie staining SDS-PAGE

Proteins were fixed in the PAGE gels by destain solution (30% [v/v] methanol and 10% [v/v] glacial acetic acid) for 10 min with gentle shaking. Destain was replaced with dH<sub>2</sub>O for 10 min with gentle shaking. Proteins were stained with Coomassie Brilliant Blue stain (14% [v/v] methanol, 10% [v/v] glacial acetic acid, 0.25% [w/v] Coomassie Brilliant Blue R250) , until proteins could be visualized, in most cases O/N with gentle shaking.

### 2.6.3 Blue native-polyacrylamide gel electrophoresis (BN-PAGE)

The fusion protein constructs were also separated by Blue Native-PAGE (BN-PAGE) as described by Schagger *et al.*, 1994. Linear gradient gels of 1.6 mm in size were cast using 23 mL of 5-18% Gradient gel: 5% gel (1.8 mL 49.5% [w/v] acrylamide:bis (48:1.5), 6 mL Gel buffer (150 mM Bistris/HCl, 1.5 M aminocaproic acid, pH 7.0), 100  $\mu$ L 10% [w/v] APS, 10  $\mu$ L TEMED); 18% gel (5.5 mL 49.5% [w/v] acrylamide:bis (48:1.5), 5 mL Gel buffer (150 mM Bistris/HCl, 1.5 M aminocaproic acid, pH 7.0), 3 g of glycerol, 50  $\mu$ L 10% [w/v] APS, 5  $\mu$ L TEMED). The gradients were poured together to create the gradient, overlaid with butanol and left to polymerize for 30 min. The Gradient gel was washed thoroughly to remove any traces of butanol after polymerization, and overlaid with 6 mL of 4% sample gel (500  $\mu$ L 49.5% [w/v] acrylamide:bis (48:1.5), 2 mL Gel buffer (150 mM Bistris/HCl, 1.5 M aminocaproic acid, pH 7.0), 50  $\mu$ L 10% [w/v] APS, 5  $\mu$ L TEMED), and a 10-well comb was inserted. The gel was left to polymerize for 30 min.

BN-PAGE requires the use of 2 Running buffers: Cathode buffer (50 mM Tricine, 15 mM Bistris/HCl, pH 7.0, and Coomassie blue G-250 (0.02 or 0.002%)), and the Anode buffer (50 mM Bistris/HCl, pH 7.0). Approximately 60  $\mu$ L samples and 10  $\mu$ L Fermentas Prestained Protein ladder were loaded and electrophoresed at 100 Volts O/N.

#### 2.6.4 Western blot analysis

Proteins separated on PAGE gels that were not stained after electrophoresis were immobilized to a solid matrix through electroblotting, using a semi-dry transfer method. A single SDS-PAGE gel was soaked for 10 min in 1X Tyler transfer buffer (25 mM Tris, 150 mM glycine, pH 8.3, 20% [v/v] methanol) along with four pieces of 3MM Whatmann paper and one 0.45 micron Hybond ECL nitrocellulose membrane (Amersham Pharmacia Biotech), each cut the same size as the gel. Using an Electrophoretic Transfer System ET-10 apparatus (Tyler Research Instruments), two pieces of soaked 3MM Whatmann paper were placed in the apparatus, overlaid with the PAGE gel, the nitrocellulose membrane, and the last two pieces of 3MM Whatmann paper, and transferred for 1 h at 52 mA. If more than one gel were transferred at the same time then the current was increased; two gels at 80 mA, three gels at 100 mA and four gels at 120 mA.

After transfer, the membranes were incubated at RT in 1X BLOTTO blocking buffer (1X TBS, 5% [w/v] skim milk powder, 0.02% [v/v] Sodium azide) for 15 min with gentle shaking. Membranes were incubated in fresh 10 mL of 1X BLOTTO containing anti-CrhR antibody(1:2000) (Chamot *et al.*, 2005) O/N at RT with gentle shaking. In some cases, for MBP-tagged CrhR and His-tagged CrhR, monoclonal anti-MBP antibody (Sigma) (1:3000) or monoclonal anti-polyhistidine antibody (Sigma) (1:3000) were used.

Membranes were washed with 20 mL 1X TBS (150 mM NaCl, 10 mM Tris-HCl, pH 8.0) for 15 minutes, 1X TBST (1X TBS containing 0.05% [v/v] Tween-20) for 15 min, and 1X TBS for 15 min with gentle shaking at RT.

Membranes were incubated in 20 mL 1X TBS containing anti-rabbit IgG antibody conjugated to horse-radish peroxidase (HRP) (1:20,000) (Sigma) or anti-mouse antibody conjugated to HRP (1:20,000) (Sigma) for the MBP- or His-tagged CrhR constructs. The washes were repeated with the final 1X TBS wash for 25 min. CrhR was visualized using the ECL Western Blotting Detection kit (Amersham Biosciences) and chemilluminescence was detected using autoradiography.

### 2.6.5 Far-Western analysis

Fusion-tagged proteins at 2, 4 and 6  $\mu$ g, were separated on SDS-PAGE and immobilized to Hybond matrix as described in section 2.6.4. After transfer, the membrane was incubated at RT for 10 min with gentle shaking in HBB buffer (25 mM HEPES/KOH, pH 7.5, 25 mM NaCl, 5 mM MgCl<sub>2</sub> and 1 mM DTT) containing 8 M urea to denature protein. Proteins were partially renatured by incubating the membrane in HBB buffer with decreasing concentrations of urea (7M, 6M, 5M, 4M, 3M, 2M, 1M, 0.5M and 0.25M) for 10 min each at RT. Membranes were blocked with HBB buffer containing 5% [w/v] skim milk for 30 min, and incubated O/N in hybridization buffer (20 mM HEPES/KOH, pH 7.5, 75 mM KCl, 2.5 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 1 mM DTT, 0.1% [v/v] Nonidet P-40, and 1% [w/v] skim milk) containing 1 $\mu$ g/ $\mu$ L purified His-R probe (section 2.3.8).

The membrane was washed 5 times with 1X TBST for 10 min each and incubated in 1X TBS containing monoclonal anti-polyhistidine antibody (Sigma) (1:3000) for 120 min. The membrane was washed 3X with 1X TBS, 1X TBST and 1X TBS for 15 min each and incubated in 20 mL 1X TBS containing anti-

mouse antibody conjugated to HRP (1:20,000) (Sigma). The same three washes were repeated with the final 1X TBS wash for 25 min. CrhR was visualized using the ECL Western Blotting Detection kit (Amersham Biosciences) and chemilluminescence was detected using autoradiography.

## 2.7 Mass Spectrometry Analysis

### 2.7.1 Sample preparation from *Synechocystis* sp. PCC 6803

Native cold shocked *Synechocystis* sp. PCC 6803 proteins that were size fractionated on a Sephacryl S300 column (section 2.4.3) eluted in a high molecular weight complex (~600 kDa). The first 16 fractions collected from gel filtration were TCA precipitated (section 2.4.4) and separated by SDS-PAGE (section 2.6.1) for 15 min and Coomassie stained O/N (section 2.6.2). Once visualized, protein bands were excised in 1 mm segments using feather surgical scalpel blades (Fisher Scientific). Each protein segment was subsequently cut into 1 mm x 1 mm cubes and placed in separate wells of a pre-sterilized, DNase/RNase free 96-well plates (Costar) (Figure 2.3). Gel segments were submerged in 100 mM ammonium bicarbonate (David Kramer, Department of Biochemistry, University of Alberta).

### 2.7.2 Trypsin digestion of samples

Each 96-well plate was subject to tryptic digestion using an automated PerkinElmer MassPrep Robotic protein handling station. Gel segments were de-stained twice using 50 mM ammonium bicarbonate in 50% acetonitrile. Gels were

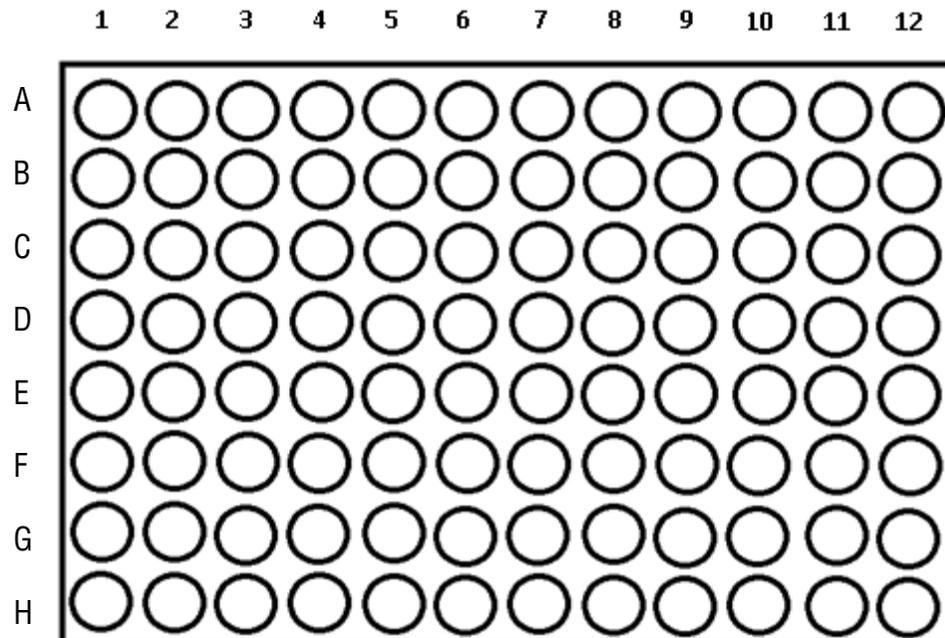


Figure 2.3: Mass spectrometry analysis of native cold shocked *Synechocystis* sp. PCC 6803 proteins in a 96 well plate. The first 16 FPLC fractions of the Sephacryl S300 column were separated by SDS-PAGE and Coomassie stained. Each of the samples were cut into mini sections and placed into two-96 well plates containing 100 mM ammonium bicarbonate, pH 8.0. Approximately 10 sections were cut per sample and a total of 2 plates were processed; A1 on the first plate and F7 on the second plate are the standards. Sample segments were deposited into each well from top to bottom, therefore lane 1 was divided into well B1, C1, D1, E1, F1, G1, H1, A2, B2 and C2. Lane 2, D2-D3; lane 3, E3-E4; lane 4, F4-G5; lane 5, H5-A7; lane 6, B7-B8; lane 7, C8-C9; lane 8, D9-D10; lane 9, E10-F11; lane 10, G11-H12; lane 11, (second plate) A1-C2; lane 12, D2-D3; lane 13, E3-D4; lane 14, E4-D5; lane 15, E5-E6; and lane 16, F6-E7.

dehydrated for 10 minutes in 100% acetonitrile and cysteine residues were reduced by adding 10 mM  $\beta$ -mercaptoethanol in 100 mM ammonium bicarbonate for 30 minutes at room temperature. Following reduction, the cysteines were alkylated (carbamidomethylated) with 55 mM iodoacetamide in 100 mM ammonium bicarbonate for 30 minutes at room temperature. Gel slices were washed twice with 100 mM ammonium bicarbonate for 5 min/wash and dehydrated twice with 100% acetonitrile for 10 minutes/dehydration to improve trypsin absorption into the gel.

Sequencing grade porcine trypsin (Promega) was diluted to 6 ng/ $\mu$ L in 50 mM ammonium bicarbonate and added to each well. The plate was incubated at 37°C for 6 hours to allow adequate trypsinization of each protein sample. Trypsin was quenched and tryptic peptides were extracted through the addition of 30  $\mu$ L 1% formic acid/2% acetonitrile to each sample for 30 minutes at room temperature. Liquid was extracted from each well and placed in a new pre-sterilized 96-well plate designed for use with nano-LC systems (Axygen). A second extraction was performed through the addition of 30  $\mu$ L 1% formic acid/50% acetonitrile for 30 minutes at room temperature. This extraction was pooled with the first and samples were analyzed by LC-MS/MS (David Kramer, Department of Biochemistry, University of Alberta).

### 2.7.3 Liquid chromatography mass spectrometry analysis (LC-MS)

Digested and extracted samples were subjected to nano-reverse phase liquid chromatography (RPLC) prior to mass spectrometry analysis. Using a

ThermoScientific Easy nle-II, 3  $\mu$ L of extract/well was subject to RPLC, using a New Objective PicoFrit™ ProteoPep II 100 mm C18 trapping column. A 100%A/0%B  $\rightarrow$  60%A/40%B (A= HPLC grade water + 0.2% formic acid; B= acetonitrile + 0.2% formic acid) gradient was run over the course of 40 minutes with two 5-minute column washes between samples. Peptides eluting off of the nLC column were directly subject to nano-electrospray ionization (nanoESI) and MS/MS analysis using a ThermoFisher LTQ Orbitrap XL mass spectrometer.

MS/MS data was saved in .RAW files. Using ThermoScientific's ProteomeDiscoverer 1.3, all RAW files generated from samples in a single SDS-PAGE gel lane were grouped and searched against a *Synechocystis* sp. PCC 6803 proteome .FASTA index (from UniprotKB) using Sequest software. Protein filters were applied: A Minimum of 1 medium-confidence peptide (95% confidence) and to include distinct protein isoforms. After generation of each protein list with the specified filters, each list was exported to an excel file and any proteins identified with only 1 unique peptide were removed from the data set (David Kramer, Department of Biochemistry, University of Alberta).

### III. Results:

#### 3.1 CrhR Self-Dimerization *in vitro* and *in vivo*

Dimerization was tested and observed in the yeast two-hybrid screen, tagged-CrhR protein exchange or swap analysis and FPLC gel filtration analysis. CrhR dimerization was investigated using overexpressed fusion constructs *in vitro* along with *in vivo* natively-expressed CrhR from *Synechocystis*.

##### 3.1.1 Yeast two-hybrid interaction in $\beta$ -Galactosidase assay

In the yeast two-hybrid protein interaction screen, CrhR was fused to both the activation domain (AD) (pEG202) and the DNA-binding domain (DBD) (pJG4-5), and tested for self-interaction. In the first instance, CrhR interaction was observed using the  $\beta$ -Gal screen and  $\beta$ -Gal filter assay. As described in the Materials and Methods, yeast containing DBD-crhR and AD-crhR in the EGY48 strain were plated on S-Gal induction plates containing X-Gal. Protein interaction is indicated by cleavage of the X-Gal in the media, and results in a blue-colored product. The  $\beta$ -Gal filter assay, also described in the Materials and Methods section, also results in a blue-colored product which is indicative of protein-protein interaction.

From both screens, it was observed that CrhR was positive in terms of self-interaction, indicated by growth in the absence of leucine and a positive  $\beta$ -Gal assay (Table 3.1). The self-interaction observed was confirmed, as evidence by

Table 3.1: Yeast two-hybrid interactions on S-Gal and (-) Leucine plates

Strain Name AD/DBD	$\beta$ -Gal Assay	Leucine Selection	AD	DBD
EGR/JGR	+++	+++		
EGR/JG5'	+++	++		
EGR/JG3'	-	-		
EG5'/JG5'	++	+		
EG5'/JG3'	-	-		
EGR/JG-END *	++	+		
EG-END/JGR *	++	+		
EG-END/JG-END	+++	-		
EGR/JGC	-	-		
EGC/JGR	-	-		
EGR/JGP	-	-		
EGR/JGQ	-	-		
EGR/JGD	-	-		
EGC/JGD	-	-		
EG-END/JGD	-	-		
EGD/JGD	-	-		
EGR/pJG4-5	-	-		
pEG202/JGR	-	-		
EG-END/pJG4-5 **	+	++		
pEG202/JG-END **	+	-		
YKR/YBR	++++	++++		

Note: EG = AD fused to a protein and JG = DBD fused to a protein

Colored barrels represent CrhR (Figure 2.1); plain barrels represent other proteins; lines

Represent empty vectors

\*EG-END and JG-END construct results inconsistent between  $\beta$ -Gal Assay and Leucine selection

\*\*EG-END and JG-END constructs give false positives as negative controls

the lack of color development when the full length CrhR was fused to the AD and paired with an empty prey vector (EGR/pJG4-5) and the reciprocal pEG202/JGR), both these negative controls showed there was non-specific interaction of CrhR and each of the pEG202 and pJG405 empty vectors. CrhR was also shown not to interact with CrhC, the cyanobacterial *Anabaena* sp. PCC 7120 RNA helicase, Rps1a, the yeast ribosomal protein of the small 40S subunit, RecQ, the *E. coli* DNA helicase, and DED1, the yeast DEAD-box RNA helicase. All these constructs were used as additional negative controls in the yeast two-hybrid system. While CrhR appears to interact on S-Gal plates and in the  $\beta$ -Gal filter assay, the blue color product resulting from the interaction was a little less intense than that observed in the Yeast positive control YKR/YBR (from Dr. Claudio DeVirgilio, University of Fribourg, Switzerland; Zervos *et al.*, 1993). For this purpose, we repeated the yeast two-hybrid assay using leucine selection to confirm this result.

### 3.1.2 Yeast two-hybrid interaction on plates lacking leucine by monitoring growth

The EGY48 yeast strain contains six LexA operators directing transcription of the LEU2 gene (Zervos *et al.*, 1993). Interaction between the AD and DBD containing CrhR would drive leucine expression and colonies would be able to grow on in the absence of leucine and indicate protein-protein interaction. Therefore, CrhR was also tested for interaction on YTH dropout media lacking leucine. Yeast two-hybrid strains grown on plates lacking leucine are also shown

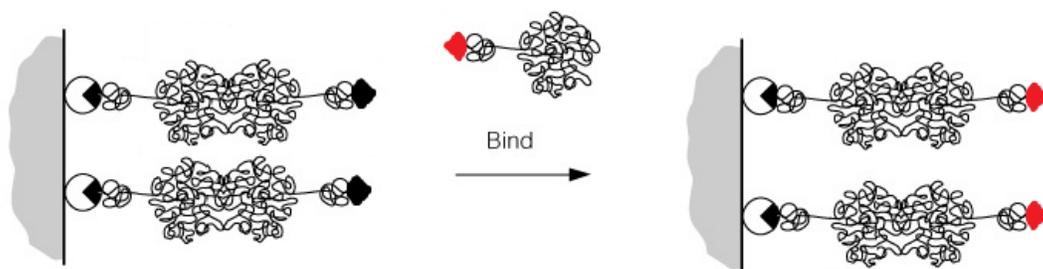
in Table 3.1. CrhR was also shown to self-interact, as there was colony growth observed on plates lacking leucine, which matched the yeast positive control YKR/YBR.

The negative controls, EGR/pJG4-5 and pEG202/JGR, did not grow in the absence leucine; and the lack of growth indicates the lack of CrhR interaction with CrhC, Rps1a, RecQ and DED1 was also confirmed, as these strains did not grow in the absence leucine. Therefore, CrhR self-interaction observed with the leucine selection confirms the results of both the  $\beta$ -Gal methods.

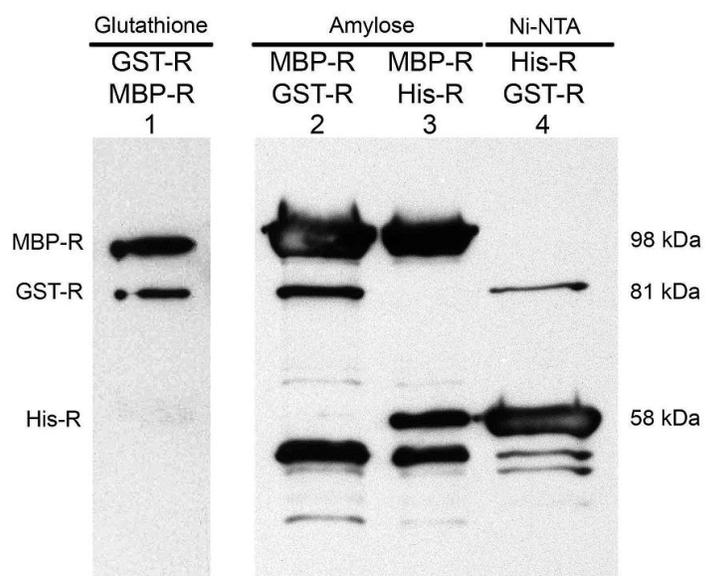
### 3.1.3 Self-interaction analysis using tagged-CrhR (swap analysis)

CrhR was tagged with three different fusion proteins, maltose binding protein, glutathione-S transferase and 6x-Histidine to generate MBP-CrhR (MBP-R, 98 kDa), GST-CrhR (GST-R, 81 kDa) and His-CrhR (His-R, 58 kDa), respectively. Fusion proteins were overexpressed in *E. coli* BL21. To further test CrhR helicase self-dimerization, what we termed swap analysis was used (Materials and Methods section 2.3.9). This can best be explained as a bait and prey system in which one fusion construct is attached to its respective affinity column as a dimer. A portion of the bound dimer will dissociate leaving monomer attached at some of the sites. The second fusion construct is passed over the column, some of which will be in monomer form which will bind to the monomers on the column. Thus, differently tagged-CrhR proteins will exchange with each other creating heterodimers. The extent of heterodimer formation will depend on the  $K_D$  of the dimer and thus it can be predicted that not all bound

A



B



C

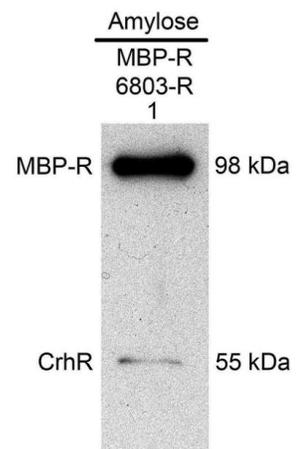


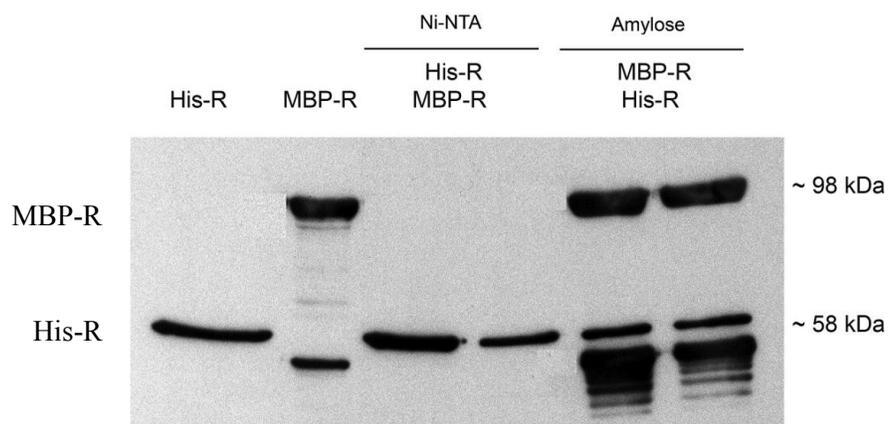
Figure 3.1: Dimerization of three differently tagged-CrhR in protein exchange analysis. The tagged-CrhR constructs were overexpressed in *E. coli* BL21, and affinity purified on the respective columns, by binding the first tagged-CrhR (for example MBP-R in black), and passing over the bound beads a second differently tagged-CrhR (His-R in red) seen in (A), which represents lane 3 (B). The purified fusion proteins were size fractionated on SDS-PAGE and analyzed on a Western, using an anti-CrhR antibody. Positive self-dimerization is seen in (B) with the reciprocal GST-R/MBP-R and MBP-R/GST-r swaps in lane 1 and 2, as well as in the MBP-R/His-R and His-R/GST-R swaps. CrhR self-dimerization was also detected in (C) when using native *Synechocystis* lysate termed 6803-R that was passed over the bound MBP-R construct. All extra bands not corresponding to the labeled protein sizes were deemed to be breakdown products.

proteins will be heterodimers. After extensive washing, eluted proteins are analyzed by western blotting. Detection of both tagged versions of CrhR provides evidence that CrhR self-dimerizes. It should be noted that the size and N-terminal location of the tags may weaken CrhR dimer formation.

