INITIAL CHARACTERIZATION OF THE FUNCTIONAL CYCLE OF YIHA

DORA CAPATOS
Master of Science, **University of Alberta, 2013**

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

MASTER OF SCIENCE

in

BIOCHEMISTRY

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

INITIAL CHARACTERIZATION OF THE FUNCTIONAL CYCLE OF YIHA

DORA CAPATOS

Date of Defence: August 20, 2021

Dr. H.-J. Wieden Thesis Supervisor	Professor	Ph.D.
Dr. U. Kothe Thesis Examination Committee Member	Professor	Ph.D.
Dr. M. Roussel Thesis Examination Committee Member	Professor	Ph.D.
Dr. T. Patel Chair, Thesis Examination Committee	Associate Professor	Ph.D.

Abstract

The worldwide shortage of new and effective antibiotics is becoming of increasing concern. New antibiotic targets are needed for testing and validation in the development of a new generation of antibiotics. This thesis investigates a potential antibiotic target, the essential GTPase YihA. The objective of this thesis was to gain a greater understanding of the mechanisms that regulate YihA's functional cycle. Small angle X-ray scattering (SAXS) was used to obtain a low resolution solution structure of YihA and to provide information about the hydrodynamic characteristics of YihA including size, shape and flexibility. Further analysis of the SAXS data by disorder prediction programs and by modelling suggested that YihA in solution contains a folded core with flexible N-terminal and C-terminal tails. The SAXS structure of YihA, the SAXS-based modelling and the solution properties of YihA all supported dimerization of YihA in solution. Additionally, analytical gel filtration chromatography at concentrations of YihA closer to physiological concentrations revealed dimerization of YihA. The characteristics of the dimerization interface in the *E. coli* YihA crystal structure suggest that dimerization of YihA may play an important role in the functional cycle of YihA. Thus, these studies lay the groundwork for future structural and biochemical studies on the mechanism and significance of dimerization of YihA.

Preface

Sections of Chapter 1 were taken from a review titled “Understanding the Properties and Cellular Roles of the Small GTPase YihA” (Tentative title) on the characteristics and roles of YihA that was written by Dora Capatos and Dr. Hans-Joachim Wieden before the writing of this thesis and was intended to be submitted to the journal Biochemistry and Cellular Biology.

Chapter 2 contains SAXS data that was analyzed in three stages. In the first stage, Dr. Trushar Patel performed buffer subtraction of the peaks in the program Scatter, performed Guinier analysis and calculated the P(r) function for all datasets. In the second stage, Dora Capatos analyzed two of the data sets from the first stage of data analysis using the programs Dammin and Damaver to build an *ab initio* model of YihA. In the third stage, Dr. Trushar Patel used the resulting Damaver file from the first data set to prepare a SAXS model superimposed with the crystal structure of YihA from the first data set. Dora Capatos used the second data set to prepare a SAXS model superimposed with the crystal structure of YihA, ran all calculations in Crysol to compare experimental X-ray scattering to X-ray scattering from the crystal structure and prepared graphs of all data. The results from the second SAXS data set are reported in this thesis. All interpretation of the SAXS data was performed by Dora Capatos.

Acknowledgements

First, I would like to thank Dr. H.-J. Wieden for giving me the opportunity to work in his lab and to learn about ribosomes. I truly appreciate the feedback on my work and the time spent discussing the results of experiments.

I would like to thank Dr. Harland Brandon for his mentorship and support and for the many creative and unique ideas that he has given me for moving my project forward.

A big thank you to my Committee members Dr. Ute Kothe and Dr. Marc Roussel for their input into my project and for their helpful suggestions and guidance.

I would like to thank Dr. Trushar Patel for his contribution to analyzing the SAXS data (see page iv) and for teaching me how to analyze my SAXS data. I would like to thank Jalyce Heller for making a very detailed SAXS tutorial and for assistance with troubleshooting using the SAXS software.

I would like to extend a huge thank you to Fan Mo for help with designing many experiments as well as for help with ordering reagents.

Thank you to Luc Roberts for providing data for control experiments in this thesis.

Finally, I would like to extend thanks to all the other past and present members of the Wieden lab including Jessica, Taylor, Will, Sydnee, Amanda, Davinder, Fabian, Kristi, Emily, Dylan, Justin, Dustin, Binod, Katherine, Preethi, Rhys, Raja, Senthil, Jumai, Andy, Graeme, Soumya, Cristina and Adam, as well as all members of the Kothe, Patel and Demeler labs and L9Q1.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

A site	Aminoacyl tRNA binding site
ADP	Adenosine diphosphate
AGC	Automatic gain control
apo	Without a cofactor bound
BME	β-mercaptoethanol
BSA	Bovine serum albumin
Cryo-EM	Cryogenic electron microscopy
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EF	Elongation factor
E site	tRNA exit site
ETS	External transcribed spacer
FPLC	Fast protein liquid chromatography
FRET	Fluorescence resonance energy transfer
G-domain	GTPase domain
GAP	GTPase activating protein
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GTP	Guanosine triphosphate
GTPase	Guanosine triphosphatase
HAS GTPase	Hydrophobic amino acid substituted guanosine triphosphatase
HPLC	High performance liquid chromatography
IPTG	Isopropyl β-D-thiogalactopyranoside
ITS1	Internal transcribed spacer 1
ITS2	Internal transcribed spacer 2
KH domain	K homology domain
LB	Luria-Bertani
Mant	2'-/3'-O-N'-methylanthraniloyl
MWCO	Molecular weight cut off
NCBI	National Center for Biotechnology Information
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PDB ID	Protein data bank identification
PEP	Phosphoenolpyruvate
P_i	Inorganic phosphate
PK	Pyruvate kinase
PMSF	Phenylmethylsulfonylfluoride
P site	Peptidyl tRNA binding site
R_g	Radius of gyration
RMF	Ribosome modulation factor
rrn operon	ribosomal RNA operon
SASBDB ID	Small angle scattering biological data bank identification
SAXS	Small angle x-ray scattering
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
TAP	Tandem affinity purification
TCA	Trichloroacetic acid

CHAPTER 1: INTRODUCTION

1.1 THE BACTERIAL RIBOSOME

The mature bacterial ribosome is composed of a large 50S subunit bound to a smaller 30S subunit, forming the 70S ribosome [1]. The 50S ribosomal subunit consists of 33 ribosomal proteins and the 23S and 5S ribosomal RNA (rRNA) while the 30S ribosomal subunit consists of 21 ribosomal proteins and the 16S rRNA [2]. Three separate domains are found in 30S ribosomal subunits, including the 5' domain or body, the central domain, or platform and the 3' domain or head [3].

1.2 BACTERIAL RIBOSOME BIOGENESIS

Bacterial ribosome biogenesis takes place when ribosomal proteins bind cooperatively to the folding rRNA and interact with other ribosomal proteins, promoting conformational remodelling of the rRNA and forming the 30S and 50S ribosomal subunits [3]. Ribosome biogenesis occurs in tandem with transcription in bacteria [4]. During the exponential growth phase, an *Escherichia coli* (*E. coli*) cell dedicates almost 40% of its energy supply towards the process of ribosome biogenesis [5].

Biogenesis of both the small and the large ribosomal subunits in bacteria is characterized by parallel and sequential events [6-9].

1.2.1 rRNA MATURATION IN BACTERIA

Bacterial rRNA is transcribed within the cell as a single RNA transcript with an external transcribed spacer (5'-ETS) at the 5' end, followed by, in 5' to 3' order, the 16S rRNA, an internal transcribed spacer (ITS1) containing at least one tRNA, the 23S rRNA, a second internal transcribed spacer (ITS2) and then the 5S rRNA at the 3' end followed by the 3' external transcribed spacer (3'-ETS) (Figure 1.1) [3, 10]. Evidence from mutational analysis and structural probing suggests that ITS1 may form a temporary secondary structure that may play a role in facilitating the folding of the 23S rRNA and promoting ribosome assembly [11]. In bacteria, the rRNA genes are located in the rRNA (rrn) operons [12]. There are seven copies of the rrn operons in *E. coli*, designated A,B,C,D,E,H and G [11, 13]. In some rrn operons but not in rrn operons A,B,E and G in *E. coli*, one tRNA is found at a position downstream of the 5S rRNA (Figure 1.1) [10, 11]. The first processing reaction to take place for bacterial rRNA is cleavage of the 30S rRNA transcript by the endonuclease RNase III, releasing the precursors to the 5S rRNA, 23S rRNA and 16S rRNA [3, 14]. RNase III cleavage of the 30S rRNA transcript occurs simultaneously with transcription and thus the 30S rRNA transcript only accumulates to detectable levels when RNase III is not expressed [15]. Maturation of the precursor 16S rRNA, giving rise to the 17S rRNA, is necessary to remove the 115 extra nucleotides attached to the 5' terminus and the 33 extra nucleotides attached to the 3' terminus [3]. Processing of the 16S rRNA follows a specific order, beginning at the 5' end and subsequently proceeding to the 3' end [16]. The endonuclease RNase E is the first enzyme to cleave near the 5' terminus of the 17S rRNA, producing a precursor transcript with a 5' terminus containing an additional 66 nucleotides [3] The maturation of the 5' and 3' termini of the 23S rRNA is linked and the processing reactions at both termini proceed at the same rate [17]. However, the enzymes that catalyze the cleavage of the 5' ends of the 23S and 5S rRNA as well as the cleavage of the 3' end of the 16S rRNA remain unknown [17].

Surprisingly, the *E. coli* 50S ribosomal subunit containing immature 5' and 3' ends of the 23S rRNA remains fully functional and may be found in translating 70S ribosomes [18].

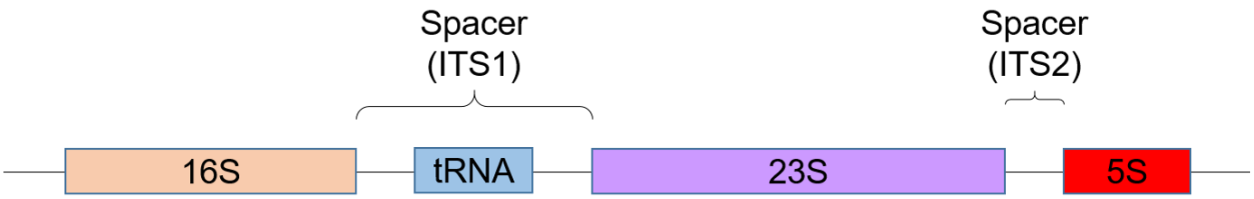


Figure 1.1. Locations of spacers for a bacterial rRNA transcript. The transcript shown is transcribed from the rrnB operon of *E. coli*. Full names for the abbreviations are given in the text.

1.2.2 rRNA MODIFICATIONS IN BACTERIA

There are 25 modifications identified to date in the 23S rRNA, including 9 pseudouridines, 1 methylated pseudouridine, 14 methylations and 1 unknown modification [3]. In contrast, the 16S rRNA contains only 11 modifications, including 10 methylations and 1 pseudouridine [3]. A general principle is that RNA modifications were identified largely on conserved nucleotides within segments of rRNA that are required for a particular function of the ribosome [3]. Emerging evidence suggests that RNA modifications regulate the conformation of rRNA and affect the role of rRNA in ribosome biogenesis or translation [3]. Moreover, 30S ribosomal subunits assembled *in vitro* from 30S ribosomal subunit proteins produced in the cell and from 16S rRNA that is produced by *in vitro* transcription and that lacks modifications are able to form 70S ribosomes capable of limited translation of a polyphenylalanine peptide and are capable of binding to tRNA [3]. Interestingly, *E. coli* 50S ribosomal subunits that are able to carry out protein synthesis are

not able to be assembled *in vitro* from 50S ribosomal subunit proteins produced in the cell and 23S rRNA transcribed *in vitro* [19]. In *E. coli* 23S rRNA, the segment from nucleotides 2445 to 2523 requires seven modifications in order for fully functional 50S ribosomal subunits to be formed *in vitro* [19]. By comparison, in *Bacillus stearothermophilus*, fully active 50S ribosomes are able to be assembled *in vitro* from 50S ribosomal subunit proteins produced in the cell and 23S rRNA transcribed *in vitro*, suggesting that the seven modifications of nucleotides 2445 to 2523 of the 23S rRNA required in *E. coli* are not universally essential among bacteria [20].

1.2.3 *IN VITRO* ASSEMBLY ORDER OF RIBOSOMAL PROTEINS IN BACTERIA

Previous studies have also demonstrated that sequential binding of ribosomal proteins at strategic stages during assembly can assist remodelling and folding of ribosomal RNA (rRNA) into the 3-dimensional structure adopted in translation-competent ribosomes [21]. *In vitro* experiments have shown that association of small subunit ribososomal proteins with 16S rRNA is hierarchical [3]. The primary binding proteins of the 30S ribosomal subunit are able to associate directly with the 16S rRNA during transcription and include uS4[1], uS7, uS8, uS15, uS17 and bS20 [3]. Secondary and tertiary binding proteins are not able to associate with assembling ribosomes if the required primary binding proteins and secondary binding proteins are not already bound to the pre-ribosomes [22]. According to the Nomura map of assembly of the 30S ribosomal subunit, secondary binding proteins are bS16, a complex of bS6 bound to bS18, uS9, uS13, and uS19 whereas the tertiary binding proteins are uS2, uS5, uS11, bS21, uS10, uS14, uS3 and uS2 [22].

[1] In the nomenclature system used here, the prefix 'u' (from universal) is attached to the *E. coli* names of ribosomal proteins with homologues from across all three domains of life. Ribosomal proteins that are found only in bacteria but not in archaea or eukaryotes are given the prefix 'b' (from bacterial) that is attached to their *E. coli* names. Similarly, ribosomal proteins found only in eukaryotes or in eukaryotes and archaea are given the prefix 'e' (from eukaryotic) and ribosomal proteins found only in archaea are given the prefix 'a' (from archaeal).

1.2.4 ASSEMBLY FACTORS FOR BACTERIAL RIBOSOMES

Current models of biogenesis of the bacterial ribosome lack detail regarding molecular interactions of assembly factors with the respective immature ribosomal particles [23]. Although principles of the 30S ribosome subunit assembly process are beginning to emerge, questions remain unanswered about why many factors are needed in the assembly process. These assembly factors include enzymes that assist in rRNA folding (acting as "placeholders"), modify RNA, or act as chaperones and helicases [24]. Both the large and small bacterial ribosomal subunits can assemble *in vitro* in the absence of assembly factors but the reactions proceed very slowly at or below room temperature [25-27].

Within the bacterial cell, assembly of a single ribosome can be completed in 2 to 3 minutes [24] whereas for *in vitro* experiments at 30^{o}C the 30S subunit takes up to 2 hours to assemble from ribosomal protein fraction and 16S rRNA [28]. This indicates that additional assembly factors such as chaperones and GTPases are needed to associate transiently with ribosomes for ribosome biogenesis to proceed efficiently.

In fact, ribosome biogenesis requires over 60 assembly factors in *E. coli* [3]. Ribosome assembly factors block ribosomal subunit joining and translation in ribosomal subunits that are not fully assembled and promote the correct assembly of subunits by acting as placeholders, molecular timers, sensors, chaperones and RNA modification enzymes. Interestingly, only a small number of ribosome assembly factors are essential for bacterial cell survival. Among the proteins that are required for survival of bacteria and that regulate ribosome biogenesis in bacteria are CgtA, EngA, Era, GroEL-GroES, and YihA [3].

1.3. INTRODUCTION TO GTPASES

GTPases are enzymes that hydrolyze guanosine triphosphate (GTP) to form guanosine diphosphate (GDP) [29]. Typically, GTPases such as Ras cycle between a functional, GTP-bound form that is able to bind to effectors in the cell, a non-functional GDP-bound form that dissociates from effectors and a short-lived form from which GDP has dissociated leaving the *apo* enzyme [30]. The catalytic mechanisms of GTPases are governed by the G-domain present in all GTPases [31]. Each G-domain includes highly conserved amino acids located in five G-motifs (G1-G5) within the active site [31]. Further details on each motif of the G-domain are provided in section 1.5.1. In many GTPases however, an amino acid adjacent to the G3 motif (glutamine in Ras) coordinates the water molecule for a nucleophilic attack on the γ-phosphate of GTP, resulting in the stabilization of the transition state [32, 33]. Upon GTP hydrolysis, a change in conformation of the GTPase occurs that promotes the release of effectors that are bound to the GTPase [32]. Small GTPases in eukaryotes, such as Ras, as well as many other GTPases both in prokaryotes and eukaryotes, often have a slow intrinsic rate of GTP hydrolysis and depend on a GTPase activating protein (GAP) to stimulate GTP hydrolysis [34]. GTPases may also utilize guanine nucleotide exchange factors (GEFs) to reduce affinity for and to promote dissociation of GDP and binding of GTP and guanine nucleotide dissociation inhibitors (GDIs) to increase the affinity for GDP [35, 36]. GAPs utilize a variety of ways to increase the rate of GTP hydrolysis, including by shielding the negative charge of the transition state, by positioning the catalytic residue of the GTPase so that it can coordinate the nucleophilic water, or by providing an amino acid that is positioned to coordinate the water molecule that can carry out a nucleophilic attack on the γ-phosphate [37]. An example of the latter mechanism is the “arginine finger” provided by the GAP for Ras referred to as RasGAP [33, 38].

After a GTPase hydrolyzes GTP, inorganic phosphate (P_i) is released and GDP remains in the active site. If a GTPase binds to GDP with a low affinity, the GTPase cycle can continue as

GDP dissociates from the GTPase after GTP hydrolysis, enabling the enzyme to bind GTP again and a GEF is not necessary. Alternatively, if the GTPase binds more tightly to GDP than it does to GTP, then a GEF is necessary to allow GDP to be released from the enzyme so that the GTPase cycle can proceed efficiently. In bacterial cells, some GTPases require a GEF, such as the GTPase EF-Tu. EF-Tu binds more tightly to GDP than to GTP and depends on its GEF, EF-Ts in order to carry out its functional cycle. EF-Ts facilitates GDP release from EF-Tu by interfering with the interaction of GDP with the magnesium ion through introducing a phenylalanine residue into the active site that disturbs an aspartate residue found in EF-Tu's G3 motif [32].

In contrast to the large body of work on GEFs, very little is known about bacterial GDIs. GDIs were first discovered in eukaryotes, where GDIs have been identified that target the Rho, Rab and Ras subfamilies but not the Ran and Arf subfamilies of small GTPases in mammals [34]. As the Rho, Rab and Ras subfamilies but not the Ran and Arf subfamilies of small GTPases can be reversibly anchored in the membrane by insertion through their prenylated C-terminus, it has been suggested that GDIs that target small GTPases primarily play the role of providing a hydrophobic cavity for the prenyl modification in order to allow prenylated GTPases to be transported into the cytosol and between membranes [34, 39]. Prenylation refers to the process whereby proteins are post-translationally modified by the addition of an isoprenoid group (farnesyl or geranylgeranyl) to a CAAX sequence at the C-terminus of the protein [40]. The idea has been put forward that the interaction of the isoprenoid group of the small GTPase from the Rho, Rab or Ras subfamilies with the β-sheet-lined cleft in the C-terminal domain of the GDI stabilizes the interaction of the N-terminal GTPase-binding domain of the GDI protein with the small GTPase, preventing dissociation of GDP and abrogating interaction with both GEFs and GAPs [34, 41]. Yeast cannot survive with a RabGDI deletion and patients from several families with X-linked mental retardation were shown to express RabGDI1 with missense mutations [42, 43]. These

mutations are in the cavity bridging the GTPase-binding domain and the lipid-binding domain of RabGDI1 and reduce the binding affinity of the GTPase Rab3A for RabGDI1 by up to 8-fold [43]. GDIs are also regulators of some heterotrimeric G proteins in eukaryotes. The bacterial proteins FeoB, C3 exoenzyme from *Clostridium botulinum* and YpkA from *Yersinia pseudotuberculosis* have all been reported to exhibit GDI activity and to possess distinct GDI domains, in addition to other domains with different activity [44-46]. It has been suggested that the GTPase FeoB requires a GDI because its high rate of GDP release and very slow GTPase activity would cause it to be constitutively active [47]. However, it remains unknown how an intrinsic GDI activity contained within the protein itself could be playing the role of the required GDI.

Thus, it appears that in both bacteria and in eukaryotes, GDIs are sometimes necessary for the functional cycle of some GTPases but may not be required for the function of other GTPases.

In this thesis, the term “universally conserved” is used to refer to proteins that are encoded by genes found in the genomes of organisms from all three domains of life. The proteins described here as “universally conserved” are encoded in the genomes of most but not necessarily all species that belong to each domain of life [29]. This usage is consistent with the definition of the term “universally conserved” in the current biochemical literature [29, 48, 49]. There are eight GTPases that are universally conserved across all domains of life including IF2, EF-G, EF-Tu, YchF, HflX, YihA, FtsY and Ffh, where all names refer to *E. coli* proteins [29, 31]. These eight proteins have been assigned different names in organisms from each domain of life (Table 1.1). IF2, EF-G and EF-Tu belong to the Translation Factor superfamily [31]. IF2 plays a role in initiation of protein translation [31, 50]. In contrast, EF-G and EF-Tu are both translation elongation factors; EF-G shifts the position of the ribosome on the mRNA thereby accelerating the translocation of tRNAs from the aminoacyl tRNA binding site (A-site) to the peptidyl tRNA binding site (P-site) and from the P-site to the tRNA exit site (E-site) of the 70S ribosome [51, 52]. The aminoacylated

tRNA is recruited to the A-site of the ribosome by EF-Tu [53, 54]. YchF and HflX are grouped together in the OBG-HflX-like superfamily [31]. The function of YchF remains unknown, although there is some evidence that YchF may act as a controller of the oxidative stress response in bacterial cells [55]. Furthermore, YchF plays a role on the 70S ribosome as the 70S ribosome acts as a GAP for YchF [31, 49]. Likewise, the 70S ribosome as well as the 50S ribosomal subunit both represent GAPs for the functional cycle of HflX. [48, 56]. However, some insight into the function of HflX has emerged as HflX has been shown to dissociate 70S ribosomes into 50S and 30S ribosomal subunits [57] and to dissociate hibernating 100S ribosome dimers [58]. YihA is a member of the TrmE-Era-EngA-YihA-Septin-like superfamily [31]. The emerging functions of YihA are discussed in sections 1.5 and 1.5.3 and include regulation of ribosome biogenesis and putatively, regulation of the cell cycle [59, 60]. Finally, Ffh and FtsY are in separate families of the signal recognition particle, MinD and BioD (SIMIBI) class of proteins [31] and are components of the signal recognition particle (SRP) and the SRP receptor, respectively [61]. The SRP and SRP receptor regulate the targeting of secretory and membrane proteins during protein synthesis [62].

Table 1.1 Names of the eight universally conserved GTPases in each domain of life. Abbreviations are given by Sso (*Sulfolobus solfataricus*), Mja (*Methanococcus jannaschii*), Pab (*Pyrococcus abyssi*), Pho (*Pyrococcus horikoshii*). The headings of each column provide the names of the GTPases in *E. coli*.

GTPase	IF2	EF-G	EF-Tu	YchF	HflX	YihA	FtsY	Ffh
Domain Bacteria								
Bacillus subtilis	IF2 [63]	EF-G [64]	EF-Tu [64]	YyaF [65]	HflX [66]	YsxC [67]	FtsY [68]	Ffh [69]
Thermotoga maritima	IF2 [70]	EF-G [64]	EF-Tu [64]	TM1240 [65]	HflX [66]	YsxC [71]	FtsY [69]	Ffh [69]
Domain Archaea	aIF5B [70]	aEF2 [64]	aEF1A [64]	Ygr210 [72]	SsGBP/ SsoHflX (Sso) [73, 74]	MJ0320 (Mja) PAB2234 (Pab) PH0200 (Pho) [65]	FtsY [69]	SRP54 [69]

Continued on next page

GTPase	IF2	EF-G	EF-Tu	YchF	HflX	YihA	FtsY	Ffh
Domain Eukarya								
Homo sapiens	hIF2/ eIF5B [75, 76]	eEF2 [77]	eEF1A1, eEF1A2 [77]	hOLA1 [78]	GTPBP6/ PGPL [73, 79, 80]	GTPBP8 Isoform 1 GTPBP8 Isoform 2 [81]	SRα [69]	SRP54 [82]
Saccharo-myces cerevisiae	eIF5B [83]	eEF2 [77]	eEF1A1, eEF1A2 [77]	Ola1p [84]	Not found in genome [85]	MRX8 [81]	SRα [69]	Srp54p [82]
Arabidop-sis thaliana	eIF5B1 [86]	AtEF-G1mt plastid EF-G [87]	Rabe1b/ SVR11 [88, 89]	AtYchF1 [90]	AtHflX [91]	At2g22870 [65]	SRα cpFtsY [68, 69]	cpSRP54 [82]

1.3.1 GTPASES THAT ACT AS BACTERIAL RIBOSOME BIOGENESIS FACTORS

Many proteins that transiently bind to the ribosome are required in biogenesis of bacterial ribosomes, some of which are GTPases. However, unlike the process of eukaryotic ribosome biogenesis for which some of the steps have been described in detail, the mechanisms that explain how prokaryotic GTPases mediate the maturation of premature ribosomes and the role that the GTPase activity of these proteins plays in the assembly process are yet to be determined. The current understanding of the roles of GTPases in bacterial ribosome assembly will be discussed next.

1.3.2 YPHC

YphC is grouped in the TEES family of GTPases that includes TrmE/MnmE, Era, EngA/YphC, YihA and Septins [31]. GTPases in the TEES family share a highly conserved sequence located between the G1 and G3 (Switch II) motifs [66]. YphC, also referred to as Der, EngA, or YfgK, is required for survival of many species of bacteria but not universally conserved as it is not found in archaea [92-94]. A unique feature of YphC is that it possesses two tandem GTP-binding domains referred to as GD1 and GD2 [92]. A short stretch of negatively charged amino acids separates GD1 from GD2 [92]. The third domain in YphC is a C-terminal K Homology domain (KH domain) that could form contacts with RNA via several successive basic residues [95].

Evidence that YphC plays a role in ribosome biogenesis was obtained from studies in *E. coli* cells with a deletion of the *rrmJ* gene which results in an abnormal ribosome profile with excess ribosomal subunits and a very long doubling time [96]. However, a normal phenotype was observed in Δ*rrmJ* cells that overexpressed YphC [97].

The cryo-EM structure of YphC bound to the *E. coli* 50S ribosomal subunit revealed contacts between arginine and lysine residues of YphC's KH domain and helix 80 and helix 93

of the peptidyl transferase centre [98]. This interaction positions the KH domain of YphC so that the last 30 amino acids of the KH domain, which have a sequence unique to each species and are not resolved in this structure, could bind directly to the peptidyl transferase centre and ribosomal protein L16 [98]. There is also a cryo-EM structure available of a eukaryotic homologue of YihA, mt-EngA, bound to a mitoribosomal large ribososomal subunit assembly intermediate from *Trypanosoma brucei* [99]. Further insight into the role of YphC in ribosome assembly has come from depletion of YphC. A 45S precursor of the mature 50S ribosomal subunit can be isolated from *B. subtilis* with a depletion of YphC [100-102]. This 45S precursor lacks a mature fold in the rRNA of the A-, P- and E-sites as well as the L7/L12 stalk and the central protuberance [103]. This evidence has led to the proposal that YphC may mediate the proper folding of rRNA at the central protuberance, tRNA binding sites and bL7/bL12 stalk of the 50S ribosomal subunit [103].

1.3.3 RBGA

Although a cryo-EM structure of RbgA bound to the mature 50S ribosomal subunit from *B. subtilis* has been published, the mechanism for promotion of maturation of the 50S ribosomal subunit by RbgA remains unknown. However, RbgA is the only bacterial ribosome biogenesis factor for which a cryo-EM structure of the assembly factor bound to a bacterial ribosome biogenesis intermediate intermediate is available [104].

1.3.4 OBG

Obg is also named as CgtA or YhbZ. Obg's binding site on the 50S ribosomal subunit is localized within the intersubunit interface of the 50S ribosomal subunit such that Obg's GTPase domain forms contacts with the sarcin-ricin loop and the bL7/bL12 stalk at the GTPase activating centre [105]. Although the C-terminal domain of Obg was not resolved in this

structure, confirming its mobility, the N-terminal domain of Obg binds to the peptidyl transferase centre and shows a surprising structural mimicry of variants of Obg with basic residues.

1.3.5 ERA

Like YphC, Era is a member of the TEES family of GTPases. Although Era-depleted cells exhibit impaired translation, the inability of Era to rescue normal translation in an *in vitro* translation system made with cell extracts from *E. coli* with a depletion in Era suggested that Era does not play a direct role in bacterial translation [106]. However, Era has been shown to be involved in the maturation of 30S ribosomal subunits [106, 107]. The N-terminus of Era interacts with helix 37 at G1099 and G1100 of the 16S rRNA while a helix-turn-helix structure located within the C-terminal KH domain of Era forms contacts with nucleotides 1530 to 1534 of helix 45 in the 3' minor domain of the 16S rRNA [108].

1.3.6 RSGA

RsgA belongs to the YlqF family of GTPases in which the GTPase motifs are circularly permutated [66]. RsgA is referred to as YjeQ in some published reports. The 30S ribosomal subunit increases RsgA's GTP hydrolysis activity by a factor of 160, suggesting that RsgA plays a role in biogenesis of the 30S ribosomal subunit [109]. Studies of the role of RsgA in ribosome biogenesis have been hampered by the finding that no precursor of the 30S ribosomal subunit accumulates when RsgA is depleted from cells. However, several cryo-EM structures of RsgA bound to the mature 30S ribosomal subunit have been published. RsgA forms a complex with the mature 30S ribosomal subunit that contains contacts between RsgA and the head, platform and helix 44 of the 30S ribosomal subunit [109].

1.3.7 YQEH

YqeH is another member of the YlqF family of GTPases containing a circular permutation [66] and YqeH is essential for cell survival [110]. When YqeH was depleted in *B. subtilis*, the 70S ribosome fraction was diminished concommitantly with an increase in the fraction of 50S ribosomal subunits [110]. Although an increase in the fraction of 30S ribosomal subunit was not observed upon depletion of expression of YqeH, northern blotting revealed the accumulation of a pre-16S rRNA in rRNA isolated from *B. subtilis* in which YqeH had been depleted, providing evidence that YqeH plays a role in assembly of the 30S ribosomal subunit [110]. In addition, the abundance of mature 16S rRNA was reduced in *B. subtilis* in which YqeH was depleted and the proportion of 23S rRNA to 16S rRNA was increased [111]. Futhermore, deletion of the potential zinc ribbon motif in YqeH, a motif involved in binding of RNA to proteins and notably present in some ribosomal proteins, abolished survival of cells with the deletion [111].

1.4 INTRODUCTION TO YIHA

The protein YihA is the only GTPase of the eight GTPases that are conserved across all domains of life (IF2, EF-G, EF-Tu, YchF, HflX, YihA, FtsY and Ffh) that has been suggested to be involved in bacterial ribosome biogenesis [31, 48]. YihA has other potential roles in the cell, including potential roles in bacterial cell division [60] and signaling to regulate protein degradation in bacteria [112] but may potentially be involved in mitochondrial translation in trypanosomes [113]. The names of eukaryotic homologues of YihA are provided in Table 1.1 and section 1.5.1. Given that mitochondrial ribosomes originated from bacterial ribosomes via endosymbiosis and that the role of YihA in eukaryotic cells remains unknown, it is interesting that YihA translocates to the mitochondria in the malaria parasite *Plasmodium falciparum (P. falciparum)* [114]. However,

mitochondrial signal sequences targeting YihA to the mitochondria are predicted to be present in the *P. falciparum* YihA protein [115], in mammalian and zebrafish YihA homologues but not in YihA proteins in plants or fungi (D. Capatos, unpublished results).

The interaction partners of YihA in the cell are gradually beginning to emerge from mass spectrometry studies [59, 67, 103]. YihA is among the smallest known GTPases, including the eukaryotic Ras and Ran GTPases [29] and consists of only a single domain. The interactome of YihA is likely to be very important for YihA's function and for potentially conferring additional functions to YihA given that YihA only possesses a single G-domain. This raises the question how YihA is able to interact with the rRNA of ribosomes during ribosome biogenesis since it lacks a recognisable RNA binding domain. In this thesis the emerging knowledge as to how YihA has adapted to and solved this problem of recognizing rRNA will be discussed. This thesis also explores what is currently known of YihA's role in bacterial ribosome biogenesis as well as its potential role as an antibiotic target in addition to what is known about the sequence and function of its homologues in eukaryotic cells.

1.4.1 NOMENCLATURE OF YIHA PROTEINS

In the domain Bacteria, homologues of YihA have been identified in both gram positive and gram negative bacteria. The homologue of YihA found in the gram-positive bacterial species *Bacillus subtilis (B. subtilis)* and *Staphylococcus aureus (S. aureus)* has been named YsxC [59, 67]. *B. subtilis* YihA will be referred to hereafter as BsYihA and *S. aureus* YihA will be referred to as SaYihA. In the gram negative bacterium *E. coli*, the YihA homologue is referred to as YihA whereas in the gram negative species *Thermotoga maritima (T. maritima)*, YihA is referred to as YsxC [59, 71]. Hereafter, *E. coli* YihA will be referred to as EcYihA and *T. maritima* YsxC will be

referred to as TmYsxC. The above mentioned species (*B. subtilis*, *S. aureus*, *E. coli*, and *T. maritima*) are the only bacterial species in which YihA homologues have been studied thus far.

Three eukaryotic genomes from model organisms have been shown to contain genes encoding homologues of YihA, namely, human, yeast and plant genomes [65]. The human homologue of the YihA protein is referred to as GTPBP8 [81], the *Saccharomyces cerevisiae (S. cerevisiae)* homologue of YihA is named MRX8 [81], and in *Arabidopsis thaliana (A. thaliana)*, the homologue of YihA is At2g22870 [65] (Table 1.1). The only additional eukaryotic homologues of YihA that have been characterized are in the eukaryotic kingdom Protozoa [116]. The protozoan *Trypanosoma brucei* (*T. brucei*) contains a homologue of YihA referred to as mtYsxC whereas the protozoan *P. falciparum* possesses three homologues of YihA named PfYihA1, PfYihA2 and PfYihA3 [113, 114].

The only information that is available about archaeal YihA to date involves the gene sequences and phylogenetic analysis for some homologues of YihA found in archaeal species whose genomes have been sequenced. Nevertheless, one comparative genomics study has identified and assigned names to homologues of YihA in three archaeal species [65]. In the archaeal species *Methanococcus jannaschii* (*M. jannaschii*), the YihA homologue is referred to as MJ0320 [65]. Additionally, in *Pyrococcus abyssii* (*P. abyssii*) and *Pyrococcus horikoshiii* (*P.horikoshii*), which have been studied extensively as model archaeal organisms [117], the YihA homologues have been designated as PAB2234 and PH0200, respectively [65]. Future studies will identify other archaeal YihA homologues.

1.4.2 SEQUENCE AND STRUCTURE OF YIHA

Structural studies of BsYsxC bound to guanosine [(β,γ)-imido] triphosphate (GMPPNP), a non-hydrolyzable analogue of GTP or to GDP [112, 118] and TmYsxC bound to GDP and in the *apo* form [71] by x-ray crystallography revealed that YsxC contains a GTP-binding domain (G domain) that consists of a core of a 6–stranded β-sheet encircled by 5 α-helices (Figure 1.1). However, a 15-20 residue α-helix is also found at the C-terminus of YihA and an additional β-strand is located at the N-terminus [71, 112]. The alignment of YihA sequences included in Figure 1.2 shows the species in the same order in which they are grouped in the phylogenetic tree for YihA (Figure 1.3) for consistency, highlighting the grouping of archaeal species with animals and yeast and the grouping of *arabidopsis thaliana* with bacterial species. Multiple sequence alignment confirms that EcYihA contains the five canonical G motifs within its G domain characteristic of the superfamily of GTPases (Figure 1.2). The G1 motif or P-loop contains the conserved sequence GxxxxGKS/T that is required to form contacts with the phosphate groups of the bound nucleotide [30, 119]. The G2 motif or Switch I and the G3 motif or Switch II interact with the magnesium cation and with the γ-phosphate of GTP and adopt alternative conformations in the presence of di-phosphate or tri-phosphate nucleotide [30, 112]. The loops corresponding to switch I and switch II are mobile in the GDP-bound state of YsxC and more rigid when bound to GTP [71]. Switch I and Switch II have been identified as interaction sites for effectors of GTPases [30]. The G4 motif consists of the sequence (N/T)(K/Q)xD and together with the G5 motif facilitates specific recognition of the guanine nucleotide [30]. YihA is part of the TrmE-Era-EngA-YihA-Septin (TEES) family of GTPases that contains two conserved asparagine residues found in the sequence xxGxx**N**$_1$xGKSxxx**N**$_2$, where the G1 motif is underlined and N_1 and N_2 represent conserved asparagine residues (Figure 1.3) [120, 121]. N_1 is almost universally conserved within the G1 motif and N_2 is commonly located eight residues downstream of N_1 in the direction towards the C-terminus of YihA (Figure 1.3, Appendix Figure 1.1A). There is also very commonly an

asparagine that follows the G2 motif (Appendix Figure 1.1A). The sequence of *Neurospora crassa* YihA is unusual in that it contains a very long insert between the G1 and G2 motifs and a G2 motif that is substituted with a histidine residue in place of the consensus lysine or threonine (Appendix Figure 1.1A, B). Similarly, in the other TEES family GTPases, the distance of N_2 from the G1 motif does not normally vary from eight residues downstream of the N_1 residue in the direction towards the C-terminus [120, 121]. Most homologues of YihA do not contain the conserved hydrophobic residue within the G3 motif that is found in some members of the hydrophobic amino-acid substituted (HAS) family of GTPases [120]. YihA contains a single conserved threonine in the G2 motif of its G domain whereas its G1 and G4 motifs contain the conserved sequences present in most other translation factor-related (TRAFAC)-class GTPases [31, 119]. A structural alignment of TmYsxC and BsYsxC was prepared (Figure 1.1). The RMSD of the alignment is 1.477 Å and the Switch I regions of TmYsxC-GDP and BsYsxC-GMPPNP are overlapping (Figure 1.1).

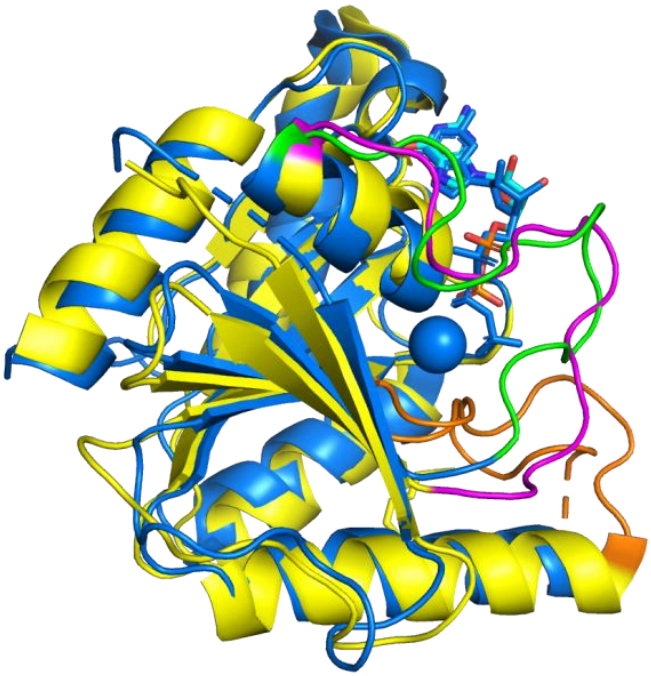
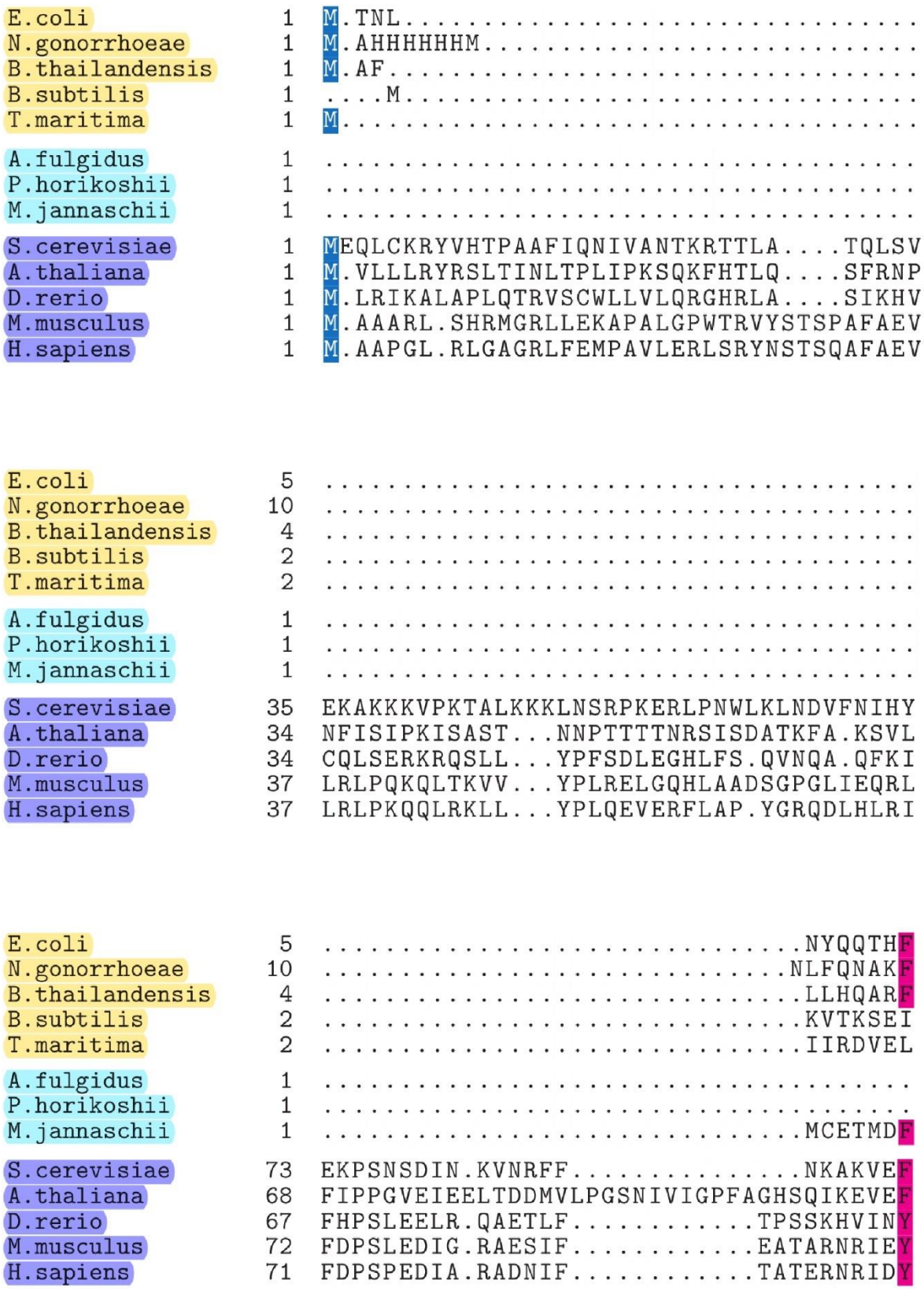
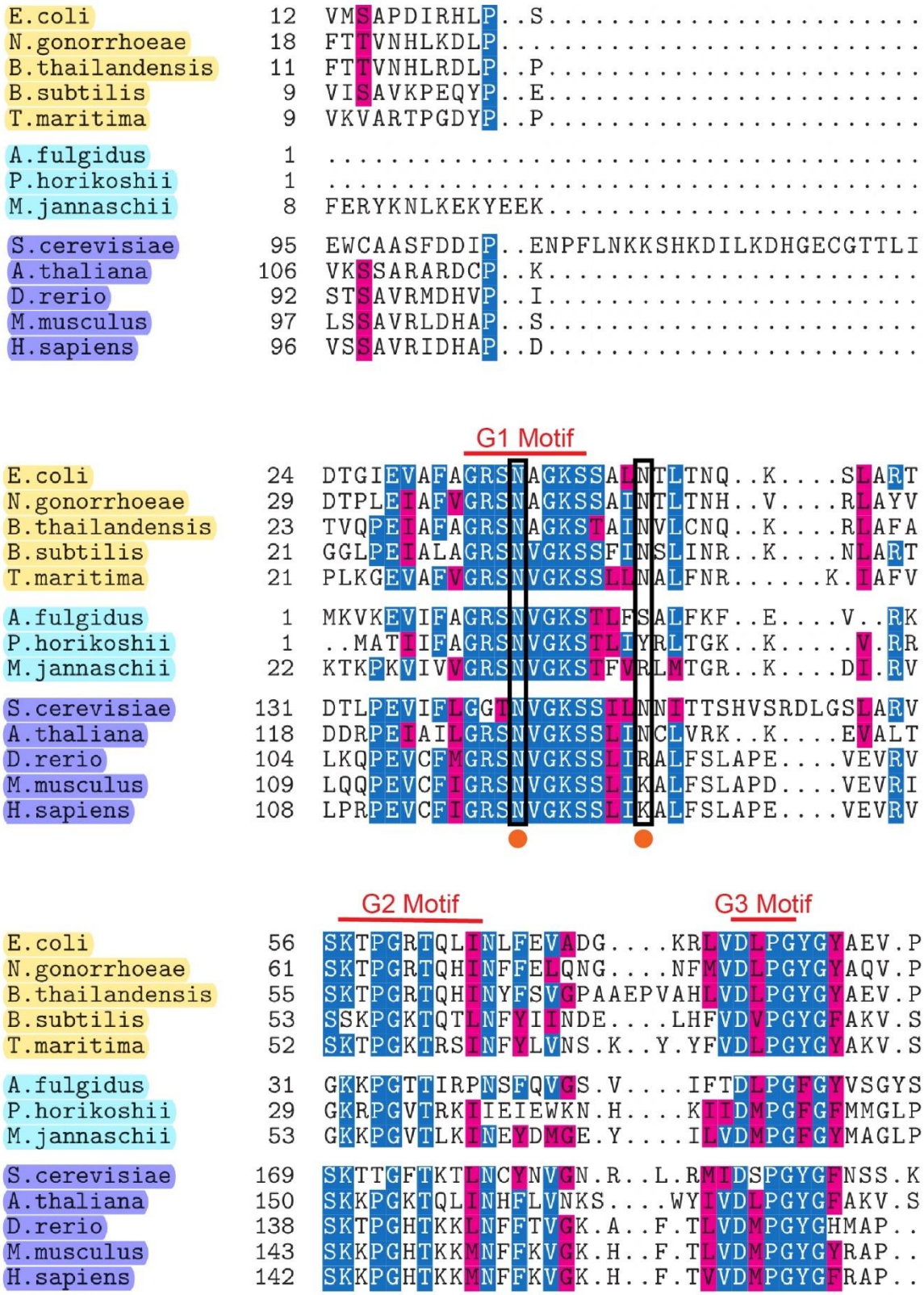


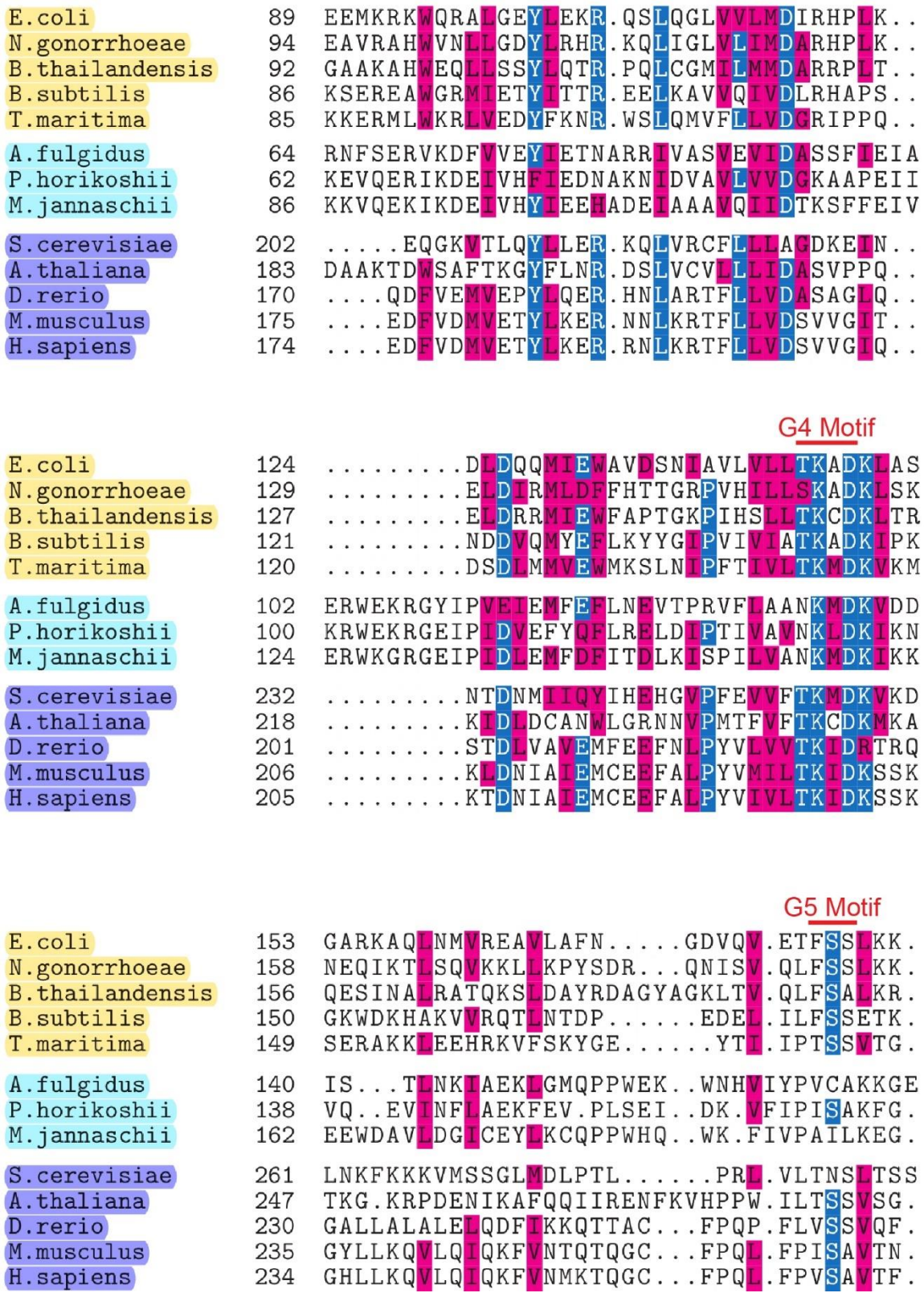
Figure 1.2. Structural alignment of TmYsxC with BsYsxC. Depicted in yellow is TmYsxC bound to GDP while blue depicts BsYsxC bound to GMPPNP. The blue sphere represents a Mg^{2+} ion. The pink region shows Switch I of TmYsxC bound to GDP and the green region represents Switch I of BsYsxC bound to GMPPNP.



Continued next page.



Continued next page.



Continued next page.

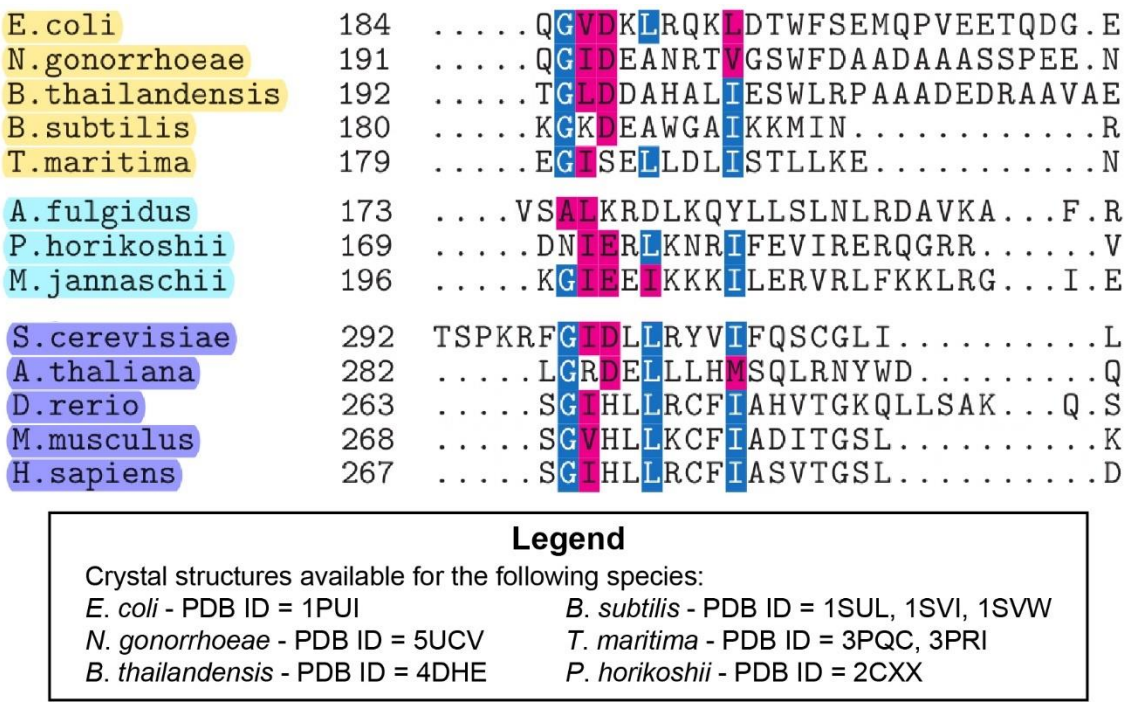


Figure 1.3. Protein sequence alignment of bacterial, eukaryotic and archaeal YihA. Motifs G1-G5 are indicated with red labels above the sequences. The conserved asparagine residues are outlined in bold black lines with an orange circle beneath the residues at each position. Dark blue shading denotes amino acid residues that are greater than or equal to 50% identical whereas pink shading denotes similar residues. Sequences are grouped according to domain, then according to phylum. Organisms that belong to the domain Bacteria are shown in yellow, organisms that belong to the domain Archaea are shown in cyan and organisms in the domain Eukaryota are shown in purple. The National Center for Biotechnology Information (NCBI) accession numbers are P0A6P7.1 for *E. coli*, 5UCV_B for *Neisseria gonorrhoeae*, AAB98305.1 for *M. jannaschii*, KUK06096.1 for *Archaeoglobus fulgidus,* NP_001119874.1 for the zebrafish *Danio rerio*, EDK98048.1 for GTP-binding protein 8 in *Mus musculus*, NP_054889.2 for GTP-binding protein 8 in *Homo sapiens*, AAB64772.1 for *S. cerevisiae*, WP_010884308.1 for *P. horikoshii*, WP_009906421.1 for *Burkholderia thailandensis*, PWI62610.1 for *B. subtilis,* 3PQC_A for *T. maritima* and O81004.2 for *A. thaliana.* The alignment was prepared with M-Coffee.

