CHARACTERIZATION OF THE UNIVERSALLY CONSERVED ATPASE YCHF USING AN *IN VITRO* AND *IN SILICO* APPROACH

Kirsten Shadoe Rosler
B. Sc. University of Lethbridge 2010

A Thesis
Submitted to the School of Graduate Studies
Of the University of Lethbridge
In Partial Fulfillment of the
Requirements for the Degree

MASTER OF SCIENCE

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

CHARACTERIZATION OF THE UNIVERSALLY CONSERVED ATPASE YCHF USING AN *IN VITRO* AND *IN SILICO* APPROACH

KIRSTEN S. ROSLER

Approved:

(Print Name) (Signature) (Rank) (Highest Degree)

Date

Hans-Joachim Wieden

Marc Roussel

Steven Mosimann

Michael Gerken

Abstract

GTPases perform a myriad of functions within the cell, from protein synthesis to cellular signaling. Of all known GTPases, only eight are conserved across all three domains of life. YchF is one of the eight universally conserved GTPases, however its cellular function is poorly understood. YchF differs from the classical GTPases in that it has a higher affinity for ATP than for GTP and functions as an ATPase. As a HAS-GTPase, YchF does not possess the canonical catalytic glutamine required for nucleotide hydrolysis, urging the question of how does YchF hydrolyze ATP. Here we have used molecular dynamics simulations (*in silico*) and biochemical experiments (*in vitro*) to identify an amino acid, histidine 114, essential for ATP hydrolysis in YchF. His 114 is located in a flexible loop of the G-domain of YchF, which shows nucleotide-dependent conformations *in silico*. The findings reported indicate that the 70S ribosome can stimulate the ATPase activity of YchF by directly participating in catalysis as well as helping to position the catalytic histidine residue. Additionally, nucleotide binding and dissociation rate constants have been determined in the presence and absence of Mg^{2+} in order to further understand the functional cycle of YchF.

Acknowledgements

I would like to thank my supervisor for allowing me to pursue this degree and always motivating me to work harder. I would also like to thank my committee members for all of their feedback during my few (whoops), but very helpful, committee meetings.

I want to thank everyone I have met in the Wieden and Kothe Labs over the past couple of years! Despite my tendency to avoid lab events, they really were some of the best times I have ever had!

I would like to acknowledge the students who have assisted in collecting some of the data presented here: Ian Andrews who helped to purify the His 114 and His 102 variants of YchF and Maja Gehre who assisted in performing *in vitro* experiments presented in Chapter 3.

The National Science and Engineering Research Council of Canada, Alberta Innovates Technology Futures and the University of Lethbridge funded the work presented here. Computational resources were provided by WestGrid (Compute Canada).

Table of Contents

List of Tables

List of Figures

List of Abbreviations

β-mEtOH	β-mercaptoethanol
30S	Small ribosomal subunit (Bacteria) (S=Svedberg)
50S	Large ribosomal subunit (Bacteria) (S=Svedberg)
60S	Large ribosomal subunit (Eukaryotes) (S=Svedberg)
70S	Associated ribosome couples (Bacteria) (30S + 50S)
aa-tRNA	Aminoacyl –tRNA
ADPNP/AMPPNP	Adenosine-5' –(β,γ-imido)-triphosphate
A-site	Aminoacyl-tRNA binding site
cpGTPase	Circularly permuted GTPase
CPG-domain	Circularly permuted GTPase domain
dpm	Decays per minute
EF	Elongation factor
E-site	tRNA exit site
fMet	Formyl Methionine
FRET	Fluorescence (Förster) Resonance Energy Transfer
GAP/GAF	GTPase activating protein (factor)
GDI	Guanine nucleotide dissociation inhibitor
GEF	Guanine nucleotide exchange factor
Gln^{cat}	Glutamine (catalytic)
HAS	Hydrophobic amino acid substituted
IF	Initiation factor
IPTG	Isopropyl-β-D-galactopyranoside
Kan	Kanamycin
LB	Luria-Bertani
mant	2'(or 3')-O-(N-methylantraniloyl)
MD	Molecular Dynamics
MES	2-(N-morpholino)ethanesulfonic acid
MWCO	Molecular weight cutoff
NAMD	Nanoscale molecular dynamics
OD_{600}	Optical Density at 600 nm
PDB ID	Protein data bank identification code
PEP	Phosphoenolpyruvate
PK	Pyruvate Kinase
P-loop	Phosphate binding loop
PMSF	Phenylmethylsulfonyfluoride
P-site	Peptidyl-tRNA binding site
PTC	Peptidyl transferase center
Rb	Ribosome
RbgA	Ribosome biogenesis GTPase A
RF	Release factor
RRF	Ribosome recycling factor
RT	Room Temperature
serRS	seryl-tRNA synthetase

SEC	Size Exclusion Chromatography
SMD	Steered molecular dynamics
SRL	Sarcin-ricin loop
TGS domain	Threonyl-tRNA synthetase, GTPase, SpoT domain
TRAFAC	Translation factor class
VMD	Visual molecular dynamics

Chapter 1 Introduction

1.1 Guanine Nucleotide Binding Proteins

G-proteins belong to a family of regulatory GTPases responsible for regulating almost every aspect of cellular life. These GTPases are involved in cellular processes such as cytoskeletal rearrangement (Rho, Rac); translation (elongation factors and the signal recognition particle); nuclear import (Ran); cellular trafficking (Rab, ARF) and receptor-mediated extracellular communication (heterotrimeric G proteins).[1] Members of this family possess a conserved G–domain whose sequences, although different, share a common structure. The domain is characterized by the presence of five distinct motifs, the G1 through G5 motifs. The G1 or Walker A motif (GxxxxGK (S/T)), better known as the P-loop, is involved in nucleotide binding through interactions with the phosphate moiety. As the P-loop is one of the most abundant folds in cellular organisms, the GTPases are members of the much larger class of P-loop NTPases.[2] The G2 motif (Switch I or effector region) is involved in magnesium binding and shows large structural rearrangements depending on whether a tri- or di-phosphate nucleotide is bound. The G2 motif is characteristic for each subfamily of GTPases as this region is frequently involved in effector interactions.[3] The G3 motif (Walker B motif or switch II; DxxG) also participates in magnesium and γ-phosphate coordination, and similarly to Switch I shows structural rearrangements dependent on the nucleotide bound. Both the G4 ((N/T)(K/Q) xD) and the G5 motifs are involved in guanine nucleotide specificity. The G1, G3 and G4 motifs are easy to recognize across the GTPase superfamilies, however the G2 and G5 motifs have varying sequence motifs that are more

conserved within a respective sub-family.[1,4] GTPases function as molecular switches, transitioning from an active GTP-bound state which allows the protein to interact with downstream effector molecules followed by GTP hydrolysis to an inactive GDP-bound state, which can no longer bind to the effector. GDP release results in a transient *apo* state of the GTPase before the protein can again rebind GTP to continue this cycle. In order to 'fine-tune' the cycle, regulatory proteins have evolved such as GTPase activating factors (GAFs, better known as GAPs when a protein factor is involved), which are capable of accelerating hydrolysis, Guanine Nucleotide Exchange Factors (GEFs), which facilitate the exchange of GTP to GDP and Guanine Nucleotide Dissociation Inhibitors (GDIs), which prevent the release of GDP.[5,6] The GTPase superclass can be divided into two distinct families: the TRAFAC class (translation factor related) which includes classical GTPases such as EF-Tu and Ras, and the SIMBI class (signal recognition GTPases, MinD and BioD superfamily), which includes SRP GTPases and a few classes of ATPases.[7]

The number of GTPases found within bacteria is considerably less than that found in eukaryotic cells, in particular those involved in cellular signaling and some of the small-GTPase families (such as heterotrimeric G proteins and the Ras family) appear to be absent. The actual number of GTPases in bacterial cells can vary anywhere from 30 to a minimal number of 11, conserved amongst bacteria.[3] Through the use of comparative genomics it has become evident that 8 universally conserved GTPases exist: EF-Tu, EF-G, IF-2, Ffh, FtsY, YihA, HflX and YchF (*E. coli* nomenclature).[8] These conserved GTPases can be placed into four main ancestral groups: The Era group (YihA), the FtsY/Ffh group, the elongation factor group (EF-

Tu, EF-G & IF-2) and the Obg group (HflX and YchF).[8] Only the translation factors and the protein secretion factors have known functions, as YihA, HflX, and YchF have not been characterized in detail (see table 1.1).

Table 1.1 Eight universally conserved GTPases and their functions

	Function	Ref.
EF-G	Elongation of protein synthesis. Catalyzes translocation of tRNA from the A to the P site.	9,10
EF-Tu	Elongation of protein synthesis. Shuttles amino-acylated tRNA to the ribosomal A site.	10,11
IF2	Initiation of protein synthesis	10,12
FfH / FtsY	Co-translational protein targeting	10,13
YihA	Assembly / processing of 50S subunit. Possibly participates in checkpoint mechanisms for cell division.	14,15
HflX	Unknown function. Binds to 70S ribosomes.	10,16,17
YchF	Unknown function. Associates with 70S ribosomes. Linked to oxidative stress.	10,18-20

Ras is often considered a prototype for the minimal GTPase domain.[21] In Ras and related proteins, a catalytic glutamine residue (Gln^{cat}) located immediately downstream of the DxxG motif in Switch II, is required for catalysis. The role of Gln^{cat} is to stabilize the transition state of the reaction by orienting a nucleophilic water molecule for an in-line attack on the γ-phosphate of the nucleotide.[21,22] The Gln^{cat} (Gln 61 in Ras) is substituted with histidine in the translation factor superfamily (EF-Tu, IF2, EF-G), threonine in Rap and arginine in Ffh.[5]

The length of time a GTPase remains in its 'on-state' is dependent upon the rate at which the G-protein can catalyze hydrolysis. Many GTPases have extremely slow intrinsic catalytic rates, on the order of minutes to hours, which must be accelerated through the use of a GAF; otherwise the protein will remain 'on' for an

extended period of time.[4,23-26] An over-active enzyme can present a major problem within the cell; for example, constitutively active Ras mutants are prevalent in many types of cancers.[27] All GTPases perform hydrolysis using a similar mechanism in which an 'activated' water molecule performs a nucleophilic attack on the γ-phosphate of the GTP molecule. Although all known GAFs essentially perform the same function, accelerating the rate of GTP hydrolysis, the mechanisms used by different GAFs can vary.[23] A GAF may speed up hydrolysis by i) stabilizing the intrinsic catalytic residue, ii) directly supplying its own catalytic residue in trans to activate the nucleophilic water, or iii) by stabilizing the developing negative charge of the transition state (Figure 1.1.1).[23] Ras and its homologs employ a mechanism in which an 'arginine finger' stabilizes the endogenous glutamine residue and neutralizes the developing negative charge.[22,23] RanGAPs and regulators of G protein signaling however simply stabilize the intrinsic catalytic machinery without the use of an Arg residue to neutralize the charge effects.[26,28,29] Overall, the purpose of the GAF is to stabilize the intrinsic catalytic machinery, and in many cases to also provide catalytic machinery in trans.[23] GTP hydrolysis can also be stimulated through dimerization, as is the case for dynamin, hGBP1 and MnmE.[30,31] Recently, it has been shown that a potassium ion is also capable of stimulating hydrolysis in a number of GTPases, in particular the HAS-GTPases as described in section 1.2.[31-35] The potassium ion is stabilized by a conserved Asn residue in the P-loop as well as an insertion in the Switch I region (K-loop) preceding the GTTRD motif. Superimposition of the transition state structures of MnmE and Ras•RasGAP have

shown that the potassium ion acts by stabilizing the negative charge of the transition state, similarly to mechanism employed by the ‘arginine’ finger’.[31]

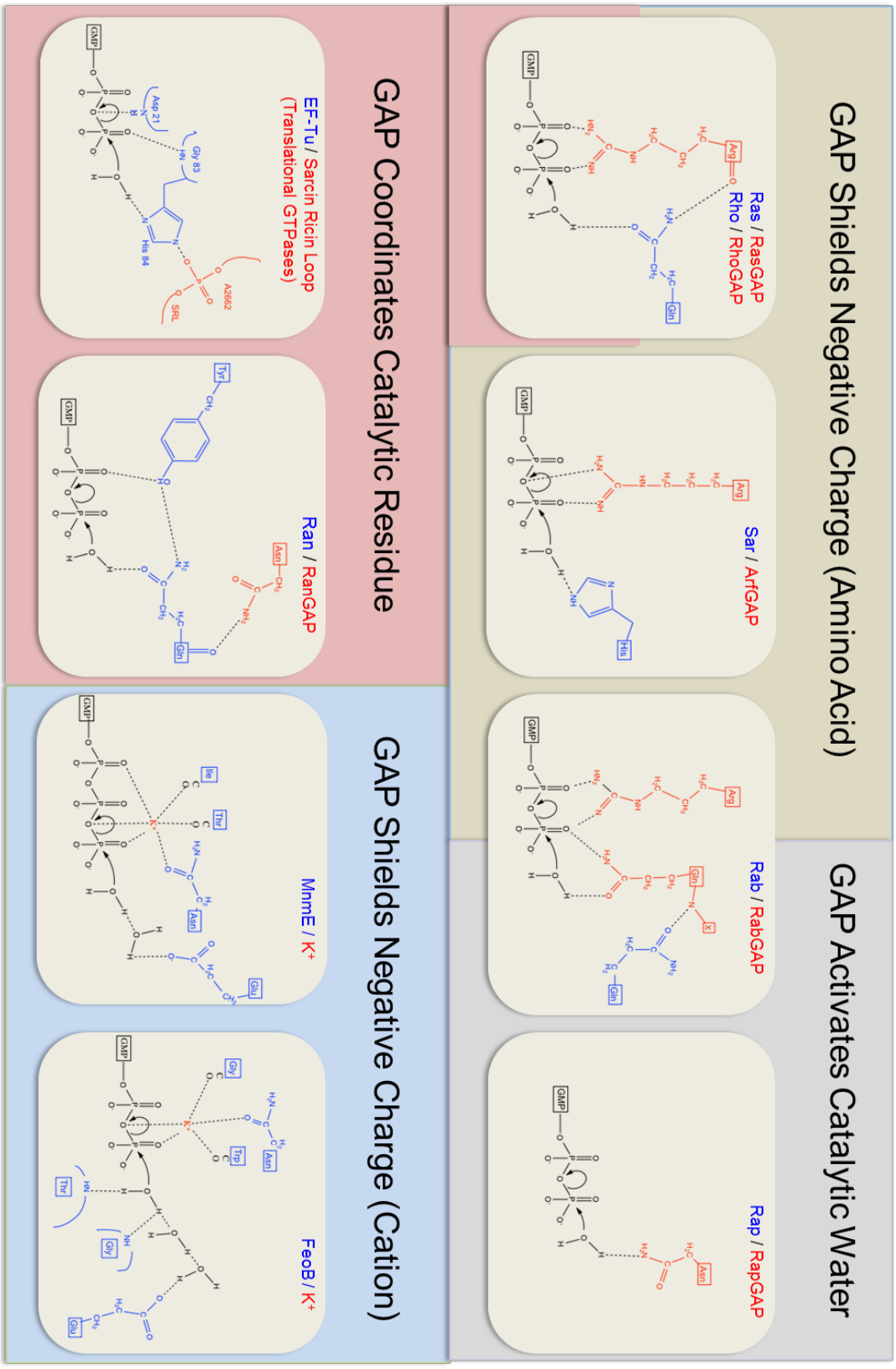


Figure 1.1.1 Catalytic mechanisms of GTPase Activating Factors

1.2 Circularly permuted GTPases

The YawG/YlqF superfamily of GTPases consists of 5 distinct sub-families, all of which show a circular permutation of their G-domain (CPG-domain). These five families can be represented by the following proteins: YlqF (*Bacillus subtilis)*, YqeH/RbgA (*B. Subtilis*), YjeQ (*Escherichia coli)*, MJ1464 (*Methanocaldococcus jannaschii*), and YawG (*Schizosaccharomyces pombe*).[7] As a result of their circular permutation the canonical sequence of motifs, G1-G2-G3-G4-G5, has been rearranged to G4-G5-G1-G2-G3 and this sequence order has been conserved in all circularly permuted GTPases, or cpGTPases. Despite the altered motif arrangement, the overall fold of the G-domain has been well conserved.[7,36] The cpGTPases are multi-domain proteins, which at a minimum possess one C-terminal domain along with an N-terminal CPG-domain. The necessity of a C-terminal domain is a result of the circular permutation of the G-domain as the Switch II region is now found at the C-terminus of the G-domain. Therefore, a C-terminal domain is able to position the Switch II region and reduce its flexibility. Also, conformational changes occurring within the switch regions can now be propagated to the C-terminal domain. When an additional domain is present it is always found at the N-terminus such that the CPG-domain is sandwiched between the other two domains, likely for regulatory reasons.[36] The YjeQ family has an N-terminal OB-fold and a C-terminal Zn-binding domain, both of which are involved in RNA binding, and it has been shown that YjeQ interacts with the ribosome *in vivo.*[36,37] The YlqF and YawG families have a similar CPG-domain and C-terminal α-helical domain, however YawG also possesses an additional N-terminal domain with a coiled-coil motif. Nonetheless these families are frequently discussed together as both have been shown to associate with the large ribosomal

subunit.[36,38] The YqeH family of proteins have a N-terminal Zn finger motif and a C-terminal domain of unknown function. Overall, it appears that members of the YawG/YlqF superfamily, along with their CPG domain, have evolved to use a C-terminal domain to bind RNA and possess a hydrophobic substitution at the position of the Gln^{cat} (described in chapter 1.3).

1.3 HAS-GTPases

A unique subset of catalytically active GTPases has recently emerged into view, whose ability to hydrolyze GTP in the absence of a Gln^{cat} has raised many questions.[39] These proteins have a hydrophobic amino acid substitution at the Gln^{cat} position, and have been termed the HAS-GTPases.[39] The orientation of the hydrophobic amino acid has been shown, through structural alignments, to be retracted away from the nucleotide binding pocket and unable to support a catalytic role during GTP hydrolysis (Figure 2.1.1).[39] Therefore, HAS-GTPases must use a different approach to facilitate catalysis of GTP hydrolysis than that found in Ras and related GTPases. However, previous studies have failed, using both sequence and structural alignments, to identify a specific residue within any of the HAS-GTPases that would be able to interact with the γ-phosphate and is conserved within a subfamily.[39]

Members of the HAS-GTPase family are not related by cellular function, for example MnmE is involved in tRNA modification, FeoB in iron acquisition whereas the functions of YchF and HflX are unknown. Members of the HAS-GTPases include the following families: HflX, EH, Era, EngA, EngB, TrmE, NogI, FeoB, Rsr1, Rb25 and YchF.[39] The HAS-GTPases are an enigma in that hydrophobic substitutions for the

Gln^{cat} in Ras are found in many cancer cell lines, as the enzyme is unable to function in the absence of its catalytic residue. Thus far the elucidation of a catalytic mechanism for the HAS-GTPase family members has been limited to a small number and will be described below.

1.3.1 MnmE

The HAS-GTPase MnmE as it is now commonly referred to, was originally known as TrmE and is a member of the TrmE-Era-EngA-YihA-Septin like (TEES) superfamily of GTPases.[7] MnmE is involved in the enzymatic modification of the wobble uridine (U34) to 5-methylaminomethyl-2-thiouridine ($mnm^5s^2U_{34}$) in tRNAs from bacteria, yeast and mammals.[40,41] The three-domain protein is conserved amongst bacteria, but has not been found in any archaeal species to date. The *mnmE* gene has shown to be essential only when other defects are found within the genome (synergistic effect), in particular those involved in the decoding process of translation.[41-43] For example, the *S. cerevisiae* homologue, Mto1P, when deleted, causes a respiratory deficiency but only when the cell contains additional mutations in the 5S rRNA conferring paromomycin resistance.[42] However, through the use of transposon mutagenesis it was shown that there is a cold-sensitive phenotype in *P. syringae* Lz4W when the *mnmE* gene is interrupted.[44] The synthesis of $mnm^5s^2U_{34}$ is complex and requires several enzymatic reactions that are not fully understood.[41,45] MnmA catalyzes the addition of the thiol group (s^2) following which MnmE interacts with the protein GidA (MnmG) (Mss1p and Mto1P in *S. cerevisiae* or GTPBP3 and MTO1 in humans) to add the cmnm group to position 5 of the U34, however the steps connecting these two events are unclear.[40,46,47]

MnmE is unique with respect to other regulatory GTPases in that it has a high intrinsic GTP hydrolysis rate and a low affinity for GTP and GDP.[41] It was initially proposed that the energy of GTP hydrolysis could be used to catalyze the cmnm group addition reaction, making it very different from Ras-like GTPases. However, MnmE does possess the G1 through G4 motifs making it more likely that MnmE continues to function as a molecular switch. Although GTPase activity is required for modification, it is not sufficient and the actual catalysis of modification may use an essential, conserved cysteine residue (Cys 451 in *E. coli*) found in the central helical domain.[48,49]

As previously mentioned, MnmE is a HAS-GTPase in which a leucine residue has replaced the Gln^{cat} (Figure 2.1.1, Table 2.1.1).[39] Crystal structures of MnmE bound to GDP and AlF_4 show that the catalytic residue is provided in *cis* by a conserved glutamate residue in helix α2 (Figure 1.3.1). In its dimer form, MnmE is able to correctly reorient and position Glu 282, which positions a bridging water responsible for activating the catalytic water molecule.[31] MnmE uses a region of Switch I, termed the K-loop, to coordinate a potassium ion, which is capable of functioning as a GAF.[31,32] In an alignment between crystal structures of MnmE and Ras•RasGAP, it becomes clear that the potassium ion is located in a position similar to that of the 'arginine finger' guanidinium group of the RasGAP. The potassium ion stabilizes the dimer formation due to conformational rearrangements and helps neutralize the negative charge of the phosphates, thereby accelerating GTP hydrolysis.[31]

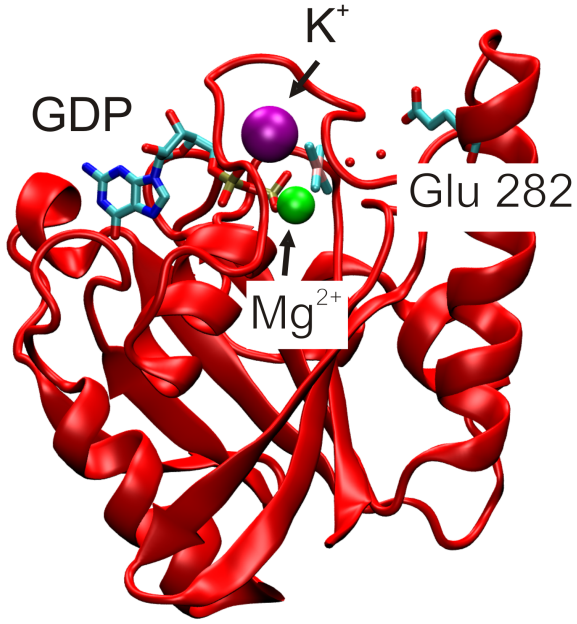


Figure 1.3.1 Cartoon representation of the structure of MnmE bound to GDP•AlF$_4$ (PDB ID 2GJ8). Glu 282 is coordinating a bridging water molecule to the catalytic water. The nucleotide is being coordinated by a Mg^{2+} and a K^+ ion.

1.3.2 YqeH

YqeH is a member of the YlqF/YawG superfamily of cpGTPases, which are also characterized as HAS-GTPases.[7] YqeH is found in mostly gram-positive bacteria, as well as some species of archaea and eukaryotes (AtNOA1 in plant species).[50,51] YqeH is a three-domain protein, which possesses a central CPG-domain with the canonical G1 to G4 motifs (Figure 1.3.2).[50] The N-terminal domain is a zinc-finger domain typically found in RNA binding proteins and the C-terminus shows no homology to any known domains.[36] YqeH has been shown to be important for the maturation of the 30S ribosomal subunit in *B. subtilis as* YqeH-depleted cells show reduced levels of 70S ribosomes, along with an accumulation of free 50S subunits but not 30S subunits.[7,50,52] This is in contrast to Era-depleted cells which

show an accumulation of a stable precursor of 30S subunits with a reduced molecular weight, suggesting that while both YqeH and Era are involved in 30S ribosome biogenesis they have different functions.[53] For YqeH from *B. subtilis,* hydrolysis rates ranging from of 0.2 to 1.2 min^{-1} have been reported, dependent on the concentration of potassium.[35,50,54] While the GTPase activity of Era is stimulated by the 30S ribosomal subunit, the GTPase activity of YqeH is not.[50] Despite this, YqeH has been shown to co-associate with the 30S ribosomal subunit and the formed complex is most stable in the presence of GTP or the non-hydrolysable analog GDPNP.[54] Both, the C-terminal and N-terminal domains are required for this interaction to occur.[54] Interestingly, even though the GTPase activity of YqeH is not stimulated by the 30S subunit, it can be stimulated by the presence of S5, a ribosomal protein known to participate in the early stages of 30S ribosomal assembly.[54] Through the use of homology modeling, in which YqeH was modeled to the transition state structure of MnmE, and a mutational analysis it has been suggested that YqeH uses a conserved aspartate residue analogous to the Glu 282 in MnmE along with a GTPase activating potassium ion.[35] Unlike MnmE however, YqeH does not require dimerization for catalysis and the catalytic residue while still presented from helix α2, it is presented from the N-terminal region, rather than the C-terminus as seen for MnmE, due to the circular permutation of the protein.[35] It is also important to note that these observations are only valid if YqeH does indeed adopt a transition state conformation (and K-loop conformation) similar to that of MnmE on which it was modeled. However, this hypothesis seems likely as a D57I mutation abolished catalytic activity of the enzyme.[35]

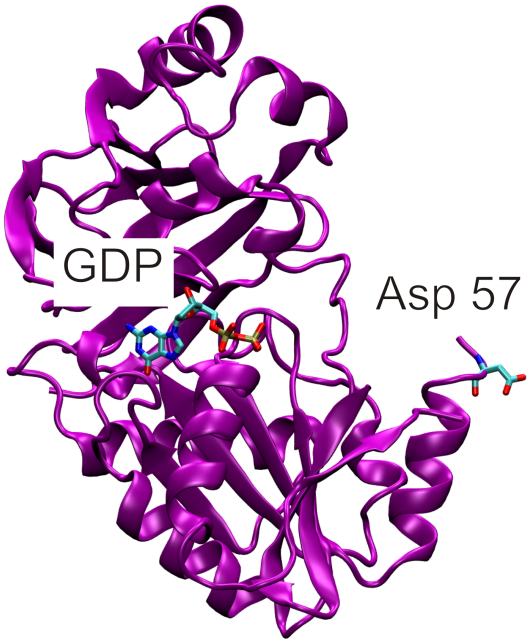


Figure 1.3.2 Cartoon representation of the crystal structure of YqeH bound to GDP (PDB ID 3H2Y). The proposed catalytic aspartate residue (Asp 57) is also shown.

1.3.3 FeoB

FeoB is a prokaryotic, membrane-associated member of the TEES superfamily of GTPases, which is responsible for the import of ferrous ions.[55] The FeoB family is widespread amongst bacteria and archaea, but not found in eukaryotes.[7] The membrane domain consists of at least seven transmembrane α-helices suggesting it can provide more than a simple anchoring role. Linking the GTPase domain to the membrane-embedded domain is a highly variable spacer domain with only a weakly conserved sequence motif R (W/Y) Lx (L/I) xxLExD.[56] Similar to the Era family of GTPases, FeoB has a low affinity for GTP with a K_D of approximately 4 μM (at 20 °C), and an extremely fast rate of GDP release (>1000 s^{-1} at 20 °C).[57,58] Using the soluble N-terminal domains (NFeoB) it has been shown that FeoB hydrolyzes GTP with a k_{cat} of 0.09 min^{-1}. Based on the determined rate constant, FeoB would be constitutively

active, suggesting FeoB requires an unknown GDI to regulate its function.[57] Crystal structures of the entire hydrophilic domain revealed that FeoB forms a trimer around a central cytoplasmic pore, which is closed in the *apo* structure and open in the GTP-bound structure. The linker domain is able to act as an effector to regulate pore opening, explaining its variable, but essential sequence.[59] Based on these findings it has been proposed that FeoB has evolutionary links to eukaryotic G protein coupled membrane processes.[57,59,60]

With more information becoming available about the function of FeoB as an iron transporter the question arose as to how the G-domain is capable of hydrolyzing GTP. As for other members of the TEES superfamily, the GTPase activity of FeoB is stimulated by the presence of potassium ions, but differs in its intrinsic catalytic mechanism. It was surprisingly noted that the co-crystal structure of NFeoB and a transition state analogue ($GDP{\bullet}AlF_4^-$) revealed the lack of a catalytic, water-aligning residue within the nucleotide-binding pocket. Furthermore, a mutational analysis also failed to identify a candidate residue (Figure 1.3.3).[61] Instead it has been proposed that the backbone amides of Thr 35 and Gly 56 coordinate the magnesium ion and position the catalytic water molecule, making a specific catalytic side chain unnecessary.[62] However, it cannot be excluded that a catalytic residue is being supplied in *cis* from another domain, such as the linker region, or in *trans* from an activating partner.

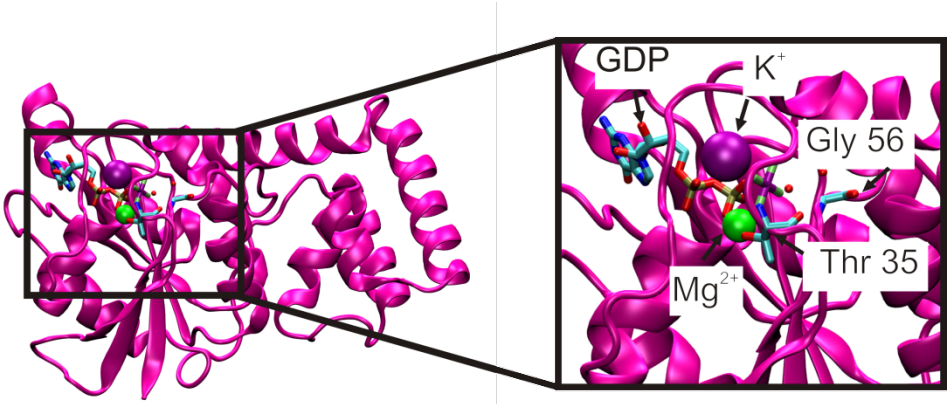


Figure 1.3.3 Cartoon representation of FeoB in complex with GDP•AlF_4^-, Mg^{2+} and a K^+ ion (PDB ID 3SS8). Also shown are Gly 56 and Thr 35 whose backbone oxygen atoms are coordinating the catalytic water molecule.

1.3.4 RbgA

Ribosome biogenesis GTPase A (RbgA), also known as YlqF, is another member of the YlqF/YawG superfamily of cpGTPases. RbgA is widely conserved and is involved in the assembly of the large ribosomal subunit in bacteria and the mitochondria in eukaryotes.[63,64,65] The assembly factor has been shown to be essential for growth in *B. subtilis, Streptococcus pneumoniae* and *Staphylococcus aureus*.[66,67] The structure of RbgA reveals a highly basic N-terminal GTPase domain connected to an acidic C-terminal ANTAR domain by a conserved linker.[68,69]

RbgA is required for the maturation or assembly of the large ribosomal subunit in *B. subtilis* and in eukaryotes (OsNug2 in *Oryza sativa*, AtNug2 in *Arabidopsis thaliana,* Mtg1 in *S. cerevisiae* mitochondria, Lsg1p/Nog2p/Nug2p in other species of Yeast*)*.[38,64,65,70-72] Cells that have been depleted of RbgA accumulate a 45S complex that is lacking three ribosomal proteins - L16, L27 and L36.[64,71] Incorporation of L16 is predicted to occur at a later stage of 50S maturation and in yeast, the RbgA homolog Lsg1 participates in the association of Rpl10p (L16 homolog) with the 60S subunit.[65]

The interaction of RbgA with the 50S ribosomal particle is nucleotide dependent, with maximal association in the presence of a non-hydrolyzable GTP analog.[63] Hydrolysis of GTP promotes RbgA dissociation from 50S subunits, similar to other translation factor GTPases such as EF-Tu.[63] The 50S subunit stimulates GTPase activity approximately 50-fold, as compared to only minor stimulation by the 45S intermediate.[63] It has been proposed that RbgA interacts with the 50S subunit before L16 incorporation, and either acts as a recruitment factor or facilitates conformational rearrangements of the ribosome to facilitate L16 binding, followed by nucleotide hydrolysis and release of RbgA.[63]

Recently, a hypothesis for the catalytic mechanism of GTP hydrolysis in RbgA was proposed.[69] Again, this mechanism relies on a potassium ion to stimulate hydrolysis by acting like an arginine finger, as seen for MnmE, FeoB and YqeH.[63] Furthermore, mutational analyses of RbgA revealed that His 9 is critical for efficient GTP hydrolysis in the presence and the absence of the 50S ribosomal subunit, suggesting its role in catalysis.[69] His 9 is located in a highly flexible N-terminal region of the protein, which is not ordered in available crystal structures. In order to efficiently perform catalysis, His 9 would need to stably interact with a catalytic water molecule, therefore, it has been hypothesized that the 50S subunit is capable of coordinating His 9 in an active conformation, similar to that seen in the translational GTPases (Figure 1.1.1).[69] Gulati *et al.* superimposed the GDPNP bound structures of EF-Tu and RbgA revealing that His 9 of RbgA is located in a similar position to His 84 of EF-Tu, and that both residues are located within a PGH motif conserved in nearly all homologs of both proteins.[69] Therefore, it has been proposed that RbgA is

activated by the 50S subunit, as in the translational GTPases in which the phosphate of residue A2662 of the sarcin-ricin loop (SRL) (23S rRNA) orders the catalytic histidine residue into its catalytic conformation (Figure 1.1.1).[69,73]

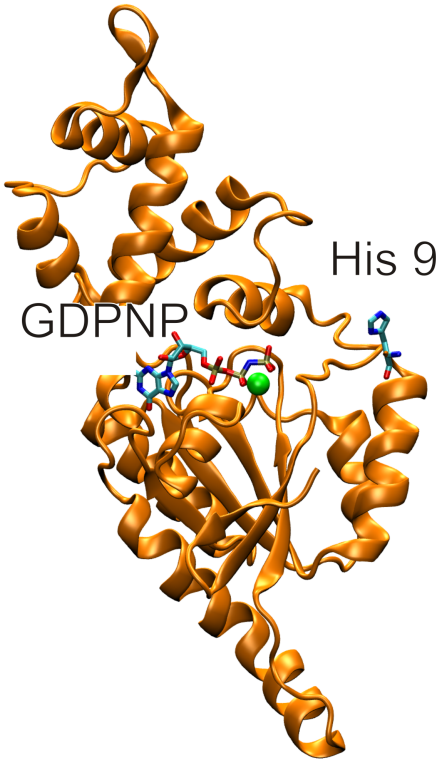


Figure 1.3.4 Cartoon representation of RbgA bound to GDPNP (PDB ID 1PUJ). Also shown is the proposed catalytic residue His 9.