In the first exchange analysis in Figure 3.1, GST-R was bound to a glutathione agarose column, and crude *E. coli* lysate containing overexpressed MBP-R was passed over the bound beads. If CrhR self-interacts, it is expected that both overexpressed proteins to be detected on the western blot analysis. Figure 3.1A in lane 1, CrhR was detected at 98 kDa and 81 kDa corresponding to MBP-R and GST-R respectively. Lane 2 shows the reciprocal exchange, with MBP-R bound to an amylose resin column, and GST-R passed over the bound beads; western results were identical to those shown in lane 1. All other possible combinations of the fusion proteins were used in the exchange or swap analysis and were also positive; MBP-R bound to amylose resin, and His-R passed over the bound beads as shown in lanes 3, shows CrhR detection at both 98 kDa and 53 kDa. And in lane 4, His-R was bound to the Ni-NTA column, and GST-R passed over the bound beads, also results in positive CrhR detection at both 81 kDa and 53 kDa. The swaps were also functional when using soluble cell lysate from *Synechocystis*. As shown in Figure 3.1C, cell soluble lysate containing the native 55 kDa CrhR protein was passed over bound MBP-R, and CrhR was detected at both 98 kDa and 55 kDa.

Reciprocals of the swaps were also positive, indicative of CrhR interaction. However, at first affinity purification of His-R on the Ni-NTA column and

A



B

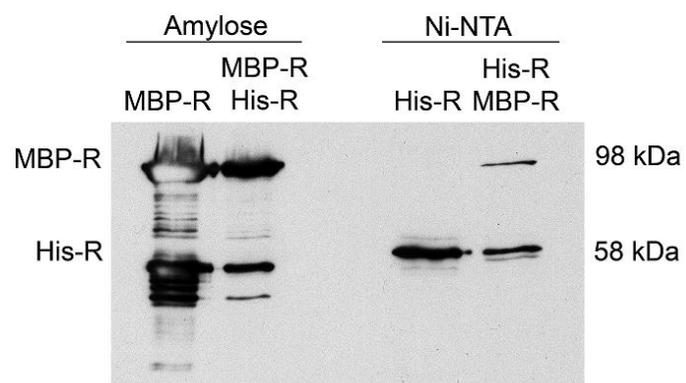


Figure 3.2: In-solution interaction of affinity purified His-R and MBP-R. His- and MBP-tagged-CrhR constructs were overexpressed in *E. coli* BL21. In (A), His-R was affinity purified on the Ni-NTA column, and MBP-R was passed over the bound beads. The western also shows purified His-R in lane 1 and purified MBP-R in lane 2. In (B), the swap analysis of both His-R and MBP-R was repeated by combining the total cell lysates of both in solution prior to binding on the Ni-NTA affinity column. CrhR detection was positive in both reciprocal swaps at 58 kDa and 98 kDa. Purified MBP-R in lane 1 and purified His-R in lane 3 also shown for confirmation. All extra bands not corresponding to the labeled protein sizes were deemed to be breakdown products.

addition of the MBP-R lysate to the bound beads was not successful. As shown in Figure 3.2A, CrhR is only detected at 53 kDa, which represents His-R, however, it was not detected at 98 kDa as expected and seen in the reciprocal MBP-R/His-R swap (Figure 3.1B). Therefore, a different approach was used for the His-R/MBP-R swap, which involved His-R and MBP-R crude cell lysates being mixed for 15 min in solution prior to binding to the Ni-NTA column. As a result, this method allowed for self-interaction, and CrhR was detected at both the expected 53 and 98 kDa shown in Figure 3.2B. All results in the protein exchange analysis were reproducible with each trial repetition.

#### 3.1.4 Swap analysis using non-denaturing BN-PAGE

In order to further differentiate between monomer and a dimer formation in solution, the protein self-interaction was also performed using Blue-native (BN)-PAGE, which does not denature the protein interaction between the differently tagged-CrhR constructs. However, since CrhR appears to dimerize in solution as shown in Figure 3.2B, it was difficult to interpret results of BN-PAGE, as single-tagged CrhR and the reciprocal swaps would appear to show the same size. The western blot in Figure 3.3 shows MBP-R alone migrates similar to both MBP-R and His-R reciprocals on the polyacrylamide gel, although some MBP-R alone also migrated faster on the gel. Protein sizes are not indicative of a monomer and a dimer but a difference in complex composition and breakdown products on the gel. The MBP-R alone shows breakdown product while the

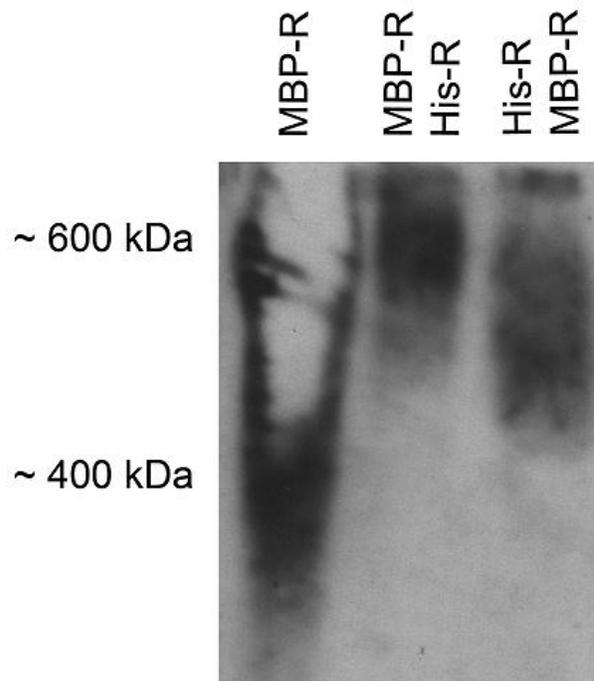


Figure 3.3: Self-dimerization of His-R and MBP-R on BN-PAGE. Protein eluates from MBP-R, MBP-R/His-R swap and His-R/MBP-R reciprocal swap were subjected to non-denaturing BN-PAGE, and analyzed on a western blot. Anti-CrhR antibody was used for detection of CrhR in all three fractionated fusion protein samples observed at 600 kDa, and 400 kDa only in MBP-R lane.

reciprocal swaps do not on the gel. The results do indicate that CrhR is capable of self-dimerization in the swap analysis when detected on non-denaturing PAGE.

### 3.1.5 Far-western analysis using purified His-R as a probe

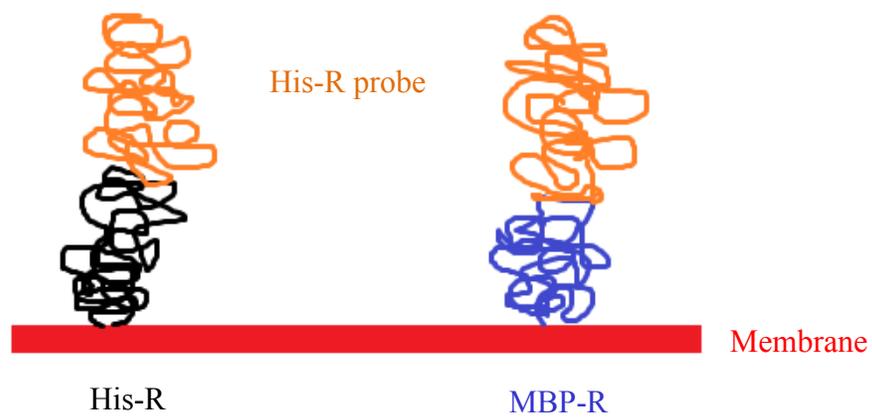
CrhR self-dimerization was also investigated through far-western analysis. His-R and MBP-R were separated by denaturing SDS-PAGE, and transferred to a nitrocellulose membrane. The membrane was incubated in buffer containing decreasing concentrations of urea to allow for denaturation and renaturation of CrhR on the membrane, as described in the Materials and Methods section 2.6.5. The membrane was then incubated O/N with a purified His-R protein as the probe (section 2.3.9) before incubation with anti-His antibody and detection by chemilluminescence on a western blot. If there is interaction between the tagged-CrhR on the membrane and the His-R probe, then His-R, as well as MBP-R, should be detected on the western blot.

Figure 3.4 shows the detection of His-R at high intensity and MBP-R at low intensity. The difference in intensity between His-R and MBP-R reflects the additional His-R on the membrane. The far-western results therefore confirm the previous analysis indicating CrhR self-dimerization.

### 3.1.6 Eliminating non-specific binding of GST-R, MBP-R and His-R

The results presented above suggest that CrhR fused to three different N-terminal tags self-dimerizes when tested by a variety of methods. In order to confirm these results, controls to eliminate the possibility of non-specific binding

A



B

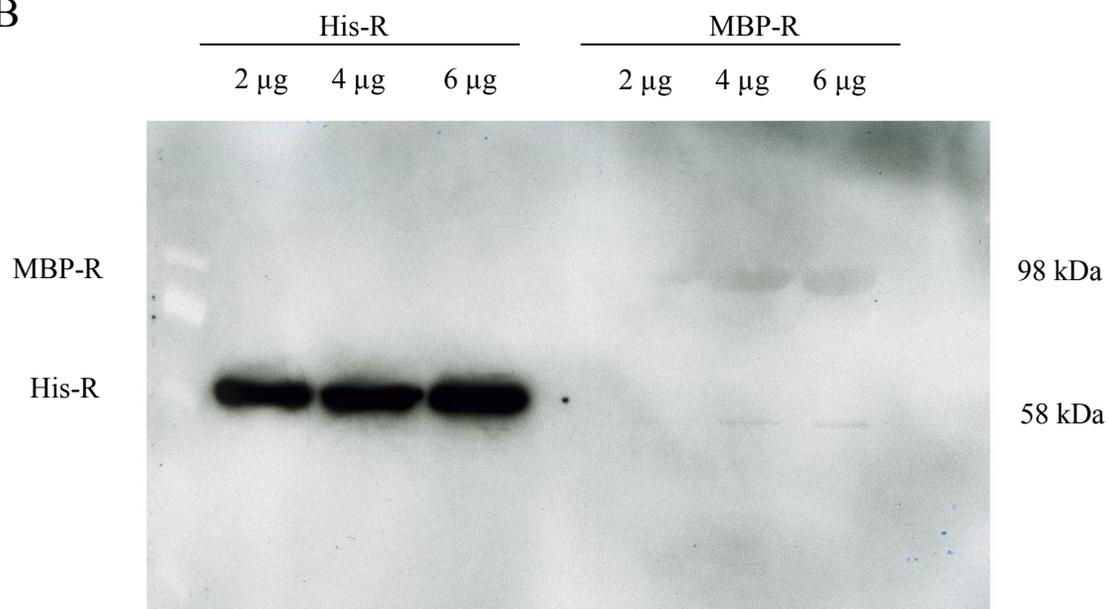
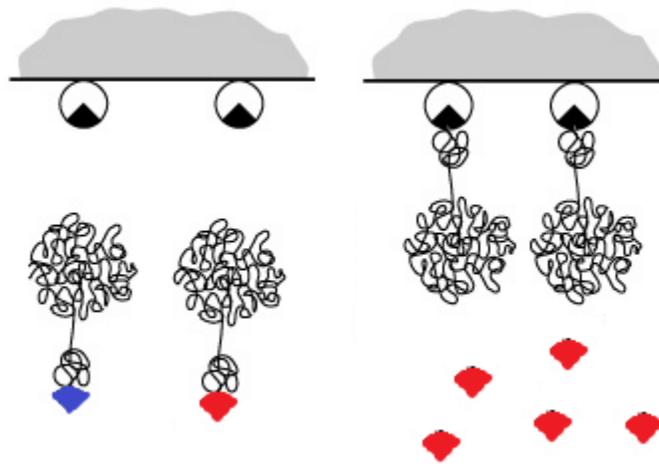


Figure 3.4: Far-western self-dimerization analysis of MBP-R using a purified His-R probe. (A) A schematic representation of His-R and MBP-R immobilized on a membrane and probed with a purified His-R in far western analysis. (B) MBP-R along with His-R at 2, 4 and 6  $\mu\text{g}$ , were size fractionated on SDS-PAGE and transferred to a nitrocellulose membrane, which was washed with decreasing concentrations of urea (described in more detail in section 2.6.5). The membrane was probed with purified His-R O/N and detected using anti-His antibody. His-R in the first 3 lanes was included as a control.

of CrhR fusion constructs were performed. Two questions were asked: 1) could CrhR be interacting with the fusion protein tags MBP, GST or His instead of itself, and 2) do the tags have an affinity to each other; in other words, is there non-specific binding between the tags (Figure 3.5A). To address the first question, the swaps described above were repeated using MBP and GST protein tags lacking the CrhR ORF. Figure 3.5B, lanes 1 and 2, show the MBP tag itself (43 kDa) was bound to the amylose resin column, with GST-R being passed over the column in lane 1, and His-R in lane 2. CrhR was not detected in either elution, indicating that GST- and His-tagged versions of CrhR were not able to bind to the MBP tag. This swap was also repeated with the GST tag which was bound to a glutathione column with the MBP-R and His-R lysates added to the beads resulting in similar results (data not shown).

The reciprocal swap was also performed in which MBP-R was bound to an amylose resin column, and the GST-tag lysate was then combined with it. In this case, eluted proteins were separated by SDS-PAGE and detected by Coomassie staining. As shown in lane 3 of Figure 3.5B, MBP-R but not the 26 kDa GST-tag protein was detected. This indicates that the 26 kDa GST tag alone does not have affinity for MBP-R. We can summarize that the CrhR self-dimerization observed in Figures 3.1, 3.2 and 3.3 does not result from CrhR having affinity for any of the fusion protein tags, or the fusion protein tags having affinity for each other.

A



B

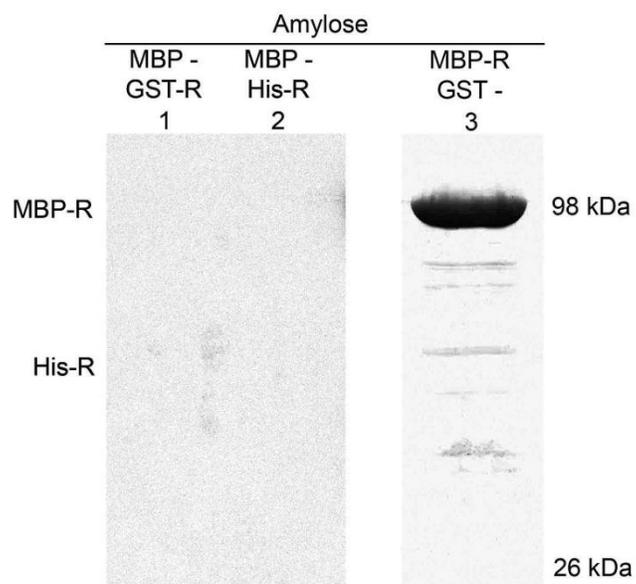


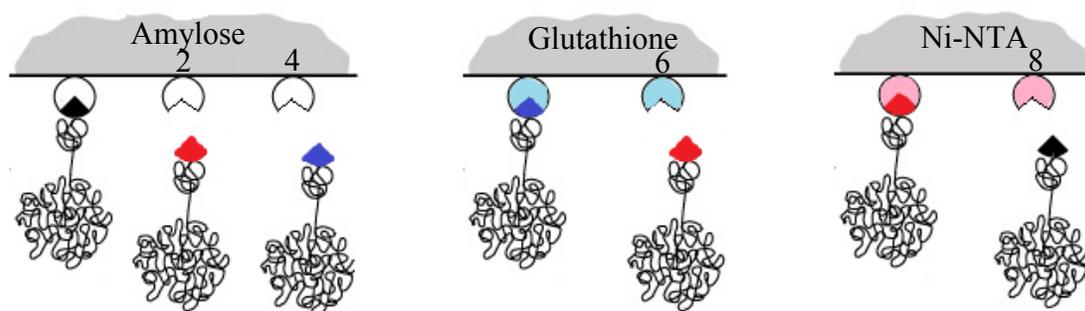
Figure 3.5: Controls to exclude non-specific binding of CrhR in protein exchange analysis. (A) A MBP construct (black) without the CrhR insert was affinity purified on the amylose resin column (white), and both GST-R (dark blue) and His-R (red) were passed over the bound beads in 1 and 2, respectively. In a reciprocal control, MBP-R (black) was affinity purified on amylose resin (white), and GST construct (dark blue) without the CrhR insert was passed over the bound beads in 3. (B) Western analysis of the controls in (A) using anti-CrhR antibody for detection of MBP/GST-R and MBP/His-R in lanes 1 and 2, respectively. MBP-R/GST was detected on a Coomassie stained gel in lane 3.

To confirm CrhR self-dimerization observed in the protein exchange analysis, the possibility of the fusion protein tags having an affinity to non-specific columns needed to be eliminated (Figure 3.6A). On the western blot shown in Figure 3.6B, binding of GST-R and His-R to the amylose resin column (lane 2 and 4, respectively), GST-R to the Ni-NTA column (lane 6), and MBP-R to the glutathione column (lane 8) was tested. The lack of CrhR detection in the elutions from these columns indicates that neither CrhR nor the protein tags have an affinity to non-specific columns. On this blot, positive swaps were repeated as internal controls in lanes 1 and 7 (MBP-R and GST-R, and reciprocal respectively); 3 (MBP-R and His-R); and 5 (GST-R and His-R). The results indicate that the CrhR self-dimerization pattern is not a result of the tagged-CrhR protein binding non-specifically to an affinity column.

### 3.1.7 FPLC-gel filtration analysis *in vitro* using His-R

In order to further examine the self-dimerization of CrhR, FPLC-gel filtration analysis was performed using the affinity purified His-R fusion construct. His-R, overexpressed and purified from *E. coli*, was separated on a Superdex 200BKL size exclusion column, and absorbance was measured at 280 nm. Protein standards were also run to determine the elution profiles of known proteins shown in Figure 3.7, and used to compare the elution profiles of the purified His-R FPLC run in Figure 3.8. The elution profile reveals the presence of 4 peaks of eluted protein.

A



B

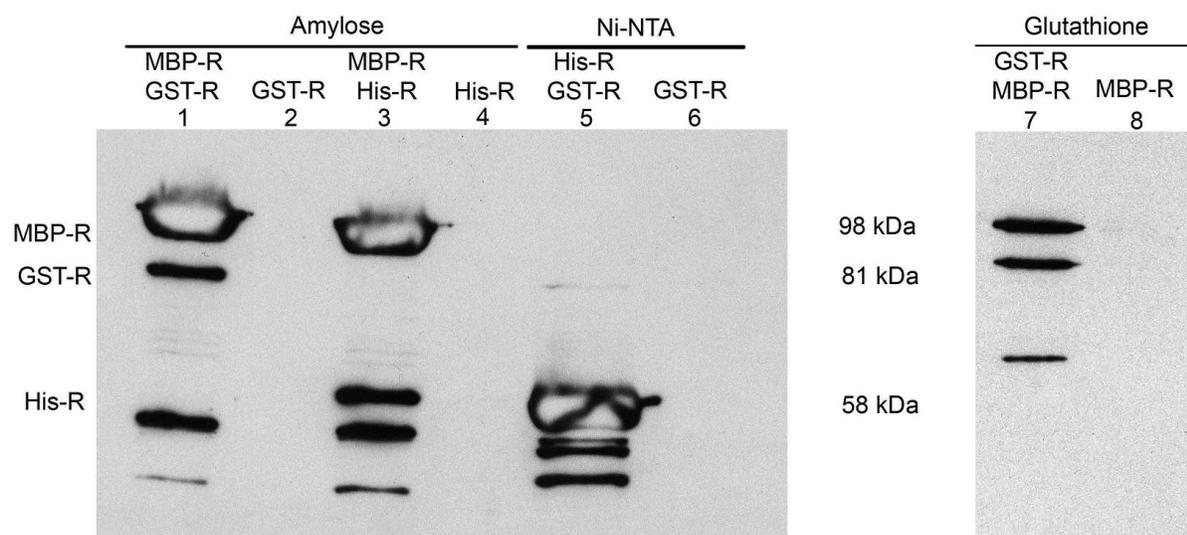
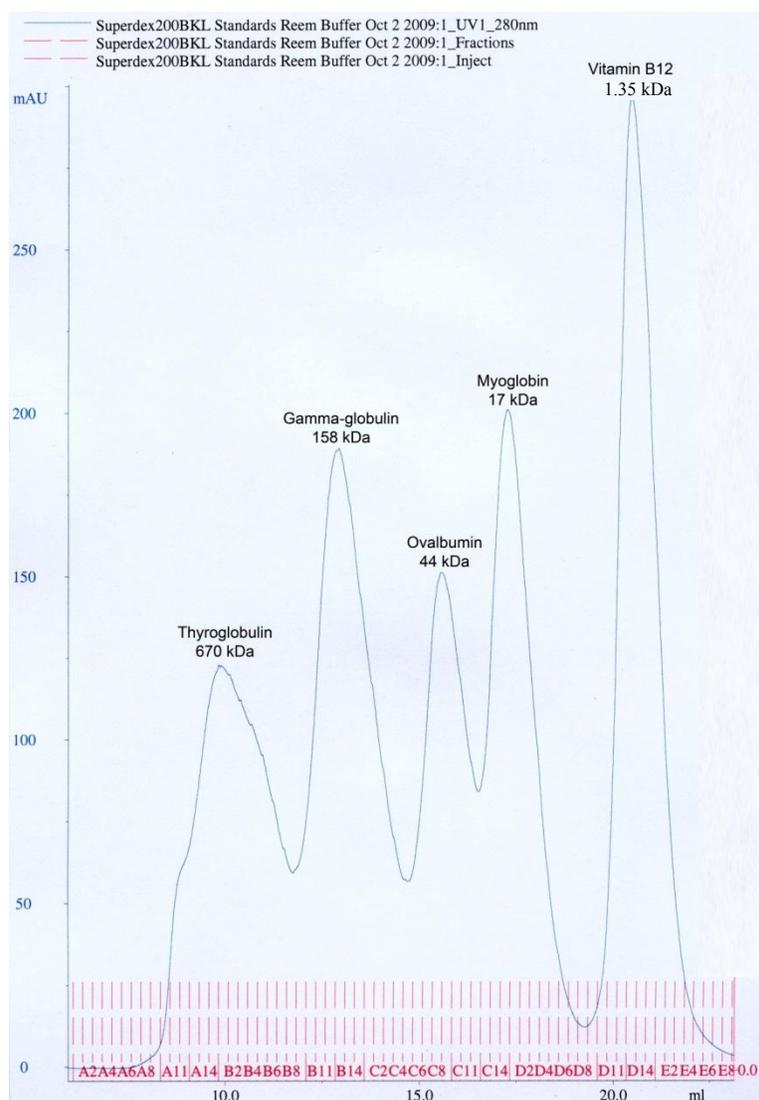


Figure 3.6: Non-specific controls to rule out binding of the three differently tagged-CrhR to non-respective affinity columns. (A) A schematic representation of the overexpressed tagged-CrhR constructs affinity purified on respective and non-respective columns to test for non-specific binding of the tags to non-specific columns. His-R (red) and GST-R (dark blue) were incubated with the amylose resin column (white), His-R (red) was incubated with the glutathione agarose column (light blue), and MBP-R (black) was incubated with the Ni-NTA column (pink). (B) Western analysis of the non-specific binding controls using anti-CrhR antibody in lanes 2, 4, 6 and 8, respectively. Positive swaps were size fractionated alongside each control for comparison in lanes 1, 3, 5 and 7. All extra bands not corresponding to the labeled protein sizes were deemed to be breakdown products.