Although YihA possesses no RNA binding domain, YihA is still able to interact with ribosomal particles. As expected for a protein that binds to the ribosome and to rRNA, YihA contains two positively charged patches of basic residues on its surface as highlighted in the electrostatic surface potential map (Figure 1.3). Residues R31, R116, R117 and K146 (*B. subtilis* numbering) form one basic patch on the surface of YihA that is in proximity to the active site

(Figure 1.3). The C-terminal tail of YihA consisting of residues K182-R195 (*B. subtilis* numbering) contains the second basic cluster of residues that could interact with RNA and are required for binding to the 50S ribosome [71, 112, 122]. The electrostatic surface of YihA shows the position of the basic C-terminal helix of YihA relative to YihA's active site (Figure 1.3).

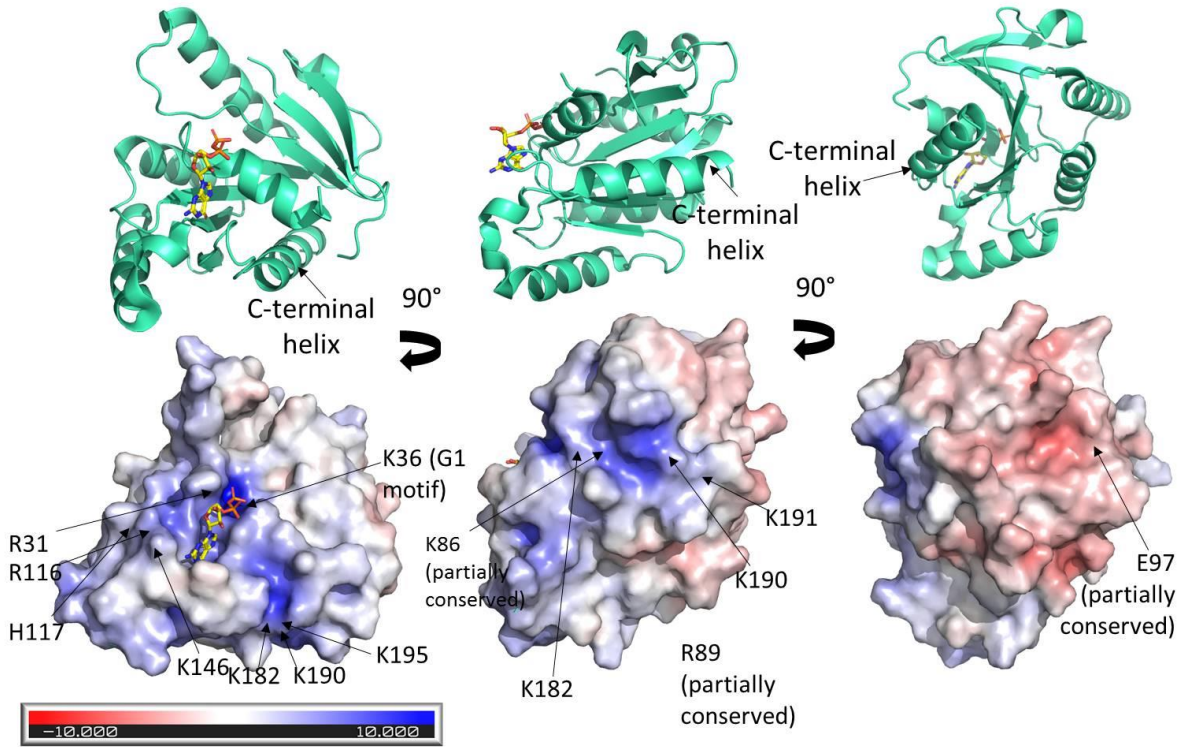


Figure 1.4. Electrostatic surface potential of BsYsxC bound to GDP. Red indicates electronegative regions and blue indicates electropositive regions. The electrostatic surface map was prepared using the APBS/PDB2PQR web server [123].

1.4.3 FUNCTIONAL CYCLE OF YIHA

The current experimental evidence on the functional cycle of YihA suggests that a ribosome assembly precursor of the 50S ribosomal subunit termed the 44.5S intermediate acts as a GAP for BsYsxC [103]. It has not yet been established how YihA's molecular mechanism differs from those of other ribosome-associated bacterial GTPases. Several methods have been used to show that YsxC binds to both GTP and GDP [124, 125]. YihA binds to 2'-/3'-O-N'-methylanthraniloyl (mant)-labeled-GDP and to GDP with an affinity close to 3 μM and to mant-

GTP with an affinity close to 27 μM [119] (Table 1.2). A model of the functional cycle of YihA is given in Figure 1.4. The affinity of YihA for GDP is higher in the presence of magnesium, suggesting that a role of magnesium ions coordinating phosphate groups in the active site of G proteins is also required in the catalytic mechanism of YihA [119]. GTP and GDP also promote the interaction of YihA with 50S ribosomal subunits [126]. The affinity of YihA for GTP and GDP differs by an order of magnitude. Small GTPases often bind very tightly to GDP and co-purify with GDP [34]. It remains unknown whether EcYihA co-purifies with GDP, but TmYsxC was found to co-purify with GDP [71]. In addition, previous work suggested that YihA contains a single nucleotide binding pocket such that only one nucleotide binds to each YihA protein molecule [119]. As summarized in Table 1.1, mutagenesis of the nucleotide binding region of YihA showed that single amino acid substitutions of mainly conserved residues in the P loop (G38V and K39T) and in the G4 motif (T145N and D148N) prevented YihA from binding mant-GDP [119]. Experiments in which the YihA-mant-GDP complex was titrated with increasing concentration of ADP and similarly titration of mant-ADP with increasing concentrations of YihA revealed that YihA does not bind to adenine nucleotides [119, 125]. Interestingly, the conserved NKXD sequence of the canonical G4 motif found in most GTPases is also present in YihA but many prokaryotic and eukaryotic homologues of YihA contain the conserved sequence of TKXD that is also found in FtsY which is involved in targeting of translating ribosomes to membranes (Figure 1.2). However, it remains unknown whether the functional cycle of YihA requires a GDI or a GEF. GTP hydrolysis activity of YihA is not stimulated by potassium even though YihA is in the TEES family of GTPases [120, 121]. GTP hydrolysis activity of YsxC is stimulated by 50S ribosomal particles and the turnover number of YsxC, k_{cat}, is increased 38-fold by 50S ribosomal particles [103, 126]. GTP hydrolysis by YsxC is stimulated by the 44.5S assembly intermediate in YsxC-depleted cells and also activated by 45S intermediates in YphC and RbgA-depleted cells [103]. Although the precise functions of the GTPases YphC and RbgA remain unknown, a general understanding of their

roles in the cell is emerging. When RbgA, YphC or YsxC is depleted in *B. subtilis*, 45S precursors of the 50S ribosomal subunit are observed that are lacking the ribosomal proteins uL16, bL28, bL35 and deficient in bL27, bL33 and bL36 when precursors are isolated with buffers containing 150 mM NH_4Cl [103]. Interestingly, uL16, bL28, bL35, bL27, bL33 and bL36 are all situated either in the bottom region of the central protuberance or adjacent to this region in the *B. subtilis* 50S ribosomal subunit [103]. The central protuberance refers to the middle bulge of three defining bulges that resemble a crown structure on the surface of the large ribosomal subunit and is used to specify the orientation of the large ribosomal subunit [127]. At the interface between the ribosomal subunits, the central protuberance forms one of the sites of contact between the large and small ribosomal subunits [128]. It has been proposed that the central protuberance functions in maintaining intersubunit bridges connecting the large ribosomal subunit with the small ribosomal subunit and in interacting with with tRNAs in the A-, P- and E-sites of the ribosome, possibly ensuring accuracy of translation, and there is limited evidence to support this claim [127, 129, 130]. *In E. coli* ribosomes, the central protuberance is composed of several ribosomal proteins including uL5, uL16, uL18, bL25, bL27, bL31, bL33 and bL35 [131]. These ribosomal proteins along with the 5S rRNA and segments of the 23S rRNA including helices 81 – 85 of domain V and helices 38, 39 and the loop joining helices 41 and 42 found in domain II make up the central protuberance of the *E. coli* ribosome [127, 131, 132]. The location of the central protuberance in the 70S *E. coli* ribosome is shown in Figure 1.7. RbgA forms a complex with the 45S precursor of the 50S ribosomal subunit from RbgA-depleted cells by binding to the P-site [104]. The current model of RbgA and YphC function suggests that just prior to the stage in which the essential GTPases RbgA, YphC and YsxC bind to precursors of the 50S ribosomal subunit, ribosomal protein uL6 interacts with these precursors that are postulated to be maturing through several different biogenesis routes [133]. The 45S precursors found in RbgA- or YphC-depleted cells and the 44.5S precursor found in YsxC-depleted cells all contain a central protuberance with

a mature fold of the 23S and 5S rRNA and bind to each of RbgA, YphC and YsxC, either in parallel or in a specific order but most probably independently of each other [133]. In the final steps of biogenesis of the 50S ribosomal subunit, RbgA remodels the A-site rRNA at helices 38, 91 and 92 and remodels the P-site rRNA at helix 93, leading to the attainment of the final, folded state for these helices [104]. The identity of the helices that are remodelled by YphC and YsxC remains unknown, but it is predicted that helix 89 within the A-site, helices 69 and 71 within the P-site and helix 68 inside the the E-site are among the helices that are remodelled at this stage [133]. Remodelling of rRNA by RbgA, YphC and YsxC is likely to provide the correct conformation for uL16, bL27, bL28, bL33, bL35 and bL36 to bind to precursors of the 50S ribosomal subunit and is one of the last steps in biogenesis of the 50S ribosomal subunit in *B. subtilis* in addition to the stimulation of GTP hydrolysis by RbgA in the presence of uL16 [133, 134].

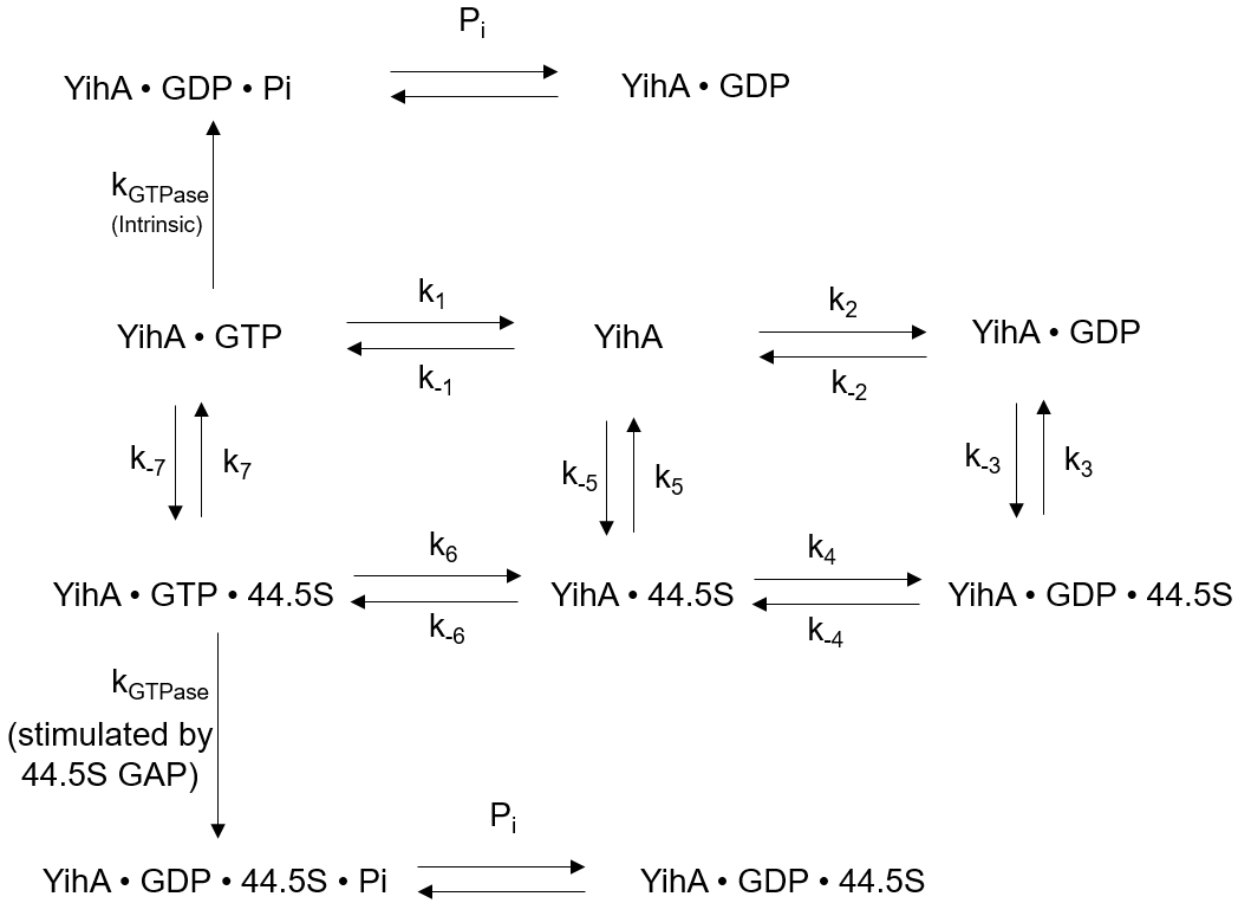


Figure 1.5. Kinetic model for the functional cycle of YihA. ON and OFF states for YihA are not indicated in this model as the roles that YihA plays in its ON state compared to in its OFF state remain unknown. Whether YihA is in its ON state in its GDP-bound or GTP-bound form also has not been investigated. Earlier work on the eukaryotic small GTPase Ran suggested, contrary to the canonical idea that a GTPase is in its ON state only in the GTP-bound form, that Ran interacts with the effector protein NTF2 when Ran is bound to GDP [41, 135]. Thus, assumptions that correlate YihA's ON state to the GTP-bound or GDP-bound form of YihA are not made in this model. Previous work indicated that the 44.5S precursor of the 50S ribosomal subunit is able to stimulate the GTPase actvity of BsYsxC, thereby functioning in the role of a GAP for BsYsxC [103]. The 44.5S ribosome assembly intermediate observed in YsxC-depleted cells is shown as the GAP for YihA (see section 1.5.3 of this thesis), with a rate constant for the GTPase reaction given by the k_{GTPase} labelled as "stimulated by 44.5S GAP". In the case of translational GTPases, the ribosome acts as a GAP in a reaction where the phosphate of A2662 forms an interaction with the catalytic histidine allowing the histidine to be situated in the active site in the correct orientation for catalysis [136]. In contrast, there is no structure of YihA bound to the 44.5S intermediate so the molecular details of how the 44.5S intermediate acts as a GAP for YihA remain unknown. The model here depicts GTP hydrolysis as analogous to the mode of GTP hydrolysis by HflX [48] and nucleotide exchange as analogous to the exchange of GDP for GTP on EF-Tu promoted by its GEF EF-Ts [137].

Table 1.2 Previously characterized variants of YihA [119, 122].

Mutation	Organism & Protein Name	Functional Region	Effect
R31A	*B. subtilis* YsxC	Surface residue that contributes to positively charged surface region	Some evidence it may promote association of YsxC with 30S ribosomal subunit
S37A	*B. subtilis* YsxC	Walker A motif	No effect found
G38V	*E. coli* YihA	G1 motif	Unable to bind guanine or adenine nucleotides
K39T	*E. coli* YihA	G1 motif	Unable to bind guanine or adenine nucleotides
Y82W	*E. coli* YihA	G3 motif	No effect on catalytic activity
R116A	*B. subtilis* YsxC	Surface residue that contributes to positively charged surface region	No effect found
H117N	*B. subtilis* YsxC	Surface residue that contributes to positively charged surface region	No effect found
T145N	*E. coli* YihA	G4 motif	Unable to bind nucleotides
K146M	*B. subtilis* YsxC	Surface residue that contributes to positively charged surface region	Unable to bind nucleotides
D148N	*E. coli* YihA	G4 motif	No effect found
R31A R116A	*B. subtilis* YsxC	Positively charged surface	No effect found
H117N K146N	*B. subtilis* YsxC	G4 motif and positively charged surface	No effect found
ΔK182-R195	*B. subtilis* YsxC	C-terminal α-helix	Unable to associate with 16S and 23S rRNA or with 50S ribosomal subunits
ΔK190-R195	*B. subtilis* YsxC	C-terminal α-helix	Unable to associate with 16S and 23S rRNA or with 50S ribosomal subunits

Table 1.3. Biophysical parameters governing the functional cycle of EcYihA and BsYsxC

	K_D (μM)			k_{cat} (h^{-1})	Apparent K_M (μM)	k_{cat}/K_M ($\mu M^{-1}\ h^{-1}$)
	Mant-GDP	GDP	Mant-GTP			
EcYihA (+ Mg^{2+})	3.6 ± 0.3	3.9 ± 0.5	–			
EcYihA (- Mg^{2+})	2.3 ± 0.2	8.6 ± 0.2	27.4 ± 9			
BsYsxC	–	–	–	5.5 ± 3.3	1268 ± 116.6	0.004
BsYsxC + 50S	–	–	–	212.3 ± 40.9	377.1 ± 166.7	0.5
BsYsxC + 44.5S	–	–	–	149 ± 25.3	126.3 ± 72.5	1.1
Method used to determine	fluorimetry (FRET to mant)	fluorimetry	fluorimetry (FRET to mant)	GTP hydrolysis assay (Malachite green)	GTP hydrolysis assay (Malachite green)	GTP hydrolysis assay (Malachite green)
Reference	[119]	[119]	[119]	[103]	[103]	[103]

1.4.4 POTENTIAL ROLE OF YIHA IN RIBOSOME BIOGENESIS

It remains unknown whether YihA's GTPase activity can serve as an rRNA folding checkpoint, a sensing device, or is required for timing events and incorporation of ribosomal proteins during the assembly of ribosomes. Current models of biogenesis of the bacterial ribosome are lacking knowledge of molecular interactions of assembly factors with the respective immature ribosomal particles [23]. Although principles of the 30S ribosome subunit assembly are beginning to emerge, questions remain unanswered with respect to why many factors, such as YihA, are needed in the assembly process. Several biophysical studies on the assembly of the bacterial 30S ribosomal

subunit monitoring binding of ribosomal proteins and rRNA have shown conformational changes in the 16S rRNA using a FRET-based assay [138-140]. Pulse chase labelling analyzed by quantitative mass spectrometry has shown that the ribosome assembly process may follow several simultaneous pathways [134]. A key *in vitro* experiment in which the purified methyltransferase RlmE was added to bacterial cell extract containing a 45S intermediate from bacterial cells depleted of RlmE showed a decrease in the 45S peak representing an assembly intermediate along with the appearance of a 50S peak dependent upon addition of RlmE [141]. This 50S peak was isolated and added to 30S ribosomal subunits and a peak corresponding to the formation of 70S ribosomes was subsequently detected [141]. Similarly, *in vitro* assembly of 30S ribosomal subunits from 16S rRNA and purified recombinant ribosomal proteins showed that 30S subunits were able to associate with tRNA and also competent to form 70S ribosomes in a sucrose gradient density ultracentrifugation experiment [142]. Together, these findings suggest that analyzing intermediates which accumulate in cells depleted of a respective ribosome assembly factor can advance the understanding of the particular factor's contribution to the formation of a functional ribosomal subunit [27].

The main evidence for YihA's involvement in ribosome biogenesis came from depletion of YsxC, which resulted in the accumulation of a 44.5S precursor of the 50S ribosomal subunit [59, 101]. In *S. aureus*, YsxC depletion also resulted in a reduction in the level of 70S ribosomes [59]. Potentially, YihA could act as a sensor of maturation on many different ribosome assembly intermediates. It has been shown experimentally that the k_{cat} of YsxC is increased 20-40-fold in the presence of the 50S ribosomal subunit, the 44.5S assembly intermediate from YsxC-depleted cells, or the 45S intermediates from YphC-depleted or RbgA-depleted cells [103]. Moreover, western blotting of fractions from sucrose gradient density ultracentrifugation showed that YsxC can be detected in the fraction containing 50S ribosomal subunits, but not to a significant extent in the fractions containing 70S ribosomes or 30S ribosomal subunits [59]. The evidence from

previous work suggested that YihA is required for the biogenesis of the large ribosomal subunit [23, 59, 67, 101, 143, 144].

To investigate a role of YsxC during ribosome assembly, studies were performed using sucrose gradient density ultracentrifugation to analyze extracts from wild type or YsxC-depleted *B. subtilis* cells in which the YsxC gene was removed from the genome but supplemented from a plasmid under the control of an inducible promoter [67, 101]. A 44.5S precursor of the 50S ribosomal subunit accumulated after the inducer was depleted in YsxC knockout cells expressing the inducible YsxC gene from a plasmid, but not in wild type cells or controls with inducer present [101]. A mass spectrometry analysis combined with a low resolution cryo-EM reconstruction of the 44.5S ribosome assembly intermediate isolated from YsxC-depleted cells revealed that this intermediate was missing the late-stage assembly ribosomal proteins uL16, bL27, bL28, bL33, bL35 and bL36 [103] (Figure 1.5). Accumulating evidence suggests that the base of the central protuberance as well as the A, P, and E tRNA-binding regions of the 50S ribosomal subunit reach their final folded state during the later stages of 50S ribosomal subunit assembly [103]. In the 44.5S intermediate found in YsxC-depleted cells, the rRNA of the central protuberance is folded differently than in the mature subunit [103]. Interestingly, helix 38 of the 23S rRNA, also referred to as the A-site finger, is shown to regulate entry of L16 into its binding site on the ribosome [103, 145]. Helix 38 forms direct contacts with the amino-acyl tRNA on the ribosome and the events of helix 38 restructuring, association with the ribosomal protein L16, followed by formation and maturation of the A-site, have been suggested to occur sequentially [1, 2, 103, 145]. This order of events in maturation of the 50S ribosomal subunit is consistent with the cryo-EM reconstruction of the 44.5S intermediate from YsxC-depleted cells, in which the A, P, and E sites were not completely folded [103]. These findings suggest that YihA may facilitate the binding of ribosomal proteins and remodelling of rRNA during the last stages of biogenesis of the 50S ribosomal subunit [23, 101].

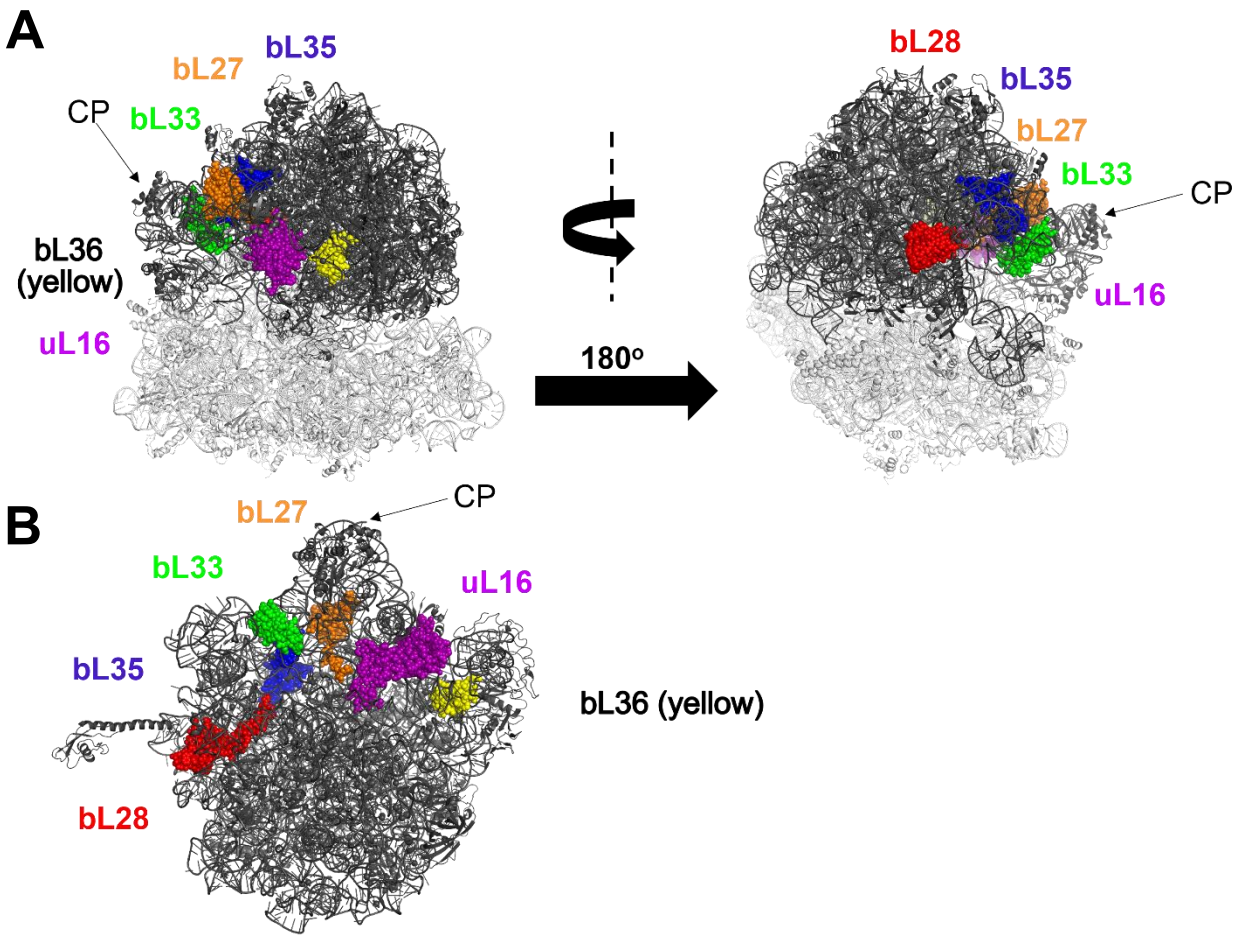


Figure 1.6. Structure of the *E. coli* 70S ribosome. The positions of ribosomal proteins that were identified to be missing from the 44.5S ribosome assembly intermediate isolated from BsYsxC-depleted cells are shown in color. The position of the central protuberance (CP) is indicated. A. Two views of the 70S ribosome with the small ribosomal subunit colored light grey and the large ribosomal subunit colored dark grey. B. Crown view of the large ribosomal subunit with the position of the central protuberance indicated by CP. The orientation shown here positions the large ribosomal subunit face that forms contacts with the small ribosomal subunit towards the reader. The cryo-EM structure is from PDB ID 4V9D.

1.4.5 YIHA AS A POTENTIAL ANTIBIOTIC AND ANTIMICROBIAL TARGET

Over the past twenty years there has been much interest in YihA as a putative target for new classes of antibiotics. Efforts to design synthetic inhibitors to specifically block YihA's activity have been hampered by a lack of understanding of YihA's functional cycle. Currently, there are very few known compounds that inhibit ribosome biogenesis directly, yet inhibition of ribosome

biogenesis provides an effective antibiotic target [146-151]. Ribosome biogenesis is a novel target for antibiotics such as lamotrigine that do not affect translation [152].

There are several criteria that YihA fulfills that make YihA a good antibiotic target. First, YihA is essential in most bacteria [31, 153]. YihA is expressed in many bacterial species including several species of pathogenic bacteria. YihA is required for survival of many species of clinically infectious bacteria but is not essential in *Mycoplasma genitalium* [154] and *Mycoplasma pneumoniae* [154]. Furthermore, earlier studies suggested that the *yihA* gene is absent from the genome of *Mycobacterium tuberculosis* [124]. However, more recent sequencing data deposited in the National Center for Biotechnology Information (NCBI) databases shows that the *yihA* gene is indeed present in the *Mycobacterium tuberculosis* genome. Alignment of multiple protein sequences from different species in the three kingdoms of life indicates that most residues of the G1-G5 motifs of YihA are conserved across all domains of life (Figure 1.2). However, this alignment also shows that there are regions that differ between the human isoforms of YihA and clinically relevant bacterial pathogens such as *S. aureus* and *Streptococcus pneumoniae*, including the N-terminal insertion sequences (Figure 1.2). Although YihA is universally conserved, bacterial YihA proteins show very low sequence identity to human YihA homologues (<30%, Figure 1.3). An effective antibiotic target protein should also not be essential in humans in order to avoid side effects of inhibiting the target. Moreover, YihA is not essential in *S. cerevisiae* [155], potentially suggesting that YihA might not be essential in most eukaryotes, including humans.

There is also some evidence that YihA is not essential for survival in trypanosomes. An RNAi depletion of mitochondrial YsxC (mtYsxC) was performed in *T. brucei* [113]. Tyc *et al.* prepared RNAi-competent cell lines in *T. brucei* [113]. The constructs were introduced into the genome under the control of a promoter by homologous recombination [113]. Depletion of mtYsxC in trypanosomes resulted in a very slight decrease in cell viability under growth conditions both in the presence and absence of glucose [113]. Growth of the cells under conditions of glucose

deprivation was tested in order to reveal a role of the mitochondria in carrying out cellular respiration in the absence of glycolysis [113]. However, a decrease in cell viability was seen to a greater extent with depletion of the two other eukaryotic proteins studied, RSM22 and PNKD-like protein, which exhibited a zero growth rate after 6-8 days of RNAi depletion [113]. This data suggests that mtYsxC might not be essential for cell survival, similar to findings for YihA in *S. cerevisiae [155].* Another consideration is that RNAi depletion of mtYsxC still resulted in a relatively high abundance of mRNA transcripts of mtYsxC in the cells, suggesting that the RNAi depletion of mtYsxC was not effective [113]. This finding also raises the question as to whether YihA could potentially be inhibited by antibiotics in humans while still retaining mitochondrial function.

YihA thus satisfies criteria for a possible antibiotic target and there is much interest in understanding its mechanistic function and binding site on the ribosome. At present, there is no published structure of bacterial YihA bound to the 70S ribosome or to any ribosomal subunits nor to any ribosomal subunit assembly precursors. Thus, understanding how YihA binds to pre-50S ribosomal intermediates could aid in the rational design of novel antibiotics that selectively inhibit bacterial ribosome biogenesis via targeting YihA. Since the G motifs of YihA are conserved across many species of bacterial and human homologues (Figure 1.2), targeting the rRNA recognition mechanisms of YihA may be a more promising and selective approach for inhibiting the cellular function of bacterial YihA. Since residues adjacent to the basic stretch of residues present in the C-terminal α-helix of YihA, and that have been shown to be important for binding of YihA to the ribosome, are not conserved across prokaryotes and eukaryotes, this region could provide a target for an antibacterial drug interfering with the binding of YihA to the 44.5S assembly intermediate (Figure 1.2). Such a drug would, in turn, likely not inhibit the human homologue of YihA, GTPBP8. Once the binding site of YihA on the 50S ribosomal precursor is identified, it will be possible to determine whether this site overlaps with sites targeted by other clinically used

antibiotics that target ribosomal particles. Although some clinically used drugs interfere with ribosome assembly as an effect of a secondary mechanism, there are currently very few antibiotics in the development pipeline whose primary target is ribosome biogenesis [152]. A drug that inhibits an essential putative ribosome biogenesis factor binding to a site on the ribosome that is not targeted by other antibiotics could provide a new strategy to combat resistance of bacterial pathogens.

Antimicrobial resistance is increasingly being acknowledged as a major public health threat. As the number of newly discovered antimicrobial drugs has declined dramatically since the 1990's, new strategies are being sought to develop antimicrobial drugs with novel mechanisms of action that may be effective against multi-drug resistant bacterial pathogens. Over 50% of antibiotics prescribed in the clinic today directly target the process of translation by the bacterial ribosome. YihA could be targeted directly by novel antibiotics as has been done for other universally conserved GTPases. EF-G is inhibited by the antibiotic fusidic acid, which is used to treat *S. aureus* skin infections in humans with the cream Fucidin H [156]. The antibiotic thiostrepton, which also inhibits EF-G, is approved for use in the veterinary clinic. Moreover, EF-Tu is inhibited by the antibiotics streptomycin, kirromycin, pulvomycin and GE2270 [157]. None of these antibiotics are approved for use in the clinic but a derivative of GE2270, LFF571 completed Phase 2 Clinical Trials in 2015 [157]. Even though EF-Tu is universally conserved, kirromycin shows much less inhibition of a mitochondrial homologue of EF-Tu and mitochondrial translation in eukaryotes compared to inhibition of translation in *E. coli* [158]. These lines of evidence suggest that some universally conserved GTPases are able to be targeted by clinically effective antibiotics.

A small library of 44,000 chemicals was screened for activity against YihA, yielding some hits [125]. Additionally, screening a library of compounds as inhibitors of YsxC has been reported [159]. YihA is a potential therapeutic target in malaria since YihA localizes to the apicoplast which

is a specialized organelle not found in many other organisms [114] Furthermore, a proteomics study of *P. falciparum* ribosomes suggest that there are significant differences in ribosomal protein composition between *P. falciparum* ribosomes and bacterial ribosomes [115]. Thus, because the YihA homologues in *Plasmodium falciparum Pf*YihA1 and *Pf*YihA3 translocate to the apicoplast, understanding the mechanism of these *Pf*YihA isoforms could be important towards the development of anti-malarial medicines. From a drug development standpoint, *Pf*YihA1, *Pf*YihA2 and *Pf*YihA3 show less than 30% sequence identity to EcYihA thus suggesting that there are significant differences in sequence and structure between bacterial YihA and its eukaryotic homologues that could be exploited in the design of new antimicrobial drugs.

1.4.6 CONSIDERATIONS IN DEVELOPMENT OF SMALL MOLECULE INHIBITORS TARGETING GTPASES

Several prokaryotic GTPases including EF-G, EF-Tu, Der and Obg have been considered as antibiotic targets and small molecule inhibitors have been developed to block their activity or function. Some of these small molecule inhibitors have been tested in various stages of clinical trials. In contrast, the development of small molecule inhibitors targeting the active site of the eukaryotic small GTPase Ras, for example, as potential chemotherapeutic agents has proven very challenging [160]. One breakthrough came when it was discovered that Ras dimerizes both *in vitro* and *in vivo* and that dimerization is important for the cellular function of Ras. Dimerization of Ras is considered a key therapeutic target and specifically it has been proposed that interfering with Ras's dimerization interface using small molecule inhibitors could abolish the ability of the effector Raf to promote cancer progression [161]. Size exclusion chromatography and Small Angle X-ray Scattering (SAXS) were used to show that Raf, the downstream effector of Ras, stimulates dimerization of Ras [162]. Molecular dynamics simulations have been used to

investigate the stabilization of two potential dimerization interfaces comprised of two different pairs of interacting helices in Ras [162].

A second model for the functional cycle of a GTPase has been demonstrated in which some GTPases exhibit dimerization-dependent GTP hydrolysis activity [163]. However, whether dimerization is part of YihA's functional cycle remains unknown.

1.5 RESEARCH OBJECTIVES

The cellular function of the universally conserved GTPase YihA has not been firmly established and very little is known about the functional cycle of YihA and its regulation. Understanding the biochemistry of YihA will provide the framework required for the systematic study of YihA's cellular role, enabling insight into the process of late-stage ribosome biogenesis. Although YsxC is a very small Ras-like protein that consists of only a single G-domain, depletion of YsxC results in the accumulation of a 44.5S intermediate that is a precursor of the 50S ribosomal subunit [103]. However, previous work led to controversy over the functional cycle of YihA, with one study reporting that the 50S ribosomal subunit and the 44.5S assembly intermediate from YsxC-depleted cells act as GAPs for BsYsxC, while another study found that EcYihA's GTP hydrolysis activity is stimulated two-fold by the 50S ribosomal subunit but ten-fold by the 70S ribosome [103, 164]. Further controversy arises over the rate of intrinsic GTP hydrolysis of YihA, with some evidence suggesting that EcYihA has no detectable intrinsic GTP hydrolysis activity while two other studies were able to measure rates of intrinsic GTP hydrolysis for EcYihA and BsYsxC [103, 119, 164]. Given the limited knowledge in the current literature regarding the regulation of YihA's functional cycle, an initial characterization of the functional cycle of YihA was undertaken here.

Initial work on EcYihA for this thesis focused on studying the kinetic parameters governing guanine nucleotide binding and hydrolysis by EcYihA (Appendix Chapter A2). Experiments within the Appendix in Chapter A2 showed that although the 50S ribosomal subunit fits the formal definition of a GAP for EcYihA in that the 50S ribosomal subunit increases the observed rate of GTP hydrolysis by EcYihA, the 50S ribosomal subunit is not a biologically relevant GAP for EcYihA (Appendix Figure A2.7A). Surprisingly, the work reported in the Appendix of this thesis revealed that EcYihA has no intrinsic GTP hydrolysis activity (Appendix Figure A2.7). The finding that EcYihA has no intrinsic GTP hydrolysis activity (Appendix Figure A2.7) agrees with one report published much earlier [119]. As the structure of a GTPase's G-domain, and specifically the G-motifs within the domain, are crucial for interaction with guanine nucleotides, further work sought to assess the YihA-nucleotide interaction using structural and biochemical studies. Examination of the crystal structure of EcYihA (PDB ID 1PUI) revealed that firstly, EcYihA is in the apo state. Secondly, the EcYihA crystal structure showed that there are two protein chains in the asymmetric unit. Thirdly, the available crystal structure of EcYihA suggested that residues of the P-loop from each monomer appear to be interacting across the dimerization interface with residues from the other monomer. As there is no research report accompanying the crystal structure of EcYihA, the approach taken was to search the literature for other GTPases having no intrinsic GTP hydrolysis activity. The literature search revealed that the β-subunit of the signal recognition particle receptor (SRβ) and ADP ribosylation factor (Arf), both of which are eukaryotic proteins, have no intrinsic GTP hydrolysis activity [165, 166]. The functional cycles of SRβ and Arf share two common characteristics. Firstly, both SRβ and Arf co-purify with nucleotide, which was suggested to provide a rationale for the lack of intrinsic GTP hydrolysis activity of these two proteins [165, 166]. Secondly, both SRβ and Arf exhibit nucleotide-dependent dimerization as part of their functional cycles [167, 168]. Therefore, two hypotheses are put forward here. The first hypothesis suggests that EcYihA co-purifies with nucleotide. The second hypothesis

suggests that EcYihA undergoes nucleotide-dependent dimerization as components of YihA's catalytic machinery for the GTP hydrolysis reaction could be provided or stabilized in trans. The technique of SAXS was chosen to investigate the second hypothesis as not only would SAXS provide a low resolution solution structure of EcYihA but it would also reveal whether EcYihA is globular, partially unfolded or completely unfolded in solution.

The main objective of this thesis is to study the solution conformation of YihA, including potential dimerization. Results obtained are provided in Chapter 2 and show that *in vitro* dimerization of YihA is supported by preliminary gel filtration chromatography results, structural predictions and small angle X-ray scattering (SAXS) data. The properties of the low resolution solution structure of EcYihA are also analyzed.

CHAPTER 2: DIMERIZATION AND STRUCTURE OF YIHA IN SOLUTION

2.1 INTRODUCTION

Instead of binding to a GAP to promote GTP hydrolysis, some G-proteins activate GTP hydrolysis through dimerization. In eukaryotes, the GTP hydrolysis activity of most members of the dynamin superfamily including guanylate-binding protein (GBP) and atlastin-1, is regulated through dimerization of the G domains [169, 170]. Nucleotide-dependent dimerization is a common mechanism as part of the functional cycle seen in several classes of nucleotide-binding proteins [171]. In GBP and atlastin-1, binding of nucleotides including GTP, GppNHp or GDP•AlF_4, promotes homodimerization and stimulates GTP hydrolysis [171]. In contrast, binding of GDP to GBP or atlastin-1 leads to monomerization of the enzyme [171].

The eukaryotic G-protein Ras is also able to dimerize, but dimerization of Ras has not been shown to affect GTP hydrolysis [172]. This is not surprising as the GTPase activity of Ras is known to be stimulated by GAPs such as p120GAP [173]. Dimerization can be an important step in the functional cycle of bacterial GTPases such as MnmE or ObgE [174-177]. The dimerization of MnmE's G domains is activated by potassium ions and positions the catalytic residue of *E. coli* MnmE (Glu282) to correctly orient the nucleophilic water molecule for catalysis [178]. In contrast, experimental evidence indicates that ObgE may dimerize *in vitro* but does not dimerize *in vivo* [179].

Moreover, GTPases can also dimerize in response to redox stress. In some GTPases and ATPases, dimerization inhibits nucleotide hydrolysis activity and could potentially provide a mechanism that regulates the cellular concentration of the enzyme in the GTP/ATP-bound "ON" state as part of the functional cycle. The bacterial NTPase YchF can form dimers upon exposure to oxidative stress by forming disulfide bridges and the ATPase activity of dimeric YchF is severely inhibited [55].

Interestingly, several of the available crystal structures of YihA homologues in the Protein Data Bank contain two or more chains in the asymmetric unit. Although several of the other GTPases in the TrmE/MnmE-Era-EngA-YihA-Septin family of GTPases, including MnmE and Septins, have been shown to dimerize or multimerize [180, 181], the dimerization of YihA has not been studied. In contrast, no evidence for dimerization of EngA/Der was demonstrated with analytical ultracentrifugation and potassium ions do not stimulate dimerization of EngA/Der [182].

The first approach used here to study the functional cycle of YihA was to investigate the interaction of EcYihA with guanine nucleotides and to measure EcYihA's intrinsic and ribosome-stimulated GTP hydrolysis (Appendix Chapter A2). Like the eukaryotic GTPases Arf and SRβ which both belong to the Arf family of proteins [183], EcYihA completely lacks intrinsic GTP hydrolysis activity, as demonstrated in Appendix Figure A2.7 [165, 166]. The finding that EcYihA has no intrinsic GTP hydrolysis activity raises the question as to how EcYihA hydrolyzes GTP. A further experiment described in the Appendix of this thesis showed that the 50S ribosomal subunit is likely not relevant as a GAP for EcYihA in a cellular context (Appendix Figure 2.7A). Interestingly, two other GTPases that exhibit no intrinsic GTP hydrolysis activity, Arf and SRβ, both co-purify with nucleotide and both form dimers that are regulated by the binding of specific guanine nucleotides [165-168]. For SRβ, dimerization is triggered by release of nucleotide [167]. There is a whole subgroup of GTPases that require dimerization to activate GTP hydrolysis rather than requiring a GAP [163]. The first evidence for dimerization comes from the crystal structure of EcYihA (PDB ID 1PUI). In the crystal structure of EcYihA, the unit cell contains two molecules of YihA and the G1-motif or P-loop, which coordinates the phosphate groups during nucleotide binding, is participating in contacts in the dimerization interface. This thesis hypothesizes that, firstly, EcYihA dimerizes in solution and that either nucleotide binding or release triggers dimerization or that dimerization regulates the interaction of EcYihA with nucleotide. Secondly, this thesis proposes that EcYihA co-purifies with nucleotide. Understanding the interplay between

potential YihA dimers and guanine nucleotides could be important for determining the functional role of YihA in the cell, especially if dimerization is an integral part of YihA's functional cycle. The oligomerization state of EcYihA can be monitored by the solution-state methods of SAXS and analytical gel filtration chromatography. In this chapter, structural and biochemical studies are used to evaluate the ability of YihA to dimerize in solution.

2.2 METHODS

2.2.1 REAGENTS, PLASMIDS AND CELL STRAINS

Cloning of *E. coli hflX* was done previously with the plasmid pET28a [48]. The plasmids containing the genes for *E. coli EF-Ts* and *E. coli hflD* in pCA24N were obtained from the ASKA Minus Collection [184]. Unless noted otherwise, all reagents were obtained from BioBasic.

The gene encoding *E. coli yiha* with an N-terminal 6-His tag was synthesized in the plasmid pET28a (GeneWiz). A kanamycin resistance cassette is encoded on pET28a. The sequence of the *E. coli yiha* gene was submitted to GeneWiz in order for GeneWiz to perform codon optimization and synthesis of the codon optimized gene.

2.2.2 TRANSFORMATION, PROTEIN OVEREXPRESSION AND PURIFICATION

E. coli HflX protein was overexpressed and purified as described previously [184]. *E. coli* EF-Ts protein was overexpressed and purified as described previously [185]. Bovine Serum Albumin (BSA) was obtained from BioBasic. *E. coli* EF-Tu protein (kindly provided by Luc Roberts) was cloned, overexpressed and purified by methods described previously [185]. *E. coli* HflD and EcYihA proteins were both overexpressed and purified as described below for EcYihA.

The plasmid expressing cloned YihA was transformed into *E. coli* BL21 DE3 cells (Invitrogen). EcYihA protein was overexpressed in *E. coli* BL21 DE3 cells in LB media in the

presence of 50 µg/mL kanamycin at 37°C. When cell cultures reached an optical density (OD) of approximately 0.6 at 600 nm, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the media to a final concentration of 1 mM. After three hours of induction with IPTG, cells were pelleted by centrifugation at 5000 x g in a Beckman Coulter centrifuge (Avanti JXN-26) using a JLA 8.1000 rotor (Beckman Coulter) and cell pellets were flash frozen and stored at -80°C. Overexpression was validated by analyzing cell culture samples collected every hour from the start of induction with IPTG and for three hours afterward. Samples were dissolved in 8 M urea and analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE). All SDS PAGE was performed at 80 V for 20 minutes followed by 150 V for 100 minutes in a Mini Protean 3 system (BioRad). Protein bands were visualized by staining of the gels with Coomassie Brilliant Blue dye.

Cells were opened by slowly stirring at 4°C with 1 mg/mL lysozyme in 7 mL of Buffer A containing 50 mM Tris-Cl pH 8.0 at 4°C, 70 mM NH_4Cl, 300 mM KCl, 7 mM $MgCl_2$, 10 mM imidazole, 15% glycerol, 7 mM β-mercaptoethanol and 1 mM phenylmethylsulfonylfluoride (PMSF) for 45 minutes. Additionally, 12.5 mg/g of sodium deoxycholate was added to the cell suspension followed by stirring for 30 minutes. Genomic DNA was digested by incubation of the cell suspension with DNase I followed by 4 cycles of sonication. The opened cells were centrifuged at 3000 x g for 30 minutes. Cellular debris was collected in the pellet and the resulting supernatant was centrifuged at 30,000 x g for 45 minutes to obtain the S30 cell extract.

A nickel affinity chromatography column was prepared by placing 5 mL of nickel sepharose resin (GE Healthcare) into a gravity flow column (BioRad) followed by equilibrating the resin with Buffer A. Nickel affinity chromatography was performed by incubating the prepared S30 extract with the nickel resin on a rotating shaker at 4°C for 1 hour. Washing of the nickel column was carried out with 15 column volumes of Buffer A followed by 20 column volumes of Buffer A containing 20 mM imidazole (Buffer B). YihA was eluted from the nickel column using ten washes

at 90% column volume per wash in Buffer A containing 300 mM imidazole (Buffer E). Fractions containing YihA were combined and concentrated using a 10 kDa molecular weight cut off (MWCO) ultrafiltration device to a final volume of 5 mL. The resulting solution was loaded on a prep grade Superdex-75 XK26/100 column (GE Healthcare) equilibrated with storage buffer either containing magnesium (50 mM Tris-Cl pH 7.5 at 4°C, 70 mM NH_4Cl, 300 mM KCl, 7 mM $MgCl_2$, 15% glycerol) or lacking magnesium (50 mM Tris-Cl pH 7.5 at 4°C, 70 mM NH_4Cl, 300 mM KCl, 15% glycerol) to exchange buffer and remove contaminant proteins by size exclusion chromatography at a flow rate of 1 mL/min. SDS PAGE was used to analyze the YihA fractions. Fractions containing high purity YihA were pooled and concentrated by ultrafiltration (see above). The concentration of YihA was determined by averaging the values determined using the Bradford Assay (BioRad) as well as from the absorbance using a spectrophotometer. For the latter, the concentration of YihA was determined at 280 nm using a molar extinction coefficient of 20 970 M^{-1} cm^{-1} calculated using the ProtParam tool of the ExPASY webserver [186]. The YihA purification was aliquoted, and aliquots were flash frozen and stored at -80°C.

2.2.3 NUCLEOTIDE-FREE PURIFICATION OF YIHA

Nucleotide-free EcYihA was prepared by incubation of 10,000 pmol of YihA in HSG TAE buffer (25 mM Tris-HCl, pH 7.5 at 4˚C, 300 mM NH_4Cl and 10 mM ethylene diamine tetraacetic acid (EDTA), 15% glycerol) at 37°C for 30 minutes. The presence of EDTA and heating allows chelation of metal ions in the buffer by EDTA, facilitating the release of bound nucleotide. Following the incubation with EDTA, YihA was centrifuged at 13,000 x g for 5 minutes at 4°C and then loaded on a Superdex-75 10/300 GL column (GE Healthcare) in order to allow separation of *apo* YihA from the released nucleotide and buffer exchange. Release of nucleotide from YihA

was confirmed by monitoring of the ratio of the absorbance at 260 nm to the absorbance at 280 nm and free nucleotide was detected as a peak that elutes last in the elution profile with a high 260 nm to 280 nm absorbance ratio. Size exclusion chromatography (SEC) was performed at a flow rate of 0.25 mL/min and was carried out on an Akta Pure fast protein liquid chromatography (FPLC) system (GE Healthcare). The buffer used for SEC was HSG TA Buffer (25 mM Tris-HCl, pH 7.5 at 4°C, 300 mM NH_4Cl, 15% glycerol).

2.2.4 DIMERIZATION ASSAY

EcYihA (10,000 pmol) was diluted to 50 µM in HSG TAKM7 buffer (50 mM Tris-Cl pH 7.5 at 4°C, 70 mM NH_4Cl, 300 mM KCl, 7 mM $MgCl_2$, 15% (v/v) glycerol). The resulting solution was either incubated at 37°C for 10 minutes, 15 minutes, or 1 hour or incubated at 4°C. A protein volume of 200 µL was subsequently loaded onto a Superdex-75 10/300 GL column (GE Healthcare) to separate monomers and dimers using an Akta Pure FPLC (GE Healthcare) operated at a flow rate of 0.25 mL/min. The column was developed at 4°C.

Areas of the peaks obtained in the chromatograms were calculated using the Unicorn 5.11 software provided by the FPLC manufacturer (Cytiva) with an automated method that required the “Calculate Baseline” peak integration function of Unicorn 5.11. The “Calculate Baseline” function was used to optimize the baseline of the chromatogram with implementation of the Classic algorithm available in Unicorn 5.11. The Classic algorithm is defined by four parameters that optimize the peak baseline, including the shortest baseline segment, the noise window, the max baseline level and the slope limit. These parameters provide the coordinates of a rectangle shape. Each region of the chromatogram peak needs to fall completely within the boundaries of the defined rectangle shape in order to be included when the baseline is optimized. The Classic

algorithm selects a shortest baseline segment and identifies regions of the peak that fall outside of the set parameters. Using the Classic algorithm, the shortest baseline segment, the noise window, the max. baseline level and the slope limit were manually adjusted until an accurate baseline was achieved. Defining the noise window appropriately prevents noisy regions along the peak baseline from being included in the peak integration while ensuring that all the peaks of interest are recognized as peaks in the peak area calculation. The slope limit is able to define when a curved baseline is necessary in order to obtain an accurate peak area. The optimal value for the shortest baseline segment was 0.05 mL, the noise window was set to 0.016 mAU, the max. baseline level was between 2.5 and 3.5 mAU and the slope limit was set to between 1.8 mAU/mL and 2.5 mAU/mL. These settings of the shortest baseline segment, noise window, max. baseline level and slope limit provided a “Valley to Valley” method to calculate peak areas for overlapping peaks where the beginning and end points of each peak are defined at the valley separating each set of two peaks.