1.4 Introduction to YchF

YchF was initially discovered while examining the nucleotide sequences flanking the peptidyl-tRNA hydrolase-encoding gene (pth) of *E. coli.*[74] It was observed that the open reading frame immediately downstream of pth had amino acid sequence homology to GTP binding proteins, and was named *gtp1*, and later renaimed *ychF*.[74] Both *pth* and *ychF* can be found as either bicistronic or monocistronic transcripts, but interactions between the two gene products have not been observed.[74] Despite its universal conservation, YchF has only been reported to

be essential in *S. aureus,* with a disruption of the *ychF* gene in *E. coli* causing growth retardation and filamentation at an elevated temperature (42 °C).[15,75] Even more intriguing, the cellular function of YchF has yet to be fully understood, but its ubiquitous nature and high levels of sequence conservation between kingdoms suggests a vital role.[76,77]

1.4.1 Structure of YchF

The first crystal structure of YchF was obtained from the bacterial species *H. influenzae* (PDB ID 1JAL).[78] This structure revealed that YchF folds into three subunits: an N-terminal G-domain flanked by an α-helical domain consisting of two long α-helices (A domain) and a C-terminal TGS (Threonyl-tRNA synthetases, GTPases and SpoT) domain of poorly characterized function.[78] The presence of the five highly conserved G motifs classified YchF as a GTP binding protein (Figure 1.4.1).[78] In particular, YchF was placed into the TRAFAC class of P-loop NTPases as it contains a conserved threonine residue in the sequence between the Walker A and Walker B motifs.[7,78] Further classification was provided by the presence of a glycine rich motif in the Switch II region, a distinctive YxFxTxxxxxG motif within Switch I, and the presence of the TGS domain, which are characteristic features of the Obg super-family.[7,78] Notably, a variation was found in the G4 motif, which has been altered from (N/T) KxD to NxxE in YchF. In the canonical G4 motif, the lysine residue is able to interact with the guanine base and coordinate P-loop residues through backbone interactions.[4,21] The aspartate residue in the G4 consensus motif forms hydrogen bonds to the N1 and N2 atoms of the guanine base, and the asparagine residue binds the N7 atom (Figure 1.4.2).[4,21] This is not the case for YchF as the glutamate

residue is positioned too far away to interact with a guanine base and the role of the lysine residue is replaced by a conserved and essential phenylalanine in the β5 loop (Figure 1.4.2).[79] The N7 of the pyrimidine ring in the adenine base is also capable of forming hydrogen bonds with the conserved asparagine residue and a side chain of one of the x residues from the NxxE motif (Val 208 in *E. coli* YchF*)*. As a result of these alterations, YchF from *E. coli, S. cerevisiae, H. influenzae and Trypanosoma cruzi* as well as the human homologue hOla1, have been reported to hydrolyze ATP more efficiently than GTP.[18,33,79,80]

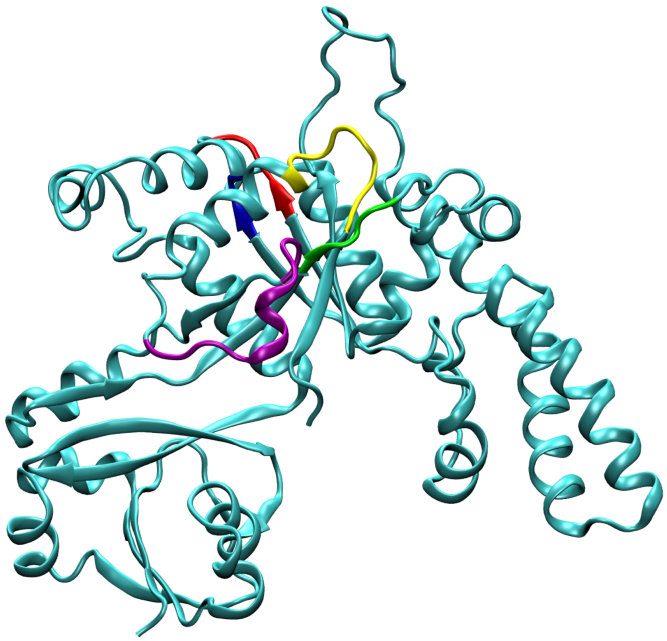


Figure 1.4.1 Homology model of YchF*E. coli* generated with SWISS-MODEL using YchF from *H. influenzae* (PDB ID 1JAL) as a template and the amino acid sequence of *E. coli* YchF (Uniprot accession code: P0ABU2). The G1 motif is shown in yellow, G2 in purple, G3 in green, G4 in red and G5 in blue.

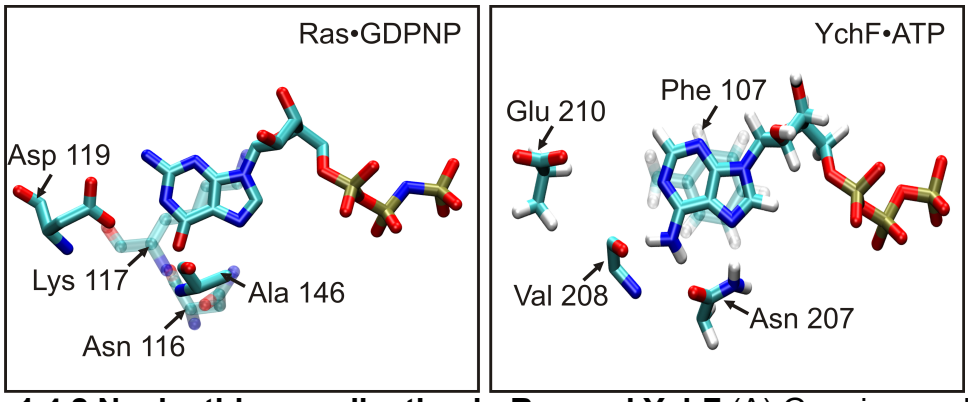


Figure 1.4.2 Nucleotide coordination in Ras and YchF (A) Guanine nucleotide coordination in Ras (PDB ID 5P21). Asn 119 of the NKxD motif and Ala 146 directly coordinate the base, whereas Lys 117 and Asn 116 make stabilizing contacts with the P-loop (not shown). (B) ATP coordination in YchF (*E. coli* homology model). Asn 207, Val 208 and Phe 107 directly stabilize the bound ATP, whereas Glu 210 of the ^{207}NVNE^{210} motif is positioned too far away.

The C-terminal domain of YchF shows homology to a domain found in threonyl-tRNA synthetases, some GTPases, and SpoT (the TGS domain).[78] The α-helical domain resembles domains required for RNA binding in the transcriptional cleavage factor GreA and in seryl-tRNA synthetases (SerRS). Together the three domains form a claw-like structure, with a positively charged cleft between the TGS domain and the α-helical domain suggesting that the function of the TGS and α-helical domain may be to bind nucleic acids.[78]

1.4.2 Insights into YchF's Catalytic Mechanism

The unique ability of YchF to preferentially bind and hydrolyze ATP has already been discussed, and this feature alone would suggest that YchF could possess a catalytic mechanism unique among the P-loop GTPases. However, YchF is also member of the class of HAS-GTPases, as the canonical Gln^{cat} is found to be substituted with a leucine residue.[39] No crystal structure of YchF bound to a transition

state analogue is available, and while there is a crystal structure of the human homolog of YchF bound to AMPPCP, the entire protein is not resolved. Therefore, it is unclear what mechanism YchF employs to achieve intrinsic ATP hydrolysis. Recently a conserved K-loop sequence, responsible for potassium GAF activity, was identified in the G2 motif of YchF species. It has since been confirmed that the ATPase activity of $YchF_{E.\ coli}$ can be stimulated by potassium ions, and that Lys 78 (*E. coli* numbering) along with the K-loop and conserved P-loop asparagine is involved in this specificity.[33] Interestingly, it has never been discussed whether stimulation (or activation) accurately describes the effect of potassium on GTPase activity. Maximal stimulation by K^+ ion has been shown around 200 mM, which incidentally is the intracellular concentration of potassium.[81] Unless faced with osmotic stress the potassium concentration is tightly regulated, therefore it seems likely that GTPases with an affinity for potassium are always coordinating this cation, rather than being transiently 'stimulated'.

1.4.3 YchF as a negative regulator of oxidative stress

Reactive oxygen species, or ROS, play dual roles within biological systems as they are critical for maintaining cell signaling and pathogen defense, but are also capable of causing extensive damage to cellular structures. The redox environment within a cell needs to be carefully balanced, so cells are required to develop defense mechanisms against ROS including repair and preventative mechanisms, physical defenses and both enzymatic and non-enzymatic antioxidants.[82] The imbalance of ROS and antioxidant defense is known as oxidative stress and plays a role in a number of cancers and neurodegenerative disorders such as cancer, aging, trauma

and Alzheimer's disease.[83-85] Recently YchF has been implicated as a negative regulator of the oxidative stress response in *E. coli* and humans.[19,20] YchF is capable of directly interacting with the stress response protein, KatG, resulting in H_2O_2 hypersensitivity when over-expressed.[19,20] Wenk *et al.* were also able to show that YchF is repressed by the transcription factor OxyR in an H_2O_2-dependent manner, elucidating how the cell is able to regulate the expression of YchF and thus KatG inhibition. The observations made in bacteria support those previously seen, in which the human homologue of YchF, hOla1, can negatively regulate the antioxidant response, likely through the regulation of cellular thiols.[19] Moreover, the enhanced ability of YchF-knockdown cells to survive oxidative stress appears to inhibit breast cancer cell migration and invasion.[86] These findings provide valuable venture points for new therapeutic opportunities as P-Loop ATPases are proving to be attractive drug targets.[86,87]

1.4.4 YchF interacts with the Ribosome

It has long been proposed that the core 11 GTPases found in bacteria are required to regulate ribosomal function or to transmit signals from the ribosome to effector pathways.[3] The *T. cruzi* homolog of YchF has been shown to co-immunoprecipitate with both ribosomal subunits and the proteasome. Furthermore a TAP-tag based purification of *in vivo* complexes in yeast revealed that YchF interacts with eEF1, as well as the proteosome.[80,88,89] More recently, it has been shown that $YchF_{E.\ coli}$ forms a much tighter interaction with the 70S ribosome and 50S subunit, as compared to the 30S subunit and that this interaction is nucleotide dependent.[18,33] Even more intriguing is that YchF is only capable of binding to approximately 40% of

the total pool of purified 70S ribosomes, suggesting that YchF in its ATP-bound conformation can bind to only a specific state of the 70S ribosome, and only this functional state of the ribosome is capable of stimulating YchF's ATPase activity approximately 10-fold.[18] Thus far it has not been determined what population of 70S ribosomes YchF is capable of interacting with or how the ribosome is able to stimulate hydrolysis in YchF. It is also unclear how this information ties into recent studies on YchF and oxidative stress.

1.5 Molecular Dynamics Simulations

Molecular dynamics (MD) simulations allow one to observe the movement of molecules on the nanosecond time scale.[90,91] The total length of such simulations are typically on the tens of nanoseconds time scale, therefore no large conformational rearrangements of the molecular system under study can be observed. However, motions of loops or excursions of the side chain network into different functional states can be studied.[92] To set up an MD simulation one must first assign coordinates to the atoms from available resources, frequently crystal structures. Next, these atoms are given velocities from which momenta for each atom can be computed. Using a chosen force field, the forces applied to each atom are computed. A short time later (the time step), new coordinates are derived for the atoms from which velocities and momenta can be calculated. The rest of the steps are repeated and iterated over a given time frame providing a full set of time dependent coordinates making up a molecule's trajectory. In order to include vibrations in the stimulation, a time step on the order of femtoseconds to tenths of femtoseconds is required.[91]

Molecular dynamics simulations make use of the principles of molecular mechanics, or a set of algebraic equations, to describe the forces between atoms and the potential energy of a system. The equations and associated constants used to determine the energy of a system are incorporated into a specific force field. Different force fields use different equations (and constants) to describe aspects of a molecule such as bond stretching, bond bending, torsion angles, hydrogen bonding, van der Waals interactions and electrostatic interactions. In order for MD simulations to be useful, a force field that is able to accurately describe the intermolecular forces and vibrations of the system away from equilibrium must be used. All-atom MD simulations assume that every atom will experience a force, as defined by the force field. The molecule being studied is treated as a ball and spring system in which the atom is the ball, and is described by the center of mass, and the bonds are the springs holding the atoms together.[91,93] Periodic boundary conditions are used to ensure that boundary effects are avoided as the atoms of the system are enclosed in a cell that is replicated to infinity by periodic translations and therefore always subject to the force field, even if they move out of one cell. As molecular systems consist of a large number of particles, MD simulations circumvent this problem using statistical ensembles that can be under different environmental constraints. Ensembles used for MD simulations can have a constant number of particles and a constant volume and energy (NVE ensemble), a constant number of particles and a constant volume and temperature (NVT ensemble) or a constant number of particles and a constant pressure and temperature (NPT ensemble). The MD simulations run for this thesis were performed with an all-atoms model with explicit water molecules and the force

field chosen was the Chemistry at Harvard Macromolecular Mechanics (CHARMM) force field.[94,95] An NPT ensemble (with periodic boundary conditions) was invoked as this ensemble most closely mimics an *in vitro* condition with flasks open to ambient temperature and pressure.

A major disadvantage to MD simulations are the time scales that can be observed (tens of nanoseconds) as transitions from one equilibrium to another in biological systems occur on a much longer time scale (such as large conformational rearrangements required for catalysis).[92] Only a special few have the resources available to produce simulations on the microsecond time scale.[96,97] Therefore, in order to view such transitions one can bias or steer the system in one direction or another through the use of external forces. These simulations in which an external force is applied to a system are called steered molecular dynamics (SMD).[93]

X-ray crystallography and cryo-electron microscopy provide us with high-resolution structures of biological molecules, and while these methods can reveal changes in conformations of a molecule they lack details on the dynamics of such movements. On the other hand, MD simulations can provide a more detailed, visual analysis of the molecular movements that are not available by another technique. Due to restricted time scales, the observed conformations are defined by new side chain conformers or loop conformations, which occur on the nanosecond to millisecond time scale, rather than large conformational changes that occur on the millisecond to second time scale.[92,98] However, MD simulations need to be correlated with experimental data, as they are simply simulations of what could happen in the selected molecular environment. To date, no MD studies have been performed with

YchF, however there are 5 available crystal structures that can be used as a starting point for such studies.

1.6 Objective

In this thesis, both *in vitro* and *in silico,* experiments are performed to study the catalytic mechanism of ATP hydrolysis and the nucleotide binding properties of YchF. The HAS-GTPase group of proteins represents an enigma in their ability to hydrolyze GTP without a conserved Gln^{cat}. As such, several attempts have been made over recent years to understand and characterize their unique mechanisms of nucleotide hydrolysis.

The **first objective** of this thesis, touched upon in Chapter 2, was to identify a catalytic residue(s) responsible for intrinsic ATPase activity in YchF, and compare this mechanism to other known HAS-GTPases. In order to function, a GTPase must bind guanine nucleotides, and the rate at which it does so can provide important information about the functional cycle of the protein. For example, a GTPase with a very slow rate of GDP release will most likely require a guanine nucleotide exchange factor if it is involved in a cellular process that is 10 times faster than the rate of GDP dissociation. Therefore, the **second objective** was to determine the rate constants describing nucleotide binding and dissociation in YchF. In chapter 3, I have examined the kinetic rate constants of nucleotide binding in YchF in both the presence and absence of Mg^{2+} in order to further understand how YchF binds adenine nucleotides.

Chapter 2. Histidine 114 is an essential catalytic residue in the universally conserved ATPase YchF

2.1 Introduction

GTPases are essential to all living organisms in all kingdoms of life as they perform critical cellular functions such as protein synthesis, intracellular signaling, cell cycle regulation, cytoskeletal rearrangement and differentiation.[1,10] The number of GTPases actually required by a given organism can vary to a large extent, but overall the number of GTPases present in archaea and bacteria tends to be smaller than that found in eukaryotes.[3] Interestingly, only eight GTPases (EF-Tu, EF-G, IF2, FtsY, Ffh, YihA, HflX, and YchF) are conserved amongst all domains of life.[8] Of these eight universally conserved GTPases, the cellular roles and functional mechanisms of three (YihA, HflX and YchF) are only poorly understood.

In order to perform their respective functions, GTPases cycle between an active, GTP-bound state in which the protein is able to interact with downstream effectors, followed by GTP hydrolysis to an inactive, GDP-bound state.[5] The mechanisms employed by these GTPases to catalyze GTP hydrolysis can differ significantly, but typically relies on a conserved catalytic residue found immediately downstream of the conserved G3 motif (DxxG), a sequence motif present in all GTPases, which is responsible for magnesium and γ-phosphate coordination.[1,4] Ras, which is frequently used as a prototypical GTPase, possesses a conserved glutamine residue that aligns a catalytic water molecule for an inline attack on the gamma phosphate of GTP.[21,22] This catalytic glutamine (Gln^{cat}) is found to be substituted with a histidine residue in the translation factor superfamily (such as EF-Tu, EF-G and

IF2), threonine in Rap and arginine in Ffh.[5] Consistent with previous literature, Gln^{cat} is referred to as the canonical catalytic residue.[5] In addition to the Gln^{cat}, many GTPases require additional factors to regulate their activity such as GTPase activating factors (GAFs). GAFs increase the rate of hydrolysis allowing the protein to transition from the GTP-bound on-state to the GDP-bound off-state, thereby regulating how long the enzyme remains in its active state.[5,6]

Recently, a novel class of GTPases has been discovered in which the Gln^{cat} is substituted with a hydrophobic amino acid (e.g. Ile or Leu) whose position is retracted away from the nucleotide binding pocket when compared to Ras (Figure 2.1.1 and Table 2.1.1).[39] To date, these so called Hydrophobic Amino acid Substituted GTPases, or HAS-GTPases, include members such as HflX, EH, Era, EngA, EngB, MnmE, NogI, FeoB, Rsr1, Rb25 and YchF, as well as the circularly permuted GTPases (cpGTPases) YqeH and RbgA (YlqF).[39] As these GTPases clearly cannot use the classical catalytic mechanism of GTP hydrolysis it is of great interest to understand how the GTPase activity of these proteins is regulated, how they perform catalysis of GTP hydrolysis and what the consequences of their cellular functions are. The first mechanistic study was done on the enzyme MnmE (TrmE family), involved in the modification of the wobble uridine (U34) in tRNAs from bacteria, yeast and mammals.[41] The Gln^{cat} of MnmE has been substituted with a leucine residue. Interestingly the three-dimensional structure of MnmE in complex with a transition state analogue obtained by X-ray crystallography revealed a glutamate residue (Glu 282) involved in positioning a bridging water capable of activating the catalytic water (Figure 1.2.1).[31] It has also been proposed, based on homology modeling and a

mutational analysis, that another HAS-GTPase YqeH, which is essential for 30S subunit maturation in *B. subtilis*, requires a conserved aspartate residue (Asp 57) which functions analogously to the Glu 282 of MnmE.[7,35,52] Another HAS-GTPase mechanism has been proposed for RbgA, a cpGTPase involved in the assembly of the large ribosomal subunit in bacteria and the mitochondria of eukaryotes.[63,64] Mutational studies of RbgA have revealed a histidine residue (His 9) critical for GTP hydrolysis activity.[69] The last HAS-GTPase whose mechanism has been studied is that of the prokaryotic membrane-associated GTPase FeoB, which shows an interesting alternative to a Gln^{cat}. Co-crystal structures of the N-terminal G-domain of FeoB in complex with a transition state analogue, as well as mutational studies, were surprisingly unable to identify a catalytic side chain. As an alternative, it has been suggested that the backbone amides of two residues, Thr 35 and Gly 56, are capable of both coordinating a magnesium ion and aligning the catalytic water molecule (Figure 1.2.3).[55,62,99] Despite the recent work focusing on the HAS-GTPase family, the catalytic mechanisms for most HAS-GTPases remain poorly understood and more examples of this family of GTPases need to be studied mechanistically in order to shed light on catalytic strategies employed.

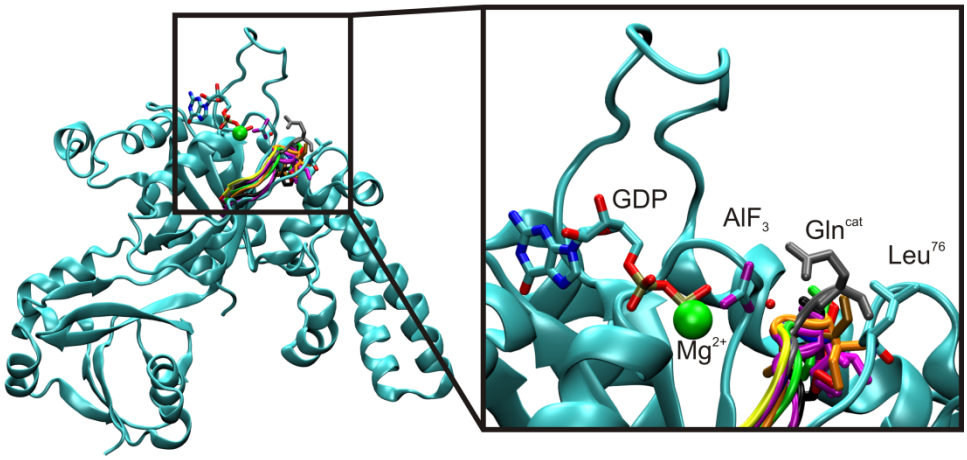


Figure 2.1.1 YchF is a member of the HAS-GTPase family Superposition of Ras (grey), YqeH (purple), MnmE (red), HflX (black), EHD2 (green), FeoB (pink), Era (blue), YsxC (brown) and YlqF (orange) to YchF (cyan). Also shown is GDP•AlF_3 and the attacking water molecule (red sphere in line with AlF_3) from Ras•RasGap (PDB ID 1WQ1). Superposition was based upon structure of the conserved G1 motif (P-loop).

Table 2.1.1 Structural alignment of the G3 motif of 9 HAS-GTPases and Ras.

Protein	G3 motif (DxxG)
YchF	DIAG**L**
FeoB (PDB ID 3lx5)	DLPG**L**
RbgA (YlqF) (PDB ID 1puj)	DTPG**I**
MnmE (PDB ID 2gj8)	DTAG**L**
Era (PDB ID 1WF3)	DTPG**I**
YqeH (PDB ID 3H2Y)	DTPG**I**
HflX (PDB ID 2QTH)	DTVG**F**
EHD2 (PDB ID 2qpt)	DTPG**I**
YsxC (EngB)(PDB ID 1SVW)	DVPG**Y**
Ras (PDB ID 1WQ1)	DTAG**Q**

The universally conserved ATPase, YchF, is also classified as a HAS-GTPase, owing to the leucine residing immediately downstream of its G3 motif (Figure 2.1.1 and Table 2.1.1).[33,39] The primary sequence of YchF adopts a three-domain structure with a central G-domain, required for nucleotide binding, flanked by an α-helical domain and a C-terminal TGS domain (named for its presence in Threonyl-tRNA synthetases, GTPases and SpoT), both predicted to play a role in nucleic acid binding.[78] Aside from being a HAS-GTPase, YchF has an altered G4 motif, the motif

specifying guanine nucleotide binding, in which the canonical (N/T) KxD sequence is changed to NxxE. This has resulted in an altered nucleotide affinity in which YchF binds and hydrolyzes ATP more efficiently than GTP.[18,78,79] YchF has also joined a number of GTPases (FeoB, MnmE, YqeH, EngA) whose GTPase activity can be stimulated by the presence of a potassium ion, which is capable of functioning similarly to an arginine finger as seen in Ras (Figure 1.1.1).[31,33,34,54,57,62,63] Thus far, YchF has been implicated in numerous cellular functions including protein synthesis, iron utilization in *Brucella melitensis* and *Vibrio vulnificus*, protein degradation, virulence in *S. pneumoniae* and *V. vulnificus*, the oxidative stress response and infection defense response in rice.[18-20,78,80,86,89,100-103] Likely due to its initial characterization as a GTPase, only one crystal structure of YchF bound to an adenine nucleotide is available, that of hOla1 in complex with AMPPCP.[79] Unfortunately the crystal structures do not provide structural information for critical regions of the protein including both switch regions and a flexible loop in the G-domain. Mutational analyses have revealed that two residues, Asn 13 and Lys 78 (*E. coli* numbering), play important roles in the potassium stimulation of ATP hydrolysis, but residues essential for ATP hydrolysis have yet to be identified.[33] As disussed above, it is clear that YchF may be important for many biological functions, highlighting the importance of understanding the catalytic mechanism of ATP hydrolysis.

In order to elucidate the catalytic mechanism of the universally conserved HAS-GTPase, YchF, we have taken a classical biochemical and *in silico* approach to characterization. The use of molecular dynamics (MD) simulations allowed us to

overcome the absence of a catalytically competent structure of YchF as well as providing a dynamical view of the protein. Using this combination of methods we have identified His 114, a highly conserved residue, as an intrinsic catalytic residue. Our data also suggests that YchF is similar to the HAS-GTPase RbgA and the translational GTPases in its mechanism of catalysis.

2.2 Materials and Methods

DH5α cells were purchased from New England Biolabs and BL21-DE3 competent cells were purchased from Novagen. PCR primers were purchased from Invitrogen and Integrated DNA technologies. Restriction enzymes were obtained from Fermentas. Radiochemicals were purchased from Perkin-Elmer. Adenine nucleotides and fluorescent nucleotide analogs were purchased from Invitrogen. All buffers were filtered through 0.45 μM Whatman nitrocellulose membranes.

2.2.1 Sequence Alignments

All YchF protein sequences were obtained from the UniProt (www.uniprot.org) and Entrez Gene (www.ncbi.nlm.nih.gov/gene) databases. Multiple sequence alignments were performed using ClustalW2 provided by EBI (http://www.ebi.ac.uk/Tools/clustalw2). GeneDoc software was used to visualize the sequence alignments. Conservation (identity) was displayed using the GeneDoc software program with four levels of shading (100% black, 80 to 100% dark grey, 60 to 80% light grey and less than 60% white).

2.2.2 Cloning and Site-directed Mutagenesis

Amino acid substitutions were introduced into pET28a containing the full-length sequence coding for an N-terminal His_6-tagged YchF using the Quikchange^{TM} method. An alanine, glutamine, lysine, arginine or phenylalanine residue was introduced into position 114 substituting the native histidine in this position in the YchF protein. A BspEI restriction site was also introduced at the DNA level to allow for restriction screening, with the exception of the alainine variant. The same residues were introduced into position 102 and removing a PvuI restriction site enabled restriction screening. Primers (Table 2.2.1) were obtained from Invitrogen. Reactions were carried out in a $T_{Gradient}$ (Biometra) thermocycler: 25 µL of the reaction mixture contained 1000 ng of template plasmid DNA, 0.4 µM primer pair, 400 µM of each dNTP, and 3 units of DNA polymerase (*Pfu*, Fermentas). For all primers the position of the mutation is denoted in bold and the new restriction site introduced or removed to enable restriction screening is underlined.

Table 2.2.1 YchF Mutagenesis Primers

Amino acid substitutions	Forward Primer (5'-3')	Reverse Primer (5'-3')	T_m
H114A	GACAACATCATT**GC**CG TTTC**G**GGCAA	TTGCCCGAAACG**GC**AA TGATGTTGTC	62.4°C
H114K	GACAACATCATT**A**A**A**G TTTCCGG**A**AAAGTTAA CCCGGC	GCCGGGTTAACTTT**T**C CGGAAAC**T**T**T**AATGAT GTTGTC	74.5°C
H114F	GACAACATCATT**TT**CGT TTCCGG**A**AAAGTTAAC CCGGC	GCCGGGTTAACTTT**T**C CGGAAACG**AA**AATGAT GTTGTC	73.0°C
H114R	GACAACATCATTC**G**CG TTTCCGG**A**AAAGTTAA CCCGGC	GCCGGGTTAACTTT**T**C CGGAAACG**C**GAATGAT GTTGTC	79.0°C
H114Q	GACAACATCATTCA**G**G TTTCCGG**A**AAAGTTAA CCCGGC	GCCGGGTTAACTTT**T**C CGGAAAC**C**TGAATGAT GTTGTC	77.8°C
H102A	CCGTGAAACCGAAGCG ATT**T**GGT**GC**CGTTGTTC GCTGC	GCAGCGAACAACG**GC**A CC**A**ATCGCTTCGGTTT CACGG	70.6°C
H102R	CCGTGAAACCGAAGCG ATT**T**GGTC**G**CGTTGTTC GCTGC	GCAGCGAACAACG**C**GA CC**A**ATCGCTTCGGTTT CACGG	70.4°C
H102F	CCGTGAAACCGAAGCG ATT**T**GGT**TT**CGTTGTTCG CTGC	GCAGCGAACAACG**AA**A CC**A**ATCGCTTCGGTTT CACGG	68.2°C
H102K	CCGTGAAACCGAAGCG ATTGGT**AAG**GTTGTTC GCTGC	GCAGCGAACAACCTTA CCAATCGCTTCGGTTT CACGG	67.9°C
H102Q	CCGTGAAACCGAAGCG ATTGGT**CAG**GTTGTTC GCTGC	GCAGCGAACAACCTGA CCAATCGCTTCGGTTT CACGG	69.1°C

The reaction was carried out by heating the reaction mixture to 95 °C for 5 min followed by 16 cycles of 95 °C for 45 sec, annealing temperature (T_m – 15 °C) for 1 min, 68 °C for 15 min and subsequent final elongation at 68 °C for 15 min. To remove the template DNA the reactions were treated with restriction enzyme *DpnI* (Fermentas) overnight at 37 °C. Two µL aliquots of a 1:10 or 1:100 dilution of the product were transformed into 20 µL *E. coli* DH5α competent cells (New England

Biolabs), which were grown on Luria-Bertani (LB) media supplemented with 50 µg/mL of kanamycin. Plasmids were isolated from selected colonies using a mini-prep purification kit (EZ-10 Spin Column Plasmid DNA Kit, BioBasic). Positive mutants were identified by restriction digestion with BspEI or PvuI (New England BioLabs Inc.), respectively. All mutants were confirmed by sequencing (Genewiz). The resulting plasmid was transformed into the *E. coli* strain BL21-DE3 (Invitrogen) for YchF over-expression.

2.2.3 Protein Expression and Purification

Overnight cultures of transformed cells (both wild type and mutant) were grown in LB medium supplemented with 50 µg/mL of kanamycin to mid-log phase (OD_{600} = 0.6) and induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG, BioBasic). Cells were grown for an additional 3 hours at 37 °C, harvested by centrifugation (5,000 x *g* for 10 min at 4 °C) and stored at -80 °C prior to use. To confirm the over-expression of YchF, 1 OD_{600} samples were taken every hour after induction and lysed with 8 M urea in $TAKM_7$ (50 mM Tris-Cl pH 7.5 (20 °C), 70 mM NH_4Cl, 30 mM KCl and 7 mM $MgCl_2$) and analyzed by 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) at 80 V for 20 min followed by 200 V for 60 min (BioRad Mini Protean 3 System). Gels were stained with Coomassie Brilliant blue; all other SDS-PAGEs were performed in a similar manner.