A



B

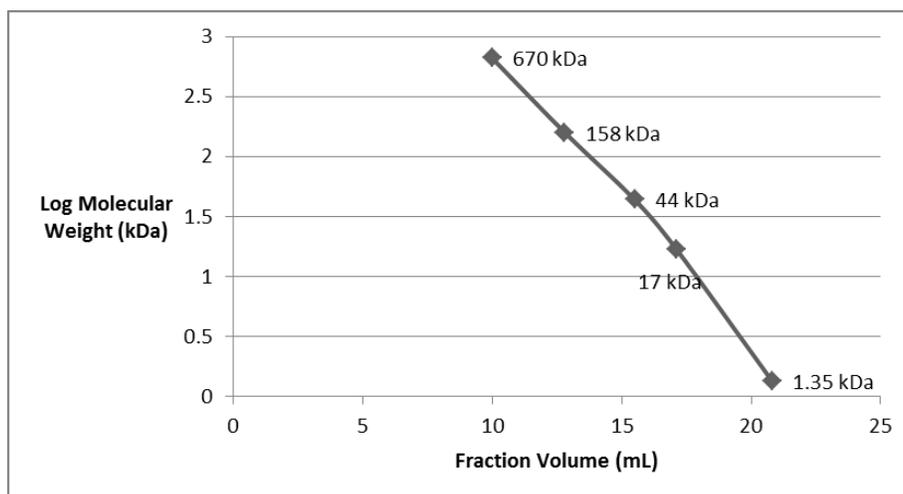


Figure 3.7: FPLC gel filtration analysis of standard molecular weight markers using the Superdex 200BKL column. (A) Molecular weight standards were size separated on a Superdex 200BKL with a fractionation range of 10,000 to 600,000 Daltons. Protein standards included thyroglobulin (670 kDa),  $\gamma$ -globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa) and vitamin B12 (1.35 kDa). Absorbance at 280 nm is shown by the blue line. Red streaks represent the 250  $\mu$ L fractions collected throughout the run at the corresponding volume (mL shown in black on the x-axis). (B) Based on the FPLC results, a graph of the molecular weight size standards was generated as a standard curve for the later determination of the sizes of the purified tagged-CrhR protein samples analyzed.

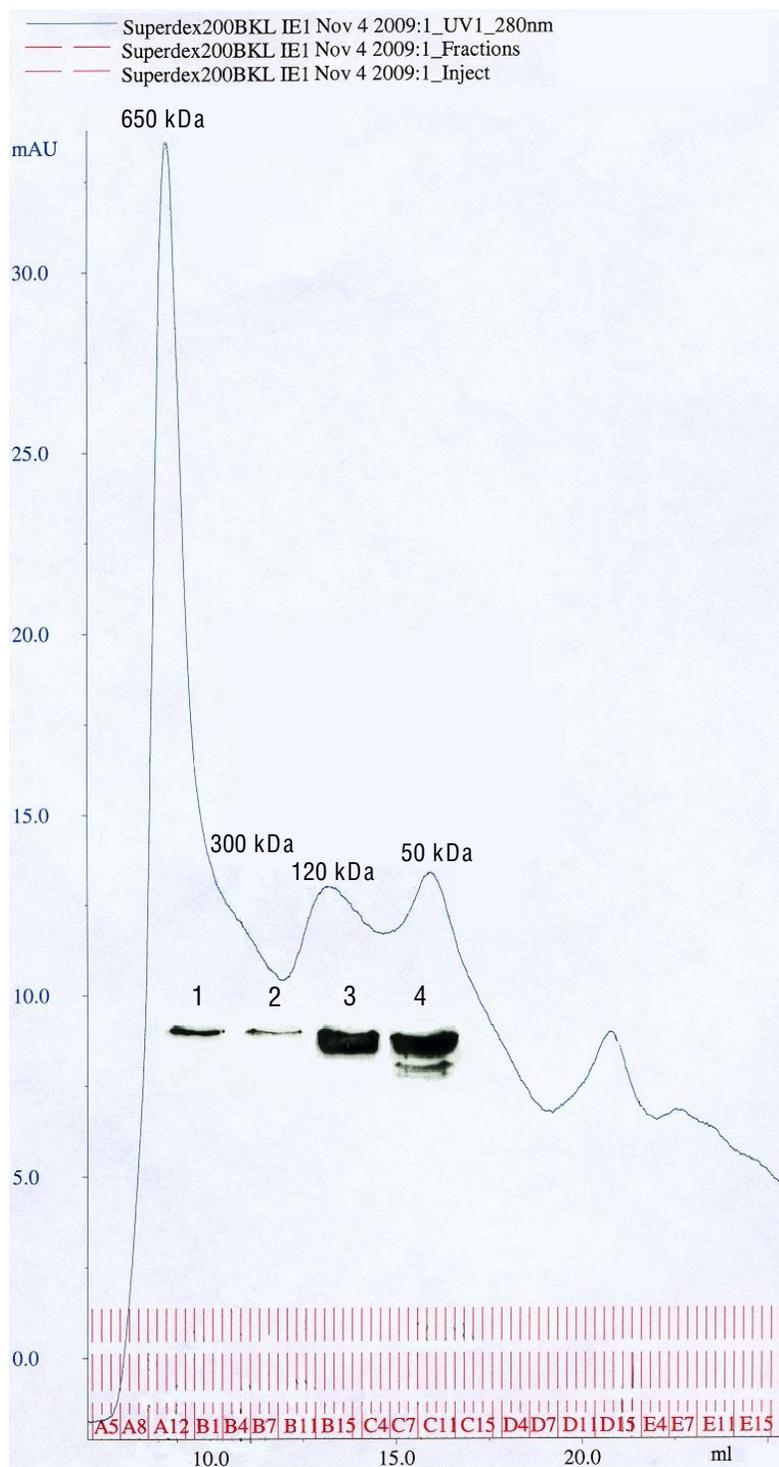


Figure 3.8: FPLC gel filtration analysis of purified His-R using the Superdex 200BKL column. Overexpressed and affinity purified His-R was separated on a Superdex 200BKL column. Fractions within each peak present were combined, TCA precipitated and size fractionated on SDS-PAGE. Superimposed below each peak on the column trace are the results from chemiluminescent detection of CrhR by western analysis. A standard curve (Figure 3.7B) was used to determine the approximate molecular weight of proteins in each peak (shown in kDa). CrhR was detected by western analysis using anti-CrhR antibody.

The presence of CrhR as detected by western analysis in peaks of ~ 50 kDa and ~120 kDa suggesting it was eluting in both monomeric and dimeric form. Fractions within each of the peaks were combined and separated by SDS-PAGE, blotted with anti-CrhR antibody, and western results confirmed the presence of CrhR within peaks 3 and 4 of Figure 3.8. Detection of low levels of CrhR high a molecular weight peak eluting between approximately 650 kDa and 300 kDa, (lane 1 and 2, respectively) suggests that CrhR also forms higher multimer complexes. These complexes may represent aggregated material resulting from overexpression of the protein, which will be discussed further in section 3.3.1. In summary, the results from the gel filtration analysis indicate that His-tagged CrhR is capable of self-dimerization, as was observed previously with the yeast two-hybrid and swap analyses.

### 3.2 Localization of the self-dimerization domain of CrhR

#### 3.2.1 Yeast two-hybrid analysis to identify the dimerization domain

Given the evidence that CrhR self-dimerizes as shown through the yeast two-hybrid, protein exchange or swap and FPLC analysis, the next goal was to determine the general location of the dimerization site within the CrhR protein. In order to locate the dimerization domain of the CrhR protein, interaction of full length CrhR with either the N-terminal domain or the C-terminal domain was examined in yeast two-hybrid analysis. Two constructs were made in an attempt to identify the dimerization domain. The first was an the N-terminal domain containing the RecA-like domain 1 of the helicase core and lacking the C-terminal

domain or RecA-like domain 2 and the C-terminal extension unique to CrhR as previously shown in Figure 2.1 of the Materials and Methods, and the second was the C-terminal domain containing the RecA-like domain 2 of the helicase core along with the extended C-terminal domain, and lacking the N-terminal domain or RecA-like domain 1. Interaction between CrhR and the N-terminal (JG5'), the C-terminal (JG3') and the extended C-terminal (JG-END) were analyzed by  $\beta$ -Gal screening and leucine selection (section 2.2.3 and 2.2.4), as reported in Table 3.1.

Interaction was positive between CrhR and the N-terminal domain (EGR/JG5'), as colonies were blue-colored on X-Gal containing plates, and there was growth observed on plates lacking leucine. The same cannot be said in regards to CrhR interaction with the C-terminal (EGR/JG3'); the colonies were white on X-Gal containing plates, and there was no growth on plates lacking leucine. This was also shown to be the case, as self-interaction of the N-terminal (EG5'/JG5') was positive, while N-terminal interaction with the C-terminal (EG5'/JG3') was negative (Table 3.1). These results were further confirmed by using the negative controls including CrhR on either of the pEG202 and pJG4-5 lacking inserts, as well as a positive control YKR/YBR. These results indicate that CrhR is capable of dimerization, and using the yeast two-hybrid analysis shows that the dimerization domain for CrhR is predominantly localized to the N-terminal.

### 3.2.2 Yeast two-hybrid analysis of the C-terminal domain and C-terminal extension

The results were not consistent when comparing CrhR interaction with the full C-terminal (Domain 2 and the extended C-terminal) (EGR/JG3'), which was negative for interaction, and CrhR interaction with just the extended C-terminal (EGR/JG-END), which was positive for interaction in both  $\beta$ -Gal screening and leucine selection. This suggests that the C-terminal extension maybe involved in CrhR self-dimerization, however, it is unknown to what extent. Perhaps the C-terminal extension is important for the stability of the dimer, but this requires a different approach and method to test other than the yeast two-hybrid system.

The JG3' and JG-END constructs contain the extended C-terminal, similar results would be expected for both constructs, but this was not the case. The extended C-terminal was also tested for interaction on itself (EG-END/JG-END) and as a negative control against empty vectors (EG-END/pJG4-5 and pEG202/JG-END), but results were also inconsistent. In terms of EG-END/JG-END,  $\beta$ -Gal screening showed interaction as colonies were blue in color, yet there was no growth on plates lacking leucine. And the negative controls EG-END/pJG4-5 and pEG202/JG-END gave false positives, as there was blue colored colonies and in most cases growth on plates lacking leucine (Table 3.1). These results therefore are regarded as false positives because any construct containing the END fragment produced a positive reaction in the absence of an interacting partner. Therefore, both the C-terminal and extended C-terminal constructs were tested to see if the expected proteins were being overexpressed.

Figure 3.9 shows results of the C-terminal protein overexpression using 3 independent EGR/JG3' constructs. In each case, the detection of both CrhR fused to the DBD (BD-CrhR; expected molecular weight of 79 kDa), and the C-terminal fused to the AD (AD-3'CrhR; expected molecular weight of 48 kDa) were detected. Therefore, we can conclude constructs containing the C-terminal domain (JG3') are functional due to the expression of a protein product in yeast.

In terms of the extended C-terminal constructs with either EG-ENG or JG-END, the conclusions differ. Shown in Figure 3.10 in lane 1, the EGR/JG-END construct was positive for interaction, however, only AD-END is detected on the western, but not the BD-CrhR, at ~79 kDa. Thus a positive yeast two-hybrid with EGR/JG-END is due solely to the presence of AD-END. This indicates AD-END is somehow activating the yeast two-hybrid system on its own. Similarly, for the negative control constructs, EG-END/pJG4-5 and pEG202/JG-END, there is no observed proteins in the empty vectors overexpressing, as only AD-END at ~30 kDa and BD-END at ~38 kDa was detected on the western blot in lanes 3 and 4, respectively. This again indicates a positive interaction signal from JG-END alone, therefore revealing that any constructs containing the extended C-terminal domain END construct gives inconclusive results in the yeast two-hybrid system.

These results potentially indicate that the END protein activity was capable of activation of the yeast two-hybrid system on its own. Therefore, rather than repeating the amplification and cloning of the END constructs in the yeast two-hybrid assay, a different approach was taken, which involved the deletion of the C-terminal extension.

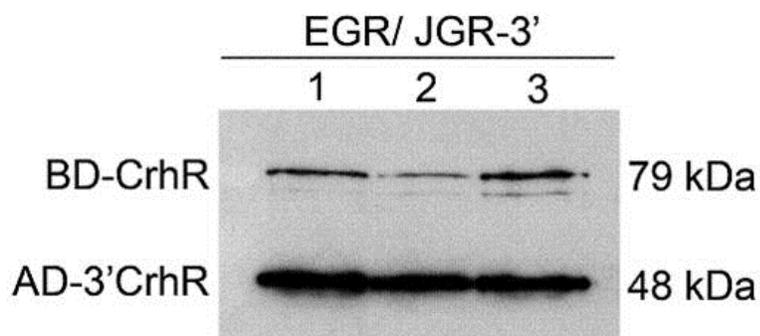


Figure 3.9: Protein overexpression analysis of yeast two-hybrid constructs containing the C-terminal truncation. Three different EGR/JG3' constructs containing full length CrhR attached to the DBD (BD-CrhR) and CrhR lacking Domain 2 and the extended C-terminal attached to the AD (AD-3'CrhR) were grown and lysed to extract proteins in order to test for overexpression of both proteins in the constructs. Proteins were size fractionated on SDS-PAGE and analyzed by Western using anti-CrhR antibody.

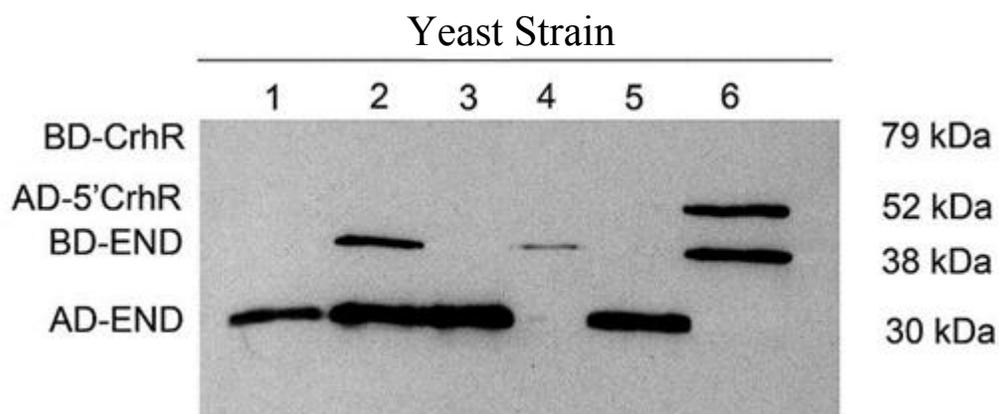


Figure 3.10: Protein overexpression analysis of yeast two-hybrid constructs containing the extended C-terminal truncation (END). Cells were lysed and proteins extracted from six constructs; EGR/JG-END in lane 1, EG-END/JG-END in lane 2, negative controls pEG202/JG-END in lane 3 and 5, and EG-END/pJG4-5 in lane 4, and finally EG5'/JG-END in lane 6. Samples were fractionated on SDS-PAGE and analyzed by western analysis using anti-CrhR antibody. Negative controls only have AD-END or BD-END. Expected molecular weight of the constructs are: BD-CrhR, expected at ~ 79 kDa, AD-5'CrhR, expected at ~ 52 kDa, BD-END, expected at ~ 38 kDa and AD-END, expected at ~ 30 kDa.

### 3.2.3 Truncation of the extended C-terminal domain, His-C2R

In order to test whether the extended C-terminal truncation is required for CrhR dimerization, the His-C2R constructs containing only the helicase core region were constructed as described in section 2.3.3. Protein exchange analysis was repeated using this construct and the MBP-R and GST-R constructs to test for interaction. Figure 3.11 shows the truncated-CrhR produced from His-C2R construct at approximately 48 kDa, in lane 2, as compared to the His-R construct which is detected at 53 kDa. His-C2R possesses the 334 aa helicase core and 53 aa downstream of the HRIGR-box; therefore excluding the C-terminal extension that is unique to CrhR.

In order to test whether the dimerization site was in the deleted C-terminus of CrhR, we performed protein exchange analyses with the truncated His-C2R, along with GST-R and MBP-R. If the C-terminus was the only domain required for dimerization we would expect that GST-R and MBP-R would not be able to bind to the truncated His-C2R. In Figure 3.12, His-C2R was attached to the Ni-NTA column, and GST-R (lanes 3 and 4), and MBP-R (lanes 5 and 6) was passed over the bound beads. Western analysis revealed that a low level of interaction was still occurring between the truncated and full length CrhR. However, the level of interaction detected in eluates was significantly reduced compared to that detected with full length versions of tagged-CrhR. An increased level of interaction was detected if the beads from the protein exchange analysis were boiled. Therefore, these results suggest that CrhR dimerization is not fully

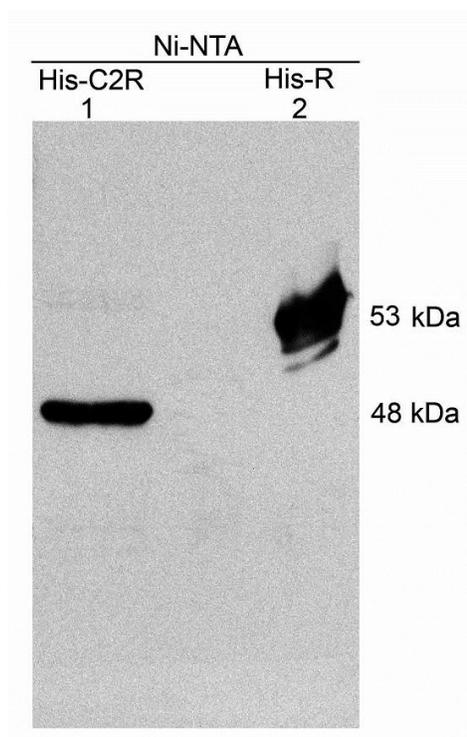


Figure 3.11: Overexpression analysis of His-tagged and truncated CrhR lacking the C-terminal extension. The CrhR C-terminal extension truncation construct, His-C2R, was overexpressed in *E. coli* BL21, and affinity purified on the Ni-NTA column. His-C2R was size fractionated on SDS-PAGE (48 kDa) in lane 1, alongside the full length CrhR, His-R (58 kDa).

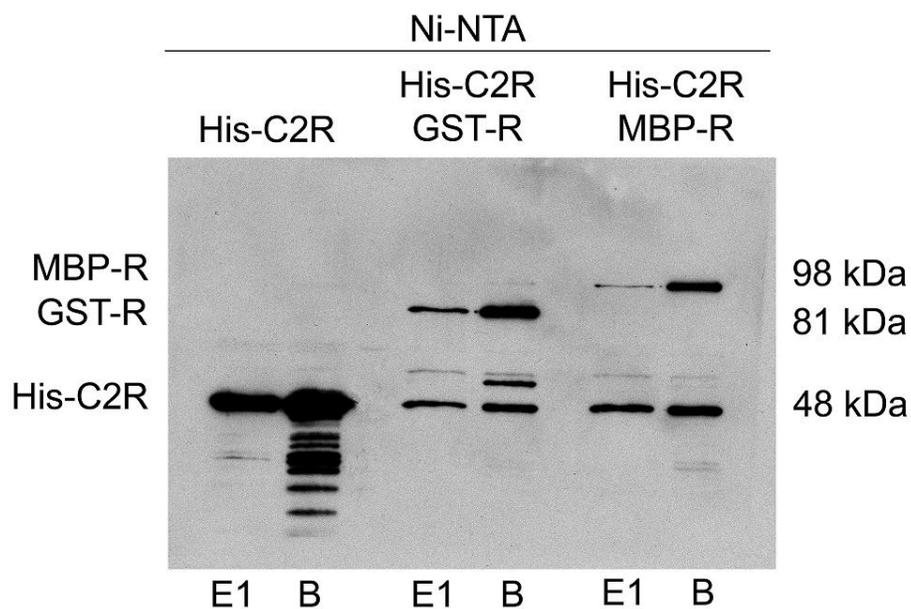


Figure 3.12: Self-dimerization of the His-tagged and truncated CrhR lacking the C-terminal extension. His-C2R was affinity purified on the Ni-NTA column, and eluted in first 2 lanes. Protein exchange analysis was done using GST-R (two middle lanes) and MBP-R (last two lanes) which were passed over the bound His-C2R and beads. E1 represents elutions from the column using elution buffer, while B represents elutions from the column by resuspending the beads in 0.1 M DTT/CO<sub>3</sub> and boiling for 4 min. All extra bands not corresponding to the labeled protein sizes are deemed to be breakdown products.

dependent on the C-terminal extension, which suggests that the N-terminal domain is the principle domain, but the C-terminal possibly plays a role.

### 3.2.4 Partial N-terminal truncation, His-N2R

To confirm the results from the His-C2R construct (section 3.2.3), an N-terminal truncation was constructed, His-N2R, and tested with protein exchange analysis. The 422 aa His-N2R construct lacks part of the N-terminal region including motif I and Ia, but containing the DEAD-box motif. His-N2R was overexpressed using different concentrations of IPTG and for different amounts of time, as shown in Figure 3.13. Western analysis revealed detection of His-N2R at ~49 kDa.