The percentages of monomer and dimer in the YihA samples were determined based on the peak area using the following equations:

$$\%\, Dimer = \frac{Area\ of\ dimer\ peak}{\{(Area\ of\ dimer\ peak) + (Area\ of\ monomer\ peak)\}} \times 100$$

$$\%\, Monomer = \frac{Area\ of\ monomer\ peak}{\{(Area\ of\ dimer\ peak) + (Area\ of\ monomer\ peak)\}} \times 100$$

Standard curves were prepared for SEC with the Superdex-75 10/300 GLC column. The corresponding elution volumes were compared to the elution volumes of protein size standards to estimate the apparent molecular mass of the respective protein eluting at the particular elution volume in the elution profile. SEC of protein standards was carried out in HSG TAKM7 buffer. The

amount (pmol) of each protein standard required to obtain an absorbance at 280 nm of approximately 20 mAU in the elution profile was calculated and the protein standards were loaded accordingly onto a Superdex-75 10/300 GL column and resolved by SEC with a flow rate of 0.25 mL/min.

2.2.5 PURIFICATION OF EF-TU FOR USE AS A MOLECULAR WEIGHT STANDARD

E. coli EF-Tu protein was purified on a prep grade Superdex-75 XK26/100 column (GE Healthcare) using a flow rate of 1 mL/min. The method used for purification of EF-Tu was described previously [185]. EcYihA protein was purified from cells as described in section 2.2.2 of this thesis, except that the same buffer was used for purification as was used for purification of EF-Tu. The elution profile for the purification of EF-Tu was superimposed with the elution profile from the purification of YihA as a molecular weight reference in order to determine the molecular weight of peaks in the YihA purification.

2.2.6 ESTIMATION OF K_D OF DIMERIZATION OF YIHA

The peaks for the YihA monomer and dimer in the chromatogram collected at a wavelength of 280 nm in Figure 3.1B were used to estimate the K_D of dimerization for YihA. The areas of the two peaks representing the YihA monomer and dimer in Figure 3.1B were measured using ImageJ [187]. Peak width was estimated by drawing a line along the width of the peak at baseline and peak height was estimated by drawing a line connecting the peak maximum with the peak baseline. The length of the line estimating peak width or peak height was determined using the Measure tool in ImageJ. Peak area was obtained with the formula

$$Area = \frac{(Peak\ width\ at\ baseline)\ (Peak\ height)}{2} \qquad \text{(Equation 2.1)}$$

Let YY represent dimeric YihA and Y represent monomeric YihA. Peak width of YY was determined to be 2.70 mL and peak width of Y was 2.28 mL. Peak height of YY was measured as 10.49 mAU and peak height of A was 8.84 mAU. Using Equation 3.1, the peak area of YY was estimated as 14.16 mAU·mL and the peak area of Y was estimated as 10.08 mAU·mL. The optical path length of the flow cell in the detector of the FPLC system is 2 mm. The Beer-Lambert law is given by

$$A = \epsilon C L$$

where A is the absorbance, ϵ is the molar extinction coefficient, C is the concentration and L represents the path length.

Peak areas of YY and Y can be determined from the Beer-Lambert law as follows:

$$Peak\ Area = \varepsilon C L v \qquad \text{Equation 2.2}$$

where v represents the peak volume.

The concentration of the dimer is obtained by rearranging Equation 2.2:

$$[YY] = \frac{Peak\ Area}{\epsilon L v} = \frac{0.01416\ AU \cdot mL}{(20\ 970\ M^{-1}cm^{-1})(0.2\ cm)(2.70\ mL)} = 1.25\ \mu M$$

The extinction coefficient is for the monomer, so it is necessary to divide the concentration of YY by 2 to obtain the concentration of the dimer:

$$[YY] = 0.125\ \mu M \div 2 = 0.63\ \mu M$$

The concentration of the monomer is determined by

$$[Y] = \frac{Peak\ Area}{\epsilon L v} = \frac{0.01008\ AU \cdot mL}{(20\ 970\ M^{-1}cm^{-1})(0.2\ cm)(2.28\ mL)} = 1.05\ \mu M$$

The equation representing the dissociation of YihA dimers is given by

$$YY \leftrightarrows Y + Y$$

From these estimations of the concentration of YihA monomers and dimers, the K_D of dimerization for YihA can be calculated as

$$K_D = \frac{[Y][Y]}{[YY]} = \frac{[Y]^2}{[YY]} = \frac{(1.05\ \mu M)^2}{0.63\ \mu M} = 1.8\ \mu M$$

2.2.7 PRE-STEADY STATE FLUORESCENCE STOPPED FLOW ASSAY

Stopped flow experiments were carried out on a Kintek SF-2004 stopped flow apparatus (Kintek Corporation) along the lines of previous work [48]. YihA's tryptophan residues were excited at 280 nm and FRET between the tryptophans and mant group of the mant-nucleotide was measured after passing LG-400-F filters (Newport Filters) to detect the fluorescence of the mant group.

EcYihA at 2 μM and 60 μM mant-GDP were incubated separately in HSG TAK Buffer (50 mM Tris-Cl pH 7.5 at 4°C, 70 mM NH_4Cl, 300 mM KCl, 15% (v/v) glycerol) either at 37°C or at 4°C for 15 minutes. The fluorescence change for the bimolecular association of mant-GDP to YihA was measured by rapidly mixing 25 μL of 2 μM YihA with 25 μL of 60 μM mant-GDP at 20°C in HSG TAK buffer. The resulting fluorescence time course was fit in Table Curve (Jandel Scientific) with a one-exponential function (Equation 3.2):

$$F = F_\infty + A\ exp(-k_{app}t) \qquad \text{(Equation 2.3)}$$

Where F is the fluorescence emission at time t, F_∞ is the final fluorescence, k_{app} is the characteristic apparent rate constant and A is the signal amplitude.

2.2.8 NUCLEOTIDE HYDROLYSIS ASSAY

The observed rate of GTP hydrolysis catalyzed by EcYihA was determined as described previously [48] using a multiple turnover GTP hydrolysis assay to measure the formation of inorganic phosphate (P_i) containing incorporated ^{32}P from [γ-^{32}P]-GTP (Perkin Elmer). In order to promote the conversion of any nucleotide diphosphate to nucleotide triphosphate present in the sample and to ensure that multiple turnover GTP hydrolysis was not inhibited by the presence of nucleotide diphosphate in the reaction, a charging solution was prepared in which [γ-^{32}P]-GTP at ~ 200 dpm/pmol was incubated at 37°C for 15 minutes with 0.25 µg/µL pyruvate kinase (PK) and 3 mM phosphoenolpyruvate (PEP). The reaction buffer was either TAKM7 (50 mM Tris-Cl pH 7.5 at 4°C, 70 mM NH_4Cl, 30 mM KCl, 7 mM $MgCl_2$) or HSG TAK (50 mM Tris-Cl pH 7.5 at 4°C, 70 mM NH_4Cl, 300 mM KCl, 15% glycerol) or HSG TAKM7 (50 mM Tris-Cl pH 7.5 at 4°C, 70 mM NH_4Cl, 300 mM KCl, 7 mM $MgCl_2$, 15% glycerol). Each reaction included 5 µM YihA and 125 µM [γ-^{32}P]-GTP. Over the time course, 5 µL aliquots of the reaction were taken and transferred into 50 µL of a solution of 3 mM K_2HPO_4 and 1 M $HClO_4$ to quench the reaction. In order to facilitate the extraction of radiolabelled inorganic phosphate into the organic phase in the form of a phosphate-molybdate complex, 300 µL of 20 mM Na_2MoO_4 and 400 µL of isopropyl acetate were added to the quenched reaction. After mixing for 1 minute in a shaker (Eppendorf MixMate) and centrifuging at 17,000 x g for 5 minutes, half of the total inorganic phosphate in the organic phase was transferred into 2 mL of scintillation cocktail (MP Ecolite) and radioactivity was measured using a liquid scintillation counter (Perkin Elmer TriCarb 2800 TR). Based on the known specific activity of the [γ-^{32}P]-GTP solution, the pmol of inorganic phosphate formed in the GTP hydrolysis reaction was calculated from the measured activity of each sample. The concentration of inorganic phosphate formed in the reaction was determined by taking into account the volume of sample collected. A correction of the concentration of inorganic phosphate formed was carried out by subtracting the background GTP hydrolysis activity measured in the absence of YihA, and

the concentration of P_i was plotted against time. The data was then fitted with a linear regression where the slope of the best fit line yielded the apparent rate of GTP hydrolysis.

2.2.9 SMALL ANGLE X-RAY SCATTERING (SAXS) ANALYSIS OF YIHA

EcYihA samples were analyzed by High Pressure Liquid Chromatography-SAXS (HPLC-SAXS) at the B21 beamline, Diamond Light Source (Didcot, Oxfordshire, UK). To ensure the monodispersity of the YihA samples from which SAXS data was collected, the purified YihA was resolved by SEC on an Agilent 1200 (Agilent Technologies, Stockport, UK) HPLC controller. The HPLC system was set up to connect to a flow cell followed by an in-line absorbance detector. The specialized flow cell was a quartz capillary with a 1.5 mm diameter and 200 um thick walls [188]. Analysis was performed by loading 50 µL of EcYihA sample (without nucleotide added) at a concentration of 10 mg/mL (413 µM) or 20 mg/mL (826 µM) onto a Shodex KW402.5-4F column for SEC (Showa Denko America Inc.) developed at a flow rate of 0.16 mL/min at 4°C and equilibrated in 1X SAXS buffer (50 mM Tris-HCl, pH 7.5 at 4°C, 70 mM NH_4Cl, 7 mM $MgCl_2$, 148 mM KCl, 8% (v/v) glycerol). A control sample of 50 µL of 10 mg/mL (144 µM) BSA was loaded onto the column in 1X SAXS buffer prior to the loading of the experimental samples onto the column. Exposure to the beam of X-rays was carried out for 3 seconds for each frame from a set of frames resulting from resolving of YihA by SEC and the data was recorded. The distance from the sample to the detector for the beamline was 4.014 m and a wavelength of 1 Å was used. The range of the scattering wavevector, q, obtained from the beamline was 0.0032 to 0.34 $Å^{-1}$ and is given by:

$$q = \frac{4\pi sin\theta}{\lambda}$$

where θ is the scattering angle and λ indicates the wavelength [189].

Data analysis was carried out as described previously [190]. Following selection of the SEC sample peak, version 3.0 of ScÅtter [191] was used to subtract the background scattering intensities and to merge the data from each peak. The PRIMUS program [192] was utilized for Guinier analysis in order to calculate the forward scattering I(0) and the radius of gyration (R_g) and for Kratky analysis to verify that YihA is folded. Analysis of the data in GNOM [193] yielded the pair-distance distribution P(r) function from which was obtained the real-space R_g and the maximum particle dimension (D_{max}), which represents the radius at which the P(r) function approaches zero after achieving its maximum value. Low resolution *ab initio* models were determined from the resulting P(r) function GNOM output file with simulated annealing in DAMMIN [194]. Twelve low resolution structures were calculated with DAMMIN run in expert mode with P2 symmetry enforced (since a dimer was observed for YihA according to the molecular weight obtained in PRIMUS) and the maximum number of annealing steps was set to 500. The *ab initio* models were aligned, averaged and filtered in DAMAVER [195], providing a final representative model of YihA. Superimposition of the crystal structure of YihA (PDB ID 1PUI) as a dimer with the representative SAXS model of YihA was performed using DAMSUP [195] and visualized in PyMOL [196]. DAMSUP was used to determine the normalized spatial discrepancy (NSD) for the superimposition of the crystal structure with the representative SAXS model as described previously [197]. The CRYSOL package [198] was used with a setting of 99 harmonics to predict the theoretical scattering curves, the theoretical R_g, the theoretical D_{max} and the theoretical P(r) distribution for the crystal structures with PDB ID 1PUI (for the monomer or dimer of YihA) and with PDB ID 2CXX (for the trimer of YihA). Comparison of the fit of the theoretical scattering curves to the experimental SAXS data was carried out by computing the chi^2 value in CRYSOL for each fit as follows:

$$reduced\ chi^2 = \frac{1}{m-1}\sum_{k=1}^{m}\left(\frac{I_{fit}(q) - I_{exp}(q)}{SE_{I_{\exp(q)}}}\right)^2$$ Equation 2.4

where $I_{fit}(s)$ represents the intensity of the fitted model, $I_{exp}(s)$ represents the experimental intensity, $SE_{I_{\exp(s)}}$ represents the error of the experimental intensity and m represents the number of experimental values. For the *ab initio* model constructed with DAMMIN and DAMAVER, the R_g and D_{max} values were calculated with the program HYDROPRO [199].

Following calculation of the Porod volume by Porod-Debye analysis, the protein density of EcYihA was determined with the following equation:

$$Protein\ Density\ = Molar\ Mass \times \frac{1}{Avogadro's\ number} \times \frac{1}{Porod\ Volume}$$ Equation 2.5

In addition to the molecular weight of YihA calculated in PRIMUS with the Volume of Correlation method, the MM_{QP} method, the MoW method and the Size and Shape method, an additional method (the Volume method) of calculating the molecular weight of YihA was carried out in which the molecular weight was estimated from the volume of dummy atoms in the *ab initio* model [200] according to the Equation 3.6.

$$Molecular\ Weight \approx \frac{DAMAVER\ model\ volume}{1.7}$$ Equation 2.6

The Ensemble Optimization method 3.0 (EOM 3.0) [201, 202] was used to model a structure of EcYihA including the flexible regions. The EcYihA dimeric crystal structure PDB ID 1PUI was used to prepare the models and the two intersubunit salt bridges found in the crystal

structure of EcYihA were used as constraints between the two monomers in the dimer. 100,000 models were calculated based on the SAXS scattering curve. The structure of EcYihA that was submitted to EOM was first made complete by modelling in the webserver AllosModFoXS [203, 204] that uses the Modeller [205] program to fill in missing loops. The models of EcYihA built by AllosModFoXS were optimized to fit the experimental SAXS scattering curve for EcYihA.

2.3 RESULTS

2.3.1 YIHA DIMERIZES IN SOLUTION

Initial experiments to purify nucleotide-free EcYihA by heating to 37°C in the presence of 10 mM EDTA yielded three peaks in the chromatogram (Figure 2.1A). The peak that eluted last had a higher 254 nm to 280 nm ratio and could potentially represent the free nucleotide (Figure 2.1A). However, the two peaks that eluted first had a higher 280 nm to 254 nm ratio, consistent with the presence of protein (Figure 2.1A). The molecular weight of EcYihA is 24.2 kDa (Figure 2.1A-B). Seeing the presence of two peaks in the elution profile that elute at the approximate volumes corresponding to a monomer and a dimer led to the hypothesis that YihA is able to dimerize in solution. There was no previously published evidence for dimerization of YihA in solution.

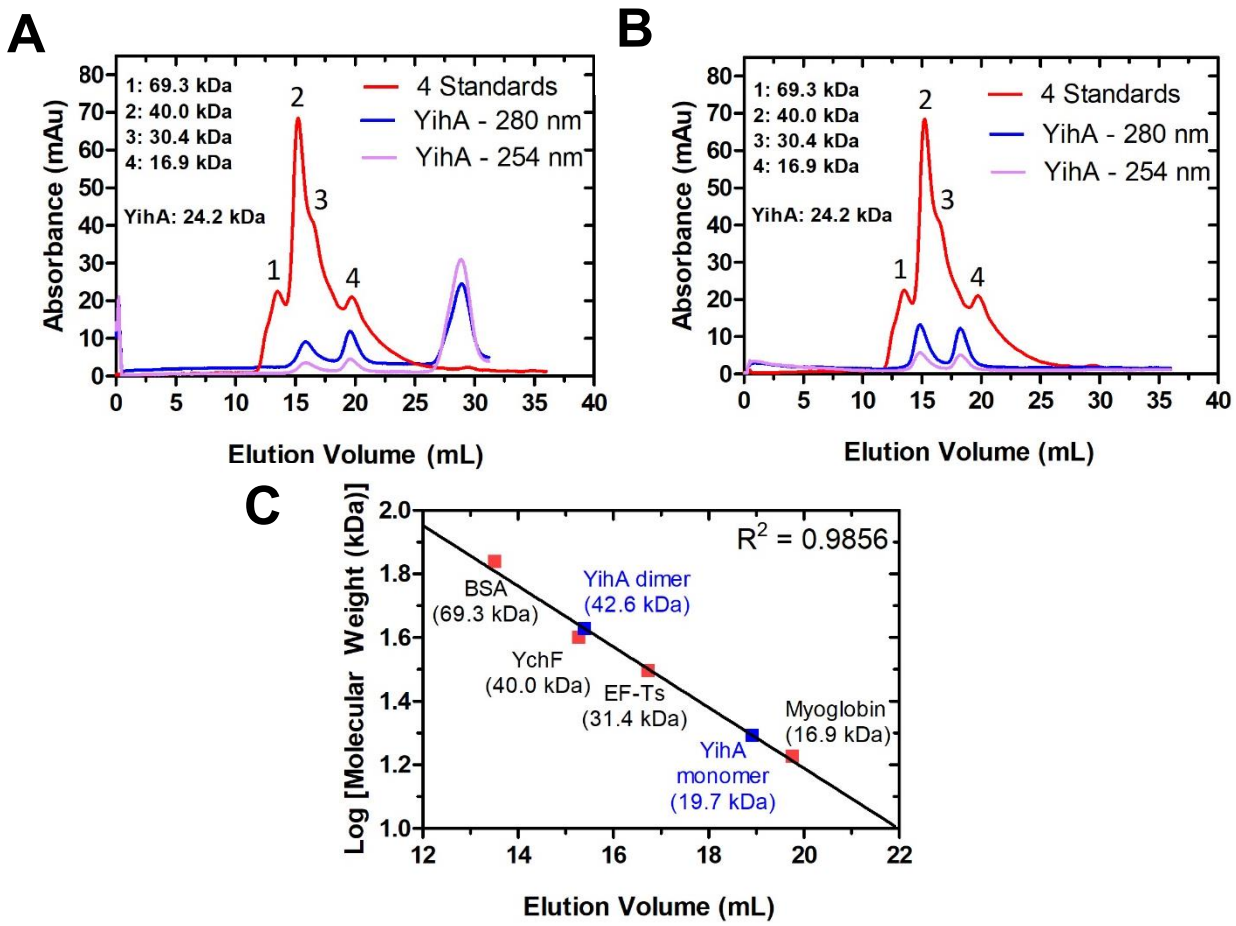


Figure 2.1. Nucleotide-free purification of YihA. This experiment provides the first evidence that YihA dimerizes in solution**.** (A) Chromatogram for 10,000 pmol of YihA (blue trace) treated with the nucleotide-free purification or protein standards (red trace) loaded onto the Superdex-75 column. (B) Chromatogram for the same assay as in (A) but no EDTA was added to the incubation buffer. The red trace is measured at an absorbance of 280 nm for A and B. For both panels A and B, peak 1 corresponds to BSA, peak 2 corresponds to YchF, peak 3 corresponds to EF-Ts and peak 4 corresponds to Myoglobin. (C) Standard curve for the SEC of the four standards shown in (A) and (B).

Furthermore, heating of YihA to 37°C without incubation with EDTA as shown in Figure 2.1B resulted in the presence of the YihA peak that elutes close to the elution volume of standard B representing YihA dimers and the presence of the YihA peak that elutes close to the elution volume of standard D representing YihA monomers. However, the peak that elutes last at approximately 28 mL in Figure 2.1A representing free nucleotide was missing in Figure 2.1B,

suggesting that nucleotide release does not occur without incubation in EDTA. Moreover, the lack of the peak that elutes at approximately 28 mL along with the presence of the dimer peak in Figure 2.1B suggested that dimerization of YihA can occur without the release of nucleotide. However, as the experiment in Figure 2.1 was not performed without heating to 37°C it remains undetermined whether heating to 37°C is part of the mechanism required for YihA dimerization.

Four protein standards were analyzed in the same buffer used in the nucleotide-free experiment, showing that the first peak was in the size range of a dimer and the second peak was in the size range of a monomer (Figure 2.1A-B). The observed molecular weight of the first peak to elute was estimated to be 42.6 kDa and the observed molecular weight of the second peak to elute was estimated to be 19.7 kDa.

A standard curve was prepared with peaks 1-4 (Figure 2.1A-B) that showed a linear relationship ($R^2 = 0.9856$) between the logarithm of the molecular weight of each standard and its elution volume (Figure 2.1C). However, peak 2 (YchF) and peak 3 (EF-Ts) were poorly resolved with peak 3 (EF-Ts) eluting as a barely detectable shoulder to peak 2 (YchF) (Figure 2.1A-B). Therefore, in subsequent experiments to investigate the conditions required for YihA dimerization, the YchF standard was replaced with HflX and the Myoglobin standard was replaced with HflD (Figure 2.1A-C, Figure 2.2A-C) and a new standard curve was prepared. All standards used in both standard curves were *E. coli* proteins. HflX is 50.5 kDa in size which is close to the theoretical size of a YihA dimer (48.4 kDa) and HflD is 25.4 kDa in size, which is very close to the theoretical size of a YihA monomer (24.2 kDa). Thus, the use of HflX and HflD as standards (Figure 2.2A-C) enabled a quick visual confirmation in each experiment that the peaks assigned as the YihA monomer and dimer were the correct size for a monomer and dimer according to the standard curve in Figure 2.2C. In Figure 2.2 and Figure 2.3, the peak corresponding to the Bovine Serum Albumin (BSA) protein standard in the black trace with a size of 69.3 kDa is designated as “Peak S”. In order to have a size marker corresponding to the size of a YihA dimer, which is

approximately 50 kDa, a standard consisting of the protein HflX was analyzed by chromatography (Figure 2.2, Figure 2.3). The HflX standard is seen in the black trace as the peak labelled “Peak H” and has a size of 50.5 kDa (Figure 2.2, Figure 2.3). Peak S and Peak H occur only in the experiments presented in Figure 2.2 and Figure 2.3.

A standard curve containing BSA, HflX, EF-Ts and HflD was constructed that displayed a linear fit (R^2 = 0.9952) of the logarithm of the molecular weight of the standards plotted against the elution volume of the standards (Figure 2.2C). The observed molecular weight of the YihA dimer peak was estimated from the standard curve to be 46.8 ± 2.5 kDa and the observed molecular weight of the YihA monomer peak was estimated to be 25.6 ± 1.5 kDa (Figure 2.2C). These observed molecular weights provide further confirmation that the two peaks obtained from SEC with purified YihA (Figure 2.2A-B) correspond to a monomer and dimer of YihA and that YihA does dimerize in solution.

SDS PAGE confirmed that the purified YihA used to study dimerization of YihA via size exclusion chromatography was free of any other major contaminating proteins (Figure 2.2D).

Because the nucleotide-free purification protocol requires a heating step and also requires incubation in a buffer containing EDTA and chromatography in a buffer lacking magnesium and potassium, it seemed possible that one or more of these conditions might be triggering dimerization of YihA. Accordingly, the ability of YihA to form dimers in response to heating at physiological temperature was tested. EcYihA was heated to 37°C for 30 minutes and a control was performed in which YihA was not heated. This experiment yielded two peaks for the heated sample, the first of which was approximately 50 kDa in size, overlapped with the peak for the standard HflX and could represent a dimer and the second of which was approximately 25 kDa in size, overlapped with the peak for the standard HflD and could represent a monomer as the size of the purified His-tagged YihA used in these experiments is 24.2 kDa (Figure 2.2B). The non-

heated control showed a single peak consistent with a lack of dimerization (Figure 2.2A). These experiments led to the conclusion that YihA dimerizes in solution at 37°C.

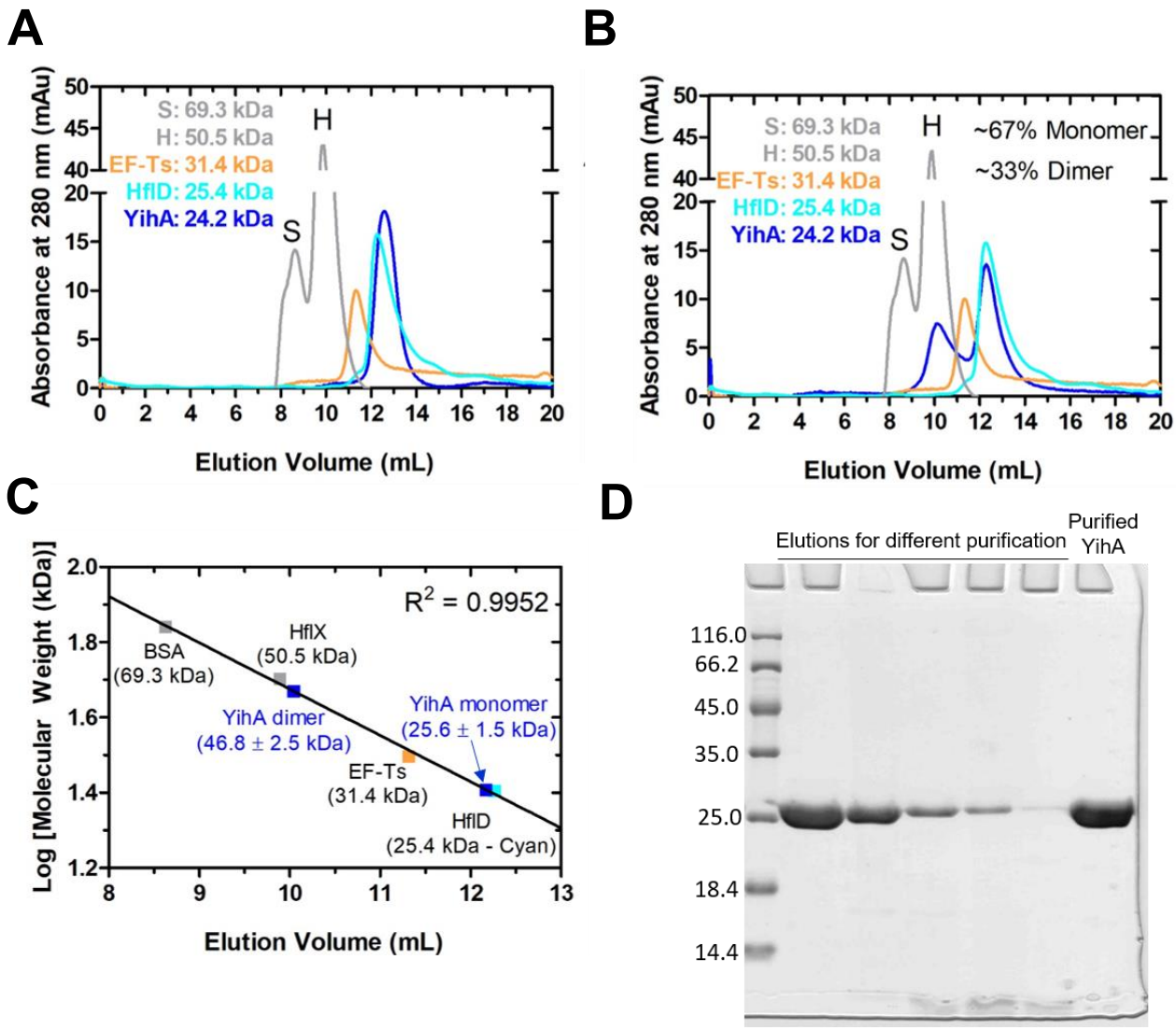


Figure 2.2. YihA dimerizes in solution at 37°C. (A) Chromatogram for 10,000 pmol of YihA (blue trace) loaded onto the Superdex-75 column. (B) Chromatogram for 10,000 pmol of YihA (blue trace) after YihA was heated at 37°C for 30 minutes then loaded onto the Superdex-75 column. The approximate percentage of monomer or dimer is given in the top right hand corner of each panel. Protein size standards are HflD (purple), EF-Ts (red), HflX (Peak H, black) and BSA (Peak S, black) and their sizes in kDa are indicated. The percentages of monomer and dimer were calculated as described in section 3.2.4. (C) Standard curve for the SEC of the four standards shown in (A) and (B). (D) A 15% SDS PAGE analysis of the elution fractions from SEC of the YihA purification used in experiments A-C. The lane marked as “Purified YihA” was loaded with 333 pmol of YihA from a different purification as a size marker.

Without further controls, it is not possible to determine whether the dimerization of YihA is a temperature-dependent switch: no dimerization under 37°C and dimerization at or above 37°C. A control for dimerization that was not done but that would provide further information towards understanding whether YihA dimerization is a temperature-dependent switch is incubation of YihA for 10 minutes at room temperature followed by SEC to detect dimerization. Incubation of YihA for 10 minutes at room temperature with analysis by SEC would show whether dimerization can occur at a temperature below 37°C that still permits growth of *E. coli* and whether dimerization of YihA could be correlated with permissive growth temperatures.

Further evidence suggests that YihA is purified from cells exclusively as a monomer, as YihA (with a size of 24.2 kDa) elutes later than EF-Tu (with a size of 44.3 kDa) in the chromatogram obtained with the Superdex-75 XK26/100 column (Figure 2.3). The later eluting peak represents free YihA and free EF-Tu, whereas the earlier eluting peak shows a higher 254 nm to 280 nm ratio and likely contains nucleic acid bound to a larger complex that includes YihA or EF-Tu, respectively (Figure 2.3). There is only one later eluting peak for each of YihA and EF-Tu in which the the absorbance at 280 nm is much higher than the absorbance at 260 nm suggesting that YihA exists entirely as a monomer at this stage of purification (Figure 2.3). The concentration of the purified YihA peak labelled Y in Figure 2.3 is estimated from the method described in section 2.2.6 of this thesis to be 11.9 μM, suggesting that the concentration of 11.9 μM is below the K_D of dimerization for YihA. The concentration of YihA in an *E. coli* cell is between 1.0 μM and 2.0 μM in log phase (Table 2.1). Therefore, YihA elutes from the column as just a monomer in peak Y, predicting that YihA does not spontaneously dimerize in an *E. coli* cell. Alternatively, there could be another cellular factor required for YihA's dimerization and the other cellular components that are required for dimerization could have been purified away prior to SEC by the nickel sepharose chromatography step. Additionally, the K_D for YihA's dimerization may be above the concentration in the YihA peak labelled Y and as such in the cell a process could occur

to increase the local concentration of YihA to a level where it is above the concentration of the purified YihA peak labelled Y and thus above the potential K_D for YihA's dimerization, allowing YihA to dimerize in the cell. For example, formation of nanoclusters by K-Ras at membranes facilitates K-Ras dimerization through increasing the concentration of K-Ras in localized regions of membranes and prevents mobility of K-Ras thus positioning K-Ras monomers to interact, forming dimers [206]. Another possible explanation for why EcYihA purifies from cells as a monomer could be that free EcYihA has a picomolar affinity for GDP and co-purifies with GDP if it is not bound to another factor or complex. Potentially, EcYihA bound to GDP may not be able to dimerize. Previous work on other GTPases that dimerize suggested that GTP binding precedes dimerization [163]. The explanation that EcYihA purifies from cells as a monomer because YihA co-purifies with GDP and the GDP-bound form of EcYihA does not dimerize is supported by preliminary evidence in Figure 2.1.

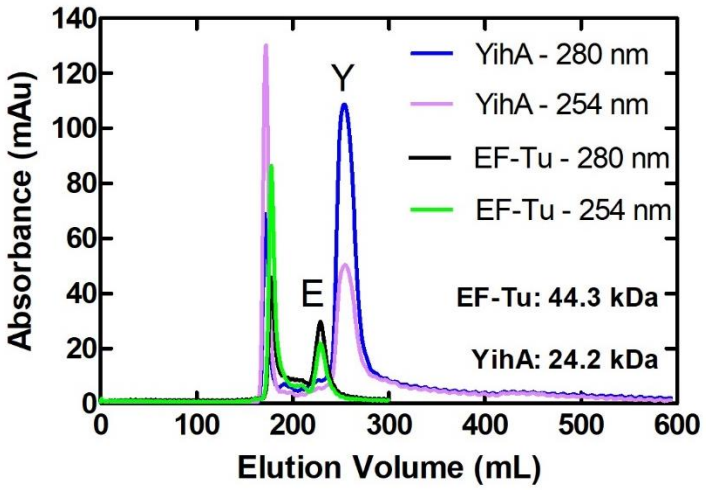


Figure 2.3. Evidence that YihA is purified as a monomer. Chromatogram for YihA where the YihA peak labelled Y has a size of 24.2 kDa. The blue trace is for 280 nm and the lilac trace is for 254 nm. The YihA purification was loaded onto a Superdex-75 XK26/100 SEC column after nickel affinity purification. EF-Tu was purified by nickel affinity chromatography and loaded onto a XK26/100 Superdex-75 SEC column. The EF-Tu peak labelled E has a size of 44.3 kDa. The chromatogram for EF-Tu is indicated as the black trace for 280 nm and the green trace for 254 nm.

2.3.2 PRELIMINARY CHARACTERIZATION OF THE BIOCHEMICAL PROPERTIES OF YIHA MONOMERS AND DIMERS

Pre-steady-state stopped flow analysis of YihA was performed without heating or with heating of YihA to 37°C for 15 minutes. These conditions were intended to promote dimerization for YihA that was heat-treated while there would be no dimer present for YihA that was not heat-treated. YihA was rapidly mixed with mant-GDP. It was found that there was no fluorescence change for the association of YihA with mant-GDP for the YihA that was heat-treated (Figure 2.4). In contrast, the YihA that was not heated to 37°C showed an increase in fluorescence over time suggesting association of mant-GDP with YihA (Figure 2.4). The increase in fluorescence represented a monophasic time course and was fitted to a single exponential function (Figure 2.4). This evidence suggests that the active form of YihA is the monomer and the inactive form is the dimer. Although the stopped flow experiments with YihA heated to 37°C for 15 minutes were performed with what appears to be 60% to 70% monomer (Figure 2.1, Figure 2.2), there was no change in fluorescence detected although a relative fluorescence increase of 0.006 that should be detectable is expected if 60% of YihA is monomeric and only monomers but not dimers are able to bind to nucleotide (Figure 2.4). In contrast, when YihA was incubated at 4°C for 15 minutes prior to performing stopped flow, a relative fluorescence increase of 0.01 was observed, suggesting the binding of YihA to nucleotide (Figure 2.4). When YihA is heated to 37°C *in vitro*, it completely loses the ability to bind nucleotide, suggesting that YihA monomers and possibly also dimers undergo a temperature-dependent conformational change at 37°C that prevents YihA from binding nucleotide and prevents FRET between the tryptophans of YihA and the mant group of mant-GDP (Figure 2.4).

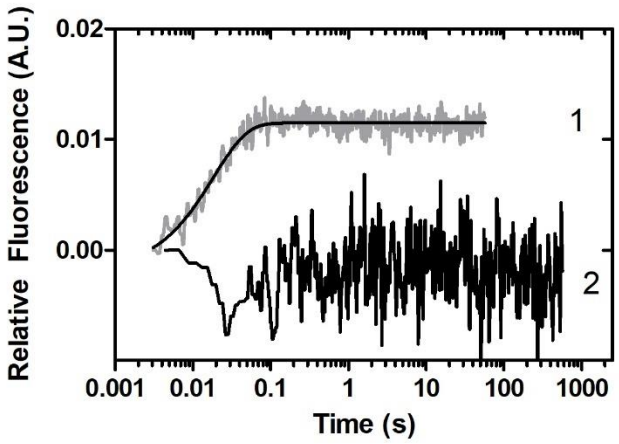


Figure 2.4. A temperature-dependent conformational change affects nucleotide association with YihA. Stopped flow assay measuring the association of 0.85 μM YihA rapidly mixed with 30 μM of mant-GDP. Excitation was at 280 nm. YihA was either not heated prior to mixing with mant-GDP (1) or was heated at 37°C for 15 minutes prior to rapid mixing with mant-GDP (2).

Initial experiments were performed to measure the GTP hydrolysis of YihA in parallel at a temperature that promotes dimerization and at a temperature well below physiological temperature. I reasoned that at 37°C there would be YihA present in the reaction as dimer and that at 20°C there should be little or no dimer present. These experiments measured the intrinsic GTP hydrolysis of YihA at 37°C or at 20°C (Figure 2.5A-B) and showed that there was no effect of decreasing the temperature on the intrinsic GTP hydrolysis catalyzed by YihA (Figure 2.5A-B) as EcYihA monomers appear to have no intrinsic GTP hydrolysis activity that is detectable *in vitro*. Since the lack of intrinsic activity could be due to the instability of YihA, the effect of using a buffer containing high salt (300 mM KCl) and also 15% glycerol was tested. The high salt with glycerol buffer either contained magnesium (HSG TAKM7) or no magnesium (HSG TAK). There was no detectable GTP hydrolysis activity for any of the buffer conditions tested (Figure 2.5A-B). Even though YihA is able to bind nucleotides when it is incubated at 4°C followed by incubation at 20°C in the stopped flow, YihA still requires a GAP to stimulate GTP hydrolysis at 20°C.

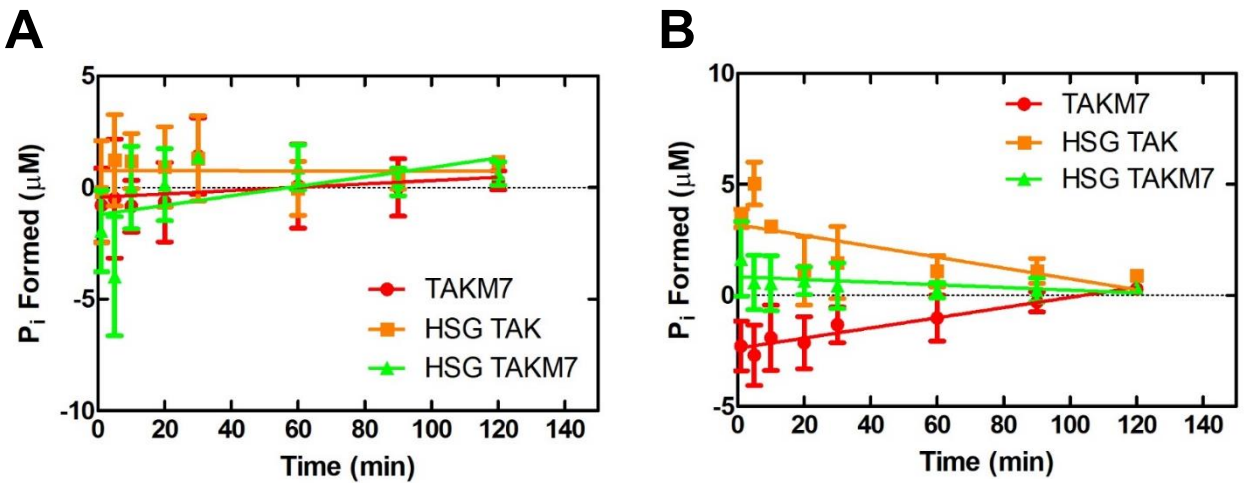


Figure 2.5. Reducing the temperature has no effect on the intrinsic GTP hydrolysis activity of YihA. (A) GTP hydrolysis assay in three different buffers performed at 37°C. (B). GTP hydrolysis assay in three different buffers performed at 20°C.

2.3.3 INITIAL PREDICTIONS THAT YIHA DIMERIZES *IN VIVO*

Data from four proteomics studies in *E. coli* and *B. subtilis* was analyzed. The average EcYihA concentration in the cell varies from 0.2 to 2 μM, depending on the phase of cell growth with a range from 0.1 to 4.9 μM (Table 2.1). The K_D of dimerization of YihA was estimated from the SEC dimerization data to be 1.8 μM (Table 2.1) (See section 2.2.6 for the method and calculations used to predict YihA's K_D of dimerization). Therefore, dimerization of YihA can be predicted to occur *in vivo* based on all of the proteomics studies carried out in exponential phase cells but not in cells in stationary phase, as the cellular concentration of YihA is predicted to be too low for dimerization to occur in stationary phase (Table 2.1). During the transition from exponential to stationary phase in bacteria, the expression of proteins required for growth such as YchF, EF-Tu and YihA, is mainly switched off (Table 2.1). Intriguingly, the expression of another bacterial ribosome assembly factor, Obg, and also the expression of bacterial ribosome proteins is regulated by the stage of cell growth, with high expression in exponential phase and

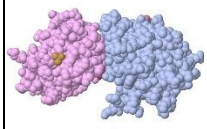
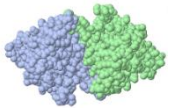
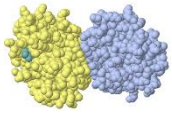
decreasing expression with the approach of and entry into stationary phase [31, 175]. YchF is most highly expressed in early exponential phase (Harland Brandon, PhD Thesis).

The stoichiometry of protein subunits in the asymmetric unit of a crystal structure does not always indicate the correct physiological quaternary structure of a protein. Moreover, the biological assembly (quaternary structure) assigned to all structures in the Protein Data Bank can in some cases be assigned wrongly without experimental evidence from studies in solution [207-209]. It is important to use other tools to predict the quarternary structure of a protein. Accordingly, predictions and models of the quarternary structure of YihA were generated for each crystal structure of YihA with the PDBePISA web server [210, 211]. PDBePISA provides an analysis of the crystal packing and prediction of the protein interface.

Table 2.1 Prediction that YihA dimerizes *in vivo.* The estimated K_D of dimerization for YihA is shown and used to predict whether YihA dimerizes *in vivo*. YihA expression in *E. coli* cells at different phases of the growth cycle is compared to that of EF-Tu, YchF, and RMF.

Estimated K_D of dimerization of YihA (μM)	1.8					
	Morimoto *et al*., 2002, Microbiol. [124]	**Wisniewski and Rakus, 2014, Data in Brief [212]**	**Schmidt *et al*., 2016, Nat. Biotech. [213]**	**Schmidt *et al*., 2016, Nat. Biotech. [213]**	**Schmidt *et al*., 2016, Nat. Biotech. [213]**	**Ishihama *et al*., 2008, BMC Genom. [214]**
Species	*B. subtilis*	*E. coli*	*E. coli*	*E. coli*	*E. coli*	*E. coli*
Average [YihA] in the cell (μM)	1.5	0.2	0.3	0.5	2.0	1.0
Range of [YihA] in the cell (μM) (Based on the range of cell volumes)	0.9 – 3.8	0.1 – 0.5	0.2 – 0.9	0.3 – 1.3	1.2 – 4.9	0.6 – 2.6
Does YihA dimerize in the cell?	Yes	No	No	No	Yes	Yes
Phase of Cell Growth	Log phase	Stationary phase	Stationary phase – 1 Day	Stationary phase – 3 Days	Log phase	Log phase
[EF-Tu] in the cell (μM)	Not available	33.8 (20.8 – 84.4)	9.3 (5.7 – 23.3)	19.6 (12.1 – 49.0)	156.5 (96.2 – 391.4)	135.0 (82.9 – 337.4)
[YchF] in the cell (μM)	Not available	0.6 (0.4 – 1.5)	0.3 (0.2 – 0.7)	0.3 (0.2 – 0.8)	7.4 (4.5 – 18.4)	1.7 (1.0 – 4.2)
[RMF] in the cell (μM) (Ribosome Modulation Factor)	Not available	Not available	5.3 (3.3 – 13.4)	2.7 (1.7 – 6.8)	0.3 (0.2 – 0.8)	Not available

Table 2.2 Prediction of the oligomeric state of YihA crystal structures in the PDB. The PDBePISA web server was used for the predictions [210, 211].

Species	PDB ID	Number of chains in unit cell of crystal structure	Oligomeric State (Assigned in PDB)	Oligomeric State Predicted with PDB ePISA	Stability of Predicted Oligomer (from PDBePISA)	Model of Oligomeric Assembly from PDBePISA
E. coli	1PUI	2	Monomer	Dimer	Stable	
Klebsiella pneumoniae	7SZS	2	Dimer	Dimer	Stable	
Neisseria gonorrhoeae	5UCV	2	Dimer	Dimer	Not Stable	
Thermotoga maritima	3PQC	2	Monomer	Monomer	N/A	N/A
Thermotoga maritima	3PR1	1	Monomer	Monomer	N/A	N/A
B. subtilis	1SUL	2	Monomer	Monomer	N/A	N/A
B. subtilis	1SVI	1	Monomer	Monomer	N/A	N/A
B. subtilis	1SVW	2	Monomer	Monomer	N/A	N/A
Burkholderia Thailandensis	4DHE	2	Monomer	Monomer	N/A	N/A
Pyrococcus horikoshii	2CXX	3	Monomer	Monomer	N/A	N/A

2.3.4 MODEL OF ECYIHA DIMER WITH NUCLEOTIDE BOUND

The nucleotide GTP and the magnesium ion were modelled into the crystal structure PDB ID 1PUI of the EcYihA dimer by performing a structural alignment in PyMOL of the EcYihA dimer structure with two monomers of the BsYsxC PDB ID 1SVW structure that is bound to GTP (Figure 2.6). Each 1SVW YihA monomer was aligned separately to the EcYihA dimer structure. Consistent with the data presented in this thesis, the crystal structure of the EcYihA dimer is not bound to nucleotide, suggesting that the dimer is not the active form of YihA. According to PDBePISA, the dimerization interface of the YihA dimer seen in the EcYihA crystal structure contains a buried surface area of 2,167 $Å^2$. The dimerization interface of this model of the crystal structure of the EcYihA dimer contains several conserved residues (Figure 2.7, Figure 2.8). Interestingly, the residues that form a conserved positively charged patch on the surface of each YihA monomer are located within the dimerization interface of the EcYihA dimeric crystal structure on both YihA monomers (Figure 2.8). Additionally, analysis of the crystal structure of the EcYihA dimer in PDBePISA predicts that R34, R119, H120 and K149 are situated within the YihA dimer interface and are among the residues in the interface that are largely buried. The basic patch on each YihA monomer is not expected to be involved in stabilizing the dimerization interface as the positive charges on each monomer would lead to electrostatic repulsion. However, the R119 residues on both monomers in the dimeric EcYihA crystal structure are involved in hydrogen bonds (Figure 2.7, Figure 2.8). These positively charged patches may be playing a role in a different, yet unknown process in YihA's functional cycle. Nevertheless, electrostatic interactions may be contributing to the interactions in the dimer interface as there are two salt bridges formed in the dimer interface between K149-D124 and between D117-K123 in the dimeric EcYihA crystal structure (Figure 2.8). These two salt bridges may function to stabilize the dimerization interface of the EcYihA dimer (Figure 2.8). Although there are potential unfavorable electrostatic interactions in the dimerization inteface of the EcYihA crystal structure, the electrostatic energy

of this dimerization interface is calculated by the PPCheck webserver to be -0.46 kJ/mol, which is slightly favorable. Another property of the dimeric crystal structure of EcYihA is that it is not completely symmetrical (Figure 2.6).

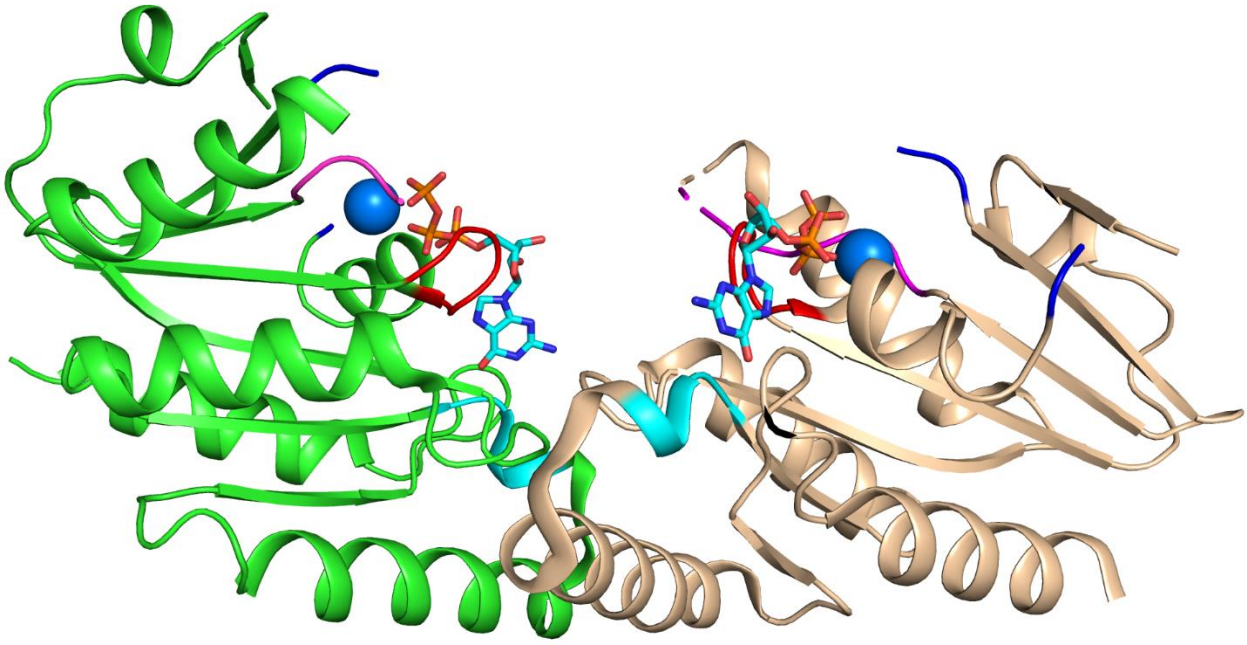


Figure 2.6. The active sites of the EcYihA dimer are accessible to nucleotide. The nucleotides which were absent from the structure were modelled into the crystal structure of EcYihA (PDB 1PUI) using PyMOL. The P-loop is shown in red and Switch I is in blue. Switch II is shown in magenta and the G4 motif is colored in cyan. EcYihA forms a symmetrical dimer in crystal structure 1PUI. B. Amino acid residues 51-62 (most of Switch I) and residues 84-86 (part of Switch II) are missing in the 1PUI structure. PDB ID 1SVW monomers were individually aligned with PDB ID 1PUI to provide the binding site of the GTP nucleotide and to model the locations of the magnesium ions. The magnesium ions are shown as blue spheres. Monomer A is shown in green and monomer B is shown in tan.

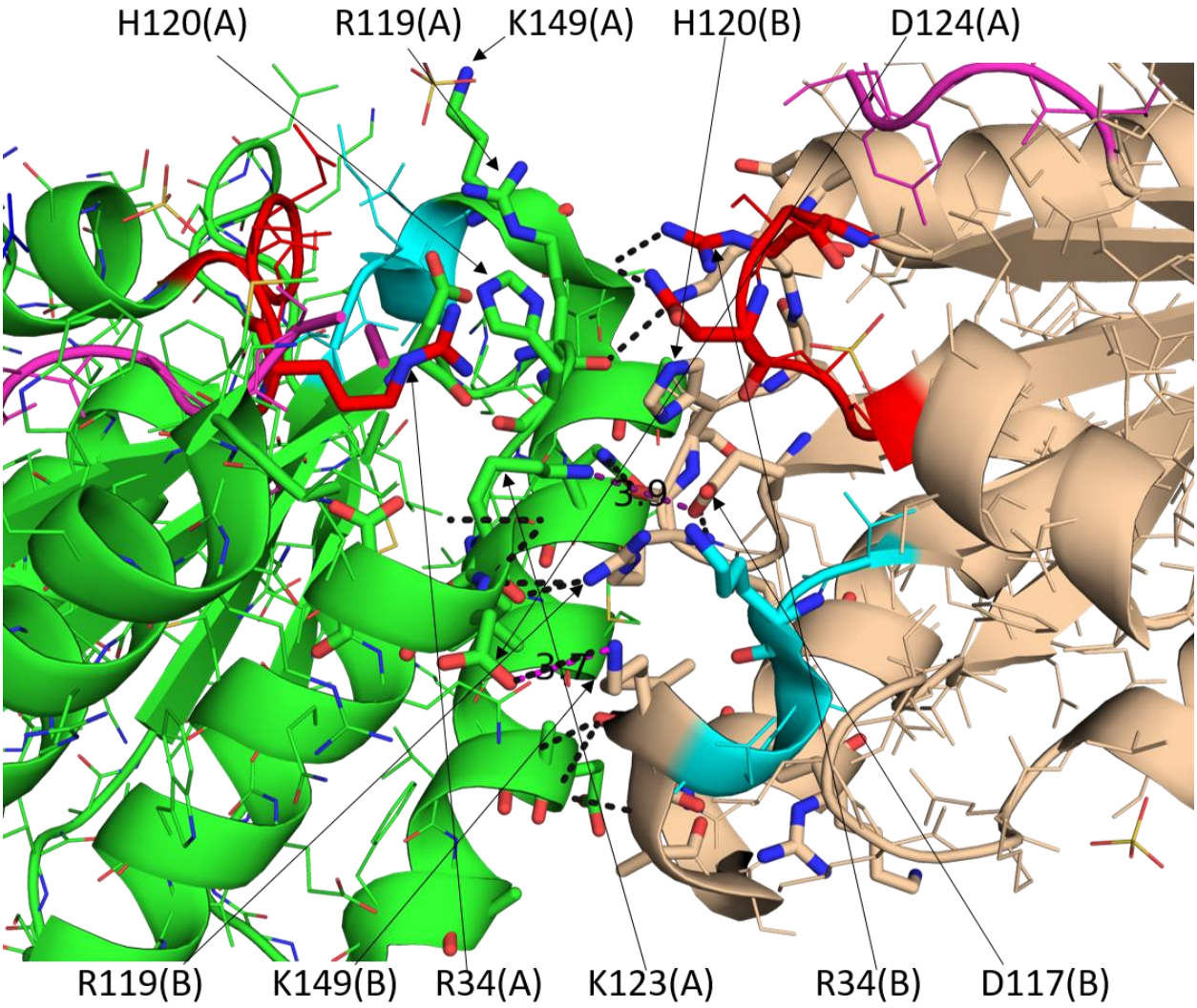


Figure 2.7. Dimerization interface of EcYihA showing the locations of salt bridges, H-bonds and conserved residues. A zoomed in view of the model in Figure 2.6 is shown. K149(B)-D124(A) forms a salt bridge of 3.7 Å and D117(B)-K123(A) forms a salt bridge of 3.9 Å. The P-loop is shown in red and Switch I is in blue. Switch II is shown in magenta and the G4 motif is colored in cyan. Amino acid residues 51-62 (most of Switch I) and residues 83-86 (part of Switch II) are missing in the structure. H-bonds are shown with black dashed lines and salt bridges are shown in magenta dashed lines. Distances are given in the structure for salt bridges in units of Ångstroms. The crystal structure is from PDB ID 1PUI. Monomer A is shown in green and monomer B is shown in tan.