Similar purification procedures were followed for wild type and mutant proteins. All purification steps were carried out at 4 °C. The harvested frozen cell pellet from over-expression was thawed on ice and re-suspended in 7 mL Buffer A (50 mM Tris-Cl (pH 7.5), 60 mM NH_4Cl, 7 mM $MgCl_2$, 7 mM β-mercaptoethanol, 1 mM

phenylmethylsulfonyl fluoride (PMSF), 300 mM KCl, 10 mM imidazole and 15% v/v glycerol) per gram of cells. Lysozyme was added to a final concentration of 1 mg/mL per gram of cells and stirred in a 4 °C fridge for 1 hr, followed by the addition of 12.5 mg sodium deoxycholate per gram of cells and a few crystals of DNAse I and further stirring for 30 min. Cell debris was removed by centrifugation at 5,000 x *g* for 30 min and 30,000 x *g* for 45 min using a JA-16 rotor (Beckman). The YchF protein was purified from the cleared lysate (S-30 extract) using affinity chromatography (5 mL Ni^{2+} -Sepharose resin (GE Healthcare) equilibrated with Buffer A. Unbound protein was washed away with 18 column volumes of binding buffer and 25 column volumes of Buffer B (Buffer A with 20 mM imidazole) Bound protein was eluted with 10 column volumes of Buffer E (Buffer A containing 250 mM imidazole). The protein was concentrated and the buffer exchanged with Q-Sepharose buffer A (50 mM Tris-Cl pH 7.5 (8 °C), 10 mM $MgCl_2$, 5 mM EDTA, 1 mM dithiothreitol (DTT)) using a Vivaspin 20 (30,000 MWCO; GE Healthcare). Further purification was done by anion exchange chromatography (Q-Sepharose XK26/10 Fast Flow (GE Healthcare)) equilibrated in Q-Sepharose Buffer A. A salt gradient of 0% to 100% Q-sepharose Buffer B (Buffer A plus 1 M KCl) was used to elute bound proteins and any RNA previously bound to the protein. Fractions containing YchF were pooled, concentrated and further purified by size exclusion chromatography (Superdex 75 XK26/100 column (GE Healthcare)) equilibrated in the final storage buffer $TAKM_7$ (50 mM Tris-Cl pH 7.5 (4 °C), 70 mM NH_4Cl, 30 mM KCl, and 7 mM $MgCl_2$)) with 1 mM DTT. All elutions were analyzed by SDS-PAGE and fractions containing pure YchF were pooled, concentrated, flash frozen in liquid nitrogen and stored at -80 °C. Protein concentration was determined

photometrically at 280 nm using a molar extinction coefficient of 16,305 M^{-1} cm^{-1} (calculated using ProtParam)[104] and confirmed using the Bradford Protein Assay (BioRad).

2.2.4 Preparation of Ribosomes

Ribosomes were prepared as described in Becker *et al*.[18]

2.2.5 Fluorescence Measurements

Fluorescence measurements were carried out using a QuantaMaster Fluorescence Spectrophotometer (Photon Technology International, Birmingham, NJ, USA) with a 2 mm x 10 mm quartz cuvette (18F-Q-1MS; Starna Cells, Inc, Atascadero, CA, USA).

$YchF_{E.coli}$ contains a single tryptophan residue located in the TGS domain, which can be used to monitor YchF's intrinsic fluorescence upon nucleotide binding.[18] One µM YchF in $TAKM_7$ was titrated with adenine nucleotides and excited at 280 nm with a slit width of 1 nm. Fluorescence emission was monitored from 295 to 400 nm at a step size of 1 nm and a slit width of 5 nm. Following the addition of ADP or ADPNP to YchF, the solution was equilibrated for 1 min prior to excitation. Equilibration was not performed when ATP was added due to the intrinsic ATPase activity of YchF. As a negative control the fluorescence of $TAKM_7$ buffer with nucleotides was subtracted from all nucleotide-binding experiments. Fluorescence intensities were also corrected for dilution of the protein. The equilibrium dissociation constant (K_D) was determined by plotting the change in fluorescence at 337 nm (F) versus the nucleotide concentration ([nt]) with respect to the initial fluorescence (F_o)

and the amplitude of the signal change (ΔF_{max}). Data were fit with a hyperbolic function (equation 1) using the TableCurve (Jandel Scientific) software and Prism (GraphPad Software). Final K_Ds and their standard deviations were calculated from at least three independent experiments.

$$F = F_O + [(\Delta F_{\max} \times [nt])/(K_D + [nt])] \quad (1)$$

2.2.6 ATP Hydrolysis Assays

The multiple turnover ATP hydrolysis (ATPase) activity of YchF and YchF variants was measured at 37 °C in $TAKM_7$ buffer, unless otherwise stated, by following the liberation of $^{32}P_i$ from [γ-^{32}P]-ATP (200 dpm/pmol). ATP was charged by incubating with 0.25 µg/µL pyruvate kinase and 3 mM phosphoenolpyruvate for 15 min at room temperature (RT) and then 15 min at 37 °C. All other components were charged by incubation for 30 min at RT.

The intrinsic ATPase activity of YchF WT and variants were determined using reactions containing 5 µM YchF and 125 µM [γ-^{32}P]-ATP. The influence of 70S ribosomes on YchF's ATPase activity was measured using 1 µM YchF, 125 µM ATP and 5 µM 70S ribosomes. Michaelis-Menten titrations were performed under the same conditions, however the concentration of 70S ribosomes varied from 0 to 15 µM.

Hydrolysis reactions were started by the addition of radiolabelled-nucleotides, then 5 µL aliquots were taken at different time points and the reaction was quenched in 50 µL of 6 M formic acid or 50 µL of 1 M $HClO_4$ with 3 mM K_2HPO_4 dependent upon the separation method used, thin layer chromatography or extraction respectively. Thin layer chromatography was used for intrinsic ATPase assays to

separate the $^{32}P_i$ from [γ-^{32}P]-ATP and visualized using a Typhoon Trio scanner (GE Healthcare). The relative amount of $^{32}P_i$ formed was determined using ImageJ.[105] For 70S stimulated hydrolysis assays inorganic phosphate was extracted using 300 µL 20 mM Na_2MoO_4 and 750 µL isopropyl acetate. Samples were vortexed for 10 min and centrifuged at 15 800 x *g* for 5 min. The upper aqueous layer containing free $^{32}P_i$ was extracted and added to 2 mL of scintillation cocktail (MP EcoLite) in 7 mL polyethylene scintillation vials (PerkinElmer). $^{32}P_i$ disintegrations per minute (dpm) were counted with a PerkinElmer TriCarb 2800TR liquid scintillation analyzer.

The pH dependent ATPase assays were performed similarly to the intrinsic ATPase assays with $TAKM_7$ buffer used at a pH range of 7 to 9. For those experiments performed below a pH of 7 an 2-(N-morpholino)ethanesulfonic acid (MES) buffer was used.

All hydrolysis experiments were corrected for the background hydrolysis of ATP at 37 °C. Background intrinsic ribosome ATPase activity was also subtracted from experiments containing YchF and 70S ribosomes. ATPase experiments were performed at least in triplicate to determine standard deviations. For Michaelis-Menten titrations the initial velocity at different 70S concentrations was determined within the first 10 min of the reaction to ensure that consumption of the substrate was negligible. The 70S concentration dependence of YchF's ATP hydrolysis rate was fit with a hyperbolic function (Michaelis-Menten) (equation 2) to determine the parameters v_{max} and K_M.

$$v = (v_{\max} \times [S])/(K_M + [S]) \quad (2)$$

2.2.7 MD Simulations

The initial model for $YchF_{E.\ coli}$ was obtained by constructing a homology model using the Swiss-Model server and the crystal structure of *H. influenzae* YchF as a template (PDB ID 1JAL).[106-109] The conformation of Switch I (residues 28 to 41 which were missing in the 1JAL structure) was modeled to be identical to Switch I of the YchF structure from *T. thermophilus* bound to GDP (PDB ID 2DBY). In order to model the adenine nucleotide in the $YchF_{E.\ coli}$ model we used the conformation of AMPPCP bound to the human homolog of YchF, hOla1 (PDB ID 2OHF). Transformation of the AMPPCP to ATP and ADP was done manually. In order to position the magnesium ion associated to the bound nucleotide we aligned the nucleotide with either EF-Tu bound to GDP or GDPNP (PDB ID 1EFC and 1EFT respectively). Hydrogen atoms were added to all models using psfgen in the NAMD software package, and histidine side chains were protonated at the ε-nitrogen only.[93] Initial models were minimized in a vacuum and then placed in a water box extending at least 10 Å from the protein in all directions. Water molecules present in the respective crystal structures were included in this box, and all other waters were added at random using the SOLVATE package in NAMD 2.7b1.[93] Relaxation of the solvated system was achieved by minimizing the positions of water molecules with no constraints, followed by minimization of the protein and ligand atoms in two iterative rounds (10,000 steps each). Sodium ions were then added in random positions by the AUTOIONIZE package in VMD 1.8.7 to neutralize the total charge of the system, followed by a final minimization of all components of the system. YchF•ATP•Mg^{2+} and YchF•ADP•Mg^{2+} had total charges of -17 and -16 and therefore 17 and 16 sodium ions were added, respectively. The system was considered to be minimized when no change in energy

was observed for at least 1,000 steps.[110] Minimizations, and subsequent equilibrations and MD simulations, were performed with periodic boundary conditions using the NAMD software package.[93] Minimized models were initially equilibrated at 300 K and 350 K for 150 ps at constant pressure (1 atm). Production phase simulations were started using velocities from the 300 K equilibration and coordinates from the 350 K equilibration. Simulations were performed over 50 ns at 300 K with a step size of 0.5 fs using the CHARM22 parameters for proteins and CHARMM27 parameters for nucleic acids as implemented in the NAMD package.[93,94] MD simulations were performed in an NPT ensemble in which the pressure was maintained at one atmosphere with a Nosé-Hoover Langevin piston and the temperature was controlled using Langevin dynamics. All simulations were performed in the NAMD software package, and visualization was carried out in VMD.[93,110]

Snapshots of each MD simulation were saved every 0.5 ps, and trajectories were fitted with the software Carma in order to remove any rotations of the protein complex or translation of the center of mass.[111] The root mean-square deviation (RMSD) and root mean-square fluctuation (RMSF) calculations were performed using in-house written scripts invoked with the VMD software package.[110]

In silico simulated annealing experiments allow the heat denaturation of a molecule (or part of a molecule) followed by a slower cooling process back to ambient temperatures thereby enabling the exploration of alternative conformations and allowing the loop to stabilize in a lower energy conformations. Simulated annealing experiments were performed on both the switch regions and the flexible His 114-containing loop, either all together or the flexible loop alone. Step one, or the heating

step, involved heating the selected residues and water to 1000 K over 200,000 steps, denaturing any secondary structure. The second step, or the cooling step, involved cooling the water and desired residues back to 300 K over 1 ns in 200 fs steps. Simulated annealing calculations were performed at least 25 times in order to sample a wider range of conformations.

To assess whether the bound nucleotide has an effect on the conformation of the 'flexible' His114-containing loop we measured the distance between the α carbons of serine 16 in the P-loop as a reference point and histidine 114. The distribution of distances during the different simulations were analyzed using a histogram plot for each simulation with bin sizes of 0.4 Å x 0.4 Å. Distance measurements were taken every 0.5 ps during the equilibrium MD simulations from the final cooled structure of the simulated annealing experiments.

Steered molecular dynamics (SMD) simulations were combined with simulated annealing in order to determine if His 114 was capable of occupying a conformation consistent with a hypothetical hydrolysis competent state. In order to determine how close to the gamma phosphate His 114 would need to be in order to align a catalytic water molecule we used the coordinates of Glu^{cat} from MnmE. The flexible loop and the switch regions of YchF•ATP•Mg^{2+} were heated up to 1000 K from the 10 ns structure and cooled back down to 300 K over 1 ns as described previously. Cooling was carried out over 1 ns as previously described, however, an SMD driving force was attached to the ε nitrogen (Nε1) of His 114 and a point in space determined by the location of the OE1 of Glu^{cat} from MnmE. A spring constant of 5 kcal/mol/$Å^2$ was used and pulled at a rate of 0.1 Å/ps.

2.3 Results

2.3.1 pH Dependence of Intrinsic ATPase Activity

The catalytic mechanism of ATP hydrolysis in YchF has yet to be determined. In an attempt to elucidate the catalytic machinery a pH titration of the intrinsic ATPase activity has been performed (Figure 2.3.1). Under conditions described in chapter 2.2.5, YchF hydrolyzes ATP at a rate of 110 ± 6 pmol/min, and this activity shows a pH dependence.[18] The pH dependence of the intrinsic ATP hydrolysis activity of YchF (Figure 2.3.1) is consistent with a single ionizable group. Intrinsic ATPase activity is maximal at a pH equal to or above 7.5 and negligible below a pH of 6. From the log plot of the specific activity as a function of pH a pKa of ≈ 6.7 ± 0.5 for the ionizable group can be obtained (Figure 2.3.1, panel B).

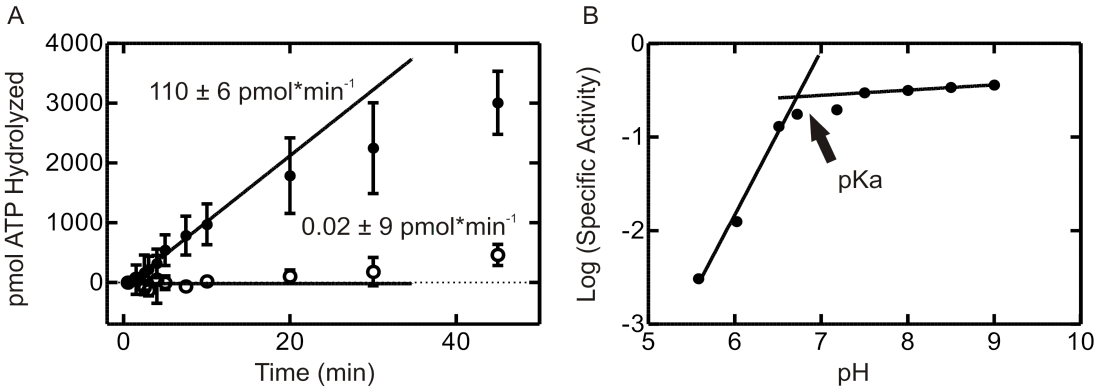


Figure 2.3.1 pH dependence of YchF's intrinsic ATPase activity (A) Time dependence of ATP hydrolysis incubated in the presence of 5 μM YchF (closed circles) or no protein (open circles). A linear function was fit to the initial phase of the reaction (first 10 minutes) to give a rate of ATP hydrolysis (pmol / min). (B) pH dependence of the specificity activity of YchF (closed circles).

A pKa of 6.7 is consistent with the participation of a histidine residue(s) (pKa of 6.04) during catalysis.[112] Based on the importance of a putative His^{cat} for the function of YchF, this histidine should be evolutionary conserved. Therefore, an alignment of YchF from 116 bacterial species was performed to determine the conservation levels

of the histidines present in YchF$_{E.coli}$ (see appendix for the detailed alignment). YchF$_{E.coli}$ has four histidine residues of varying conservation: 76%, 98%, 17% and 100% identity for residues 102, 114, 145 and 308 respectively (Figure 2.3.2). Of these, only His 102 and 114, are located within the G-domain and therefore in a position capable of interacting with a catalytic water molecule located close to the γ-phosphate of the bound nucleotide. Interestingly, *Wigglesworthia glossinidia* has a lysine residue in position 114 which could act as a general base, or the glutamate residue in position 116 could play this role as well. *Streptococcus pneumoniae* shows an insertion in the loop region (residues 105 to 121) with an arginine residue in position 114, glutamate in position 117 and glutamine in position 118, all of which could act as a catalytic residue. Therefore, although two bacterial species of the 116 examined did not have the conserved histidine residue at position 114, each displayed alternative residues capable of acting analogously during catalysis. The residue found in position 102 shows more variety, with the histidine being replaced with an asparagine, methionine, glutamine or tyrosine approximately 8%, 1%, 13% and 2% of the time respectively. However, His 114 and His 102 are 10.7 Å and 14.1 Å away from a position that would be compatible with the function in aligning the catalytic water (Gln 61 in Ras and His 85 in EF-Tu) or the stabilization of the developing negative charge in the transition state (Arg 178 in $G_{i\alpha1}$ and 174 in $G_{t\alpha}$).[21,113,114]

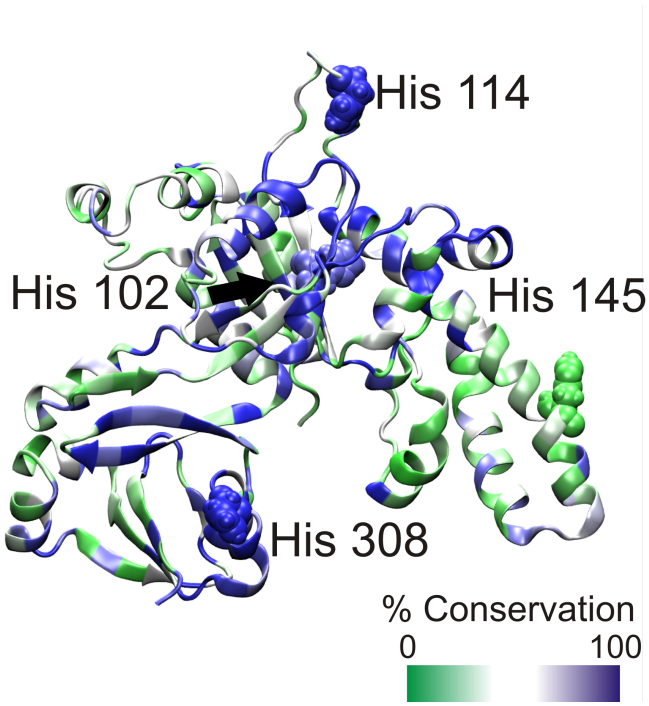


Figure 2.3.2 YchF$_{E.coli}$ contains four histidine residues Homology model of YchF$_{E. coli}$ generated with SWISS-MODEL using YchF$_{H. influenzae}$ (PDB ID 1JAL) as template and the amino acid sequence of YchF$_{E. coli}$ (Uniprot accession code: P0ABU2), represented in cartoon. The level of conservation of each residue is depicted using a green/white/blue color scale based on the alignment of YchF from 116 bacterial species. The four histidine residues present in YchF$_{E coli}$ are indicated and shown in space fill: His 102 (76% conserved), His 114 (98%), His 145 (17%) and His 308 (100%).

2.3.2 Molecular Dynamics Simulations of YchF

Previous work has demonstrated that the intrinsic dynamics of proteins contribute to their function.[115] Structural information obtained by x-ray crystallography only provides limited insight into these features of the molecular system under study. To this end, the use of molecular dynamics simulations can provide important information on the flexibility and dynamic properties of biomolecules. In order to investigate the potential of His 102 and His 114 to participate in the catalysis of ATP hydrolysis in YchF by entering the catalytic shell around the bound nucleotide, three homology models of YchF$_{E. coli,WT}$ in complex with either ATP, ADP or in the *apo* state

were constructed (see 2.2.7 for details). These homology models were subsequently hydrated with explicit waters and subjected to energy minimization. The stability of these models during simulations was scored by measuring the RMSD of the backbone atoms with respect to the initial structure throughout the whole simulation (Figure 2.3.3 A). After an initial increase in RMSD by approximately 2-3 Å, dependent upon the simulation, the models appear to stabilize after 5 ns, however the YchF•ADP•Mg^{2+} simulation does show a 1-2 Å fluctuation around 35 ns. The fluctuation stems from conformational differences mainly in the TGS domain, the flexible loop, the alpha helical domain as well as in helix α3 and Switch I of the G-domain (Figure 2.3.4). The flexibility of the protein in the different complexes as determined by measuring the Cα RMSF values over the final 40 ns varies only marginally (Figure 2.3.3 B). The *apo* state of YchF shows greater flexibility in residues 28 to 40, which makes up a majority of the Switch I region, as well as in residues 246 to 265 of helix α10. In order to determine the role of the conserved G-domain histidines we narrowed our focus and observed that in all three simulations His 102 showed very little flexibility. In fact, the side chain of His 102 is pointing away from the nucleotide binding pocket due to a stable interaction between Nδ1 of His 102 and Hε1/Hε2 of Tyr 204 (the interaction flips between Hε1 and Hε2 around 22 ns during the YchF•ATP•Mg^{2+} simulation but is stable with Hε1 for the YchF•ADP•Mg^{2+} and the *apo* state simulation) (Figure 2.3.5) making it an unlikely candidate for a catalytic residue. This interaction is conserved when YchF is bound to ADP and in its *apo* state (not shown).

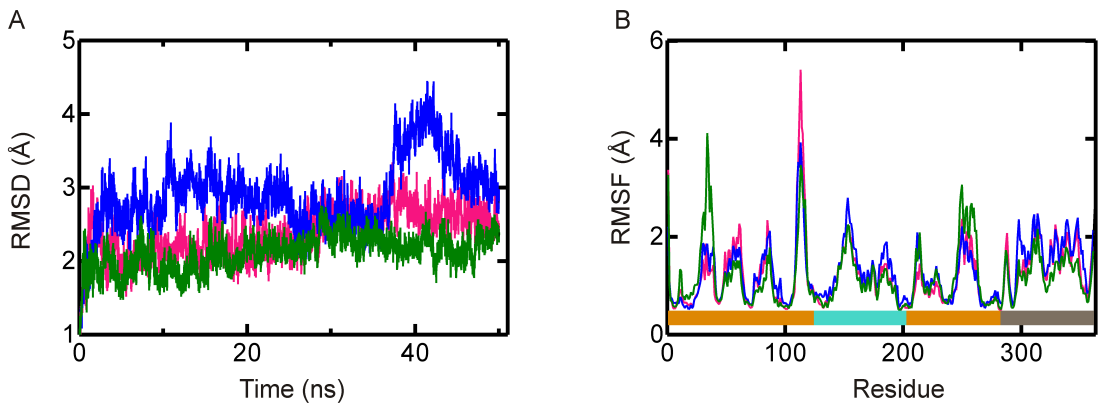


Figure 2.3.3 Structural changes of YchF during MD simulations (A) RMSDs of $YchF_{WT}$ in complex with ATP (pink), ADP (blue) and in its *apo* state (green) from its initial conformation at 300 K. RMSDs were calculated using all backbone atoms. (B) Cα RMSF for $YchF_{WT}$ in complex with ATP (pink), ADP (blue) or in its *apo* state (green). Colored bars at the bottom indicate the domain of the protein (G-domain in orange, alpha-helical domain in teal and the TGS domain in grey).

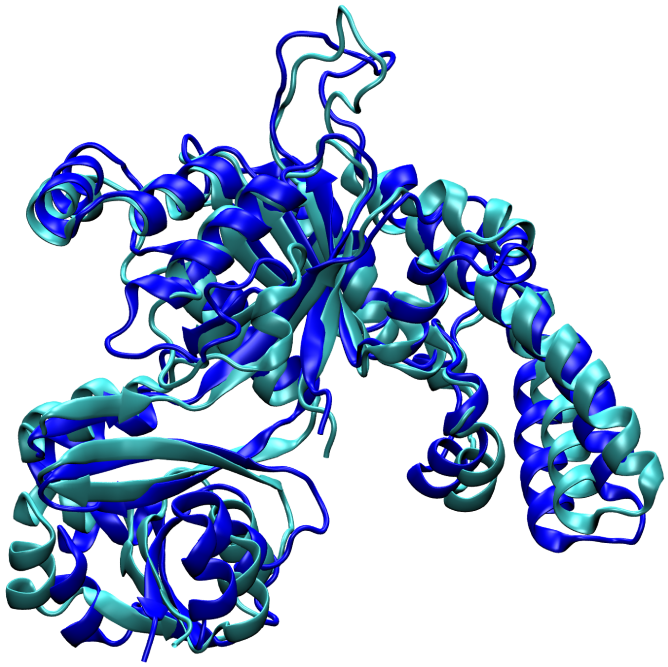


Figure 2.3.4 Structural change of YchF•ADP•Mg^{2+} during 50 ns simulation Superimposition of YchF•ADP•Mg^{2+} snapshots before 50 ns long simulations (cyan) and at 40 ns of simulations (blue).

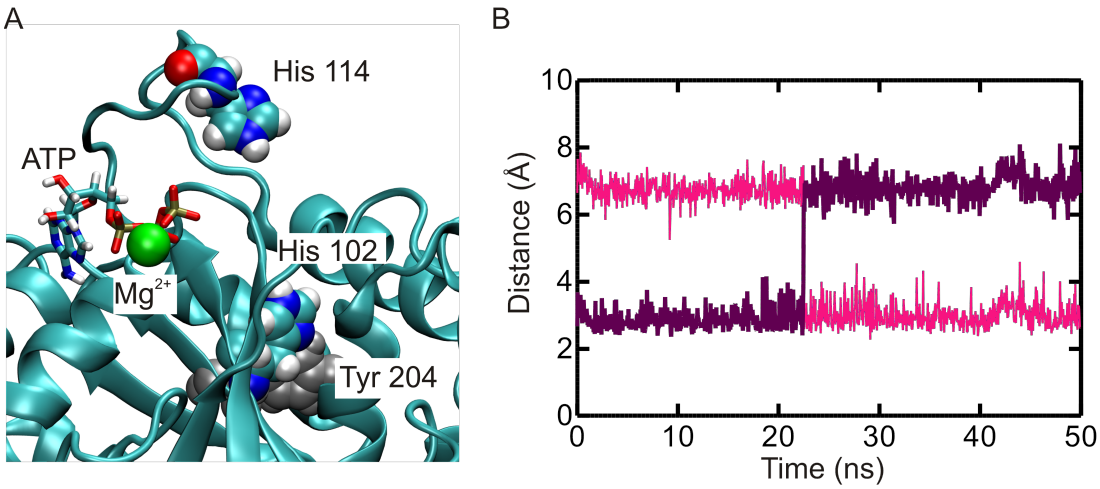


Figure 2.3.5 His 102 forms a stable interaction with Tyr 204 A stable interaction between His 102 and Tyr 204 positions His 201 away from nucleotide binding pocket (A) YchF in complex with ATP after 50 ns of MD simulations. (B) Hydrogen bond distance between Nδ1 of His 102 and Hε2 (pink) or Hε1 (purple) of Tyr 204.

A closer look at His 114 reveals that it is located in an extended loop structure (residues 105 to 121) within the G-domain that shows a greater flexibility when bound to ATP as opposed to ADP and *apo* (Figure 2.3.3 B). The highest level of flexibility is seen in the ATP complex (RMSF of 4.2 Å for His 114) followed by the ADP complex (RMSF of 3.4 Å for His 114) and the *apo* state (RMSF of 3.3 Å for His 114). This high level of conformational flexibility is consistent with the fact that electron density for this loop could only be observed in one of the five experimentally determined structures of YchF that are available in the protein data base (PDB IDs 2DBY, 1JAL, 1NI3, 2DWQ and 2OHF), prompting reference to this loop as the 'flexible' loop.[78,79] A visual comparison of loop conformations before and after 50 ns of simulation revealed that the 'flexible' loop is found in a more open conformation in the YchF•ADP•Mg^{2+} complex and in a closed conformation (closer to the nucleotide binding pocket) in the YchF•ATP•Mg^{2+} complex or *apo* state (Figure 2.3.6). To describe the movement of the loop during the simulation the distance between the

alpha carbons of His 114 and Ser 16 were plotted over the respective simulation time (0 to 50 ns) (Figure 2.3.6. Panel B). Ser 16 was chosen because it is a conserved amino acid in the P-loop responsible for coordinating the magnesium ion and shows no fluctuations in its position over a 50 ns simulation. A similar plot was constructed for each of the remaining histidine residues (His 102, His 145 and His 309). None of these residues show nucleotide-dependent behaviour over the 50 ns as they shown an average distance range between 2-4 Å (Figure 2.3.7) whereas the differences for His 114 in the 'flexible' loop are on the order of 15 Å.

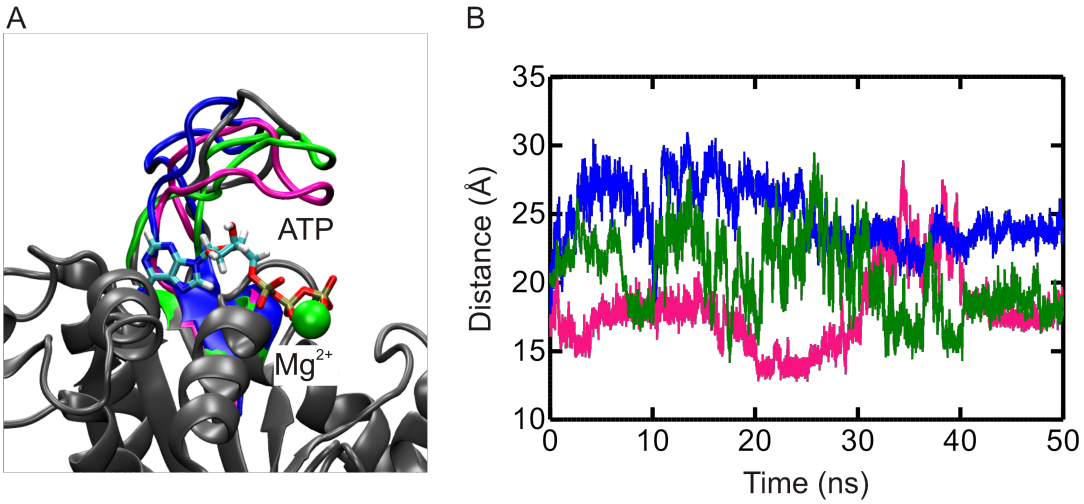


Figure 2.3.6 MD simulations reveal different conformations of the 'flexible' loop
(A) YchF•ATP•Mg^{2+} (pink), YchF•ADP•Mg^{2+} (blue) and YchF *apo* (green) final structures after 50 ns of simulation aligned to pre-simulation structure (grey). (B) Distance between the alpha carbons of His 114 and Ser 16 over a 50 ns simulation of YchF•ATP•Mg^{2+} (pink), YchF•ADP•Mg^{2+} (blue) and *apo* YchF (green).

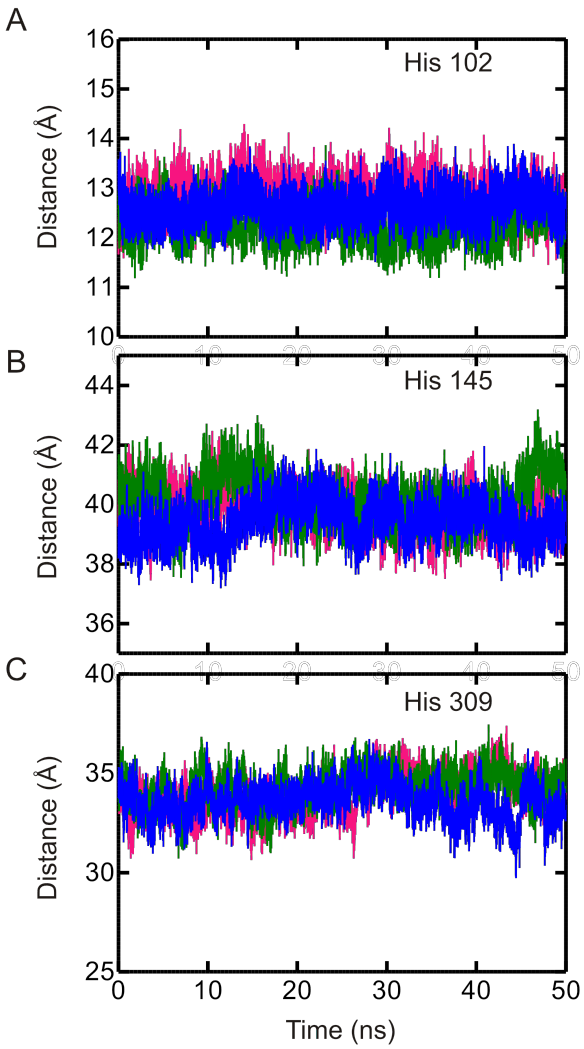


Figure 2.3.7 His 102, His 145 and His 308 do not show nucleotide dependent behaviour Distance between the alpha carbons of His 102 (A), His 145 (B) or His 308 (C) and Ser 16 of the P-loop over a 50 ns MD simulation. YchF•ATP•Mg^{2+} is always shown in pink, YchF•ADP•Mg^{2+} in blue and *apo* YchF in green.