His-N2R was tested in the protein exchange analysis for dimerization using MBP-R. Interactions were not detected in either protein exchange or swap in elution fractions (lanes E1 in Figure 3.14) nor in boiled beads (lanes B in Figure 3.14) from the MBP-R and His-N2R swap. Some interaction was observed in the boiled beads elution from the His-N2R/MBP-R swap. Both results indicate that the N-terminal is important for interaction. Since this is only a partial N-terminal truncation, the dimerization site of CrhR may require the DEAD-box motif and the following motifs in the RecA-like domain of the N-terminal, which is still present in the His-N2R construct. This could mean the entire N-termini or all of Domain 1 needs to be truncated to completely abolish interaction. Another possibility is that the C-terminal, Domain 2 and the C-terminal extension or just the C-terminal extension, is still involved in the dimerization of CrhR. These

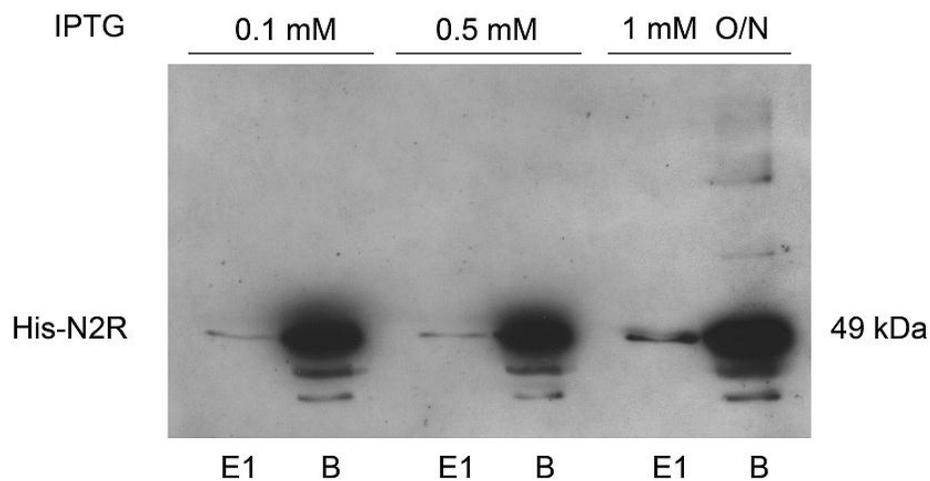


Figure 3.13: Overexpression analysis of His-tagged and truncated CrhR lacking part of the N-terminal domain. The N-terminal truncation of CrhR, His-N2R (~ 49 kDa), was induced with 0.1 mM and 0.5 mM IPTG for 2 hours and with 1 mM IPTG O/N. Cells were harvested, lysed and affinity purified on the Ni-NTA column. E1 represents elutions of the bound His-N2R to beads using elution buffer, and B represents elutions from the column by resuspending the beads in 0.1 M DTT/CO<sub>3</sub> and boiling for 4 min.

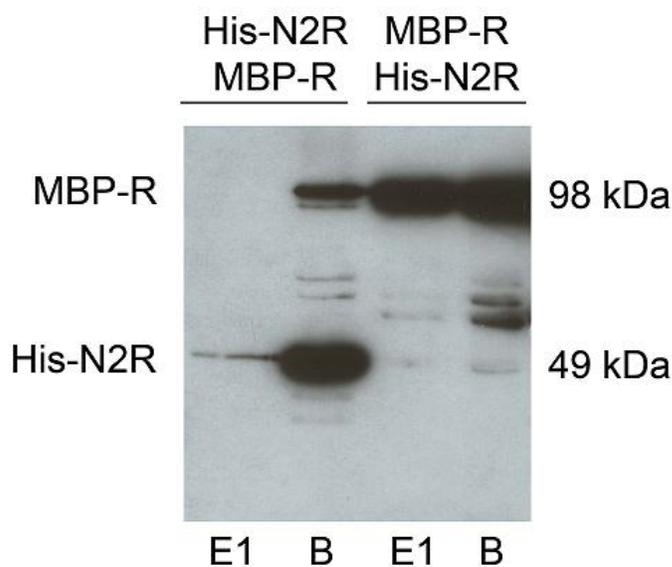


Figure 3.14: Self-dimerization of His-tagged and truncated CrhR lacking part of the N-terminal domain. His-N2R was affinity purified on the Ni-NTA column in the first 2 lanes, and MBP-R was passed over the bound beads. In the reciprocal protein exchange or swap in the last two lanes, MBP-R was affinity purified to the amylose resin column, and His-N2R was passed over the bound beads. E1 represents elutions from the column using elution buffer, while B represents elutions from the column by resuspending the beads in 0.1 M DTT/CO<sub>3</sub> and boiling for 4 min. All extra bands not corresponding to the labeled protein sizes are deemed to be breakdown products.

possibilities need to be further examined in order to make a final conclusion about the dimerization site of CrhR, but overall the data indicates that the N-terminal does appear to have a major role in CrhR dimerization.

### 3.3 CrhR Interaction with Protein Complexes

#### 3.3.1 RNA-independent multimerization of CrhR *in vitro* and *in vivo*

The dimerization results reported above could result from two CrhR monomers interacting independently with RNA. Lower levels of CrhR were detected in the peak eluting with the highest apparent molecular weight at approximately 650 kDa, the center of peak 1, and 300 kDa, the tailing end of peak 1 in Figure 3.8 (lane 1 and 2, respectively). To test if this is aggregated CrhR due to interaction of RNA and CrhR, FPLC analysis was repeated on samples treated with RNase A before the chromatography.

Affinity purified His-R samples were incubated with RNase A and analyzed by gel filtration analysis as described in section 2.4.3. This was expected to reveal whether CrhR binding to RNA is required for dimerization, and if more than one dimer binds to the RNA. In Figure 3.15, the RNase A treated His-R resulted in the presence of two peaks and a shoulder, in FPLC analysis. Fractions covering each peak were combined and separated by SDS-PAGE. Detection of CrhR within each peak determined if there was any shift detectable by western blot due to the presence or absence of RNA in the samples. Western blot analysis reveals the detection of CrhR in the 3 peaks as shown previously representing the monomer, dimer and multimer range, regardless of the presence of RNase A in the sample (Figure 3.15). Although there appears to be a slight difference in the

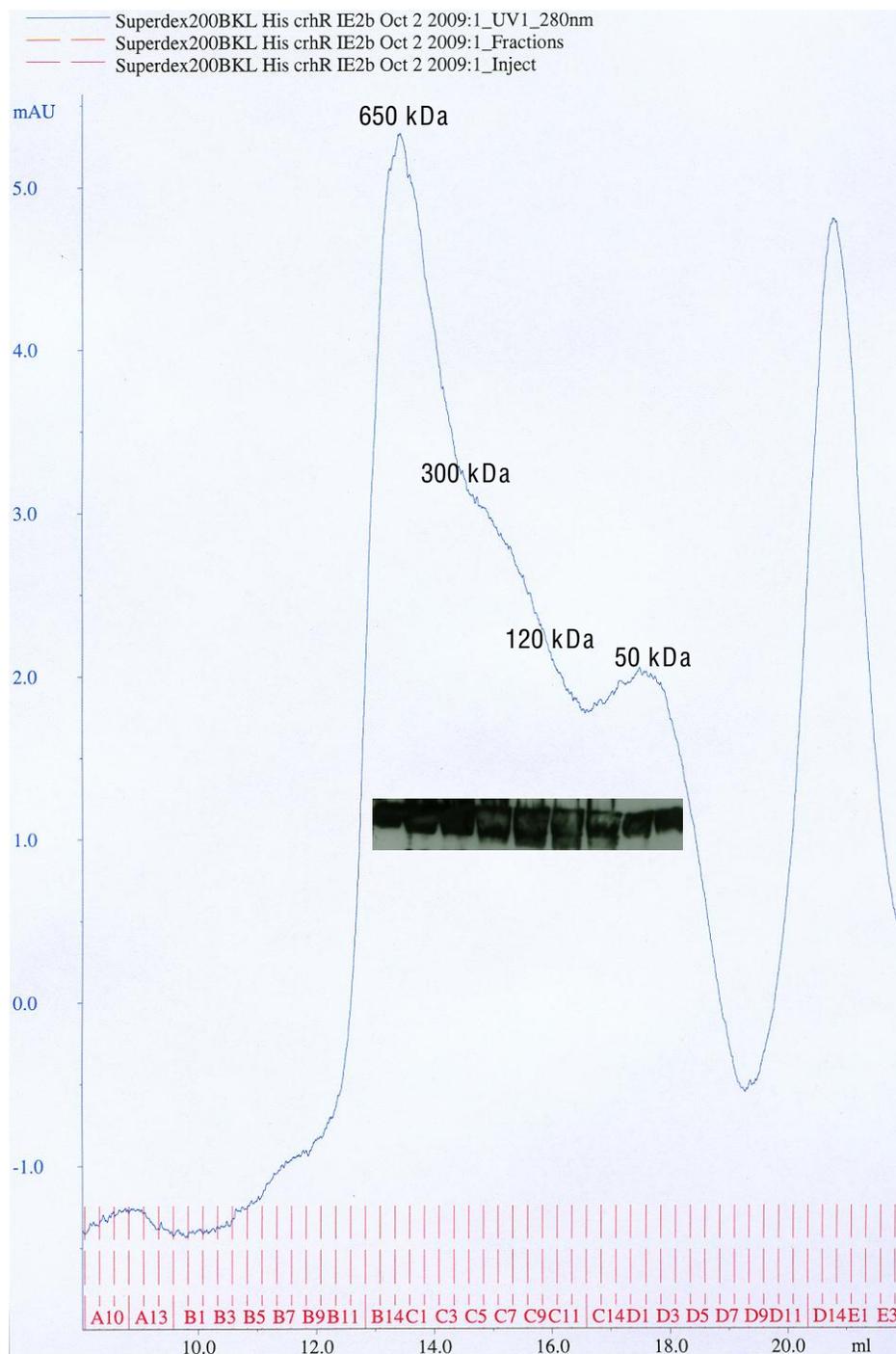


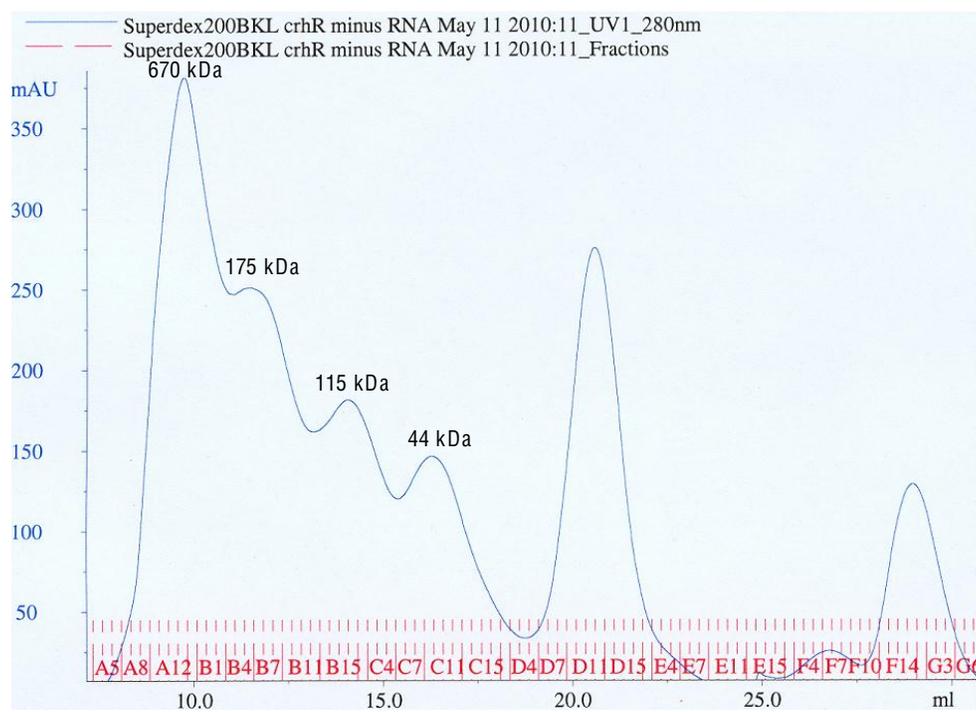
Figure 3.15: FPLC gel filtration analysis of the purified His-R treated with RNase A. Overexpressed and affinity purified His-R treated with 10 mg/mL RNase for 10 min at 37°C was separated on a Superdex 200BKL column. Five fraction samples across the range of the first two peaks were combined, corresponding to the bands on the graph, and TCA precipitated, size fractionated on SDS-PAGE, and analyzed by western using anti-CrhR antibody. A standard curve (Figure 3.7B) was used to determine the approximate molecular weight of proteins in each peak (shown in kDa) and superimposed below the peaks on the column trace. CrhR was detected by western analysis using anti-CrhR antibody. Untreated samples were in Figure 3.8.

peaks observed between the treated and non-treated His-R sample (Figure 3.8), there was no difference in CrhR detection within the monomer, dimer, and multimer range.

This finding suggests that CrhR was not forming multimers due to the presence of RNA in the samples. However, this was tested on the overexpressed His-R in *E. coli*. Therefore, this could also mean that CrhR does not interact with *E. coli* RNA. In order to confirm CrhR eluting at high molecular weight in FPLC analysis is RNA-independent, the experiment was repeated using native cyanobacterial *Synechocystis* sp. PCC6803 6803 samples, first to see whether there is higher multimer formation, and then test if it is RNA-independent.

Native cyanobacterial *Synechocystis* cells were grown at 30°C, and cold shocked at 20°C O/N, as CrhR expression is significantly induced by exposure to low temperature. Proteins were extracted and analyzed by gel-filtration to test whether CrhR is capable of multimerization *in vivo*. The profile of protein elution from a Superdex 200BKL column for the native cell lysate sample without RNase A treatment is shown in Figure 3.16A, and the native cell lysate sample with RNase A treatment is shown in Figure 3.16B. Cell lysates containing 10 mg/mL RNase A were treated for 30 min at 37°C. Since this is a native cell lysate, the graph represents the profile of all *Synechocystis* proteins; therefore, presence of peaks does not indicate whether it is due to the presence of CrhR in the corresponding fractions. Western blot analysis was used to detect CrhR, by combining 5 consecutive fractions across the range of the observed peaks. If CrhR is binding to RNA, and thus forming a higher molecular weight complex, then the

A



B

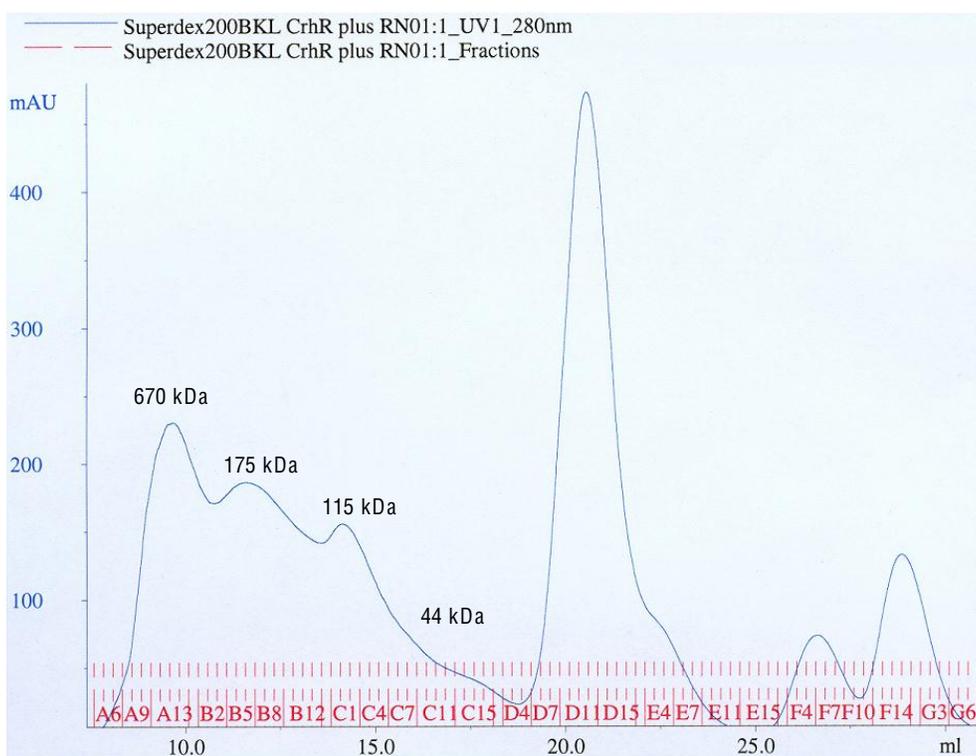


Figure 3.16: FPLC gel filtration analysis of native *Synechocystis* sp. PCC 6803 proteins with/ or without RNase A treatment. (A) Native proteins were extracted from cold-shocked *Synechocystis* sp. at 20°C O/N, and were separated on a Superdex 200BKL column. (B) Samples were treated with 10 mg/mL RNase A for 30 min at 37°C before FPLC analysis. Five fractions each across the four peaks (A) and the three peaks (B) were combined, TCA precipitated, size fractionated on SDS-PAGE, and either Coomassie stained or analyzed on a western using anti-CrhR antibody. Peaks observed in each (A) and (B) are not solely CrhR because whole cell lysate was applied to the column.

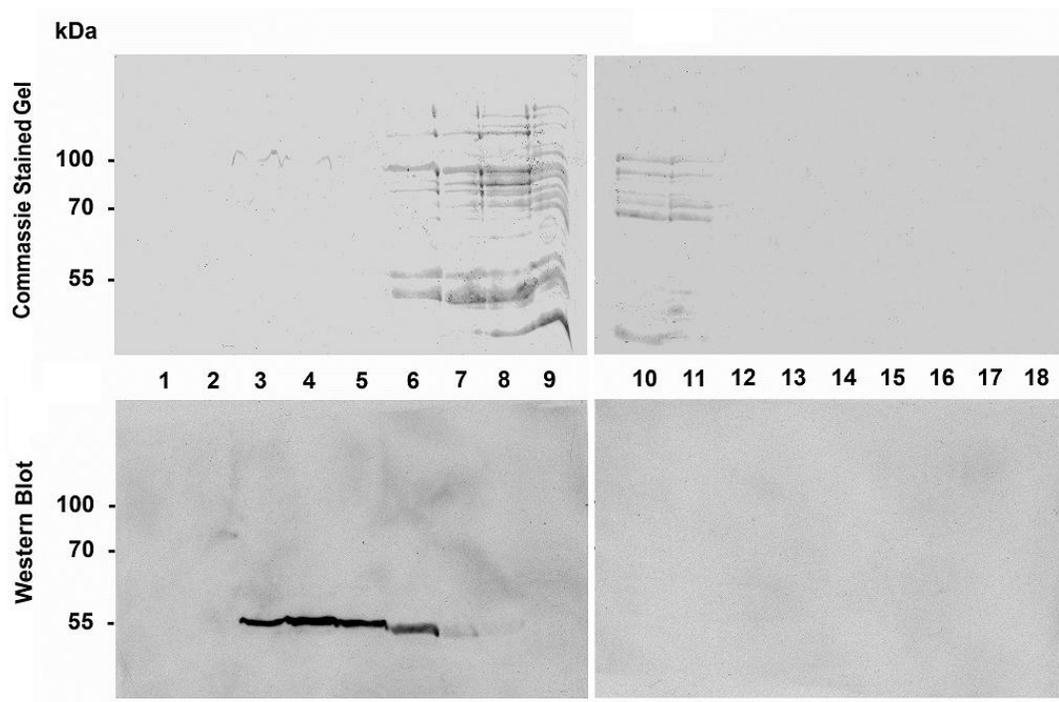


Figure 3.17: Coomassie and western blot analysis of native *Synechocystis* sp. PCC 6803 proteins from FPLC gel filtration analysis. Proteins from native *Synechocystis* lysate not treated with RNase A were analyzed on Coomassie stained gels represented in the top two panels, and by western blot analysis in the bottom two panels. Each lane represents a sample containing 5 combined fractions corresponding to FPLC graph in Fig. 3.18A, starting from the left hand side.

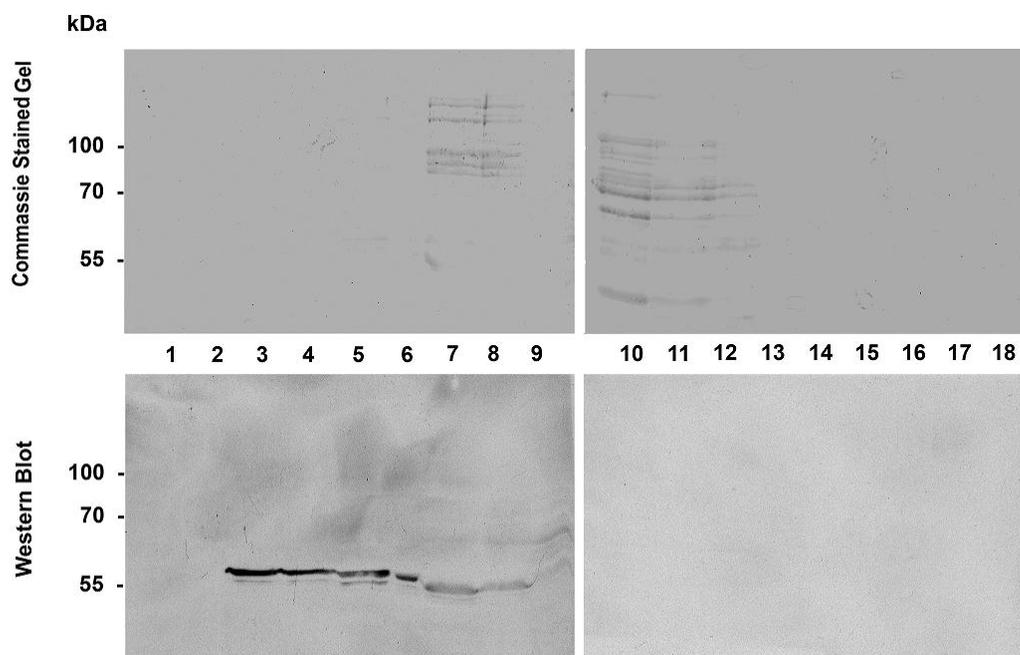


Figure 3.18: Coomassie and western blot analysis of native *Synechocystis* sp. PCC 6803 proteins from FPLC gel filtration analysis treated with RNase A. Proteins from native *Synechocystis* lysate treated with RNase A for 30 minutes at 37°C prior to FPLC analysis. Fractions were analyzed on Coomassie stained gels represented in the top two panels, and by western blot analysis in the bottom two panels. Each lane represents a sample containing 5 combined fractions corresponding to FPLC graph in Fig. 3.18B, starting from the left hand side.

addition of RNase A would result in CrhR migration at a lower molecular weight. As observed in Figure 3.17 and Figure 3.18, western blot analysis indicates similar CrhR elution profiles with or without RNase A treatment. A minor detection at lower molecular weight is observed regardless of the absence or presence of RNase in the samples, respectively.