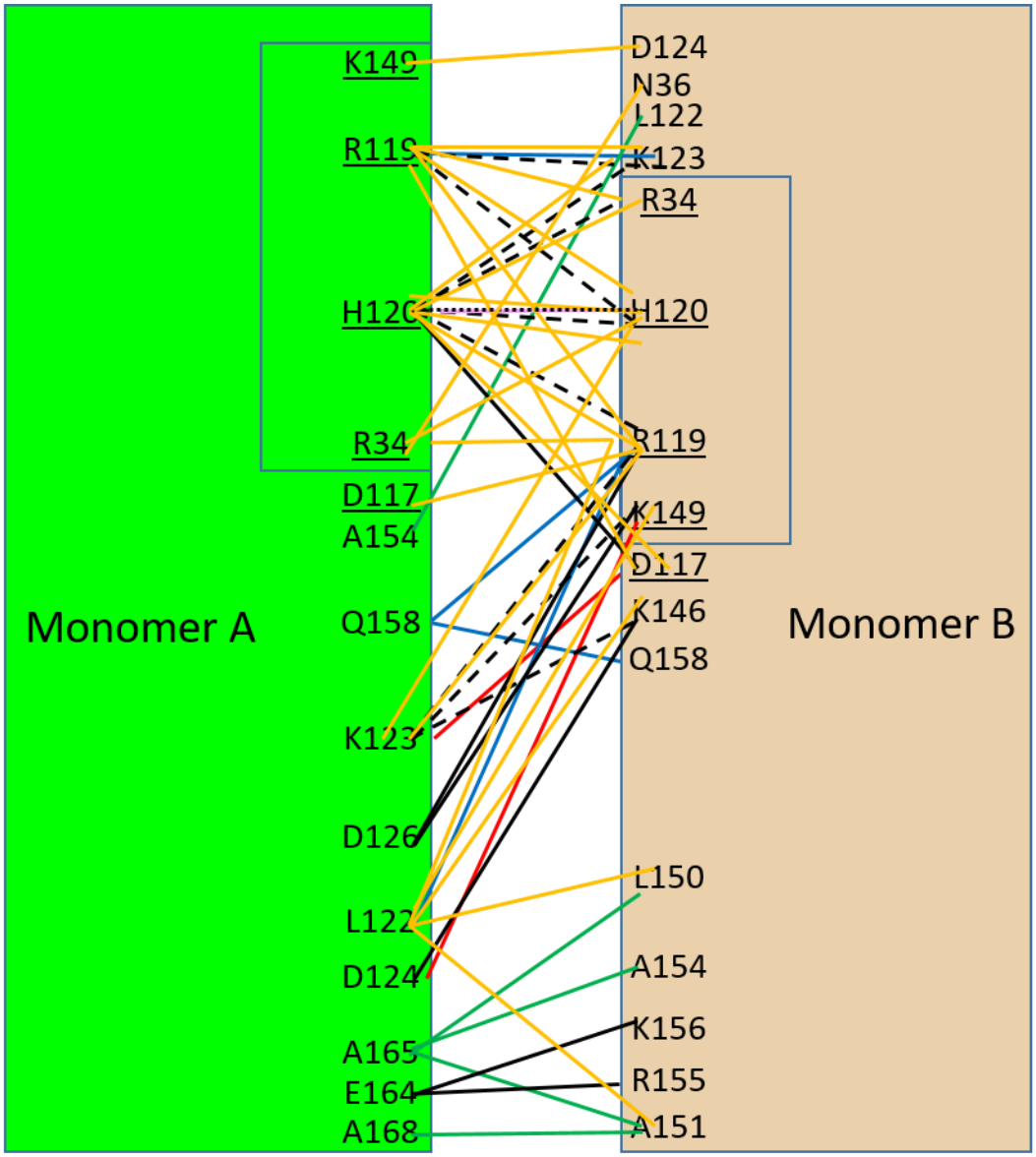


Figure 2.8. Potential intermolecular interactions between the two monomers in the EcYihA crystal structure. The diagram is based on the PDB ID 1PUI crystal structure. K149(B)-D124(A) forms a potential salt bridge of 3.7 Å and D117(B)-K123(A) forms a potential salt bridge of 3.9 Å. Potential H-bonds are shown with blue lines and salt bridges are shown with red lines. Potential hydrophobic interactions are indicated with green lines. Dotted black lines show potential π-π stacking interactions and solid black lines represent potential favorable electrostatic interactions. Potential unfavorable electrostatic interactions are indicated by dashed black lines and potential Van der Waals interactions are shown in orange lines. The boxed residues form a conserved patch of positive charge on the surface of YihA. Underlined residues are conserved. Potential interactions were calculated with the PPCheck [215], RING 3.0 [216] and PDBePISA [210, 211] web servers. Monomer A is shown in green and monomer B is shown in tan.

2.3.5 BIOPHYSICAL CHARACTERIZATION OF YIHA IN SOLUTION BY SAXS

Given the controversy between the PDB and the PDBePISA web server regarding their predictions as to whether the crystal structure of EcYihA represents a monomer, dimer, or trimer (Table 2.2), a biophysical method was needed that would provide a model or structure of EcYihA in solution that the EcYihA crystal structure of YihA could be docked into for comparison. However, EcYihA is too small in size for structural determination by cryo-EM [217]. Alternatively, Small Angle X-ray Scattering (SAXS) is an important tool for determining the oligomeric state of a macromolecule in solution [208]. Therefore, the solution properties of EcYihA were analyzed by SAXS (Figure 2.9). SAXS provides information about the shape, size, conformational changes and flexibility of proteins or nucleic acids in solution [218].

Initially, EcYihA samples in the concentration range of 1.5 mg/mL (62 μM) to 2.5 mg/mL (103 μM) were analyzed by SAXS. However, as scattering of the X-rays was not sufficient for a strong signal above background noise to be achieved at these concentrations, attempts were made to collect scattering data for EcYihA concentrations of 10 mg/mL (413 μM) and 20 mg/mL (826 μM). The data from the 20 mg/mL samples was not analyzed further as initial analysis suggested that these samples contained aggregation and all further analyses were carried out with the 10 mg/mL samples.

SEC coupled in-line to SAXS was utilized here to investigate the oligomerization state of EcYihA. Using a specialized flow cell (refer to section 2.2.9 for details), SAXS measurements were taken to determine the scattering intensity corresponding to frames that eluted following SEC (Figure 2.9). Each frame was exposed to the beam of X-rays for 3 seconds (Figure 2.9). EcYihA was observed in the eluate from the SEC column in the form of one major peak eluting at a frame number of ~ 350 (Figure 2.9). Previous work with the SEC-SAXS setup at the B21 beamline at the Diamond Light Source using the same SEC column used here (Shodex KW402.5-4F) has estimated that samples are diluted by a factor of approximately 3 prior to entering the

flow cell, when analyzed with the Shodex KW402.5-4F SEC column [219]. Therefore, the concentration of EcYihA upon exposure to X-rays in the SEC-SAXS experiments was ~ 3.3 mg/mL (~ 138 μM) for the 10 mg/mL samples analyzed here [219].

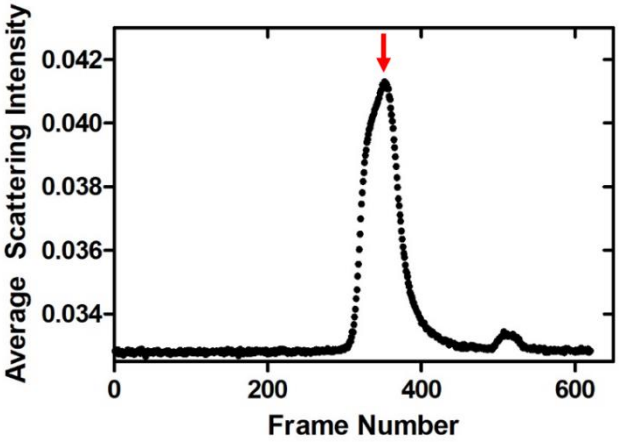


Figure 2.9. SEC-SAXS profile of EcYihA. The integrated intensities obtained over the course of the SEC-SAXS experiment are shown for each elution time point, which is represented as a frame number. The red arrow indicates the peak that was analyzed and was found to correspond to an EcYihA dimer as explained in the text. The exposure time was 3 seconds per frame.

The scattering curve for the experimental EcYihA sample is shown in Figure 2.10A where the scattering intensity of the X-rays, (*I(q)*) is shown as a function of the magnitude of the scattering wavevector, q, in units of $Å^{-1}$. I(0) represents the forward scattering, which is the scattering intensity at a scattering wavevector, q, of zero (Table 2.3). A Guinier plot of *lnI(q)* vs q^2 (Figure 2.10B) is used to assess the monodispersity of the sample. The Guinier plot demonstrates a linear fit (R^2 = 0.9703) in the low-q region as seen in Figure 2.10B, with the linearity of the Guinier plot indicating that the EcYihA SAXS sample does not contain aggregated material [220]. Moreover, a linear Guinier fit confirms that the background was subtracted properly and that there are no repulsive or attractive inter-particle interactions that strongly affect the X-ray scattering of the sample [221-223]. As aggregation was not detected in the sample by the Guinier analysis, validating that the SAXS data quality was acceptable, the data was analyzed further starting with

calculation of the radius of gyration (R_g) from the Guinier analysis. The R_g represents the average of the root mean square distance from the molecule's electrons to the centre of mass of the molecule's electron density [224] and can be calculated based on the slope of the Guinier plot [222, 225]. If the Guinier plot is fitted with a linear regression, the linear regression line can be extrapolated in order to obtain I(0) at q = 0 as the beginning of the scattering curve cannot be recorded experimentally due to obstruction from the beamstop [222, 226]. From the Guinier analysis, the R_g was determined to be 29.76 ± 0.18 Å with the I(0) found to be (1.1×10^{-2}) ± (3.5×10^{-5}) (Table 2.3).

The one-dimensional information contained in the inverse space representation of the X-ray scattering intensity data in the *logI*(*q*) vs. *q* plot (Figure 2.10A) was transposed into the real space electron pair distance distribution P(r) function (Figure 2.10C) using an indirect Fourier transformation executed by the program GNOM [193]. Although the Guinier approximation utilizes only the data in the low q region of the Guinier plot (Figure 2.10B) to calculate the R_g, the P(r) function is derived from an indirect Fourier transformation of most of the scattering dataset at both low and higher scattering angles and contributes towards facilitating accurate calculation of I(0), R_g and D_{max} [190, 227]. D_{max}, which is the maximum particle dimension of YihA, represented by the largest radius *(r)* where the P(r) function is greater than zero [228], was 85.0 Å for EcYihA and the real space R_g for EcYihA determined from the P(r) function was 28.47 ± 0.06 Å (Table 2.3, Figure 2.10C). The I(0) calculated from the P(r) function was (1.1×10^{-2}) ± (2.4×10^{-5}) (Table 2.3). It is important to note that the error in the D_{max} value is not calculated or reported here as per convention [223, 229-232]. Since SAXS experiments are able to collect data only for a limited set of values of *q*, the indirect Fourier transformation required to obtain the P(r) function assumes that both radius = D_{max} and radius = 0 will have a corresponding P(r) value of P(r) = 0 [223] (Figure 2.10C). The resulting D_{max} has an error that is determined largely by how reliable the SAXS

experimental data is [223]. Therefore, calculation of the uncertainty in D_{max} is normally considered to be invalid [223].

As the P(r) function represents all of the distances between each electron pair in the molecule in the format of a histogram weighted by the electron density at each position, the symmetry or asymmetry and form of the P(r) plot reveals information about the size and overall shape of the molecule in solution [218, 225]. In the case of a globular, compact protein such as the 66 kDa protein monomeric BSA, the P(r) distribution is anticipated to be bell-shaped and symmetrical about the maximum that is approximately half of D_{max} [233], as shown in Figure 2.10D. In contrast, for the 61 kDa monomeric human protein p67phox (neutrophil cytosol factor 2) which contains somewhat flexible, incompletely structured linker regions connecting multiple folded domains [234], the P(r) function is asymmetric with an extended tail and a maximum that is much less than half of D_{max} (Figure 2.10D). Although p67phox contains less amino acids in its protein chain than monomeric BSA, the D_{max} of BSA is 82.3 Å whereas the D_{max} of p67phox is 160.0 Å based on data taken from the SAXS Biological Data Bank (SASBDB), consistent with the extended conformation of p67phox (Figure 2.10D). The program CRYSOL [198] was used to compute the theoretical P(r) functions for three crystal structures of YihA including 1PUI (*E. coli* monomer), 1PUI (*E. coli* dimer) and 2CXX (*Pyrococcus horikoshii* YihA (PhYihA) trimer) (Table 2.4). The theoretical P(r) function for the monomeric EcYihA crystal structure is approximately symmetrical as expected for a globular spherical protein (Figure 2.10C). In contrast, the dimeric EcYihA crystal structure has a P(r) function that is asymmetrical and skewed towards larger values for the inter-particle distance represented by the radius (Figure 2.10C). This shape of P(r) function is consistent with the rod-like shape of the dimeric crystal structure [221] which forms an extended dimer (Figure 2.6). However, the experimental P(r) is also asymmetrical but more elongated than the crystal structure of the dimer with a shoulder at a radius of ~ 65 Å indicating that the solution structure of EcYihA is more extended and rod-shaped than the dimeric crystal

structure of EcYihA (Figure 2.10C). Comparison of the experimental P(r) distribution with the P(r) distributions of the monomeric, dimeric and trimeric crystal structures of YihA suggests that EcYihA in solution is most similar in shape and size to the dimeric crystal structure (Figure 2.10C). Furthermore, the presence of a clear shoulder peak at ~ 65 Å in the experimental P(r) function for EcYihA in addition to the maximum at ~ 30 Å (Figure 2.10C) has been shown previously to be diagnostic of a multi-subunit macromolecule displaying a bi-lobed configuration with the two subunits situated slightly apart [235]. Thus, the shape of the experimental P(r) function for EcYihA suggests that YihA forms a dimer in solution (Figure 2.10C). The shape of the P(r) function for EcYihA in solution is completely different from the shape of the P(r) function of the trimeric crystal structure (Figure 2.10C). Rather than forming a trimer of YihA subunits joined end to end in a rod-like or chain-like shape, the trimeric crystal structure of YihA resembles a ring shape. For a ring-like assembly of a trimeric biological macromolecule, such as that of the Z-α1-Anti-trypsin trimer, previous SAXS studies have demonstrated that its P(r) function consists of two peaks (bimodal) with one peak located at a shorter distance than half of D_{max} that represents the breadth of one subunit in the ring [236, 237]. The other peak has been determined to be located at a distance larger than half of D_{max} of the P(r) function, representing the greater distance of the internal diameter of the ring [236, 237]. These two characteristic peaks are both present in the theoretical P(r) function calculated for the trimeric crystal structure of YihA but the peak at a distance farther than $D_{max}/2$ is absent from the P(r) function calculated from the experimental scattering data for EcYihA, indicating that EcYihA in solution does not adopt the ring-shaped biological assembly found in the trimeric crystal structure of YihA (Figure 2.10C). The possible reasons explaining why EcYihA forms a dimer in solution as shown by analytical gel filtration chromatography experiments (Figure 2.1, Figure 2.2) even though the shape of EcYihA in solution given by the P(r) function is more extended than the shape of the dimeric crystal structure are presented in subsequent paragraphs. The overall shape of the P(r) function of EcYihA most closely resembles the shape

of the theoretical P(r) function for the crystal structure of the dimer and is more extended than the crystal structure of the monomer (Figure 2.10C). The P(r) function for the monomeric crystal structure is approximately symmetrical with a maximum at approximately half of D_{max} (Figure 2.10C).

The P(r) function for EcYihA is characterized by an asymmetric peak with a maximum that is smaller than half of D_{max} along with a considerable shoulder starting at a radius of ~ 65 Å (Figure 2.10C). The presence of a shoulder along with the asymmetry in the P(r) function of EcYihA provided the first suggestion that EcYihA does not have a spherical or completely compact structure in solution [238]. To test this assessment, the SAXS scattering intensity curve was analyzed with the algorithm DATCLASS [239]. DATCLASS predicts the shape of biological macromolecules by normalizing the log(*I*) vs. *q* scattering intensity plot to a dimensionless Kratky plot, which is then integrated to yield an apparent volume based on three different limits of the product qR_g that quantitatively describe the normalized apparent volume space in 3 dimensions, as compared to the normalized apparent volume space of a set of pre-defined geometric shapes [240, 241]. DATCLASS estimated that EcYihA has an extended shape in solution instead of the other six available shape categories including compact, flat, ring, compact-hollow, hollow-sphere and random chain. However, in order to draw conclusions about the solution conformation of EcYihA, it was necessary to perform flexibility analysis and therefore dimensionless Kratky analysis and Porod-Debye analysis was carried out.

A dimensionless Kratky plot enables a direct comparison of the folding of proteins or nucleic acids with varying conformations and molecular weights [234]. The advantage of the dimensionless Kratky plot is that it can differentiate entirely folded proteins from proteins that have both folded domains and unstructured elements [189, 234]. A dimensionless Kratky plot normalizes both the scattering wavevector, *q*, and the scattering intensity, *I(q)*, by a transformation such that the product qR_g is depicted on the x-axis and $[I(q)/I(0)](qR_g)^2$ is presented on the y-axis

[189, 242]. Taking the product qR_g ensures that qR_g is independent of the size of the macromolecule and taking the quotient *I(q)/I(0)* ensures that *I(q)/I(0)* is independent of the molecular weight of the macromolecule according to the equation $I(0) = \frac{cM_w}{N}\left[\left(\rho_p - \rho_s\right)V_p\right]^2$ where N represents Avogadro's number, c represents the concentration of the protein, ρ_p represents the scattering length density of the protein, ρ_s represents the scattering length density of the solvent, M_w represents the molecular weight and V_p represents the specific volume of the protein [189]. Thus, dimensionless Kratky plots can qualitatively compare folding and the degree of disorder in any proteins regardless of their size [189]. Comparison of dimensionless Kratky plots for many different proteins demonstrates that these plots feature a maximum at $qR_g = \sqrt{3} \approx 1.732$ and $[I(q)/I(0)](qRg)^2 = 1.104$ for fully folded, globular proteins [243] such as chymotrypsinogen A shown in Figure 2.10F. In contrast, the N-terminal domain of the Estrogen Receptor alpha (ER-α NTD), a disordered domain, has a dimensionless Kratky curve shown in the yellow trace in Figure 2.10F, that does not reach a maximum but instead $[I(q)/I(0)](qRg)^2$ continues to increase with increasing qR_g [244]. The intrinsically disordered PIR (phosphorylated Insulin Receptor interaction region) domain shows a similar dimensionless Kratky plot as the ER-α NTD and it is estimated from nuclear magnetic resonance (NMR) experiments and circular dichroism (CD) spectroscopy data that secondary structure is only found in under 10% of the sequence of the PIR domain [234, 245, 246] Furthermore, an example of the dimensionless Kratky plot of a partially folded protein is given by the green trace for the Bromodomain Adjacent to Zinc Finger Domain protein 2A in Figure 2.10F, which is bell-shaped with a maximum at $qR_g > 1.732$ and $[I(q)/I(0)](qRg)^2 > 1.104$.

Dimers of globular proteins such as the human protein LRRK2 or BSA as well as tetramers of the globular protein yeast alcohol dehydrogenase-1 and 24-mers of the globular protein Apoferritin Light Chain from horse spleen all have dimensionless Kratky plots that are bell-shaped with a maximum at $qR_g = \sqrt{3} \approx 1.732$ and $[I(q)/I(0)](qRg)^2 = 1.104$, which is referred to as the

Guinier-Kratky point [247-249]. The theoretical dimensionless Kratky curves for the atomic resolution crystal structures of the YihA monomer, dimer and trimer were calculated in CRYSOL (Figure 2.10E). The theoretical dimensionless Kratky plots for the monomeric, dimeric and trimeric YihA crystal structures all contain a bell-shaped curve at low *qRg* values demonstrating that all three of these structures consist of a fully folded core (Figure 2.10E). Furthermore, the dimensionless Kratky plots of the monomeric and trimeric YihA crystal structures are characterized by maxima at $qR_g \approx 1.732$ and a value of $[I(q)/I(0)](qRg)^2 \approx 1.104$, confirming that the monomeric and trimeric crystal structures of YihA are globular (Figure 2.10E). By contrast, the dimensionless Kratky plot for the dimeric crystal structure of YihA evidences a maximum at $qR_g \approx 1.9$ and a value of $[I(q)/I(0)](qRg)^2 \approx 1.2$, indicating that the dimeric crystal structure of YihA is non-globular, consistent with the extended rod-like shape observed for the dimeric crystal structure of YihA (Figure 2.6, Figure 2.10E). The dimensionless Kratky plot for the trimeric crystal structure of YihA can also be distinguished from the dimensionless Kratky plots of the monomeric and dimeric YihA crystal structures by the presence of an additional peak at approximately *qRg* = 4.5, a peak also observed in dimensionless Kratky plots from other oligomeric proteins whose quaternary structure is ring-shaped and hollow [250, 251]. Because of noise in the scattering data in the region surrounding *qRg* = 4.5 in the dimensionless Kratky plot for the experimental EcYihA sample, it is not possible to detect whether the additional peak at roughly *qRg* = 4.5, characteristic of the trimeric YihA crystal structure, is present in the dimensionless Kratky plot prepared from the scattering data of EcYihA in solution (Figure 2.10E). The dimensionless Kratky plot from the scattering data for EcYihA is bell-shaped with a maximum at approximately $qR_g = 2.4$ and a value of approximately $[I(q)/I(0)](qRg)^2 = 1.3$ (Figure 2.10E). Therefore, the shape of the dimensionless Kratky plot in Figure 2.10E indicates that EcYihA used for SAXS structure determination contains a completely folded core structure but also contains one or more flexible elements such that its overall structure is incompletely folded [238]. In addition, the dimensionless Kratky plot derived

from the scattering curve for EcYihA in solution displays increasing values of $[I(q)/I(0)](qRg)^2$ for large values of qRg, a trend that has been previously predicted to signal that a protein contains flexible regions with an extended conformation [252]. An understanding of the structural elements of EcYihA that could be unfolded will be discussed in subsequent paragraphs.

The Porod-Debye law stipulates that for a globular protein, the scattering intensity I(q) will be proportional to q^{-4}, multiplied by $\Delta\rho$ (the electron density scattering contrast), c (protein concentration) and S (surface area of the protein) [242]. In order to facilitate the preparation of a Porod-Debye plot of $q^4I(q)$ vs q^4 with the scattering data from EcYihA in solution, the Porod-Debye region, which is the region including all q values for which the Porod-Debye law is valid, was determined by plotting a Porod curve depicting $q^4I(q)$ vs q (Figure 2.10G) [243]. Based on the q range up to the maximum of the first peak in the Porod plot (Figure 2.10G), the q range which should be included in the Porod-Debye plot is from approximately q = 0.03 $Å^{-1}$ (q^4 = 8.10 x 10^{-7} $Å^{-4}$) to approximately q = 0.13 $Å^{-1}$ (q^4 = 2.86 x 10^{-4} $Å^{-4}$) (Figure 2.10G, Figure 2.10H) [243]. To clarify the extent of flexibility in the structure of EcYihA in solution, the data from the scattering intensity plot shown in Figure 2.10A was transformed to generate a Porod-Debye plot (Figure 2.10H). Macromolecules that have folded structure show a well-defined X-ray scattering electron density contrast between the scattering from the particle and the scattering from the solvent, with a Porod-Debye plot that approaches a characteristic plateau at increasing q values [242, 253]. This plateau is referred to as the Porod Plateau (Figure 2.10H) [242]. Conversely, disordered macromolecules that lack folded structure do not have a sharp electron density scattering contrast between solvent scattering and sample scattering [242]. The Porod-Debye plot for disordered macromolecules increases at high q values rather than approaching a plateau [253]. A Porod Plateau within the Porod-Debye region of the scattering data can be clearly seen in the Porod-Debye plot for EcYihA (Figure 2.10H), indicating that EcYihA contains mostly folded structure but not excluding the presence of some disordered segments in EcYihA [235].

In general, the scattering intensity I(q) is proportional to q^{-PE}, with PE representing a constant, referred to as the Porod Exponent, that provides a quantitative measure of the flexibility of a particle [254, 255]. The Porod Exponent can be computed from the negative value of the slope of a line fitted to the linear range of a *log(I(q))* vs. *log(q)* plot [254, 255]. For a completely unfolded chain, the Porod Exponent will be 2 whereas in accordance with the Porod-Debye law, a globular particle will display a Porod Exponent of 4 [255, 256]. Likewise, a particle containing both compact, folded regions and disordered chains, such as a configuration of multiple folded domains connected by flexible regions, will exhibit a Porod Exponent near 3 [256]. In the case of EcYihA, the Porod Exponent was determined from the SAXS data, using the program Scatter, to be 3.7, suggesting that EcYihA has mainly folded structure consisting of a folded core attached to short unstructured segments [257].

Further evidence that EcYihA dimerizes in solution comes from Porod-Debye analysis of the EcYihA scattering curve, which reveals that the Porod volume for EcYihA is 65,891 Å^3. The Porod volume refers to the particle volume, including the hydration shell, where the particle volume is described by $2\pi^2 \frac{I(0)}{Q}$ with Q obtained by computing the area under a Kratky plot of $q^2I(q)$ vs q [242, 258]. If EcYihA adopts a dimeric state in solution, then its density can be estimated to be 1.22 g/mL from the Porod volume, which is a reasonable estimate of the protein density given that an earlier SAXS study of a set of 31 proteins yielded protein densities ranging from 0.9 g/mL up to 1.5 g/mL based on the scattering results [242]. The alternative hypotheses that EcYihA exists as a monomer or as a trimer in solution would result in a protein density for EcYihA of 0.61 g/mL for the monomer or 1.83 g/mL for the trimer and both of these protein densities fall well outside of the expected range for proteins containing some folded structure [235]. As the average protein density is between 1.35 g/mL and 1.37 g/mL for a globular protein [242], a protein density of 1.22 g/mL calculated for EcYihA indicates that there are flexible

elements present in EcYihA's structure, consistent with the findings from dimensionless Kratky analysis.

A significant difference between the reciprocal space R_g based on Guinier analysis and the real space R_g may indicate that attractive or repulsive interactions between particles are influencing only the Guinier analysis [222]. The R_g and I(0) values calculated from the reciprocal space Guinier analysis and the real space P(r) function were similar with a % difference of less than 4.5% for the R_g values, suggesting that the dataset is of good quality and that the data quality is high enough for low resolution structural models to be built [227, 229]. Therefore, as the SAXS analysis suggested that YihA contains a folded core, *ab initio* modelling was performed with the data from the P(r) using the program DAMMIN and the models were averaged in DAMAVER. *Ab initio* modelling is executed without the need for information about the structure of the macromolecule to be known beforehand [239]. Twelve low resolution *ab initio* three-dimensional models of YihA were constructed and averaged. As an extended bar/rod-shaped SAXS envelope was obtained, the crystal structure of the dimer was docked in the model using DAMSUP [195] (Figure 2.11A). The resolution of SAXS structural models of RNA obtained with the program DAMAVER, which was used here to obtain the model of YihA, has been estimated to range from 13 Å to 37 Å [230], consistent with the YihA model obtained here being a low resolution model. In agreement with the P(r) distribution function and Porod-Debye analysis, the model of YihA resembles a short rod or a short bar (Figure 2.11A). The SAXS model of EcYihA is elongated and the crystal structure of the dimer is able to be superimposed on the model with a good normalized spatial discrepancy (NSD) of 1.05, whereas a structure of a monomer would not fit into the SAXS envelope (Figure 2.11A). In the case of two objects resembling each other that can be perfectly superimposed, the NSD will be closer to zero, whereas if two objects are systematically dissimilar, the NSD will be greater than 1 [259]. An NSD of 1.05 for the low resolution model, which is very

close to 1, suggests that the superimposition of the SAXS model with the YihA dimeric crystal structure is good and that the model agrees well with the dimeric crystal structure of YihA.

In order to validate the structural modelling of EcYihA using SAXS data before attempting to build any ensemble-based atomistic models of EcYihA that may fit the SAXS data better, the molecular weight of EcYihA was estimated from the SAXS data [223]. The molecular weight of a protein or nucleic acid can be directly calculated from SAXS data [189]. In the case of a protein solution that displays monodispersity, the molecular weight approximated from SAXS data can be compared to the sequence-based expected molecular weight as a check for aggregation and to identify the oligomeric state of homo-oligomers [241, 260]. Two concentration-dependent methods are available to determine the molecular weight of a protein using SAXS. Firstly, calibration of the forward scattering intensity I(0) to an absolute scale based on the known scattering of water or of a similar standard enables I(0) to be given in absolute units, in which case the molecular weight of the protein is proportional to $\frac{I(0)}{C}$ scaled by Avogadro's number and normalized against ${\Delta\rho_M}^2$, where C represents the concentration of the protein and ${\Delta\rho_M}^2$ represents the electron density scattering contrast per weight of dry protein [241, 260]. Alternatively, if protein standards (for instance, lysozyme or glucose isomerase) are analyzed by SAXS, the molecular weight of an experimental protein sample can be calculated knowing that the ratio of the molecular weight of the experimental protein to the known molecular weight of the standard is equal to the ratio of $\frac{I(0)_{sample}}{C_{sample}}$ to $\frac{I(0)_{standard}}{C_{standard}}$ [260-262].

Because the concentration of the EcYihA peak from the SEC-SAXS experiment performed in order to collect SAXS data for EcYihA was not made available, concentration-independent methods were used to determine the molecular weight of EcYihA from the scattering data. The molecular weight of EcYihA was approximated from estimates of the protein volume computed from SAXS data using five different concentration-independent methods including the Volume of

Correlation method, the Volume method, the MoW method, the MM_{QP} method, and the Size and Shape method. There is an advantage in reporting molecular weights determined by different methods since the error in each of these approaches is due to different causes. According to the Volume method, the molecular weight of a protein or nucleic acid can be estimated by taking its hydrated volume calculated from an *ab initio* model reconstructed by the program DAMMIN (with averaging and filtration in the DAMAVER software) and dividing that hydrated volume by a factor of 1.7 [200, 208]. The Volume method yielded a molecular weight of 41.9 kDa for EcYihA (Table 2.3).

In the MoW method, Q', representing the area under a Kratky curve of $q^2I(q)$ vs q is integrated up to the user-provided limit of q_{max} [263]. An apparent volume, V', is calculated from Q' since $V' = \frac{2\pi^2 I(0)}{Q'}$ [263]. Given that $Q = \int_0^\infty q^2 I(q)\, dq$ and Q is defined as integrated from 0 to ∞ where the Porod volume, V_p, is equal to $V_p = \frac{2\pi^2 I(0)}{Q}$, the Porod volume is estimated from the apparent volume using correction factors computed from theoretical SAXS curves for 1145 available protein structures [241, 263]. Taking the product of the Porod volume and the average mass density of a protein in solution (0.83 x 10^{-3} kDa $Å^{-3}$), without accounting for post-translational modifications, yields the estimated molecular weight [263]. With the MoW method, EcYihA's molecular weight was estimated to be 44.9 kDa (Table 2.3). In contrast to the MoW method, the MM_{QP} method calculates Q by extrapolating the SAXS intensity curve [241]. The *I(q)* value corresponding to the lower limit of the integral for Q, *I(q)* = 0, is computed as the I(0) in Guinier analysis [241]. For the upper limit of the integral for Q, *I(q)* = ∞, the I(q) value of q_{max} = 8/R_g is used, since a previous study showed that an upper limit of q_{max} = 8/R_g provides an accurate estimation of the Porod volume [200, 241]. The scattering curve is also extrapolated to infinity as a final check on the reliability of the estimate of the Porod volume and the molecular weight is

equal to the quotient $\frac{Porod\ volume}{1.37}$ [241]. From analysis of the scattering data for EcYihA, the MM_{QP} method provided a molecular weight approximation for EcYihA of 42.4 kDa (Table 2.3).

Protein volume can also be estimated by the Volume of Correlation method. The Volume of Correlation, V_c, is defined as $V_c = \frac{V_p}{2\pi l_c}$, where V_p is the Porod volume and l_c is the correlation length, which describes the mean length of a line between two positions in the electron density pair distribution function, indicating that V_c is dependent on a protein's conformation [241, 264, 265]. According to the Volume of Correlation method, the Volume of Correlation is calculated from the quotient of *I(0)* divided by the molecule's scattering intensity, with intensity given as the integral of the area under a plot of *qI(q)* vs *q* where the integral is first evaluated up to an upper limit of q_{max} = 0.5 Å^{-1} [241, 264]. There is a linear relationship between $\frac{V_c^2}{R_g}$ and molecular weight when plotted on a log-log scale [264]. Based upon the Volume of Correlation analysis, a molecular weight of 39.1 kDa was obtained for EcYihA (Table 2.3).

Unlike the four approaches to estimate molecular weight discussed above, the Size and Shape method is based upon a machine learning strategy that calculates shape and size parameters from the SAXS analysis of a protein and compares them to shape and size parameters for 165,982 proteins to determine the molecular weight of a protein solution [241]. The SAXS scattering intensity curve for the protein solution is transformed to a dimensionless Kratky plot [$I(q)/I(0)(qRg)^2$ vs *qRg*] and *Q'*, the area under the dimensionless Kratky plot, is integrated to three particular upper limits of *qRg* yielding three apparent volumes V', where $V' = \frac{2\pi^2}{Q'}$ [240]. An approximation of the molecular weight of the unknown protein is obtained by averaging the five classes of neighbors that are closest to the intersection of the three values of *V'* and the R_g of the unknown protein, representing a space of four dimensions [240, 241]. Applying the Size and Shape analysis to the SAXS data for EcYihA resulted in an estimated

molecular weight for EcYihA of 47.5 kDa (Table 2.3). While molecular weights calculated from SAXS data are not as precise as molecular weights obtained by mass spectrometry or SEC multi-angle light scattering (SEC-MALS), calculations with SAXS data typically yield molecular weights with a deviation of approximately 10% from the theoretical value, which is accurate enough to indicate the oligomeric state adopted by a monodisperse protein solution in a SAXS experiment [226, 260, 264].

If the estimates for EcYihA's molecular weight from SAXS data are interpreted as indicating that EcYihA is a dimer in solution, the % deviation from the theoretical molecular weight of an EcYihA dimer, 48.4 kDa, can be computed. The largest % deviation from the theoretical molecular weight of EcYihA was found for the Volume of Correlation method, which gave a % deviation of 19.2%. In contrast, the Volume method yielded a molecular weight with a % deviation of 13.5%. With the MM_{QP} and MoW methods for determining molecular weight, the % deviation values for molecular weight were 12.4% and 7.2%, respectively. The smallest % deviation from the theoretical molecular weight of an EcYihA dimer was obtained with the Size and Shape method for calculating molecular weight, which provided a % deviation of only 1.9 %.

Results from one study demonstrated that the Volume method of determining molecular weight from SAXS data provides an approximation of a protein's molecular weight with an average % deviation of 23% in the case of proteins whose molecular weight is greater than 100 kDa [208]. Moreover, the % deviation in molecular weight is greater than 50% for many proteins with a molecular weight smaller than 100 kDa [208]. There is agreement in the literature that the Volume method is somewhat accurate for proteins that are greater than 70 kDa in molecular weight and globular in shape [266]. However, the Volume method often cannot provide reliable estimates of protein molecular weight for proteins with a compact form or for proteins that are smaller than 70 kDa [266].

An explanation for the very high % deviation for the molecular weight of EcYihA determined with the Volume of Correlation method is that previous work showed that the Volume of Correlation method is unable to accurately approximate the molecular weight of proteins categorized as extended by the DATCLASS algorithm [240, 241]. Since the Volume of Correlation, V_c, is proportional to the inverse of the correlation length l_c, V_c (and therefore the calculated molecular weight) will be altered if a protein has a conformation that is extended in the direction of one axis only [241]. As EcYihA was classified as extended by DATCLASS and *ab initio* modelling suggests that EcYihA forms an extended dimer in solution (Figure 2.11A), a high % deviation for molecular weight, well above 10%, is expected from the Volume of Correlation method, consistent with the 19.2% deviation from EcYihA's dimeric molecular weight obtained from the Volume of Correlation method.

Furthermore, one research group that examined the accuracy of methods to estimate moelcular weight from SAXS data found that the Size and Shape method was the most reliable method among the five methods discussed here, followed by the MoW method [241]. However, the MoW method provides a larger uncertainty in the molecular weight for SAXS data where $q < 0.45$ Å^{-1} [267], which is the case for the data presented here where $q_{max} = 0.34$ Å^{-1}, consistent with the larger % deviation in the molecular weight of 7.2% for EcYihA with the MoW method. Where proteins in the PDB were used as the test set, the Size and Shape method provided an estimate of molecular weight with a deviation of less than 10% for about 90% of proteins in the test set [240].

Taken together, the conclusions that the Volume method of determining molecular weight is not often reliable for proteins with a molecular weight less than 70 kDa, that the Volume of Correlation method cannot reliably determine the molecular weight of extended proteins and that the Size and Shape method is the most accurate of the five methods presented here with the MoW method being found to be the next most accurate method, all provide support for the

conclusion that EcYihA forms a dimer in solution. As agreement exists in the literature that a moelcular weight that is obtained from SAXS data and has a % deviation of less than 10% from the sequence-based molecular weight can be used to demonstrate homo-oligomerization of a monodisperse sample [260, 264], the molecular weights obtained by both the Size and Shape method and the MoW method were well below a 10% deviation from the theoretical value for an EcYihA dimer, suggesting that the EcYihA sample dimerizes in solution. Consistently, the Size and Shape method and the MoW method are the two most reliable methods available to estimate molecular weight from SAXS data among the methods discussed here and the Size and Shape method and the MoW method are the best methods of the methods presented here for determining molecular weights of proteins with an extended shape [241]. The interpretation of the molecular weights obtained for EcYihA from the SAXS data as identifying EcYihA as a homodimer in solution is supported by the values for R_g for EcYihA obtained from the Guinier and P(r) function analysis (Table 2.3). Using the formula $R_{g\ globular\ protein} = \left(\frac{3}{5}\right)^{\frac{1}{2}}\left[\frac{Molecular\ Weight\ (Daltons)}{\left\{\left(0.44\frac{Daltons}{Å^3}\right)\left(\frac{3}{4}\pi\right)\right\}}\right]^{\frac{1}{3}}$, the theoretical R_g of an EcYihA monomer is predicted to be 22.1 Å [266], which is much smaller than the R_g of 29.76 ± 0.18 Å from Guinier analysis or than the R_g of 28.47 ± 0.06 Å obtained from P(r) function analysis (Table 2.3).

Previous work has suggested that addition of a tag of even 7 amino acids to a protein, including a His tag, results in a greater D_{max} [268]. This could partially explain why the theoretical D_{max} for the YihA dimeric crystal structure (D_{max} = 82.6 Å) is smaller than the D_{max} determined experimentally by SAXS for YihA (D_{max} = 85.0 Å), as the YihA used in SAXS experiments contained an N-terminal His tag whereas the YihA used in the crystal structure of the dimer had no His tag (Figure 2.12). The SAXS model of YihA is most likely not monomeric since the approximate dimensions of the BsYsxC crystal structure monomer (PDB ID 1SVW) are 35 Å x 40 Å x 45 Å and the D_{max} for the SAXS model of YihA is 85.0 Å (Table 2.4).

From the PDB coordinates, theoretical scattering curves were calculated in CRYSOL for the crystal structure of a single subunit of EcYihA (PDB ID 1PUI) (Figure 2.11D), for the dimeric EcYihA crystal structure (PDB ID 1PUI) (Figure 2.11C) and for the crystal structure of YihA with three chains in the asymmetric unit (PDB ID 2CXX) (Figure 2.11E). The theoretical scattering curves for the monomeric, dimeric and trimeric crystal structures were fitted to the experimental scattering data (Figure 2.11C-E). The reduced chi^2 test statistically describes the degree of agreement between the experimental scattering intensities and the scattering intensities of the fitted model for *m* experimental values [269]. It is expected that the calculated chi^2 value will fall within the range $0.9 \leq chi^2 \leq 1.1$ if the model satisfactorily fits the experimental data [270]. Furthermore, the discrepancy between the theoretical calculated value and the experimental value given by the chi^2 value calculated for the theoretical scattering from the dimeric crystal structure was only 2.0 for comparison of the theoretical scattering from the crystal structure of the dimer, suggesting that there is good agreement between the experimental data and the theoretical scattering from the dimeric crystal structure (Table 2.4). In contrast, a chi^2 value of 3.5 was obtained for the crystal structure of the trimer and a chi^2 value of 12.6 was obtained for the monomer, suggesting a poor agreement between the theoretical scattering of the crystal structures of the trimer and monomer and the experimental data (Table 2.4). The theoretical scattering curve of the trimeric crystal structure of YihA contains a wide peak at approximately q = 0.15 Å^{-1} that is indicative of a hollow object [268, 271], consistent with the ring-shaped crystal structure of trimeric YihA (Figure 2.11E). However, as the experimental scattering curve is very broad in the region of and adjacent to q = 0.15 Å^{-1}, it is not possible to determine whether this wide peak at approximately q = 0.15 Å^{-1} is present in the experimental scattering curve for EcYihA, although as expected this peak is not observed in the theoretical scattering curves for the EcYihA monomeric crystal structure (Figure 2.11D), EcYihA dimeric crystal structure (Figure 2.11C) or in the scattering curve for the *ab initio* SAXS model of YihA (Figure 2.11B).

A limitation of the low resolution *ab initio* model calculated from the SAXS data (Figure 2.11A) is the very high concentration of YihA used to obtain the scattering data (10 mg/mL). As a larger R_g corresponds to a larger total size of a macromolecule [218], the experimental R_g of EcYihA is larger than the theoretical R_g of the EcYihA dimer and larger than the theoretical R_g of the EcYihA trimer (Table 2.4), consistent with structural flexibility detected from YihA's scattering data that gives YihA an extended shape and length (Figure 2.10E-H). For two proteins with the same number of residues but a different fold, the protein with an extended conformation will have a larger R_g whereas the fully folded protein or globular protein will have a decreased R_g. The experimental D_{max} calculated for EcYihA is also larger than the theoretical D_{max} of the dimeric crystal structure or of the monomeric crystal structure, but smaller than the theoretical D_{max} of the trimeric crystal structure (Table 2.4), likely because of the much greater number of amino acids in the protein chain of purified EcYihA compared to the number of amino acids in the two protein chains in the unit cell of the EcYihA crystal structure (Figure 2.12). The two protein chains in the EcYihA crystal structure were truncated as the N-terminal and C-terminal tail regions were not resolved (Figure 2.12).

The R_g calculated in CRYSOL for chain A of the crystal structure of the EcYihA monomer of 18.6 Å is very close to the approximation of the R_g of a globular protein with 169 amino acids (Figure 2.12) given by $R_g \approx 3(n)^{\frac{1}{3}} \approx 17$ Å [234]. Short loops such as are found in EcYihA's structure in the Switch I and Switch II regions and not resolved in its crystal structure (PDB ID 1PUI), are not predicted to affect the SAXS scattering profile [272].

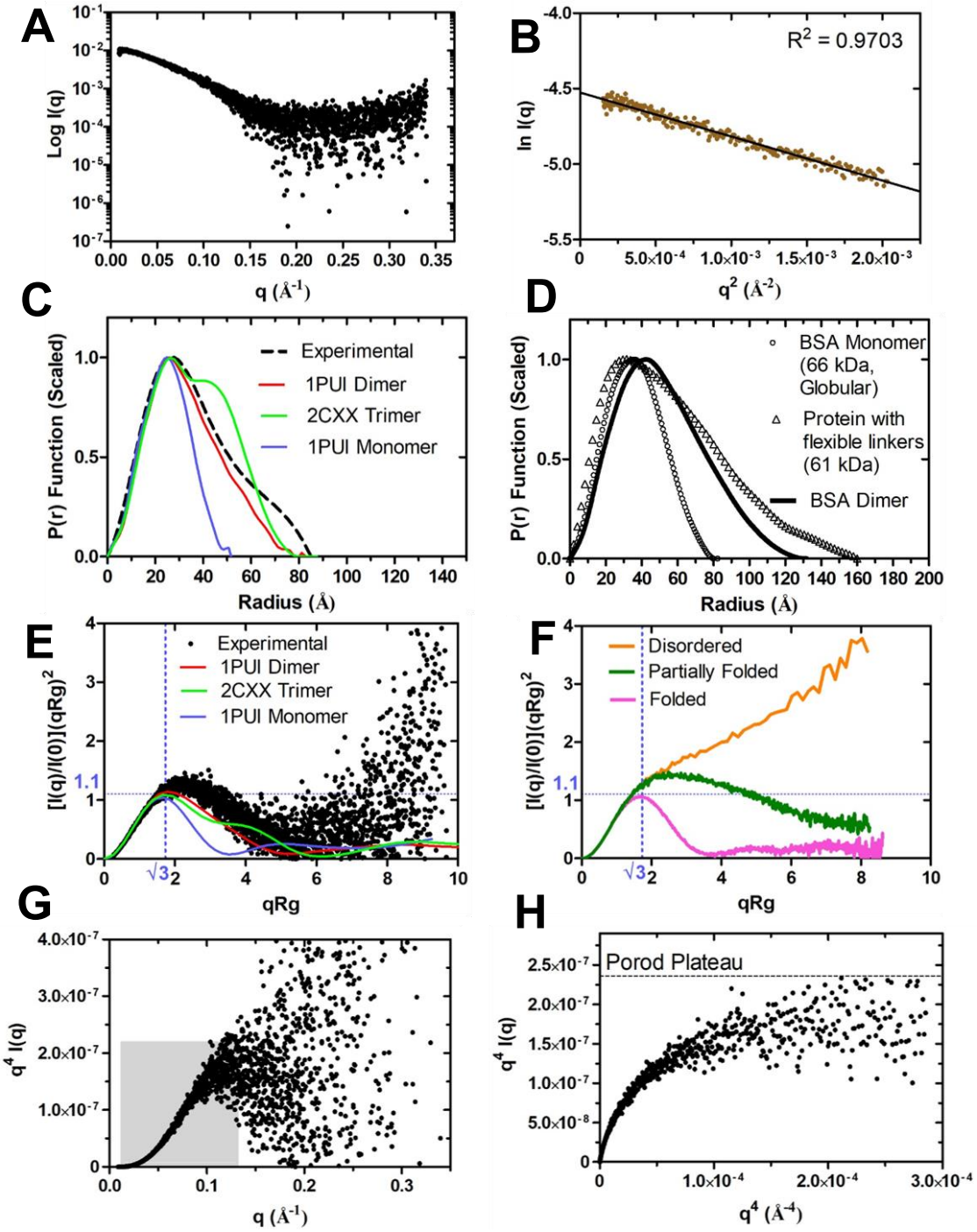


Figure 2.10. Analysis of the structure of EcYihA in solution by SAXS. (A) Intensity plot from the SAXS scattering data. (B) Guinier plot used to calculate the R_g. (C) P(r) function from SAXS data (dashed line) or calculated in CRYSOL [198] for the crystal structure of EcYihA dimer PDB ID 1PUI, EcYihA monomer PDB ID 1PUI or PhYihA trimer PDB ID 2CXX. (D) P(r) function for a globular protein (BSA monomer, SASBDB ID SASDBJ3) and for a protein with flexible linkers (p67phox, SASBDB ID SASDEL3). (E) Dimensionless Kratky plots based on the scattering data and calculated in CRYSOL for crystal structures of the YihA monomer, dimer and trimer. (F) Dimensionless Kratky plot (G) Porod plot. (H) Porod-Debye plot.

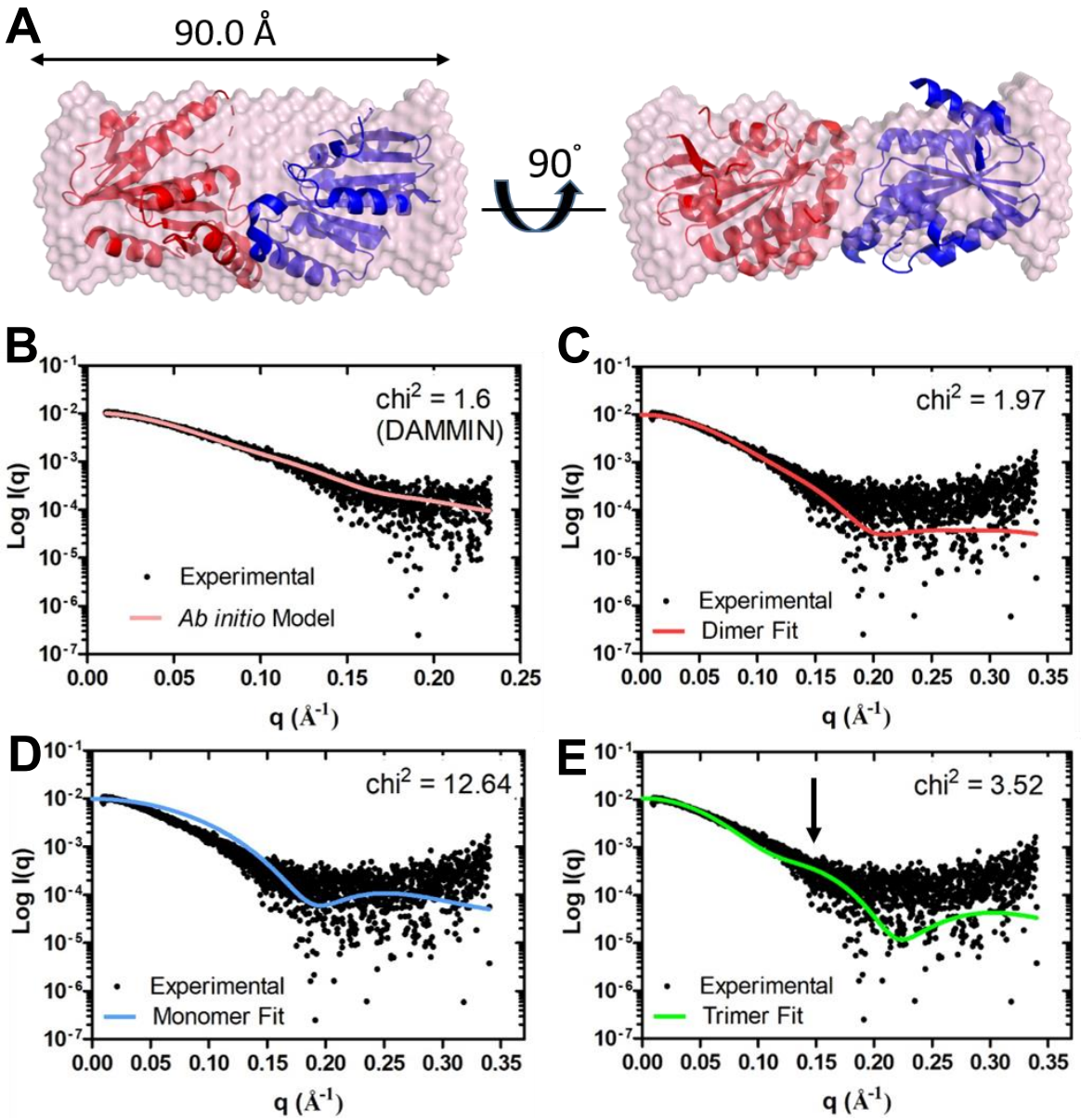


Figure 2.11. Models of YihA and their fits to the experimental SAXS curve. (A) Averaged model of YihA built from the *ab initio* SAXS envelopes. The low resolution model was superimposed on the crystal structure of EcYihA using a dimer present in the asymmetric unit. The dimeric crystal structure of EcYihA (PDB 1PUI) was docked in the SAXS model using DAMSUP [195] and a normalized spatial discrepancy (NSD) of 1.05 was obtained for the superimposition of the SAXS model with the crystal structure. The model is rotated about the x-axis by 90°. The maximum particle dimension is indicated based on the D_{max} for the model calculated in HYDROPRO (See Table 3.4). PYMOL was used to visualize the model. C-E. Fits of models of EcYihA were calculated in CRYSOL [198]. (B) Fit of the low resolution model shown in A to the SAXS data. (C) Fit of the dimeric crystal structure 1PUI of EcYihA to the SAXS data. (D) Fit of one monomer from the EcYihA crystal structure 1PUI to the SAXS data. (E) Fit of the PhYihA trimer crystal structure PDB ID 2CXX to the SAXS data. The black vertical arrow indicates the wide peak commonly seen in the scattering curves of hollow objects such as the ring-shaped trimeric crystal structure of YihA.