In order to evaluate the distribution of different conformations explored by the 'flexible' loop, the distances between the Cα of His 114 and Ser 16 in each of the simulated annealing simulations and the three 50 ns MD simulations, were calculated and plotted as histograms counting the number of times the Cα-Cα distance fell into a particular distance bin (Figure 2.3.8). Analysis of the three 50 ns MD trajectories (YchF•ATP•Mg^{2+}, YchF•ADP•Mg^{2+} and YchF *apo,* respectively) revealed distinct

distribution profiles indicating the prevalence of certain conformations of the 'flexible' loop as a function of the nucleotide bound state. In the YchF•ATP•Mg^{2+} MD simulation the 'flexible' loop is capable of visiting approximately 5 independent conformations characterized by a Gaussian distribution (see appendix for fitted Gaussian distributions) centered around 13.7 Å, 15.2 Å, 17.6 Å, 22 Å and 26.2 Å. The YchF•ADP•Mg^{2+} trajectory the distributions are significantly different as the 'flexible' loop mainly occupies only three conformations centered around 22.7 A and 23.7 Å and 27 Å. Interestingly, when compared with the YchF•ATP•Mg^{2+} simulation data these three states are also observed, however the conformations described by 13.7 Å and 15.2 Å are essentially inaccessible to the 'flexible' loop of YchF when bound to ADP (Figure 2.3.8). The *apo* state MD simulation of YchF remarkably resembles the five state distribution observed in the YchF•ATP•Mg^{2+} simulation, however, the respective states are less discrete and mainly center around 17 Å but also 20.1 Å and 21.7 Å. This is consistent with rapid fluctuations between states seen in Figure 2.3.6 (B).

These observations clearly support the nucleotide dependent behavior of the 'flexible' loop, suggesting two major conformations, an open conformation of the loop in the ADP-bound form of YchF that converts into a closed conformation in the ATP-bound and *apo* state of YchF. Although the 'flexible'-loop primarily exists in the closed conformation in these forms, it is also capable of visiting one of its two open ADP-conformations closest to the ATP/*apo* conformations, providing a route for interconversion between the configurations enabling for a nucleotide dependent functional cycle of YchF.

In order to confirm the nucleotide-dependent conformations (open and closed) of the 'flexible' loop (residues 105 to 121) *in silico* simulated annealing experiments of the 'flexible loop' were performed. The 10 ns structure of the three homology models (YchF•ATP•Mg^{2+}, YchF•ADP•Mg^{2+} and *apo* YchF) were heated over 300 ps to 1000 K to melt any existing secondary structures that might bias the conformational space accessible to the loop during the 50 ns simulations. Subsequently the 'flexible' loop residues were slowly cooled down to 300 K, allowing the loop to explore possible alternate conformations. This method ensures that the flexible loop from each simulation begins in an unbiased conformation. In a separate set of simulated annealing simulations the switch regions, Switch I (residues 15 to 42) and Switch II (residues 73 to 98), of YchF•ATP•Mg^{2+}, YchF•ADP•Mg^{2+} and *apo* YchF models were also heated and cooled in parallel with the 'flexible' loop to investigate a potential influence on the nucleotide-dependent behavior of the 'flexible' loop, complementing the 50 ns MD simulation data for the YchF•ATP•Mg^{2+}, YchF•ADP•Mg^{2+} and YchF *apo* molecular systems. Interestingly, a similar segregation of the 'flexible' loop conformations into closed (ATP/*apo*) versus open (ADP) conformation was only observed in the simulated annealing simulations when the switch regions were included in the simulation. Surprisingly, these simulations also allowed that YchF•ADP•Mg^{2+} complex to explore the ATP-like conformation (4 of 25 simulations). When the simulated annealing was set up to include only the 'flexible' loop, the nucleotide-dependent conformations are less pronounced with the YchF•ATP•Mg^{2+} simulation resulting in a distribution similar to the YchF•ADP•Mg^{2+}. These observations support the existence of two major nucleotide-dependent states or

ensembles of YchF and are consistent with the role of switch I and II for conformation selection in a nucleotide-dependent manner, similar to other translational GTPases (such as EF-Tu and Ras).[1,4,116]

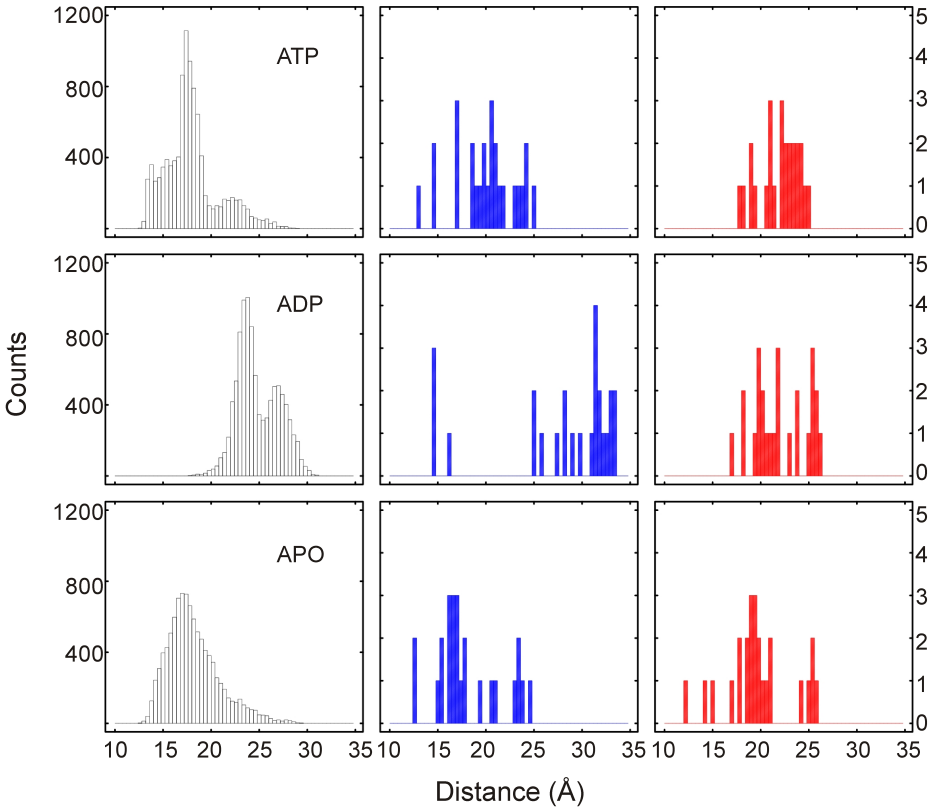


Figure 2.3.8 Nucleotide dependent conformations of the 'flexible' loop in YchF Distribution of Cα-Cα (His 114 to Ser 16) distances plotted as histograms (bin size 0.4 Å). MD simulations (50 ns) of the YchF•ATP•Mg^{2+}, YchF•ADP•Mg^{2+} and YchF *apo* models (left column). Histograms of distance distributions after simulated annealing of the flexible loop (residues 105 to 121), Switch I (residues 25 to 42) and Switch II (residues 75 to 93) are shown in blue (middle column) and of only the 'flexible' loop in red (right column).

2.3.3 Nucleotide Binding Properties of YchF Variants

MD simulations have shown that His 114 is located in a novel nucleotide sensing structural element, the flexible loop, suggesting a role as the catalytic histidine. Therefore, His 114 was substituted with alanine, phenylalanine arginine,

lysine and glutamine (H114A, H114F, H114R, H114K and H114Q) in order to probe whether His 114 is essential for catalysis, and if so what properties of histidine are responsible. Alanine is a nonpolar, aliphatic amino acid which should not be able to support catalysis, whereas phenylalanine is a similar shape to hisitidne but uncharged. Alternatively, both arginine and lysine are basic amino acids and are acceptable substitutions in position 114 in two of the 116 bacterial species examined in this work. As glutamine is the canonical catalytic residue seen in ras-like GTPases it was of interest to see if this amino acid was an acceptable substitution in YchF. However, as His 102 is 76% conserved and appears to be making a stable interaction with Tyr 204 in the MD simulations reported here we also constructed the variants H102A, H102Q and H102K to see if His 102 is important for the structure of YchF. First we confirmed adenine nucleotide binding using steady-state fluorescence titrations, as variants that cannot bind nucleotides are likely folded incorrectly and will be unable to function catalytically. The TGS domain of YchF contains a single tryptophan residue which was successfully used previously to probe nucleotide binding, upon the addition of increasing concentrations of nucleotide.[18] Monitoring the tryptophan fluorescence from 295 to 400 nm, a decrease in fluorescence can be observed with the addition of increasing concentrations of ADP and ATP indicating nucleotide binding (Figure 2.39A and 2.39C). A shoulder around 305 nm appears indicating the seven tyrosines found within YchF are not fully quenched by fluorescence resonance energy transfer from the tyrosine residues to the tryptophan residue.[117] This fluorescence decrease can be plotted as a function of nucleotide concentration and fit with a hyperbolic equation in order to obtain a K_D (Figure 2.3.9B

and 2.3.9D; Table 2.3.1). All variants tested are capable of binding to adenine nucleotides in the low micromolar range (1 to 10 μM), which is comparable to wild type.[18] Interestingly, while $YchF_{WT}$ was reported to show no signal change upon the addition of ATP this was not the case for variants $YchF_{H114A}$, $YchF_{H114F}$, $YchF_{H114Q}$, $YchF_{H114R}$ or $YchF_{H102A.}$ The dissociation constants for $YchF_{H114A}$, $YchF_{H114F}$, $YchF_{H114Q}$, $YchF_{H114R}$, and $YchF_{H102A}$ for ATP are in the low micromolar range, similar to that seen for ADP. $YchF_{H102Q}$ and $YchF_{H102K}$ did not show a signal change for ATP binding, therefore ADPNP was used and dissociation constants in the low micromolar range were observed.

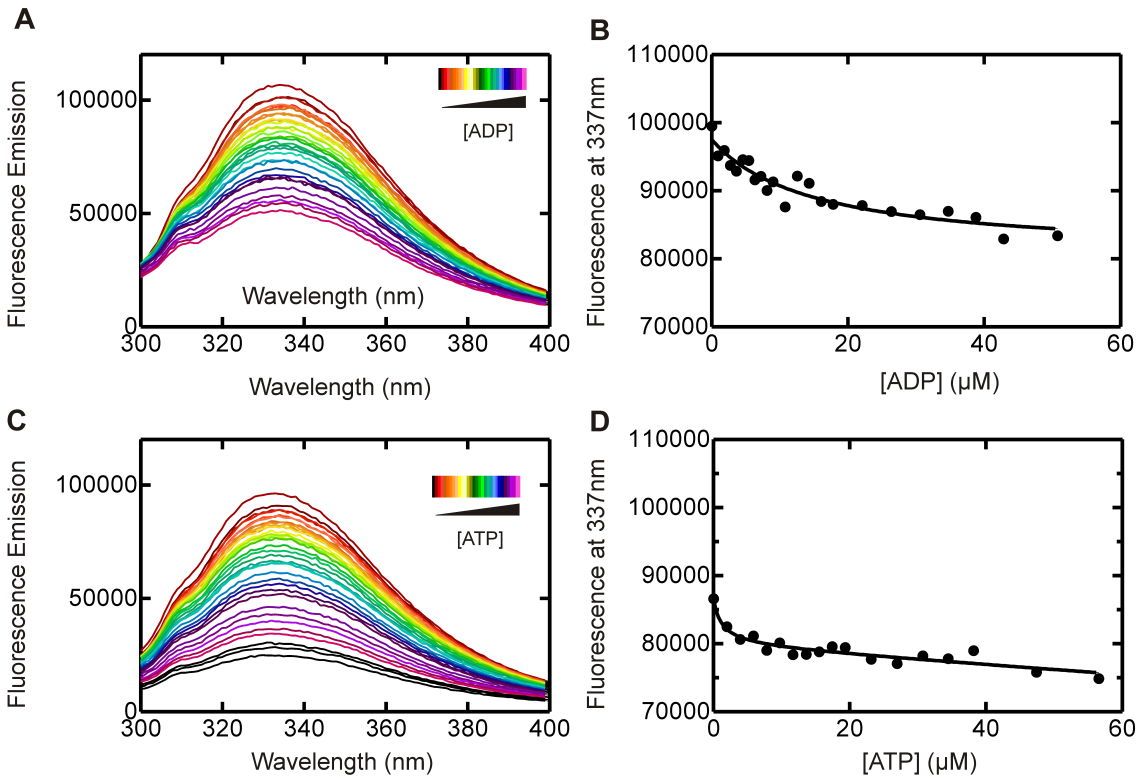


Figure 2.3.9 $YchF_{H114A}$ binds adenine di- and tri-phosphates Equilibrium fluorescence titrations of 1 μM H114A with increasing concentrations of nucleotide. Tyrosine and tryptophan fluorophores of YchF were excited at 280 nm and fluorescence emission spectra were detected from 295 to 400 nm in the presence of increasing concentrations of ADP (A) and ATP (C). Plots of the fluorescence signal at 337 nm against the concentrations of ADP (B) and ATP (D). K_Ds were obtained by fitting the data with a hyperbolic function.

Table 2.3.1 Equilibrium dissociation constants for $YchF_{E.\ coli,WT}$ and variants for adenine nucleotides.

	ADP		ATP		ADPNP	
	K_D [µM]	SD [µM]	K_D [µM]	SD [µM]	K_D [µM]	SD [µM]
WT[25]	14	5	n.d		9	8
H114A	9	4	2.5	2		
H114F	8	5	5	2		
H114R	8	7	3	2		
H114Q	4	4	0.6	0.3		
H114K	10	5	n.d.			
H102A	17	6	2	2		
H102Q	3	2	n.d.		7	10
H102K	1	1	n.d		22	5

2.3.4 Nucleotide Hydrolysis Activity

Next the ability of the YchF variants to hydrolyze ATP was examined as described in section 2.2.6. Five µM protein was incubated with 125 µM [γ-^{32}P]-ATP and samples were removed at varying time points. The amount of $^{32}P_i$ released due to ATP hydrolysis by YchF was determined using thin layer chromatography and plotted as a function of time in order to determine a rate of ATP hydrolysis. The obtained hydrolysis rates reveal that, although $YchF_{H114A}$ is able to bind adenine nucleotides like wild-type, it is not able to hydrolyze ATP (Figure 2.3.10 and Table 2.3.2). Interestingly, among the other H114 variants only H114R is capable of hydrolyzing ATP (0.15 min^{-1}), however with approximately half the rate observed for the wild type (0.28 min^{-1}). Surprisingly, only one of the constructed His 102 variants (H102Q) is capable of hydrolyzing ATP (0.24 min^{-1}) and it does so at a wild-type rate (0.28 min^{-1}) (Table 2.3.2).

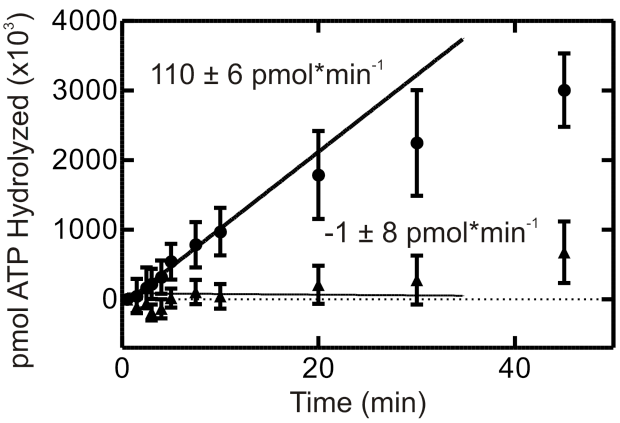


Figure 2.3.10 $YchF_{H114A}$ does not catalyze ATP hydrolysis Time dependence of ATP hydrolysis in the presence of YchF WT (closed circles) and YchF H114A (closed triangles). A linear function was fit to the initial phase of the reaction (first 10 minutes) to give a rate of ATP hydrolysis (pmol / min).

Table 2.3.2 Specific ATPase activity of $YchF_{WT}$ and YchF variants

	Rate (min^{-1})
$YchF_{WT}$	0.28 ± 0.02
$YchF_{H114A}$	0
$YchF_{H114F}$	0
$YchF_{H114R}$	0.15 ± 0.04
$YchF_{H114K}$	0
$YchF_{H114Q}$	0
$YchF_{H102A}$	0
$YchF_{H102Q}$	0.24 ± 0.05
$YchF_{H102K}$	0

It has been shown that YchF is capable of interacting with both polysomes and ribosomal subunits, and that only the 70S ribosome acts as an ATPase activating factor for $YchF_{E.\ coli}$.[18,33,80] GAF's may catalyze nucleotide hydrolysis by providing parts of the necessary catalytic machinery (in trans) to stabilize the developing negative charge on the transition state, directly activate the catalytic water molecule, or alternatively by positioning the enzyme's intrinsic catalytic machinery (Figure 1.1.1).[21] Therefore, we wished to characterize the 70S-stimulated ATPase activity of

YchF for a catalytically inactive and marginally active variants, $YchF_{H114A}$ and $YchF_{H114R}$ respectively, in order to provide insight into what role the 70S ribosome is playing in catalysis. Michaelis-Menten titrations were performed by measuring the initial velocities of ATP hydrolysis at 1 μM YchF, 125 μM [γ-^{32}P]-ATP and varying concentrations of 70S ribosomes (Figure 2.3.11). The resulting catalytic constant (k_{cat}) for $YchF_{H114A}$ on the ribosome is $k_{cat,ATP,70S}$ = 0.7 ± 0.2 min^{-1} and the corresponding Michaelis-Menten constant ($K_{M,70S}$) is 9 ± 7 μM (Figure 2.3.11 and Table 2.3.3). The resulting k_{cat} for $YchF_{H114R}$ on the ribosome is $k_{cat,ATP,70S}$ = 0.3 ± 0.2 min^{-1} and the corresponding Michaelis-Menten constant ($K_{M,70S}$) is 13 ± 13 μM. When we compare the determined k_{cat} values to those previously published for $YchF_{WT}$ ($k_{cat,ATP}$ = 0.36 ± 0.02 min^{-1}, $k_{cat,\ ATP,70S}$ = 3.1 ± 0.2 min^{-1}) we can see that the intrinsically inactive $YchF_{H114A}$ and $YchF_{H114R}$ have $k_{cat,ATP,70S}$ values approximately 10-fold lower than that see for $YchF_{WT}$, and more similar to the unstimulated WT ($k_{cat,ATP}$).[18] The previously determined K_M values for $YchF_{WT}$ ($K_{M,ATP}$ = 41.3 ± 5.8 μM and $K_{M,\ ATP,70s}$ = 7.7 ±1.1 μM) agree well with those determined for the $YchF_{H114A}$ and $YchF_{H114R}$.[18]

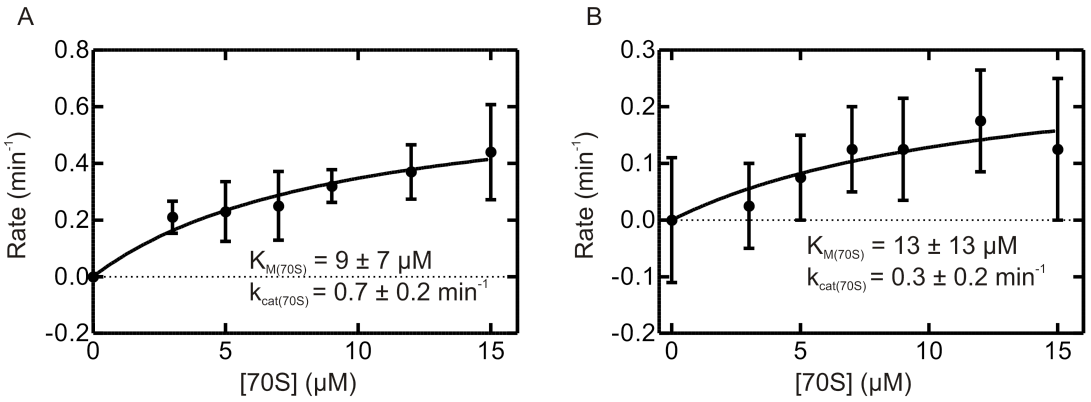


Figure 2.3.11 Ribosome-stimulated ATPase activity of $YchF_{H114A}$ and $YchF_{H114R}$
(A) Michaelis-Menten titration of $YchF_{H114A}$ with 70S ribosomes. (B) Michaelis-Menten titration of $YchF_{H114R}$ with 70S ribosomes. 1 μM YchF and 125μM [γ-^{32}P]-ATP were incubated with increasing concentrations of 70S ribosomes. 70S concentration dependence of the initial rates was fit with the Michaelis-Menten equation to obtain and a v_{max} of 0.7 ± 0.2 min^{-1} and a K_M of 9 ± 7 μM for $YchF_{H114A}$ and a v_{max} of 0.3 ± 0.2 min^{-1} and a K_M of 13 ± 13 μM for $YchF_{H114R}$.

Table 2.3.3 Michaelis-Menten kinetic parameters of $YchF_{E.\ coli,WT}$ and variants.

	v_{max} ($\mu M*min^{-1}$)	**K_M** (μM)	**k_{cat}** (min^{-1})
$YchF_{WT}$•ATP + 70S[18]	3.1 ± 0.2	7.7 ± 1.1	3.1 ± 0.2
$YchF_{H114A}$•ATP + 70S	0.7 ± 0.2	9 ± 7	0.7 ± 0.2
$YchF_{H114R}$•ATP + 70S	0.3 ± 0.2	13 ± 13	0.3 ± 0.2

2.4 Discussion and Future Directions

In order to elucidate the catalytic mechanism of ATP hydrolysis in YchF we have performed ATPase assays over a range of pH values. The activity of YchF shows a clear pH dependence in which one ionizable group with a pKa of approximately 6.7 is responsible for catalysis, suggesting a catalytic histidine residue(s). Sequence alignments reveal that three of the four histidines in $YchF_{E.\ coli}$ are highly conserved, of which two are located in the G-domain. In order to further determine which histidine residue was involved in catalysis we performed an analysis of $YchF_{E.\ coli}$ homology models in complex with ATP, ADP or in the *apo* state.

A plot of the RMSF highlights the fact that the flexibility of His 102 does not vary over a 50 ns simulation with respect to the nucleotide bound. On the other hand, His 114 shows flexibility in the presence of ATP and ADP but less flexibility in the *apo* state. A closer look at His 102 reveals that it is engaged in a stable interaction with Tyr 204, whether the protein is in complex with ATP or ADP or the nucleotide-free state. This interaction is stable throughout the entire 50 ns simulations and essentially locks His 102 in a conformation in which it points away from the nucleotide-binding pocket. The movement of His 102 is also restricted due to a vast number of surrounding hydrophobic interactions (for example Leu 9 of the P-loop with Ile 127), suggesting it would not be able to visit a catalytically active conformation. His 114 is located within an unstructured loop protruding away from the G-domain, whose inherent flexibility is demonstrated by the fact that 4 of the 5 crystal structures of YchF homologues did not provide resolution of this loop. MD simulations have shown that this 'flexible' loop exhibits a nucleotide-dependent behavior, and that can be described by the distance between His 114 in the flexible loop and Ser 16 in the nucleotide binding pocket. The performed MD simulations reveal 5 independent states between the 'flexible' loop and the binding pocket. In the YchF•ATP•Mg^{2+} complex the 'flexible' loop is capable of visiting all of the states. However two of these conformations are essentially unavailable to the loop in the YchF•ADP•Mg^{2+} model, which only visits the two longest distance states and occasionally the mid-state. In the *apo* state on the other hand the 'flexible' loop is also able to visit all states. However it is interesting to note that the loop shows a preference for those 'ATP-like' states. This suggests that the *apo* conformation may be more similar to the ATP conformation, which would explain

why Becker *et al* observed no signal change for ATP binding to YchF when performing steady-state nucleotide titrations monitoring only the intrinsic Trp fluorescence.[18] As a way to confirm that the other histidine residues were not involved in catalysis we have also plotted the distances between them and Ser 16. No nucleotide dependent behavior was observed.

To confirm the nucleotide-dependent states seen for the 'flexible' loop we used *in silico* simulated annealing experiments. These have shown that the 'flexible' loop is only capable of exhibiting a nucleotide-dependent response when the switch regions are included in the annealing procedure. When the switch regions are kept rigid during the simulated annealing, all YchF models show a similar loop conformation, centered on the middle state. Although YchF hydrolyzes ATP more efficiently than GTP it was originally classified as a GTPase and might have retained certain functional aspects associated with GTPases. GTPases function as molecular switches in which the two switch regions show nucleotide-dependent conformations. No such behavior has been reported for YchF thus far. However, it could be that the nucleotide-dependent 'flexible' loop conformations of YchF are correlated to the changes seen in the Switch regions of other GTPases. The nucleotide dependence of the 'flexible' loop suggests that His 114 is required for intrinsic catalysis in YchF. However, His 114 is still on average 20 Å away from the nucleotide-binding pocket, indicating that this model is not in its active conformation. It is easy to envision a movement of the 'flexible' loop inwards to a catalytically competent state. In order to assess if the 'flexible' loop would be capable of moving into an active conformation we attached a spring to the delta nitrogen of His 114 and effectively pulled it to the

analogous position occupied by the epsilon oxygen of Glu 282 in MnmE (Figure 2.4.1). Backbone phi and psi angles of the flexible loop remained favorable despite the strained conformation, and a conserved interaction between Phe 107 and the adenine base was not affected, supporting the hypothesis that His114 can reach a position during YchF's catalytic cycle that could support catalysis.

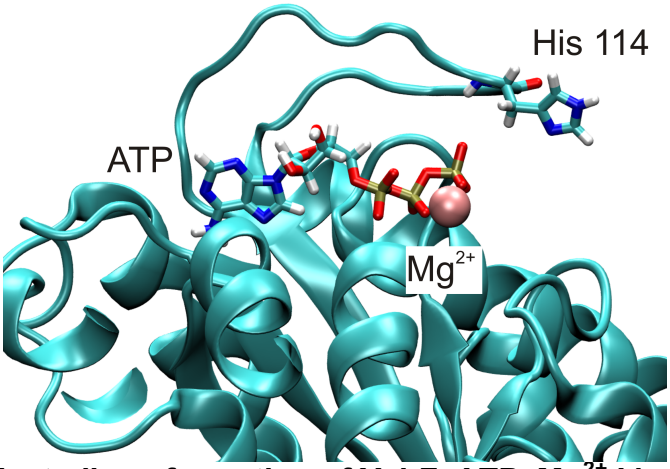


Figure 2.4.1 'Activated' conformation of YchF•ATP•Mg^{2+} Model of $YchF_{E.\ coli}$ in which His 114 has been pulled by a spring to within 7 Å of the γ-phosphate into a position analogous to that seen by Glu 282 in MnmE.

In order to confirm that His 114 is indeed responsible for catalysis, YchF H114 variants were designed and characterized with respect to their nucleotide binding properties and their ATP-hydrolysis activity. As His 102 is 76% conserved we also wished to observe the effect of amino acid substitutions in position 102. All His 114 and His 102 variants are capable of binding adenine nucleotides with similar affinities. Interestingly, and in contrast to the wild type enzyme, equilibrium ATP dissociation constants for the YchF variants H114A, H114F, H114R, H114Q and H102A could be determined using intrinsic tryptophan fluorescence.[18] This might be explained by the presence of two different conformations of YchF in solution, as it has been observed

that Ras in complex with ATP has two different conformations, one of which resembles the *apo* state of the enzyme.[118-121] If, like Ras, YchF has two ATP-bound conformations and the conformation that resembles the *apo* state is more prevalent, we would not expect a signal change upon nucleotide binding. As a signal change was observed for the variants, it is possible that these single amino acid substitutions have resulted in a different ATP-bound conformation or stabilized YchF in the conformation that does not resemble the *apo* state. Although there are possible changes in conformation, all variants were able to bind adenine nucleotides in the low micromolar range. Despite their abilities to bind adenine nucleotides all H114 variants, except $YchF_{H114R}$, are catalytically inactive with respect to intrinsic ATP hydrolysis activity. These results suggest that His 114 might either act as a general base, capable of either abstracting a proton from the catalytic water molecule to generate a better nucleophile, or positioning the attacking water molecule through hydrogen bonding similarly to the suggested role of Gln in Ras-like enzymes.[21] Had $YchF_{H114F}$ been able to hydrolyze ATP it could have hypothesized that the shape of His 114 (planar and aromatic) was responsible for catalysis, but this was not the case. Alternatively, replacing histidine with the canonical catalytic residue, glutamine, did not rescue catalysis. As glutamine is uncharged compared to the positively charged histidine and longer, either of these factors could be playing a role. In the 116 bacterial species examined, only two species showed a substitution of histidine 114, either lysine or arginine. An arginine residue can partially rescue ATPase activity, however it is not ideal as $YchF_{H114R}$ hydrolyzed ATP at rate that was two-fold slower than that of wild type YchF. The lysine variant was not active at all, suggesting

it is not just a charge that is important for catalysis. A lysine residue is found in position 114 of YchF from *Wigglesworthia glossinidia*, however the 'flexible' loop of this species is not well conserved. For example, positions 109, 111, 112 and 115 are all over 80% conserved amongst the 116 bacterial species observed but none of the conserved residues are present in *W. glossinidia,* suggesting that this loop may have a different conformation, allowing Lys 114 to perform a catalytic role.

Recently it was shown that the 70S ribosome is capable as acting as an ATPase activating factor for YchF, raising the question of whether the 70S ribosome simply stabilizes an activated conformation of YchF bound to ATP thereby facilitating the observed enhancement in hydrolysis rate, or whether the ribosome contributes components (e.g. protein or RNA) to the catalytic machinery required for the observed 10-fold stimulation. Based on the distance of His 114 from the active site in the crystal structures we had predicted that the 70S ribosome was enhancing hydrolysis by 'pushing' the loop into an active conformation. However, the ability of the 70S ribosome to stimulate hydrolysis in the catalytically inactive $YchF_{H114A}$ suggests that the ribosome might be donating parts to the actual catalytic machinery, as well as likely helping position the 'flexible' loop. The k_{cat} of the stimulation for both $YchF_{H114A}$ and $YchF_{H114R}$ is about 10-fold less than that seen for 70S stimulated $YchF_{WT}$ activity. Since the substitution of histidine with arginine results in only a 2-fold decrease in the intrinsic ATPase activity, but a 10-fold decrease in the 70S stimulated k_{cat}, it appears that maximal stimulation by the 70S ribosome may be dependent on the presence of a His in position 114, which cannot be compensated by an alternative residue (arginine) in that position.

The substitution variants of His 102 reported here are able to bind adenine nucleotides. Surprisingly, only $YchF_{H102Q}$ was able to hydrolyze ATP efficiently. This is consistent with the observations that Histidine 102, although 76% conserved, is replaced with glutamine in 13% of the aligned sequences (Figure A.1 and Table A.1). The lack of hydrolysis from H102A and H102K however was surprising in that MD simulations suggested this residue would have no part in ATP hydrolysis. Interestingly, during the steady-state nucleotide binding experiments the starting fluorescence, and subsequent amplitude change, is significantly lower (50% in some cases) for the H102K and H102A variants despite using the same protein concentrations and an identical experimental set up as the His 114 variants. This suggests that either the environment of the single Trp is different or that the FRET efficiency from the seven tyrosine residues in YchF that likely contribute to the observed change in tryptophan fluorescence upon nucleotide binding is changed, suggesting a different overall conformation of the YchF variants. As the tryptophan residue is located in the TGS domain, over 30 Å from the bound nucleotide, it seems more likely that the observed changes are in the environment of Tyr 218 and Tyr 204, which are located in the G domain. A structural role for His 102 is supported by the fact that the hydrogen bond between His 102 and Tyr 204 is stable in all nucleotide conformations of YchF. Tyr 204 is 97% conserved in the species examined (with three hydrophobic substitutions), supporting the conservation and importance of this interaction.