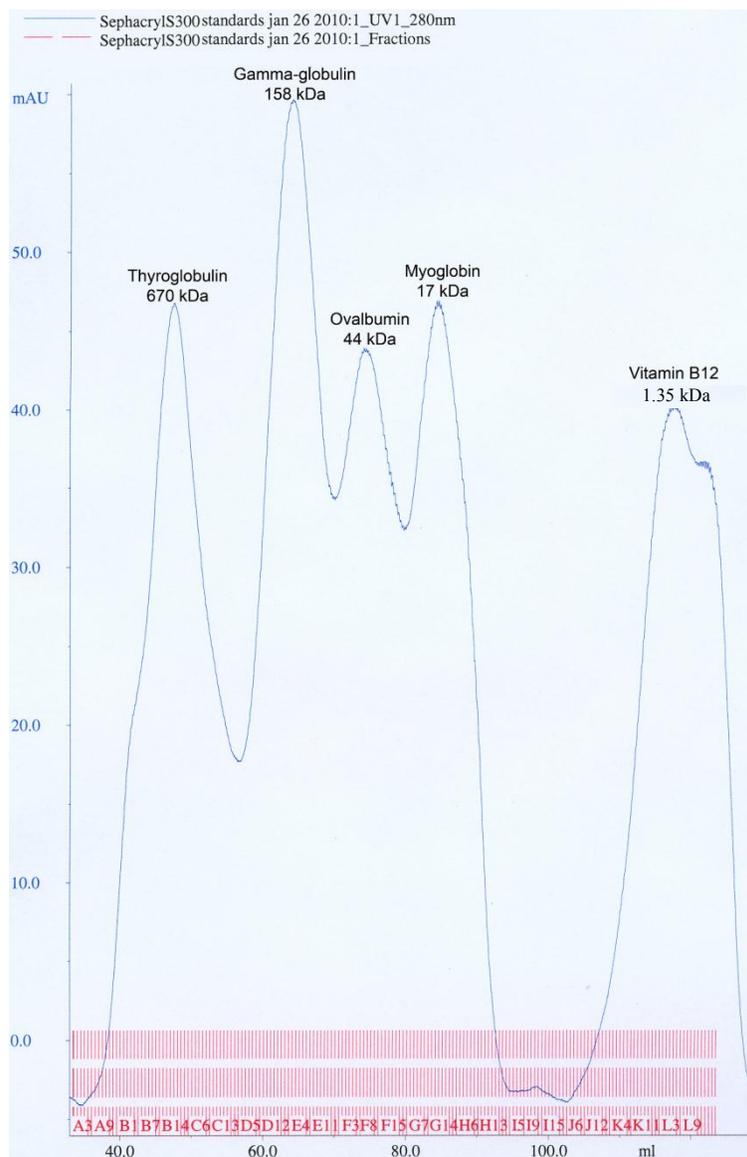
These results indicate that CrhR is associated with high molecular weight complexes *in vivo*, which is not RNA dependent. Whether the elution profile results from CrhR self-assembly or CrhR presence in a multi-subunit complex will be further investigated in Section 3.3.3.

### 3.3.2 Temperature-dependent multimerization of CrhR *in vivo*

Since it is known that CrhR expression is significantly induced by exposure to low temperature, we repeated detection of CrhR in cells cold shocked, as well as grown at 30°C, which is referred to as the warm treatment. Native cyanobacterial *Synechocystis* cells were grown at 30°C, and either cold-shocked O/N at 20°C, or left at 30°C. Proteins were extracted and analyzed by gel-filtration in this case on a Sephacryl S300 size exclusion column, as it provides separation over an extended molecular weight range. The profile of protein elution from a Sephacryl S300 column for the molecular weight standards is shown in Figure 3.19. Elution profiles of CrhR expressed at 20°C and 30°C are shown in Figures 3.20 and 3.21, respectively.

Cold-shocked protein samples resulted in the presence of 2 major protein peaks that span the monomer to multimer range, as shown in Figure 3.20. The

A



B

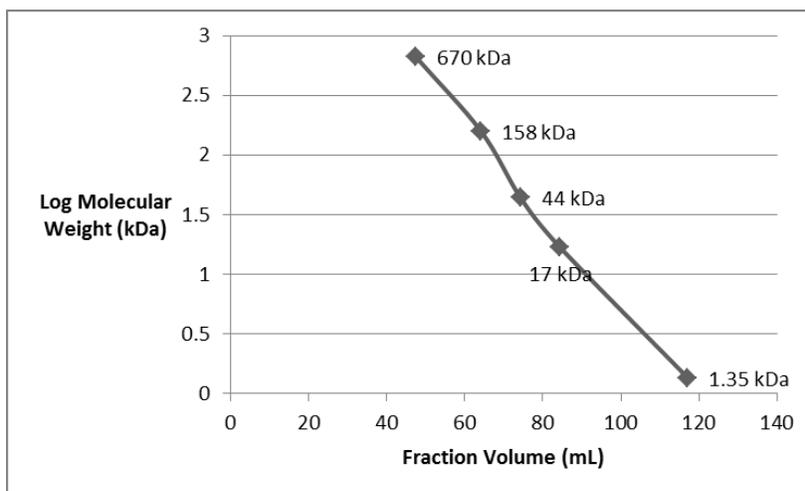


Figure 3.19: FPLC gel filtration analysis of standard molecular weight markers using the Sephacryl S300 column. (A) Molecular weight standards were size separated on a Sephacryl S300 with a fractionation range of 10,000 to 1,500,000 Daltons. Protein standards included thyroglobulin (670 kDa),  $\gamma$ -globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa) and vitamin B12 (1.35 kDa). Absorbance at 280 nm is shown by the blue line. Red streaks represent the 250  $\mu$ L fractions collected throughout the run at the corresponding volume (mL shown in black on the x-axis). (B) Based on the FPLC results, a graph of the molecular weight size standards was generated as a standard curve for the determination of the apparent molecular weights of the native protein lysate extracted from *Synechocystis* sp. PCC 6803.

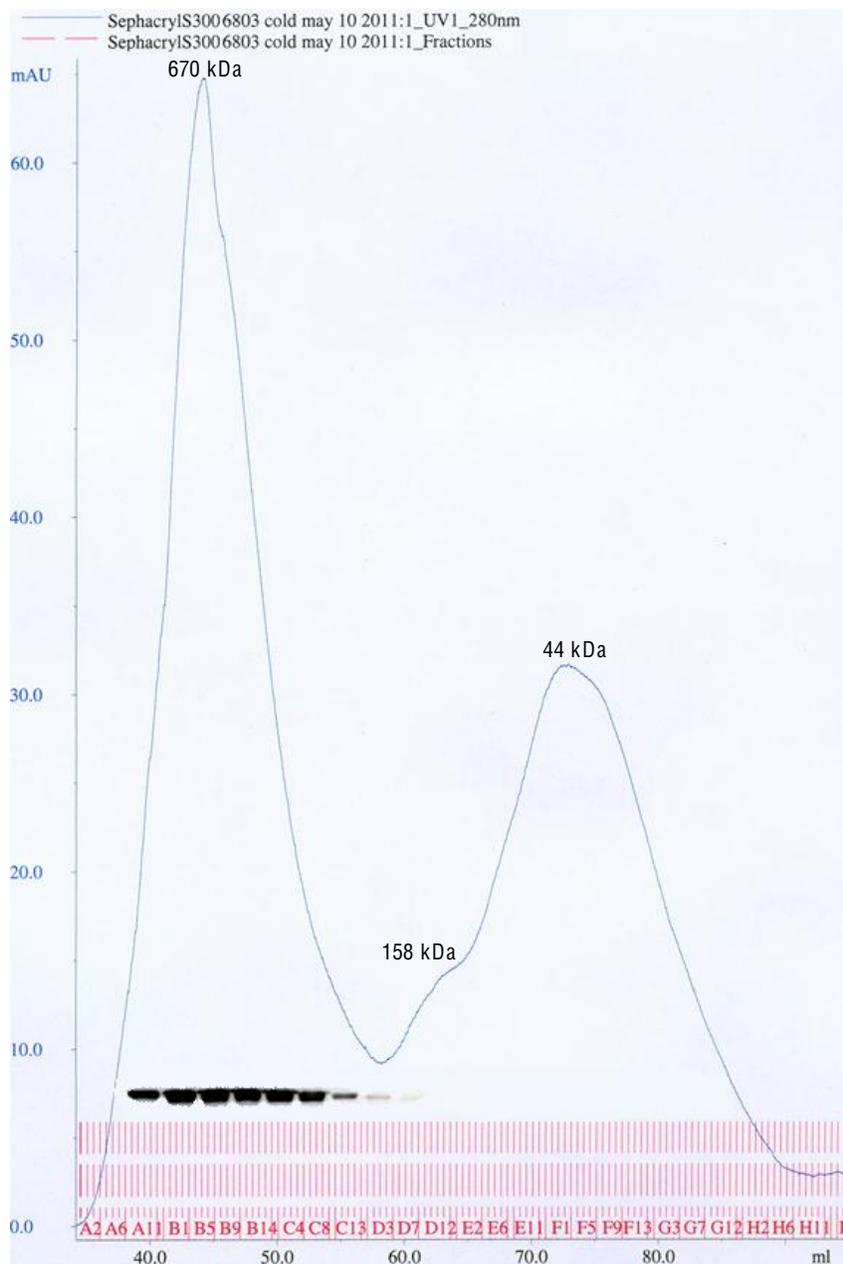


Figure 3.20: FPLC gel filtration analysis of native cold-shocked *Synechocystis* sp. PCC 6803 proteins using a Sephacryl S300 column. Native protein lysate was extracted from cold-shocked *Synechocystis* sp. grown at 20°C O/N, and fractionated on a Sephacryl S300 column. Five fractions each across the two peaks were combined, TCA precipitated, size fractionated on SDS-PAGE, and analyzed by western using anti-CrhR antibody. A standard curve (Figure 3.19B) was used to determine the approximate molecular weight of proteins in each peak (shown in kDa) and superimposed below the peaks on the column trace.

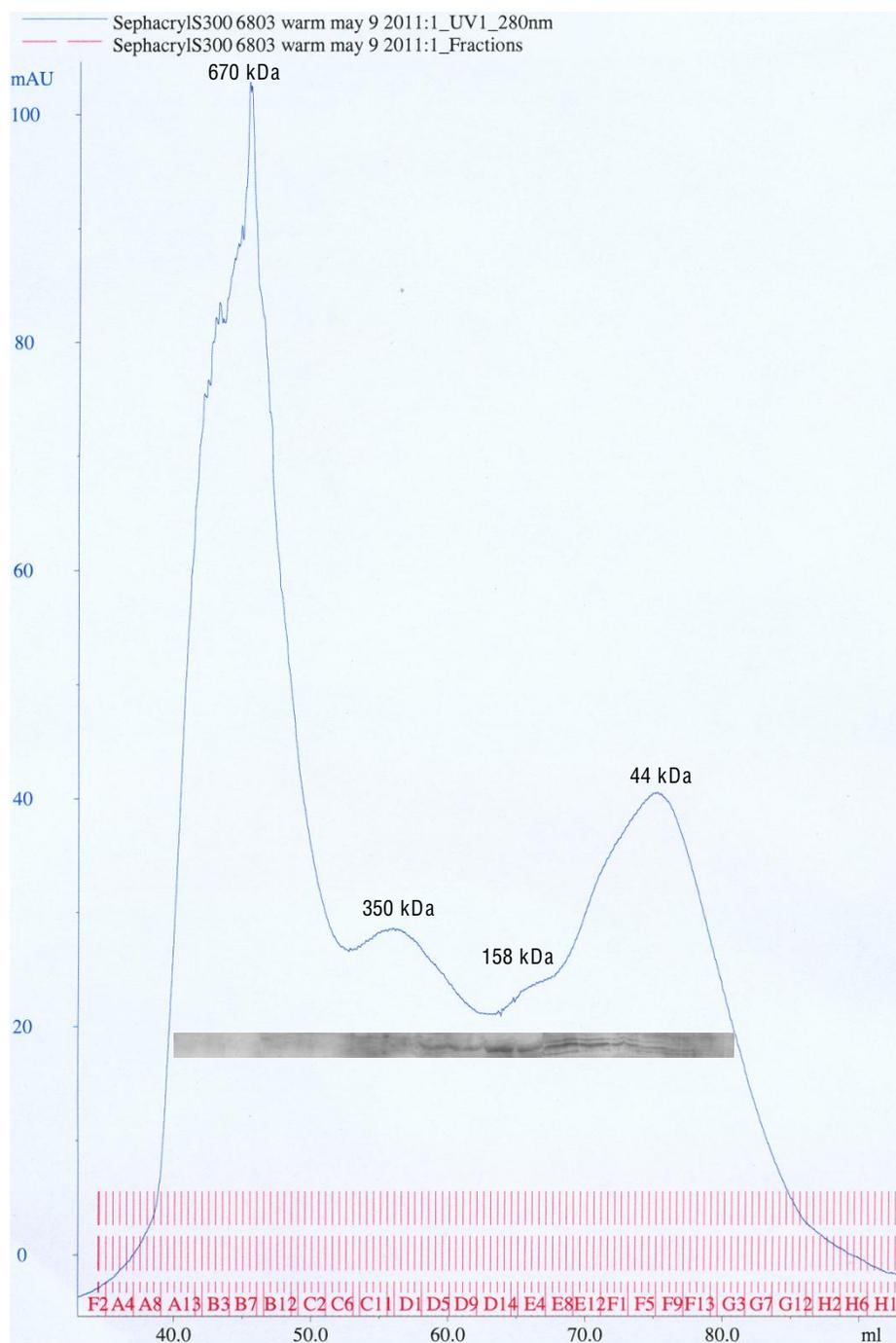


Figure 3.21: FPLC gel filtration analysis of native non cold-shocked *Synechocystis* sp. PCC 6803 proteins using a Sephacryl S300 column. Native protein lysate was extracted from *Synechocystis* sp. grown at 30°C, and separated on a Sephacryl S300 column. Five fractions each across the three peaks were combined, TCA precipitated, size fractionated on SDS-PAGE, and analyzed on a western using anti-CrhR antibody. A standard curve (Figure 3.19B) was used to determine the approximate molecular weight of proteins in each peak (shown in kDa) and superimposed below the peaks on the column trace. CrhR was detected by western analysis using anti-CrhR antibody. Untreated samples were in Figure 3.8.

warm samples resulted in the presence of 3 peaks that represent this range, shown in Figure 3.21. Western blot analysis was used to detect CrhR, by combining 5 fractions across the range of the observed peaks.

CrhR elution with the cold-shock treatment resulted in detection primarily in the higher molecular weight fractions as shown in Figure 3.20, present in the first peak. CrhR was not detected in the second peak. In Figure 3.21, CrhR from cells grown at 30°C eluted at a lower molecular weight corresponding to monomer/dimer forms. As expected, less CrhR was detected in cells grown at 30°C. These results were consistent and reproducible when repeated.

Therefore, elution profiles differ when CrhR is expressed at higher and lower temperatures; at 30°C CrhR eluted at a lower molecular weight corresponding to monomer/dimer forms, while at 20°C it eluted at higher molecular weight. It was shown in Figures 3.17 and 3.18 that CrhR was not aggregating as a result of RNA interaction, therefore indicating that CrhR is associating with a high molecular weight multi-subunit protein complex at lower temperatures.

### 3.3.3 Potential CrhR and ribosomal protein interaction

To investigate and identify the proteins co-eluting with CrhR at high molecular weight in FPLC at 20°C, mass spectrometry was applied to samples from gel filtration analysis. Samples submitted were from a native protein lysate from cold-shocked *Synechocystis* that had been separated on a Sephacryl S300 size exclusion column. Only samples corresponding to the higher multimer

molecular weight were submitted, and analyzed. The mass spectrometry IDs were obtained using the Sequest search from Uniprot KB (described in section 2.7), under the *Synechocystis* proteome FASTA index.

The majority of the hits from mass spectrometry corresponded to 50S and 30S ribosomal protein subunits from *Synechocystis*. The remnants, accounting to 20% of the hits, were other *Synechocystis* proteins including Photosystem I P700 chlorophyll *a* apoprotein A1, an SOS function regulatory protein, bicarbonate binding protein CmpA, and Phycobilisome linker-polypeptide.

CrhR was not detected in this mass spectrometry analysis. This was surprising as with the numerous FPLC gel filtration analysis that were performed, it was always detected in the high molecular weight range when the cells were cold-shocked. However, western analysis can detect smaller amounts of proteins than required for mass spectrometry; therefore, this could mean that there was not enough CrhR in the sample to be detected. Another reason could be due to the trypsin digest, which could have cleaved the protein at several sites, resulting in smaller polypeptides that were not able to be detected by mass spectrometry analysis, which could be determined by looking at the sequence. However, in order to distinguish between the two reasons, the mass spectrometry analysis would need to be repeated because it was attempted once.

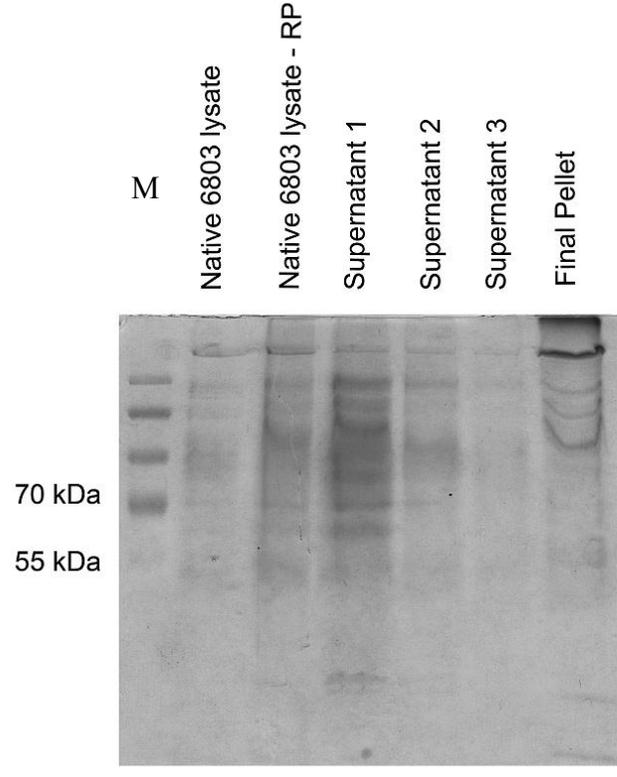
Although the majority of proteins detected in mass spectrometry analysis were the 50S and 30S ribosomal subunits, this only indicates CrhR and the ribosomal subunits co-elute in the same fractions in FPLC and do not necessarily associate with each other. In order to test whether CrhR associates with ribosomal

proteins, a ribosomal protein preparation was performed. As described in the Methods and Materials, native cyanobacterial *Synechocystis* cells were grown at 30°C, cold shocked at 20°C O/N and proteins extracted using ribosomal protein isolation buffer.

The sample was ultracentrifuged O/N to pellet the crude ribosomes, then layered onto a 10% sucrose gradient and ultracentrifuged O/N, and then finally layered onto a 30% sucrose gradient for a final ultracentrifugation O/N. As shown in Figure 3.22, aliquots of each supernatant and the final pellet were size fractionated on SDS-PAGE. Figure 3.22A shows a Coomassie stained gel, while Figure 3.22B shows a western blot. While it is difficult to identify CrhR on the Coomassie stained gel, the western blot analysis in Figure 3.22b shows CrhR detection in the pellet observed in the final crude ribosomal pellet. CrhR was not detected in the three supernatants collected after each spin, which is expected since it is present in the pellet with the ribosomal proteins.

Detection of CrhR was not successful in mass spectrometry analysis, perhaps due to the helicase not being present in a large enough amounts to be detected on mass spectrometry. Yet, it was detected in the ribosomal protein preparation, which indicates that CrhR maybe associating with a ribosomal protein complex, and maybe involved in ribosome biogenesis, especially of the 50S subunit, peptides of which are detected in mass spectrometry analysis of FPLC gel filtration fractions.

A



B

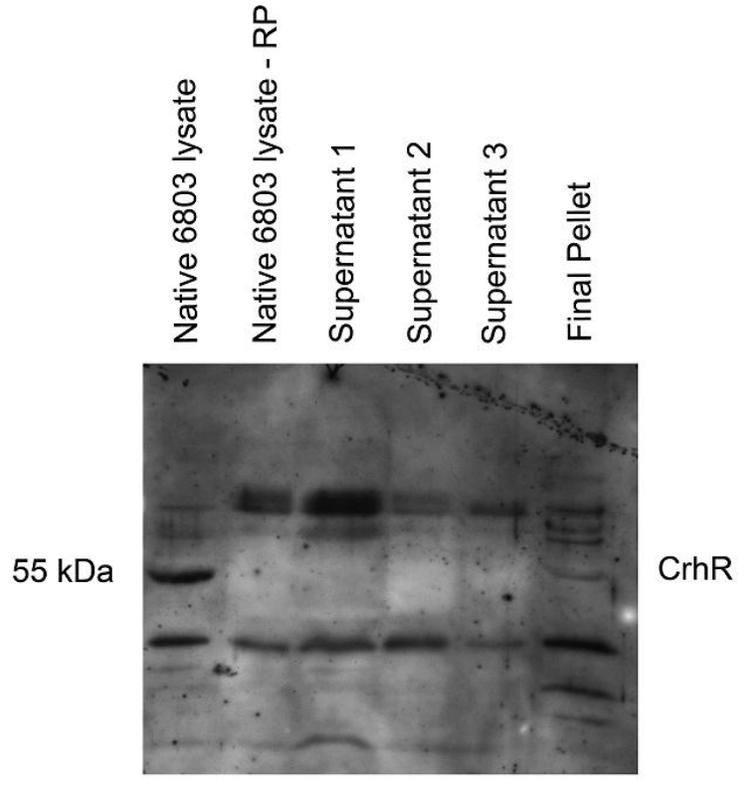


Figure 3.22: Coomassie stained and western blot analysis of ultracentrifuged ribosomal proteins extracted from native *Synechocystis* sp. PCC 6803. Native proteins were extracted from cold-shocked *Synechocystis* sp. using ribosomal extraction buffer, and ultracentrifuged to collect ribosomal proteins and any interacting partners. (A) Native protein lysate was size fractionated in lane 2 as a control, while the ribosomal extraction was fractionated in lane 3. Lanes 4, 5 and 6 represent the three supernatants from the three separate O/N ultracentrifugations, and lane 7 contains the final pellet from the last ultracentrifugation step. (B) In western analysis using anti-CrhR antibody, native protein lysate is in lane 1, and the ribosomal extraction in lane 2. Lanes 3, 4 and 5 represent the three supernatants, and lane 6 contains the final pellet. Other protein complexes reacting with the antibody were also observed in all the lanes.

#### IV. Discussion:

Structural data for several RNA helicases have become available in recent years, providing a much better insight into the structure–function relationship (Rocak and Linder, 2004). Structures for members of the SF2 families and groups, DEAD-box helicases in particular, have been reported and a broad view of the domain architecture is now possible (Fairman-Williams *et al.*, 2010). Apart from the highly conserved helicase core, DEAD-box proteins are surrounded by N- and C-terminal domains, which often exceed the helicase core in size. These extensions have frequently been proposed to contain domains for RNA binding, which in a few cases have led to the helicase exhibiting an additional biochemical activity like annealing and sister strand exchange. The human RNA helicase II/Gu is an example of a DEAD-box helicase that was observed to anneal ssRNA, and this activity was shown to be linked to the C-terminal domain (Valdez *et al.*, 1997).

Other studies have proposed N- and C-terminal extensions as sites for protein-protein interactions. This can include interactions with proteins in RNP complexes, but in a rare few cases for DEAD-box helicases, it has been connected to the helicase interacting with itself to form a dimer. This was observed in four DEAD-box proteins, which include the eukaryotic p68 and p72, which self-dimerize through the N-terminal domain (Ogilvie *et al.*, 2003), and the prokaryotic HerA and CshA, observed to self-dimerize in each case through the C-terminal domain (Klostermeier and Rudolph, 2008; Lehnik-Habrink *et al.*, 2010).

Previous analysis in the Owttrim lab on the cyanobacterial DEAD-box RNA helicase, CrhR, in the *Synechocystis* sp. strain PCC 6803, showed CrhR to efficiently catalyze strand exchange, and annealing, more processively than unwinding (Chamot *et al.*, 2005), activities also observed with both p68 and p72 (Rössler *et al.*, 2001). While the extent to which the biochemical features translating into a physiological function for CrhR, has not been elucidated, it is important to understand the structural features of the helicase and whether it is linked to certain biochemical activities, in order to elucidate certain functions (Jankowsky, 2011). In this study, the self-dimerization aspect of CrhR was investigated in order to further characterize the structure and function of CrhR.