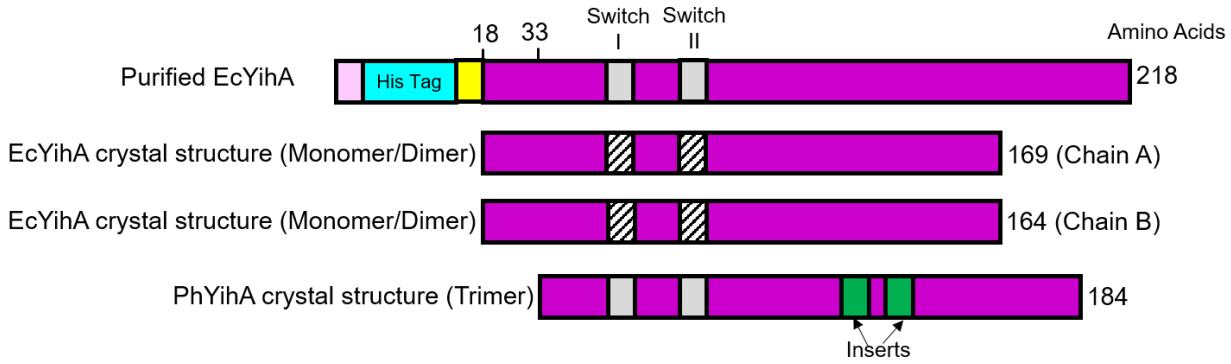


Figure 2.12. Alignment of YihA crystal structures with purified EcYihA. Domain structure of purified YihA indicating the sequences that were added to the N-terminus. The truncations in the crystal structure due to unresolved regions at the N-terminus and C-terminus are indicated.

Table 2.3 SEC-SAXS parameters for EcYihA. R_g: radius of gyration, D_{max}: maximum particle dimension, I(0): forward scattering intensity, NSD: normalized spatial discrepancy

Data Collection		SEC-SAXS
Guinier	R_g (Å)	29.76 ± 0.18
	$q.R_g$ range	0.3 – 1.3
	I(0) (Intensity extrapolated to q = 0)	$(1.1 \times 10^{-2}) \pm (3.5 \times 10^{-5})$
	Points used	7 - 251
P(r) distribution (GNOM)	R_g (Å)	28.47 ± 0.06
	I(0) (Intensity extrapolated to q = 0)	$(1.1 \times 10^{-2}) \pm (2.4 \times 10^{-5})$
	D_{max} (Å)	85.0 [b]
	Points used	7 - 1686
Ab initio modelling (DAMMIN)	Models calculated	12
	chi^2	1.6
	NSD – (discrepancy between models)	0.601 ± 0.019
Molecular weight (kDa)	Sequence	24.2
	Volume of Correlation method	39.1
	Volume method	41.9
	MM_{QP} method	42.4
	MoW method	44.9
	Size and Shape method	47.5
Ensemble Optimization Method (EOM)	chi^2	1.4
	Ensemble (pool) average R_g (Å)	28.7 (26.6)
	Ensemble (pool) D_{max} (Å)	130.7 (102.0)
	R_{flex} ensemble (pool)	73.4% (79.5%)
	R_{sigma}	0.9

[b] The D_{max} corresponds to the value where the P(r) plot sets to the x-axis after reaching a maximum. In the field, it is not established as customary to report the error in D_{max}. See the main text (section 2.3.5 and p.100) for details.

Table 2.4 Theoretical scattering parameters of available models compared to experimental SAXS parameters. Theoretical parameters were calculated in CRYSOL [198].

	Experimental	**Theoretical**				
		Monomer	**Dimer (in asymmetric unit)**	**Trimer (in asymmetric unit)**	***Ab Initio* model (Dimer)**	**Ensemble of Models (Dimer)**
Species	*E. coli*	*E. coli*	*E. coli*	*P. horikoshii*	Based on *E. coli* scattering curve	Based on *E. coli* sequence and structure
PDB ID	--	1PUI	1PUI	2CXX	--	--
R_g (Å)	28.47 ± 0.06 (P(r) distribution) 29.76 ± 0.18 (Guinier)	18.6	25.8	27.7	28.7	28.7 (average)
D_{max} (Å)	85.0	51.7	82.9	87.6	90.0	130.7 (average)
Chi^2	--	12.6	2.0	3.5	1.6	1.4

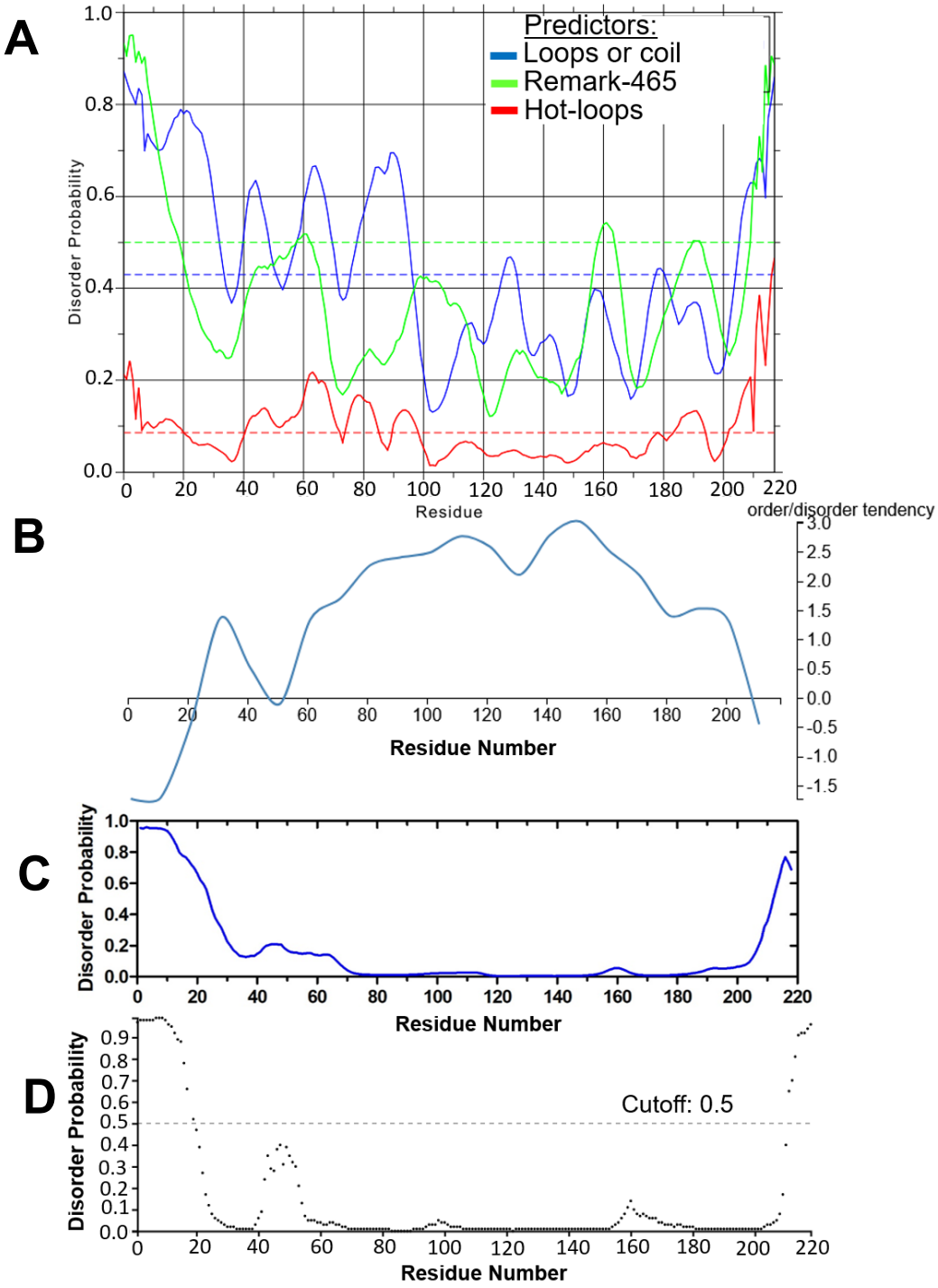


Figure 2.13. Prediction of disorder in YihA's structure. The sequence of purified EcYihA was used to predict disordered regions in YihA's structure. Dashed lines indicate the cutoff value. (A) Prediction with DisEMBL web server [273]. (B) Prediction with DICHOT web server [274]. (C) Prediction with Espritz web server [275]. (D) Prediction with DISOPRED3 web server [276].

In order to predict which regions of YihA's structure are flexible, the EcYihA crystal structure was analyzed with four different web servers (Figure 2.13). All four of the webservers predicted that the greatest disorder in YihA's structure that was above the threshold for significance was found in the N-terminus and C-terminus (Figure 2.13). The DisEMBL web server predicts structural disorder based on regions of loops (which are also referred to as coils), hot loops, which are highly mobile loops based on their B-factors, and Remark-465, which refers to regions with absent coordinates in structures found in the PDB [273].

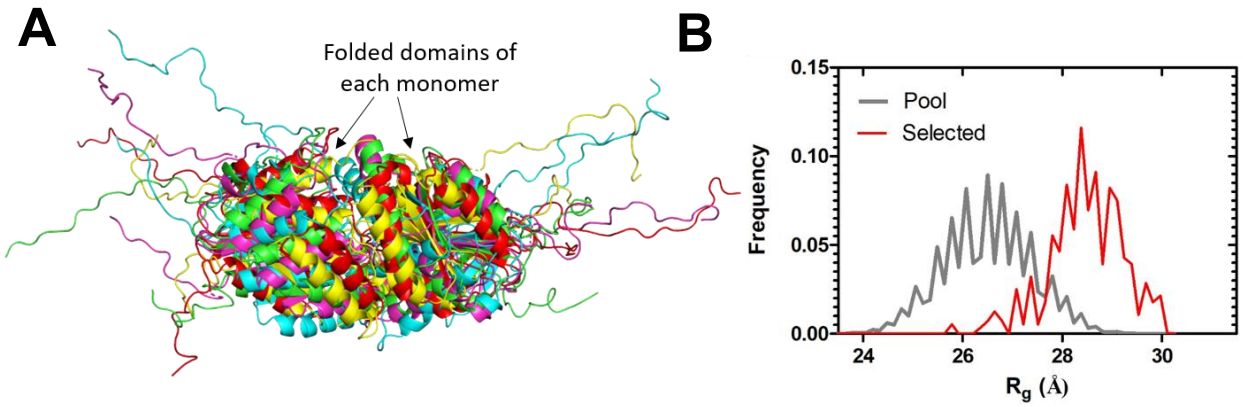


Figure 2.14. Analysis of flexibility in EcYihA using EOM. (A) Structural alignment of the top 5 representative atomistic models from the ensemble of conformations for EcYihA calculated with EOM based on the SAXS scattering curve. Alignment was performed in PyMOL and each model is colored a different color. The long tails are the N-terminal tails and the short tails are the C-terminal tails. (B) Histogram of R_g values of the pool and selected ensemble. The red trace represents the selected ensemble of models and the grey trace represents the pool of 100,000 models.

Given the prediction of structural flexibility in the N-terminal and C-terminal regions of EcYihA, an ensemble set of models was prepared from the sequence for purified EcYihA shown in Figure 2.12 using the Ensemble Optimization method (EOM) program (Figure 2.14A,B). From the pool, the EOM chooses an ensemble containing between 10 and 50 configurations that best fits the SAXS scattering profile [218]. Although the small ensemble size is inadequate to

represent the entire conformational space sampled by the protein, such an ensemble can provide an accurate distribution of the D_{max} and R_g values of the ensemble for a flexible protein [218]. The ensemble of models with flexible N-terminal and C-terminal tails fits the SAXS scattering curve with a chi^2 value of 1.4, which represents an improvement in fit over the *ab initio* model which had a chi^2 value of 1.6 (Table 2.4). The average R_g for the selected ensemble was 28.7 Å whereas the average R_g for the pool of calculated structures was only 26.6 Å (Table 2.3, Figure 2.14B). Similarly, the average D_{max} for the selected ensemble was 130.7 Å whereas the average D_{max} for the pool of structures was 102.0 Å (Table 2.3). The R_{flex} value was 73.4% for the selected ensemble and 79.5% for the pool of structures (Table 2.3). The R_{flex} value for the selected ensemble is not much smaller than the R_{flex} value for the pool of structures, which would suggest that EcYihA is fairly flexible (Table 2.3). Because the R_{sigma} value represents the variance of the selected ensemble over the variance of the pool of structures, R_{sigma} should not be above 1. The R_{sigma} was 0.9, suggesting significant flexibility compared to the pool of structures and highlighting the high quality of the data, as a value above 1 may imply a low quality SAXS data set [277] (Table 2.3). The EOM analysis suggests that EcYihA contains flexible conformations in solution (Table 2.3, Figure 2.14B). The selected ensemble is much more elongated than the conformations in the pool of structures and the selected conformations are not random (Figure 2.14B). The data shows that the D_{max} calculated from ensemble analysis of the SAXS data for EcYihA with EOM was signficantly larger than the experimental D_{max} of 85.0 Å calculated in GNOM from the logI(q) vs. q intensity plot (Table 2.4). This finding is supported by previous studies showing that D_{max} values determined by indirect Fourier transformation have been found to be much smaller than the D_{max} values characterizing the set of configurations in an ensemble of a flexible protein [278, 279].

Taken together, the SAXS data indicates that EcYihA forms dimers but not trimers or monomers in solution. There are 5 main lines of evidence from the analysis of the SAXS data

supporting the dimerization of EcYihA in solution. Firstly, the shape of the experimental P(r) distribution for EcYihA is most similar to the shape of the P(r) distribution of the dimeric crystal structure of YihA (Figure 2.10C). Secondly, the analysis of the experimental molecular weights of EcYihA from five different methods indicates that EcYihA homo-dimerizes in solution (Table 2.3). Thirdly, only the dimeric crystal structure of YihA but not the monomeric or trimeric crystal structures, will fit when superimposed on the *ab initio* model built from the SAXS data (Figure 2.11A). Fourthly, the fit of crystal structures to the experimental SAXS curve shows that the dimeric crystal structure and the *ab initio* model fit the data best with the lowest chi^2 values (Figure 2.11B-E and Table 2.4). Fifthly, comparison of the experimental R_g values for EcYihA with the theoretical R_g values for the two SAXS-based models shows that the dimeric *ab initio* model and the dimeric ensemble model have R_g values that are closest to the experimental R_g values for EcYihA (Table 2.4). As the protein chains in the monomeric, dimeric and trimeric crystal structures of YihA are much shorter than the protein chain of purified YihA (Figure 2.12) due to unresolved N-terminal and C-terminal tails in the crystal structures (Figure 2.13) and because of the 18 amino acid tag that was added to the N-terminus of purified YihA, the R_g values obtained for the crystal structures cannot be compared to the experimental R_g. Similarly, the experimental D_{max} is expected to be 5-10 Å greater than the D_{max} of the crystal structures because of the effect of the solvent shell in solution alone [280]. The error in the D_{max} values calculated from the P(r) distribution is at least 3-5 Å and may be higher [280]. Additionally, as the uncertainty in the D_{max} is normally not determined, the D_{max} values of the crystal structures and models were not compared to the experimental D_{max} values [223].

2.4 DISCUSSION

In this work, the direct trigger of YihA dimerization remains undetermined, although the preliminary evidence presented here suggests that heating to 37°C may be involved in the mechanism (Figure 2.1, Figure 2.2). There are many potential factors that could directly induce the dimerization of YihA, including the concentration of specific cations in the buffer or the dissociation of nucleotide from the nucleotide-binding pocket. As EcYihA does not contain intrinsic cysteine residues (Figure 1.3), the dimerization is not occuring via disulfide bridges.

Interestingly, the crystal structure of EcYihA (PDB 1PUI) is a dimer in the asymmetric unit and it is in the *apo* form, supporting the idea that the EcYihA dimer does not bind nucleotide (Figure 2.5). However, crystal structures available of the YihA dimer in other species are bound to nucleotide. The trimer of YihA found in *P. horikoshii* (Table 2.2) contains one nucleotide in one of the active sites (PDB ID 2CXX). Furthermore, initial evidence was shown in this chapter suggesting that YihA co-purifies with nucleotide. Previous evidence that YihA has a very high affinity for nucleotides comes from crystallization experiments with TmYihA [71]. Attempts to obtain a crystal structure of purified *apo*-TmYihA by not adding nucleotide to the crystallization buffer resulted in a structure containing a bound GDP, consistent with the suggestion that YihA co-purifies with GDP [71]. Dissociation of the GDP bound to purified TmYihA was achieved by denaturation of TmYihA in a buffer containing guanine hydrochloride followed by renaturation by dialysis in folding buffer [71]. After denaturation of purified TmYihA followed by renaturation, a crystal structure was obtained of TmYihA in the *apo* form [71].

Potentially, dimerization could be part of YihA's functional cycle. Incubation at 37°C can potentially induce formation of dimers *in vitro* (Figure 2.1, Figure 2.2). Under cellular conditions, preliminary calculations suggest that dimerization might occur *in vivo* during exponential phase (Table 2.1).

Potentially, within the cell YihA could carry out nucleotide association and dissociation when it is bound to ribosomal particles, and binding of YihA to ribosomal particles could prevent the conformational change or conformational instability in YihA that is observed in the lack of fluorescence increase in the stopped flow when YihA is heated to 37°C but not when YihA is incubated at 4°C (Figure 2.5). The conformational instability might be occuring in Switch I and Switch II of EcYihA, as these structures were not resolved in the crystal structure of EcYihA, suggesting that Switch I and Switch II are highly mobile (Figure 2.7). Switch I and Switch II are the sites where GAPs interact with GTPases and potentially the 50S ribosomal subunit could interact with Switch I and Switch II of YihA, allowing stabilization and facilitating GTP hydrolysis at 37°C (Figure A2.7A).

The data in this chapter suggests that YihA is purified from *E. coli* cells as a monomer, consistent with a previous finding that although there was a dimer of BsYsxC present in the crystal unit cell (PDB ID 1SVW) during crystallization there was no evidence of a dimer during purification [112]. If it can be predicted based on the initial evidence presented here that YihA exists partially as a dimer and partially as a monomer at 37°C, a potential explanation for why YihA exists exclusively as a monomer in the purification could be that the cells were harvested during stationary phase when it is predicted that YihA does not dimerize (Table 2.1). More work is needed to test this hypothesis.

It is interesting that, based on the Gibbs Helmhotz equation, there are two classes of proteins with the ability to dimerize in response to temperature changes. The first are exothermic binding proteins that show decreased dimerization as the temperature increases, with $\Delta H < 0$ [281]. An example of this class of proteins would be the viral envelope protein DENV2 [281]. The other class of proteins that dimerizes would be the endothermic binding proteins with $\Delta H > 0$ and that show increased dimerization as the temperature increases [282]. An example of this class of proteins are beta-crystallin found in the lens of the eye and tubulin [282]. Although more work is

needed to determine the ΔH of dimerization for YihA, it is possible that YihA belongs to the second class of dimerizing proteins. If the dimerization of YihA occurs such that the monomer and dimer are in equilibrium and the reaction is reversible, then the ΔH of dimerization for YihA could be determined with isothermal titration calorimetry (ITC) as described previously for the viral envelope protein DENV2 [281]. However, if the dimerization of YihA is irreversible, then the ΔH of dimerization for YihA cannot be determined. As enthalpy is a state function, the method that is used to calculate the ΔH for an irreversible reaction is to find a reversible path [283].

The conserved basic patch on YihA's surface consisting of residues R34, R119, H120 and K149 (*E. coli* numbering) is located within the dimerization interface of YihA (Figure 2.8, Figure 2.9). Potentially, dimerization could function to disrupt YihA's interaction with immaturely folded 23S rRNA by sequestering the conserved basic patch that could mediate specific recognition of residues in rRNA. Although a previous study found that a variant containing amino acid substitutions at each of the four residues in the basic patch retained the ability to bind to mature 50S and 30S ribosomal subunits [122], it is possible that the variant would not be able to recognize a target immaturely folded region of the 23S rRNA in a precursor of the 50S ribosomal subunit that differs from the mature ribosomal subunits in its conformation.

2.5 CONCLUSIONS

The work presented here provides structural and biochemical evidence that YihA forms a homodimer in solution. A structural model of the EcYihA dimer was determined based on SAXS analysis, yet whether the dimer forms under physiological conditions is unknown. Preliminary data

suggests that heating to 37°C may be part of the dimerization mechanism although heating to 37°C may only indirectly affect dimerization. Initial experiments indicate that YihA co-purifies with nucleotide. Dimerization of YihA does not require the presence of EDTA and preliminary evidence was provided here that dimerization occurs in the absence of release of nucleotide that co-purifies with YihA.

Furthermore, YihA purifies from the cell as a monomer. Preliminary predictions based on calculations estimated that YihA's K_D of dimerization of 1.8 μM, suggesting that YihA is able to form dimers in the cell based on the known cellular concentration of YihA. Predictions from previous data and calculations show that YihA does not dimerize during stationary phase. However, additional studies are needed to confirm that dimerization of YihA is concentration-dependent. Two salt bridges are formed in the dimer interface between K149 and D124 and between D117 and K123 and likely stabilize the dimer. Analysis of the SAXS structure data and analysis of available crystal structures of YihA monomers, dimers and trimers in the program CRYSOL suggests that the SAXS structure of YihA represents a dimer and that at high, non-physiological concentrations of YihA such as the concentrations required for SAXS analysis of YihA, YihA is able to form dimers in the absence of heating. Analysis of the SAXS data suggests that in solution YihA has a fully folded core and that both the N-terminal and C-terminal tails are flexible.

CHAPTER 3: APPENDIX DISCUSSION, FUTURE DIRECTIONS AND CONCLUSIONS

3.1 DISCUSSION AND LIMITATIONS OF APPENDIX DATA

The Appendix describes experiments that were done to elucidate the affinity of EcYihA for nucleotides, including experiments to measure intrinsic tryptophan fluorescence and FRET-based studies, performed using titrations (Figure A2.2, Figure A2.3, Figure A2.4, Figure A2.5). However, a closer look at the structural model prepared from the crystal structure of EcYihA (PDB ID 1PUI) showed that W132 and W196 are positioned on a different face of EcYihA from that of YihA's nucleotide binding site (Figure A2.1). Only one of YihA's three tryptophans, W95, is within the distance for efficient FRET of less than 25 Å from YihA's nucleotide binding site [284] (Figure A2.1). Furthermore, the structural model of EcYihA and the position of W95 in relation to the nucleotide binding site along with the conservation of tryptophan residues at position W95 in EcYihA for several bacterial species (Figure 1.3) leads to the prediction that the environment of YihA's tryptophans will be altered upon nucleotide binding and there will be a change in EcYihA's intrinsic tryptophan fluorescence in response to interaction of EcYihA with unlabelled nucleotide (Figure A2.1). YihA's tryptophans are likely not located in positions where they could act as FRET donors to the mant acceptor group of mant-labelled nucleotide or have their environment altered by binding of unlabelled nucleotide (Figure A2.1), providing an explanation as to why the fluorescence titration experiments in the Appendix did not enable the determination of a K_D of YihA for nucleotides (Figure A2.2, Figure A2.3, Figure A2.4, Figure A2.5). The structural model of EcYihA also indicates that if mant-labelled nucleotide binds to EcYihA, there could be a stabilization of the mant group leading to a measurable change in mant fluorescence (Figure A2.1). The fluorescence of YihA and mant-labelled nucleotide was measured following excitation of tryptophan and indeed an increase in mant fluorescence was

seen in the presence of YihA without the occurrence of FRET, indicating nucleotide binding (Figure A2.6). Thus, experiments can be done in the future using the stabilization of the mant group and increase in mant fluorescence upon nucleotide binding to EcYihA to obtain the K_D of EcYihA for nucleotides (Figure A2.6). Whether an assay to remove nucleotide that may co-purify with YihA would need to be done prior to measuring the K_D of EcYihA using the increase in mant fluorescence upon stabilization of the mant group remains unclear.

In order to determine whether the 50S ribosomal subunit or the 70S ribosome can act as a GAP for EcYihA, multiple turnover GTPase activity of EcYihA was measured (Figure A2.7). In comparison to HflX, for which the 50S ribosomal subunit has been shown to be a GAP [48], the observed rate constant for GTP hydrolysis by EcYihA in the presence of the 50S ribosomal subunit is approximately 1100-fold smaller than the observed rate constant for GTP hydrolysis by HflX in the presence of the 50S ribosomal subunit and no more than 2 rounds of GTP hydrolysis were carried out in 2 hours by EcYihA assuming every molecule of EcYihA was active, suggesting that the 50S ribosomal subunit is not one of YihA's physiological GAPs (Table A2.3, Figure A2.7A). Slot blotting experiments confirmed that EcYihA is able to bind to the 50S ribosomal subunit (Figure A2.8). Furthermore, EcYihA was shown to possess no intrinsic GTPase activity (Figure A2.7), in agreement with one earlier study [119]. However, one other study was able to measure intrinsic GTP hydrolysis activity for EcYihA [164], which may potentially be attributed to the use of much less highly purified EcYihA in the GTP hydrolysis assay as compared to the highly purified preparations of EcYihA used in the GTP hydrolysis assays reported in this thesis. In the case of the 70S ribosome, the ability of the 70S ribosome to provide a GAP functionality for EcYihA could not be evaluated because of the high intrinsic GTP hydrolysis activity of the 70S ribosomes used (Figure A2.7B). Previous studies with GTPases have discussed pre-treatment of 70S ribosomes with N-ethylmaleimide (NEM) in order to alkylate and inactivate any GTPases bound to the 70S ribosomes to lower or even

abolish the intrinsic GTPase activity of the 70S ribosomes [165]. If incubation of 70S ribosomes with NEM is followed by buffer exchange to remove the NEM, following NEM removal the pre-treated 70S ribosomes could potentially be used in a GTP hydrolysis assay. The technique of carrying out GTP hydrolysis assays with YihA in the presence of NEM-treated 70S ribosomes could in the future potentially be used in order to determine whether the 70S ribosome represents a GAP for EYihA.

3.2 FUTURE DIRECTIONS FOR INVESTIGATING THE MECHANISM OF DIMERIZATION OF YIHA

The mechanism of dimerization of YihA remains unknown. Many questions remain unanswered as to how dimerization affects GTP hydrolysis and nucleotide binding. Under the experimental conditions of buffer and temperature that were associated with 100% YihA monomer in Figure 2.2A, mant-labelled nucleotide association with YihA was also detected using the stopped flow assay in Figure 2.4, suggesting that the YihA monomer might not co-purify with nucleotide. Consistently with Figure 2.4, however, if YihA co-purifies with nucleotide as a dimer there would be no association of YihA with mant-labelled nucleotide observed in the stopped flow assay, as seen under the conditions of Figure 2.4, which involved heating to 37°C for 15 minutes. Thus, nucleotide dissociation from YihA under the appropriate experimental conditions *in vitro* (or due to association with a factor such as a GEF *in vivo*) might potentially provide the trigger for YihA dimers to dissociate into monomers. To investigate these questions, first, experiments could be performed in order to establish if YihA dimers or monomers co-purify with nucleotide. Fractions containing exclusively monomeric or dimeric YihA could be heated at 90°C to denature the protein and release any bound nucleotide. Alternatively, YihA could be denatured by trichloroacetic acid (TCA) precipitation. After pelleting the denatured protein, the supernatant potentially containing

released nucleotide could be injected onto a reverse phase column connected to an HPLC and the presence of a peak corresponding to GTP or GDP could be monitored by comparison with standards.

Future experiments could test whether 50S ribosomal subunit-stimulated GTP hydrolysis is increased at temperatures well below 37°C at which YihA could be active as a cold stress GTPase. Furthermore, the ability of exclusively dimeric YihA to bind the 50S precursor ribosome could be tested, providing insight into the functional role of the YihA dimer.

Future experiments could isolate the dimeric form of YihA from the monomeric YihA using size exclusion chromatography and could concentrate the dimeric YihA. If the dimeric YihA is injected back onto the Superdex-75 column and shows 100% dimer present on the column suggesting that dimerization is irreversible, the dimeric YihA could be used in the stopped flow apparatus to characterize the nucleotide binding properties of YihA that is exclusively dimeric.

Furthermore, in order to validate a temperature-dependence of dimerization, it would be necessary to establish the effect of temperature on the K_D of dimerization for YihA. Dissociation ITC experiments could be done in which YihA at a concentration 3-fold higher than its K_D of dimerization could be titrated into buffer using an injection syringe and sample cell kept at a constant temperature. ITC enables the temperature of the sample cell and titration syringe to be maintained at a selected temperature within the range from 2˚C to 80˚C. Although more noise is seen in the baseline for ITC when the temperature is set above 50°C leading to lower quality data, potentially good quality data could be collected for temperatures at and below 50°C. Dissociation ITC experiments would allow the K_D of dimerization of YihA to be determined at several different temperatures over a temperature range from 15°C to 60°C. These experiments would reveal whether temperature significantly affects the K_D of dimerization of YihA. There may be a correlation between the regulation of YihA's functional role and the heat shock response. It may

be worthwhile to investigate whether increased YihA dimerization is associated with induction of heat shock conditions in *E. coli* cells. The heat shock response is induced in *E. coli* by growing cells at 37°C followed by growth at 42°C for a transient heat shock or growth between 45°C to 50°C for a sustained heat shock response [285]. If YihA functions as a heat shock protein, the percentage of dimer should be expected to increase from about 30% to 40% observed at 37°C to up to 80% or 90% at 42°C to 46°C.

Another future direction could be to determine whether YihA dimerizes *in vivo*, as dimerization could be an artifact of the *in vitro* conditions. To do this, an *E. coli* cell strain could be prepared that contains flag-tagged YihA in the genome and a plasmid encoding for a his-tagged YihA under the control of an IPTG-inducible promoter could be transformed into the same cell strain. Affinity purification with nickel affinity chromatography would allow the His-tagged YihA to be purified from cells. Western blotting could be done with the nickel affinity purified YihA to probe for both his-tagged and flag-tagged YihA. If YihA dimerizes *in vivo,* it should be possible to detect the presence of flag-tagged YihA in the nickel affinity purified YihA with an anti-flag tag antibody. Alternatively, *in vivo* crosslinking with glutaraldehyde, which is cell permeable and has a linker length of approximately 5 Å, could be used to demonstrate whether YihA dimers are formed within the cell without the need for both flag-tagged and his-tagged YihA. Glutaraldehyde could potentially be used to crosslink YihA dimers as the salt bridges seen in the crystal structure of the dimer were 3.7 Å and 3.9 Å apart. YihA could be purified by nickel affinity chromatography and crosslinked YihA dimers could be detected by western blotting with an anti-His tag antibody.

The potential role of the YihA dimer in inhibiting the GTP hydrolysis activity of YihA has been mentioned in Chapter 3. It would be interesting to measure the GTP hydrolysis of YihA that is exclusively dimeric stimulated in the presence of 50S ribosomal subunits and to show that, as predicted, dimeric YihA has no ribosome-stimulated activity. Furthermore, it would be informative to measure the GTP hydrolysis activity of exclusively monomeric YihA as stimulated by

ribosomes. If the K_D of dimerization of YihA could be measured by analytical ultracentrifugation, then ribosome-stimulated GTP hydrolysis assays could be performed at concentrations well below the K_D of dimerization to ensure that most of YihA in the reaction would be monomeric. If the observed rate of ribosome-stimulated GTP hydrolysis by YihA were significantly increased compared to the observed rate measured in this thesis, this evidence would support the hypothesis that YihA dimers are unable to hydrolyze GTP.

In order to validate that temperature affects dimerization, experiments could be done to investigate structural changes that occur in YihA over the temperature range of dimerization. These experiments could be done with circular dichroism spectroscopy. It is predicted that changes in secondary structure of YihA would be observed over the temperature range from approximately 30°C to 45°C, consistent with a potential role for the YihA dimer in the heat shock response.

Finally, in order to validate the model of EcYihA dimerization based on the crystal structure of the dimer that suggests that two intermolecular salt bridges K149-D124 and D117-K123 are mediating the dimer interaction, charge reversal mutagenesis of the lysine residues could be done to yield a K149D K123D variant. If the variant YihA does not dimerize or shows a reduction in dimerization, this would indicate that these two salt bridges are required for formation of the dimer. To ensure that the effect of lack of dimerization is due to the loss of these salt bridges and not to a change in the structure of wild type YihA such as a folding defect, it would be necessary to demonstrate that the variant has similar properties to a wild type monomer, except for the ability to dimerize. Analytical ultracentrifugation could provide the K_D of dimerization for wildtype and variant YihA. It might be expected that the variant YihA would have a much higher K_D of dimerization than the wild type. Circular dichroism (CD) spectra could be collected for wild type YihA at concentrations at, above and below the K_D of dimerization to determine if there is any effect of dimerization on the CD spectra. CD spectra could also be obtained for the YihA variant

at concentrations well below its K_D of dimerization. The percentages of secondary structure elements could then be compared between the wild type and variant to ensure that the variant is folded the same as the wild type. For functional assays that could be performed at concentrations well below the K_D of dimerization so that most of the protein would be monomeric thus facilitating comparison of predominantly monomeric wild type to monomeric variant, the results of these assays could be compared for wild type to the variant and should be very similar. Functional assays that could be done to ensure that the variant is correctly folded (depending on the concentrations required) would include measuring GTP hydrolysis activity, binding affinity for GDP and GTP, binding affinity for ribosomes and ribosome-stimulated GTP hydrolysis. Thermal denaturation assays could be used to verify that wild type and variant YihA have similar melting temperatures and similar thermodynamic stability. The aim of all these experiments would be to determine whether the K149D K123D EcYihA variants show decreased or absent dimerization while also demonstrating that that the conformation of the variant monomer is not affected.

Furthermore, if the K149D K123D variant is not competent in dimerization, a cell strain could be made in which the K149D K123D *yihA* is inserted into the genome. Growth of the K149D K123D strain at 16°C, 37°C, and 42°C in comparison to the wild type strain would reveal whether the YihA dimer is required *in vivo* in the heat shock response. Moreover, if the formation of YihA dimers *in vivo* could be validated, then it would be important to determine whether YihA dimers are required for the role that YihA plays in ribosome biogenesis. Ribosome profiling could be performed using cultures of the strain with K149D K123D *yihA* inserted into the genome in place of wild type *yihA*. If a precursor of a ribosomal subunit accumulates in the strain with K149D K123D *yihA* in the genome but not in the wild type strain, this would indicate that YihA dimers have a role in ribosome biogenesis.

3.3 CONCLUSIONS

In **Chapter 2** of this thesis, preliminary evidence was obtained that EcYihA co-purifies with nucleotide, suggesting that the actual affinity of EcYihA for nucleotides might be in the nanomolar or picomolar range. Furthermore, the data indicates that monomeric YihA binds to nucleotides when it is incubated at 4°C followed by incubation at 20°C in the stopped flow. However, the experiments that were performed were not able to reveal whether exclusively dimeric YihA is able to bind to nucleotide. Furthermore, SAXS analysis demonstrated that the crystal structure of EcYihA that forms a dimer in the asymmetric unit is able to be docked into the SAXS envelope of EcYihA, confirming that YihA forms a dimer in solution. Heating to 37°C may be part of the mechanism that triggers dimerization of YihA under the conditions used in the dimerization assay. The mechanism of dimerization of YihA remains undetermined.

REFERENCES

[1] N. Ban, P. Nissen, J. Hansen, P.B. Moore, T.A. Steitz, The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution, Science 289(5481) (2000) 905-920.

[2] M.M. Yusupov, G.Z. Yusupova, A. Baucom, K. Lieberman, T.N. Earnest, J.H.D. Cate, H.F. Noller, Crystal structure of the ribosome at 5.5 Å resolution, Science 292(5518) (2001) 883-896.

[3] Z. Shajani, M.T. Sykes, J.R. Williamson, Assembly of bacterial ribosomes, Annual Review of Biochemistry 80 (2011) 501-526.

[4] M.L. Rodgers, S.A. Woodson, Transcription increases the cooperativity of ribonucleoprotein assembly, Cell 179(6) (2019) 1370-1381.

[5] B.A. Maguire, Inhibition of bacterial ribosome assembly: a suitable drug target?, Microbiology and Molecular Biology Reviews 73(1) (2009) 22-35.

[6] K.H. Nierhaus, F. Dohme, Total reconstitution of functionally active 50S ribosomal subunits from *Escherichia coli*, Procedings of the National Academy of Sciences of the United States of America 71(12) (1974) 4713-4717.

[7] P. Traub, M. Nomura, Structure and function of *E. coli* ribosomes, V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins, Proceedings of the National Academy of Sciences of the United States of America 59(3) (1968) 777-784.

[8] A.M. Mulder, C. Yoshioka, A.H. Beck, A.E. Bunner, R.A. Milligan, C.S. Potter, B. Carragher, J.R. Williamson, Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit, Science 330(6004) (2010) 673-677.

[9] J.H. Davis, Y.Z. Tan, B. Carragher, C.S. Potter, D. Lyumkis, J.R. Williamson, Modular assembly of the bacterial large ribosomal subunit, Cell 167(6) (2016) 1610-1622.

[10] D.L.J. Lafontaine, D. Tollervey, Ribosomal RNA, Encyclopedia of Life Sciences, 2006.

[11] A. Liiv, T. Tenson, T. Margus, J. Remme, Multiple functions of the transcribed spacers in ribosomal RNA operons, Biological Chemistry 379(7) (1998) 783-793.

[12] R. Rastogi, M. Wu, I. DasGupta, G.E. Fox, Visualization of ribosomal RNA operon copy number distribution, BMC (BioMed Central) Microbiology 9 (2009) 208.

[13] C. Condon, J. Philips, Z.Y. Fu, C. Squires, C.L. Squires, Comparison of the expression of the seven ribosomal RNA operons in *Escherichia coli*, EMBO (European Molecular Biology Organization) Journal 11(11) (1992) 4175-4185.

[14] M.P. Deutscher, Maturation and degradation of ribosomal RNA in bacteria, Progress in Molecular Biology and Translational Science 85 (2009) 369-391.

[15] P. Gegenheimer, D. Apirion, *Escherichia coli* ribosomal ribonucleic acids are not cut from an intact precursor molecule, Journal of Biological Chemistry 250(6) (1975) 2407-2409.

[16] O. Rene, J.H. Alix, Late steps of ribosome assembly in *E. coli* are sensitive to a severe heat stress but are assisted by the HSP70 chaperone machine, Nucleic Acids Research 39(5) (2011) 1855-1867.

[17] N.S. Gutgsell, C. Jain, Coordinated regulation of 23S rRNA maturation in *Escherichia coli*, Journal of Bacteriology 192(5) (2010) 1405-1409.

[18] R. Sirdeshmukh, D. Schlessinger, Why is processing of 23S ribosomal RNA in *Escherichia coli* not obligate for its function?, Journal of Molecular Biology 186(3) (1985) 669-672.

[19] R. Green, H.F. Noller, In vitro complementation analysis localizes 23S rRNA posttranscriptional modifications that are required for *Escherichia coli* 50S ribosomal subunit assembly and function, RNA 2(10) (1996) 1011-1021.

[20] R. Green, H.F. Noller, Reconstitution of functional 50S ribosomes from in vitro transcripts of *Bacillus stearothermophilus* 23S rRNA, Biochemistry 38(6) (1999) 1772-1779.

[21] D. Moazed, S. Stern, H.F. Noller, Rapid chemical probing of conformation in 16S ribosomal RNA and 30S ribosomal subunits using primer extension, Journal of Molecular Biology 187(3) (1986) 399-416.

[22] S. Mizushima, M. Nomura, Assembly mapping of 30S ribosomal proteins from *E. coli*, Nature 226(5252) (1970) 1214-1218.

[23] R.A. Britton, Role of GTPases in bacterial ribosome assembly, Annual Review of Microbiology 63 (2009) 155-176.

[24] A.E. Bunner, S. Nord, M. Wikstrom, J.R. Williamson, The effect of ribosome assembly cofactors on *in vitro* 30S subunit reconstitution, Journal of Molecular Biology 398(1) (2010) 1-7.

[25] P. Traub, M. Nomura, Structure and function of *Escherichia coli* ribosomes VI. Mechanism of assembly of 30S ribosomes studied *in vitro*, Journal of Molecular Biology 40(3) (1969) 391-413.

[26] W.A. Held, M. Nomura, Structure and function of bacterial ribosomes. Rate-determining step in the reconstitution of *Escherichia coli* 30S ribosomal subunits, Biochemistry 12(17) (1973) 3273-3281.

[27] K.L. Holmes, G.M. Culver, Mapping structural differences between 30S ribosomal subunit assembly intermediates, Nature Structural & Molecular Biology 11(2) (2004) 179-186.

[28] T. Powers, G. Daubresse, H.F. Noller, Dynamics of *in vitro* assembly of 16S ribosomal RNA into 30S ribosomal subunits, Journal of Molecular Biology 232(2) (1993) 362-374.

[29] C.E. Caldon, P.E. March, Function of the universally conserved bacterial GTPases, Current Opinion in Microbiology 6(2) (2003) 135-139.

[30] C. Maracci, M.V. Rodnina, Review: translational GTPases, Biopolymers 105(8) (2016) 463-475.

[31] N. Verstraeten, M. Fauvart, W. Versees, J. Michiels, The universally conserved prokaryotic GTPases, Microbiology and Molecular Biology Reviews 75(3) (2011) 507-542.

[32] S.R. Sprang, G protein mechanisms: Insights from structural analysis, Annual Review of Biochemistry 66 (1997) 639-678.

[33] K. Scheffzek, M.R. Ahmadian, W. Kabsch, L. Wiesmuller, A. Lautwein, F. Schmitz, A. Wittinghofer, The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants, Science 277(5324) (1997) 333-338.

[34] J. Cherfils, M. Zeghouf, Regulation of small GTPases by GEFs, GAPs, and GDIs, Physiological Reviews 93 (2013) 269-309.

[35] S.R. Sprang, G proteins, effectors and GAPs: structure and mechanism, Current Opinion in Structural Biology 7(6) (1997) 849-856.

[36] J. Cherfils, P. Chardin, GEFs: structural basis for their activation of small GTP-binding proteins, Trends in Biochemical Sciences 24(8) (1999) 306-311.

[37] J.L. Bos, H. Rehmann, A. Wittinghofer, GEFs and GAPs: critical elements in the control of small G proteins, Cell 129(5) (2007) 865-877.

[38] S. Xue, M. Barna, Specialized ribosomes: a new frontier in gene regulation and organismal biology, Nature Reviews Molecular Cell Biology 13(6) (2012) 355-369.

[39] M. Paduch, F. Jelen, J. Otlewski, Structure of small G proteins and their regulators, Acta Biochimica Polonica 48(4) (2001) 829-850.

[40] C. Robert, S. Vagner, Boosting immunity by targeting post-translational prenylation of small GTPases, Cell 175(4) (2018) 901-902.

[41] M. Geyer, A. Wittinghofer, GEFs, GAPs, GDIs and effectors: taking a closer (3D) look at the regulation of Ras-related GTP-binding proteins, Current Opinion in Structural Biology 7(6) (1997) 786-792.

[42] A. Rak, O. Pylypenko, T. Durek, A. Watzke, S. Kushnir, L. Brunsveld, H. Waldmann, R.S. Goody, K. Alexandrov, Structure of Rab GDP-dissociation inhibitor in complex with prenylated YPT1 GTPase, Science 302(5645) (2003) 646-650.

[43] D. Toniolo, P. D'Adamo, X-linked non-specific mental retardation, Current Opinion in Genetics & Development 10(3) (2000) 280-285.

[44] E.T. Eng, A.R. Jalilian, K.A. Spasov, V.M. Unger, Characterization of a novel prokaryotic GDP dissociation inhibitor domain from the G protein coupled membrane protein FeoB, Journal of Molecular Biology 375(4) (2008) 1086-1097.

[45] A. Pautsch, M. Vogelsgesang, J. Trankle, C. Herrmann, K. Aktories, Crystal structure of the C3bot-RalA complex reveals a novel type of action of a bacterial exoenzyme, EMBO (European Molecular Biology Organization) Journal 24(20) (2005) 3670-3680.

[46] G. Prehna, M.I. Ivanov, J.B. Bliska, C.E. Stebbins, *Yersinia* virulence depends on mimicry of host Rho-family nucleotide dissociation inhibitors, Cell 126(5) (2006) 869-880.

[47] T.C. Marlovits, W. Haase, C. Herrmann, S.G. Aller, V.M. Unger, The membrane protein FeoB contains an intramolecular G protein essential for Fe(II) uptake in bacteria, Proceedings of the National Academy of Sciences of the United States of America 99(25) (2002) 16243-16248.

[48] M.J. Shields, J.J. Fischer, H.J. Wieden, Toward understanding the function of the universally conserved GTPase HflX from *Escherichia coli*: a kinetic approach, Biochemistry 48(45) (2009) 10793-10802.

[49] M. Becker, K.E. Gzyl, A.M. Altamirano, A. Vuong, K. Urban, H.J. Wieden, The 70S ribosome modulates the ATPase activity of *Escherichia coli* YchF, RNA Biology 9(10) (2012) 1288-1301.

[50] J.S. Anderson, M.S. Bretscher, B.F.C. Clark, K.A. Marcker, A GTP requirement for binding initiator tRNA to ribosomes, Nature 215 (1967) 490-492.

[51] M.V. Rodnina, F. Peske, B.Z. Peng, R. Belardinelli, W. Wintermeyer, Converting GTP hydrolysis into motion: versatile translational elongation factor G, Biological Chemistry 401(1) (2020) 131-142.

[52] B. Cabrer, M.J. San-Millian, D. Vazquez, J. Modolell, Stoichiometry of polypeptide chain elongation, Journal of Biological Chemistry 251(6) (1976) 1718-1722.

[53] T. Pape, W. Wintermeyer, M.V. Rodnina, Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the *E. coli* ribosome, EMBO (European Molecular Biology Organization) Journal 17(24) (1998) 7490-7497.

[54] M.V. Rodnina, Translation in prokaryotes, Cold Spring Harbor Perspectives in Biology 10 (2018) a032664.

[55] L. Hannemann, I. Suppanz, Q.R. Ba, K. MacInnes, F. Drepper, B. Warscheid, H.G. Koch, Redox activation of the universally conserved ATPase YchF by thioredoxin 1, Antioxidants & Redox Signaling 24(3) (2016) 141-156.

[56] J.J. Fischer, M.L. Coatham, S.E. Bear, H.E. Brandon, E.I. De Laurentiis, M.J. Shields, H.J. Wieden, The ribosome modulates the structural dynamics of the conserved GTPase HflX and triggers tight nucleotide binding, Biochimie 94(8) (2012) 1647-1659.

[57] M.L. Coatham, H.E. Brandon, J.J. Fischer, T. Schummer, H.J. Wieden, The conserved GTPase HflX is a ribosome splitting factor that binds to the E-site of the bacterial ribosome, Nucleic Acids Research 44(4) (2016) 1952-1961.

[58] A. Basu, M.N. Yap, Disassembly of the *Staphylococcus aureus* hibernating 100S ribosome by an evolutionarily conserved GTPase, Proceedings of the National Academy of Sciences of the United States of America 114(39) (2017) E8165-E8173.

[59] E.L. Cooper, J. Garcia-Lara, S.J. Foster, YsxC, an essential protein in *Staphylococcus aureus* crucial for ribosome assembly/stability, BMC (BioMed Central) Microbiology 9 (2009) 266.

[60] M. Dassain, A. Leroy, L. Colosetti, S. Carolé, J.P. Bouché, A new essential gene of the 'minimal genome' affecting cell division, Biochimie 81 (1999) 889-895.

[61] T. Powers, P. Walter, Co-translational protein targeting catalyzed by the *Escherichia coli* signal recognition particle and its receptor, EMBO (European Molecular Biology Organization) Journal 16(16) (1997) 4880-4886.

[62] D. Akopian, K. Shen, X. Zhang, S.O. Shan, Signal recognition particle: an essential protein-targeting machine, Annual Review of Biochemistry 82 (2013) 693-721.

[63] K. Shazand, J. Tucker, R. Chiang, K. Stansmore, H.U. Sperling-Petersen, M. Grunberg-Manago, J.C. Rabinowitz, T. Leighton, Isolation and molecular genetic characterization of the *Bacillus subtilis* gene (*infB*) encoding protein synthesis initiation factor 2, Journal of Bacteriology 172(5) (1990) 2675-2687.

[64] G.C. Atkinson, The evolutionary and functional diversity of classical and lesser-known cytoplasmic and organellar translational GTPases across the tree of life, BMC Genomics 16(1) (2015) 78-92.

[65] G. Mittenhuber, Comparative genomics of prokaryotic GTP-binding proteins (the Era, Obg, EngA, ThdF (TrmE), YchF and YihA families) and their relationship to eukaryotic GTP-binding proteins (the DRG, ARF, RAB, RAN, RAS and RHO families), Journal of Molecular Microbiology and Biotechnology 3(1) (2001) 21-35.

[66] D.D. Leipe, Y.I. Wolf, E.V. Koonin, L. Aravind, Classification and evolution of P-loop GTPases and related ATPases, Journal of Molecular Biology 317(1) (2002) 41-72.

[67] C. Wicker-Planquart, A.E. Foucher, M. Louwagie, R.A. Britton, J.M. Jault, Interactions of an essential *Bacillus subtilis* GTPase, YsxC, with ribosomes, Journal of Bacteriology 190(2) (2008) 681-690.

[68] M.K. Kellogg, S.C. Miller, E.B. Tikhonova, A.L. Karamyshev, SRPassing co-translational targeting: the role of the signal recognition particle in protein targeting and mRNA protection, International Journal of Molecular Sciences 22(12) (2021).

[69] C. Zwieb, J. Eichler, Getting on target: the archaeal signal recognition particle, Archaea 1(1) (2002) 27-34.

[70] B.S. Laursen, K.K. Mortensen, H.U. Sperling-Petersen, D.W. Hoffman, A conserved structural motif at the N terminus of bacterial translation initiation factor IF2, Journal of Biological Chemistry 278(18) (2003) 16320-16328.

[71] K.H. Chan, K.B. Wong, Structure of an essential GTPase, YsxC, from *Thermotoga maritima*, Acta Crystallographica Section F Structural Biology and Crystallization Communications 67 (2011) 640-646.

[72] A. Teplyakov, G. Obmolova, S.Y. Chu, J. Toedt, E. Eisenstein, A.J. Howard, G.L. Gilliland, Crystal structure of the YchF protein reveals binding sites for GTP and nucleic acid, Journal of Bacteriology 185(14) (2003) 4031-4037.

[73] H. Wu, L. Sun, F. Blombach, S.J.J. Brouns, A.P.L. Snijders, K. Lorenzen, R.H.H. van den Heuvel, A.J.R. Heck, S. Fu, X. Li, X.C. Zhang, Z. Rao, J. van der Oost, Structure of the ribosome associating GTPase HflX, PROTEINS: Structure, Function, and Bioinformatics 78(3) (2010) 705-713.

[74] F. Blombach, H. Launay, V. Zorraquino, D.C. Swarts, L.D. Cabrita, D. Benelli, J. Christodoulou, P. Londei, J. van der Oost, An HflX-type GTPase from *Sulfolobus solfataricus* binds to the 50S ribosomal subunit in all nucleotide-bound states, Journal of Bacteriology 193(11) (2011) 2861-2867.

[75] J.H. Lee, S.K. Choi, A. Roll-Mecak, S.K. Burley, T.E. Dever, Universal conservation in translation initiation revealed by human and archaeal homologs of bacterial translation initiation factor IF2, Proceedings of the National Academy of Sciences of the United States of America 96(8) (1999) 4342-4347.

[76] P.A.R. Chukka, S.D. Wetmore, N. Thakor, Established and emerging regulatory roles of eukaryotic translation initiation factor 5B (eIF5B), Frontiers in Genetics 12 (2021) 737433.

[77] K.L. Harvey, V.M. Jarocki, I.G. Charles, S.P. Djordjevic, The diverse functional roles of elongation factor Tu (EF-Tu) in microbial pathogenesis, Frontiers in Microbiology 10 (2019) 2351-2370.

[78] N. Balasingam, H.E. Brandon, J.A. Ross, H.J. Wieden, N. Thakor, Cellular roles of the human Obg-like ATPase 1 (hOLA1) and its YchF homologs, Biochemistry and Cell Biology 98(1) (2020) 1-11.