The catalytic mechanisms of only a few HAS-GTPases have been proposed to date, including MnmE, YqeH, RbgA and FeoB. Interestingly, with the exception of

FeoB, which does not appear to use a catalytic side chain to align an attacking water molecule, each GTPase provides a catalytic residue from helix α2, or in the case of RbgA in a flexible linker region, which is located in a position analogous to helix α2. Similar to RbgA, YchF has helix α2 relocated closer to the TGS domain and helix α4 is located in this position instead. In this respect YchF shares similarities only with RbgA with respect to its catalytic mechanisms when compared to the other HAS GTPases, as YchF and RbgA are the only proteins that use a catalytic histidine residue situated in a flexible loop located in proximity to the nucleotide binding pocket (Figure 2.4.2). It is plausible that upon activation, due to binding of ATP or the 70S ribosome, the flexible loop of YchF undergoes a conformational change leading to the histidine residue being positioned analogously to either the Gln^{cat} or Glu^{cat} within the active site. The 70S ribosome likely helps push this flexible loop into a catalytic position, while also donating catalytic machinery. This mechanism suggests that YchF is most similar to RbgA in its mode of intrinsic catalysis, since this protein also has an essential catalytic histidine residue presented from a flexible linker region of the protein.[69]

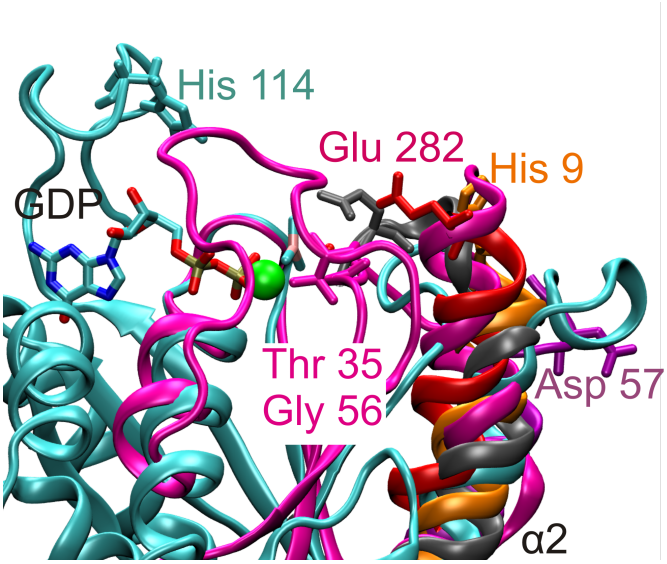


Figure 2.4.2 Alignment of HAS-GTPases of known catalytic mechanisms and YchF Alignment of MnmE (red), FeoB (pink), RbgA (orange), YqeH (purple) and Ras (grey) to YchF (cyan) based upon P-loop structure. Proposed catalytic residues are labeled.

2.5 Conclusion

Using classical biochemical techniques and MD simulations it has been shown that His 114 is required for intrinsic ATPase activity and for maximal 70S ribosome stimulated activity. This residue is located in a highly flexible loop located over 10 Å away from the nucleotide-binding pocket when in its inactive conformation. It is possible that binding to ATP slowly induces a conformational change to the active state, and that this process is accelerated by interactions made with the 70S ribosome, which in turn also provides catalytic machinery in trans. Overall it has been shown that YchF uses an intrinsic catalytic mechanism similar to that observed in RbgA, rather than Mnme, YqeH or FeoB, but may present a novel mechanism of GAP stimulated hydrolysis as the 70S ribosome appears to be involved in stabilizing the catalytic residue while also contributing directly to catalysis.

Chapter 3. Pre-Steady State Kinetics of Nucleotide Binding and the Effect of Mg^{2+}

3.1 Introduction

GTPases are regulatory proteins whose function as a molecular switch is defined by the conversion between a GTP-bound on-state and a GDP-bound off-state facilitated by nucleotide hydrolysis.[1,4,21] As such, G-proteins are responsible for regulating many cellular processes such as protein synthesis, nuclear export and extracellular signaling.[1] Like many enzymes, GTPases need to be strictly regulated and this can be achieved by different mechanisms. Guanine nucleotide Activating Proteins or Factors (GAPs or GEFs) are capable of enhancing the rate of hydrolysis thereby speeding up inactivation of the protein.[5,21] GTPases can be reactivated with the help of a Guanine nucleotide Exchange Factor (GEF), which will facilitate the exchange of GDP for GTP, or GTPases can be stabilized in its inactive state by a Guanine nucleotide Dissociation Inhibitor (GDI).[5,6]

In order to bind and hydrolyze guanine nucleotides, all G-proteins have a conserved G-domain, which contains 5 distinct G motifs (G1 – G5).[4] The G1 motif (also known as the Walker A motif or P-loop) directly interacts with the phosphate moiety, whereas the G2 motif or Switch I region coordinates a bound Mg^{2+} ion (using the hydroxyl group from a conserved Ser or Thr residue).[4,21] The G3 motif (also known as the Walker B motif or Switch II region) also coordinates the Mg^{2+} ion (using a conserved Thr residue) and the γ-phosphate, whereas the G4 and G5 motifs are responsible for guanine nucleotide specificity.[4,21] The conformation of a GTPase will switch depending on which nucleotide is bound and these conformational changes

stem from the Switch I and II regions but can be propagated to other domains of the G-protein.[21,122]

Knowing the rate at which G-proteins are able to associate with and dissociate from guanine nucleotides is important to understanding the functional cycle of the protein. For example, some G-proteins, like Ras and EF-Tu, have very high affinities for guanine nucleotides but have very slow rates of GDP dissociation, and therefore GEFs are required to activate these proteins.[123-125] Alternatively, GTPases like EF-G do not require a GEF as they have fast rates of GDP release and the higher cellular levels of GTP over GDP ensure protein activation.[122,126]

Only eight GTPases can be found in all domains of life: EF-Tu, EF-G, IF-2, Ffh, FtsY, YihA, HflX and YchF (*E. coli* naming).[8] Despite their universal conservation, YihA, HflX and YchF have been relatively unstudied. YchF adopts a three-domain structure with a central G-domain flanked by an alpha helical domain and a C-terminal TGS domain (Threonyl-tRNA synthetase, GTPase and SpoT).[78] Within the G-domain YchF has an altered G4 motif (NxxE instead of N/T KxD) which has resulted in YchF binding and hydrolyzing ATP more efficiently than GTP.[18,33,79] YchF has also been classified as a HAS-GTPase owing to the replacement of the canonical Gln^{cat} with a leucine residue in the G3 motif.[39] Previously in this work I have shown that His 114 is required for intrinsic ATP hydrolysis in YchF, explaining how YchF is able to function (hydrolyze ATP) in the absence of the canonical Gln^{cat} (Chapter 2). Recently, Becker *et al.* reported nucleotide equilibrium dissociation constants (K_D) for $YchF_{E.\ coli}$, however a detailed kinetic mechanism for adenine nucleotide binding and dissociation has yet to be determined.[18] Despite its

implications in numerous cellular functions including protein synthesis, iron utilization in *B. meletensis* and *V. vulnificus*, oxidative stress response, virulence in *S. neumoniae* and *V. vulnificus*, and infection defense in rice, the exact cellular function of YchF is poorly understood. [18-20,78,80,86,100-103] In order to better understand the functional cycle of YchF within the cell I have determined rate constants describing nucleotide association and dissociation in YchF using fluorescence based rapid kinetics. These rates have been determined in the presence and absence of Mg^{2+} in order to further understand the mode of nucleotide binding by YchF.

3.2 Materials and Methods

DH5α cells were purchased from New England Biolabs and BL21-DE3 competent cells were purchased from Novagen. PCR primers were purchased from Invitrogen and Integrated DNA technologies. Restriction enzymes were obtained from Fermentas. Radiochemicals were purchased from Perkin-Elmer. Nucleotides and fluorescent nucleotide analogs were purchased from Invitrogen. All buffers were filtered through 0.45 μM Whatman nitrocellulose membranes.

3.2.1 Sequence Alignments

Refer to section 2.2.1

3.2.2 Cloning and Site-directed Mutagenesis

Refer to section 2.2.2.

3.2.3 Protein Expression and Purification

Refer to section 2.2.3.

3.2.4 Fluorescence Measurements

Refer to section 2.2.4.

3.2.5 Stopped-Flow Kinetics

All fluorescence stopped-flow measurements were performed using a KinTek SF-2004 stopped-flow apparatus (KinTek, Austin, TX) at 20 °C. Mant-nucleotides were excited via a Fluorescence/Förster Resonance Energy Transfer (FRET) from the single tryptophan (λ_{ex} = 280 nm) present in $YchF_{E.\ coli}$ and measured after passing through LG-400-F cut off filters (NewPort Filters, Irvine, CA).

For dissociation experiments 20 µM mant-nucleotide and 2 µM YchF were pre-incubated in $TAKM_7$ for 30 min at 37 °C. Pyruvate kinase (PK, 0.25 µg/uL final; Roche Applied Science) and phosphoenolpyruvate (PEP; 3 mM final) were also added to the mixture when mant-ATP was used. To observe dissociation, 25 µL of YchF•mant-ATP/mant-ADP (1 µM after mixing) was rapidly mixed with 25 µL ATP/ADP (100 µM after mixing). Due to the excess of unlabeled nucleotide present, only the dissociation of the mant-labeled nucleotide contributed to the observed fluorescence change and rebinding of mant-nucleotide is negligible. The observed time courses were fit with a one- or two-exponential binding equation (equations 2 and 3, respectively), where F is the mant fluorescence at time t, F_0 is the starting fluorescence signal and A is the amplitude change and the apparent rate corresponds to the nucleotide dissociation constant (k_{app} or k_{-1})

Calculations were performed using TableCurve (Jandel Scientific) and Prism (GraphPad Software).

$$F = F_O + Ae^{(k_{app} \times t)} \quad (3)$$

$$F = F_O + A_1 e^{(k_{app1} \times t)} + A_2 e^{(k_{app2} \times t)} \quad (4)$$

For association experiments 25 µL of 2 µM YchF in $TAKM_7$ was rapidly mixed with 25 µL of increasing concentrations of mant-nucleotide (1 to 10 µM mant-ADP and 4 to 40 µM mant-ATP). As in the nucleotide dissociation experiments, PEP and PK were added to mant-ATP solutions and the mixtures were pre-incubated for 30 min at 37 °C. Fluorescence data was fit with a one-exponential binding function in the case of mant-ADP (equation 2) and a two- or three-exponential binding function (equations 3 and 4 respectively) in the case of mant-ATP to obtain apparent rates of bimolecular association (k_{app}). The bimolecular association rate constant, k_1, was determined from the slope of the linear concentration dependence of k_{app} on the mant-ADP/mant-ATP concentration.

$$F = F_O + A_1 e^{(k_{app1} \times t)} + A_2 e^{(k_{app2} \times t)} + A_3 e^{(k_{app3} \times t)} \quad (5)$$

$$K_D = k_{-1} / k_1 \quad (6)$$

$$K_D = k_{-1} k_{-2} / k_1 \quad (7)$$

$$K_D = k_{-1} k_{-2} k_{-3} / k_1 k_2 \quad (8)$$

3.2.6 MD Simulations

The initial model for $YchF_{E.\ coli}$ was obtained by constructing a homology model using the Swiss-Model server and the crystal structure of *H. influenzae* as a template (PDB ID 1JAL).[106] The conformation of Switch I (residues 28 to 41, which were missing in the 1JAL structure) was modeled to be identical to Switch I of the YchF structure from *T. thermophilus* bound to GDP (PDB ID 2DBY). In order to model the adenine nucleotide in the $YchF_{E.\ coli}$ model we used the conformation of AMPPCP bound to the human homolog of YchF, hOla1 (PDB ID 2OHF). Transformation of the AMPPCP to ATP and ADP was done manually. In order to position the magnesium

ion associated to the bound nucleotide we aligned the nucleotide with either EF-Tu bound to GDP or GDPNP (PDB ID 1EFC and 1EFT respectively). Hydrogen atoms were added to all models using psfgen in the NAMD software package, and histidine side chains were protonated at the ε-nitrogen only.[93] Initial models were minimized in a vacuum and then placed in a water box extending at least 10 Å from the protein in all directions. Water molecules present in the respective crystal structures were included in this box, and all other waters were added at random using the SOLVATE package in NAMD.[93] Relaxation of the solvated system was achieved by minimizing the positions of water molecules with no constraints, followed by minimization of the protein and ligand atoms in two iterative rounds (10,000 steps each). Sodium ions were then added in random positions by the AUTOIONIZE package in VMD to neutralize the total charge of the system, followed by a final minimization of all components of the system. YchF•ATP•Mg^{2+}, YchF•ADP•Mg^{2+}, YchF•ATP and YchF•ADP had total charges of -17, -16, -20 and -19 and therefore 17, 16, 20 and 19 sodium ions were added respectively. The system was considered to be minimized when no change in energy was observed for at least 1,000 steps.[110] Minimizations, and subsequent equilibrations and MD simulations, were performed with periodic boundary conditions using the NAMD software package.[93] Minimized models were initially equilibrated at 300 K and 350 K for 150 ps at constant pressure (1 atm). Production phase simulations were started using velocities from the 300 K equilibration and coordinates from the 350 K equilibration. Simulations were performed over 50 ns at 300 K with a step size of 0.5 fs using the CHARMM22 parameters for proteins and CHARMM27 parameters for nucleic acids as

implemented in the NAMD package.[93,94] MD simulations were performed in an NPT ensemble in which the pressure was maintained at one atmosphere with a Nosé-Hoover Langevin piston and the temperature was controlled using Langevin dynamics. All simulations were performed in the NAMD software package, and visualization was carried out in VMD.[93,110]

Snapshots of each MD simulation were saved every 0.5 ps, and trajectories were fitted with the software Carma in order to remove any rotations of the protein complex or translation of the center of mass.[111] The root mean-square deviation (RMSD) and root mean-square fluctuations (RMSF) were performed using in-house written scripts invoked with the VMD software package. [110]

3.3 Results

3.3.1 Pre-Steady State Nucleotide Binding to $YchF_{WT}$

Previously, dissociation constants for adenine and guanine nucleotides to YchF were determined using equilibrium methods.[18] However, unlike steady-state nucleotide binding, rapid kinetics techniques can provide information on the rates of association and dissociation. Using a stopped-flow device the rates of association and dissociation of mant-ATP and mant-ADP to $YchF_{WT}$ and YchF variants were determined. For dissociation experiments a YchF•mant•nt complex (1 μM final) was rapidly mixed with excess unlabelled nucleotide (100 μM final) and FRET between the tryptophan residue and the mant group was monitored. Dissociation of the mant-nucleotide resulted in an exponential decrease in mant fluorescence that could be fit with either a single or double exponential function (Equation 3 or 4), yielding the rate constant of nucleotide dissociation (k_{-1}). Nucleotide association experiments were

performed by rapidly mixing YchF (1 μM final) with increasing concentrations of mant-nucleotide resulting in an exponential increase in mant fluorescence that could be fit with an exponential function (Equation 3, 4 or 5). The calculated k_{app} was plotted as a function of the nucleotide concentration and fit with a linear equation to obtain the rate constant of nucleotide association (k_1). Nucleotide dissociation constants (K_D) were determined using Equations 6 to 8. Rate constants were determined for $YchF_{WT}$ (Figure 3.3.1 and Table 3.3.1) and $YchF_{H114A}$ (Figure 3.3.2 and Table 3.3.1). ADP dissociation and association rate constants were also determined for YchF variants $YchF_{H114Q}$, $YchF_{H114K}$ and $YchF_{H114R}$ (Table 3.3.1).

The dissociation of mant-ADP from YchF was fit with a one-exponential function with a k_{-1} of 30 ± 2 s^{-1}. Association traces were also fit with a one-exponential function and the rate constant of association was determined to be 3.3 ± 0.5 $\mu M^{-1}s^{-1}$. The two experimental rates give a dissociation constant of 9 ± 1 μM (using equation 6), which is comparable to that reported previously by Becker *et al* (4 ± 2 μM).[18] The dissociation constant can also be determined by plotting the amplitude change of nucleotide association against the concentration of nucleotide (Figure 3.3.3) and fit with a hyperbolic function (equation 1). This method gives a K_D for mant-ADP of 10 ± 5 μM.

The dissociation of mant-ATP from YchF was fit with a two-exponential function yielding two rate constants: k_{-1} = 0.33 ± 0.05 s^{-1} and k_{-2} = 5 ± 2 s^{-1} (Figure 3.3.1). The association of mant-ATP to YchF was fit with a three exponential function, in which only one rate constant was concentration dependent, which we have taken to be the apparent rate of binding, as a linear concentration dependence is expected

for a bimolecular binding reaction (Figure 3.3.1). The two remaining rate constants observed were not concentration dependent and could be a result of i) a mixed population of mant-ATP and mant-ADP both of which are binding to YchF ii) a mixed population of different conformations of YchF in solution or iii) YchF is undergoing two conformational changes as a result of nucleotide binding. It is possible to rule out a mixture of mant-ADP and mant-ATP as neither observed dissociation rate constant (0.33 ± 0.05 s^{-1} and 5 ± 2 s^{-1}) is consistent with the observed dissociation rate constant of mant-ADP, 30 ± 2 s^{-1}, not to mention the use of PEP and PK in order to convert any ADP back to ATP. The remaining alternatives are hard to differentiate between, however it seems more likely that binding to mant-ATP induces conformational changes within the protein. It has been shown for Ras and Rap that binding to GTP induces state 1 (considered the 'off' state where Thr 35 is not coordinating the Mg^{2+} ion), which must be converted state 2, the 'on' state (Thr 35 in contact with Mg^{2+}) and this occurs on the millisecond time scale.[127-129] Therefore we can envision a model in which binding of mant-ATP to YchF induces two unknown conformational changes within the protein (Figure 3.3.9). The apparent rate constants associated with the bimolecular binding event were used to determine a k_1 of 0.59 ± 0.04 $\mu M^{-1}s^{-1}$. The rate constants, k_2 and k_3, were determined by taking the average of the observed rates for each concentration of mant-ATP used. The dissociation constant was determined using equation 8, yielding a K_D of 19 ± 11. Using the amplitude plot a K_D of 4 ± 1 μM was determined for mant-ATP, which is consistent with that determined using the rate constants of association and dissociation.

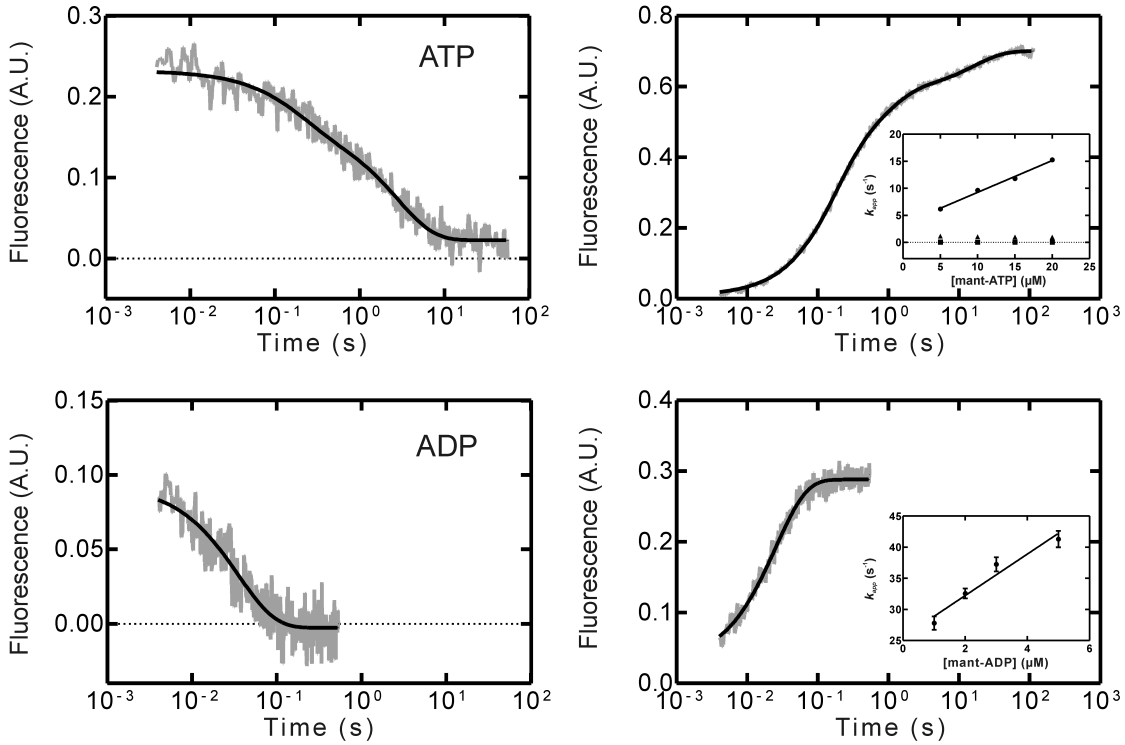


Figure 3.3.1 Pre-steady state kinetics of adenine nucleotide binding and dissociation for YchF$_{WT}$ Representative time courses of the dissociation of a YchF mant-nucleotide complex (1 μM) in the presence of excess unlabelled nucleotide (100 μM) are shown in the left panels. Representative time courses of mant-nucleotide (5 μM) association to YchF (1 μM) are shown in the right panels with the concentration dependence of k_{app} on mant-nucleotide association to YchF shown in the insert. The k_{app} values were calculated by single-exponential or three-exponential fitting of the time courses. For all time courses the fluorescence of the mant group was monitored.

3.3.2 Pre-Steady State Nucleotide Binding of YchF variants

Dissociation of mant-ADP from the catalytically inactive variant YchF$_{H114A}$ was fit to a one-exponential function with a k_{-1} of 58 ± 9 s^{-1}, which is 2-fold higher than wild type. The association of mant-ADP to YchF$_{H114A}$ was fit with a one-exponential equation yielding a k_1 rate constant of 8 ± 0.4 μM^{-1}s^{-1}, which is also 2-fold higher than wild type. Therefore using equation 6, a dissociation constant, K_D, of mant-ADP from YchF$_{H114A}$ was determined to be 7 ± 1 μM, therefore both YchF$_{WT}$ and YchF$_{H114A}$ bind

mant-ADP with the same affinity. The dissociation constant, determined using the amplitude plot is 13 ± 4 μM.

The dissociation of mant-ATP from $YchF_{H114A}$ varied from wild type in that it was fit with a one-exponential function to give a k_{-1} of 14 ± 3 s^{-1}, a rate constant which is 2.5-fold greater than wild type k_{-2} (rate constant of mant-ATP dissociation in $YchF_{WT}$). The association of mant-ATP to $YchF_{H114A}$ was fit to a two-exponential function in which only one apparent rate constant was concentration dependent and showed a k_1 of 0.9 ± 0.2 $\mu M^{-1}s^{-1}$, which is approximately 2-fold higher than wild type. The rate constant associated with a conformational chanage was determined to be 0.05 ± 0.02 s^{-1}. The change from a triple exponential function used to fit mant-ATP binding to $YchF_{WT}$ to a double exponential function to fit mant-ATP binding to $YchF_{H114A}$ supports that the substitution of His 114 with Ala has resulted in the loss of one conformational change within YchF. Nonetheless the dissociation constant of mant-ATP for $YchF_{H114A}$ determined using equation 8 is 311 ± 157 μM, which is significantly higher that that determined for wild type 19 ± 11 μM, due to the much slower rate constant of the conformational change. The dissociation constant determined using the amplitude plot was 4 ± 2 μM, which does not take into account the rate of conformational changes.

The kinetics of mant-ADP binding to $YchF_{H114Q}$, $YchF_{H114R}$ and $YchF_{H114K}$ was also determined. All traces were fit with a one-exponential function. Mant-ADP dissociation rates from $YchF_{H114Q}$, $YchF_{H114R}$ and $YchF_{H114K}$ showed similar rate constants and were 62 ± 11 s^{-1}, 65 ± 10 s^{-1} and 59 ± 10 s^{-1} respectively. Mant-ATP association rate constants for $YchF_{H114Q}$, $YchF_{H114R}$ and $YchF_{H114K}$ were 8 ± 1 $\mu M^{-1}s^{-1}$,

3.5 ± 0.3 $\mu M^{-1}s^{-1}$ and 10 ± 1 $\mu M^{-1}s^{-1}$, yielding dissociation constants of 8 ± 2 μM, 19 ± 3 μM and 6 ± 1 μM respectively using equation 5, whereas the amplitude plots yielded dissociation constants of 4 ± 1 μM, 4 ± 1 μM and 2.2 ± 0.4 μM.

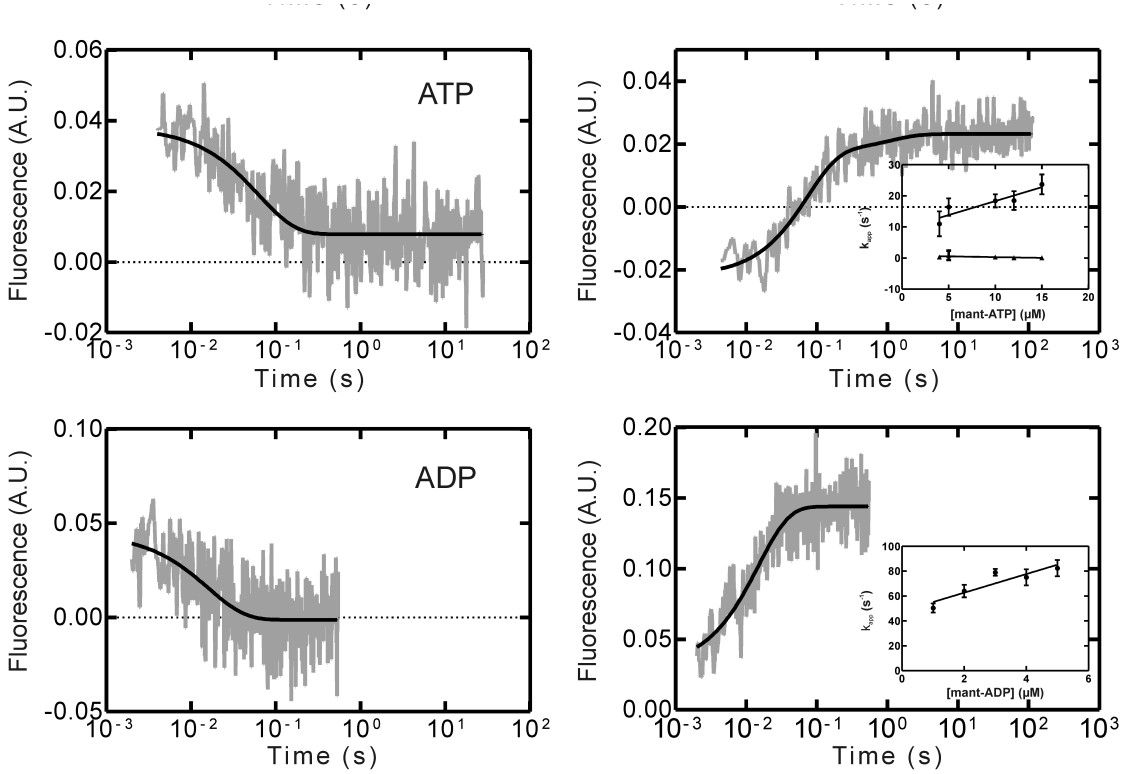


Figure 3.3.2 Pre-steady state kinetics of adenine nucleotide binding and dissociation for $YchF_{H114A}$ Representative time courses of the dissociation of a $YchF_{H114A}$ mant-nucleotide complex (1 μM) in the presence of excess unlabelled nucleotide (100 μM) are shown in the left panels. Representative time courses of mant-nucleotide (5 μM) association to $YchF_{H114A}$ (1 μM) are shown in the right panels with the concentration dependence of k_{app} on mant-nucleotide association to $YchF_{H114A}$ shown in the insert. The k_{app} values were calculated by single-exponential or two-exponential fitting of the time courses. The fluorescence of the mant group was monitored in all time courses.

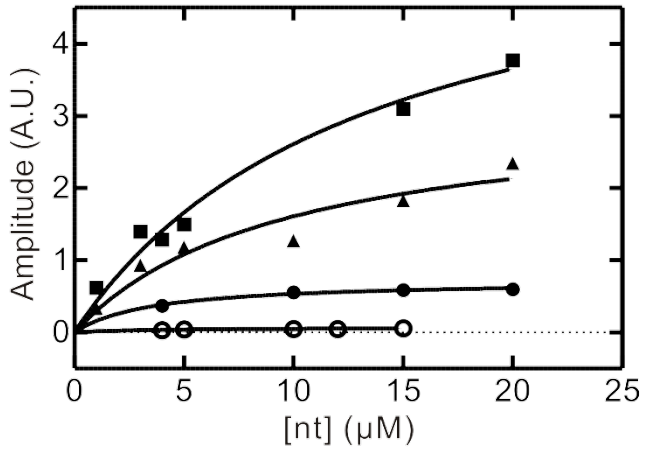


Figure 3.3.3 K_D determination using amplitude plot. Amplitude of the signal change was plotted as a function of nucleotide concentration. Traces describe mant-ATP association to $YchF_{WT}$ (filled circles), $YchF_{H114A}$ (open circles) and mant-ADP association to $YchF_{WT}$ (filled triangles), $YchF_{H114A}$ (filled squares).

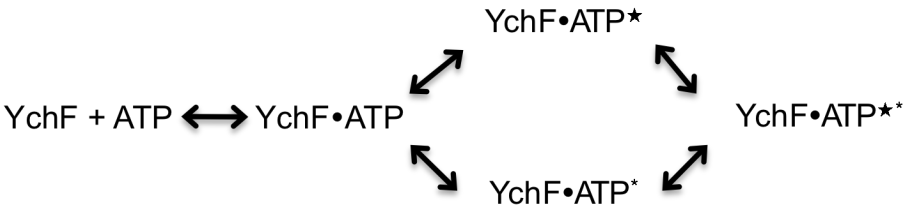


Figure 3.3.4. Proposed mechanism of ATP binding to $YchF_{WT}$ Binding of ATP to YchF induces two conformational changes (indicated with a * and ★) within the protein.

Table 3.3.1 Summary of experimentally determined rate constants of YchF and YchF variants with adenine nucleotides

	Rate Constant	Wild Type	H114A	H114Q	H114K	H114R
mant-ADP	k_1 ($\mu M^{-1}s^{-1}$)	3.3 ± 0.5	8.0 ± 0.4	7.5 ± 0.9	3.5 ± 0.3	10 ± 1
	k_{-1} (s^{-1})	30 ± 2	58 ± 9	62 ± 11	65 ± 10	59 ± 10
	k_{-1} (s^{-1}) (k_1 plot)	26 ± 2	45 ± 4	39 ± 3	53 ± 3	44 ± 4
	K_D (μM)	9 ± 1	7 ± 1	8 ± 2	19 ± 3	6 ± 1
	K_D (μM) (k_1 plot)	8 ± 1	6 ± 1	5 ± 1	15 ± 2	4 ± 1
	K_D (Amp plot)	10 ± 5	13 ± 4	4 ± 1	2.2 ± 0.4	4 ± 1
mant-ATP	k_1 ($\mu M^{-1}s^{-1}$)	0.59 ± 0.04	0.9 ± 0.2			
	k_2 (s^{-1})	0.15 ± 0.01	0.05 ± 0.02			
	k_3 (s^{-1})	1.0 ± 0.2	n/a			
	k_{-1} (s^{-1})	0.33 ± 0.05	14 ± 3			
	k_{-2} (s^{-1})	5 ± 2	n/a			
	k_{-1} (s^{-1}) (k_1 plot)	3.4 ± 0.5	9 ± 2			
	K_D (μM)	19 ± 11	311 ± 157			
	K_D (μM) (k_1 plot)	13 ± 4	200 ± 102			
	K_D (Amp plot)	4 ± 1	4 ± 2			

3.3.3 Effect of Mg^{2+} on Nucleotide Binding

The coordination of Mg^{2+} by GTPases is essential for their catalytic activity.[21] In the GTP-bound conformation of a GTPase such as EF-Tu, the coordination of the Mg^{2+} ion is hexacoordinate in which two of the ligands are the β- and γ-phosphates. Two additional ligands are provided from within the protein itself in the form of hydroxyl groups; one from a Ser or Thr in the P-loop (Ser 17 in Ras, Thr 25 in EF-Tu) and a second Thr (Thr 35 in Ras and Thr 61 in EF-Tu) located within the G-2 motif. Lastly, two water molecules can act as the final two ligands, which can be aligned by other residues (such as a conserved Asp in Ras).[21] In contrast to GTP-bound structures, the Mg^{2+} in GDP bound complexes is coordinated by the β-phosphate, a

Ser/Thr from the P-loop and four water molecules.[21] Crystal structures of YchF have not been able to confirm whether Mg^{2+} is coordinated in a similar way as in other GTPases, however YchF does possess a conserved Ser in position 16 (P-loop) and Thr in position 36 (Switch I). In GTPases the binding of GTP and Mg^{2+} is tightly coupled, as seen by the fact that in G-proteins GTP dissociation is tenfold enhanced in the absence of Mg^{2+}.[21,130,131] However, the effect of Mg^{2+} on GDP binding can differ. For example in Ras the rate constant of GDP dissociation in the absence of magnesium is increased 20-fold, and association is 2 to 3-fold faster.[132] On the other hand Mg^{2+} has little effect on GDP binding by some G-proteins such as $G_{i\alpha}$, $G_{o\alpha}$, and $G_{s\alpha}$.[21,131,133] Stopped-flow experiments were performed as previously described in the presence of varying concentrations of Mg^{2+} and EDTA in order to analyze the effect of Mg^{2+} on nucleotide binding by YchF (Table 3.3.2). The absence of Mg^{2+} results in a 2 to 4-fold increase in the rate constant of mant-ADP association and an increase in the rate constant of mant-ADP dissociation up to 3-fold. Therefore Mg^{2+} has no major effect on the dissociation constant of mant-ADP from YchF.