The different experiments carried out in this study confirm CrhR dimerization, which was shown to occur in an RNA-independent manner. CrhR was also observed to elute with a multi-subunit complex in the cold, composed mostly of ribosomal subunits. These results suggest CrhR may potentially be associated with ribosomes in a macro-molecular complex, but its elution may also be an independent event. Due to these observations models for CrhR function have been proposed.

#### 4.1 CrhR RNA-Independent Dimerization *In Vitro* and *In Vivo*

Yeast two-hybrid interaction studies along with the protein exchange or swap analysis and FPLC gel filtration analysis indicated CrhR can interact with itself, which confirmed the results of CrhR dimerization observed previously in the lab. In the yeast two-hybrid screen, CrhR bound to an activation domain (AD)

and to the DNA binding domain (DBD) (EGR/JGR strain) resulted in blue colonies when plated on media containing X-Gal, as a result of protein-protein interaction (Table 3.1). The EGR/JGR strain also resulted in growth when plated on media lacking leucine. This result is also indicative of CrhR interaction, since the EGY48 parental strain transformed has the LexA promoter, which will express LEU2 if the proteins, CrhR in this case, attached to the AD and the DBD interact, and this will complement the leu auxotrophy of the strain.

Yeast two-hybrid controls validated CrhR self-interaction observations. The YKR/YBR positive control showed blue colonies on X-Gal containing media, and growth on plates lacking leucine. Negative controls with either an AD attached to a protein and a DBD on its own, or vice versa showed no interaction between the domains or proteins and domains, except in the case of the EG-END/JG-END constructs. These constructs containing the CrhR C-terminal extension only when combined with AD or DBD only appeared to give false-positives on X-Gal plates and plates lacking leucine, indicating the END constructs capable of an interaction signal on their own. Due to these unexpected results, CrhR dimerization was further investigated using protein exchange analysis.

Protein exchange analysis using all combinations of differently-tagged CrhR confirmed dimerization results observed in the yeast two-hybrid system (Figures 3.1 and 3.2). The controls indicated CrhR was not interacting with the tags non-specifically, and neither were the tags interacting with each other or non-specific columns, demonstrating CrhR self-interaction was specific between the

differently-tagged constructs. Additionally, Far-western analysis validated this result (Figure 3.4), although His-R bound to MBP-R was very faint on the western in comparison to His-R control. This can simply be explained by His-R being used as both the control and the probe, and the antibody was also against His-tag.

The tagged-CrhR protein exchange in non-denaturing conditions revealed a difference in size between a single tagged-CrhR (MBP-R) and two differently tagged-CrhR (reciprocal protein exchange or swaps MBP-R/His-R and His-R/MBP-R), when size fractionated using BN-PAGE (Figure 3.3). Since CrhR self-interaction was observed in the tagged-CrhR protein exchange, CrhR is expected to dimerize in solution, hence the presence of all three constructs at a higher molecular weight. However, in the single tagged-CrhR, a majority of the proteins is also observed at a lower molecular weight, which indicates some denaturation in solution of the aggregates not observed for the reciprocal swaps. This does not show a monomer and dimer, however BN-PAGE does show a difference in aggregation when using a single tagged-CrhR and two differently tagged-CrhR.

FPLC gel filtration analysis was another method used to confirm dimerization of CrhR, and some conditions under which CrhR might dimerize. Using both His-R (Figure 3.8) and native 6803 lysate (Figure 3.17), CrhR was eluted within the dimerization range at ~ 120 kDa. In order to determine whether CrhR dimerization was a result of interaction with RNA, CrhR dimerization was tested *in vitro* and *in vivo* in the presence of RNase A (Figure 3.15 and Figures

3.18, respectively). Regardless of whether RNA was present in the samples or not, dimerization was still observed in all samples.

All of the *in vitro* and *in vivo* analysis performed in this study involving His-R and native CrhR, respectively, has confirmed CrhR dimerization, which takes place in an RNA-independent manner. To date, there have only been four other DEAD-box RNA helicases that have been shown to dimerize; the highly related p68 and p72, with 90% homology (Ogilvie *et al.*, 2003), and prokaryotic HerA and CshA (Klostermeier and Rudolph, 2008; Lehnik-Habrink *et al.*, 2010). Both p68 and p72 interact with each other to form heterodimers through the C-terminal domains, but they have also been shown to self-associate and form homodimers through the N-terminal domain. It has been proposed that formation of homodimers of these proteins could possibly occur in a somewhat different way to that of heterodimers and have separate functions in the cell (Ogilvie *et al.*, 2003). Dimerization for HerA and CshA has been linked to the C-terminal domain (Klostermeier and Rudolph, 2008; Lehnik-Habrink *et al.*, 2010). From the results presented here, CrhR dimerization appears to resemble p68 and p72 dimerization, which occurs through the N-terminal domain. Although the C-terminal of CrhR could contribute, possibly through dimer stabilization, which can be facilitated through RNA binding to the C-terminal.

Domains required for dimerization were investigated using both N- and C-terminal truncations of CrhR. A truncation of the CrhR C-terminal extension in protein exchange analysis (Figure 3.12) as well as a full truncation of the C-terminal domain (domain 2 and the C-terminal extension) in yeast-two hybrid

analysis (Table 3.1) marginally affected dimerization of the protein. The N-terminal truncation in both the yeast two-hybrid (Table 3.1) and protein exchange analysis (Figure 3.14), on the other hand, significantly decreased CrhR self-interaction levels. This suggests that the N-terminal domain, rather than the C-terminal domain, is more important for dimerization. This result was unexpected, because the dimerization domain was thought to be linked to the C-terminal domain in our original hypothesis, mainly due to CrhR having a unique C-terminal extension. However, this result may imply that the C-terminal extension may involve an extra RNA binding site or domain that aids CrhR's helicase unwinding and/or annealing activity, as observed with eukaryotic RNA helicase II/Gu (Valdez *et al.*, 1997).

Protein-nucleic acid annealing seems to require more than one nucleic acid binding site within a given enzyme (Kurumizaka *et al.*, 1999), which could be provided by multiple subunits of an oligomeric complex. Furthermore, dimerization may contribute to the different biochemical activities exhibited by CrhR. For example, monomer and dimer forms may perform different functions. A monomer may catalyze unwinding, as observed for the majority of DEAD-box helicases tested, while the dimer may promote annealing. In addition, branch migration catalyzed by a helicase in complex with an annealing activity can mediate strand exchange, and in turn enhance the processivity of the helicase.

CrhR has been previously shown to efficiently catalyze strand exchange, annealing more processively than unwinding (Chamot *et al.*, 2005), as observed with both p68 and p72 (Rössler *et al.*, 2001). RNA helicases, DEAD-box in

general, are frequently observed to only unwind RNA duplexes over very short distances, and are assumed to be involved in processes in which unwinding of very short stretches of dsRNA is required (Linder and Jankowsky, 2011). This represents most monomeric helicases. Conversely, strand exchange as a role in transcription may require extensive duplex rearrangements and thus extended processivity, in this case requiring a dimer.

Unfortunately, the mechanistic manner of how p68 and p72 undergo these additional enzymatic activities has yet to be elucidated, but previous findings have alluded to dimer formation being required to achieve subtle alterations in p68/p72 function in the cell depending on how these proteins are interacting with their RNA substrates, and on the relative amounts of each present at a certain time (Ogilvie *et al.*, 2003). Helicase assays of the truncated C-terminal extension of CrhR have shown a significant decrease in both unwinding and annealing activities (Chamot and Owttrim, unpublished). This could allude to the possibility that the C-terminal extension is involved in the CrhR catalyzing these additional biochemical activities, which could also require a dimer formation in order to occur concomitantly. Therefore, it is possible to consider CrhR capable of functioning as a single subunit/monomer usually within larger complexes (Yang and Jankowsky, 2005) or a dimer, perhaps depending on which helicase activity it needs to achieve in the required rearrangement.

Two other DEAD-box RNA helicases, Ded1 (Yang and Jankowsky, 2005) and the human RNA helicase II/Gu (Valdez *et al.*, 1997), have been shown to anneal ssRNA in an ATP-independent manner, with RNA helicase II/Gu RNA-

folding activities being linked to the C-terminal domain (Valdez *et al.*, 1997). Both of these helicases are incapable of RNA strand exchange, and structurally have not been shown to confer self-interaction capabilities. These data indicate that a range of structure-function activities exist within the RNA helicase family outside of the traditional roles in translation or RNA degradation. This ability to unwind and anneal is thought to be pivotal for the function of DEAD-box proteins in facilitating RNA remodeling of large RNP complexes (Pyle, 2008; Linder and Jankowsky, 2011). In fact in many cases, the accessory N- and C-terminal domains are also considered sites for protein-protein interactions like those taking place in an RNP complex. Thus data from other helicases indicates CrhR may also function to catalyze RNP rearrangements in addition to traditional RNA duplex unwinding. This intriguing possibility remains to be investigated, as this may indicate CrhR dimerization is required to carry out RNA strand exchange through the concomitant action of both the unwinding and annealing activities within RNP complexes.

Ribosomes constitute the major RNP complex in cells. FPLC gel filtration was utilized to investigate both the oligomer state of CrhR from affinity purified CrhR but also native CrhR from total *Synechocystis* extracts. This FPLC analysis of CrhR was referred to as *in vitro* and *in vivo*, respectively. In both cases, CrhR eluted in higher molecular weight fractions, corresponding to multimer formation, which also appeared to be unaffected by the presence or absence of RNA in the samples. However, like the SF1 DNA helicases, to which they are structurally and phylogenetically similar, SF2 helicases function as monomers or dimers on

nucleic acids, and they have not been shown to form the hexameric rings that typify other families classified into SF3-SF6 helicases (Pyle, 2008). It is very rare for enzymes exhibiting NTP-dependent RNA helicase activity to form homo-oligomeric helicase toroids. In fact, to date there have only been two RNA helicases that have been characterized structurally as ring-shaped hexamers; P4 proteins, the packaging motors of dsRNA bacteriophages, and the transcription termination factor Rho (Rabhi *et al.*, 2010). Neither of these RNA helicases are DEAD-box proteins or even part of SF2 RNA helicase family. Since there has not been one example of a DEAD-box protein functioning as a hexamer, it is difficult to suggest that CrhR may be the first.

While this means FPLC analysis of CrhR in the higher multimer range is not a strong indication of this protein's capabilities to form a ring-shaped hexamer, it suggests instead that CrhR may be associating with a macromolecular complex. Mass spectrometry analysis of CrhR elution profiles in the high molecular weight range in gel filtration analysis revealed most of the *Synechocystis* proteins also eluting within this range consist of ribosomal proteins. And ribosomal extracts from *Synechocystis* also indicate CrhR association as observed in western analysis. Therefore, a more likely explanation would be a functional RNA-independent CrhR dimer that is interacting with a higher RNP complex, in which the C-terminal extension of CrhR may be involved in promoting this interaction.

In many instances, the highly conserved helicase core is responsible for helicase function, conferring ATP-hydrolysis and RNA-binding activities required for RNA duplex unwinding. The complexity of their diverse functions, however,

may be largely due to amino- and carboxy-terminal sequences or domains that some of these helicases possess. The *in vitro* and *in vivo* analyses, presented here, demonstrate that the cyanobacterial DEAD-box RNA helicase CrhR is capable of self-interaction in an RNA-independent manner, possibly through the N-terminal domain, and may be involved in an RNP complex. Results in this study point to a complex consisting mainly of ribosomal proteins, associating with CrhR at low temperatures.

#### 4.2 Potential Functions for CrhR DEAD-Box RNA Helicase

To date, there has been a remarkable range of ATP-dependent and ATP-independent biochemical activities reported for the DEAD-box protein family of SF2 helicases (Linder and Jankowsky, 2011). Several established crystal structures for DEAD-box proteins have also been reported regarding structural folding and organization of conserved domains. However, many of these studies lack structures of full length proteins with N- or C-terminal extensions included and thus folding. Potential protein-protein interaction information is lacking. In addition, in most cases, it is still not well understood how structural and biochemical features translate into physiological function, but knowledge of biochemical activities and structural organization of helicases has been essential for devising physical models of how DEAD-box proteins function (Del Campo *et al.*, 2009; Linder and Jankowsky, 2011).

CrhR is proposed to catalyze its helicase and annealing activity concurrently to promote RNA strand exchange most probably through a branch

migration mechanism, indicating that CrhR might catalyze a diverse range of RNA secondary structure rearrangements, and might rapidly and efficiently convert one RNA structure into another, possibly rearranging the structure of RNP complexes (Chamot *et al.*, 2005). RNA helicases are thought to function in large part by facilitating structural transitions of RNA as well as RNP complexes, whose transitions would otherwise be too slow to allow the complexes to form or function (Grohman *et al.*, 2007). There have been several cases in recent years showcasing DEAD-box helicases as having a broader role in remodeling RNA structures either by assisting in correct RNA folding or helping to dissociate RNA-protein interactions (Charollais *et al.*, 2004). CrhR could be involved in either or both, as results indicate its association in a macromolecular complex.

In this study, gel filtration analysis of native CrhR from *Synechocystis* revealed the detection of CrhR not only within the dimer range, but also at a high molecular weight corresponding to the formation of multimers (Figure 3.20). This was only observed in samples cold-shocked at 20°C, but not with samples grown at 30°C (Figure 3.21), and was not affected by the presence or absence of RNA in the sample. CrhR protein expression has been shown to greatly increase in a downshift in temperature (unpublished data). Therefore, perhaps the *in vivo* role performed by CrhR involves a functional dimer interacting in an RNP complex in order to alter RNA-RNA and RNA-protein configurations required for adaptation to environmental stress (Chamot *et al.*, 2005).

In this context, it can be argued that bacterial RNA helicases are necessary under certain growth conditions, which include low temperature, probably to

overcome enhanced stability of secondary structures under temperatures below the growth optimum (Rocak and Linder, 2004). A well-known example is in *E. coli*, which has been shown to induce the synthesis of the DEAD-box protein CsdA, specifically under cold-shock conditions (Awano *et al.*, 2010), which was reputed to be ribosome associated, with the 50S subunit (Charollais *et al.*, 2004). The elongated C-terminal domain in CsdA possesses an RNA binding motif that has been shown to be sufficient for its role in ribosome biogenesis (Dr. Jens Georg, University of Freiburg, Germany). RhIE is another example of a DEAD-box protein in *E. coli*, even though it has not been interpreted as a cold-shock protein, it has been shown to compensate for CsdA function at low temperatures (Awano *et al.*, 2010).

CrhR is closely related to both CsdA and RhIE DEAD-box proteins; however, it is not a homologue of either CsdA or RhIE. A CrhR, CsdA and RhIE phylogenetic analysis reveals that all three DEAD-box helicases form separate branches due to structural sequence homology (Figure 4.1) (Dr. Jens Georg, University of Freiburg, Germany). Cyanobacterial proteins from the different species analyzed were divided accordingly into each of the three branches of the cold-shock DEAD-box proteins based on structural similarities, mainly of the accessory C-terminal domains. The C-terminal extension of CrhR clearly separates CrhR into a different clade than both CsdA and RhIE. Interestingly, only proteins from cyanobacterial species group into the CrhR-related clade. This suggests that proteins in this clade might perform functions different from CsdA and RhIE in similar cellular processes.



Figure 4.1: Phylogenetic analysis of DEAD-box RNA helicases CrhR, CsdA and RhIE. Phylogenetic neighbor joining tree of 72 cyanobacterial, 5 *E. coli* and 3 *A. thaliana* DEAD-box RNA helicases. Three sub clusters are highlighted, the CsdA cluster (blue), the RhIE cluster (orange) and the CrhR cluster (green). Non-cyanobacterial helicases and CrhR are underlined and written in bold face type. Bootstrap values are given for important branch points (500 repetitions). The black stars in the CrhR cluster mark strains where the *crhR* gene is located directly 3' of *rimO*. Identical lower case letters of the same color mark different helicases which are present in the same strain (Generated by Dr. Jens Georg, University of Freiburg, Germany).

Four out of the five *E. coli* DEAD-box RNA helicases have been implicated in ribosome biogenesis. Prokaryotic ribosome biogenesis has been closely investigated in *E. coli*; a single rRNA precursor is synthesized and processed into the 23S, 16S and 5S rRNAs. In an ordered pathway, both the 23S and 5S rRNAs are assembled into the mature 50S ribosomal subunit, the larger subunit of the 70S ribosome of prokaryotes, while the 16S rRNA is assembled into the mature 30S ribosomal subunit, or the smaller subunit of the 70S ribosome (Srivastava and Schlessinger, 1990). Along with the CsdA DEAD-box protein, which has been implicated in the assembly of the mature 50S ribosomal subunit (Charollais *et al.*, 2003; Charollais *et al.*, 2004), RhlE has also been shown to be involved in ribosome assembly through an association or interaction with CsdA in order to modulate its function during ribosome maturation, by regulating the accumulation of immature ribosomal RNA or ribosome precursors (Jain, 2008). The role performed by helicases in ribosome maturation remains to be determined, but could involve rRNA maturation or folding/assembly of the ribosomal subunits in an RNP complex.

Several microarray studies have demonstrated that low temperature stress induces the expression of a large number of genes, which includes genes for ribosomal proteins (Prakash *et al.*, 2010). Mass spectrometry analysis of *Synechocystis* sp. PCC 6803 FPLC fractions in this study reveals the association of CrhR in the higher molecular weight range is possibly by interaction with the 50S ribosomal subunit when cold-stressed. The majority of proteins analyzed within this multimer complex are 50S ribosomal proteins, followed by 30S

ribosomal proteins. And while CrhR was not detected in this complex in mass spectrometry analysis, it has been successfully detected through western blot analysis of these FPLC fractions. This can be explained by the need for a greater concentration of proteins in mass spectrometry, as compared with western analysis. To investigate the potential for CrhR-ribosome association, density centrifugation was performed. Density centrifugation of ribosomal extracts does reveal the presence of CrhR among ribosomal proteins from *Synechocystis* (Figure 3.22). Although preliminary, these results provide initial evidence that CrhR may be ribosome associated in *Synechocystis*.

In eukaryotic systems, it has also been observed that the frog homologue of the nucleolar RNA helicase II/Gu (DDX21), which, like CrhR, is capable of annealing ssRNA substrates. RNA helicase II/Gu (DDX21) down-regulation results in the depletion of 18 and 28S ribosomal RNAs (rRNAs), suggesting its involvement in the processing of 18S rRNA and contribution to the stability of 28S rRNA in *Xenopus* (Yang *et al.*, 2003). Although it is not yet elucidated, RNA helicase II/Gu's unwinding and annealing activity may be required to carry out this cellular role. Some of the protein-catalyzed rRNA structural rearrangements have been suggested to proceed via a branch migration type of mechanism, in which portions of two competing secondary structures transiently coexists (Jalal *et al.*, 2007). While RNA helicase II/Gu has been shown to anneal ssRNA, it has not been shown to rearrange RNA structures through a branch migration mechanism.

Branch migration structures mimic Holliday junction structures, which are well-characterized intermediates in homologous recombination of DNA molecules (Singleton *et al.*, 2007; Rössler *et al.*, 2001). In *E. coli*, homologous DNA strand exchange during repair and maintenance of DNA is induced and processed by RecA. The RecA protein possesses both a DNA-dependent ATPase and DNA annealing activity essential for these processes. Structural analysis of DNA and RNA helicases have shown the catalytic core to have very similar if not identical folds to RecA, indicating these helicases to be structural homologues of RecA (Rössler *et al.*, 2001).

Recombinational activities have been shown to be inherent in the eukaryotic DEAD-box proteins p68 and p72 (Rössler *et al.*, 2001). Both p68 and p72 have been proposed to catalyze such RNA structural rearrangements via a branch migration mechanism within the pre-60S ribosomal subunit (Jalal *et al.*, 2007). Since mass spectrometry analysis alludes to CrhR interaction with a ribosomal complex, mainly 50S and 30S rRNA, and *Synechocystis* CrhR is detected in the ribosomal extracts, this could indicate a role for CrhR in ribosome biogenesis, perhaps in the assembly of the 50S or 30S rRNA subunits.

DEAD-box proteins usually function within complexes containing dozens or even hundreds of components, such as the nascent ribosome (Linder and Jankowsky, 2011). However, the ribosome is an energy consuming machine that does not appear to require NTP-hydrolysis for the basic chemical reactions it catalyzes (Staley and Guthrie, 1998). While this makes DEAD-box proteins a favorable candidate for RNA helicase activity during ribosome biogenesis, since

they carry out their helicase activities independent of ATP-hydrolysis, CrhR-catalyzed helicase activities do require ATP-hydrolysis (Chamot *et al.*, 2005). This is in contrast to all other DEAD-box proteins, including RNA helicase II/Gu (Valdez *et al.*, 1997) and both p68 and p72 (Rössler *et al.*, 2001).

As coordination among ATP-binding, strand separation, ATP-hydrolysis and phosphate or ADP release probably varies for different DEAD box proteins, strand separation may be coupled to ATP-hydrolysis in some cases and to ATP-binding in others (Linder and Jankowsky, 2011). Nevertheless, this could indicate CrhR involvement at multiple steps in a cellular process like ribosome biogenesis. This could also indicate CrhR involvement in more than one cellular process, such as translation initiation or RNA degradation, during acclimatization of *Synechocystis* cells to low temperatures.

Additionally, FPLC analysis of CrhR does display a shift from RNA association in the higher protein complex in cold-shocked cells (Figure 3.20), to just a self-association or dimer in cells grown at 30°C (Figure 3.21). This implies that CrhR functions in a multi-subunit complex at low temperature, a complex which does not form or to which CrhR does not associate at 30°C. Thus this may be an indication that CrhR performs different tasks at the two temperatures. For example, CrhR may participate in the assembly of ribosomes, most likely in the case of the 50S ribosomal subunit during acclimatization to lower temperatures. CrhR could also potentially function as a dimer in other processes involving RNA-RNA rearrangements at higher temperatures, and its requirement for both its unwinding and annealing activities can come into play in several processes.

If CrhR is one of the DEAD-box helicases with several possible roles in RNA metabolism, it is also important to understand whether it is like other DEAD-box oligomers, and is capable of functioning as either a monomer or a dimer. This could depend on the process and whether CrhR is associated with a higher protein complex or not, and if possible, how many dsRNA binding domains are present in the CrhR polypeptide. Since the C-terminal does not appear to be the main site of CrhR dimerization, perhaps it is required for RNA binding in order to carry out the additional helicase activities. It may also be required to stabilize the dimer formation of CrhR, since a basal level of CrhR interaction was observed with the N-terminal truncation. However, this could also suggest that the N-terminal and part of the helicase core are required for dimerization, and carrying out its biochemical activities, which may not be limited to performing the helicase activities for an RNP complex in a single cellular process. And since CrhR is the only DEAD-box RNA helicase in *Synechocystis*, perhaps its increased expression levels at low temperatures, as well as higher temperatures, is required for more than one cellular process, which may be temperature related.