[79] F. Gianfrancesco, T. Esposito, L. Montanini, A. Ciccodicola, S. Mumm, R. Mazzarella, E. Rao, S. Giglio, G. Rappold, A. Forabosco, A novel pseudoautosomal gene encoding a putative GTP-binding protein resides in the vicinity of the Xp/Yp telomere, Human Molecular Genetics 7(3) (1998) 407-414.

[80] S. Weng, S.A. Stoner, D.E. Zhang, Sex chromosome loss and the pseudoautosomal region genes in hematological malignancies, Oncotarget 7(44) (2016) 72356-72372.

[81] Y. Verma, U. Mehra, D.K. Pandey, J. Kar, X. Pérez-Martinez, S.S. Jana, K. Datta, MRX8, the conserved mitochondrial YihA GTPase family member is required for de novo Cox1 synthesis at suboptimal temperatures in *Saccharomyces cerevisiae*, Molecular Biology of the Cell 32 (2021) 21.

[82] M.R. Pool, Signal recognition particles in chloroplasts, bacteria, yeast and mammals (Review), Molecular Membrane Biology 22(1-2) (2005) 3-15.

[83] J.M. Fringer, M.G. Acker, C.A. Fekete, J.R. Lorsch, T.E. Dever, Coupled release of eukaryotic translation initiation factors 5B and 1A from 80S ribosomes following subunit joining, Molecular and Cellular Biology 27(6) (2007) 2384-2397.

[84] S. Dannenmaier, C.D. Altamirano, L. Schuler, Y. Zhang, J. Hummel, M. Milanov, S. Oeljeklaus, H.G. Koch, S. Rospert, S. Alberti, B. Warscheid, Quantitative proteomics identifies the universally conserved ATPase Ola1p as a positive regulator of heat shock response in *Saccharomyces cerevisiae*, Journal of Biological Chemistry 297(5) (2021) 101050.

[85] I.N. Suwastika, R.L. Ohniwa, K. Takeyasu, T. Shiina, Plant Drg proteins are cytoplasmic small GTPase Obg homologue, Procedia Environmental Sciences 20 (2014) 357-364.

[86] L. Zhang, X. Liu, K. Gaikwad, X. Kou, F. Wang, X. Tian, M. Xin, Z. Ni, Q. Sun, H. Peng, E. Vierling, Mutations in eIF5B confer thermosensitive and pleiotropic phenotypes via translation defects in *Arabidopsis thaliana*, The Plant Cell 29(8) (2017) 1952-1969.

[87] T. Suematsu, O. Watanabe, K. Kita, S. Yokobori, Y. Watanabe, *Arabidopsis thaliana* mitochondrial EF-G1 functions in two different translation steps, Journal of Biochemistry 155(2) (2014) 107-114.

[88] X.F. Li, C. Cai, Z. Wang, B.F. Fan, C. Zhu, Z.X. Chen, Plastid translation elongation factor Tu is prone to heat-induced aggregation despite its critical role in plant heat tolerance, Plant Physiology 176(4) (2018) 3027-3045.

[89] S.Y. Liu, L. Zheng, J. Jia, J. Guo, M.D. Zheng, J. Zhao, J.X. Shao, X.Y. Liu, L.J. An, F. Yu, Y.F. Qi, Chloroplast translation elongation factor EF-Tu/SVR11 is involved in *var2*-mediated leaf variegation and leaf development in *Arabidopsis*, Frontiers in Plant Science 10 (2019) 295.

[90] M.Y. Cheung, M.W. Li, Y.L. Yung, C.Q. Wen, H.M. Lam, The unconventional P-loop NTPase OsYchF1 and its regulator OsGAP1 play opposite roles in salinity stress tolerance, Plant, Cell & Environment 36(11) (2013) 2008-2020.

[91] I.N. Suwastika, M. Denawa, S. Yomogihara, C.H. Im, W.Y. Bang, R.L. Ohniwa, J.D. Bahk, K. Takeyasu, T. Shiina, Evidence for lateral gene transfer (LGT) in the evolution of eubacteria-derived small GTPases in plant organelles, Frontiers in Plant Science 5 (2014) 678.

[92] J. Hwang, M. Inouye, An essential GTPase, Der, containing double GTP-binding domains from *Escherichia coli* and *Thermotoga maritima*, Journal of Biological Chemistry 276(33) (2001) 31415-31421.

[93] I.J. Mehr, C.D. Long, C.D. Serkin, H.S. Seifert, A homologue of the recombination-dependent growth gene, *rdgC*, is involved in gonococcal pilin antigenic variation, Genetics 154(2) (2000) 523-532.

[94] M. Zalacain, S. Biswas, K.A. Ingraham, J. Ambrad, A. Bryant, A.F. Chalker, S. Iordanescu, J. Fan, F. Fan, R.D. Lunsford, K. O'Dwyer, L.M. Palmer, C. So, D. Sylvester, C. Volker, P. Warren, D. McDevitt, J.R. Brown, D.J. Holmes, M.K.R. Burnham, A global approach to identify novel broad-spectrum antibacterial targets among proteins of unknown function, Journal of Molecular Microbiology and Biotechnology 6(2) (2003) 109-126.

[95] V.L. Robinson, J. Hwang, E. Fox, M. Inouye, A.M. Stock, Domain arrangement of Der, a switch protein containing two GTPase domains, Structure 10(12) (2002) 1649-1658.

[96] H. Bugl, E.B. Fauman, B.L. Staker, F.H. Zheng, S.R. Kushner, M.A. Saper, J.C.A. Bardwell, U. Jakob, RNA methylation under heat shock control, Molecular Cell 6(2) (2000) 349-360.

[97] J. Tan, U. Jakob, J.C.A. Bardwell, Overexpression of two different GTPases rescues a null mutation in a heat-induced rRNA methyltransferase, Journal of Bacteriology 184(10) (2002) 2692-2698.

[98] X. Zhang, K. Yan, Y. Zhang, N. Li, C. Ma, Z. Li, Y. Zhang, B. Feng, J. Liu, Y. Sun, Y. Xu, J. Lei, N. Gao, Structural insights into the function of a unique tandem GTPase EngA in bacterial ribosome assembly, Nucleic Acids Research 42(21) (2014) 13430-13439.

[99] M. Jaskolowski, D.J.F. Ramrath, P. Bieri, M. Niemann, S. Mattei, S. Calderaro, M. Leibundgut, E.K. Horn, D. Boehringer, A. Schneider, N. Ban, Structural insights into the mechanism of mitoribosomal large subunit biogenesis, Molecular Cell 79(4) (2020) 629-644.

[100] J. Hwang, M. Inouye, The tandem GTPase, Der, is essential for the biogenesis of 50S ribosomal subunits in *Escherichia coli*, Molecular Microbiology 61(6) (2006) 1660-1672.

[101] L. Schaefer, W.C. Uicker, C. Wicker-Planquart, A.E. Foucher, J.M. Jault, R.A. Britton, Multiple GTPases participate in the assembly of the large ribosomal subunit in *Bacillus subtilis*, Journal of Bacteriology 188(23) (2006) 8252-8258.

[102] A. Bharat, M. Jiang, S.M. Sullivan, J.R. Maddock, E.D. Brown, Cooperative and critical roles for both G domains in the GTPase activity and cellular function of ribosome-associated *Escherichia coli* EngA, Journal of Bacteriology 188(22) (2006) 7992-7996.

[103] X. Ni, J.H. Davis, N. Jain, A. Razi, S. Benlekbir, A.G. McArthur, J.L. Rubinstein, R.A. Britton, J.R. Williamson, J. Ortega, YphC and YsxC GTPases assist the maturation of the central protuberance, GTPase associated region and functional core of the 50S ribosomal subunit, Nucleic Acids Research 44(17) (2016) 8442-8455.

[104] A. Seffouh, N. Jain, D. Jahagirdar, K. Basu, A. Razi, X.D. Ni, A. Guarne, R.A. Britton, J. Ortega, Structural consequences of the interaction of RbgA with a 50S ribosomal subunit assembly intermediate, Nucleic Acids Research 47(19) (2019) 10414-10425.

[105] B.Y. Feng, C.S. Mandava, Q. Guo, J. Wang, W. Cao, N.N. Li, Y.X. Zhang, Y.Q. Zhang, Z.X. Wang, J.W. Wu, S. Sanyal, J.L. Lei, N. Gao, Structural and functional insights into the mode of action of a universally conserved Obg GTPase, PLoS Biology 12(5) (2014) e1001866.

[106] A. Sayed, S. Matsuyama, M. Inouye, Era, an essential *Escherichia coli* small G-protein, binds to the 30S ribosomal subunit, Biochemical and Biophysical Research Communications 264(1) (1999) 51-54.

[107] K. Inoue, J. Alsina, J.Q. Chen, M. Inouye, Suppression of defective ribosome assembly in a rbfA deletion mutant by overexpression of Era, an essential GTPase in *Escherichia coli*, Molecular Microbiology 48(4) (2003) 1005-1016.

[108] M.R. Sharma, C. Barat, D.N. Wilson, T.M. Booth, M. Kawazoe, C. Hori-Takemoto, M. Shirouzu, S. Yokoyama, P. Fucini, R.K. Agrawal, Interaction of Era with the 30S ribosomal subunit: implications for 30S subunit assembly, Molecular Cell 18(3) (2005) 319-329.

[109] A. Jomaa, G. Stewart, J.A. Mears, I. Kireeva, E.D. Brown, J. Ortega, Cryo-electron microscopy structure of the 30S subunit in complex with the YjeQ biogenesis factor, RNA 17(11) (2011) 2026-2038.

[110] P.C. Loh, T. Morimoto, Y. Matsuo, T. Oshima, N. Ogasawara, The GTP-binding protein YqeH participates in biogenesis of the 30S ribosome subunit in *Bacillus subtilis*, Genes & Genetic Systems 82(4) (2007) 281-289.

[111] W.C. Uicker, L. Schaefer, M. Koenigsknecht, R.A. Britton, The essential GTPase YqeH is required for proper ribosome assembly in *Bacillus subtilis*, Journal of Bacteriology 189(7) (2007) 2926-2929.

[112] S.N. Ruzheinikov, S.K. Das, S.E. Sedelnikova, P.J. Baker, P.J. Artymiuk, J. Garcia-Lara, S.J. Foster, D.W. Rice, Analysis of the open and closed conformations of the GTP-binding protein YsxC from *Bacillus subtilis*, Journal of Molecular Biology 339(2) (2004) 265-278.

[113] J. Tyc, L. Novotna, P. Pena-Diaz, D.A. Maslov, J. Lukes, RSM22, mtYsxC and PNKD-like proteins are required for mitochondrial translation in *Trypanosoma brucei*, Mitochondrion 34 (2017) 67-74.

[114] A. Gupta, K. Gupta, S. Habib, YihA GTPases localize to the apicoplast and mitochondrion of the malaria parasite and interact with LSU of organellar ribosomes, Molecular and Biochemical Parasitology 236 (2020) 111265.

[115] A. Gupta, P. Shah, A. Haider, K. Gupta, M.I. Siddiqi, S.A. Ralph, S. Habib, Reduced ribosomes of the apicoplast and mitochondrion of *Plasmodium* spp. and predicted interactions with antibiotics, Open Biology 4(5) (2014) 140045.

[116] T. Cavalier-Smith, Protist phylogeny and the high-level classification of Protozoa, European Journal of Protistology 39(4) (2003) 338-348.

[117] G.N. Cohen, V. Barbe, D. Flament, M. Galperin, R. Heilig, O. Lecompte, O. Poch, D. Prieur, J. Querellou, R. Ripp, J.C. Thierry, J. Van der Oost, J. Weissenbach, Y. Zivanovic, P. Forterre, An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi*, Molecular Microbiology 47(6) (2003) 1495-1512.

[118] S.K. Das, S.E. Sedelnikova, P.J. Baker, S.N. Ruzheinikov, S.J. Foster, D.W. Rice, Expression, purification, crystallization and preliminary crystallographic analysis of a putative GTP-binding protein, YsxC, from *Bacillus subtilis*, Acta Crystallographica Section D Biological Crystallography 60 (2004) 166-168.

[119] I.E. Lehoux, M.J. Mazzulla, A. Baker, C.M. Petit, Purification and characterization of YihA, an essential GTP-binding protein from *Escherichia coli*, Protein Expression and Purification 30(2) (2003) 203-209.

[120] A. Rafay, S. Majumdar, B. Prakash, Exploring potassium-dependent GTP hydrolysis in TEES family GTPases, FEBS (Federation of European Biochemical Societies) Open Bio 2 (2012) 173-177.

[121] M.R. Ash, M.J. Maher, J. Mitchell Guss, M. Jormakka, The cation-dependent G-proteins: in a class of their own, FEBS (Federation of European Biochemical Societies) Letters 586(16) (2012) 2218-2224.

[122] C. Wicker-Planquart, N. Ceres, J.M. Jault, The C-terminal α-helix of YsxC is essential for its binding to 50S ribosome and rRNAs, FEBS (Federation of European Biochemical Societies) Letters 589(16) (2015) 2080-2086.

[123] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L.E. Felberg, D.H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D.W. Gohara, T. Dolinsky, R. Konecny, D.R. Koes, J.E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G.W. Wei, M.J. Holst, J.A. McCammon, N.A. Baker, Improvements to the APBS biomolecular solvation software suite, Protein Science 27(1) (2018) 112-128.

[124] T. Morimoto, P.C. Loh, T. Hirai, K. Asai, K. Kobayashi, S. Moriya, N. Ogasawara, Six GTP-binding proteins of the Era/Obg family are essential for cell growth in *Bacillus subtilis*, Microbiology 148 (2002) 3539-3552.

[125] L.M. Lewis, L.J. Engle, W.E. Pierceall, D.E. Hughes, K.J. Shaw, Affinity capillary electrophoresis for the screening of novel antimicrobial targets, Journal of Biomolecular Screening 9(4) (2004) 303-308.

[126] C. Wicker-Planquart, J.M. Jault, Interaction between *Bacillus subtilis* YsxC and ribosomes (or rRNAs), FEBS (Federation of European Biochemical Societies) Letters 589(9) (2015) 1026-1032.

[127] G.M. Gongadze, 5S rRNA and ribosome, Biochemistry (Moscow) 76(13) (2011) 1450-1464.

[128] A.S. Spirin, Ribosome as a molecular machine, FEBS (Federation of European Biochemical Societies) Letters 514(1) (2002) 2-10.

[129] A. Brown, A. Amunts, X.C. Bai, Y. Sugimoto, P.C. Edwards, G. Murshudov, S.H.W. Scheres, V. Ramakrishnan, Structure of the large ribosomal subunit from human mitochondria, Science 346(6210) (2014) 718-722.

[130] M.H.J. Rhodin, R. Rakauskaite, J.D. Dinman, The central core region of yeast ribosomal protein L11 is important for subunit joining and translational fidelity, Molecular Genetics and Genomics 285(6) (2011) 505-516.

[131] A.P. Korepanov, A.V. Korobeinikova, S.A. Shestakov, M.B. Garber, G.M. Gongadze, Protein L5 is crucial for *in vivo* assembly of the bacterial 50S ribosomal subunit central protuberance, Nucleic Acids Research 40(18) (2012) 9153-9159.

[132] S.J. Huang, N.A. Aleksashin, A.B. Loveland, D. Klepacki, K. Reier, A. Kefi, T. Szal, J. Remme, L. Jaeger, N. Vazquez-Laslop, A.A. Korostelev, A.S. Mankin, Ribosome engineering reveals the importance of 5S rRNA autonomy for ribosome assembly, Nature Communications 11(1) (2020) 2900.

[133] A. Seffouh, C. Trahan, T. Wasi, N. Jain, K. Basu, R.A. Britton, M. Oeffinger, J. Ortega, RbgA ensures the correct timing in the maturation of the 50S subunits functional sites, Nucleic Acids Research 50(19) (2022) 10801-10816.

[134] A. Jomaa, N. Jain, J.H. Davis, J.R. Williamson, R.A. Britton, J. Ortega, Functional domains of the 50S subunit mature late in the assembly process, Nucleic Acids Research 42(5) (2014) 3419-3435.

[135] A. Cangiani, R. Natalini, A spatial model of cellular molecular trafficking including active transport along microtubules, Journal of Theoretical Biology 267(4) (2010) 614-625.

[136] R.M. Voorhees, T.M. Schmeing, A.C. Kelley, V. Ramakrishnan, The mechanism for activation of GTP hydrolysis on the ribosome, Science 330(6005) (2010) 835-838.

[137] K.B. Gromadski, H.J. Wieden, M.V. Rodnina, Kinetic mechanism of elongation factor Ts-catalyzed nucleotide exchange in elongation factor Tu, Biochemistry 41(1) (2002) 162-169.

[138] H. Kim, S.C. Abeysirigunawarden, K. Chen, M. Mayerle, K. Ragunathan, Z. Luthey-Schulten, T. Ha, S.A. Woodson, Protein-guided RNA dynamics during early ribosome assembly, Nature 506(7488) (2014) 334-338.

[139] D. Klostermeier, P. Sears, C.H. Wong, D.P. Millar, J.R. Williamson, A three-fluorophore FRET assay for high-throughput screening of small-molecule inhibitors of ribosome assembly, Nucleic Acids Research 32(9) (2004) 2707-2715.

[140] H.D. Kim, G.U. Nienhaus, T. Ha, J.W. Orr, J.R. Williamson, S. Chu, Mg^{2+}-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized

single molecules, Proceedings of the National Academy of Sciences of the United States of America 99(7) (2002) 4284-4289.

[141] T. Arai, K. Ishiguro, S. Kimura, Y. Sakaguchi, T. Suzuki, T. Suzuki, Single methylation of 23S rRNA triggers late steps of 50S ribosomal subunit assembly, Proceedings of the National Academy of Sciences of the United States of America 112(34) (2015) E4707-E4716.

[142] G.M. Culver, H.F. Noller, Efficient reconstitution of functional *Escherichia coli* 30S ribosomal subunits from a complete set of recombinant small subunit ribosomal proteins, RNA 5(6) (1999) 832-843.

[143] K. Karbstein, Role of GTPases in ribosome assembly, Biopolymers 87(1) (2007) 1-11.

[144] A. Dönhöfer, M.R. Sharma, P.P. Datta, K.H. Nierhaus, R.K. Agrawal, D.N. Wilson, Factor-mediated ribosome assembly in bacteria, Encyclopedia of Life Sciences John Wiley & Sons, Ltd: Chichester, 2009.

[145] K. Reblova, F. Razga, W. Li, H.X. Gao, J. Frank, J. Sponer, Dynamics of the base of ribosomal A-site finger revealed by molecular dynamics simulations and cryo-EM, Nucleic Acids Research 38(4) (2010) 1325-1340.

[146] J.M. Stokes, J.H. Davis, C.S. Mangat, J.R. Williamson, E.D. Brown, Discovery of a small molecule that inhibits bacterial ribosome biogenesis, Elife 3 (2014) e03574.

[147] J.M. Stokes, C. Selin, S.T. Cardona, E.D. Brown, Chemical inhibition of bacterial ribosome biogenesis shows efficacy in a worm infection model, Antimicrobial Agents and Chemotherapy 59(5) (2015) 2918-2920.

[148] D.J. Comartin, E.D. Brown, Non-ribosomal factors in ribosome subunit assembly are emerging targets for new antibacterial drugs, Current Opinion in Pharmacology 6(5) (2006) 453-458.

[149] T.L. Campbell, J. Henderson, D.E. Heinrichs, E.D. Brown, The *yjeQ* gene is required for virulence of *Staphylococcus aureus*, Infection and Immunity 74(8) (2006) 4918-4921.

[150] R.A. Britton, B.S. Powell, S. Dasgupta, Q. Sun, W. Margolin, J.R. Lupski, D.L. Court, Cell cycle arrest in Era GTPase mutants: a potential growth rate-regulated checkpoint in *Escherichia coli*, Molecular Microbiology 27(4) (1998) 739-750.

[151] R. Bhabhra, M.D. Miley, E. Mylonakis, D. Boettner, J. Fortwendel, J.C. Panepinto, M. Postow, J.C. Rhodes, D.S. Askew, Disruption of the *Aspergillus fumigatus* gene encoding nucleolar protein CgrA impairs thermotolerant growth and reduces virulence, Infection and Immunity 72(8) (2004) 4731-4740.

[152] R. Nikolay, S. Schmidt, R. Schlomer, E. Deuerling, K.H. Nierhaus, Ribosome assembly as antimicrobial target, Antibiotics 5(2) (2016) 18.

[153] Z. Pragai, C.R. Harwood, YsxC, a putative GTP-binding protein essential for growth of *Bacillus subtilis* 168, Journal of Bacteriology 182(23) (2000) 6819-6823.

[154] C.A. Hutchison, S.N. Peterson, S.R. Gill, R.T. Cline, O. White, C.M. Fraser, H.O. Smith, J.C. Venter, Global transposon mutagenesis and a minimal mycoplasma genome, Science 286(5447) (1999) 2165-2169.

[155] F. Arigoni, F. Talabot, M. Peitsch, M.D. Edgerton, E. Meldrum, E. Allet, R. Fish, T. Jamotte, M.L. Curchod, H. Loferer, A genome-based approach for the identification of essential bacterial genes, Nature Biotechnology 16(9) (1998) 851-856.

[156] L.L. Silver, Appropriate targets for antibacterial drugs, Cold Spring Harbor Perspectives in Medicine 6(12) (2016) a030239.

[157] S.M. Prezioso, N.E. Brown, J.B. Goldberg, Elfamycins: inhibitors of elongation factor-Tu, Molecular Microbiology 106(1) (2017) 22-34.

[158] L. Zhang, N.C. Ging, T. Komoda, T. Hanada, T. Suzuki, K. Watanabe, Antibiotic susceptibility of mammalian mitochondrial translation, FEBS (Federation of European Biochemical Societies) Letters 579(28) (2005) 6423-6427.

[159] A. Goyal, K. Muthu, M. Panneerselvam, A.K. Pole, K. Ramadas, Molecular dynamics simulation of the *Staphylococcus aureus* YsxC protein: molecular insights into ribosome assembly and allosteric inhibition of the protein, Journal of Molecular Modeling 17(12) (2011) 3129-3149.

[160] T. Rudack, C. Teuber, M. Scherlo, J. Guldenhaupt, J. Schartner, M. Lubben, J. Klare, K. Gerwert, C. Kotting, The Ras dimer structure, Chemical Science 12(23) (2021) 8178-8189.

[161] S. Muratcioglu, T.S. Chavan, B.C. Freed, H. Jang, L. Khavrutskii, R.N. Freed, M.A. Dyba, K. Stefanisko, S.G. Tarasov, A. Gursoy, O. Keskin, N.I. Tarasova, V. Gaponenko, R. Nussinov, GTP-dependent K-Ras dimerization, Structure 23(7) (2015) 1325-1335.

[162] M.R. Packer, J.A. Parker, J.K. Chung, Z.L. Li, Y.K. Lee, T. Cookis, H. Guterres, S. Alvarez, M.A. Hossain, D.P. Donnelly, J.N. Agar, L. Makowski, M. Buck, J.T. Groves, C. Mattos, Raf promotes dimerization of the Ras G-domain with increased allosteric connections, Proceedings of the National Academy of Sciences of the United States of America 118(10) (2021) e2015648118.

[163] R. Gasper, S. Meyer, K. Gotthardt, M. Sirajuddin, A. Wittinghofer, It takes two to tango: regulation of G proteins by dimerization, Nature Reviews Molecular Cell Biology 10(6) (2009) 423-429.

[164] M.Y. Kong, K.G. Yan, C.Y. Ma, N. Gao, Distinct binding and enzymatic activities of two ribosome-dependent NTPases YchF and YihA, Progress in Biochemistry and Biophysics 43(6) (2016) 570-578.

[165] K.R. Legate, D.W. Andrews, The β-subunit of the signal recognition particle receptor is a novel GTP-binding protein without intrinsic GTPase activity, Journal of Biological Chemistry 278(30) (2003) 27712-27720.

[166] R.A. Kahn, A.G. Gilman, The protein cofactor necessary for ADP-ribosylation of G_s by cholera toxin is itself a GTP binding protein, Journal of Biological Chemistry 261(17) (1986) 7906-7911.

[167] T.U. Schwartz, D. Schmidt, S.G. Brohawn, G. Blobel, Homodimerization of the G protein SRβ in the nucleotide-free state involves proline cis/trans isomerization in the switch II region, Proceedings of the National Academy of Sciences of the United States of America 103(18) (2006) 6823-6828.

[168] R. Beck, Z. Sun, F. Adolf, C. Rutz, J. Bassler, K. Wild, I. Sinning, E. Hurt, B. Brugger, J. Bethune, F. Wieland, Membrane curvature induced by Arf1-GTP is essential for vesicle formation, Proceedings of the National Academy of Sciences of the United States of America 105(33) (2008) 11731-11736.

[169] J.S. Chappie, S. Acharya, M. Leonard, S.L. Schmid, F. Dyda, G domain dimerization controls dynamin's assembly-stimulated GTPase activity, Nature 465 (2010) 435-440.

[170] G.J.K. Praefcke, H.T. McMahon, The dynamin superfamily: universal membrane tubulation and fission molecules?, Nature Reviews Molecular Cell Biology 5(2) (2004) 133-147.

[171] L.J. Byrnes, H. Sondermann, Structural basis for the nucleotide-dependent dimerization of the large G protein atlastin-1/SPG3A, Proceedings of the National Academy of Sciences of the United States of America 108(6) (2011) 2216-2221.

[172] M. Chen, A. Peters, T. Huang, X. Nan, Ras dimer formation as a new signaling mechanism and potential cancer therapeutic target, Mini-Reviews in Medicinal Chemistry 16(5) (2016) 391-403.

[173] K. Scheffzek, A. Lautwein, W. Kabsch, M.R. Ahmadian, A. Wittinghofer, Crystal structure of the GTPase-activating domain of human p120GAP and implications for the interaction with Ras, Nature 384(6609) (1996) 591-596.

[174] M. Kukimoto-Niino, K. Murayama, M. Inoue, T. Terada, J.R.H. Tame, S. Kuramitsu, M. Shirouzu, S. Yokoyama, Crystal structure of the GTP-binding protein Obg from *Thermus thermophilus* HB8, Journal of Molecular Biology 337(3) (2004) 761-770.

[175] A. Sato, G. Kobayashi, H. Hayashi, H. Yoshida, A. Wada, M. Maeda, S. Hiraga, K. Takeyasu, C. Wada, The GTP binding protein Obg homolog ObgE is involved in ribosome maturation, Genes to Cells 10(5) (2005) 393-408.

[176] P. Wout, K. Pu, S.M. Sullivan, V. Reese, S. Zhou, B. Lin, J.R. Maddock, The *Escherichia coli* GTPase $CgtA_E$ cofractionates with the 50S ribosomal subunit and interacts with SpoT, a ppGpp synthetase/hydrolase, Journal of Bacteriology 186(16) (2004) 5249-5257.

[177] D.M. Raskin, N. Judson, J.J. Mekalanos, Regulation of the stringent response is the essential function of the conserved bacterial G protein CgtA in *Vibrio cholerae*, Proceedings of the National Academy of Sciences of the United States of America 104(11) (2007) 4636-4641.

[178] A. Scrima, A. Wittinghofer, Dimerisation-dependent GTPase reaction of MnmE: how potassium acts as GTPase-activating element, EMBO (European Molecular Biology Organization) Journal 25(12) (2006) 2940-2951.

[179] G. Kobayashi, S. Moriya, C. Wada, Deficiency of essential GTP-binding protein ObgE in *Escherichia coli* inhibits chromosome partition, Molecular Microbiology 41(5) (2001) 1037-1051.

[180] A. Scrima, I.R. Vetter, M.E. Armengod, A. Wittinghofer, The structure of the TrmE GTP-binding protein and its implications for tRNA modification, EMBO (European Molecular Biology Organization) Journal 24(1) (2005) 23-33.

[181] Y. Barral, M. Kinoshita, Structural insights shed light onto septin assemblies and function, Current Opinion in Cell Biology 20(1) (2008) 12-18.

[182] A.E. Foucher, J.B. Reiser, C. Ebel, D. Housset, J.M. Jault, Potassium acts as a GTPase-activating element on each nucleotide-binding domain of the essential *Bacillus subtilis* EngA, PLoS ONE 7(10) (2012) e46795.

[183] R. Vargova, J.G. Wideman, R. Derelle, V. Klimes, R.A. Kahn, J.B. Dacks, M. Elias, A eukaryote-wide perspective on the diversity and evolution of the ARF GTPase protein family, Genome Biology and Evolution 13(8) (2021) doi: 10.1093/gbe/evab157.

[184] M. Kitagawa, T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga, H. Mori, Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research, DNA Research 12(5) (2005) 291-299.

[185] E.I. De Laurentiis, E. Mercier, H.J. Wieden, The C-terminal helix of *Pseudomonas aeruginosa* Elongation Factor Ts tunes EF-Tu dynamics to modulate nucleotide exchange, Journal of Biological Chemistry 291(44) (2016) 23136-23148.

[186] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Research 31(13) (2003) 3784-3788.

[187] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, Nature Methods 9(7) (2012) 671-675.

[188] N.P. Cowieson, C.J.C. Edwards-Gayle, K. Inoue, N.S. Khunti, J. Doutch, E. Williams, S. Daniels, G. Preece, N.A. Krumpa, J.P. Sutter, M.D. Tully, N.J. Terrill, R.P. Rambo, Beamline B21: high-throughput small-angle X-ray scattering at Diamond Light Source, Journal of Synchrotron Radiation 27 (2020) 1438-1446.

[189] V. Receveur-Brechot, D. Durand, How random are intrinsically disordered proteins? a small angle scattering perspective, Current Protein and Peptide Science 13(1) (2012) 55-75.

[190] T. Mrozowich, S. McLennan, M. Overduin, T.R. Patel, Structural studies of macromolecules in solution using small angle X-Ray scattering, Journal of Visualized Experiments (141) (2018) e58538.

[191] S. Förster, L. Apostol, W. Bras, Scatter: software for the analysis of nano- and mesoscale small-angle scattering, Journal of Applied Crystallography 43(3) (2010) 639-646.

[192] P.V. Konarev, V.V. Volkov, A.V. Sokolova, M.H.J. Koch, D.I. Svergun, PRIMUS: a Windows PC-based system for small-angle scattering data analysis, Journal of Applied Crystallography 36(5) (2003) 1277-1282.

[193] D.I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, Journal of Applied Crystallography 25 (1992) 495-503.

[194] D.I. Svergun, Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing, Biophysical Journal 77(5) (1999) 2896-2896.

[195] V.V. Volkov, D.I. Svergun, Uniqueness of *ab initio* shape determination in small-angle scattering, Journal of Applied Crystallography 36 (2003) 860-864.

[196] The PyMOL molecular graphics system, version 4.3.0 Schrodinger, LLC.

[197] M.J. Gajda, D.M. Zapien, E. Uchikawa, A.C. Dock-Bregeon, Modeling the structure of RNA molecules with small-angle X-Ray scattering data, PLoS ONE 8(11) (2013) e78007.

[198] D. Svergun, C. Barberato, M.H.J. Koch, CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, Journal of Applied Crystallography 28 (1995) 768-773.

[199] J.G. de la Torre, M.L. Huertas, B. Carrasco, Calculation of hydrodynamic properties of globular proteins from their atomic-level structure, Biophysical Journal 78(2) (2000) 719-730.

[200] M.V. Petoukhov, D. Franke, A.V. Shkumatov, G. Tria, A.G. Kikhney, M. Gajda, C. Gorba, H.D.T. Mertens, P.V. Konarev, D.I. Svergun, New developments in the ATSAS program package for small-angle scattering data analysis, Journal of Applied Crystallography 45 (2012) 342-350.

[201] G. Tria, H.D.T. Mertens, M. Kachala, D.I. Svergun, Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering, International Union of Crystallography Journal 2 (2015) 207-217.

[202] P. Bernado, E. Mylonas, M.V. Petoukhov, M. Blackledge, D.I. Svergun, Structural characterization of flexible proteins using small-angle X-ray scattering, Journal of the American Chemical Society 129(17) (2007) 5656-5664.

[203] P. Weinkam, J. Pons, A. Sali, Structure-based model of allostery predicts coupling between distant sites, Proceedings of the National Academy of Sciences of the United States of America 109(13) (2012) 4875-4880.

[204] D. Schneidman-Duhovny, M. Hammel, A. Sali, FoXS: a web server for rapid computation and fitting of SAXS profiles, Nucleic Acids Research 38 (2010) W540-W544.

[205] B. Webb, A. Sali, Comparative protein structure modeling using MODELLER, Current Protocols in Bioinformatics 54 (2016) 5.6.1-5.6.37.

[206] V.P. Mysore, Z.W. Zhou, C. Ambrogio, L.B. Li, J.N. Kapp, C.Y. Lu, Q. Wang, M.R. Tucker, J.J. Okoro, G. Nagy-Davidescu, X.C. Bai, A. Pluckthun, P.A. Janne, K.D. Westover, Y.B. Shan, D.E. Shaw, A structural model of a Ras-Raf signalosome, Nature Structural & Molecular Biology 28(10) (2021) 847-857.

[207] H. Ponstingl, T. Kabir, J.M. Thornton, Automatic inference of protein quaternary structure from crystals, Journal of Applied Crystallography 36 (2003) 1116-1122.

[208] D.A. Korasick, J.J. Tanner, Determination of protein oligomeric structure from small-angle X-ray scattering, Protein Science 27(4) (2018) 814-824.

[209] E.D. Levy, PiQSi: Protein quaternary structure investigation, Structure 15(11) (2007) 1364-1367.

[210] E. Krissinel, K. Henrick, Detection of protein assemblies in crystals, Computational Life Sciences. CompLife 2005. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2005, pp. 163-174.

[211] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, Journal of Molecular Biology 372(3) (2007) 774-797.

[212] J.R. Wisniewski, D. Rakus, Quantitative analysis of the *Escherichia coli* proteome, Data in Brief 1 (2014) 7-11.

[213] A. Schmidt, K. Kochanowski, S. Vedelaar, E. Ahrne, B. Volkmer, L. Callipo, K. Knoops, M. Bauer, R. Aebersold, M. Heinemann, The quantitative and condition-dependent *Escherichia coli* proteome, Nature Biotechnology 34(1) (2016) 104-110.

[214] Y. Ishihama, T. Schmidt, J. Rappsilber, M. Mann, F.U. Hartl, M.J. Kerner, D. Frishman, Protein abundance profiling of the *Escherichia coli* cytosol, BMC (BioMed Central) Genomics 9 (2008) 102.

[215] A. Sukhwal, R. Sowdhamini, Oligomerisation status and evolutionary conservation of interfaces of protein structural domain superfamilies, Molecular Biosystems 9(7) (2013) 1652-1661.

[216] D. Clementel, A. Del Conte, A.M. Monzon, G.F. Camagni, G. Minervini, D. Piovesan, S.C.E. Tosatto, RING 3.0: fast generation of probabilistic residue interaction networks from structural ensembles, Nucleic Acids Research 50(W1) (2022) W651-W656.

[217] Y.X. Liu, D.T. Huynh, T.O. Yeates, A 3.8 Å resolution cryo-EM structure of a small protein bound to an imaging scaffold, Nature Communications 10 (2019) 1864.

[218] A.G. Kikhney, D.I. Svergun, A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins, FEBS (Federation of European Biochemical Societies) Letters 589(19) (2015) 2570-2577.

[219] M. Meier, A. Moya-Torres, N.J. Krahn, M.D. McDougall, G.L. Orriss, E.K.S. McRae, E.P. Booy, K. McEleney, T.R. Patel, S.A. McKenna, J. Stetefeld, Structure and hydrodynamics of a DNA G-quadruplex with a cytosine bulge, Nucleic Acids Research 46(10) (2018) 5319-5331.

[220] A. Guinier, G. Fournet, Small-angle scattering of X-rays, John Wiley & Sons, New York, 1955.

[221] H.D.T. Mertens, D.I. Svergun, Structural characterization of proteins and complexes using small-angle X-ray solution scattering, Journal of Structural Biology 172(1) (2010) 128-141.

[222] T.D. Grant, J.R. Luft, L.G. Carter, T. Matsui, T.M. Weiss, A. Martel, E.H. Snell, The accurate assessment of small-angle X-ray scattering data, Acta Crystallographica Section D Biological Crystallography 71 (2015) 45-56.

[223] D.A. Jacques, J. Trewhella, Small-angle scattering for structural biology - Expanding the frontier while avoiding the pitfalls, Protein Science 19(4) (2010) 642-657.

[224] C. Alsaker, F.J. Breidt, M.J. van der Woerd, Minimum mean squared error estimation of the radius of gyration in small-angle X-ray scattering experiments, Journal of the American Statistical Association 114(525) (2019) 39-47.

[225] R. Balu, R. Knott, N.P. Cowieson, C.M. Elvin, A.J. Hill, N.R. Choudhury, N.K. Dutta, Structural ensembles reveal intrinsic disorder for the multi-stimuli responsive bio-mimetic protein Rec1-resilin, Scientific Reports 5 (2015) 10896.

[226] S.P. Meisburger, W.C. Thomas, M.B. Watkins, N. Ando, X-ray scattering studies of protein structural dynamics, Chemical Reviews 117(12) (2017) 7615-7672.

[227] C.D. Putnam, Guinier peak analysis for visual and automated inspection of small-angle X-ray scattering data, Journal of Applied Crystallography 49 (2016) 1412-1419.

[228] S. Doniach, J. Lipfert, Use of small angle X-ray scattering (SAXS) to characterize conformational states of functional RNAs, Methods in Enzymology 469 (2009) 237-251.

[229] T. Mrozowich, A. Henrickson, B. Demeler, T.R. Patel, Nanoscale structure determination of Murray Valley encephalitis and Powassan virus non-coding RNAs, Viruses 12 (2020) 190.

[230] D.N. Kim, B.C. Thiel, T. Mrozowich, S.P. Hennelly, I.L. Hofacker, T.R. Patel, K.Y. Sanbonmatsu, Zinc-finger protein CNBP alters the 3-D structure of lncRNA *Braveheart* in solution, Nature Communications 11(1) (2020) 148.

[231] M.H. D'Souza, T. Mrozowich, M.D. Badmalia, M. Geeraert, A. Frederickson, A. Henrickson, B. Demeler, M.T. Wolfinger, T.R. Patel, Biophysical characterisation of human LincRNA-p21 sense and antisense Alu inverted repeats, Nucleic Acids Research 50(10) (2022) 5881-5898.

[232] R. Aguilar, K.B. Spencer, B. Kesner, N.F. Rizvi, M.D. Badmalia, T. Mrozowich, J.D. Mortison, C. Rivera, G.F. Smith, J. Burchard, P.J. Dandliker, T.R. Patel, E.B. Nickbarg, J.T. Lee, Targeting *Xist* with compounds that disrupt RNA structure and X inactivation, Nature 604(7904) (2022) 160-166.

[233] D.I. Svergun, M.H.J. Koch, Small-angle scattering studies of biological macromolecules in solution, Reports on Progress in Physics 66(10) (2003) 1735-1782.

[234] D. Durand, C. Vives, D. Cannella, J. Perez, E. Pebay-Peyroula, P. Vachette, F. Fieschi, NADPH oxidase activator P67(phox) behaves in solution as a multidomain protein with semi-flexible linkers, Journal of Structural Biology 169(1) (2010) 45-53.

[235] R.K. Singh, J.D. Larson, W.D. Zhu, R.P. Rambo, G.L. Hura, D.F. Becker, J.J. Tanner, Small-angle X-ray scattering studies of the oligomeric state and quaternary structure of the trifunctional proline utilization A (PutA) flavoprotein from *Escherichia coli*, Journal of Biological Chemistry 286(50) (2011) 43144-43153.

[236] M.A. Behrens, T.J. Sendall, J.S. Pedersen, M. Kjeldgaard, J.A. Huntington, J.K. Jensen, The shapes of z-alpha(1)-antitrypsin polymers in solution support the C-terminal domain-swap mechanism of polymerization, Biophysical Journal 107(8) (2014) 1905-1912.

[237] A.S. Solovyova, D.T. Peters, G. Dura, H. Waller, J.H. Lakey, D.A. Fulton, Probing the oligomeric re-assembling of bacterial fimbriae *in vitro*: a small-angle X-ray scattering and analytical ultracentrifugation study, European Biophysics Journal 50 (2021) 597-611.

[238] K.T. Halloran, Y.M. Wang, K. Arora, S. Chakravarthy, T.C. Irving, O. Bilsel, C.L. Brooks, C.R. Matthews, Frustration and folding of a TIM barrel protein, Proceedings of the National Academy of Sciences of the United States of America 116(33) (2019) 16378-16383.

[239] D. Franke, M.V. Petoukhov, P.V. Konarev, A. Panjkovich, A. Tuukkanen, H.D.T. Mertens, A.G. Kikhney, N.R. Hajizadeh, J.M. Franklin, C.M. Jeffries, D.I. Svergun, ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions, Journal of Applied Crystallography 50 (2017) 1212-1225.

[240] D. Franke, C.M. Jeffries, D.I. Svergun, Machine learning methods for X-ray scattering data analysis from biomacromolecular solutions, Biophysical Journal 114(11) (2018) 2485-2492.

[241] N.R. Hajizadeh, D. Franke, C.M. Jeffries, D.I. Svergun, Consensus Bayesian assessment of protein molecular mass from solution X-ray scattering data, Scientific Reports 8 (2018) 7204.

[242] R.P. Rambo, J.A. Tainer, Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law, Biopolymers 95(8) (2011) 559-571.

[243] V.M. Burger, D.J. Arenas, C.M. Stultz, A structure-free method for quantifying conformational flexibility in proteins, Scientific Reports 6 (2016) 29040.

[244] Y. Peng, S.F. Cao, J. Kiselar, X.Z. Xiao, Z.W. Du, A. Hsieh, S. Ko, Y.H. Chen, P. Agrawal, W.W. Zheng, W.X. Shi, W. Jiang, L. Yang, M.R. Chance, W.K. Surewicz, M. Buck, S.C. Yang, A metastable contact and structural disorder in the estrogen receptor transactivation domain, Structure 27(2) (2019) 229-240.

[245] K. Moncoq, I. Broutin, C.T. Craescu, P. Vachette, A. Ducruix, D. Durand, SAXS study of the PIR domain from the Grb14 molecular adaptor: a natively unfolded protein with a transient structure primer?, Biophysical Journal 87(6) (2004) 4056-4064.

[246] K. Moncoq, I. Broutin, V. Larue, D. Perdereau, K. Cailliau, E. Browaeys-Poly, A.F. Burnol, A. Ducruix, The PIR domain of Grb14 is an intrinsically unstructured protein: implication in insulin signaling, FEBS (Federation of European Biochemical Societies) Letters 554(3) (2003) 240-246.

[247] E. Deyaert, L. Wauters, G. Guaitoli, A. Konijnenberg, M. Leemans, S. Terheyden, A. Petrovic, R. Gallardo, L.M. Nederveen-Schippers, P.S. Athanasopoulos, H. Pots, P.J.M. Van Haastert, F. Sobott, C.J. Gloeckner, R. Efremov, A. Kortholt, W. Versees, A homologue of the Parkinson's disease-associated protein LRRK2 undergoes a monomer-dimer transition during GTP turnover, Nature Communications 8 (2017) 1008.

[248] M.A. Graewert, S. Da Vela, T.W. Grawert, D.S. Molodenskiy, C.E. Blanchet, D.I. Svergun, C.M. Jeffries, Adding size exclusion chromatography (SEC) and light scattering (LS) devices to obtain high-quality small angle X-ray scattering (SAXS) data, Crystals 10(11) (2020) 975.

[249] R.P. Rambo, Considerations for sample preparation using size-exclusion chromatography for home and synchrotron sources, in: B. Chaudhuri, I. Munoz, S. Qian, V. Urban (Eds.) Biological Small Angle Scattering: Techniques, Strategies and Tips, Springer, Singapore, 2017.

[250] A. Koul, D. Gemmill, N. Lubna, M. Meier, N. Krahn, E.P. Booy, J. Stetefeld, T.R. Patel, S.A. McKenna, Structural and hydrodynamic characterization of dimeric human oligoadenylate synthetase 2, Biophysical Journal 118(11) (2020) 2726-2740.

[251] K.T. Powers, E.D. Layering, M.T. Washington, Conformational flexibility of ubiquitin-modified and SUMO-modified PCNA shown by full-ensemble hybrid methods, Journal of Molecular Biology 430(24) (2018) 5294-5303.

[252] W. Bunnak, A.J. Winter, C.M. Lazarus, M.P. Crump, P.R. Race, P. Wattana-Amorn, SAXS reveals highly flexible interdomain linkers of tandem acyl carrier protein-thioesterase domains from a fungal nonreducing polyketide synthase, FEBS (Federation of European Biochemical Societies) Letters 595(1) (2021) 133-144.

[253] C.S. Tome, R.R. Lopes, P.M.F. Sousa, M.P. Amaro, J. Leandro, H.D.T. Mertens, P. Leandro, J.B. Vicente, Structure of full-length wild-type human phenylalanine hydroxylase by small angle X-ray scattering reveals substrate-induced conformational stability, Scientific Reports 9 (2019) 13615.

[254] D. Johansen, J. Trewhella, D.P. Goldenberg, Fractal dimension of an intrinsically disordered protein: Small-angle X-ray scattering and computational study of the bacteriophage λ N protein, Protein Science 20(12) (2011) 1955-1970.

[255] F.E. Reyes, C.R. Schwartz, J.A. Tainer, R.P. Rambo, Methods for using new conceptual tools and parameters to assess RNA structure by small-angle X-ray scattering, Methods in Enzymology 549 (2014) 235-263.

[256] C.A. Brosey, J.A. Tainer, Evolving SAXS versatility: solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology, Current Opinion in Structural Biology 58 (2019) 197-213.

[257] W.G. Saw, G. Tria, A. Gruber, M.S. Subramanian Manimekalai, Y.Q. Zhao, A. Chandramohan, G.S. Anand, T. Matsui, T.M. Weiss, S.G. Vasudevan, G. Gruber, Structural insight and flexible features of NS5 proteins from all four serotypes of *Dengue virus* in solution, Acta Crystallographica Section D Biological Crystallography 71 (2015) 2309-2327.

[258] M.H.J. Koch, P. Vachette, D.I. Svergun, Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution, Quarterly Reviews of Biophysics 36(2) (2003) 147-227.

[259] M.B. Kozin, D.I. Svergun, Automated matching of high- and low-resolution structural models, Journal of Applied Crystallography 34 (2001) 33-41.

[260] E. Mylonas, D.I. Svergun, Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering, Journal of Applied Crystallography 40 (2007) s245-s249.

[261] M. Hammel, M. Kriechbaum, A. Gries, G.M. Kostner, P. Laggner, R. Prassl, Solution structure of human and bovine beta(2)-glycoprotein I revealed by small-angle X-ray scattering, Journal of Molecular Biology 321(1) (2002) 85-97.

[262] M. Kozak, Glucose isomerase from *Streptomyces rubiginosus* - potential molecular weight standard for small-angle X-ray scattering, Journal of Applied Crystallography 38 (2005) 555-558.

[263] H. Fischer, M.D. Neto, H.B. Napolitano, I. Polikarpov, A.F. Craievich, Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale, Journal of Applied Crystallography 43 (2010) 101-109.

[264] R.P. Rambo, J.A. Tainer, Accurate assessment of mass, models and resolution by small-angle scattering, Nature 496(7446) (2013) 477-481.

[265] D.I. Svergun, L.A. Feĭgin, G.W. Taylor, Structure analysis by small-angle X-ray and neutron scattering, Plenum Press, New York, 1987.

[266] C.D. Putnam, M. Hammel, G.L. Hura, J.A. Tainer, X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution, Quarterly Reviews of Biophysics 40(3) (2007) 191-285.

[267] M. Guttman, P. Weinkam, A. Sali, K.K. Lee, All-atom ensemble modeling to analyze small-angle X-ray scattering of glycosylated proteins, Structure 21(3) (2013) 321-331.

[268] G.L. Hura, A.L. Menon, M. Hammel, R.P. Rambo, F.L. Poole, S.E. Tsutakawa, F.E. Jenney, S. Classen, K.A. Frankel, R.C. Hopkins, S.J. Yang, J.W. Scott, B.D. Dillard, M.W.W. Adams, J.A. Tainer, Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS), Nature Methods 6(8) (2009) 606-612.

[269] K. Pearson, On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have

arisen from random sampling, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 50(302) (1900) 157-175.

[270] D. Franke, C.M. Jeffries, D.I. Svergun, Correlation Map, a goodness-of-fit test for one-dimensional X-ray scattering spectra, Nature Methods 12(5) (2015) 419-422.

[271] M. Honda, T. Fujisawa, T. Shibata, T. Mikawa, RecR forms a ring-like tetramer that encircles dsDNA by forming a complex with RecF, Nucleic Acids Research 36(15) (2008) 5013-5020.

[272] T. Bizien, D. Durand, P. Roblin, A. Thureau, P. Vachette, J. Perez, A brief survey of state-of-the-art BioSAXS, Protein and Peptide Letters 23(3) (2016) 217-231.

[273] R. Linding, L.J. Jensen, F. Diella, P. Bork, T.J. Gibson, R.B. Russell, Protein disorder prediction: implications for structural proteomics, Structure 11(11) (2003) 1453-1459.

[274] S. Fukuchi, K. Hosoda, K. Homma, T. Gojobori, K. Nishikawa, Binary classification of protein molecules into intrinsically disordered and ordered segments, BMC (BioMed Central) Structural Biology 11 (2011) 29.

[275] I. Walsh, A.J.M. Martin, T. Di Domenico, S.C.E. Tosatto, ESpritz: accurate and fast prediction of protein disorder, Bioinformatics 28(4) (2012) 503-509.

[276] D.T. Jones, D. Cozzetto, DISOPRED3: precise disordered region predictions with annotated protein-binding activity, Bioinformatics 31(6) (2015) 857-863.

[277] S.K. Ravala, J.B. Hopkins, C.B. Plescia, S.R. Allgood, M.A. Kane, J.N. Cash, R.V. Stahelin, J.J.G. Tesmer, The first DEP domain of the RhoGEF P-Rex1 autoinhibits activity and contributes to membrane binding, Journal of Biological Chemistry 295(36) (2020) 12635-12647.

[278] P. Bernado, Effect of interdomain dynamics on the structure determination of modular proteins by small-angle scattering, European Biophysics Journal 39(5) (2010) 769-780.

[279] W.T. Heller, Influence of multiple well defined conformations on small-angle scattering of proteins in solution, Acta Crystallographica Section D Biological Crystallography 61 (2005) 33-44.

[280] A. Grishaev, Sample preparation, data collection, and preliminary data analysis in biomolecular solution X-ray scattering, Current Protocols in Protein Science 70 (2012) 17.14.1-17.14.18.

[281] S.T. Kudlacek, L. Premkumar, S.W. Metz, A. Tripathy, A.A. Bobkov, A.M. Payne, S. Graham, J.A. Brackbill, M.J. Miley, A.M. de Silva, B. Kuhlman, Physiological temperatures reduce dimerization of dengue and Zika virus recombinant envelope proteins, Journal of Biological Chemistry 293(23) (2018) 8922-8933.

[282] Y.V. Sergeev, J.F. Hejtmancik, P.T. Wingfield, Energetics of domain-domain interactions and entropy driven association of β-crystallins, Biochemistry 43(2) (2004) 415-424.

[283] G.G. Hammes, Thermodynamics and kinetics for the biological sciences, John Wiley & Sons, New York, 2000.

[284] A. Wagner, I. Simon, M. Sprinzl, R.S. Goody, Interaction of guanosine nucleotides and their analogs with elongation factor Tu from *Thermus thermophilus*, Biochemistry 34(39) (1995) 12535-12542.

[285] S. Lindquist, The heat-shock response, Annual Review of Biochemistry 55 (1986) 1151-1191.

[286] C. Lenzen, R.H. Cool, H. Prinz, J. Kuhlmann, A. Wittinghofer, Kinetic analysis by fluorescence of the interaction between Ras and the catalytic domain of the guanine nucleotide exchange factor $Cdc25^{Mm}$, Biochemistry 37(20) (1998) 7420-7430.

[287] J. John, R. Sohmen, J. Feuerstein, R. Linke, A. Wittinghofer, R.S. Goody, Kinetics of interaction of nucleotides with nucleotide-free H-Ras P21, Biochemistry 29(25) (1990) 6058-6065.

[288] B.R. Bochner, B.N. Ames, Complete analysis of cellular nucleotides by two-dimensional thin-layer chromatography, Journal of Biological Chemistry 257(16) (1982) 9759-9769.

[289] B. Wilden, A. Savelsbergh, M.V. Rodnina, W. Wintermeyer, Role and timing of GTP binding and hydrolysis during EF-G-dependent tRNA translocation on the ribosome, Proceedings of the National Academy of Sciences of the United States of America 103(37) (2006) 13670-13675.

[290] J. Hwang, M. Inouye, A bacterial GAP-like protein, YihI, regulating the GTPase of Der, an essential GTP-binding protein in *Escherichia coli*, Journal of Molecular Biology 399(5) (2010) 759-772.