Table 3.3.2 Effect of Mg^{2+} on the binding kinetics of mant-ADP to $YchF_{WT}$

	K_1 ($\mu M^{-1}s^{-1}$)	k_{-1} (s^{-1})	k_{-1} (s^{-1}) (k_1 plot)	K_D (μM)	K_D (μM) (k_1 plot)
$TAKM_7$	3.3 ± 0.5	30 ± 2	26 ± 2	9 ± 2	8 ± 2
10 mM EDTA	8 ± 1	44 ± 4	50 ± 2	5 ± 1	6.3 ± 0.1
TAK		57 ± 3			
$TAKM_{0.1}$		33 ± 5			
$TAKM_1$		31 ± 6			
$TAKM_{20}$	2.0 ± 0.04	22 ± 4	16 ± 0.1	11 ± 2	7.5 ± 0.2
$TAKM_{100}$		19 ± 5			
1 mM EDTA		44 ± 2			
200 µM EDTA		47 ± 2			
100 µM EDTA		50 ± 2			

Next we looked at the ability of mant-ATP to bind to YchF in the absence of Mg^{2+} ($TAKM_7$ + 10 mM EDTA buffer). The rate constant of mant-ATP association is 3.5-fold slower in the absence of Mg^{2+}. There is also a 44-fold increase in one of the rate constants associated with a conformational change (k_3). The two observed rate constants for mant-ATP dissociation have been affected differently in which k_{-1} (associated with a conformational change) has increased 190-fold in the absence of Mg^{2+} and k_{-2} (the rate constant of nucleotide dissociation) has decreased 2-fold. The K_D has increased approximately 6-fold in the absence of magnesium.

Table 3.3.3 Effect of Mg^{2+} on the binding kinetics of mant-ATP to $YchF_{WT}$

Rate Constant	**$YchF_{WT}$** ($TAKM_7$)	**$YchF_{WT}$** ($TAKM_7$ + 10 mM EDTA)
k_1 ($\mu M^{-1}s^{-1}$)	0.59 ± 0.04	0.17 ± 0.03
k_2 (s^{-1})	0.15 ± 0.01	0.1 ± 0.1
k_3 (s^{-1})	1.0 ± 0.2	44 ± 71
k_{-1} (s^{-1})	0.33 ± 0.05	63 ± 21
k_{-2} (s^{-1})	5 ± 2	2 ± 1
k_{-1} (s^{-1}) (k_1 plot)	3.4 ± 0.5	3.2 ± 0.4
K_D (μM)	19 ± 11	122 ± 244
K_D (μM) (k_1 plot)	13 ± 4	75.6 ± 147

3.3.4 Molecular Dynamics Simulations

To gain a better understanding of the role of Mg^{2+} in nucleotide binding we have constructed homology models of $YchF_{E.\ coli\ WT}$ in complex with either ATP or ADP, with or without Mg^{2+} (see 3.2.6 for details). The location of the Mg^{2+} was modeled based on its location in EF-Tu bound to either GTP or GDP (chapter 3.1.6). These homology models were subsequently hydrated with explicit waters and subjected to energy minimization. The stability of these models during simulations

was scored by measuring the RMSD of the backbone atoms with respect to the initial structure throughout the whole simulation (Figure 3.3.5 A). After an initial increase in RMSD by approximately 2 to 3 Å, dependent upon the simulation, the models appear to stabilize after 5 ns, however the YchF•ADP•Mg^{2+} simulation does show a 1 to 2 Å fluctuation around 35 ns (see figure 2.3.4 and Chapter 2.3 for details). YchF•ATP and YchF•ATP•Mg^{2+} simulations both stabilized around 2.6 Å, whereas YchF•ADP stabilized around 2 Å and YchF•ADP•Mg^{2+} stabilized around 3 Å (Figure 3.3.5 A). The flexibility of Switch I (residues 31 to 41) is higher in those simulations in which no Mg^{2+} is present for both nucleotides (Figure 3.3.5 B). The 'flexible' His 114-containing loop also shows varied flexibility in the different simulations with YchF bound to ADP in the absence of Mg^{2+} being the least flexible and YchF bound to ATP in the absence of Mg^{2+} being the most flexible (Figure 3.3.5 B). To observe any differences in the flexibility of the bound nucleotide the RMSF of each heavy atom was calculated and analyzed (Figure 3.3.6). In all simulations the nucleotide shows little fluctuation with an RMSF of 1.5 Å or less. The biggest change is in the flexibility of the γ-phosphate in the ATP-bound simulations in the absence of Mg^{2+}. The base of the nucleotide also appears to be more flexible when YchF is bound to ADP in the presence of Mg^{2+} and ATP in the absence of Mg^{2+}. As the overall RMSF is low we can assume that nucleotide binding in the absence of Mg^{2+} is feasible, which is also confirmed by the fact that over the 50 ns simulations the nucleotide position changes only minimally (Figure 3.3.7).

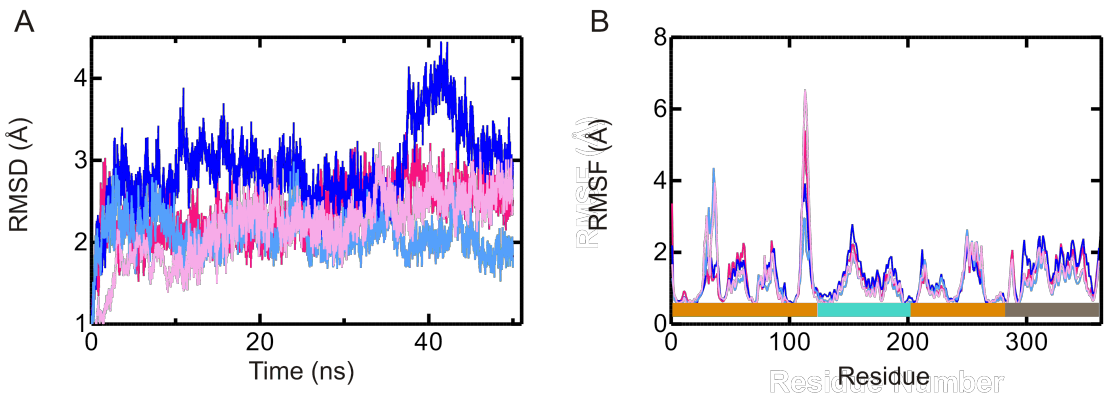


Figure 3.3.5 Conformational dynamics of YchF in the presence and absence of Mg^{2+} (A) RMSDs of $YchF_{WT}$ in complex with ATP or ADP in the presence of Mg^{2+} (dark pink and blue respectively) or in the absence of Mg^{2+} (light pink and blue). (B) Cα RMSF for $YchF_{WT}$ in complex with ATP or ADP in the presence of Mg^{2+} (dark pink and blue respectively) or in the absence of Mg^{2+} (light pink and blue). Colored bars at the bottom indicate the domain of the protein (G-domain in orange, alpha-helical domain in teal and the TGS domain in grey).

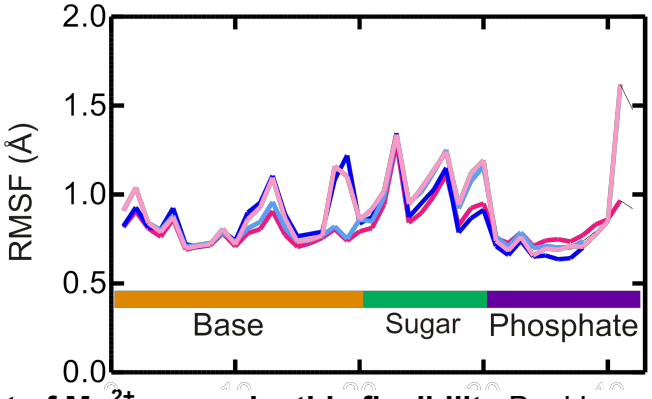


Figure 3.3.6 Effect of Mg^{2+} on nucleotide flexibility Backbone RMSF of the bound nucleotide in complex with ATP or ADP in the presence (dark pink and blue) and absence of Mg^{2+} (light pink and blue).

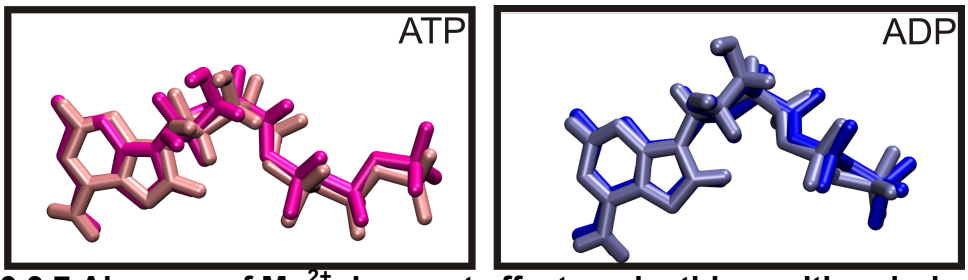


Figure 3.3.7 Absence of Mg^{2+} does not affect nucleotide position during 50 ns MD simulation Final 50 ns structures of YchF•ATP (light pink) and YchF•ATP•Mg^{2+} (dark pink) were aligned based on the P-loop (residues 9 to 16) and nucleotide positions were compared. Final 50 ns structures of YchF•ADP (light blue) and YchF•ADP•Mg^{2+} (dark blue) were aligned based on the P-loop (residues 9 to 16) and nucleotide positions were compared.

Next we used the MD simulations to observed the Mg^{2+} coordination in YchF, as the available crystal structures do not show resolution of a bound Mg^{2+} ion. As seen in other GTPases, Ser 16 of the G1 motif in YchF does coordinate the Mg^{2+} in both the GTP and GDP bound state. The remaining coordination spheres surrounding the Mg^{2+} ion are occupied by oxygen atoms provided by the β- and γ-phosphates (if present) and three (or four) water molecules (Figure 3.3.8). Noticeably, at no point does Thr 36 coordinate the Mg^{2+} for any simulation.

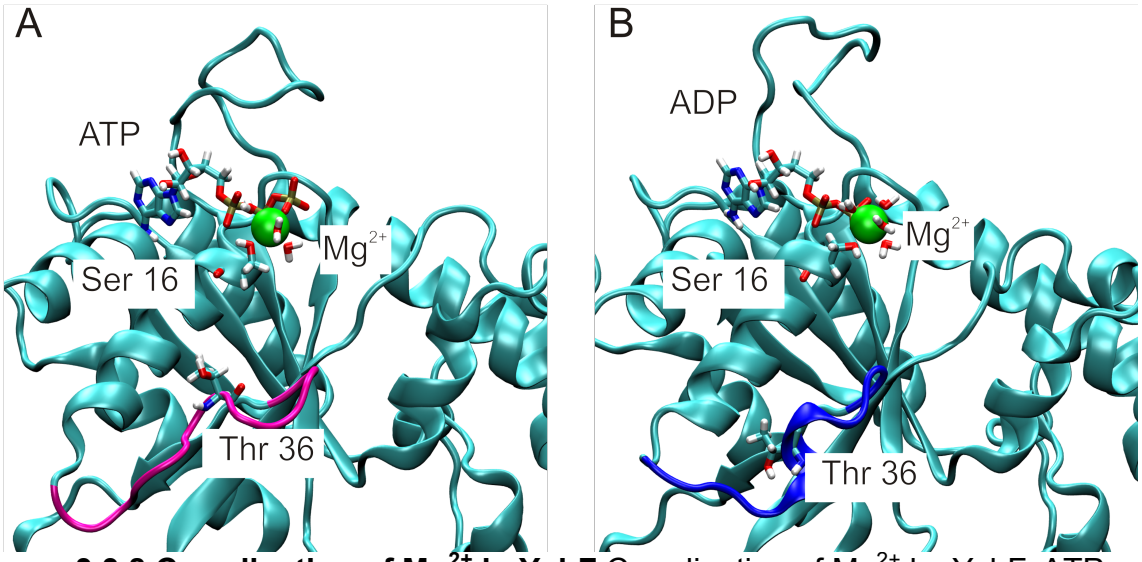


Figure 3.3.8 Coordination of Mg^{2+} in YchF Coordination of Mg^{2+} by YchF•ATP and YchF•ADP after 50 ns MD simulations. (A) Coordination of Mg^{2+} in YchF•ATP by oxygen atoms of the β- and γ-phosphates of ATP, the hydroxyl group of Ser 16 and three water molecules. (B) Coordination of Mg^{2+} in YchF•ADP by an oxygen atom of the β-phosphate, the hydroxyl group of ser 16 and four water molecules. Also shown is Thr 36.

3.4 Discussion and Future Directions

3.4.1 Pre-Steady State Nucleotide Binding

Despite its initial characterization as a GTPase, it has been shown that YchF binds and hydrolyzes ATP more efficiently than it does GTP.[18,33,79,80] Recent work by Becker *et al.* has provided equilibrium binding (K_D) constants for adenine and guanine nucleotides using a fluorescence based equilibrium binding assay.[18] This approach does not provide the individual rate constants of association and dissociation. Therefore, the binding of adenine nucleotides was studied using rapid kinetics with a stopped-flow device. The dissociation and association of mant-ATP to YchF required fitting with either a two- or three-exponential function respectively, whereas mant-ADP dissociation and association traces were fit with a one-exponential function. As we are observing a bimolecular binding event, the apparent rate constant for

association should be concentration dependent. However, of the three apparent rate constants determined for mant-ATP association, only one was nucleotide dependent, and therefore corresponded to the nucleotide-binding step. As YchF is capable of hydrolyzing ATP it was possible that one rate constant was due to the presence of mant-ADP in solution, however the rate constants did not match those of mant-ADP. It is evident from crystal structures that Ras shows conformational changes upon binding to GTP, and these changes occur mainly in Switch I (residues 32 to 38) and Switch II (residues 60 to 75).[26] It was also revealed through NMR studies that Ras can exist in two different states when bound to GDPNP or GTP and these states undergo rapid conversion on the millisecond time scale.[119-121,129] State one of the GTP complex resembles the *apo* state of Ras in which Thr 45 (Thr 35 in H-Ras) is not in direct contact with the Mg^{2+} and therefore not in indirect contact with the γ-phosphate.[128] State one is considered to be the 'off' state and conversion to state two gives the enzyme's 'on' state. It is possible that Switch I of YchF also has two conformational states and the interconversion of these states accounts for one of the signals observed. We have also previously described a model in which a 'flexible' loop of the G-domain contains an essential His residue for intrinsic ATP hydrolysis by YchF (see chapter 2). Based on available crystal structures and MD simulations it was clear that catalytically active conformation the 'flexible' loop would require a conformational change in order to position His 114 in the active site. Therefore, it was hypothesized that the second conformational change observed is a result of the 'flexible' loop. Pre-steady-state nucleotide binding kinetics of $YchF_{H114A}$ were determined to confirm that the mutation of His 114 lowers the affinity of YchF for ATP

but does not prevent binding, as well as to test our hypothesis that a conformational change in the His 114-containing loop is responsible for a concentration independent rate constant for mant-ATP binding. Interestingly, mant-ATP dissociation and association traces were fit with a one-exponential and two-exponential function respectively, as compared to a two-exponential and three-exponential function for $YchF_{WT}$. The loss of one observed phase confirms our hypothesis that it was a result of a conformational change in the His 114-containing loop. These results also suggest that the nucleotide-dependent movement of the 'flexible' loop is a direct result of the presence of the His^{cat}. Thus, the model of ATP binding to YchF was updated such that upon ATP binding YchF can undergo two conformational changes, one of which can be attributed to the movement of the 'flexible' loop containing the catalytic His 114. It has also been hypothesized that the second conformational change is a result of two conformations in Switch I, similar to that seen for Ras. The proposed model of ATP binding to YchF is shown in Figure 3.4.1. It is important to note however that the amplitude of nucleotide binding to $YchF_{H114A}$ is significantly less than that see for $YchF_{WT}$, therefore resolution between the phases may be the reason we lose the extra phase. A reduction in the amplitude could suggest an overall reduction in nucleotide binding such that while nucleotides are still able to bind with the same affinity they can only bind to a subset of enzyme. Alternatively, we could also be seeing a decrease in FRET efficiency, which would suggest either the nucleotide is binding differently or in a different location. Overall however, $YchF_{H114A}$ is able to bind both mant-ADP and mant-ATP, confirming that only nucleotide hydrolysis is affected by the mutation of His 114 to Ala.

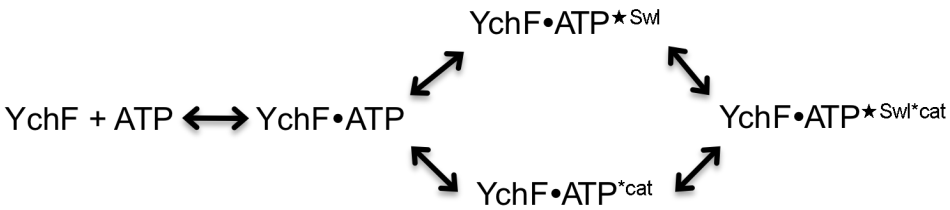


Figure 3.4.1 Proposed model of ATP binding to $YchF_{WT}$ ATP binding to YchF can induce two conformational changes: one in the Switch I region ($\star^{Swl}$) and one in the 'flexible' loop containing His^{cat} (^{*cat})

Next we compared the rate constants of nucleotide association and dissociation between the $YchF_{WT}$ and YchF variants. The dissociation rate constant of mant-ADP from $YchF_{WT}$ is 6-fold faster than that for mant-ATP (k_{-2} 30 ± 2 s^{-1} and 5 ± 2 s^{-1} respectively). The association of mant-ADP is 6-fold faster than the association of mant-ATP to $YchF_{WT}$ (3.3 ± 0.5 $\mu M^{-1}s^{-1}$ and 0.59 ± 0.04 $\mu M^{-1}s^{-1}$ respectively) resulting in equivalent dissociation constants (within error) for both mant-ADP and mant-ATP (8 ± 1 μM and 9 ± 3 μM respectively). YchF binds nucleotides with a similar affinity to other HAS-GTPases such as FeoB, Era and EngA, which have affinities for both GTP and GDP in the low micromolar range.[34,57,58] Since YchF has a similar affinity for both ATP and ADP and ADP dissociates quickly, it is unlikely that YchF requires a nucleotide exchange factor as the cellular concentration of ATP is 30-times greater than ADP during log-phase.[134] $YchF_{H114A}$ binds mant-ADP (7 ± 1 μM) with a similar affinity as wild type, but $YchF_{H114A}$ binds mant-ATP with a an affinity that is 16-fold lower than wild type due to an approximately 3-fold increase in k_{-1} and 42-fold decrease in k_2.

$YchF_{H114Q}$, $YchF_{H114K}$ and $YchF_{H114R}$ were also tested for their ability to bind mant-ADP. $YchF_{H114Q}$, $YchF_{H114R}$ and $YchF_{H114K}$ bind mant-ADP with an affinity similar to wild type (5 ± 1 μM, 4 ± 1 μM and 15 ± 2 μM respectively).

3.4.2 Effect of Mg^{2+} on Nucleotide Binding

Using pre-steady-state kinetics we have seen that Mg^{2+} stabilizes the bound nucleotide and slows down mant-ADP release from YchF, however this effect is not reflected in the equilibrium binding constants as the association rate constant of mant-ADP is also slowed down. Therefore, mant-ADP binding to YchF is not affected by the presence of Mg^{2+}, which is similar to that of $G_{i\alpha}$, $G_{o\alpha}$, and $G_{s\alpha}$, but different from Ras.[21,131,133] More surprising was that while mant-ATP association was slowed down in the presence of Mg^{2+}, the two observed rate constants of dissociation showed drastically different effects in which one was increased 190-fold and the other decreased by a factor of 2. Overall the affinity of YchF for mant-ATP has been reduced 6-fold in the absence of Mg^{2+}, which is similar to that observed in other GTPases. The reduced affinity is a result of the rate-limiting conformational change step during mant-ATP dissociation, and has increased 190-fold in the absence of Mg^{2+}.

3.4.3 MD Simulations

GTPases require a Mg^{2+} ion for GTP hydrolysis, and coordination of the ion is achieved through the use of the β- and γ-phosphates, three water molecules and two hydroxyl groups from the protein itself (either two Thr residues or a Ser and a Thr). Thus far it is unknown how Mg^{2+} is coordinated in YchF as crystal structures have been unable to provide this information. MD simulations of YchF bound to nucleotide and Mg^{2+} show that Ser 16 but not Thr 36, is coordinating the bound Mg^{2+} ion in either the ADP or ATP complex. As the homology model was based on the crystal structure of *apo* YchF from *H. influenzae*, it is likely that a conformational change in Switch I

would occur in order to facilitate Mg^{2+} coordination by Thr 35 for at least the ATP complex but this large of a conformational change would not be seen on the nanosecond time scale. It is interesting however that the absence of a Mg^{2+} ion already allows the Switch I region more flexibility to move. Overall, the absence of Mg^{2+} did not affect the simulations in such a way that the protein complexes were destabilized, as shown by the small, stable RMSD values for each simulation. The flexibility of the nucleotide in each simulation was also small, proving that the absence of Mg^{2+} does not completely destabilize the interaction of the nucleotide with YchF.

3.5 Conclusions

$YchF_{WT}$ binds to mant-ATP and mant-ADP with similar affinities around 10 μM, suggesting that YchF does not require a nucleotide exchange factor. However, binding of mant-ATP to YchF appears to induce two conformational changes, one of which can be attributed to a conformational change in the 'flexible' loop of YchF as evidenced by a loss of one conformational change when His 114 is mutated to alanine. Based on these findings we have proposed a model of ATP binding to YchF (Figure 3.4.1). It has also been confirmed that the mutation of His 114 to alanine does not affect nucleotide binding to YchF.

The coordination of Mg^{2+} by GTPases such as the Ras and translation factor superfamily has been well characterized, but little is known about the effect of Mg^{2+} on nucleotide binding by YchF. Fluorescent stopped-flow experiments were used to show that the binding of mant-ADP to YchF is not affected by the absence of Mg^{2+}, however the affinity of YchF for mant-ATP is decreased 6-fold.

While crystal structures have not been able to confirm how a Mg^{2+} ion is coordinated in YchF I have shown through the use of homology modeling and molecular dynamics simulations that a conserved serine in the P-loop, along with the β- and Y-phosphates of ATP and three water molecules are coordinating the ion. This is the first time a model for Mg^{2+} coordination in YchF has been described.

Chapter 4. Conclusions

4.1 Catalytic Mechanisms of GTP Hydrolysis

GTPases perform a myriad of functions within the cell, from protein synthesis to cellular signaling.[1,4] In order to perform their critical cellular functions, they must hydrolyze a bound GTP molecule to GDP and P_i. In all GTPases, a catalytic water molecule is activated for a nucleophillic attack on the γ-phosphate of GTP. Ras, often considered the model GTPase, activates the water molecule using a conserved, catalytic glutamine residue (Gln 61), which has been replaced with a histidine in the translational GTPases, threonine in Rap and arginine in Ffh.[21,22] A novel class of GTPases, the HAS-GTPases, do not possess a catalytic residue in the Gln^{cat} position, and have developed alternative mechanisms to catalyze GTP hydrolysis.[39] In both MnmE and YqeH, a catalytic residue is provided from helix α2 in the form of a glutamate and aspartate residue respectively, which aligns a bridging water molecule to the catalytic water.[31,35] RbgA, similarly to the translational GTPases, uses a His^{cat}, whereas the prokaryotic membrane-bound protein FeoB appears to lack a specific catalytic side chain altogether, and uses two backbone amides instead.[62,69] At this time very little is known about the mechanisms of the HAS-GTPases but it appears that they do not share a conserved catalytic mechanism, likely due to their divergent functions. The results in the present work provide support for a catalytic mechanism of ATP hydrolysis in the universally conserved HAS-GTPase, YchF that is similar to the mechanism employed by RbgA.

4.2 The role of His 114 in YchF

In this work I have proposed that His 114 is essential for ATP hydrolysis by YchF (**Chapter 2**). The pH dependence of ATP hydrolysis by YchF suggested that a hisitidine residue facilitated catalysis. YchF from *E. coli* possesses two histidine residues within the G-domain, His 102 and His 114, which are 76% and 98% conserved, respectively. MD simulations of $YchF_{E.\ coli}$ homology models in the *apo* state, or in complex with ATP or ADP, were performed over 50 ns to gain insight into the roles of these two conserved residues. These simulations revealed that His 114 is located in a 'flexible' loop region, which shows nucleotide dependent conformations, such that when YchF is bound to ATP, His 114 and the 'flexible' loop are in closer proximity to the nucleotide binding pocket. In comparison, His 102 makes a stable interaction with Tyr 204 in all simulations observed, suggesting it may play a mainly structural role but is unlikely to directly participate in catalysis of ATP hydrolysis. Several YchF variants bearing a single amino acid substitution at position 102 or 114 were studied in this work. Both equilibrium and pre-steady-state nucleotide binding experiments revealed that all His 114 variants are capable of binding adenine nucleotides, and do so with an affinity similar to wild type. However, while His 114 variants of YchF can bind adenine nucleotides, they are impaired in their ability to hydrolyze ATP, with the exception of $YchF_{H114R}$. The conformational changes observed in the 'flexible' loop during MD simulations were confirmed using pre-steady-state kinetics as two conformational changes could be observed upon mant-ATP binding to $YchF_{WT}$, one of which is abolished if His 114 is replaced with an alanine residue (**Chapter 3**).

Chapter *3* investigated the effect of Mg^{2+} on nucleotide binding in YchF. While Mg^{2+} is known to coordinate the β- and γ-phosphates of GTPases, this had not been confirmed for YchF as crystal structures could not resolve its position in the nucleotide-binding pocket. MD simulations revealed that YchF coordinates the Mg^{2+} with the β- and γ-phosphates, the hydroxyl group of Ser 16 and three water molecules. Interestingly, the Mg^{2+} ion was not essential for mant-ATP of mant-ADP binding to YchF or the YchF variants, but decreased the rate constant of mant-ATP dissociation by slowing down a conformational change preceding dissociation, thereby reducing the affinity of YchF for mant-ATP approximately 6-fold.

Recently, it has been shown that the 70S ribosome is capable of stimulating ATP hydrolysis in YchF by an unknown mechanism.[18] Results presented in **Chapter 2** reveal that the 70S ribosome is likely helping to position His 114 and the 'flexible' loop, while also contributing to the catalytic machinery. This hypothesis is based on the observation that a catalytically inactive variant of YchF ($YchF_{H114A}$) can be activated to hydrolyze ATP in the presence of the 70S ribosome, however not to the extent of wild type. These findings suggest that optimal ATPase activity by YchF requires catalytic machinery both in cis and in trans (70S ribosome). In presently available crystal structures His 114 is not in a catalytically active position supporting our hypothesis that the ribosome may facilitate hydrolysis by aiding in the positioning of the 'flexible' loop and therefore His 114.

References

1 Bourne, H. R. S., D.A.; McCormick, F. The GTPase superfamily: a conserved switch for diverse cell functions. *Nature* **348** (1990).

2 Saraste, M., Sibbald, P. R. & Wittinghofer, A. The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* **15**, 430-434 (1990).

3 Caldon, C. E., Yoong, P. & March, P. E. Evolution of a molecular switch: universal bacterial GTPases regulate ribosome function. *Mol Microbiol* **41**, 289-297 (2001).

4 Bourne, H. R., Sanders, D. A. & McCormick, F. The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**, 117-127 (1991).

5 Sprang, S. R. G proteins, effectors and GAPs: structure and mechanism. *Curr Opin Struct Biol* **7**, 849-856 (1997).

6 Cherfils, J. & Chardin, P. GEFs: structural basis for their activation of small GTP-binding proteins. *Trends Biochem Sci* **24**, 306-311 (1999).

7 Leipe, D. D., Wolf, Y. I., Koonin, E. V. & Aravind, L. Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol* **317**, 41-72 (2002).

8 Caldon, C. E. & March, P. E. Function of the universally conserved bacterial GTPases. *Curr Opin Microbiol* **6**, 135-139 (2003).

9 Cabrer, B., San-Millian, M. J., Vazquez, D. & Modolell, J. Stoichiometry of polypeptide chain elongation. *J Biol Chem* **251**, 1718-1722 (1976).

10 Verstraeten, N., Fauvart, M., Versees, W. & Michiels, J. The universally conserved prokaryotic GTPases. *Microbiol Mol Biol Rev* **75**, 507-542, second and third pages of table of contents (2011).

11 Pape, T., Wintermeyer, W. & Rodnina, M. V. Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the *E. coli* ribosome. *EMBO J* **17**, 7490-7497 (1998).

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

13 Powers, T. & Walter, P. Co-translational protein targeting catalyzed by the *Escherichia coli* signal recognition particle and its receptor. *EMBO J* **16**, 4880-4886 (1997).

14 Wicker-Planquart, C., Foucher, A. E., Louwagie, M., Britton, R. A. & Jault, J. M. Interactions of an essential *Bacillus subtilis* GTPase, YsxC, with ribosomes. *J Bacteriol* **190**, 681-690 (2008).

15 Dassain, M., Leroy, A., Colosetti, L., Carole, S. & Bouche, J. P. A new essential gene of the 'minimal genome' affecting cell division. *Biochimie* **81**, 889-895 (1999).

16 Fischer, J. J. *et al.* The ribosome modulates the structural dynamics of the conserved GTPase HflX and triggers tight nucleotide binding. *Biochimie* **94**, 1647-1659 (2012).

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

18 Becker, M. *et al.* The 70S ribosome modulates the ATPase activity of *Escherichia coli* YchF. *RNA Biol* **9** (2012).
19 Zhang, J., Rubio, V., Lieberman, M. W. & Shi, Z. Z. OLA1, an Obg-like ATPase, suppresses antioxidant response via nontranscriptional mechanisms. *Proc Natl Acad Sci U S A* **106**, 15356-15361 (2009).
20 Wenk, M. *et al.* A Universally Conserved ATPase Regulates the Oxidative Stress Response in *Escherichia coli*. *J Biol Chem* **287**, 43585-43598 (2012).
21 Sprang, S. R. G protein mechanisms: insights from structural analysis. *Annu Rev Biochem* **66**, 639-678 (1997).
22 Scheffzek, K. *et al.* The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* **277**, 333-338 (1997).
23 Bos, J. L., Rehmann, H. & Wittinghofer, A. GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **129**, 865-877 (2007).
24 Neal, S. E., Eccleston, J. F., Hall, A. & Webb, M. R. Kinetic analysis of the hydrolysis of GTP by p21N-ras. The basal GTPase mechanism. *J Biol Chem* **263**, 19718-19722 (1988).
25 Siderovski, D. P. & Willard, F. S. The GAPs, GEFs, and GDIs of heterotrimeric G-protein alpha subunits. *Int J Biol Sci* **1**, 51-66 (2005).
26 Vetter, I. R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299-1304 (2001).
27 Karnoub, A. E. & Weinberg, R. A. Ras oncogenes: split personalities. *Nat Rev Mol Cell Biol* **9**, 517-531 (2008).
28 Seewald, M. J., Korner, C., Wittinghofer, A. & Vetter, I. R. RanGAP mediates GTP hydrolysis without an arginine finger. *Nature* **415**, 662-666 (2002).
29 Rehmann, H. & Bos, J. L. Signal transduction: thumbs up for inactivation. *Nature* **429**, 138-139 (2004).
30 Praefcke, G. J. & McMahon, H. T. The dynamin superfamily: universal membrane tubulation and fission molecules? *Nat Rev Mol Cell Biol* **5**, 133-147 (2004).
31 Scrima, A. & Wittinghofer, A. Dimerisation-dependent GTPase reaction of MnmE: how potassium acts as GTPase-activating element. *EMBO J* **25**, 2940-2951 (2006).
32 Yamanaka, K., Hwang, J. & Inouye, M. Characterization of GTPase activity of TrmE, a member of a novel GTPase superfamily, from *Thermotoga maritima*. *J Bacteriol* **182**, 7078-7082 (2000).
33 Tomar, S. K., Kumar, P. & Prakash, B. Deciphering the catalytic machinery in a universally conserved ribosome binding ATPase YchF. *Biochem Biophys Res Commun* **408**, 459-464 (2011).
34 Foucher, A. E., Reiser, J. B., Ebel, C., Housset, D. & Jault, J. M. Potassium acts as a GTPase-activating element on each nucleotide-binding domain of the essential *Bacillus subtilis* EngA. *PLoS One* **7**, e46795 (2012).
35 Anand, B., Surana, P. & Prakash, B. Deciphering the catalytic machinery in 30S ribosome assembly GTPase YqeH. *PLoS One* **5**, e9944 (2010).