Traditional roles for DEAD-box helicases include translation initiation and RNA degradation; however, several DEAD-box proteins have been found to interact or co-localize with the transcription machinery (Rocak and Linder, 2004). Along with the many elucidated cellular roles, DEAD-box proteins p68 and p72 also function as transcriptional co-activators for nuclear oestrogen receptor- $\alpha$  (ER $\alpha$ ) (Endoh *et al.*, 1999; Watanabe *et al.*, 2001). On its own, p68 has been

shown to also act as a potent co-activator of the tumour suppressor p53 (Bates *et al.*, 2005). In terms of co-activation of the nuclear ER $\alpha$ , the interacting region between p68/p72 and ER $\alpha$  includes part of the conserved helicase core (Endoh *et al.*, 1999). In another case, the accessory C-terminal domain in RNA helicase II/Gu was shown to interact with the N-terminal domain of c-Jun, a transcription factor that plays a pivotal role in stress response of mammalian cells, suggesting RNA helicase II/Gu is a co-factor for c-Jun-activated transcription (Westermarck *et al.*, 2002). Since CrhR expression is increased at low temperatures, a role for it in stress response in *Synechocystis* through transcription regulation as a co-activator could be conceivable.

The three DEAD-box RNA helicases discussed above are all capable of catalyzing both unwinding and annealing activities. In all three cases, RNA helicase enzymatic activity is not required for their function in transcriptional regulation (Endoh *et al.*, 1999; Watanabe *et al.*, 2001; Westermarck *et al.*, 2002; Rocak and Linder, 2004); however, RNA binding does appear to be important for ER $\alpha$  co-activation by p68 and p72 (Endoh *et al.*, 1999; Watanabe *et al.*, 2001). These activities are not the general rule and further work is required to elucidate how these roles are fulfilled and whether other DEAD-box proteins display similar activities (Linder and Jankowsky, 2011).

Another potential function for CrhR elucidated from the data provided in this thesis, may involve translation initiation. The helicase activity of many DEAD-box proteins has been shown to be required for translation initiation, a pivotal step in gene expression (Rocak and Linder, 2004; Linder and Jankowsky,

2011). The translation initiation factor eIF4A is the prototypical DEAD-box protein, first elucidated for ATP-dependent RNA helicase activity and RNA remodelling by an RNA helicase in eukaryotes. DEAD-box protein eIF4A seems to be targeted by a wide range of factors that regulate translation initiation (Linder and Jankowsky, 2011). It has been shown to unwind and rearrange RNA-duplex structures at the 5' end of the eukaryotic mRNA to prepare it for scanning by the small ribosomal subunit (Rocak and Linder, 2004). It also requires the accessory protein eIF-4B to be maximally active (Lu *et al.*, 1999). The C-terminal extension of CrhR may fulfill the role performed by eIF-4B, and allude to CrhR having multiple functions, perhaps in the more traditional roles including translation initiation and RNA degradation.

The function of another DEAD-box protein, Ded1, has been implicated in translation initiation through several independent genetic and biochemical experiments, although the extent of its involvement is not well known (Rocak and Linder, 2004). The previously mentioned CsdA also plays overlapping roles in translation initiation and RNA degradation. It has been shown to promote translation initiation of structured mRNAs (Lu *et al.*, 1999) and the translational activation of the *rpoS* mRNA at low temperatures (Resch *et al.*, 2010). CsdA has also been implicated in the stabilization and degradation of mRNAs (Jones *et al.*, 1996).

Many DEAD-box proteins are required for the degradation of RNA molecules, by assisting the degradosome, a complex of RNase E and polynucleotidophosphorylase (PNPase) that degrades RNA (Rocak and Linder,

2004). The degradosome of the Gram-positive bacterium *Bacillus subtilis* has recently been described, and the DEAD-box RNA helicase CshA, which is also a dimer, was identified as the main helicase interacting with components of this complex. An analysis of the contribution of the individual domains of CshA to RNA degradation revealed that the C-terminal domain is not only required for dimerization, but also for all interactions with components of the RNA degradosome (Lehnik-Habrink *et al.*, 2010). This could implicate a role for CrhR dimerization in RNA degradation in *Synechocystis*.

These examples of DEAD-box proteins mentioned so far have shown the accessory domains to be required in some cases for dimerization and in other cases for catalyzing the helicase activity or even both. But the majority of accessory domains seem to play a major role in protein-protein interaction in multi-protein complexes involved in major cellular processes that include ribosome biogenesis, transcription regulation, translation initiation and RNA degradation. While results in this study may suggest a major role for CrhR in ribosome biogenesis in *Synechocystis* due to the observation that CrhR co-elutes with 50S and 30S ribosomal subunits, it does not limit its role for interaction with another RNP complex.

Several models can be devised for CrhR dimerization and interaction with a multi-protein complex involved in transcription or translation. To date, the molecular targets of CrhR have not been elucidated; nonetheless, knockout studies have revealed the involvement of CrhR in transcript and protein level regulation, noticeably of the low-temperature inducibility of heat-shock genes *groESL* and

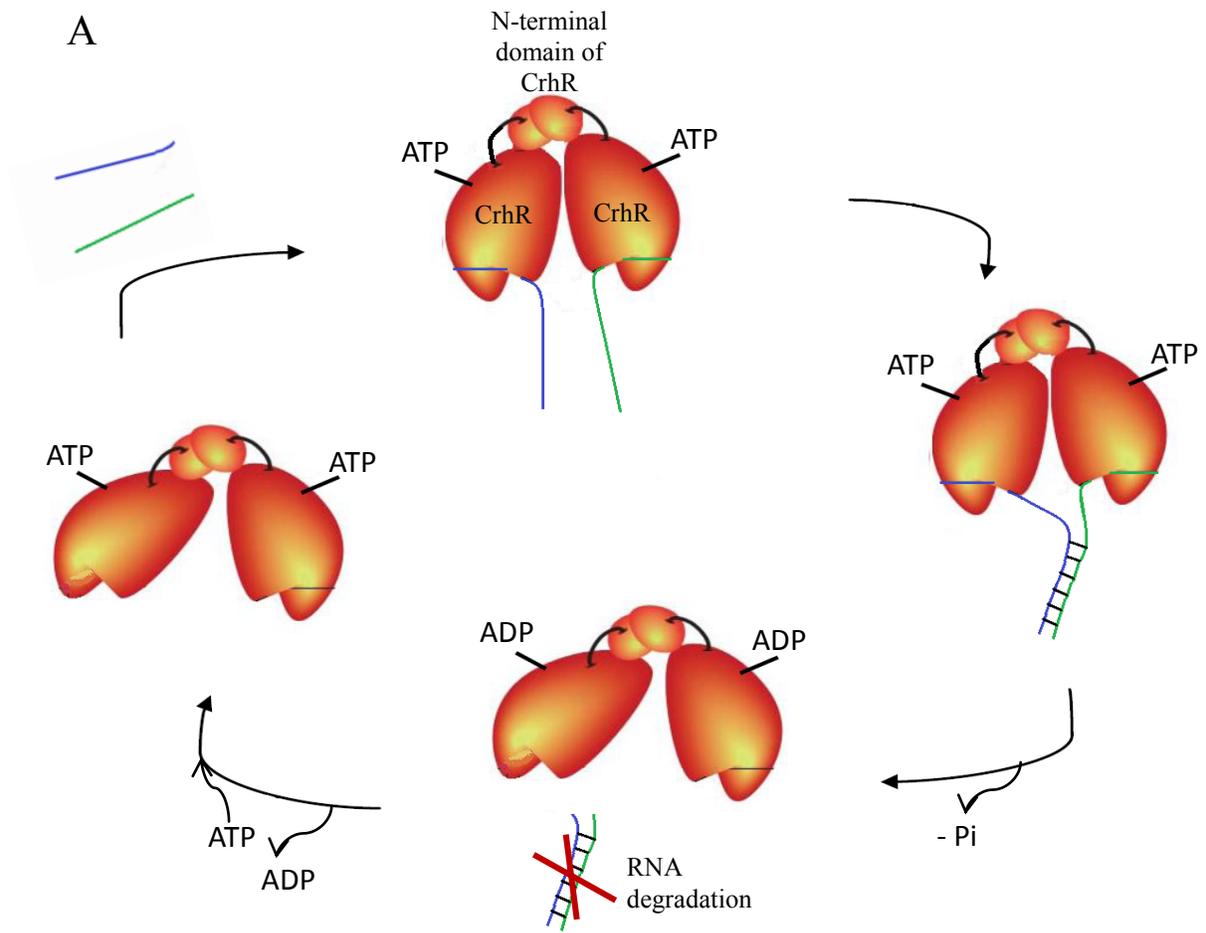
*groEL2* (Prakash *et al.*, 2010). This suggests there may be a series of the events associated with acclimation of *Synechocystis* cells to low temperatures and possible roles played by CrhR in these processes (Rowland *et al.*, 2011). Despite previous results of microarray analysis of CrhR, which does not show detection of either *groESL* or *groEL2* (George Owtrim, unpublished data), CrhR is the only DEAD-box protein in *Synechocystis*, therefore, this helicase may have to perform several different roles in *Synechocystis*.

In addition, microarray analysis did detect interaction of CrhR with small non-coding RNAs (sRNA) (George Owtrim, unpublished data). A common regulatory effect mediated by sRNAs is RNA degradation. Many sRNAs require the aid of an RNA chaperone, for example, the prokaryotic Hfq RNA chaperone. Hfq mediates interactions between the small sRNAs and specific mRNA targets, altering the stability of the target transcript (Lenz *et al.*, 2004). RNA chaperones are usually nonspecific, and a vast majority of DEAD box proteins lack inherent substrate specificity (Jankowsky, 2007; Linder and Jankowsky, 2011). By acting as an RNA clamp, the helicase can bind to both the sRNA and the target mRNA, and bring both strands together in order to anneal. Although opposite to a canonical translocating helicase, many DEAD-box proteins may be able to function as RNA chaperones, linking sRNA metabolism with regulation of mRNA transcript levels. This exciting possibility needs to be further examined in the future.

CrhR's ability to promote RNA unwinding, annealing and strand exchange, along with its interaction with sRNAs, make it an excellent candidate

for functioning as an RNA chaperone. The question that remains is why CrhR would need to dimerize. The vast majority of RNA helicase are monomeric, and this could be because they are required to interact with a single RNA in order to unwind it. Therefore in the case of CrhR, it could be required to dimerize in order to interact with two different RNAs so that it can catalyze annealing. This model is proposed in Figure 4.2A. The dimer, which is due to the interaction of the N-terminal domains of both CrhR monomer subunit, binds to the RNAs, in this case a target mRNA and sRNA through the C-terminal extensions in each monomer, and results in annealing of the two RNAs in an ATP-dependent manner. Since DEAD-box proteins have been shown to require ATP-hydrolysis only to release the products and not to catalyze their enzymatic activities, ATP-hydrolysis would result in the release of the newly formed dsRNA, which would most likely be a target for RNA degradation.

On the other hand, CrhR dimerization could also be required to catalyze both unwinding and annealing. RNAs often misfold into inactive structures, which can persist for a long time. RNA chaperone proteins give misfolded RNAs the opportunity to refold correctly (Owtrim, 2006; Jankowsky, 2007). If the CrhR dimer is acting as an RNA chaperone, it could be involved in clamping the two different RNAs. All RNAs fold into complex secondary structures which would have to be unwound before annealing could proceed. In the model in figure 4.2B, CrhR may be required to unwind both RNA substrates before it anneals them into an intermolecular RNA duplex, in a reaction similar to RNA strand exchange.



**B**

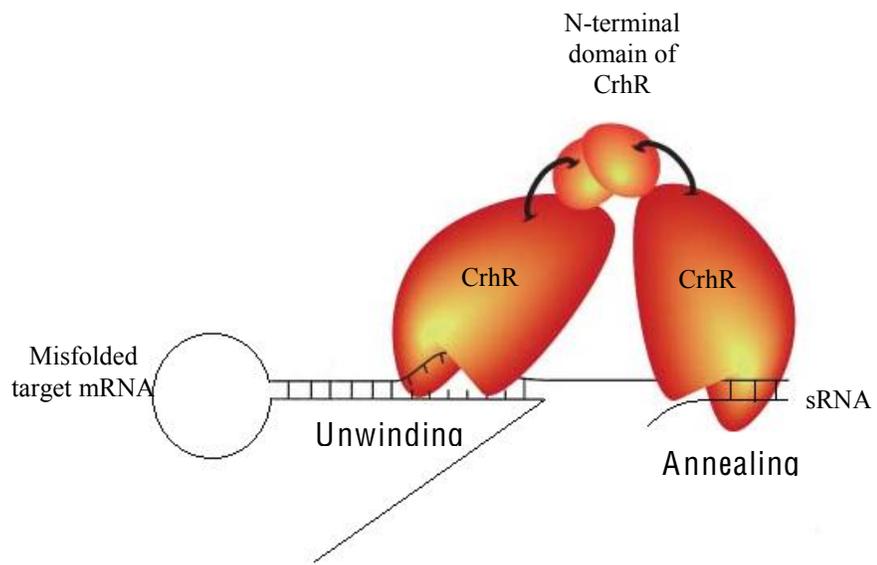


Figure 4.2: A model for CrhR dimerization. (A) ATP-bound dimeric CrhR is in an open conformation (Adapted from Jarmoskaite and Russell, 2010). Each CrhR monomer binds to either the sRNA (indicated by the green RNA strand), or the target mRNA (indicated by the blue RNA strand) through the C-terminal extensions of each monomer subunit in the dimer. The dimer changes to the closed conformation when RNA and ATP-bound resulting in annealing of the two RNAs. ATP-hydrolysis of the CrhR dimer leads to the release of the newly formed dsRNA and the recycling of the components. The newly formed dsRNA is than most likely targeted for degradation. (B) If the CrhR dimer is required for catalyzing both unwinding and annealing activities concomitantly, than one monomer can be involved in unwinding of the misfolded target mRNA while the other is involved in annealing of the sRNA. Since CrhR dimerization appears to be occurring through the N-terminal domain, the extended C-terminal domain is proposed to carry out the enzymatic activity in both models.

Therefore, the same cycle of ATP-driven changes in RNA affinity can give rise to both a high-affinity RNA clamp and duplex unwinding and annealing activities. This could also suggest the need for CrhR processivity to translocate a long RNA strand, as opposed to the short duplexes most DEAD-box helicases unwind. Thus, CrhR might traverse a clamp-like state of high RNA affinity during unwinding and/or annealing, before the products of ATP-hydrolysis dissociate and RNA is released (Linder and Jankowsky, 2011).

Available evidence suggests that the most critical DEAD-box protein region for RNA chaperone activity is the helicase core, which binds RNAs non-specifically and catalyzes the helicase activity, rather than translocation. Many DEAD-box proteins, including Ded1 and SrmB that have been shown to act as RNA chaperones, have accessory N- or C-terminal domains that are linked to the helicase core and may enhance or modulate its RNA chaperone function (Del Campo *et al.*, 2009). These accessory N- and C-terminal domains may not be required for ATPase or enzyme activity per se, but could provide a second non-specific RNA binding site that likely enhances RNA chaperone activity by tethering the helicase core to large RNA substrates (Del Campo *et al.*, 2009; Linder and Jankowsky, 2011).

CrhR dimerization, which appears to occur through the N-terminal, could be required in order to clamp both the sRNA and its target. And the accessory C-terminal domain in CrhR may contain a non-specific RNA binding site that can modulate this RNA chaperone or clamping, and allow it to catalyze its enzymatic activities; for instance, unwinding the misfolded target RNA to allow for the

annealing of the sRNA (Figure 4.2b). This role for CrhR would be equivalent to the Hfq RNA chaperone, which functions as a hexamer. Hfq mediates interactions between sRNA and target mRNA, which then typically alters the stability of the target transcript (Lenz *et al.*, 2004).

Despite the findings and conclusions made in trying to understand the structure, function and biological roles for CrhR in *Synechocystis*, there are still many unanswered questions that remain. Experiments in this thesis confirmed CrhR dimerization, most likely through the N-terminal domain, and its involvement in a multi-protein complex with preliminary evidence presented for potential ribosome association during acclimatization to low temperatures. However, a role for the extended C-terminal domain has not yet been elucidated, although helicase assays have alluded to the C-terminal extension being required for helicase activity, as the truncation shows decreased levels of both annealing and unwinding (Chamot and Owtrim, unpublished). Also, it is still not clear whether CrhR can function as a monomer and a dimer, and whether the dimer is involved in interacting with the multi-protein complex.

When it becomes possible to determine physiological RNA-binding sites, for CrhR and many other DEAD-box RNA helicases, it will be important to define whether and how RNA structure is changed by DEAD-box proteins, what roles ATP has in these reactions and when the helicase requires its enzymatic activity (Linder and Jankowsky, 2011). However, the great progress in studies on other DEAD-box RNA helicases, has led to the development of several models

which provide excellent ground work for understanding structure and function of CrhR.

## V. References:

- Amaratunga, M., and Lohman, T.M. 1993. *Escherichia coli* rep helicase unwinds DNA by an active mechanism. *Biochemistry* 32:6815–6820
- Andersen, C.B., Ballut, L., Johansen, J.S., Chamieh, H., Nielsen, K.H., Oliveira, C.L., Pedersen, J.S., Seraphin, B., Le Hir, H., and Andersen, G.R. 2006. Structure of the exon junction complex with a trapped DEAD-box ATPase bound to RNA. *Sci.* 313:1968-1972
- Arthur, H.M., and Lloyd, R.G. 1980. Hyper-recombination in *uvrD* mutants in *Escherichia coli* K-12. *Mol. Gen. Genet.* 180:185-191
- Awano, N., Rajagopal, V., Arbing, M., Patel, S., Hunt, J., Inouye, M., and Phadetare, S. 2010. *Escherichia coli* RNase R has dual activities, helicase and RNase. *J. Bact.* 192:1344-1352
- Banroques, J., Cordin, O., Doère, M., Linder, P., and Tanner, N.K. 2008. A conserved phenylalanine of motif IV in superfamily 2 helicases is required for cooperative, ATP-dependent binding of RNA substrates in DEAD-box proteins. *Mol. Cell Biol.* 28:3359-3371
- Bates, G.J., Nicol, S.M., Wilson, B.J., Jacobs, A.M.F., Bourdon, J.C., Wardrop, J., Gregory, D.J., Lane, D.P., Perkins, N.D., and Fuller-Pace, F.V. 2005. The DEAD box protein p68: a novel transcriptional co-activator of the p53 tumour suppressor. *EMBO J.* 24:543-553
- Bhaya, D., Bianco, N.R., Bryant, D., and Grossman, A. 2000. Type IV pilus biogenesis and motility in the cyanobacterium *Synechocystis* sp. PCC6803. *Mol. Microbiol.* 37:941-951
- Bird, L.E., Subramanya, H.S., and Wigley, D.B. 1998. Helicases: a unifying structural theme? *Curr. Opin. Struct. Biol.* 8:14-18
- Bizebard, T., Ferlenghi, I., Iost, I., and Dreyfus, M. 2004. Studies on three *E. coli* DEAD-box helicases point to an unwinding mechanism different from that of model DNA helicases. *Biochem.* 43:7857-7866
- Bleichert, F., and Baserga, S.J. 2007. The long unwinding road of RNA helicases. *Mol. Cell* 27:339-352
- Brennan, C.A., Dombroski, A.J., and Platt, T. 1987. Transcription termination factor Rho is an RNA-DNA helicase. *Cell* 48:945-952
- Büttner, K., Nehring, S., and Hopfner, K.P. 2007. Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat. Struc. Mol. Biol.* 14:647-652

- Byrd, A.K., Raney, K.D. 2004. Protein displacement by an assembly of helicase molecules aligned along single-stranded DNA. *Nat. Struct. Mol. Biol.* 11:531-538.
- Caruthers, J.M., and McKay, D.B. 2002. Helicase structure and mechanism. *Curr. Opin. Struct. Biol.* 12:123-133
- Castenholz, R. 1988. Culturing methods for cyanobacteria. *Meth. Enz.* 167:68-93
- Chamot, D., Magee, W.C., Yu, E., and Owttrim, G.W. 1999. A cold-shock induced cyanobacterial RNA helicase. *J. Bact.* 181:1728-1732
- Chamot, D., Colvin, K.R., Kujot-Choy, S.L., and Owttrim, G.W. 2005. RNA structural rearrangement via unwinding and annealing by the cyanobacterial RNA helicase, CrhR. *J. Biol. Chem.* 280:2036-2044
- Chao, K.L., and Lohman, T.M. 1991. DNA-induced dimerization of the *Escherichia coli* Rep helicase. *J. Mol. Biol.* 221:1165-1181
- Charollais, J., Pfeieger, D., Vinh, J., Dreyfus, M., and Iost, I. 2003. The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in *Escherichia coli*. *Mol. Microbiol.* 48:1253-1265
- Charollais, J., Dreyfus, M., and Iost, I. 2004. CsdA, a cold-shock RNA helicase from *Escherichia coli*, is involved in the biogenesis of 50S ribosomal subunit. *Nuc. Acids Res.* 32:2751-2759
- Chen, B., Doucleff, M., Wemmer, D.E., De Carlo, S., Huang, H.H., Nogales, E., Kondrashkina, E., Guo, L., and Nixon, B.T. 2007. ATP ground- and transition states of bacterial enhancer binding AAA<sup>+</sup> ATPases support complex formation with their target protein, sigma54. *Struc.* 15:429-440
- Chen, Y.F., Potratz, J.P., Tijerina, P., Del Campo, M., Lambowitz, A.M., and Russell, R. 2008. DEAD-box proteins can completely separate an RNA duplex using a single ATP. *Proc. Natl. Acad. Sci. U. S. A.* 105:20203-20208
- Chi, W., He, B., Mao, J., Li, Q., Ma, J., Ji, D., Zou, M., and Zhang, L. 2011. The function of RH22, A DEAD-box RNA helicase, in the biogenesis of the 50S ribosomal subunits of *Arabidopsis thaliana* chloroplasts. *Plant physiol. Prev.* 158:693-707
- Chu, W.K., and Hickson, I.D. 2009. RecQ helicases: multifunctional genome caretakers. *Nat. Rev. Cancer* 9:644-654
- Collins, R., Karlberg, T., Lehtio, L., Schutz, P., van der Berg, S., Dahlgren, L.G., Hammarstrom, M., Weigelt, J., and Schuler, H. 2009. The DExD/H-box RNA helicase DDX19 is regulated by an  $\alpha$ -helical switch. *J. Biol. Chem.* 284:10296-10300