[291] P. Milon, A.L. Konevega, F. Peske, A. Fabbretti, C.O. Gualerzi, M.V. Rodnina, Transient kinetics, fluorescence, and FRET in studies of initiation of translation in bacteria, Methods in Enzymology 430 (2007) 1-30.

[292] J.E. Lusk, R.J. Williams, E.P. Kennedy, Magnesium and the growth of *Escherichia coli*, Journal of Biological Chemistry 243(10) (1968) 2618-2624.

[293] J.R. Jagath, M.V. Rodnina, G. Lentzen, W. Wintermeyer, Interaction of guanine nucleotides with the signal recognition particle from *Escherichia coli*, Biochemistry 37(44) (1998) 15408-15413.

[294] D. Elseviers, L.A. Petrullo, P.J. Gallagher, Novel *Escherichia coli* mutants deficient in biosynthesis of 5-methylaminomethyl-2-thiouridine, Nucleic Acids Research 12(8) (1984) 3521-3534.

[295] H. Cabedo, F. Macian, M. Villarroya, J.C. Escudero, M. Martinez-Vicente, E. Knecht, M.E. Armengod, The *Escherichia coli trmE* (*mnmE*) gene, involved in tRNA modification, codes for an evolutionarily conserved GTPase with unusual biochemical properties, EMBO (European Molecular Biology Organization) Journal 18(24) (1999) 7063-7076.

[296] M. Fislage, L. Wauters, W. Versees, MnmE, a GTPase that drives a complex tRNA modification reaction, Biopolymers 105(8) (2016) 568-579.

[297] T. Kawashima, C. Berthet-Colominas, M. Wulff, S. Cusack, R. Leberman, The structure of the *Escherichia coli* EF-Tu · EF-Ts complex at 2.5 Å resolution, Nature 379(6565) (1996) 511-518.

[298] Y. Wang, Y.X. Jiang, M. Meyering-Voss, M. Sprinzl, P.B. Sigler, Crystal structure of the EF-Tu · EF-Ts complex from *Thermus thermophilus*, Nature Structural Biology 4(8) (1997) 650-656.

[299] M.H. Buckstein, J. He, H. Rubin, Characterization of nucleotide pools as a function of physiological state in *Escherichia coli*, Journal of Bacteriology 190(2) (2008) 718-726.

[300] A.T.Y. Lo, H.Y. Tan, P.R. Bianco, Advanced uses of IMAC affinity chromatography, Austin Chromatography 3(1) (2016) 1042.

[301] R.L. Woodbury, S.J.S. Hardy, L.L. Randall, Complex behavior in solution of homodimeric SecA, Protein Science 11(4) (2002) 875-882.

[302] K. Scheffzek, M.R. Ahmadian, A. Wittinghofer, GTPase-activating proteins: helping hands to complement an active site, Trends in Biochemical Sciences 23(7) (1998) 257-262.

[303] K. Ogata, R. Ohno, R. Morishita, Y. Endo, A. Wada, Studies on ATPase(GTPase) intrinsic to *E. coli* ribosomes, Journal of Biochemistry 128(2) (2000) 309-313.

[304] K.E. Gzyl, H.J. Wieden, Tetracycline does not directly inhibit the function of bacterial elongation factor Tu, PLoS ONE 12(5) (2017) e0178523.

[305] K.S. Rosler, E. Mercier, I.C. Andrews, H.J. Wieden, Histidine 114 is critical for ATP hydrolysis by the universally conserved ATPase YchF, Journal of Biological Chemistry 290(30) (2015) 18650-18661.

[306] J.B. Gibbs, I.S. Sigal, M. Poe, E.M. Scolnick, Intrinsic GTPase activity distinguishes normal and oncogenic ras P21 molecules, Proceedings of the National Academy of Sciences of the United States of America 81(18) (1984) 5704-5708.

[307] S.M. Chen, H.E. Takiff, A.M. Barber, G.C. Dubois, J.C. Bardwell, D.L. Court, Expression and characterization of RNase III and Era proteins. Products of the rnc operon of *Escherichia coli*, Journal of Biological Chemistry 265(5) (1990) 2888-2895.

[308] F. Pompeo, C. Freton, C. Wicker-Planquart, C. Grangeasse, J.M. Jault, A. Galinier, Phosphorylation of CpgA protein enhances both its GTPase activity and its affinity for ribosome and is crucial for *Bacillus subtilis* growth and morphology, Journal of Biological Chemistry 287(25) (2012) 20830-20838.

[309] M.V. Rodnina, A. Savelsbergh, V.I. Katunin, W. Wintermeyer, Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome, Nature 385(6611) (1997) 37-41.

[310] D. Voet, J.G. Voet, Biochemistry, 3rd ed., John Wiley & Sons, New York, 2004.

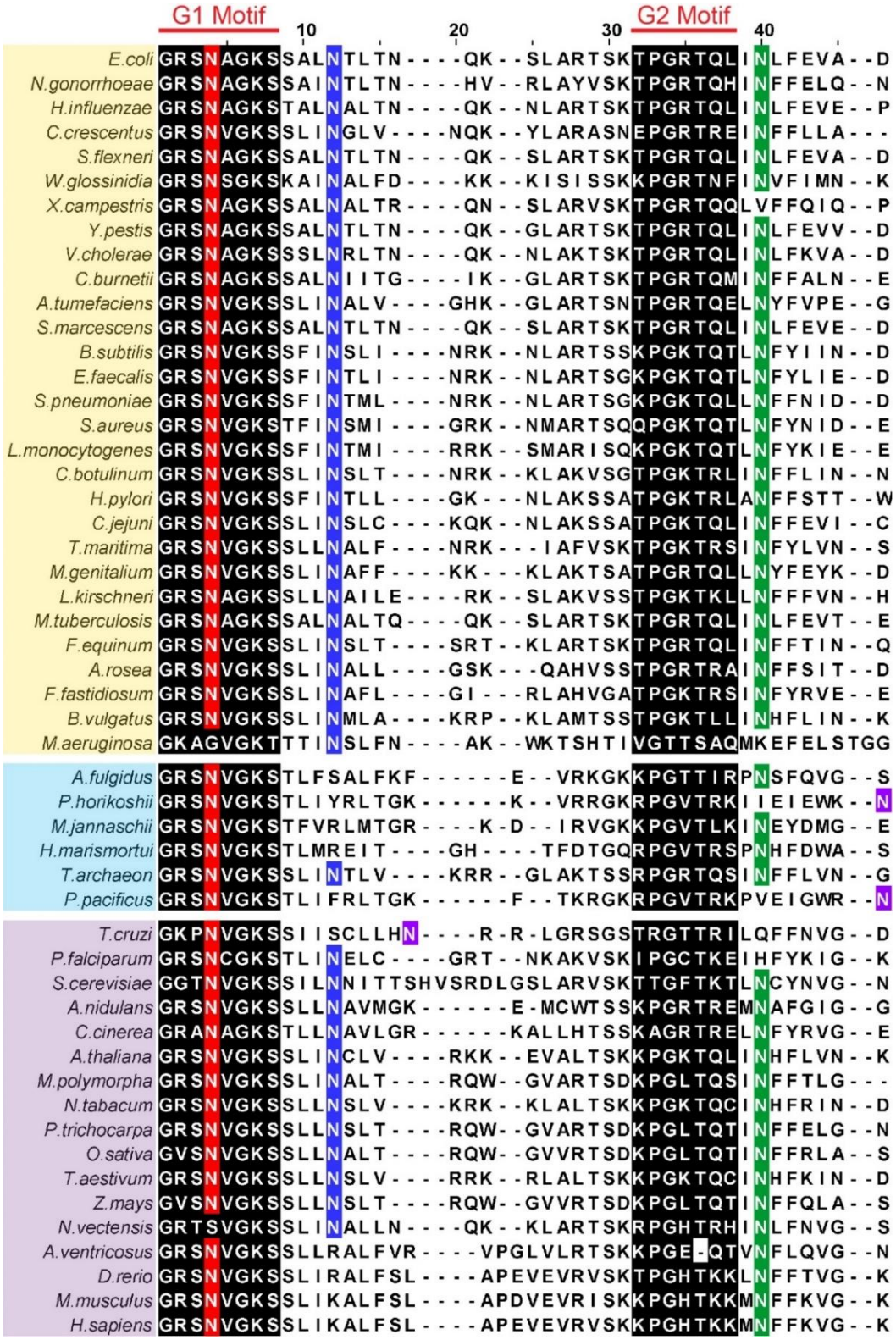
APPENDIX TO

INITIAL CHARACTERIZATION OF THE FUNCTIONAL CYCLE OF YIHA

BY DORA CAPATOS

SECTION A1 – INTRODUCTION APPENDIX

A



[Figure A on previous page]

B

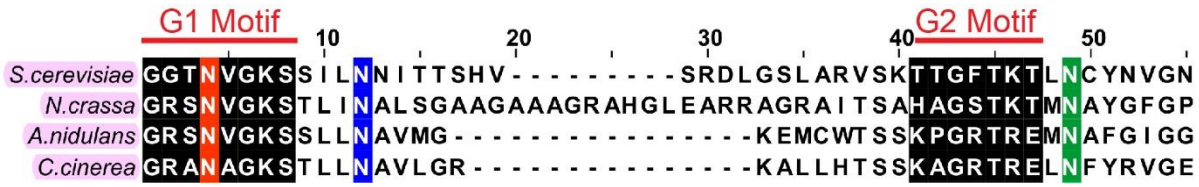


Figure A1.1 Sequence alignment of the G1 and G2 motifs of YihA homologues. Organisms in the domain Bacteria are highlighted in pale yellow, archaeal organisms are highlighted in blue and eukaryotes are highlighted in purple. (A) Homologues of YihA from across all three domains of life were aligned. The G1 and G2 motifs are shaded in black. The first conserved asparagine residue (closer to the N-terminus and N36 in *E. coli*) is shaded red. The second conserved asparagine residue (closer to the C-terminus) is shaded blue. Green shading indicates the asparagine residues following the G2 motif and other asparagine residues located nearby but in different positions are shaded purple. The species are listed in the alignment in an order based on the order that they are shown in the phylogenetic tree in Appendix Figure 1.2. (B) YihA homologues in the Kingdom fungi were aligned with the sequence of *Neurospora crassa*. The sequence alignment was generated using M-Coffee.

Table A1.1. **Species names, phyla, common names if used and accession numbers for the multiple sequence alignment in Figure A1.1**. The abbreviation UP indicates a UniProt database accession number whereas NC indicates a NCBI database accession number. Organisms in the domain Bacteria are highlighted in pale yellow, archaeal organisms are highlighted in blue and eukaryotes are highlighted in purple. Grey highlighting indicates that a common name is not used.

Species Name	Phylum	Common Name	Accession Number
Escherichia coli	Pseudomonadota		P0A6P7.1 (NC)
Neisseria gonorrhoeae	Pseudomonadota		5UCV-B (NC)
Haemophilus influenzae	Pseudomonadota		P46453 (UP)
Caulobacter crescentus	Pseudomonadota		Q9ZG89 (UP)
Shigella flexneri	Pseudomonadota		Q83PF4 (UP)
Wigglesworthia glossinidia	Pseudomonadota		Q8D338 (UP)
Xanthomonas campestris	Pseudomonadota		Q8P5E4 (UP)
Yersinia pestis	Pseudomonadota		Q8ZJS0 (UP)
Vibrio cholerae	Pseudomonadota		Q9KVN0 (UP)
Coxiella burnetii	Pseudomonadota		B6J2X4 (UP)
Agrobacterium tumefaciens	Pseudomonadota		Q8UIB4 (UP)
Serratia marcescens	Pseudomonadota		A0A0G8BBD0 (UP)
Bacillus subtilis	Bacillota		PWI62610.1 (NC)
Enterococcus faecalis	Bacillota		Q833M8 (UP)
Streptococcus pneumoniae	Bacillota		P64073 (UP)
Staphylococcus aureus	Bacillota		W8TSV3 (UP)
Listeria monocytogenes	Bacillota		Q8Y6X3 (UP)
Clostridium botulinum	Bacillota		B2TPB6 (UP)
Helicobacter pylori	Campylobacterota		O26087 (UP)
Campylobacter jejuni	Campylobacterota		Q9PHL7 (UP)
Thermotoga maritima	Thermotogota		3PQC-A (NC)
Mycoplasma genitalium	Mycoplasmatota		P47577 (UP)
Leptospira kirschneri	Spirochaetota		A0A1T1DGV8 (UP)
Mycobacterium tuberculosis	Actinomycetota		SGD52286.1 (NC)
Fusobacterium equinum	Fusobacteriota		A0A133NH15 (UP)
Acidipila rosea	Acidobacteriota		A0A4R1LDE7_9BACT (UP)
Fretibacterium fastidiosum	Synergistota		D4MAH3_9BACT (UP)
Bacteroides vulgatus	Bacteroidota		D4VDM0 (UP)
Microcystis aeruginosa	Cyanobacteria		A0A402DC49 (UP)

Continued on next page

Species Name	Phylum	Common Name	Accession Number
Archeoglobus fulgidus	Euryarchaeota		KUK06096.1 (NC)
Pyrococcus horikoshii	Euryarchaeota		WP_010884308.1 (NC)
Methanocaldococcus jannaschii	Euryarchaeota		AAB98305.1 (NC)
Haloarcula marismortui	Euryarchaeota		Q5UZW0 (UP)
Thermoplasmata archaeon	Euryarchaeota		A0A662L476 (UP)
Palaeococcus pacificus	Euryarchaeota		A0A075LUL9 (UP)
Trypanosoma cruzi	Euglenozoa		Q4D4S5 (UP)
Plasmodium falciparum	Apicomplexa		A0A0L7KKQ3 (UP)
Saccharomyces cerevisiae	Ascomycota		AAB64772.1 (NC)
Aspergillus nidulans	Ascomycota		Q5B1R2 (UP)
Neurospora crassa	Ascomycota	Bread mold	A0A0B0DTY8 (UP)
Coprinopsis cinerea	Basidiomycota	Inky cap fungus	A8NAA1 (UP)
Arabidopsis thaliana	Tracheophyta	Thale cress	O81004.2 (NC)
Marchantia polymorpha	Hepatophyta	Common liverwort	A0A176WGW6 (UP)
Nicotiana tabacum	Magnoliophyta	Tobacco	A0A1S3YA83 (UP)
Populus trichocarpa	Spermatophyta	Black cottonwood	A0A2K1WTA6 (UP)
Oryza sativa	Spermatophyta	Rice	B9F6Y6 (UP)
Triticum aestivum	Spermatophyta	Common wheat	A0A3B6GVW0 (UP)
Zea mays L.	Spermatophyta	Maize	A0A317Y2J6 (UP)
Nematostella vectensis	Cnidaria	Starlet sea anemone	A7SLL1 (UP)
Araneus ventricosa	Arthropoda	Orbweaver spider	A0A4Y2CKE6 (UP)
Danio rerio	Chordata	Zebrafish	NP_001119874.1 (NC)
Mus musculus	Chordata	Mouse	EDK98048.1 (NC)
Homo sapiens	Chordata	Human	NP_054889.2 (NC)

CHAPTER A2: INSIGHT INTO THE FUNCTIONAL CYCLE OF YIHA

A2.1 INTRODUCTION

The superfamily of guanine nucleotide triphosphatases exhibit much variation in their functional cycle based on their differing cellular roles. EF-Tu and Ras depend on GEFs as part of their functional cycle since they have low rates of nucleotide exchange and very low K_Ds for binding guanine nucleotides (in the nanomolar to picomolar range) [137, 286, 287]. On the other hand, several other G-proteins have a high rate of GDP dissociation and they can depend on the much higher cellular concentration of GTP compared to GDP to achieve the GTP-bound "ON" state such that a GEF is not needed [288]. Examples of the latter class of GTPases are EF-G [289] and HflX [48].

Nucleotide binding to GTPases depends on the coordination of a magnesium ion. GTPases contain a G3 motif shown in some GTPases to contain a threonine residue that coordinates the magnesium ion. There is also an interaction between the magnesium ion and the γ-phosphate of the nucleotide. The available crystal structure of YsxC bound to GDP contains a coordinated magnesium ion, but the corresponding crystal structure of EcYihA does not have a magnesium ion bound [112]. Examination of the crystal structures of BsYsxC reveals magnesium in the GDPNP-bound state [112]. In contrast, TmYsxC with GDP bound has no magnesium ion present in the active site [71]. Comparison of the structures of several GTPase-GDP complexes with and without magnesium bound indicates that Switch II demonstrates a structural rearrangement that is dependent on the presence of the magnesium cation [112].

Previous work has demonstrated that addition of 300 mM KCl, NH_4Cl or NaCl in the buffer yielded the same GTP hydrolysis activity of EcYihA, suggesting that catalysis of GTP hydrolysis

by YihA is not dependent on a particular monovalent cation [164]. However, there is no evidence for the effect of the divalent cation magnesium on GTP hydrolysis or nucleotide binding by YihA.

Some GTPases, like Ras, have very low intrinsic GTP hydrolysis activity that is ultimately stimulated by signalling proteins that act as GAPs as described in Chapter 1. For example, YihI acts as the GAP for EngA [290].

Previously, there were conflicting reports in the literature concerning whether EcYihA possesses intrinsic GTP hydrolysis activity. EcYihA was shown in one report to have no intrinsic GTP hydrolysis activity [119]. However, another study reported that the GTP hydrolysis activity of YihA is stimulated by 70S ribosomes [164]. This finding was in direct opposition to their finding that YihA interacts with 50S ribosomes but not with 30S ribosomal subunits or with 70S ribosomes *in vitro* [164]. Moreover, previous studies with BsYsxC showed that YihA's GTP hydrolysis activity is stimulated by the 50S ribosomal subunit [103]. Here, the objective was to test these findings and to provide a preliminary characterization of the activity and nucleotide binding properties of YihA towards understanding the functional cycle of YihA. It is hypothesized in this chapter that EcYihA`s GTP hydrolysis activity can be stimulated by the 50S ribosomal subunit and the 70S ribosome and that these serve as GAPs. Here, work is presented to test this hypothesis.

A2.2 MATERIALS AND METHODS

A2.2.1 REAGENTS, PLASMIDS AND CELL STRAINS

2'-/3'-O-N'-methylanthraniloyl-GDP/GTP/ADP (mantGDP/GTP/ADP) were purchased from Jena Bioscience. GDP and GTP were purchased from Sigma. Radiolabelled GTP was obtained from Perkin Elmer. Unless specified otherwise, reagents were obtained from BioBasic.

A2.2.2 TRANSFORMATION, PROTEIN OVEREXPRESSION AND PURIFICATION

Refer to section 2.2.2 of the main body of this thesis.

A2.2.3 EQUILIBRIUM FLUORESCENCE EXPERIMENTS

All measurements were collected on a Photon Technology International (PTI) QuantaMaster Spectrofluorometer. For excitation of intrinsic tryptophan, 1 μM EcYihA was added to a 18F-Q-1-MS 2 mm x 10 mm quartz cuvette (Starna Cells) in TAKM7 buffer (50 mM Tris-Cl, pH 7.5 at 4°C, 70 mM NH_4Cl, 30 mM KCl, 7 mM $MgCl_2$). Increasing concentrations of unlabelled GDP or GTP were added stepwise to the cuvette followed by excitation at 280 nm and recording of fluorescence emission spectra at room temperature.

The following quadratic function was fitted to the data:

$$F = 0.5\left(\frac{F_{max} - F_0}{X}\right)\left\{K_D + X + C - \left(\sqrt{(K_D + X + C)^2 - 4XC}\right)\right\} \quad \text{(Equation A2.1)}$$

where F_{max} indicates the maximum fluorescence, F_0 indicates the initial fluorescence, X indicates the concentration of nucleotide, C indicates the protein concentration and K_D indicates the equilibrium dissociation constant.

Fitting of the data was performed with Prism (GraphPad) and TableCurve (Jandel Scientific). Each fluorescence value obtained was corrected for dilution.

For fluorescence resonance energy transfer (FRET) experiments, mant-labelled nucleotide was titrated into 0.5 μM YihA in a 0.3 x 0.3 mm quartz cuvette (Starna Cells). YihA's three intrinsic tryptophans were excited at 280 nm leading to FRET between the tryptophans and the mant

group of mant-GDP/GTP/ADP. Fluorescence emission spectra were collected at room temperature.

The following hyperbolic function was fitted to the data:

$$F = \frac{F_0 + \Delta F_{max} C}{K_D + C} \qquad \text{(Equation A2.2)}$$

where C indicates the concentration of mant-labelled nucleotide, ΔF_{max} indicates the amplitude of the fluorescence change, F_0 indicates the initial fluorescence and K_D indicates the equilibrium dissociation constant.

For the method of measuring the fluorescence of individual samples, the following quadratic equation was fitted to the data:

$$Y = F_0 + (F_{max} - F_0)\frac{(C + X + K_D) - \sqrt{(C + X + K_D)^2 - 4CX}}{2C} \qquad \text{(Equation A2.3)}$$

where F_{max} represents the maximum fluorescence, F_0 represents the initial fluorescence, C represents concentration of nucleotide, X represents the total protein concentration and K_D represents the dissociation constant. Each fluorescence value obtained was corrected for dilution.

To control for the fluorescence due to YihA in the absence of nucleotide, buffer alone was titrated into YihA and the resulting data were subtracted from the data for the titration of nucleotide

into YihA. Similarly, a control for nucleotide alone without YihA was also subtracted from the data for the titration of nucleotide into YihA.

A2.2.4 PREPARATION OF RIBOSOMES

Purification of vacant ribosomes from *E. coli* MRE600 cells was performed as described previously [291].

A2.2.5 NUCLEOTIDE HYDROLYSIS ASSAYS

Refer to section 2.2.8 of the main body of this thesis. Where indicated, each reaction contained 1 μM 70S *E. coli* ribosomes or 1 μM 50S *E. coli* ribosomal subunits.

A2.2.6 MICROFILTRATION BINDING ASSAYS

Reactions were set up in which 5 μM YihA purified in magnesium or 5 μM YihA purified without magnesium was incubated with 1 μM 50S ribosomal subunits for 15 minutes at 37°C in TAKM7 buffer (50 mM Tris-HCl, 70 mM NH_4Cl, 30 mM KCl, 7 mM $MgCl_2$). Reactions were transferred to Vivaspin 500 microfiltration devices (GE Healthcare) and centrifuged at 10,000 x g for 10 minutes. The flow through was collected and the reactions above the filter were washed once by addition of 500 μL of TAKM7 buffer and centrifuged again at 10,000 x g for 10 minutes. The reaction remaining above the filter and the flow through from the wash were collected and used for western blotting.

A2.2.7 WESTERN BLOTTING WITH SLOT BLOT APPARATUS

Samples from above the filter or from the flow through for two washes as well as the initial samples before microfiltration were applied to a nitrocellulose membrane using a slot blot apparatus (BioRad Biodot SF). The binding of YihA to the ribosomes was determined by incubation of the blot with primary anti-His tag antibody to detect the His tag at the N-terminus of YihA. The anti-His tag antibody was diluted 1:3000 in Pacific® brand skim milk powder.

A2.3 RESULTS

A2.3.1 NUCLEOTIDE BINDING TO YIHA

As YihA is a member of the family of P-loop GTPases, the hypothesis is put forward in this thesis chapter that EcYihA should be able to bind to nucleotides. The objective of this Appendix chapter was to determine the guanine nucleotide binding affinity of EcYihA and to measure observed rates for intrinsic and GAP-stimulated GTP hydrolysis by EcYihA. EcYihA possesses three intrinsic tryptophan residues, which might enable measurement of its ligand binding affinity through changes in the intrinsic tryptophan fluorescence of YihA at different concentrations of unlabelled nucleotide (Figure A2.1). Alternatively, if there is no change in the environment of YihA's intrinsic tryptophans upon binding of unlabelled nucleotide, Fluorescence Resonance Energy Transfer (FRET) from YihA's tryptophans to mant-labelled nucleotide might potentially provide a method to demonstrate that EcYihA is able to interact with ligand and to measure the affinity of EcYihA's interaction with nucleotides.

To investigate the feasibility of using changes in intrinsic tryptophan fluorescence of YihA or the possibility of using Fluorescence Resonance Energy Transfer (FRET) to determine the affinity of EcYihA for nucleotides, a space filling model was prepared showing the positions of YihA's three tryptophans relative to the position of GDP bound to the nucleotide binding site (Figure A2.1). In the model, two of the tryptophans W132 and W196 are located on the back side

of EcYihA such that they are on the side of YihA opposite to the nucleotide binding site and are separated from the nucleotide binding site by distances greater than 28 Å (Figure A2.1), suggesting that detection of a FRET signal from W132 or W196 to a bound mant-nucleotide will not be possible. YihA's third tryptophan, W95, is located on the same face of YihA as the nucleotide binding site (Figure A2.1). Examination of the space filling model of EcYihA in Figure A2.1 revealed that W95 is located less than 25 Å from YihA's nucleotide binding site which is within the distance for efficient FRET from W95 to a mant-nucleotide bound in YihA's nucleotide binding pocket [284]. Therefore, this thesis chapter hypothesizes that there will be a FRET signal measured from W95 of YihA to a mant-labelled nucleotide bound to YihA's nucleotide binding pocket. Based on the conservation of tryptophan residues at position W95 in EcYihA for all the species of bacteria represented in the sequence alignment in Figure 1.3 along with a tryptophan at that position in *Arabidopsis thaliana*, the additional hypothesis was formulated that changes in the intrinsic fluorescence of W95 in response to interaction with nucleotide might be used to determine the affinity of EcYihA for nucleotides.

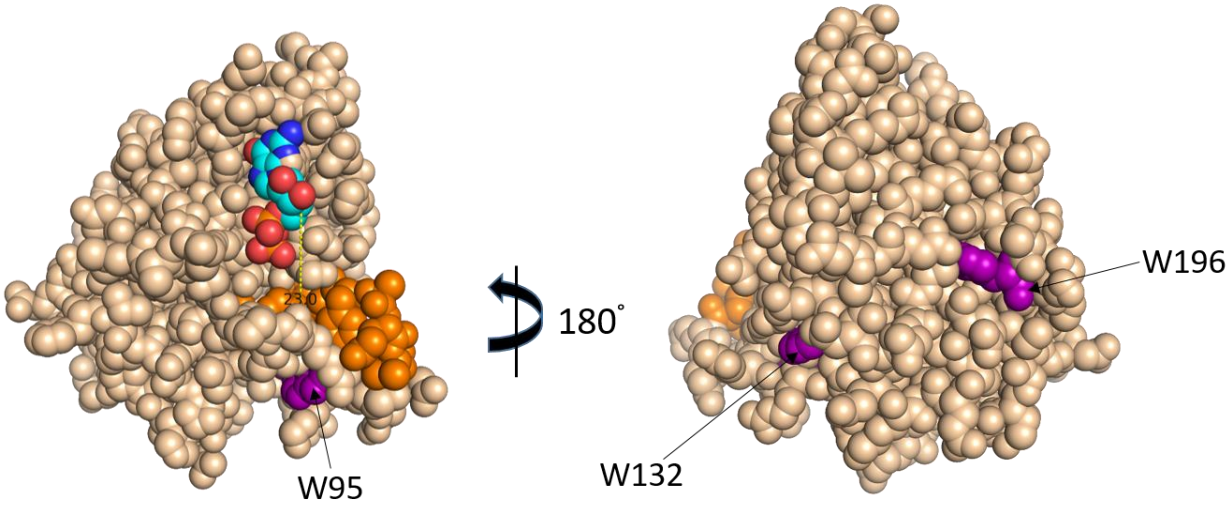


Figure A2.1. Space filling model of EcYihA predicts that FRET from YihA's W95 to mant-nucleotide will occur. Two different orientations of YihA. The front face of YihA shows the nucleotide binding pocket and a rotation of the model by 180° gives the back face of YihA that is opposite to the nucleotide binding site. EcYihA's three intrinsic tryptophan residues, W95, W132 and W196 are labelled and shown in purple. GDP is shown with carbons colored cyan, nitrogens colored blue and oxygens colored red. The Switch II loop is depicted in orange and all remaining residues in EcYihA are colored tan. The model was prepared by superimposing the structure of apo-EcYihA (PDB ID 1PUI) on the structure of *Neisseria gonorrhoeae* YihA with GDP bound (PDB ID 5UCV). Missing regions of the EcYihA structure were modelled with the program Modeller using Chimera.

In order to test the hypothesis that binding of unlabelled nucleotide to YihA might cause a change in the intrinsic tryptophan fluorescence of YihA, the first approach was to design and carry out equilibrium fluorescence titrations of EcYihA with unlabelled nucleotides. Although this approach has not been reported previously in the literature to be applied to study the nucleotide binding affinity of YihA, this approach has been used to study the binding affinity for nucleotides of other GTPases [48, 49] and here this approach was attempted in order to determine equilibrium dissociation constants (K_Ds) for EcYihA for guanine nucleotides. The emission intensity maximum of 331 nm was determined in the emission spectrum for an excitation wavelength of 280 nm when GDP or GTP was titrated into YihA (Figure A2.2A-B). A concentration of 7 mM magnesium was included in the TAKM7 buffer containing magnesium in order to be consistent with experimental

work performed previously on the GTPases HflX and YchF [48, 49] and because 7 mM magnesium is relatively close to the cellular concentration of free magnesium in the *E. coli* cell which is approximately 4 mM [292]. For titrations of both GDP and GTP, plotting the fluorescence intensity at 331 nm against the concentration of nucleotide resulted in a concentration dependent decrease in tryptophan fluorescence (Figure A2.2). However, after subtracting the background fluorescence due to titrating buffer into YihA from the fluorescence intensity, the resulting concentration dependence could not be fit with a quadratic function (Equation A2.1) and showed a nearly linear relationship (Figure A2.2C). Thus, it was not possible to determine the equilibrium dissociation constant of YihA for GDP and GTP using changes in the intrinsic tryptophan fluorescence of YihA upon binding of unlabelled nucleotide to YihA.

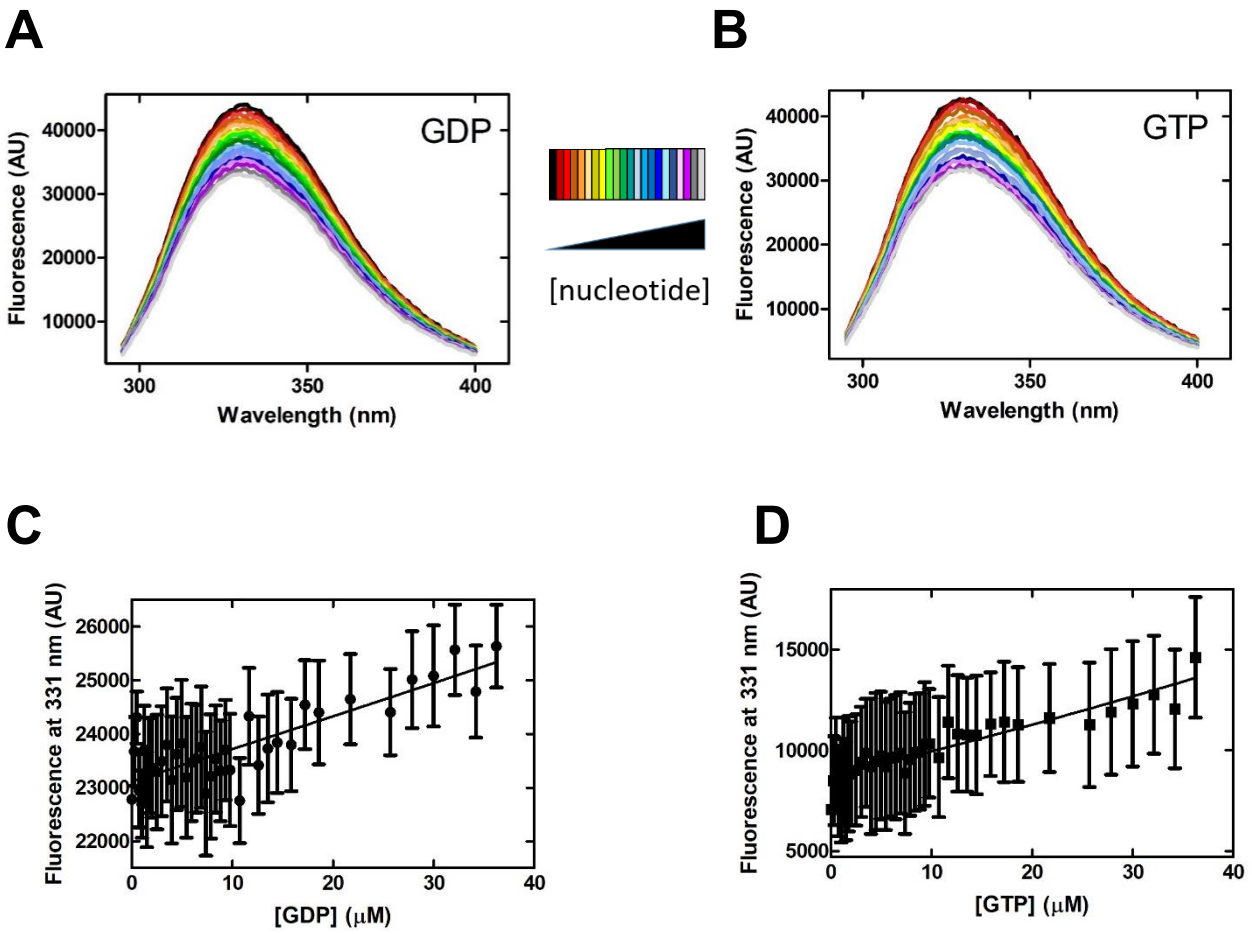


Figure A2.2. Fluorescence titrations measuring intrinsic tryptophan fluorescence fail to detect binding of YihA to GDP or to GTP. (A) Fluorescence emission spectra for the addition of increasing concentrations of GDP into 1 μM YihA followed by excitation of the intrinsic tryptophans of YihA at 280 nm. Spectra were recorded from 295 to 400 nm. (B) Fluorescence emission spectra for the addition of increasing concentrations of GTP into 1 μM YihA followed by excitation of the intrinsic tryptophan of YihA at 280 nm. Spectra were measured from 295 to 400 nm. (C) The effect of titrating increasing concentrations of GDP into YihA on the intrinsic tryptophan fluorescence at 331 nm. (D) The effect of titrating increasing concentrations of GTP into YihA on the intrinsic tryptophan fluorescence at 331 nm. Both titrations in (C) and (D) were performed in TAKM7 buffer.

The next method that was attempted in order to determine the equilibrium dissociation constant of YihA for nucleotides was to titrate mant-GDP into YihA (Figure A2.3). When YihA's intrinsic tryptophans are excited at 280 nm as performed in this experiment and mant-GDP is titrated into the reaction resulting in an increased concentration of mant-labelled nucleotide,

fluorescence resonance energy transfer (FRET) may occur which would result in energy transfer from the tryptophan donor to the mant-labelled nucleotide acceptor such that there is a decrease in the tryptophan donor fluorescence at 331 nm and a corresponding increase in the fluorescence of the mant acceptor group, which gave a fluorescence maximum at 441 nm in these experiments. Emission spectra were measured from 295 nm to 500 nm and both the tryptophan and mant group fluorescence maxima were plotted on separate graphs against the concentration of mant-GDP (Figure A2.3A-B). In contrast to previous work showing that binding of EcYihA to mant-GDP results in a concentration-dependent change in mant fluorescence that can be fit using a hyperbolic function [119], the obtained concentration dependence was linear (Figure A2.3D, F) and could not be fit with a hyperbolic function (Equation A2.2). For Figures A2.3 D,E,F, some of the data points have larger error bars while the rest of the data points have smaller error bars that are not visible in the figures. In Figure A2.3D, at the data point when mant-GDP was added to the cuvette to a final concentration of 75 μM there was more variability amongst the three trials that were averaged compared to at the other data points. The error was only greater than ± 2.9 x 10^4 for that one point at X = 75 μM and all other points had errors that were less than ± 2.9 x 10^4. The tryptophan fluorescence maxima showed a decrease in fluorescence with increasing concentration of mant-GDP that could be fit with a hyperbolic function (Figure A2.3C). However, when a control for background fluorescence was performed in which buffer was titrated into 0.5 μM YihA alone, the resulting fluorescence change for the tryptophan fluoresence (Figure A2.3E) showed a curve that could be fit with a hyperbolic function (Equation A2.2) with fluorescence values that were similar to those in Figure A2.3C except that the values for fluorescence at 331 nm were slightly higher in the control. Even in the absence of nucleotide the tryptophan fluorescence at 331 showed a hyperbolic decrease (Figure A2.3E) suggesting that the experimental design is not feasible for measuring FRET from YihA's tryptophans to the mant group. The similarity of the data in Figure A2.3C and Figure A2.3E suggests that if the control in

Figure A2.3E were subtracted from the data in Figure A2.3C, the result would be approximately a flat line, suggesting that there was no change in tryptophan fluorescence. As a control for background mant fluorescence at 441 nm, mant-GDP was titrated into buffer and the data could be fit with a linear regression (Figure A2.3F). Likewise, the values of the background mant fluorescence at 441 nm due to titration of mant-GDP into buffer (Figure A2.3F) were similar to the values of the mant fluorescence at 441 nm due to titration of mant-GDP into YihA (Figure A2.3D). Therefore, a value for the K_D of YihA for GDP could not be determined from a FRET-based equilibrium titration (Figure A2.3C-F).

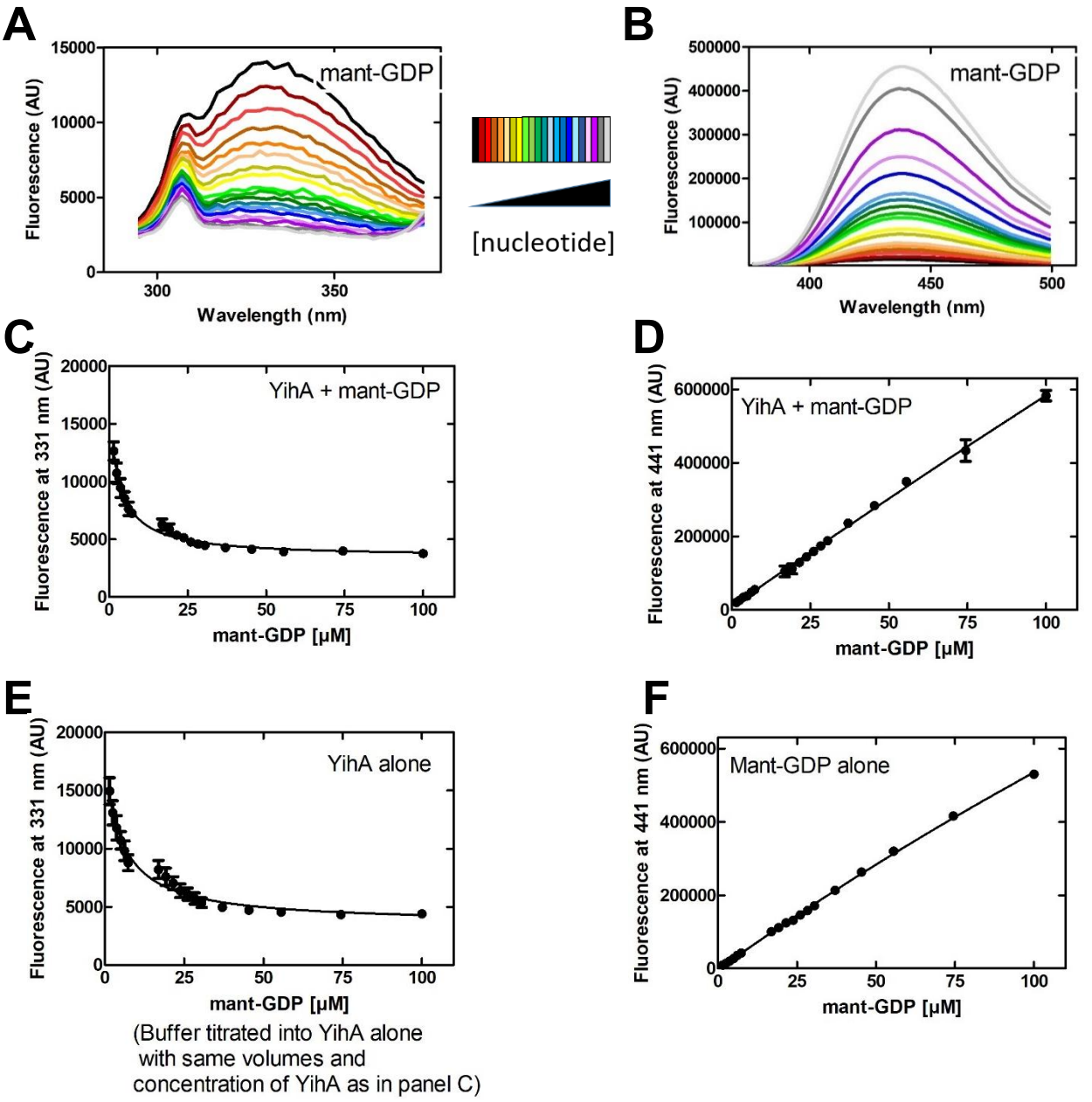


Figure A2.3. Fluorescence titrations measuring FRET from YihA's tryptophans to the mant group of mant-GDP fail to detect binding of YihA to mant-GDP. Excitation was at 280 nm for each sample. (A) Fluorescence emission spectra from 295 nm to 375 nm for the addition of increasing concentrations of mant-GDP into 0.5 μM YihA. (B) Fluorescence emission spectra from 375 nm to 500 nm for the addition of increasing concentrations of mant-GDP into 0.5 μM YihA. The spectra in (A) and (B) are plotted on two separate graphs for clearer visualization of the tryptophan fluorescence. (C) The effect of titrating increasing concentrations of mant-GDP into 0.5 μM YihA on the fluorescence of YihA's tryptophans at 331 nm. (D) The effect of titrating increasing concentrations of mant-GDP into 0.5 μM YihA on the fluorescence of the mant group at 441 nm. (E) Fluorescence at 331 for buffer titrated into YihA alone as a control. (F) Control showing the fluorescence at 441 nm for the titration of mant-GDP into buffer.

The issue with the data in FigureA2.3C,E could be due to photobleaching as it was repeatedly excited in the fluorimeter. In addition, heating of the sample by the fluorimeter lamp and incubation of the sample at 25°C during the entire course of the titration could potentially lead to unfolding of YihA if YihA is not stable at room temperature. The controls from the experiments with unlabelled GDP (Figure A2.2) and the FRET-based binding assays (Figure A2.3) both resembled a binding curve showing a decrease in fluorescence with increasing volume of buffer added to YihA, suggesting that photobleaching may be contributing to a decrease in the fluorescence signal in the absence of addition of nucleotide. To eliminate the problem of photobleaching, the fluorescence was recorded for separate samples for each data point. Using this method, the fluorescence was measured following excitation at 280 nm for samples containing only YihA in the absence of magnesium in the buffer (Figure A2.4A). The control in the absence of nucleotide provided the noise or variability in the measurement due to separate samples and as expected, resulted in no fluorescence change (Figure A2.4A). The fluorescence emission at 331 nm was recorded for samples containing EcYihA and increasing concentrations of unlabelled GDP in the absence of magnesium (Figure A2.4B,C). Again, the results did not show a concentration dependence consistent with the explanation that there is no change in the conformation of EcYihA or the fluorescent properties of EcYihA during binding of the unlabelled GDP (Figure A2.4B,C). Thus, this method of using separate samples could not be utilized to determine the equilibrium dissociation constant of YihA for unlabelled GDP.

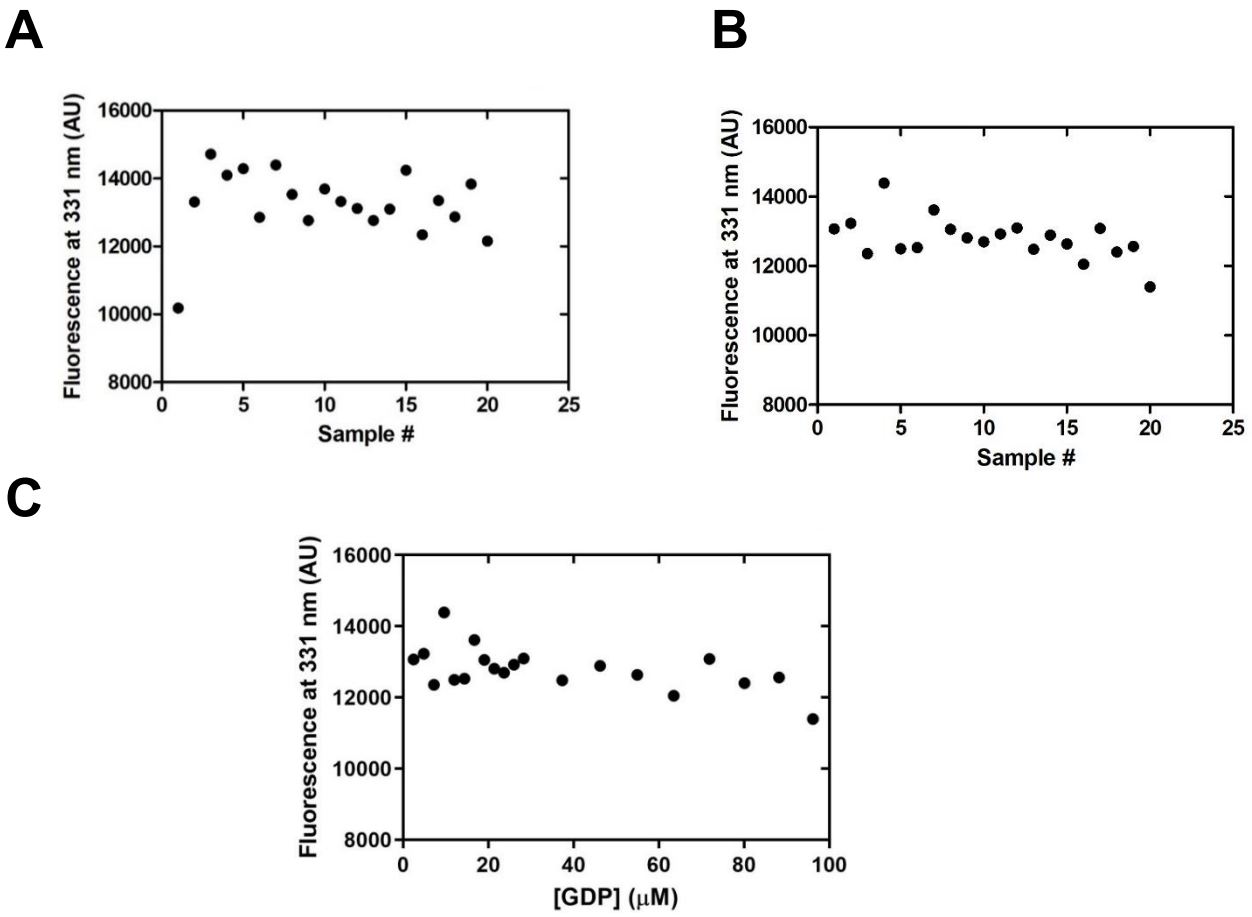


Figure A2.4. Binding of YihA to unlabelled GDP remains undetected with the method of measuring the fluorescence of separate samples. (A) Plot of the tryptophan fluorescence for identical samples containing 0.5 μM of YihA in buffer without magnesium. (C) Plot of the tryptophan fluorescence as a function of increasing concentrations of GDP. (B) Same data as in panel C but plotted with the same axes as in panel A. For (B) and (C) increasing concentrations of GDP were present in each sample containing 0.5 μM YihA followed by excitation of the intrinsic tryptophans of YihA at 280 nm. Each data point in panels A-C represents a separate sample. The buffer without magnesium used for reactions A-C was TAK (50 mM Tris-Cl pH 7.5 at 4°C, 70 mM NH_4Cl, 30 mM KCl).

A survey of the literature revealed a common method previously used for determining the equilibrium dissociation constant for binding of many GTPases to nucleotides, such as the GTPases MnmE and Ffh [178, 293]. MnmE also functions as an RNA modification enzyme within a complex that bacterial MnmE forms with the protein factor MnmG. Bacterial MnmE and MnmG

modify position 5 of the wobble uridine (U34) of some tRNAs to 5-aminomethyluridine (nm^5U) if the substrate available is the ammonium ion [294, 295]. Alternatively, if the substrate available is glycine, MnmE and MnmG interact with each other and modify the wobble uridine of some tRNAs to 5-carboxymethylaminomethyluridine ($cmnm^5U$) [31, 296]. The MnmEG complex is conserved across bacteria and eukaryotes and is able to modify tRNAs in eukaryotes [31]. The method used to establish the affinity of MnmE and many other GTPases for nucleotides was to titrate the GTPase into mant-labelled nucleotide and to measure the fluorescence emission of the mant group by exciting the mant group directly at 360 nm.

The method of titration of the GTPase into mant-labelled nucleotide was used here along with separate samples to avoid photobleaching and separate samples were prepared containing a constant concentration of mant-GDP and increasing concentrations of YihA. Samples were excited at 360 nm and emission spectra were recorded from 380 nm to 550 nm (Figure A2.5A). The mant group fluorescence (λ_{max} = 438 nm) was plotted against the concentration of YihA. The resulting concentration dependence of the mant fluorescence showed an increase in fluorescence as the concentration of YihA increased (Figure A2.5B). A control was performed by titrating YihA into buffer with excitation at 360 nm. This control resulted in no fluorescence change, as expected (Figure A2.5C). However, the data for the concentration dependence of the mant fluorescence resulted in a curve that was almost linear (Figure A2.5B). Consistently, the data for the concentration dependence of the mant fluorescence was unable to be fitted properly with a quadratic function (Equation A2.3) with the fit yielding R^2 = 0.5449 such that the fitting software was unable to solve for the fitting parameters and the K_D could not be established (Figure A2.5B). The conclusion reached was that this method was not able to determine the equilibrium dissociation constant for YihA for mant-GDP.

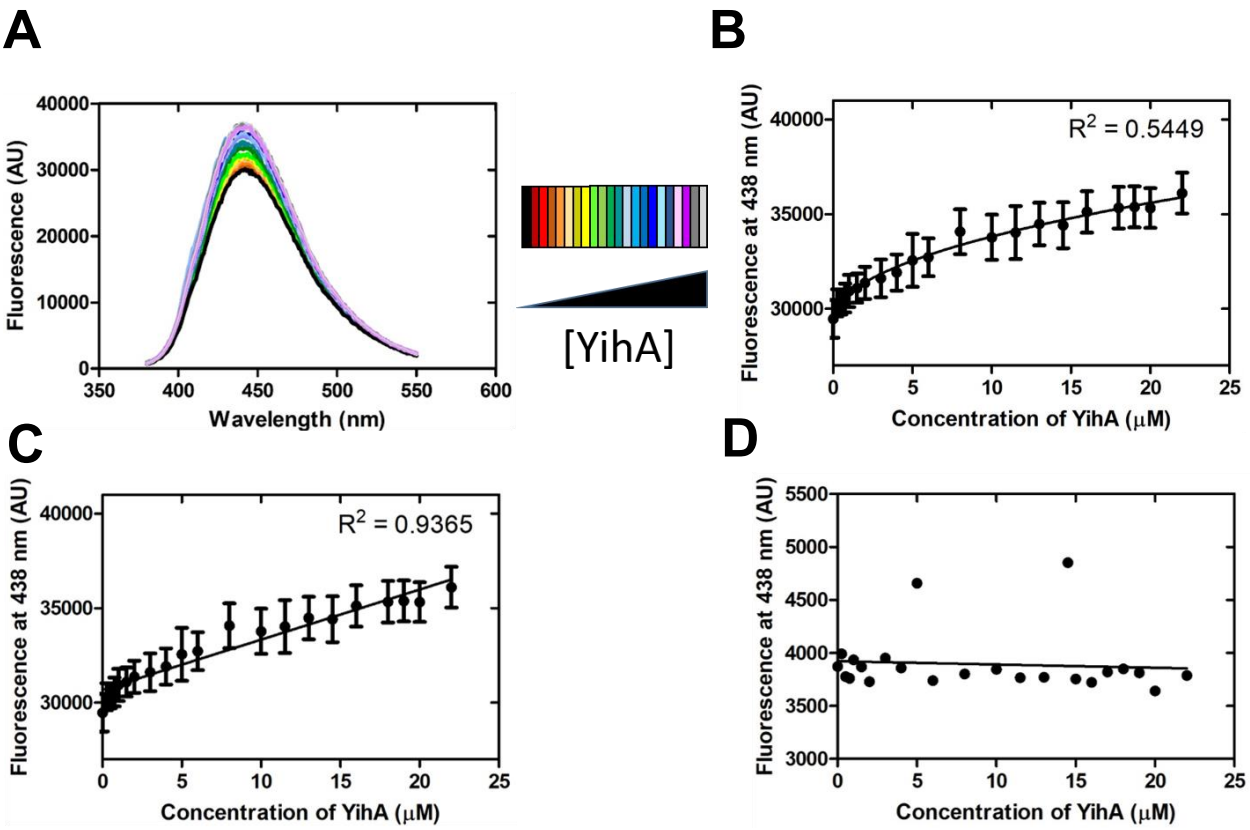


Figure A2.5. Fluorescence titrations measuring fluorescence emission following direct excitation of the mant group of mant-GDP. Each data point represents a separate sample. Samples were excited at 360 nm. (A) Fluorescence emission spectra from 380 nm to 550 nm for the addition of increasing concentrations of YihA into 0.66 μM mant-GDP. (B) The effect of titrating increasing concentrations of YihA into 0.66 μM mant-GDP on the fluorescence of the mant group at 438 nm. The fit of a quadratic equation (Equation A2.3) to the data is shown. (C) Same data as in (B) showing the goodness of fit when a linear regression was fitted to the data. (D) Control in which the fluorescence was measured for separate samples containing increasing concentrations of YihA in the absence of mant-GDP.