36 Anand, B., Verma, S. K. & Prakash, B. Structural stabilization of GTP-binding domains in circularly permuted GTPases: implications for RNA binding. *Nucleic Acids Res* **34**, 2196-2205 (2006).
37 Campbell, T. L., Daigle, D. M. & Brown, E. D. Characterization of the *Bacillus subtilis* GTPase YloQ and its role in ribosome function. *Biochem J* **389**, 843-852 (2005).
38 Im, C. H. *et al.* Nuclear/nucleolar GTPase 2 proteins as a subfamily of YlqF/YawG GTPases function in pre-60S ribosomal subunit maturation of mono- and dicotyledonous plants. *J Biol Chem* **286**, 8620-8632 (2011).
39 Mishra, R., Gara, S. K., Mishra, S. & Prakash, B. Analysis of GTPases carrying hydrophobic amino acid substitutions in lieu of the catalytic glutamine: implications for GTP hydrolysis. *Proteins* **59**, 332-338 (2005).
40 Elseviers, D., Petrullo, L. A. & Gallagher, P. J. Novel *E. coli* mutants deficient in biosynthesis of 5-methylaminomethyl-2-thiouridine. *Nucleic Acids Res* **12**, 3521-3534 (1984).
41 Cabedo, H. *et al.* The *Escherichia coli* trmE (mnmE) gene, involved in tRNA modification, codes for an evolutionarily conserved GTPase with unusual biochemical properties. *EMBO J* **18**, 7063-7076 (1999).
42 Colby, G., Wu, M. & Tzagoloff, A. MTO1 codes for a mitochondrial protein required for respiration in paromomycin-resistant mutants of *Saccharomyces cerevisiae*. *J Biol Chem* **273**, 27945-27952 (1998).
43 Decoster, E., Vassal, A. & Faye, G. MSS1, a nuclear-encoded mitochondrial GTPase involved in the expression of COX1 subunit of cytochrome c oxidase. *J Mol Biol* **232**, 79-88 (1993).
44 Singh, A. K., Pindi, P. K., Dube, S., Sundareswaran, V. R. & Shivaji, S. Importance of trmE for growth of the psychrophile *Pseudomonas syringae* at low temperatures. *Appl Environ Microbiol* **75**, 4419-4426 (2009).
45 Hagervall, T. G., Edmonds, C. G., McCloskey, J. A. & Bjork, G. R. Transfer RNA(5-methylaminomethyl-2-thiouridine)-methyltransferase from *Escherichia coli* K-12 has two enzymatic activities. *J Biol Chem* **262**, 8488-8495 (1987).
46 Hagervall, T. G., Pomerantz, S. C. & McCloskey, J. A. Reduced misreading of asparagine codons by *Escherichia coli* tRNALys with hypomodified derivatives of 5-methylaminomethyl-2-thiouridine in the wobble position. *J Mol Biol* **284**, 33-42 (1998).
47 Bregeon, D., Colot, V., Radman, M. & Taddei, F. Translational misreading: a tRNA modification counteracts a +2 ribosomal frameshift. *Genes Dev* **15**, 2295-2306 (2001).
48 Yim, L. *et al.* The GTPase activity and C-terminal cysteine of the *Escherichia coli* MnmE protein are essential for its tRNA modifying function. *J Biol Chem* **278**, 28378-28387 (2003).
49 Martinez-Vicente, M. *et al.* Effects of mutagenesis in the switch I region and conserved arginines of *Escherichia coli* MnmE protein, a GTPase involved in tRNA modification. *J Biol Chem* **280**, 30660-30670 (2005).
50 Loh, P. C., Morimoto, T., Matsuo, Y., Oshima, T. & Ogasawara, N. The GTP-binding protein YqeH participates in biogenesis of the 30S ribosome subunit in *Bacillus subtilis*. *Genes Genet Syst* **82**, 281-289 (2007).

51 Sudhamsu, J., Lee, G. I., Klessig, D. F. & Crane, B. R. The structure of YqeH. An AtNOS1/AtNOA1 ortholog that couples GTP hydrolysis to molecular recognition. *J Biol Chem* **283**, 32968-32976 (2008).
52 Uicker, W. C., Schaefer, L., Koenigsknecht, M. & Britton, R. A. The essential GTPase YqeH is required for proper ribosome assembly in *Bacillus subtilis*. *J Bacteriol* **189**, 2926-2929 (2007).
53 Sharma, M. R. *et al.* Interaction of Era with the 30S ribosomal subunit implications for 30S subunit assembly. *Mol Cell* **18**, 319-329 (2005).
54 Anand, B., Surana, P., Bhogaraju, S., Pahari, S. & Prakash, B. Circularly permuted GTPase YqeH binds 30S ribosomal subunit: Implications for its role in ribosome assembly. *Biochem Biophys Res Commun* **386**, 602-606 (2009).
55 Kammler, M., Schon, C. & Hantke, K. Characterization of the ferrous iron uptake system of *Escherichia coli*. *J Bacteriol* **175**, 6212-6219 (1993).
56 Eng, E. T., Jalilian, A. R., Spasov, K. A. & Unger, V. M. Characterization of a novel prokaryotic GDP dissociation inhibitor domain from the G protein coupled membrane protein FeoB. *J Mol Biol* **375**, 1086-1097 (2008).
57 Marlovits, T. C., Haase, W., Herrmann, C., Aller, S. G. & Unger, V. M. The membrane protein FeoB contains an intramolecular G protein essential for Fe(II) uptake in bacteria. *Proc Natl Acad Sci U S A* **99**, 16243-16248 (2002).
58 Sullivan, S. M., Mishra, R., Neubig, R. R. & Maddock, J. R. Analysis of guanine nucleotide binding and exchange kinetics of the *Escherichia coli* GTPase Era. *J Bacteriol* **182**, 3460-3466 (2000).
59 Guilfoyle, A. *et al.* Structural basis of GDP release and gating in G protein coupled Fe2+ transport. *EMBO J* **28**, 2677-2685 (2009).
60 Cho, K. H. & Caparon, M. G. tRNA modification by GidA/MnmE is necessary for Streptococcus pyogenes virulence: a new strategy to make live attenuated strains. *Infect Immun* **76**, 3176-3186 (2008).
61 Ash, M. R. *et al.* Potassium-activated GTPase reaction in the G Protein-coupled ferrous iron transporter B. *J Biol Chem* **285**, 14594-14602 (2010).
62 Ash, M. R., Maher, M. J., Guss, J. M. & Jormakka, M. The initiation of GTP hydrolysis by the G-domain of FeoB: insights from a transition-state complex structure. *PLoS One* **6**, e23355 (2011).
63 Achila, D., Gulati, M., Jain, N. & Britton, R. A. Biochemical characterization of ribosome assembly GTPase RbgA in *Bacillus subtilis*. *J Biol Chem* **287**, 8417-8423 (2012).
64 Uicker, W. C., Schaefer, L. & Britton, R. A. The essential GTPase RbgA (YlqF) is required for 50S ribosome assembly in *Bacillus subtilis*. *Mol Microbiol* **59**, 528-540 (2006).
65 Kallstrom, G., Hedges, J. & Johnson, A. The putative GTPases Nog1p and Lsg1p are required for 60S ribosomal subunit biogenesis and are localized to the nucleus and cytoplasm, respectively. *Mol Cell Biol* **23**, 4344-4355 (2003).
66 Morimoto, T. *et al.* Six GTP-binding proteins of the Era/Obg family are essential for cell growth in *Bacillus subtilis*. *Microbiology* **148**, 3539-3552 (2002).

67 Zalacain, M. *et al.* A global approach to identify novel broad-spectrum antibacterial targets among proteins of unknown function. *J Mol Microbiol Biotechnol* **6**, 109-126 (2003).
68 Kniewel, R., Buglino, J., and Lima, C.D. (New York Structural Genomics Research Consortium 2003).
69 Gulati, M., Jain, N., Anand, B., Prakash, B. & Britton, R. A. Mutational analysis of the ribosome assembly GTPase RbgA provides insight into ribosome interaction and ribosome-stimulated GTPase activation. *Nucleic Acids Res* (2013).
70 Jensen, B. C., Wang, Q., Kifer, C. T. & Parsons, M. The NOG1 GTP-binding protein is required for biogenesis of the 60 S ribosomal subunit. *J Biol Chem* **278**, 32204-32211 (2003).
71 Matsuo, Y. *et al.* The GTP-binding protein YlqF participates in the late step of 50 S ribosomal subunit assembly in *Bacillus subtilis*. *J Biol Chem* **281**, 8110-8117 (2006).
72 Barrientos, A. *et al.* MTG1 codes for a conserved protein required for mitochondrial translation. *Mol Biol Cell* **14**, 2292-2302 (2003).
73 Voorhees, R. M., Schmeing, T. M., Kelley, A. C. & Ramakrishnan, V. The mechanism for activation of GTP hydrolysis on the ribosome. *Science* **330**, 835-838 (2010).
74 Cruz-Vera, L. R., Galindo, J. M. & Guarneros, G. Transcriptional analysis of the gene encoding peptidyl-tRNA hydrolase in *Escherichia coli*. *Microbiology* **148**, 3457-3466 (2002).
75 Galperin, M. Y. & Koonin, E. V. 'Conserved hypothetical' proteins: prioritization of targets for experimental study. *Nucleic Acids Res* **32**, 5452-5463 (2004).
76 Mittenhuber, G. 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). *J Mol Microbiol Biotechnol* **3**, 21-35 (2001).
77 Adekambi, T. *et al.* Core gene set as the basis of multilocus sequence analysis of the subclass Actinobacteridae. *PLoS One* **6**, e14792 (2011).
78 Teplyakov, A. *et al.* Crystal structure of the YchF protein reveals binding sites for GTP and nucleic acid. *J Bacteriol* **185**, 4031-4037 (2003).
79 Koller-Eichhorn, R. *et al.* Human OLA1 defines an ATPase subfamily in the Obg family of GTP-binding proteins. *J Biol Chem* **282**, 19928-19937 (2007).
80 Gradia, D. F. *et al.* Characterization of a novel Obg-like ATPase in the protozoan *Trypanosoma cruzi*. *Int J Parasitol* **39**, 49-58 (2009).
81 Meury, J. & Kepes, A. The regulation of potassium fluxes for the adjustment and maintenance of potassium levels in *Escherichia coli*. *Eur J Biochem* **119**, 165-170 (1981).
82 Cadenas, E. Basic mechanisms of antioxidant activity. *Biofactors* **6**, 391-397 (1997).
83 Sayre, L. M., Smith, M. A. & Perry, G. Chemistry and biochemistry of oxidative stress in neurodegenerative disease. *Curr Med Chem* **8**, 721-738 (2001).
84 Floyd, R. A. Antioxidants, oxidative stress, and degenerative neurological disorders. *Proc Soc Exp Biol Med* **222**, 236-245, doi:pse22217 [pii] (1999).

85 Valko, M. *et al.* Free radicals and antioxidants in normal physiological functions and human disease. *Int J Biochem Cell Biol* **39**, 44-84 (2007).
86 Zhang, J. W., Rubio, V., Zheng, S. & Shi, Z. Z. Knockdown of OLA1, a regulator of oxidative stress response, inhibits motility and invasion of breast cancer cells. *J Zhejiang Univ Sci B* **10**, 796-804, doi:10.1631/jzus.B0910009 (2009).
87 Chene, P. ATPases as drug targets: learning from their structure. *Nat Rev Drug Discov* **1**, 665-673 (2002).
88 Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-147 (2002).
89 Guerrero, C., Tagwerker, C., Kaiser, P. & Huang, L. An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. *Mol Cell Proteomics* **5**, 366-378 (2006).
90 Wainwright, B. J. A. a. T. E. Studies in Molecular Dynamics. I. General Method. *J. Chem. Phys.* **31**, 459-467 (1959).
91 Young, D. C. *Computational Chemistry: a practical guide for applying techniques to real world problems*. (John Wiley & Sons, 2001).
92 Henzler-Wildman, K. A. *et al.* A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913-916 (2007).
93 Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781-1802 (2005).
94 A. D. MacKerell *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **102**, 3586-3616 (1998).
95 MacKerell, A. D., Jr., Banavali, N. & Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **56**, 257-265 (2000).
96 Dror, R. O. *et al.* Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc Natl Acad Sci U S A* **108**, 13118-13123 (2011).
97 Freddolino, P. L., Liu, F., Gruebele, M. & Schulten, K. Ten-microsecond molecular dynamics simulation of a fast-folding WW domain. *Biophys J* **94**, L75-77 (2008).
98 Wolf-Watz, M. *et al.* Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat Struct Mol Biol* **11**, 945-949 (2004).
99 Velayudhan, J. *et al.* Iron acquisition and virulence in *Helicobacter pylori*: a major role for FeoB, a high-affinity ferrous iron transporter. *Mol Microbiol* **37**, 274-286, doi:mmi1987 [pii] (2000).
100 Cheung, M. Y. *et al.* An ancient P-loop GTPase in rice is regulated by a higher plant-specific regulatory protein. *J Biol Chem* **285**, 37359-37369 (2010).
101 Chen, Y. C. & Chung, Y. T. A conserved GTPase YchF of *Vibrio vulnificus* is involved in macrophage cytotoxicity, iron acquisition, and mouse virulence. *Int J Med Microbiol* **301**, 469-474 (2011).
102 Fernebro, J. *et al.* The influence of in vitro fitness defects on pneumococcal ability to colonize and to cause invasive disease. *BMC Microbiol* **8**, 65 (2008).

103 Danese, I. *et al.* The Ton system, an ABC transporter, and a universally conserved GTPase are involved in iron utilization by *Brucella melitensis* 16M. *Infect Immun* **72**, 5783-5790 (2004).
104 Gasteiger, E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D. and Bairoch A. *Protein Identification and Analysis Tools on the ExPASy Server*. 571-607 (Humana Press, 2005).
105 Girish, V. & Vijayalakshmi, A. Affordable image analysis using NIH Image/ImageJ. *Indian J Cancer* **41**, 47 (2004).
106 Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* **31**, 3381-3385 (2003).
107 Kiefer F., A. K., Kunzli M., Bordoli L. and Schwede T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* **37**, D387-D392 (2009).
108 Arnold, K., Bordoli L., Kopp J., and Schwede T. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201 (2006).
109 Peitsch, M. C. Protein modeling by E-mail *Bio/Technology* **13**, 658-660 (1995).
110 Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).
111 Glykos, N. M. Software news and updates. Carma: a molecular dynamics analysis program. *J Comput Chem* **27**, 1765-1768 (2006).
112 Voet, D. V. a. J. G. (ed David Harris and Patrick Fitzgerald) (John Wiley & Sons, Inc, United States of America, 2004).
113 Noel, J. P., Hamm, H. E. & Sigler, P. B. The 2.2 A crystal structure of transducin-alpha complexed with GTP gamma S. *Nature* **366**, 654-663 (1993).
114 Coleman, D. E. *et al.* Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis. *Science* **265**, 1405-1412 (1994).
115 Wieden, H. J., Mercier, E., Gray, J., Steed, B. & Yawney, D. A combined molecular dynamics and rapid kinetics approach to identify conserved three-dimensional communication networks in elongation factor Tu. *Biophys J* **99**, 3735-3743 (2010).
116 Milburn, M. V. *et al.* Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science* **247**, 939-945 (1990).
117 Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. 3rd Edition edn, (Springer Sciene + Business Media, n. , 2006).
118 Spoerner, M., Herrmann, C., Vetter, I. R., Kalbitzer, H. R. & Wittinghofer, A. Dynamic properties of the Ras switch I region and its importance for binding to effectors. *Proc Natl Acad Sci U S A* **98**, 4944-4949 (2001).
119 Spoerner, M., Wittinghofer, A. & Kalbitzer, H. R. Perturbation of the conformational equilibria in Ras by selective mutations as studied by 31P NMR spectroscopy. *FEBS Lett* **578**, 305-310 (2004).
120 Geyer, M. *et al.* Conformational transitions in p21ras and in its complexes with the effector protein Raf-RBD and the GTPase activating protein GAP. *Biochemistry* **35**, 10308-10320 (1996).

121 Linnemann, T. *et al.* Thermodynamic and kinetic characterization of the interaction between the Ras binding domain of AF6 and members of the Ras subfamily. *J Biol Chem* **274**, 13556-13562 (1999).
122 Wilden, B., Savelsbergh, A., Rodnina, M. V. & Wintermeyer, W. Role and timing of GTP binding and hydrolysis during EF-G-dependent tRNA translocation on the ribosome. *Proc Natl Acad Sci U S A* **103**, 13670-13675 (2006).
123 Lenzen, C., Cool, R. H., Prinz, H., Kuhlmann, J. & Wittinghofer, A. Kinetic analysis by fluorescence of the interaction between Ras and the catalytic domain of the guanine nucleotide exchange factor Cdc25Mm. *Biochemistry* **37**, 7420-7430 (1998).
124 John, J. *et al.* Kinetics of interaction of nucleotides with nucleotide-free H-ras p21. *Biochemistry* **29**, 6058-6065 (1990).
125 Gromadski, K. B., Wieden, H. J. & Rodnina, M. V. Kinetic mechanism of elongation factor Ts-catalyzed nucleotide exchange in elongation factor Tu. *Biochemistry* **41**, 162-169 (2002).
126 Bochner, B. R. & Ames, B. N. Complete analysis of cellular nucleotides by two-dimensional thin layer chromatography. *J Biol Chem* **257**, 9759-9769 (1982).
127 Liao, J. *et al.* Two conformational states of Ras GTPase exhibit differential GTP-binding kinetics. *Biochem Biophys Res Commun* **369**, 327-332 (2008).
128 Ye, M. *et al.* Crystal structure of M-Ras reveals a GTP-bound "off" state conformation of Ras family small GTPases. *J Biol Chem* **280**, 31267-31275 (2005).
129 Spoerner, M. *et al.* Conformational states of Ras complexed with the GTP analogue GppNHp or GppCH2p: implications for the interaction with effector proteins. *Biochemistry* **44**, 2225-2236, doi:10.1021/bi0488000 (2005).
130 Higashijima, T., Ferguson, K. M., Sternweis, P. C., Smigel, M. D. & Gilman, A. G. Effects of Mg2+ and the beta gamma-subunit complex on the interactions of guanine nucleotides with G proteins. *J Biol Chem* **262**, 762-766 (1987).
131 Higashijima, T. *et al.* The effect of activating ligands on the intrinsic fluorescence of guanine nucleotide-binding regulatory proteins. *J Biol Chem* **262**, 752-756 (1987).
132 Feuerstein, J., Goody, R. S. & Wittinghofer, A. Preparation and characterization of nucleotide-free and metal ion-free p21 "apoprotein". *J Biol Chem* **262**, 8455-8458 (1987).
133 Feuerstein, J., Kalbitzer, H. R., John, J., Goody, R. S. & Wittinghofer, A. Characterisation of the metal-ion-GDP complex at the active sites of transforming and nontransforming p21 proteins by observation of the 17O-Mn superhyperfine coupling and by kinetic methods. *Eur J Biochem* **162**, 49-55 (1987).
134 Buckstein, M. H., He, J. & Rubin, H. Characterization of nucleotide pools as a function of physiological state in *Escherichia coli*. *J Bacteriol* **190**, 718-726 (2008).

Appendix 1. Supplemental Figures

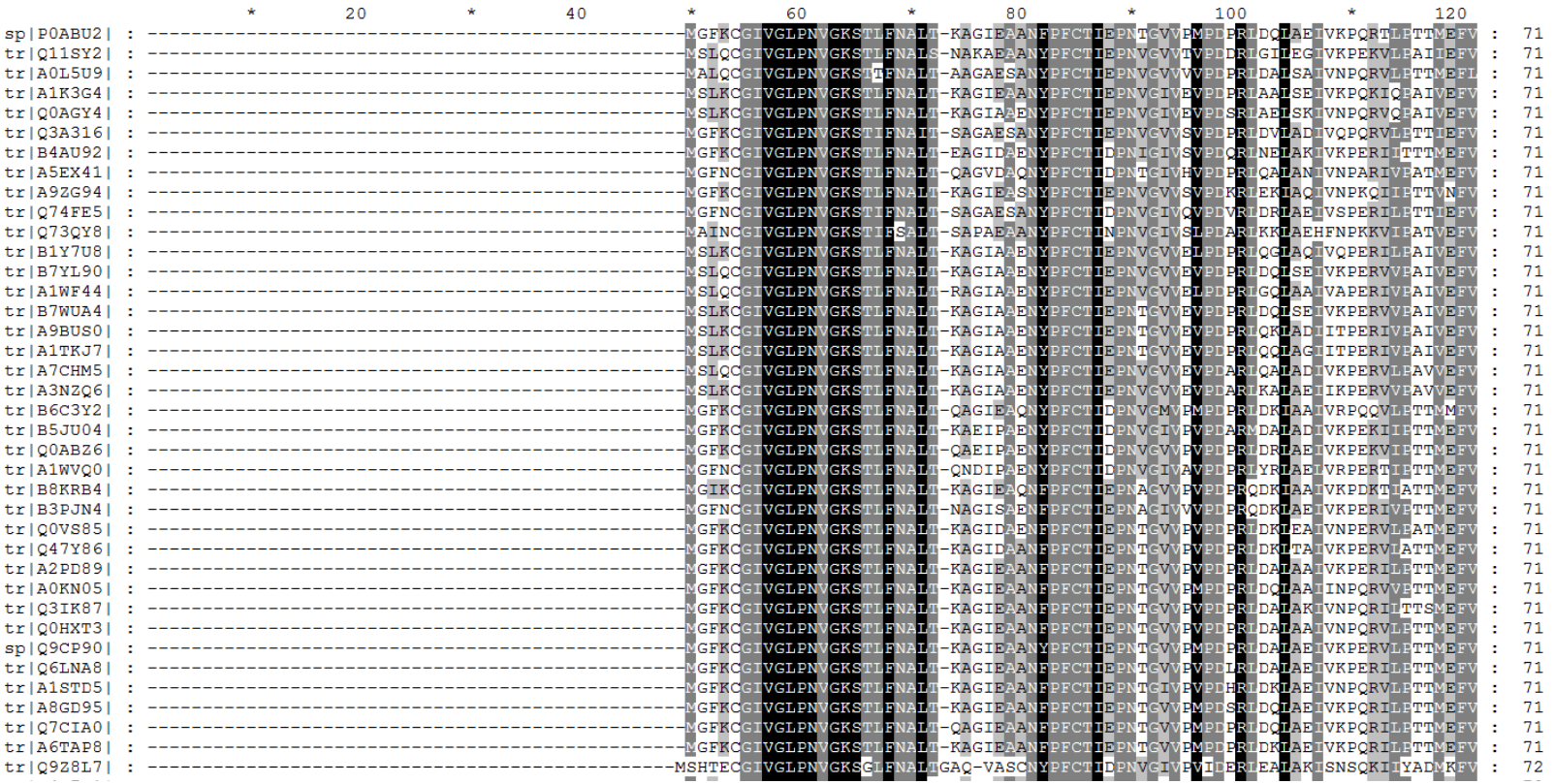
Table A.1 Accession numbers and organisms used to align YchF primary sequence Accession numbers were obtained from the UniProt and Entrez Gene databases. Residue at positions 102 (76% conserved), 114 (98% conserved) and 204 (97% conserved) are shown.

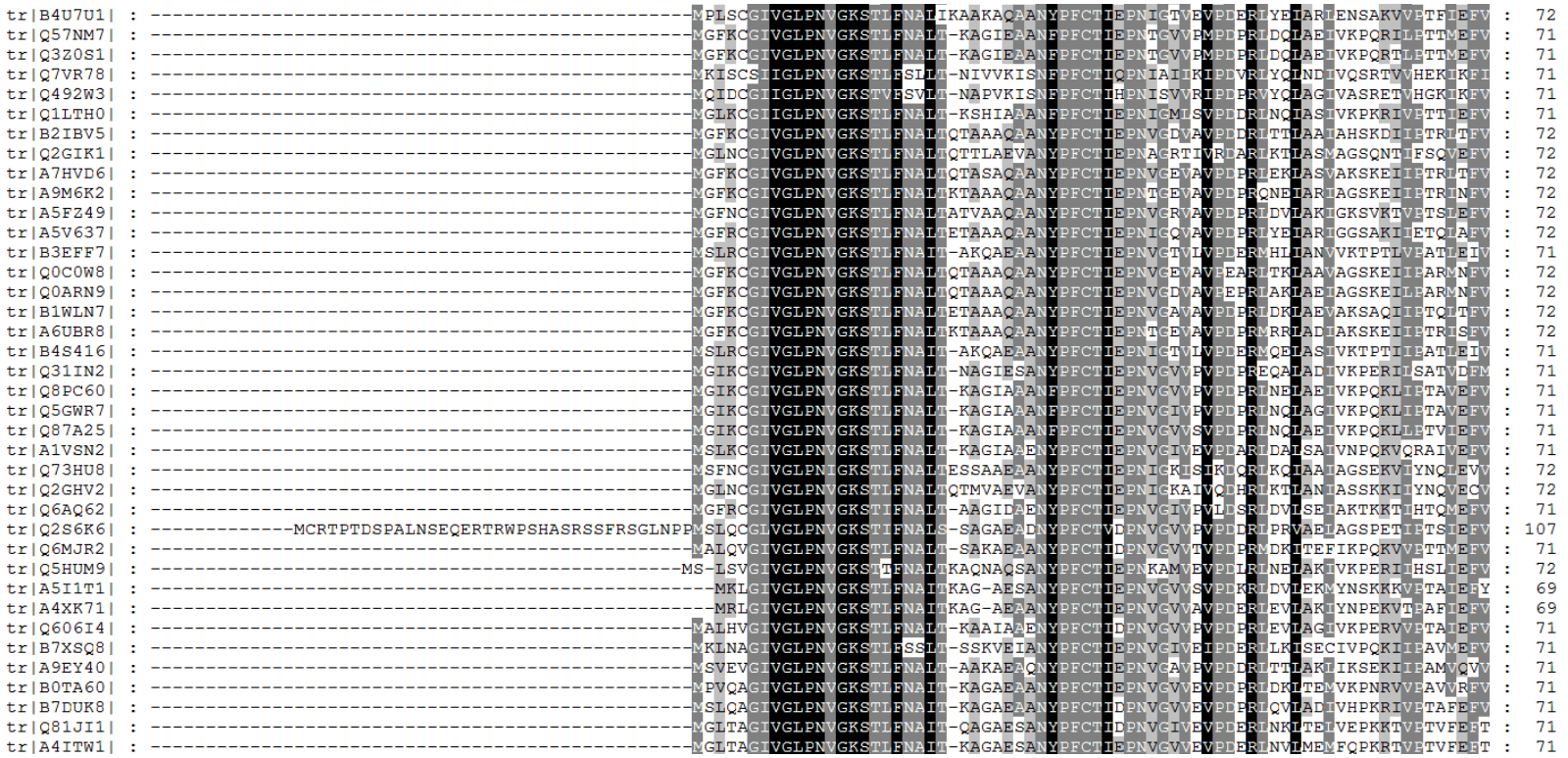
Accession Number	Organism	Residue at position 102	Residue at position 114	Residue at position 204
P8ABU2	*Escherichia coli*	His	His	Tyr
Q11SY2	*Cytophaga hutchinsonii*	His	His	Tyr
A0L5U9	*Magnetococcus sp.*	His	His	Tyr
A1K3G4	*Azoarcus sp.*	His	His	Tyr
Q0AGY4	*Nitrosomonas eutropha*	Asn	His	Tyr
Q3A316	*Pelobacter carbinolicus*	Asn	His	Tyr
B4AU92	*Francisella novicida*	His	His	Tyr
A5EX41	*Dichelobacter nodosus*	Gln	His	Tyr
A9ZG94	*Coxiella burnetii*	His	His	Tyr
Q73QY8	*Treponema denticola*	His	His	Tyr
B1Y7U8	*Leptothrix cholodnii*	Asn	His	Tyr
B7YL90	*Variovorax paradoxus*	Asn	His	Tyr
A1W744	*Verminephrobacter eiseniae*	Asn	His	Tyr
B7WUA4	*Comamonas testosteroni*	Asn	His	Tyr
A9BUS0	*Delftia acidovorans*	Asn	His	Tyr
A1TKJ7	*Acidovorax citrulli*	Asn	His	Tyr
A7CHM5	*Ralstonia pickettii*	His	His	Tyr
A3NZQ6	*Burkholderia pseudomallei*	His	His	Tyr
B6C3Y2	*Nitrosococcus oceani*	His	His	Tyr
B5JU04	*Gamma proteobacterium* ATC5015	His	His	Tyr
Q0ABZ6	*Alkalilimnicola ehrlichei*	Gln	His	Tyr
A1WVQ0	*Halorhodospira halophila*	Met	His	Tyr
B8KRB4	*Gamma proteobacterium* NORS1-B	His	His	Tyr
B3PJN4	*Cellvibrio japonicus*	His	His	Tyr
Q0VS85	*Alcanivorax borkumensis*	His	His	Tyr
Q47Y86	*Colwellia psycherythraea*	His	His	Tyr
A2PD89	*Vibrio cholerae*	His	His	Tyr
A0KN05	*Aeromonas hydrophila*	His	His	Tyr
Q3IK87	*Pseudoalteromonas haloplanktis*	His	His	Tyr
Q0HXT3	*Shewanella sp.*	His	His	Tyr

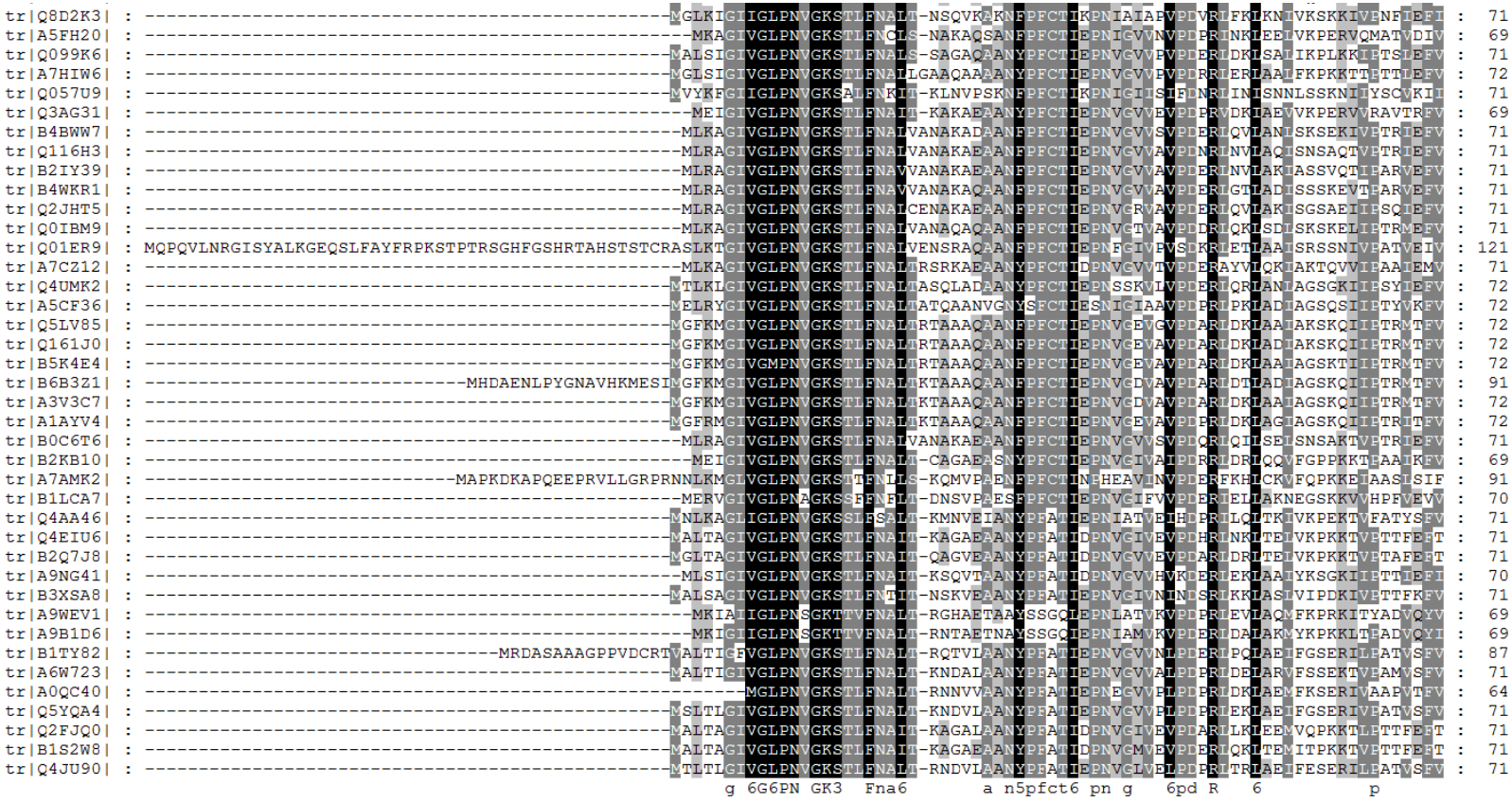
Q9CP90	*Pasteurella multocida*	His	His	Tyr
Q6LNA8	*Photobacterium profundum*	His	His	Tyr
A1STD5	*Psychromonas ingrahamii*	His	His	Tyr
A8GD95	*Serratia proteamaculans*	His	His	Tyr
Q7CIA0	*Yersinia pestis*	His	His	Tyr
A6TAP8	*Klebsiella pneumoniae*	His	His	Tyr
Q9Z8L7	*Chlamydia pneumoniae*	His	His	Tyr
B4U7U1	*Hydrogenobaculum sp.*	Gln	His	Tyr
Q57NM7	*Salmonella choleraesuis*	His	His	Tyr
Q3Z0S1	*Shigella sonnei*	His	His	Tyr
Q7VR78	*Blochmania floridanus*	His	His	Tyr
Q492W3	*Blochmannia pennsylvanicus*	His	His	Tyr
QILTH0	*Baumannia cicadellinicola*	His	His	Tyr
B2IBV5	*Beijerinckia indica*	His	His	Tyr
Q2GIK1	*Anaplasma phagocytophilum*	His	His	Tyr
A7HVD6	*Parvibaculum lavamentivorans*	Tyr	His	Tyr
A9M6K2	*Brucella canis*	His	His	Tyr
A5FZ49	*Acidiphilium cryptum*	His	His	Tyr
A5V637	*Sphingomonas wittichii*	His	His	Tyr
B3EFF7	*Chlorobium limicola*	His	His	Tyr
Q0C0W8	*Hyphomonas neptunium*	Tyr	His	Tyr
Q0ARN9	*Maricaulis maris*	Tyr	His	Tyr
B1WLN7	*Oligotropha carboxidovorans*	His	His	Tyr
A6UBR8	*Sinorhizobium medicae*	His	His	Tyr
B4S416	*Prosthecochloris aestuarii*	His	His	Tyr
Q31IN2	*Thiomicrospira crunogena*	Gln	His	Tyr
Q8PC60	*Xanthomonas campestris*	His	His	Tyr
Q5GWR7	*Xanthomonas oryzae*	His	His	Tyr
Q87A25	*Xylella fastidiosa*	His	His	Tyr
A1VSN2	*Polaromonas naptnthalenivorans*	Asn	His	Tyr
Q73HU8	*Wolbachia pipentis*	His	His	Tyr
Q2GHV2	*Ehrlichia chaffeensis*	His	His	Tyr
Q6AQ62	*Desulfotaelea psychrophile*	His	His	Tyr
Q2S6K6	*Salinibacter ruber*	His	His	Tyr
Q6MJR2	*Bdellovibrio bacteriovorus*	His	His	Tyr
Q5HUM9	*Campylobacter jejuni*	His	His	Tyr