- Cordin, O., Banroques, J., Tanner, N.K., and Linder, P. 2006. The DEAD-box protein family of RNA helicases. *Gene* 367:17-37
- Del Campo, M., Mohr, S., Jiang, Y., Jia, H., Jankowsky, E., and Lambowitz, A.M. 2009. Unwinding by local strand separation is critical for the function of DEAD-box proteins as RNA chaperones. *J. Mol. Biol.* 389:674-693
- De La Cruz, E.M., Kressler, J.D., and Linder, P. 1999. Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *TIBS* 24:192-198
- Dillingham, M.S., Spies, M., and Kowalczykowski, S.C. 2003. RecBCD enzyme is a bipolar DNA helicase. *Nature* 423:893-897
- Dong, F., Weitzel, S.E., and von Hippel, P.H. 1996. Mechanisms of helicases. *Proc. Natl. Acad. Sci. U. S. A.* 93:14456-14461
- Dürr, H., Körner, J., Hänzelmann, P., Truglio, J.J., Croteau, D.L., Van Houten, B., and Kisker, C. 2005. X-ray structures of the *Sulfolobus solfataricus* SWI2/SNF2 ATPase core and its complex with DNA. *Cell* 121:363-373
- Dürr, H., Flaus, A., Owen-Hughes, T., Hopfner, K.P. 2006. Snf2 family ATPases and DExx box helicases: differences and unifying concepts from high-resolution crystal structures. *Nucleic Acids Res.* 34:4160-4167
- Endoh, H., Maruyama, K., Masuhiro, Y., Kobayashi, Y., Goto, M., Tai, H., Yanagisawa, J., Metzger, D., Hashimoto, S. and Kato, S. 1999. Purification and identification of p68 RNA helicase acting as a transcriptional coactivator specific for the activation function 1 of human estrogen receptor  $\alpha$ . *Mol. Cell. Biol.* 19:5363-5372
- Enemark, E.J., and Joshua-Tor, L. 2008. On helicases and other motor proteins. *Curr. Opin. Struct. Biol.* 18:243-257
- Erzberger, J.P., and Berger, J.M. 2006. Evolutionary relationships and structural mechanisms of AAA<sup>+</sup> proteins. *Annu. Rev. Biophys. Biomol. Struct.* 35:93-114
- Fairman-Williams, M.E., Guenther, U.P., and Jankowsky, E. 2010. SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.* 20:313-324
- Fan, L., Fuss, J.O., Cheng, Q.J., Arvai, A.S., Hammel, M., Roberts, V.A., Cooper, P.K., and Tainer, J.A. 2008. XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. *Cell* 133:789-800

- Fisher, A.J., Smith, C.A., Thoden, J.B., Smith, R., Sutoh, K., Holden, H.M., and Rayment, I. 1995. X-ray structures of the myosin motor domain of *Dictyostelium discoideum* complexed with MgADPBeF<sub>x</sub>, and MgADPAIF<sub>4</sub>. *Biochem.* 34:8960-8972
- Flores-Rozas, H., and Hurwitz, J. 1993. Characterization of a new RNA helicase from nuclear extracts of HeLa cells which translocates in the 5' to 3' direction. *J. Biol. Chem.* 268:21372-21383
- Fuller-Pace, F.V. 2006. DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Res.* 34:4206-4215
- Geiselman, J., Wang, Y., Seifried, S.E., and von Hippel, P.H. 1993. A physical model for the translocation and helicase activities of *Escherichia coli* transcription termination protein Rho. *Proc. Natl. Acad. Sci. USA* 90:7754-7758
- Geisy, R.M. 1964. A light and electron microscope study of interlamellar polyglucositle bodies of *Oscillatoria chaly-* bin. *Am. J. Bot.* 51:588-596.
- Geourjon, C., Orelle, C., Steinfels, E., Blanchet, C., Deléage, G., Di Pietro, A., and Jault, J.M. 2001. A common mechanism for ATP-hydrolysis in ABC transporters and helicase superfamilies. *TRENDS in Biochem. Sci.* 26:539-544
- Gorbalenya, A.E., and Koonin, E.V. 1993. Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* 3:419-429
- Grohman, J.K., Del Campo, M., Bhaskaran, H., Tijerina, P., Lambowitz, A.M., and Russell, R. 2007. Probing the mechanisms of DEAD-box proteins as general RNA chaperones: the c-terminal domain of CYT-19 mediates general recognition of RNA. *Biochemistry* 46: 3013-3022
- Guenther, U.P., Laggerbauer, B., Jablonka, S., Chari, A., Alzheimer, M., Ohmer, J., Plöttner, O., Gehring, N., Sickmann, A., von Au, K., Schuelke, M., and Fischer, U. 2009. IGHMBP2 is a ribosome-associated helicase inactive in the neuromuscular disorder distal SMA type 1 (DSMA1). *Hum. Mol. Genet.* 18:1288-1300
- Guil, S., Gattoni, R., Carrascal, M., Abian, J., Stevenin, J., and Bach-Elias, M. 2003. Roles of hnRNP A1, SR proteins, and p68 helicase in c-H-ras alternative splicing regulation. *Mol. Cell. Biol.* 23:2927-2941
- He, Y., Andersen, G.R., and Nielsen, K.H. 2010. Structural basis for the function of DEAH helicases. *EMBO Rep.* 11:180-186

Hickman, A.B., and Dyda, F. 2005. Binding and unwinding: SF3 viral helicases. *Curr. Opin. Struct. Biol.* 15:77-85

Hilbert, M., Karow, A.R., and Klostermeier, D. 2009. The mechanism of ATP-dependent RNA unwinding by DEAD-box proteins. *Biol. Chem.* 12:1237-1250

**Honig, A., Auboeuf, D., Parker, M.M., O'Malley, B.W., and Berget, S.M.** 2002. Regulation of alternative splicing by the ATP-dependent DEAD-box RNA helicase p72. *Mol. Cell. Biol.* 22:5698-5707

Ikeuchi, M., and Tabata, S. 2006. *Synechocystis* sp. PCC 6803: a useful tool in the study of genetics of cyanobacteria. *Photosyn. Res.* 70:73-83

Jain, C. 2008. The *E. coli* RhlE RNA helicase regulates the function of related RNA helicases during ribosome assembly. *RNA* 14:381-389

Jalal, C., Uhlmann-Schiffler, H., and Stahl, H. 2007. Redundant role of DEAD box proteins p68 (Ddx5) and p72/p82 (Ddx17) in ribosome biogenesis and cell proliferation. *Nuc. Acids Res.* 35:3590-3601

Jankowsky, E. 2007. Indifferent Chaperones. *Nature* 449:999-1000

Jankowsky, E. 2011. RNA helicases at work: binding and rearranging. *Trends Biochem. Sci.* 36:19-29

Jankowsky, E., and Fairman, M.E. 2007. RNA helicases—one fold for many functions. *Curr. Opin. Struct. Biol.* 17:316-24

Jankowsky, E., and Putnam, A. 2010. Duplex unwinding with DEAD-box proteins. *Meth. Mol. Biol.* 587:245-264

Jarmoskaite, I., and Russell, R. 2010. DEAD-box proteins as RNA helicases and chaperones. *WIREs RNA* 2:135-152

Jennings, T.A., Mackintosh, S.G., Harrison, M.K., Sikora, D., Sikora, B., Dave, B., Tackett, A.J., Cameron, C.E., and Raney, K.D. 2009. NS3 helicase from the hepatitis C virus can function as a monomer or oligomer depending on enzyme and substrate concentrations. *J. Biol. Chem.* 284:4806-4814

Jones, P.G., Mitta, M., Kim, Y., Jiang, W., and Inouye, M. 1996. Cold shock induces a major ribosomal-associated protein that unwinds double stranded RNA in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 93:76-80

Kainov, D.E., Tuma, R., and Mancini, E.J. 2006. Hexameric molecular motors: P4 packaging ATPase unravels the mechanism. *Cell. Mol. Life Sci.* 63:1095–1105

- Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirose, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M. and Tabata, S. 1996b. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* 3:109-136
- Kaneko, T., and Tabata, S. 1997. Complete Genome Structure of the Unicellular Cyanobacterium *Synechocystis* sp. PCC6803. *Plant Cell Physiol.* 38:1171-1176
- Kawaga, R., Montgomery, M., Graig, K., Leslie, A.G., and Walker, J.E. 2004. The structure of bovine F1-ATPase inhibited by ADP and beryllium fluoride. *EMBO J.* 23:2734-2744
- Klostermeier, D., and Rudolph, M.G. 2008. A novel dimerization domain in the c-terminal domain of the *Thermus thermophilus* DEAD-box helicase Hera confers substantial flexibility. *Nucl. Acids Res.* 37:421-430
- Kujat, S.L., and Owttrim, G.W. 2000. Redox-regulated RNA helicase Expression. *Plant Physiol.* 124:703-713
- Kurumizaka, H., Ikawa, S., Sarai, A., and Shibata, T. 1999. The mutant RecA proteins, RecAR243Q and RecAK245N, exhibit defective DNA binding in homologous pairing. *Arch. Biochem. Biophys.* 365:83-91
- Lapage, S.P., Sneath, P.H.A., Lessel, E.F., Skerman, V.B.D., Seeliger, H.P.R., and Clark, W.A. 1975. International code of nomenclature of bacteria. *American Society for Microbiology, Washington, D.C.*
- Lee, J.Y., and Yang, W. 2006. UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell* 127:1349-1360
- Lehnik-Habrink, M., Pförtner, H., Rempeters, L., Pietack, N., Herzberg, C., and Stülke, J. 2010. The RNA degradosome in *Bacillus subtilis*: identification of CshA as the major RNA helicase in the multiprotein complex. *Mol. Microbiol.* 77:958-971
- Lenz, D.H., Mok, K.C., Lilley, B.N., Kulkarni, R.V., Wingreen, N.S., Bassler, B.L. 2004. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* 118:69-82
- Levin, M.K., Gurjar, M., and Patel, S.S. 2005. A Brownian motor mechanism of translocation and strand separation by hepatitis C virus helicase. *Nat. Struct. Mol. Biol.* 12:429-435

- Li, J., Li, W.X., and Gelbart, W.M. 2005. A genetic screen for maternal effect suppressors of decapentaplegic identifies the eukaryotic translation initiation factor 4A in *Drosophila*. *Genetics* 171:1629-1641
- Li, J., and Li, W.X. 2006. A novel function of *Drosophila* eIF4A as a negative regulator of Dpp/BMP signalling that mediates SMAD degradation. *Nat. Cell Biol.* 8:1407-1414
- Linder, P., Lasko, P.F., Ashburner, M., Leroy, P., Nielsen, P.J., Nishi, K., Schnier, J., and Slonimski, P.P. 1989. Birth of the D-E-A-D box. *Nature* 337:121-122
- Linder, P. 2003. Yeast RNA helicases of the DEAD-box family involved in translation initiation. *Biol. Cell* 95: 157-167
- Linder, P. 2006. DEAD-box proteins: a family affair-active and passive players in RNP remodelling. *Nucl. Acids Res.* 34:4168-4180
- Linder, P., and Jankowsky, E. 2011. From unwinding to clamping - the DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol.* 12:505-516
- Linder, P., and Lasko, P.F. 2006. Bent out of shape: RNA unwinding by the DEAD-box helicase vasa. *Cell* 125:219-221
- Liu, F., Putnam, A., and Jankowsky, E. 2008. ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. *Proc. Natl. Acad. Sci. USA* 105:20209-20214
- Liu, Z.R. 2002. p68 RNA helicase is an essential human splicing factor that acts at the U1 snRNA-50 splice site duplex. *Mol. Cell. Biol.* 22:5443-5450
- Lohman, T.M., and Bjornson, K.P. 1996. Mechanisms of helicase-catalyzed DNA unwinding. *Annu. Rev. Biochem.* 65:169-214
- Lüking, A., Stahl, U., and Schmidt, U. 1998. The Protein Family of RNA Helicases. *Crit. Rev. Biochem. Mol. Biol.* 33:259-296
- Lu, J., Aoki, H., and Ganoza, M.C. 1999. Molecular characterization of a prokaryotic translation factor homologous to the eukaryotic initiation factor eIF4A. *Int. J. Biochem. Cell Biol.* 31:215-229
- Lupas, A.N., and Martin, J. 2002. AAA proteins. *Curr. Opin. Struct. Biol.* 12:746-753
- Murray, H.L., and Jarrell, K.A. 1999. Flipping the switch to an active spliceosome. *Cell* 96:599-602

- Nakhaei, P., Genin, P., Civas, A., and Hiscott, J. 2009. RIG-I-like receptors: sensing and responding to RNA virus infection. *Sem. Immun.* 21:215-222
- Nicol, S.M., and Fuller-Pace, F.V. 1995. The “DEAD box” protein DbpA interacts specifically with the peptidyltransferase center in 23S rRNA. *Proc. Natl. Acad. Sci. USA* 5:11681-11685
- Ogilvie, V.C., Wilson, B.J., Nicol, S.M., Morrice, N.A., Saunders, L.R., Barber, G.N., and Fuller-Pace, F.V. 2003. The highly related DEAD-box RNA helicases p68 and p72 exist as heterodimers in cells. *Nucl. Acids Res.* 31:1470-1480
- Owtrim, G.W. 2006. RNA helicases and abiotic stress. *Nucl. Acids Res.* 34:3220-3230
- Paerl, H.W., Pinckney, J.L., and Steppe, T.F. 2002. Cyanobacterial–bacterial mat consortia: examining the functional unit of microbial survival and growth in extreme environments. *Env. Microbiol.* 2:11-26
- Pakrasi, H.B. 1995. Genetic analysis of the form and function of photosystem I and photosystem II. *Annu. Rev. Genet.* 29:755-776
- Pang, P.S., Jankowsky, E., Planet, P., and Pyle, A.M. 2002. The hepatitis C viral NS3 protein is a processive DNA helicase with co-factor enhanced RNA unwinding. *EMBO J.* 21:1168-1176
- Patel, S.S., and Domnez, I. 2006. Mechanisms of Helicases. *J. Biol. Chem.* 281:18265-18268
- Pause, A., Methot, N., and Sonenberg, N. 1993. The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP-hydrolysis. *Mol. Cell Biol.* 13:6789-6798
- Prakash, J.S.S., Krishna, P.S., Sirisha, K., Kanasaki, Y., Suzuki, i., Shivaji, S., and Murata, N. 2010. An RNA helicase, CrhR, regulates the low temperature-inducible expression of heat-shock genes groES, groEL1 and groEL2 in *Synechocystis* sp. PCC 6803. *Microbiol.* 156:442-451
- Pyle, A.M. 2008. Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu. Rev. Biophys.* 37:317-336
- Rabhi, M., Tuma, R., and Boudvillain, M. 2010. RNA remodelling by hexameric RNA helicases. *RNA Biology* 7:655-666

- Resch, A., Večerek, B., Palavra, K., and Bläsi, U. 2010. Requirement of the CsdA DEAD-box helicase for low temperature riboregulation of *rpoS* mRNA. *RNA Biology* 7:1-7
- Ritchie, D.B., Schellenberg, M.J., Gesner, E.M., Raithatha, S.A., Stuart, D.T., and MacMillan, A.M. 2008. Structural elucidation of a PRP8 core domain from the heart of the spliceosome. *Nature Struct. Mol. Biol.* 15:1199-1205
- Rocak, S., and Linder, P. 2004. DEAD-box proteins: the driving forces behind RNA metabolism. *Nat. Rev. Mol. Biol.* 5:232-241
- Rössler, O.G., Straka, A., and Stahl, H. 2001. Rearrangement of structured RNA via branch migration structures catalysed by the highly related DEAD-box proteins p68 and p72. *Nucl. Acids Res.* 29:2088-2096
- Rowland, J.G., Simon, W.J., Prakash, J.S.S., and Slabas, A.R. 2011. Proteomics reveals a role for the RNA helicase CrhR in the modulation of multiple metabolic pathways during cold acclimation of *Synechocystis* sp. PCC6803. *J. Prot. Res.* 10:3674-3689
- Runyon, G.T., Wong, I., and Lohman, T.M. 1993. Overexpression, purification, DNA binding, and dimerization of the *Escherichia coli* *UvrD* gene product (helicase II). *Biochemistry* 32:602-612
- Schägger, H., Cramer, W.A., and von Jagow, G. 1994. Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal. Biochem.* 217:220-230.
- Shively, J.M., Ball, F., Brown, D.H., and Saunders, R.E. 1973. Electron microscopy of the carboxysome (polyhedral bodies) of *Thiobacillus neopolitanus*. *Science* 82:584-586
- Sikora, B., Chen, Y., Lichti, C.F., Harrison, M.K., Jennings, T.A., Tang, Y., Tackett, A.J., Jordan, J.B., Sakon, J., Cameron, C.E., and Raney, K.D. 2008. Hepatitis C virus NS3 helicase forms oligomeric structures that exhibit optimal DNA unwinding activity *in Vitro*. *J. Biol. Chem.* 283:11516-11525
- Singleton, M.R., Scaife, S., and Wigley, D.B. 2001. Structural analysis of DNA replication fork reversal by RecG. *Cell* 107:79-89
- Singleton, M.R., and Wigley, D.B. 2002. Modularity and specialization in superfamily 1 and 2 helicases. *J. Bacteriol.* 184:1819-1826

- Singleton, M.R., Dillingham, M.S., and Wigley, D.B. 2007. Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* 76:23-50
- Srivastava, A.K., and Schlessinger, D. 1990. Mechanisms and regulations of ribosomal RNA processing. *Annu. Rev. Microbiol.* 44:105-129
- Staley, J.P., and Guthrie, C. 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* 92:315-326
- Stanier, R.Y., and Cohen-Bazire, G. 1977. Phototrophic prokaryotes: the cyanobacteria. *Ann. Rev. Microbiol.* 31:225-274
- Stanier, R.Y., Siström, W.R., Hansen, T.A., Whitton, B.A., Castenholz, R.W., Pfennig, N., Gorlenko, V.N., Kondratieva, E.N., Eimhjellen, K. E., Whittenbury, R., Gherna, R. L., and Truper, H. G. 1978. Proposal to place the nomenclature of the cyanobacteria (blue-green algae) under the rules of the international code of nomenclature of bacteria. *Int. J. Sys. Bacteriol.* 28:335-336
- Statlen, F. A. 1972. International Code of Botanical Nomenclature. *Utrecht, Netherlands: A. Oosthoek.*
- Sun, B., Wei, K.J., Zhang, B., Zhang, X.H., Dou, S.X., Li, M., Xi, X.G. 2008. Impediment of *E. coli* UvrD by DNA-destabilizing force reveals a strained-inchworm mechanism of DNA unwinding. *EMBO J.* 27:3279-3287
- Tackett, A.J., Chen, Y., Cameron, C.E., and Raney, K.D. 2005. Multiple full-length NS3 molecules are required for optimal unwinding of oligonucleotide DNA *in Vitro*. *J. Biol. Chem.* 280:10797-10806
- Tanner, N.K. 2003. The newly identified Q motif of DEAD box helicases is involved in adenine recognition. *Cell Cycle* 2:18-19
- Tanner, N.K., Cordin, O., Banroques, J., Doere, M., and Linder, P. 2003. The Q motif: a newly identified motif in the DEAD box helicases may regulate ATP binding and hydrolysis. *Mol. Cell* 11:127-138
- Tanner, N.K., and Linder, P. 2001. DExD/H box RNA helicases: review from generic motors to specific dissociation functions. *Mol. Cell* 8:251-262
- Taylor, S.D., Solem, A., Kawaoka, J., and Pyle, M. 2010. The NPH-II helicase displays efficient DNA-RNA helicase activity and a pronounced purine sequence bias. *J. Biol. Chem.* 285:11692-11703

Theissen, B., Karow, A.R., Kohler, J., Gubaev, A., and Klostermeier, D. 2008. Cooperative binding of ATP and RNA induces a closed conformation in a DEAD box RNA helicase. *Proc. Natl. Acad. Sci. USA* 105:548-553

Tye, B.K. 1999. MCM proteins in DNA replication. *Annu. Rev. Biochem.* 68:649–686

Valdez, B.C., Henning, D., Perumal, K. and Busch, H. 1997. RNA-unwinding and RNA-folding activities of RNA helicase II/Gu: two activities in separate domains of the same protein. *Eur. J. Biochem.* 250:800-807

Velankar, S.S., Soutanas, P., Dillingham, M.S., Subramanya, H.S., and Wigley, D.B. 1999. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* 97:75–84

Vinnemeier, J., and Hagemann, M. 1999. Identification of salt-regulated genes in the genome of the cyanobacterium *Synechocystis* sp. strain PCC 6803 by subtractive RNA hybridization. *Arch. Microbiol.* 172:377-386

Von Moeller, H., Basquin, C., and Conti, E. 2009. The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. *Nat. Struct. Mol. Biol.* 16:247-254

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1:945-951

Watanabe, M., Yanagisawa, J., Kitagawa, H., Takeyama, K.I., Ogawa, S., Arao, Y., Suzawa, M., Kobayashi, Y., Yano, Y., Yoshikawa, H., Masuhiro, Y., and Kato, S. 2001. A subfamily of RNA-binding DEAD-box proteins acts as an estrogen receptor  $\alpha$  coactivator through the N-terminal activation domain (AF-1) with an RNA coactivator, SRA. *EMBO J.* 20:1341-1352

Westermarck, J., Weiss, C., Saffrich, R., Kast, J., Musti, A.M., Wessely, M., Ansorge, W., Séraphin, B., Wilm, M., Valdez, B.C., and Bohmann, D. 2002. The DEXD/H-box RNA helicase RHII/Gu is a co-factor for c-Jun-activated transcription. *EMBO J.* 21:451-460

Wilson, B.J., Bates, G.J., Nicol, S.M., Gregory, D.J., Perkins, N.D. and Fuller-Pace, F.V. 2004. The p68 and p72 DEAD box RNA helicases interact with HDAC1 and repress transcription in a promoter-specific manner. *BMC Mol. Biol.* 5:11-18

- Wong, I., Chao, K.L., Bujalowski, W., and Lohman, T.M. 1992. DNA-induced dimerization of the *Escherichia coli* Rep helicase: allosteric effects of single stranded and duplex DNA. *J. Biol. Chem.* 267:7596-7610
- Wong, I., and Lohman, T.M. 1992. Allosteric effects of nucleotide cofactors on *Escherichia coli* Rep helicase and DNA binding. *Science* 256:350-355
- Wong, T.N., and Pan, T. 2009. RNA folding during transcription: protocols and studies. *Methods in Enz.* 468:167-193
- Xu, H.Q., Deprez, E., Zhang, A.H., Tauc, P., Ladjimi, M.M., Brochon, J.C., Auclair, C., and Xi, X.G. 2003. The *Escherichia coli* RecQ Helicase Functions as a Monomer. *J. Biol. Chem.* 278:34925-34933
- Yang, H., Zhou, J., Ochs, R.L., Henning, D., Jin, R., and Valdez, B.C. 2003. Down-regulation of RNA helicase II/Gu results in the depletion of 18 and 28 S rRNAs in *Xenopus* oocyte. *J. Biol. Chem.* 278: 28847-38859
- Yang, Q., and Jankowsky, E. 2005. ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. *Biochemistry* 44:13591-13601
- Yang, Q., and Jankowsky, E. 2005. The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. *Nat. Struct. Mol. Biol.* 13: 981-986
- Yarranton, G.T., and Gefter, M.L. 1979. Enzyme-catalyzed DNA unwinding: studies on *Escherichia coli* Rep protein. *Proc. Natl. Acad. Sci. USA* 76:1658-1662
- Ye, J., Osborne, A.R., Groll, M., Rapoport, T.A. 2004. RecA-like motor ATPases – lessons from structures. *Biochim. Biophys. Acta* 1659:1–18
- Yodh, J.G., Schlierf, M., Ha, T. 2010. Insight into helicase mechanism and function revealed through single molecule approaches. *Quart. Rev. Biophys.* 43:185-217
- Zhang, S., and Grosse, F. 2004. Multiple functions of nuclear DNA helicase II (RNA helicase A) in nucleic acid metabolism. *Acta. Biochem. Biophys. Sin.* 36:177-183
- Zervos, A.S., Gyuris, J., and Brent, R. 1993. Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites. *Cell* 72:223-232