Although the data obtained in this chapter did not enable the equilibrium dissociation constant of EcYihA for guanine nucleotides to be determined, further work did confirm that EcYihA is able to bind to guanine nucleotides. An experiment was designed in which the fluorescence of separate samples of mant-GDP or mant-GTP alone, of YihA alone, and the fluorescence of mant-GDP with YihA or the fluorescence of mant-GTP with YihA was measured following excitation at

280 nm in TAKM7 buffer (50 mM Tris-Cl pH 7.5 at 4°C, 70 mM NH_4Cl, 30 mM KCl, 7mM $MgCl_2$) or in HEPES buffer (50 mM HEPES, 100 mM NaCl, pH 7.5 at 4°C) (Figure A2.6A-E). Separate samples were used in order to avoid the problems encountered with photobleaching and heating of YihA due to incubation of the sample at room temperature for long periods of time. The purpose of measuring the fluorescence of mant-labelled nucleotide alone, of YihA alone, and of YihA with mant-labelled nucleotide was to test the hypothesis that FRET from YihA's tryptophans to a mant-labelled nucleotide bound to YihA is able to occur because the conserved W95 is within the distance for efficient FRET. Contrary to this hypothesis, no change was seen in the tryptophan FRET donor fluorescence at 331 nm for YihA alone compared to the donor fluorescence at 331 nm of YihA and mant-labelled nucleotide with the fluorescence background at 331 nm due to mant-labelled nucleotide alone subtracted, for all conditions tested in Panels A-E (Figure A2.6). However, unexpectedly, the mant group acceptor fluorescence at 441 nm for YihA in the presence of mant-labelled nucleotide with the background fluorescence at 441 nm due to mant-labelled nucleotide alone subtracted increased compared to the fluorescence at 441 nm for mant-labelled nucleotide alone, for all conditions tested in Panels A-E (Figure A2.6). The increase in mant group acceptor fluorescence at 441 nm along with no change in fluorescence of the tryptophan donor led to the conclusion that although no FRET from YihA's tryptophans to the mant-labelled nucleotide bound to YihA is occuring, there is a stabilization of the mant fluorophore that occurs upon binding of the fluorophore to YihA that is causing the acceptor increase in fluorescence at 441 nm (Figure A2.6A-E).

Nucleotide binding to the nucleotide binding pocket of YihA may cause a slight conformational change that may result in movement of one or more of YihA's three tryptophans, altering the local environment around the tryptophan and resulting in a change in fluorescence at 331 nm. If there is no mant group present to accept FRET from YihA's tryptophans, the change in tryptophan fluorescence due to conformational changes triggered by binding of unlabelled

nucleotide can be measured. Therefore, similar experiments to the FRET experiment described in Figure A2.6 were done with unlabelled nucleotides to ensure that there was no increase in tryptophan fluorescence upon binding of nucleotide that would be masked by energy transfer to the mant acceptor and the corresponding decrease in tryptophan donor fluorescence, giving the appearance of no change in tryptophan fluorescence. Specifically, the titration of unlabelled GDP or GTP into YihA with excitation at 280 nm in buffer containing magnesium shown in Figure A2.2 provides a control showing that unlabelled GDP or GTP binding to YihA (in the absence of the mant fluorophore) has no effect on the fluorescence of YihA's tryptophans in a buffer containing magnesium. When the error bars in Figure A2.2C,D are taken into account, there is no overall change in the fluorescence at 331 nm upon addition of unlabelled nucleotide. Similarly, measurement of the fluorescence of separate samples containing YihA with increasing concentrations of unlabelled GDP upon excitation at 280 nm provides a control indicating that there is no change in tryptophan fluorescence upon binding of unlabelled GDP to YihA (in the absence of the mant fluorophore) and in buffer without magnesium. Therefore, consideration of the results in Figure A2.2 and Figure A2.3 together with the result from the FRET experiment in Figure A2.6 leads to the conclusion that the lack of change seen in the tryptophan fluorescence of YihA in the presence of mant-labelled upon excitation at 280 indicates that FRET is not occuring.

Whether the mant-labelled nucleotide is binding to YihA's nucleotide binding site or to some other site on YihA cannot be determined from this experiment (Figure A2.6A-E). However, due to the small size of YihA, there are very few regions on YihA where a mant-labelled nucleotide could bind without having FRET occur from at least one of YihA's tryptophans upon excitation of YihA's tryptophans at 280 nm (Figure A2.1) other than YihA's nucleotide binding site. YihA has three tryptophans that are spatially separated from one another and located on different faces of YihA so that in most regions of YihA's structure one of YihA's three tryptophans will be located

less than 25 Å away (Figure A2.1). The conclusions from this dataset are that there is no FRET upon excitation at 280 nm from YihA's tryptophans to mant-labelled nucleotide that binds to YihA, that a stabilization of the mant group of the mant-labelled nucleotide can be detected and that therefore YihA binds to guanine nucleotides, (Figure A2.6A-E).

There is an increase in the mant fluorescence at 441 nm upon binding of mant-labelled nucleotide to YihA (Figure A2.6A-E). Likely, the stabilization of the mant fluorophore of mant-labelled nucleotide upon binding to YihA is an effect that could be seen in a stopped flow experiment where YihA could be mixed with mant-labelled nucleotide with excitation at 280 nm and could potentially be used to calculate the rate constant of association of mant-labelled nucleotide to YihA and the rate constant of dissociation of mant-labelled nucleotide from YihA·mant-nucleotide, enabling a K_D of YihA for guanine nucleotides to be calculated. Therefore, this experiment confirms that YihA is able to bind to both mant-GDP and mant-GTP (Figure A2.6A-E) and that the binding site of guanine nucleotide on YihA is distant from any of the tryptophans in YihA (Figure A2.1).

The experiment to measure FRET between YihA's tryptophans and mant-labelled nucleotide was done in both TAKM7 buffer and in HEPES buffer (Figure A2.6A-E). HEPES buffer was used in the FRET experiment (Figure A2.6B) in order to be able to compare the result to a previous study in which EcYihA's tryptophans were excited at 295 nm and mant fluorescence at 450 nm of mant-labelled nucleotide in the presence and absence of YihA was measured in HEPES buffer using the same concentrations of EcYihA and mant-labelled as were used in Figure A2.6A-E [119]. The data from the previously available study [119] showed that there is an increase in mant fluorescence of the mant-labelled nucleotide at 450 nm upon excitation at 295 nm, consistent with the results given in Figure A2.6A-E. However, this previous study did not show the fluorescence at 331 nm of tryptophan upon excitation of YihA's tryptophans in the presence and absence of mant-labelled nucleotide [119]. As the tryptophan fluorescence was missing from

the previously published FRET data reporting an increase in mant group fluorescence of the mant-labelled nucleotide upon excitation of YihA's tryptophans, the data provided in Figure A2.6A-E now provides the first evidence that FRET from YihA's tryptophans to the mant group of mant-labelled nucleotide does not occur upon excitation of YihA's tryptophans but that there is a stabilization of the mant fluorophore of the mant-labelled nucleotide upon interaction with YihA. The effect observed of no change in tryptophan donor fluorescence at 331 nm along with an increase in mant group acceptor fluorescence at 441 nm was seen in all buffers tested and in both TAKM7 buffer which contains magnesium and in HEPES buffer which does not contain magnesium (Figure A2.6A-E).

Furthermore, the occurrence of FRET from YihA's tryptophans to mant-labelled nucleotide was investigated in YihA purified with magnesium (Figure A2.6A-C) and in YihA purified without magnesium (Figure A2.6D). Preliminary evidence presented in the main body of this thesis suggested that YihA may co-purify with nucleotide and that treatment of YihA with EDTA to remove magnesium by chelation led to release of nucleotide from YihA (Figure 2.1). Earlier work introduced the concept that removal of magnesium from EF-Tu is one of the main mechanisms whereby EF-Tu's GEF EF-Ts eliminates coordination of nucleotide in the active site and promotes nucleotide release from EF-Tu [297, 298]. Later work demonstrated an increase by a factor of 6 x 10^4 in the rate constant for GDP dissociation from EF-Tu·GDP in response to addition of EF-Ts compared to an increase by a factor of 150 to 300 in the rate constant for GDP dissociation from EF-Tu·GDP in response to treatment with EDTA to disrupt magnesium binding, nevertheless indicating that removal of magnesium is an effective method to faciliate release of bound nucleotide from EF-Tu [137].

Therefore, the hypothesis was formulated that purification of YihA without magnesium removes the nucleotide that co-purifies with YihA and that conversely, purification of YihA with magnesium stabilizes a bound nucleotide that co-purifies with YihA, and thus prevents binding of

YihA to nucleotide in an *in vitro* binding assay. However, Figure A2.6C-D shows that there is an increase in mant group fluorescence at 441 nm of mant-labelled nucleotide for both YihA purified with and without magnesium, suggesting that purification of YihA with or without magnesium does not affect binding of guanine nucleotide to YihA and indicating that both the purification of YihA with magnesium and the purification of YihA without magnesium are able to bind to guanine nucleotide. Both purifications of YihA with and without magnesium show a stabilization of the mant fluorophore and an increase in mant fluorescence upon excitation of YihA's tryptophans at 280 nm (Figure A2.6C-D). One possible rationale for the increase in the mant fluorescence observed for both purifications of YihA with and without magnesium could be that neither YihA purification co-purifies with nucleotide and thus binding of mant-labelled nucleotide is observed as expected if YihA has an affinity for guanine nucleotide in the micromolar range. A second possible rationale for the increase in the mant fluorescence observed for both purifications of YihA with and without magnesium could be that both purifications have a high proportion of YihA that co-purifies with nucleotide because YihA has an affinity for guanine nucleotides in the picomolar range and that the remaining portion of YihA molecules do not co-purify with nucleotide and are binding to mant-labelled nucleotide to give rise to the increase in mant fluorescence seen in this FRET experiment for both YihA purifications. Consistently, there is also an increase in mant group fluorescence at 441 upon excitation of YihA's tryptophans in both TAKM7 buffer (that contains magnesium) and in HEPES buffer (that does not contain magnesium) (Figure A2.6A-E).

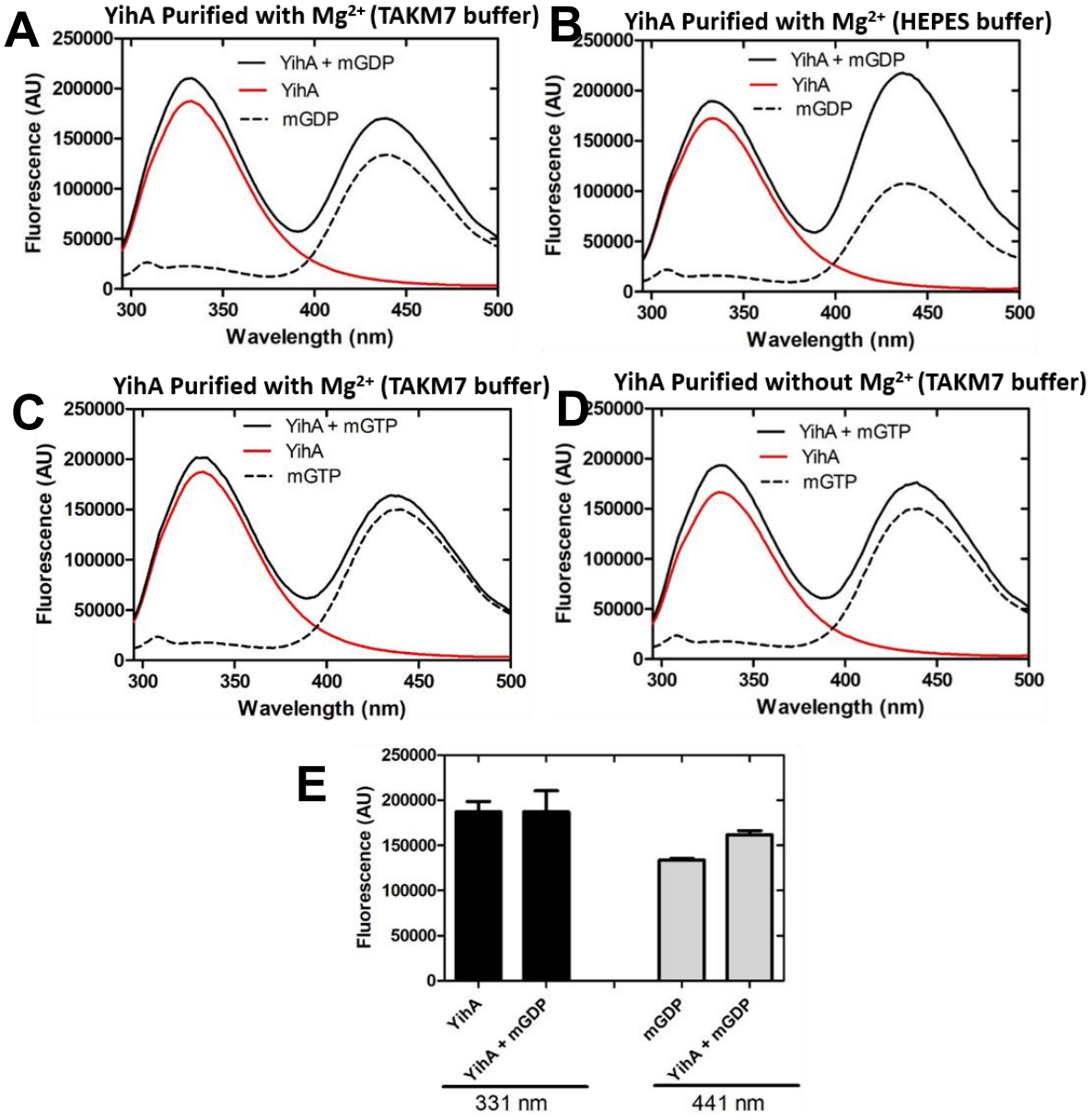


Figure A2.6. YihA binds to mant-GDP and mant-GTP. Fluorescence emission spectra obtained by excitation at 280 nm of 10 μM mant-GDP or mant-GTP alone (dashed line), 10 μM mant-GDP or mant-GTP with 0.85 μM YihA added (black line) or 0.85 μM YihA alone (red line). For each sample measured, the λ_{em} range was 295 nm to 500 nm. Each emission spectrum represents a separate sample. (A) Fluorescence was measured in TAKM7 buffer using YihA purified with Mg^{2+}. (B) Fluorescence was measured in HEPES buffer using YihA purified with Mg^{2+}. (C) Fluorescence was measured in TAKM7 buffer using YihA purified with Mg^{2+}. (D) Fluorescence was measured in TAKM7 buffer using YihA purified without Mg^{2+}. (E) The data is plotted for the experiment in Panel (A) as average ± standard deviation values. The fluorescence at 331 nm of YihA alone compared to fluorescence at 331 nm for YihA + mant-GDP with the fluorescence at 331 nm for mant-GDP alone subtracted are shown in the black bars. The grey bars show the fluorescence at 441 nm for mant-GDP alone compared to the fluorescence at 441 nm for YihA + mant-GDP with the fluorescence at 441 nm for YihA alone subtracted. The composition of each buffer is given in the text.

A2.3.2 INTRINSIC AND RIBOSOME-STIMULATED GTP HYDROLYSIS ACTIVITY OF ECYIHA

Previous work on the functional cycle of YihA led to conflicting evidence regarding the stimulation of YihA's GTP hydrolysis activity by GAPs [103, 119, 164]. One report found that the GTP hydrolysis activity of BsYsxC is stimulated by 50S ribosomal subunits [103]. However, another study reported stimulation of the GTP hydrolysis activity of EcYihA by the 50S ribosomal subunit but a much greater stimulation of the GTP hydrolysis activity of EcYihA by the 70S ribosome, suggesting that the 70S ribosome might be the GAP in EcYihA's functional cycle [164]. Additionally, there were differing conclusions in the literature about the intrinsic GTP hydrolysis activtiy of YihA. Earlier work demonstrated that EcYihA has no detectable intrinsic GTP hydrolysis activity [119]. In contrast, two later studies were able to measure intrinsic GTP hydrolysis activity of EcYihA and BsYsxC [103, 164].

In order to investigate whether the 50S ribosomal subunit or the 70S ribosome or both represent GAPs in YihA's functional cycle and to determine whether YihA has intrinsic GTP hydrolysis activity, the intrinsic and ribosome-stimulated GTP hydrolysis activity of EcYihA was measured using a mutiple turnover GTP hydrolysis assay (Figure A2.7A,B). The results showed that intrinsic GTP hydrolysis by YihA is not able to be detected, regardless of whether YihA is purified in the presence or absence of magnesium (Figure A2.7A and Table A2.1). Furthermore, a stimulatory effect of 70S ribosomes on the GTPase activity of YihA was not able to be observed because of the high intrinsic GTPase activity of 70S ribosomes (Figure A2.7B and Table A2.1). Thus, the observed rate of GTP hydrolysis of the 70S ribosome alone $(4.1 \pm 0.2) \times 10^{-3}$ μM s^{-1} was similar to the observed rate of GTP hydrolysis of YihA stimulated by the 70S ribosome of $(4.3 \pm 0.2) \times 10^{-3}$ μM s^{-1} (Table A2.1). In contrast, 50S ribosomal subunits were able to stimulate the GTP hydrolysis activity of YihA and there was approximately a ten-fold increase in the observed rate of GTP hydrolysis for YihA (purified in the presence of magnesium) stimulated by the 50S

ribosomal subunit [(1.2 ± 0.2) x 10^{-3} μM s^{-1}] when compared to the observed rate of GTP hydrolysis of the 50S ribosomal subunit alone [(3.0 ± 1.4) x 10^{-4} μM s^{-1}] (Figure A2.7A and Table A2.1). The approximately ten-fold increase in the observed rate of GTP hydrolysis for YihA stimulated by the 50S ribosomal subunit compared to the observed rate of GTP hydrolysis of the 50S ribosomal subunit alone provides preliminary evidence that the biologically relevant GAP is a 50S ribosome-like particle.

The concentration of YihA in the GTP hydrolysis assay was 5 μM (Figure A2.7). Therefore, if every YihA molecule in the GTP hydrolysis assay was active, there would be a total of 2 rounds of 50S ribosome-stimulated GTP hydrolysis in 2 hours as the concentration of P_i formed in the reaction at the end of two hours was approximately 10 μM (Figure A2.7A). The formal definition of a GAP is a protein that promotes efficient GTP hydrolysis, allowing the signal to be turned off and often supplies a catalytic element that is required for GTP hydrolysis [30, 37]. When comparing YihA's intrinsic and ribosome-stimulated observed rate of GTP hydrolysis and the turnover of GTP by YihA, the data in Figure A2.7 and Table A2.1 shows that YihA has no detectable intrinsic GTP hydrolysis whereas the observed rate of GTP hydrolysis for YihA (purified in the presence of magnesium) stimulated by the 50S ribosomal subunit was (1.2 ± 0.2) x 10^{-3} μM s^{-1}. Therefore, the 50S ribosomal subunit is by definition a GAP for YihA but the measured rate constant of 50S ribosome-stimulated GTP hydrolysis by YihA (purified with magnesium) of 0.01 min^{-1} (Table A2.2) is not biologically relevant compared to the rate constant of 50S ribosome-stimulated GTP hydrolysis by HflX of 11.40 min^{-1}, or to the rate constant for 70S ribosome-stimulated GTP hydrolysis by HflX of 14.40 min^{-1} or to the rate constant for 70S ribosome-stimulated GTP hydrolysis by YchF of 1.32 min^{-1} (Table A2.3). Indeed, the intrinsic rate constant of GTP hydrolysis by Era, a GTPase involved in ribosome biogenesis, is 0.01 – 0.02 min^{-1}, which is similar to the rate constant of 50S ribosome-stimulated GTP hydrolysis by YihA (Table A2.2, Table A2.3), suggesting that there are other factors required for efficient GTP hydrolysis by YihA

either in combination with the 50S ribosomal subunit or to stimulate GTP hydrolysis without the 50S ribosomal subunit. The 50S ribosomal subunit might not be the biologically relevant GAP for YihA but rather a precursor of the 50S ribosomal subunit might be the GAP that facilitates efficient GTP hydrolysis by YihA. The 50S ribosomal subunit is not an ideal GAP for YihA hence explaining why the measured rate of GTP hydrolysis in the presence of the 50S ribosomal subunit is not ideal compared to the observed rates of GTP hydrolysis by other GTPases in the presence of the 50S ribosomal subunit or 70S ribosome (Figure A2.7A, Table A2.1, Table A2.3).

As the observed rates of 50S ribosomal subunit-stimulated GTP hydrolysis by YihA were very slow compared to the previously published rates of many other GTPases (Table A2.2, Table A2.3), the hypothesis was put forward that YihA might co-purify with nucleotide, thus preventing YihA from binding to nucleotide in the GTP hydrolysis assay. To test this hypothesis, YihA was purified in HSG TAK buffer lacking magnesium (50 mM Tris-Cl pH 7.5 at 4°C, 70 mM NH_4Cl, 300 mM KCl, 15% glycerol) in order to attempt to remove co-purified nucleotide from YihA's nucleotide binding pocket, as the magnesium cation is required to coordinate the nucleotide in the nucleotide binding pocket. Subsequently, GTP hydrolysis assays with YihA purified in the presence or absence of magnesium were carried out with magnesium in the buffer (TAKM7 buffer). The results revealed that there is no difference in the 50S ribosomal subunit-stimulated observed rates of GTP hydrolysis for YihA purified in the presence of magnesium [$(1.2 \pm 0.2) \times 10^{-3}$ μM s^{-1}] or in the absence of magnesium [$(1.4 \pm 0.2) \times 10^{-3}$ μM s^{-1}], suggesting that either YihA does not co-purify with nucleotide or that purification without magnesium is not able to release nucleotide that co-purifies with YihA (Figure A2.7 and Table A2.1). Slot blotting confirmed that both YihA purified in the presence of magnesium and YihA purified in the absence of magnesium are able to bind to 50S ribosomal subunits (Figure A2.8).

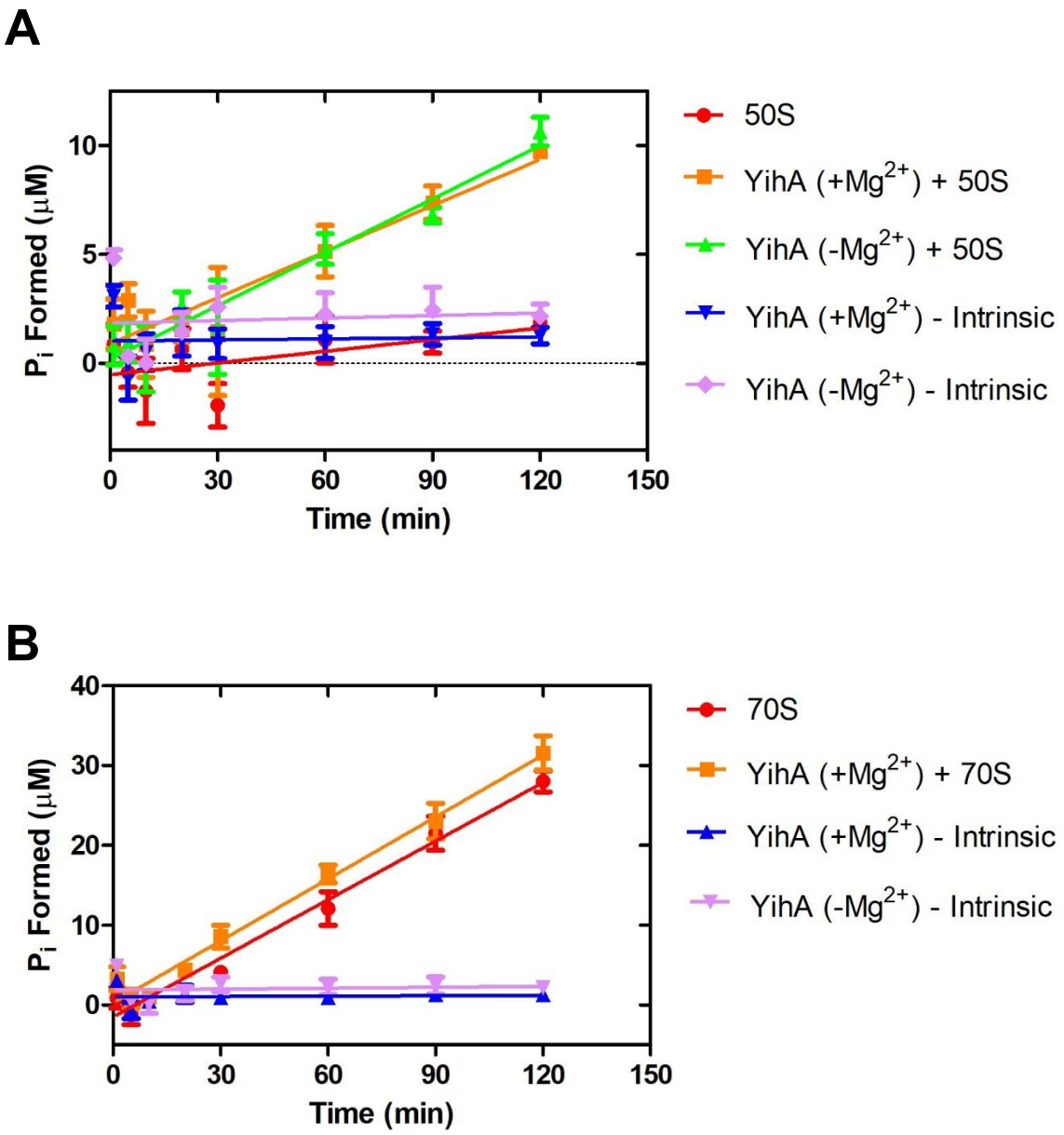


Figure A2.7. Multiple turnover GTPase activity of YihA. (A) Intrinsic and 50S ribosome-stimulated GTP hydrolysis by YihA in TAKM7 buffer. GTP hydrolysis by YihA stimulated by 1 μM 50S ribosomal subunits was measured for 5 μM YihA purified in the presence (orange line) or absence of magnesium (green line). GTP hydrolysis by 50S ribosomes alone was measured as a control (red line). Intrinsic GTP hydrolysis was measured for 5 μM YihA purified in the presence of magnesium (blue line) or purified in the absence of magnesium (lilac line). (B) Intrinsic and 70S ribosome-stimulated GTP hydrolysis by YihA in TAKM7 buffer. GTP hydrolysis stimulated by 1 μM 70S ribosomes was measured for 5 μM YihA purified in the presence of magnesium (orange line). GTP hydrolysis by 70S ribosomes alone was measured as a control (red line). Intrinsic GTP hydrolysis was measured for 5 μM YihA purified in the presence of magnesium (blue line) or purified in the absence of magnesium (lilac line).

Table A2.1 Intrinsic and ribosome-stimulated GTP hydrolysis

Sample	GTP Hydrolysis Activity ($\mu M\ s^{-1}$)
50S Alone	$(3.0 \pm 1.4) \times 10^{-4}$
YihA ($+Mg^{2+}$) + 50S	$(1.2 \pm 0.2) \times 10^{-3}$
YihA ($-Mg^{2+}$) + 50S	$(1.4 \pm 0.2) \times 10^{-3}$
70S Alone	$(4.1 \pm 0.2) \times 10^{-3}$
YihA ($+Mg^{2+}$) +70S	$(4.3 \pm 0.2) \times 10^{-3}$
YihA ($+Mg^{2+}$) – Intrinsic activity	Not detected
YihA ($-Mg^{2+}$) – Intrinsic activity	Not detected

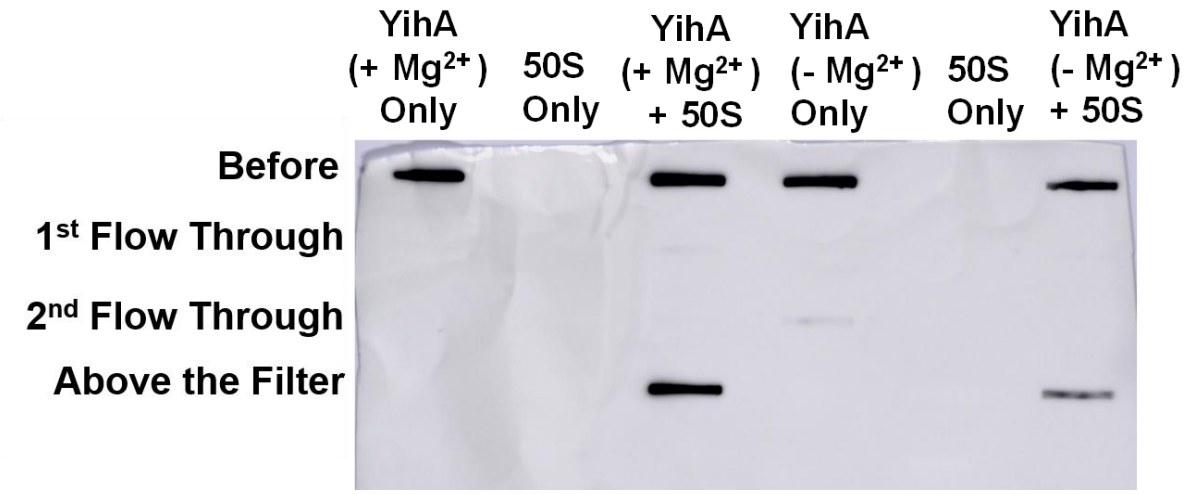


Figure A2.8. YihA purified in buffer with or without magnesium binds to the 50S ribosomal subunit. YihA in the *apo* state was incubated with 50S ribosomal subunits and microfiltration was performed. The samples from before microfiltration, the 1st flow through and the 2nd flow through and the sample remaining above the filter were quantified by slot blotting followed by western blotting with the anti-His tag antibody.

A2.4 DISCUSSION

The increase in mant fluorescence at 441 nm of the mant-labelled nucleotide without a corresponding decrease in YihA tryptophan fluorescence at 331 nm upon excitation at 280 nm (Figure A2.6A-E) supports the proposed model (Figure A2.1) where mant-labelled guanine nucleotide binds to the nucleotide-binding pocket of the G-domain that comprises YihA's structure. In this model, the tryptophan residues are at too great a distance from the mant group of the mant-labelled nucleotide.

The only data available on nucleotide specificity of YihA is for EcYihA [119]. Previous results indicated that a chase with unlabelled ADP of a reaction containing EcYihA and mant-GDP did not result in a decrease in fluorescence whereas a chase with GDP resulted in a significant decrease in fluorescence [119]. The previously reported equilibrium dissociation constants for EcYihA are 3.6 ± 0.3 μM for mant-GDP in the presence of 1 mM magnesium and and 27.4 ± 9.0 μM for mant-GTP in the absence of magnesium [119]. Based on the known concentrations of GDP and GTP in the *E. coli* cell of 200 μM for GDP and 1700 μM for GTP in mid-log phase [299], approximately 52.3% of YihA is predicted to be bound to GTP in the cell, whereas approximately 46.8% of YihA will be bound to GDP in the cell, and 0.9% of YihA will be in the *apo* state (see Section 2 of the appendix of this thesis for the calculations).

The intrinsic rate of GTP hydrolysis for YihA was unable to be measured in the current experimental setup (Figure A2.7, Table A2.1). Nucleotide-binding studies performed in this chapter show that YihA is able to bind to guanine nucleotides (Figure A2.6). Therefore, the lack of measurable intrinsic GTP hydrolysis by YihA is due to the inability of intrinsic YihA to hydrolyze bound GTP rather than due to the inability of intrinsic YihA to bind guanine nucleotides. Previous reports have indicated that as compared to BsYsxC, whose intrinsic GTP hydrolysis can be measured, EcYihA has no detectable intrinsic GTP hydrolysis activity [119] (Table A2.2).

However, the 50S ribosomal subunit and the 44.5S intermediate isolated from YsxC-depleted cells were both shown to stimulate GTP hydrolysis by BsYsxC [103] (Table A2.2). Results from one study suggested that YihA's GTPase activity is stimulated by 70S ribosomes but not by 50S or 30S ribosomal subunits [164] (Table A2.2). The observed rate constant for 50S-stimulated GTP hydrolysis by EcYihA measured by Kong *et al*, 2016, was 24-fold greater than the observed rate constant for 50S-stimulated GTP hydrolysis measured in the experiments reported here (Table A2.2) [164]. Moreover, intrinsic GTP hydrolysis activity of EcYihA could not be detected in the experiments reported here but could be measured with an observed rate constant of 0.12 min^{-1} by Kong *et al*, 2016 (Table A2.2) [164]. Additionally, intrinsic activity could also be measured previously for BsYihA, with a rate constant of 0.09 min^{-1} (Table A2.2) [126]. A possible reason for the discrepancy between the previously reported rate constants for YihA and the rate constants reported here could be the differences in salt concentration in the buffer. The activity assays reported here for EcYihA used 30 mM KCl and 70 mM NH_4Cl whereas Kong *et al*, 2016 used 300 mM of either KCl, NH_4Cl or NaCl in the activity assay buffer [164]. Lehoux *et al*, 2003 reported using a GTP hydrolysis activity assay buffer containing 150 mM NaCl and consistent with the results presented in this thesis, were unable to measure any intrinsic activity for EcYihA [119]. As is shown in Chapter 2 of this thesis, EcYihA is able to form homodimers. Potentially, at a lower concentration of salt in the buffer such as that used in the activity assays communicated in this work, YihA exists almost exclusively as a dimer and YihA dimers may be inactive. However, a higher concentration of salt in the buffer, such as 300 mM salt, could largely disrupt the YihA dimer complex. Interestingly, previous work showed that multimer formation *in vitro* can be dependent upon salt concentration [300]. Indeed, a multimer containing the *E. coli* single-stranded DNA binding protein (SSB) along with the DNA repair enzymes RecG and PriA was only able to be observed as an intact complex *in vitro* when the assay buffer contained a NaCl concentration at or below 100 mM [300]. Similarly, the K_D of dimerization for the *E. coli* ATPase SecA decreased

by approximately 5-fold to 20-fold when the assay buffer contained 200 mM K^+ acetate instead of 300 mM K^+ acetate, indicating that there was significantly increased dimerization of SecA with a salt concentration of 200 mM in the buffer compared to a buffer with 300 mM salt [301].

Alternatively, a potential rationale for the result obtained previously that EcYihA's intrinsic GTP hydrolysis activity can be measured with an observed rate constant of 0.12 min^{-1} could be that the previous GTP hydrolysis assays to measure EcYihA's intrinsic activity were done with less highly purified preparations of YihA than those used to obtain the results in this thesis [164]. If the EcYihA preparation used in previous work [164] contained contaminating cellular GTPases from the purification, a measurable rate constant for EcYihA's intrinsic GTP hydrolysis activity would be expected to be obtained.

Another possible explanation for the finding reported here that YihA possesses no intrinsic GTP hydrolysis activity could be that YihA's lack of ability to hydrolyze GTP intrinsically allows for efficient and tight regulation of YihA's functional cycle within the cell, as seen for other GTPases. For example, EF-Tu has been described as having an intrinsic GTPase activity that is barely detectable [302] and Ras also has a very low level of intrinsic GTP hydrolysis activity compared to prokaryotic GTPases (Table A2.3). Specifically, the rate constant for intrinsic GTP hydrolysis by EF-Tu is 4.60 x 10^{-3} min^{-1} and the rate constant for intrinsic GTP hydrolysis by Ras is 4.81 x 10^{-3} – 5.96 x 10^{-3} min^{-1} (Table A2.3). In contrast, HflX has a rate constant for intrinsic GTP hydrolysis of 0.05 min^{-1} and the intrinsic rate of GTP hydrolysis by YchF is 0.20 min^{-1} (Table A2.3).

Here it is shown that 50S ribosomal subunits are capable of stimulating the multiple turnover GTP hydrolysis activity of EcYihA (Figure A2.7, Table A2.1). However, as a very high intrinsic GTPase activity was measured for 70S ribosomes, it was not possible to determine whether 70S ribosomes are able to stimulate the GTP hydrolysis activity of YihA (Figure A2.7, Table A2.1). Consistent with earlier work, it was observed that 70S ribosomes and 50S ribosomal subunits have intrinsic GTP hydrolysis activity in the absence of YihA [303]. 70S ribosomes have

a high observed rate of intrinsic GTP hydrolysis which may be due to GTPases such as EF-Tu and EF-G that co-purify with the ribosomal particles. However, the observed rate of 50S ribosomal subunit-stimulated GTP hydrolysis by YihA is very slow compared to other ribosome binding GTPases (Table A2.2), suggesting that a factor that acts as a GAP *in vivo* could be missing from the reaction. The rate constant for the YihA's 50S ribosomal subunit-stimulated GTP hydrolysis (0.01 min^{-1}) is close in value to the rate constant for EF-Tu's 50S ribosomal subunit-stimulated GTP hydrolysis in the absence of EF-Ts (Table A2.3). Which protein factor or RNA could be missing from the *in vitro* GTP hydrolysis reaction remains unclear. YihA could be a ribosome biogenesis factor and the GAP for YihA might be expected to be an immature ribosomal particle that is a precursor of the 50S ribosomal subunit, such as the 44.5S intermediate found in YsxC-depleted cells [101]. YihA might bind to a precursor of the 50S ribosomal subunit alongside other factors that contribute to the GAP functionality of a precursor such as the 44.5S intermediate purified from YsxC-depleted cells. However, the previously published observed rates of GTP hydrolysis for 50S ribosomal subunit-stimulated GTP hydrolysis by BsYsxC confirm that a fully assembled 50S ribosomal subunit can be a GAP for BsYsxC [103]. Moreover, the *B. subtilis* 50S ribosomal subunits and the 44.5S intermediate used previously to stimulate YsxC's GTP hydrolysis activity may have been subjected to a method of purifying ribosomal particles involving fewer steps than the method of purifying ribosomal particles used here, such that additional factors might still be bound to the *B. subtilis* ribosomal particles that could act as GAPs for YsxC. Alternatively, YihA could co-purify with nucleotide, preventing the binding of YihA to nucleotides and thereby GTP hydrolysis.

The observed rate of 50S ribosomal subunit-activated GTP hydrolysis by YihA purified in the presence of magnesium was (1.2 ± 0.2) x 10^{-3} μM s^{-1} and for YihA purified in the absence of magnesium was (1.4 ± 0.2) x 10^{-3} μM s^{-1} (Table A2.1). In comparison, the 50S-ribosome stimulated GTP hydrolysis activity of the universally conserved GTPase and potential translation

factor HflX from *E. coli* is 0.19 ± 0.01 μM s^{-1} [48], which is two orders of magnitude greater than the observed rate of GTP hydrolysis by EcYihA (Table A2.3).

Table A2.2. Comparison of GTP hydrolysis by YihA with previously published rate constants. YihA's rate constants for GTP hydrolysis are compared to previously published studies where rate constants of GTP hydrolysis of YihA are available.

	Species	Intrinsic YihA (min^{-1})	YihA + 50S ribosomal subunits (min^{-1})	YihA + 70S ribosomal subunits (min^{-1})
This thesis	*E. coli*	Not detected	0.01[c]	0.05
Kong *et al.*, 2016, Prog. Biochem. Biophys. (In 300 mM KCl buffer) [164]	*E. coli*	0.12	0.24	1.03
Wicker-Planquart and Jault, 2015, FEBS Lett. [126]	*B. subtilis*	0.09	0.33 – 0.83	Not Reported
Lehoux *et al.*, 2003 Prot. Expression Purif. [119]	*E. coli*	Not detected	Not Reported	Not Reported

Table A2.3. Comparison of GTP hydrolysis by YihA with previously published rate constants for NTP hydrolysis by other GTPases and NTPases. YihA's rate constants of GTP hydrolysis are compared to rate constants of other GTPases from previously published work where rate constants of GTP hydrolysis are available. Unless indicated otherwise, rate constants are for GTPase activity. All data was obtained with incubation at 37°C and 50S and 70S ribosomal particles used were empty ribosomes.

	Intrinsic (min^{-1})	50S-stimulated (min^{-1})	70S-stimulated (min^{-1})
***E. coli* YihA (This thesis)**	Not detected	0.01	N/A
***E. coli* EF-Tu (without EF-Ts) [304]**	4.60×10^{-3}	6.00×10^{-3}	0.02
***E. coli* HflX [48]**	0.05	11.40	14.40
***E. coli* YchF (ATPase activity)**	0.20 [305]	0.41 [164]	1.32 [305]
***H. sapiens* Ras [306]**	4.81×10^{-3} – 5.96×10^{-3}	Not Reported	Not Reported
***E. coli* Era [307]**	0.01 – 0.02	Not Reported	Not Reported
***B. subtilis* CpgA [308]**	0.05	Not Reported	2.50
***E. coli* EF-G [309]**	Not Reported	Not reported	4.40×10^{3}
***E. coli* 70S ribosomes alone (This thesis)**	0.25		
***E. coli* 70S ribosomes alone, GTPase activity [303]**	0.60		

[c] This rate constant is for YihA (+Mg^{2+}) + 50S ribosomal subunits.

Intrinsic GTP hydrolysis activity of EcYihA was not detected in the GTP hydrolysis assay results presented here (Figure A2.7A, Table A2.1). Consistently with the results shown in this Appendix chapter, another study was not able to measure intrinsic GTP hydrolysis activity of EcYihA [119]. Interestingly, much earlier work demonstrated that a *B. subtilis* strain in which the wild type gene encoding BsYsxC on the chromosome was replaced with a gene encoding a variant BsYsxC with a truncation at the last 23 amino acids of the C-terminus could not survive [153]. As this truncation of the C-terminus of BsYsxC removed the highly basic region of the C-terminus of YsxC that interacts with the 50S ribosomal subunit [122], this experiment indicates that the interaction of BsYsxC with ribosomal particles is required for survival of *B. subtilis* [153].

The data presented here and previous findings can be summarized by a minimal model describing EcYihA's functional cycle (Figure A2.8). Earlier work demonstrated that the EcYihA-GTP complex has a higher affinity for the 50S ribosomal subunit compared to EcYihA bound to GDP [164]. Therefore, YihA could preferentially bind to the 44.5S intermediate isolated from YsxC-depleted cells or another precursor of the 50S ribosomal subunit in the GTP bound form (Figure A2.8). Previous studies suggest that YihA plays a role in the maturation of a precursor of the 50S ribosomal subunit, illustrated as a series of unknown steps in the model of the functional cycle (Figure A2.8) [59, 101, 103]. The work shown in this chapter suggests that the 50S ribosomal subunit stimulates GTP hydrolysis by YihA, potentially indicating that YihA could remain bound to the precursor of the 50S ribosomal subunit until it reaches the mature state as a 50S ribosomal subunit (Figure A2.7, Table A2.1, Figure A2.8). As demonstrated by previous work, the affinity of YihA for the 50S ribosomal subunit is reduced when YihA is bound to GDP [164], leading to dissociation of the 50S ribosomal subunit from the YihA-GDP complex (Figure A2.8). The next step in YihA's functional cycle would be release of GDP from YihA•GDP, followed by association of YihA with GTP facilitated by the greater concentration of GTP within the cell as compared to

the concentration of GDP within the cell (Figure A2.8). Thus, YihA•GTP is regenerated and available to repeat the functional cycle.

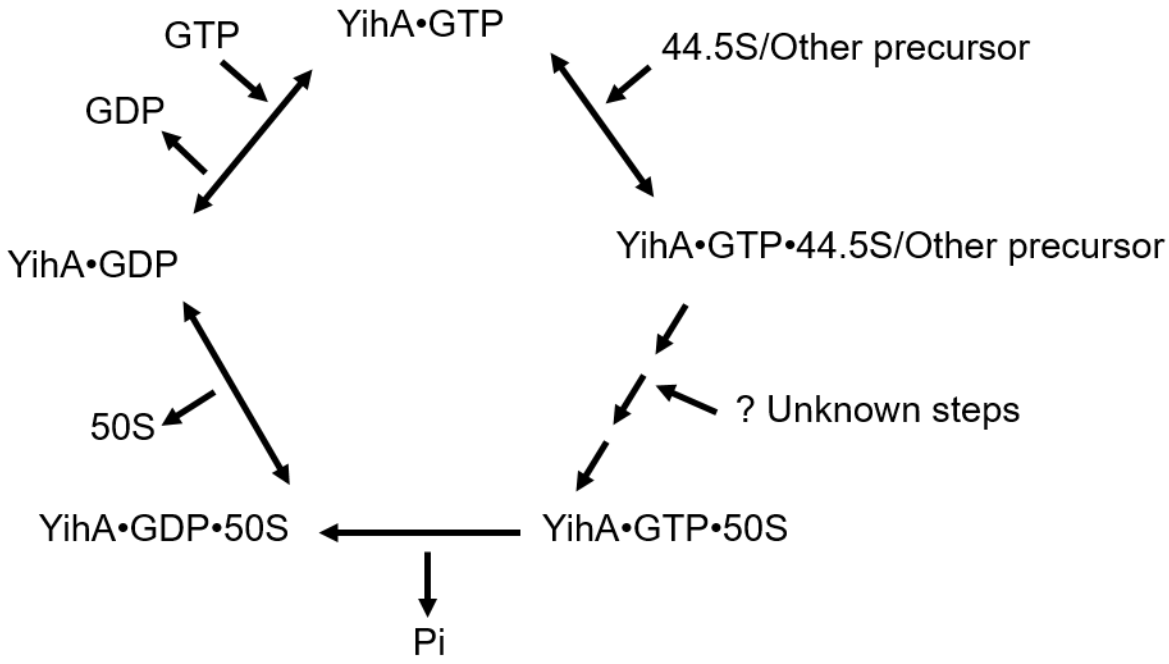


Figure A2.9 Minimal model for the functional cycle of YihA. The model is based on previously reported data and data presented here.

A2.5 CONCLUSIONS

The FRET-based assays that were performed did not yield the expected results as there was no FRET from YihA's tryptophans to the mant group of mant-labelled nucleotide upon excitation of YihA's tryptophans at 280 nm. Instead, an increase in mant fluorescence has confirmed that EcYihA binds to guanine nucleotides. The lack of FRET suggests that either W95 of EcYihA is not located within the distance for FRET from the mant-labelled nucleotide or that

the dipoles of the donor and accceptor are oriented perpendicular to one another. However, an increase in the fluorescence of mant-labelled nucleotide in the presence of YihA upon excitation at 280 nm is a result of stabilization of the mant fluorophore upon mant-nucleotide binding to a site on YihA. The conserved G-motifs of YihA and the lack of FRET signal supported by available structural information (Figure A2.1) suggest that the nucleotide is likely binding to YihA's nucleotide binding site. The increase in the mant fluorescence of mant-labelled nucleotides in the presence of YihA upon excitation at 280 nm will be used in future stopped flow experiments to determine the K_D of YihA for guanine nucleotides. Furthermore, the lack of FRET corresponding to no change in YihA's tryptophan fluorescence and an increase in mant fluorescence upon excitation of YihA's tryptophans was observed for both YihA purified with and without magnesium and was seen in FRET reactions carried out in both buffer with magnesium and in buffer without magnesium.

Moreover, purification of YihA with or without magnesium has no effect on the observed rate of GTP hydrolysis by YihA stimulated by the 50S ribosomal subunit. In addition, it has been shown here that EcYihA's multiple turnover GTP hydrolysis activity is stimulated by 50S ribosomal particles. However, the observed rate of GTP hydrolysis activity of YihA stimulated by 50S ribosomal particles is very slow compared to previously reported observed rates of GTP hydrolysis for other GTPases, potentially indicating that a GAP other than the 50S ribosomal subunit is required for YihA's functional cycle. Finally, this work suggests that EcYihA does not possess intrinsic multiple turnover GTP hydrolysis activity that is detectable *in vitro*. The finding that YihA does not possess the ability to hydrolyze GTP intrinsically along with the result that the 50S ribsoomal subunit alone does not increase the rate of YihA's GTP hydrolysis to a physiologically relevant rate shows the tight regulation of YihA's functional role in the cell.

SECTION A2 – CHAPTER A2 APPENDIX

Appendix A2.1 Calculation of the percentage of YihA bound to guanine nucleotides in the cell

Assuming ligand is in far excess compared to enzyme in the cell, the total GTP and GDP concentrations are the same as the free GTP and free GDP concentrations.

Given that YihA does not bind to adenine nucleotides as has been demonstrated previously [119], $[YihA]_T$ represents the total concentration of YihA in the cell during mid-log phase, which is 1.5 μM, taking the average of the previously reported values [213, 214]. The *E. coli* cellular concentration of GTP is 1700 μM and the cellular concentration of GDP is 200 μM during mid-log phase [299]. The K_D for YihA binding to GDP is 3.6 μM and the K_D for the interaction of YihA with GTP is 27.4 μM [119].

Binding of guanine nucleotides to YihA is analogous to a system of competitive inhibition since GDP and GTP are competing for the same binding site.

Therefore, using the derivation given in Voet and Voet, 2004 [310] for competitive inhibition gives

$$[YihA]_T = [YihA - GTP] + [YihA - GDP] + [apo - YihA] \quad \text{(Equation 2.1-A)}$$

Assuming that GTP and GDP are in rapid equilibrium with YihA and that guanine nucleotide binding to YihA is reversible,

$$YihA - GTP \rightleftharpoons apo - YihA + GTP$$

$$K_{D,GTP} = \frac{[apo - YihA][GTP]}{[YihA - GTP]} \quad \text{(Equation 2.2-A)}$$

Rearranging Equation 2.2-A gives

$$[apo - YihA] = \frac{K_{D,GTP}[YihA - GTP]}{[GTP]} \quad \text{(Equation 2.3-A)}$$

$$YihA - GDP \rightleftharpoons apo - YihA + GDP$$

$$K_{D,GDP} = \frac{[Apo - YihA][GDP]}{[YihA - GDP]} \quad \text{(Equation 2.4-A)}$$

Rearranging Equation A2.4-A yields

$$[YihA - GDP] = \frac{[apo - YihA][GDP]}{K_{D,GDP}} \quad \text{(Equation 2.5-A)}$$

Substituting Equation 2.3-A into Equation 2.5-A gives

$$[YihA - GDP] = \frac{K_{D,GDP}[YihA - GTP][GDP]}{[GTP]K_{D,GDP}} \quad \text{(Equation 2.6-A)}$$

Substituting Equation 2.6-A and Equation 2.3-A into Equation 2.4-A results in

$$[YihA]_T = K_{D,GTP}\frac{[YihA - GTP]}{[GTP]} + \frac{K_{D,GTP}[YihA - GTP][GDP]}{[GTP]K_{D,GDP}} + [YihA - GTP]$$

which can be simplified to

$$[YihA]_T = [YihA - GTP]\left\{\frac{K_{D,GTP}}{[GTP]}\left(1 + \frac{[GDP]}{K_{D,GDP}}\right) + 1\right\}$$

and solved for [YihA-GTP] by rearranging it to

$$[YihA - GTP] = \frac{[YihA]_T}{\frac{K_{D,GTP}}{[GTP]}\left(1 + \frac{[GDP]}{K_{D,GDP}}\right) + 1} = \frac{1.5\ \mu M}{\frac{27.4\ \mu M}{1700\ \mu M}\left(1 + \frac{200\ \mu M}{3.6\ \mu M}\right) + 1} = 0.8\ \mu M$$

The value for [YihA-GTP] can be inserted into Equation 2.3-A to solve for [apo-YihA]:

$$[apo - YihA] = \frac{K_{D,GTP}\,[YihA - GTP]}{[GTP]} = \frac{(27.4\ \mu M)(0.8\ \mu M)}{1700\ \mu M} = 1.3\ \times 10^{-2}\ \mu M$$

Equation 2.1-A can be rearranged to solve for [YihA-GDP]:

$$[YihA - GDP] = [YihA]_T - ([YihA - GTP] + [apo - YihA])$$
$$= 1.5\ \mu M - (0.8\ \mu M + \{1.3\ \times 10^{-2}\ \mu M\}) = 0.7\ \mu M$$

% YihA bound to GTP is given by

$$\%\ YihA - GTP = \frac{[YihA - GTP]}{[YihA]_T} \times 100 = \frac{0.8\ \mu M}{1.5\ \mu M} \times 100 = 52.3\%$$

Likewise, % YihA bound to GDP can be calculated from

$$\%\ YihA - GDP = \frac{[YihA - GDP]}{[YihA]_T} \times 100 = \frac{0.7\ \mu M}{1.5\ \mu M} \times 100 = 46.8\%$$

Similarly, % YihA in the *apo* form is determined by

$$\%\ apo - YihA = \frac{[apo - YihA]}{[YihA]_T} \times 100 = \frac{1.3\ \times 10^{-2}\ \mu M}{1.5\ \mu M} \times 100 = 0.9\%$$