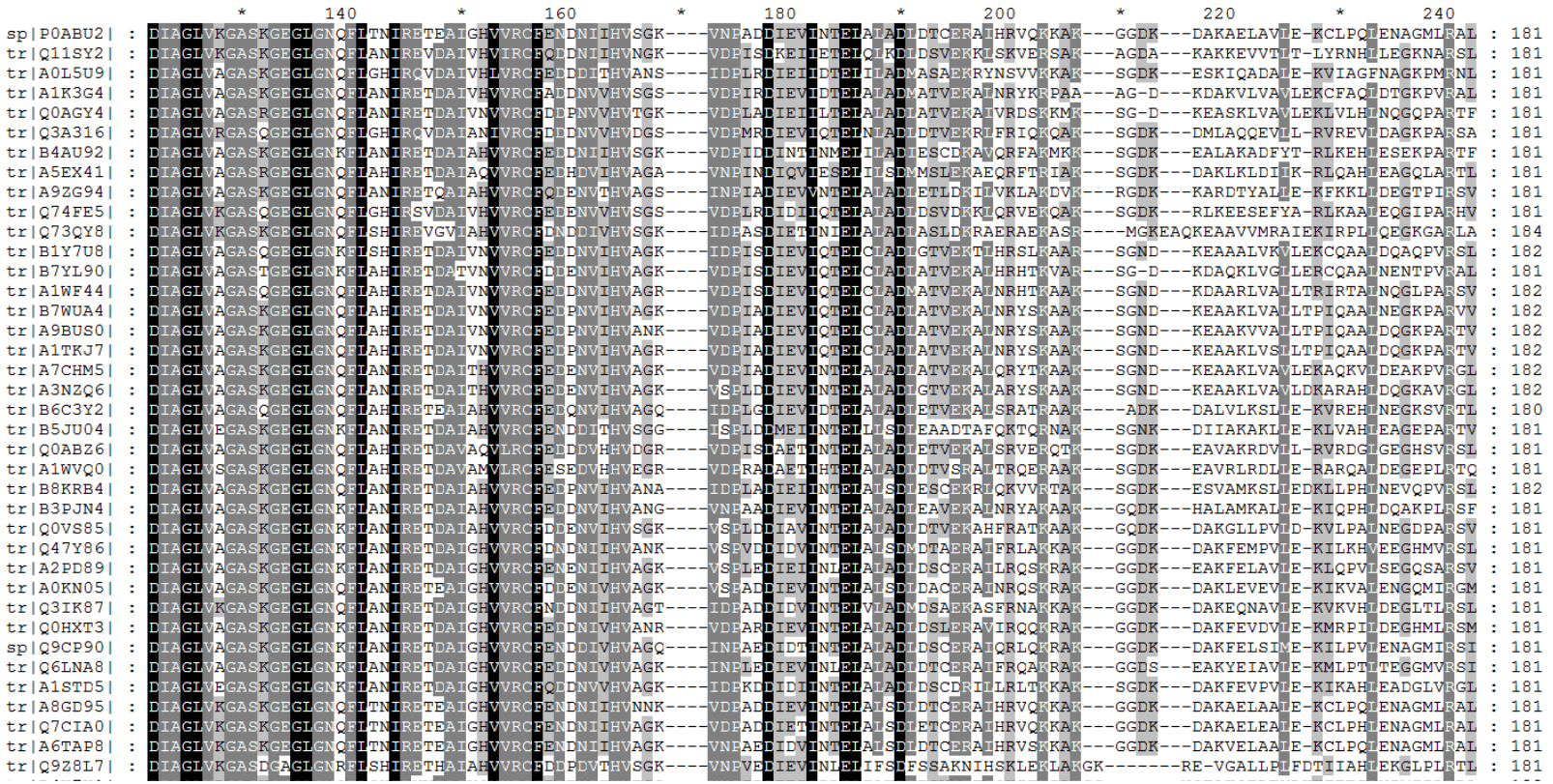
A5I1T1	*Clostridium botulinum*	His	His	Tyr
A4XK71	*Caldicellulosiruptor saccharolyticus*	His	His	Tyr
Q606I4	*Methylococcus capsulatus*	His	His	Tyr
B7XSQ8	*Borrelia garinii*	His	His	Tyr
A9EY40	*Sorangium cellulosum*	His	His	Tyr
B0TA60	*Heliobacterium modesticaldum*	His	His	Tyr
B7DUK8	*Alicyclobacillus acidocaldarius*	His	His	Tyr
Q81JI1	*Bacillus anthracis*	Gln	His	Tyr
A4ITW1	*Geobacillus thermodenitrificans*	Gln	His	Tyr
Q8D2K3	*Wigglesworthia glossinidia*	His	Lys	Tyr
A5FH20	*Flavobacterium johnsoniae*	His	His	Tyr
Q099K6	*Stigmatella aurantiaca*	His	His	Tyr
A7HIW6	*Anaeromyxobacter sp.*	His	His	Tyr
Q057U9	*Buchnera aphidicola*	His	His	Tyr
Q3AG31	*Carboxydothermus hydrogenoformans*	His	His	Tyr
B4BWW7	*Cyanothece sp.*	His	His	Tyr
Q116H3	*Trichodesmium erythraeum*	Gln	His	Tyr
B2IY39	*Nostoc puntiforme*	His	His	Tyr
B4WKR1	*Synechococcus sp*. PCC7445	His	His	Tyr
Q2JHT5	*Synechococcus sp.* JA-2-3B'aL2-B	His	His	Tyr
Q0IBM9	*Synechococcus sp. CC9311*	His	His	Tyr
Q01ER9	*Ostreococcus tauri*	His	His	Tyr
A7CZ12	*Ôpitutaceae bacterium*	Gln	His	Tyr
Q4UMK2	*Rickettsia felis*	His	His	Tyr
A5CF36	*Orientia tsutsugamushi*	His	His	Tyr
Q5LV85	*Ruegeria pomeroyi*	His	His	Tyr
Q161J0	*Roseobacter denitificians*	His	His	Tyr
B5K4E4	*Octadecabacter arcticus*	His	His	Tyr
B6B3Z1	*Rhodobacteraeceae bacterium*	His	His	Tyr
A3V3C7	*Loktanella vestfoldensis*	His	His	Tyr
A1AYV4	*Paracoccus denitrificans*	His	His	Tyr
B0C6T6	*Acaryochloris marina*	His	His	Tyr
B2KB10	*Elusimicrobium mintum*	His	His	Tyr

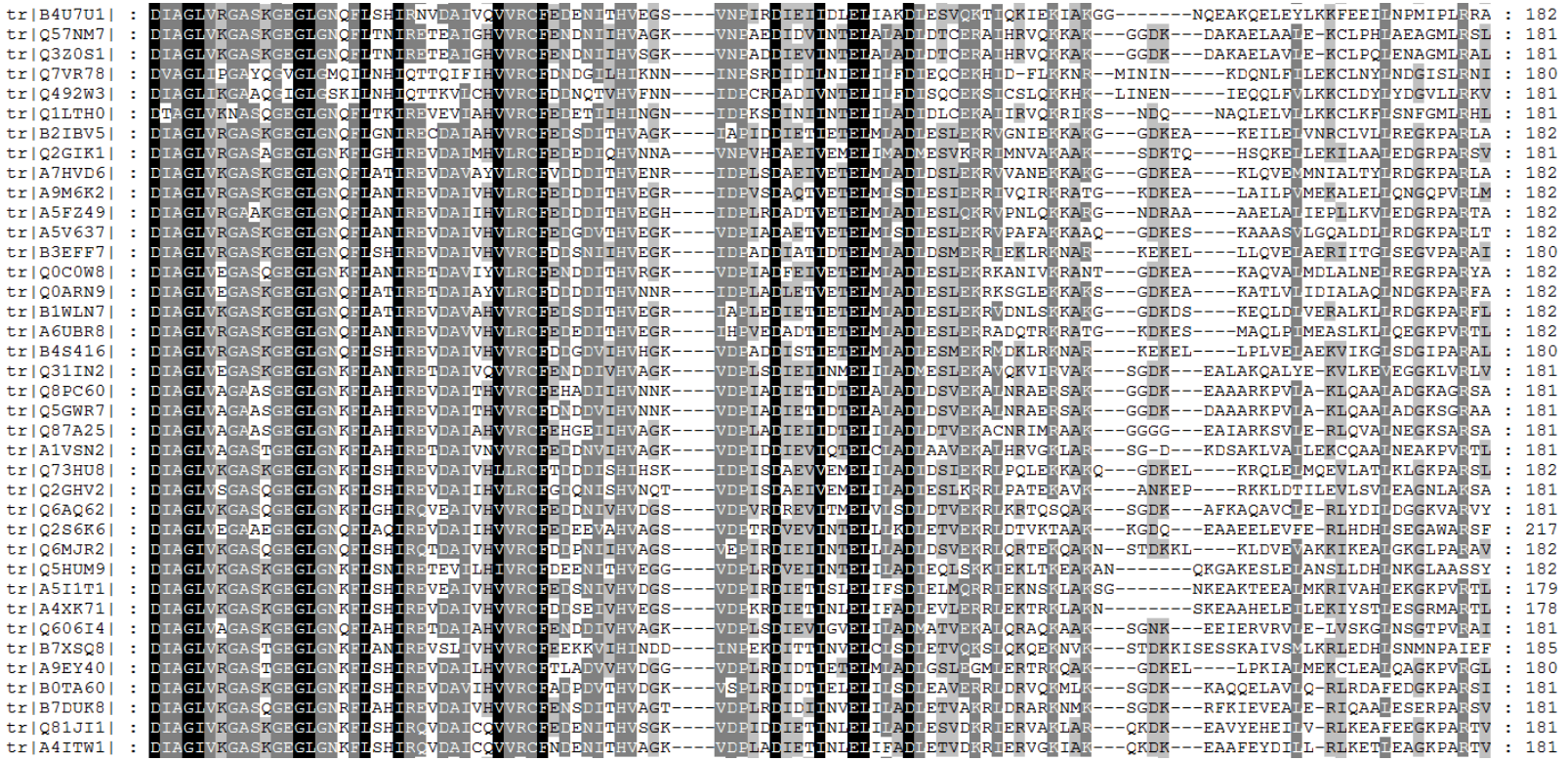
A7AMK2	*Babesia bovis*	His	His	Tyr
B1LCA7	*Thermatoga sp.*	His	His	Leu
Q4AA46	*Mycoplasma hyopneumoniae*	His	His	Tyr
Q4EIU6	*Listeria monocytogenes*	His	His	Tyr
B2Q7J8	*Exiguobacterium sp.*	Gln	His	Tyr
A9NG41	*Acholeplasma laidlawii*	His	His	Tyr
B3XSA8	*Ureaplasma urealyticum*	His	His	Tyr
A9WEV1	*Chloroflexus aurantiacus*	His	His	Ile
A9B1D6	*Herpetosiphon aurantiacus*	His	His	Val
B1TY82	*Mirococcus luteus*	Gln	His	Tyr
A6W723	*Kineococcus radiotolerans*	Gln	His	Tyr
A0QC40	*Mycobacterium avium*	Gln	His	Tyr
Q5YQA4	*Norcardia farcinica*	Gln	His	Tyr
Q2FJQ0	*Staphylococcus aureus*	Gln	His	Tyr
B1S2W8	*Streptococcus pneumoniae*	His	Arg	Tyr
Q4JU90	*Corynebacterium jeikeium*	Gln	His	Tyr

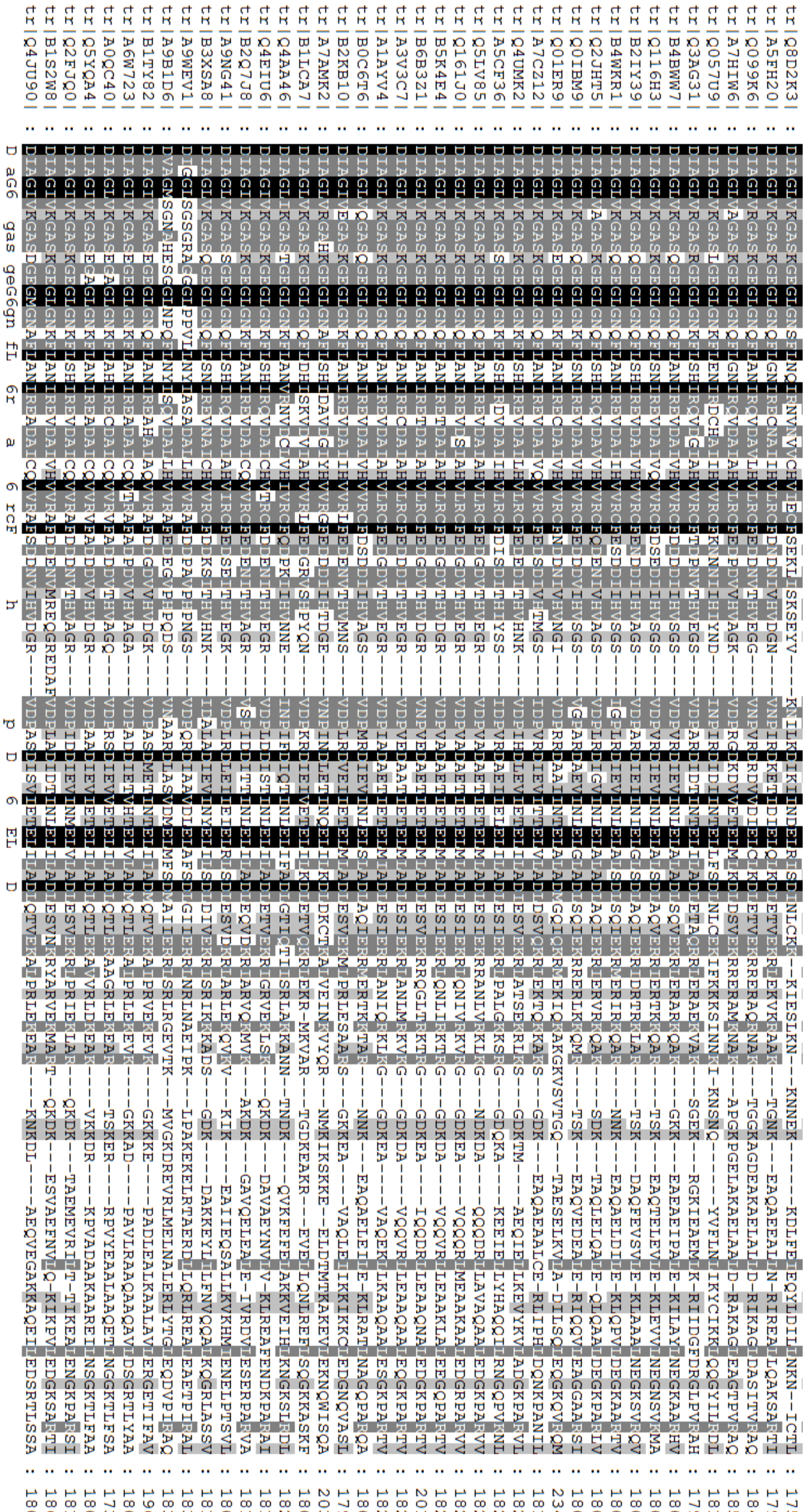


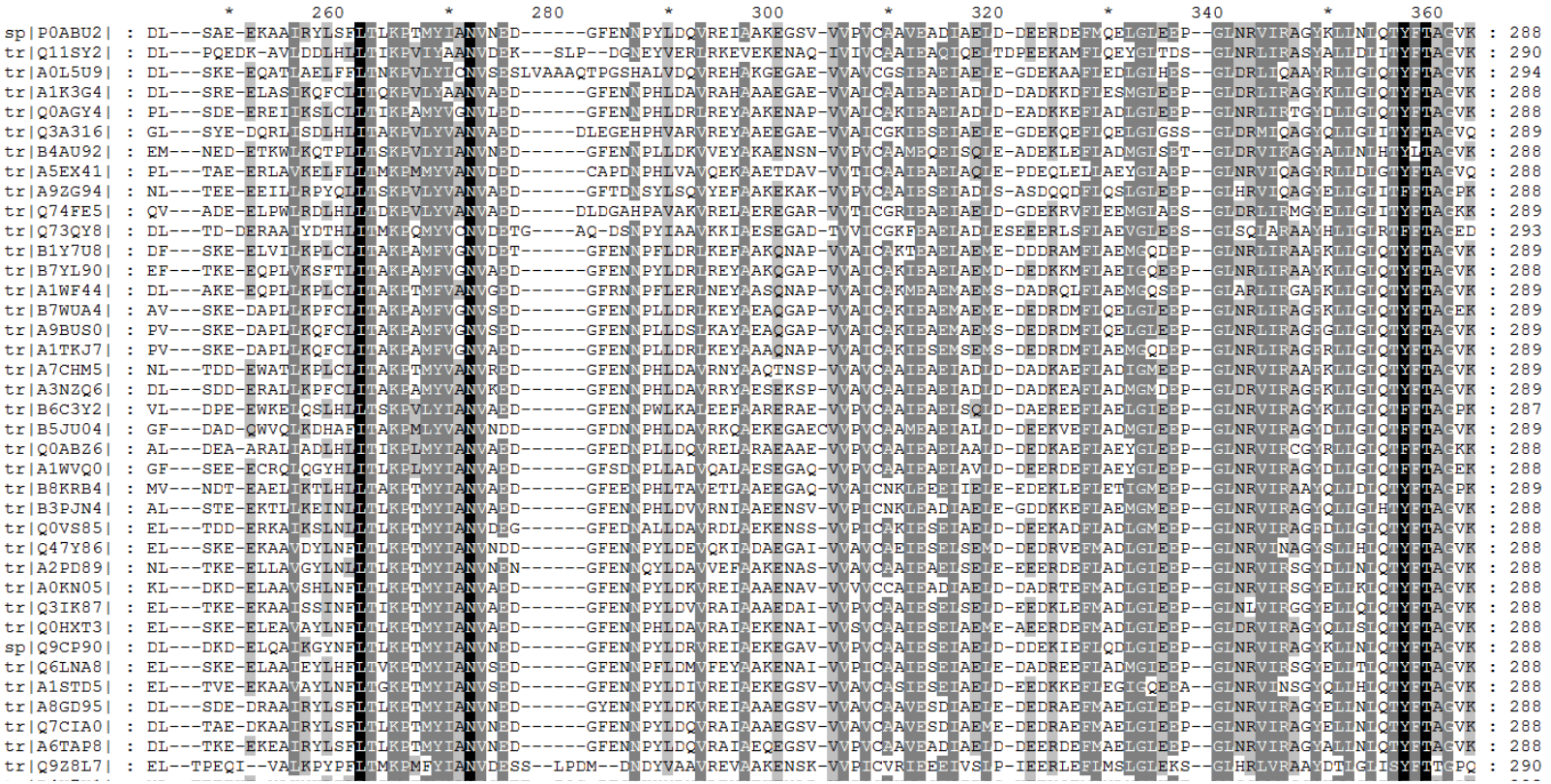


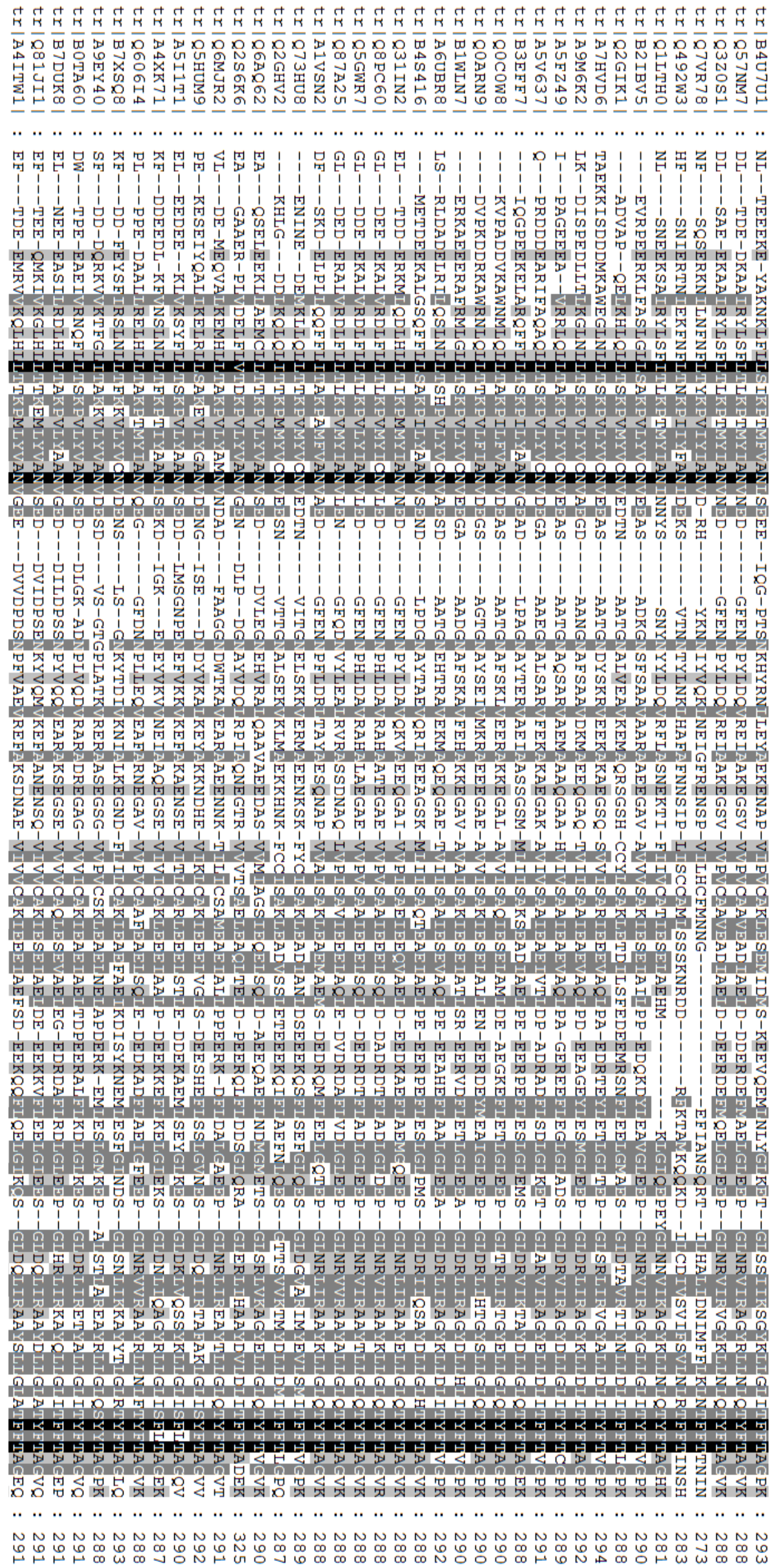


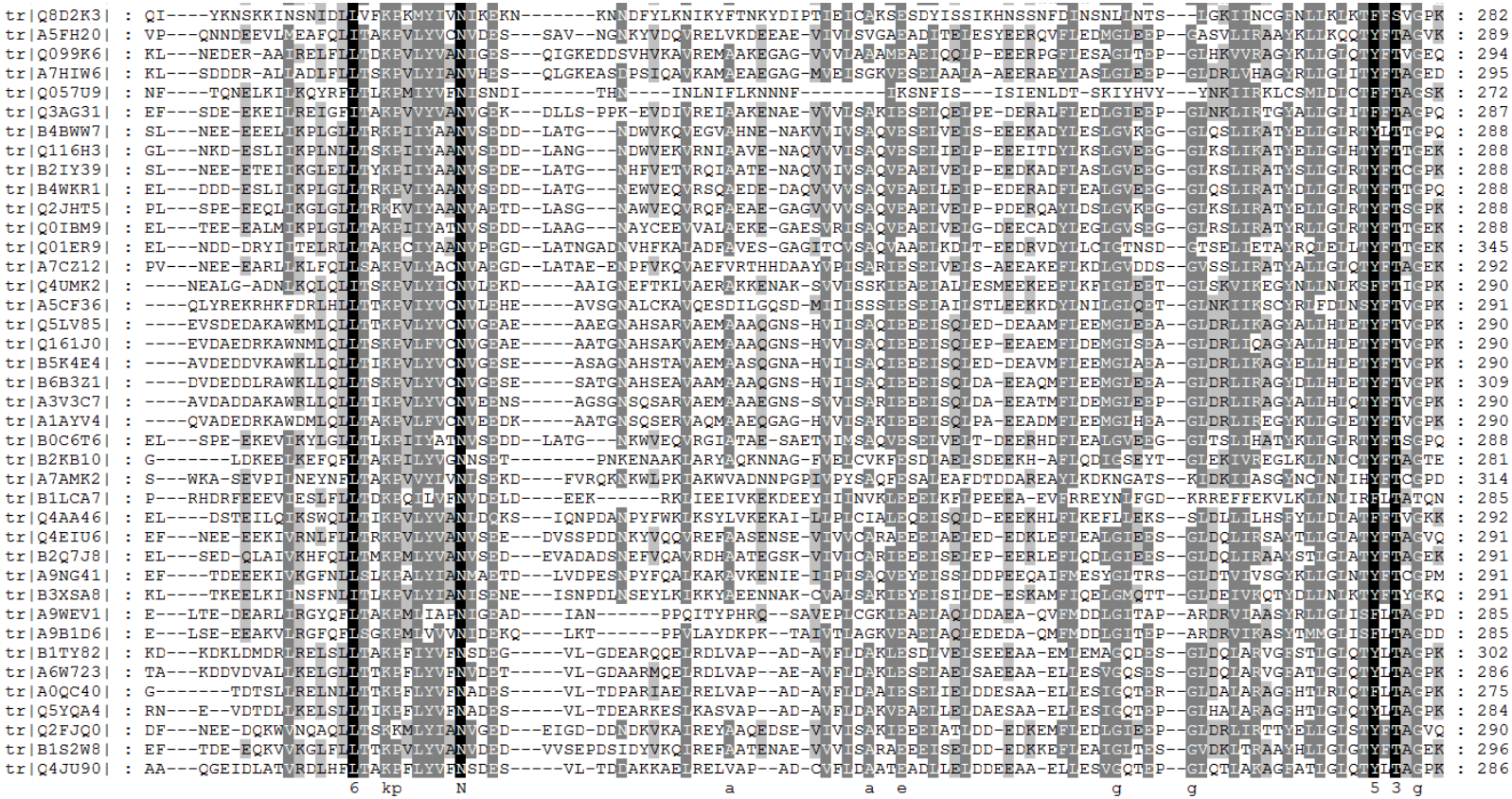

* 140 * 160 * 180 * 200 * 220 * 240
sp|P0ABU2| : 181
tr|Q11SY2| : 181
tr|A0L5U9| : 181
tr|A1K3G4| : 181
tr|Q0AGY4| : 181
tr|Q3A316| : 181
tr|B4AU92| : 181
tr|A5EX41| : 181
tr|A9ZG94| : 181
tr|Q74FE5| : 181
tr|Q73QY8| : 184
tr|B1Y7U8| : 182
tr|B7YL90| : 181
tr|A1WF44| : 182
tr|B7WUA4| : 182
tr|A9BUS0| : 182
tr|A1TKJ7| : 182
tr|A7CHM5| : 182
tr|A3NZQ6| : 182
tr|B6C3Y2| : 180
tr|B5JU04| : 181
tr|Q0ABZ6| : 181
tr|A1WVQ0| : 181
tr|B8KRB4| : 182
tr|B3PJN4| : 181
tr|Q0VS85| : 181
tr|Q47Y86| : 181
tr|A2PD89| : 181
tr|A0KN05| : 181
tr|Q3IK87| : 181
tr|Q0HXT3| : 181
sp|Q9CP90| : 181
tr|Q6LNA8| : 181
tr|A1STD5| : 181
tr|A8GD95| : 181
tr|Q7CIA0| : 181
tr|A6TAP8| : 181
tr|Q9Z8L7| : 181

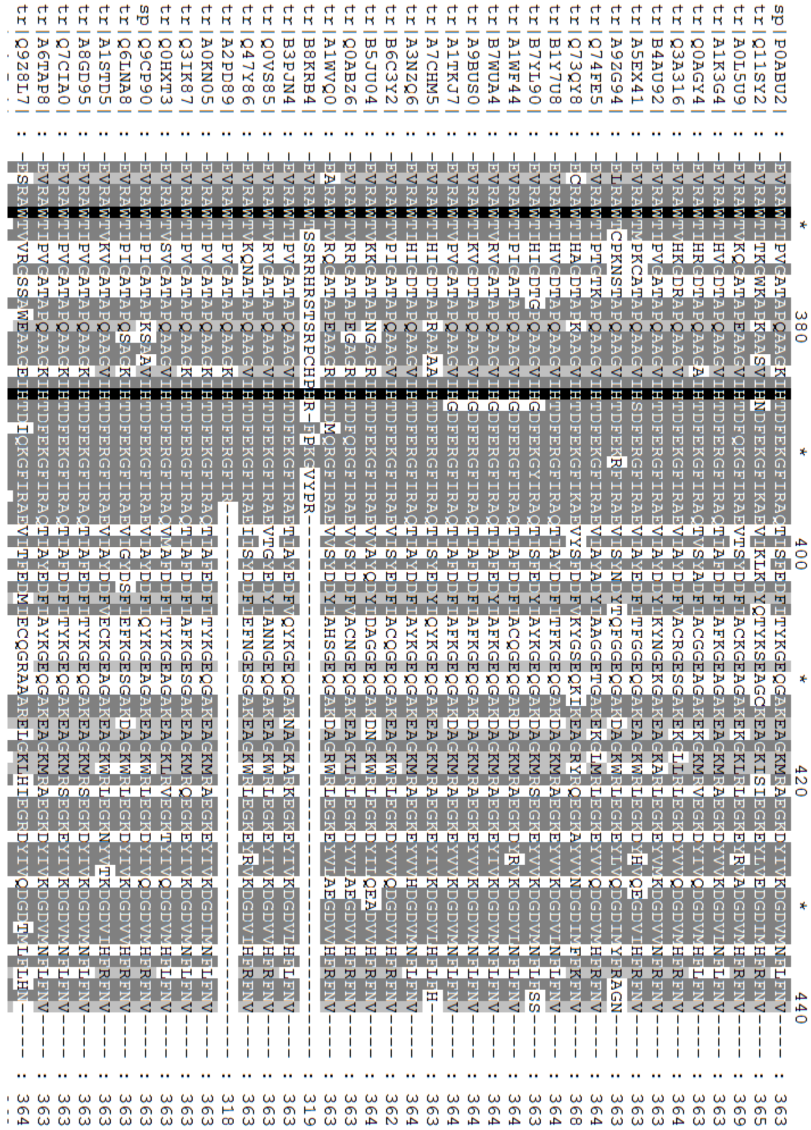

tr|B4U7U1| : DIAGLVRGASKGEGLGNQFLSHIRNVDAIVQVVRGFEDENITHVEGS----VNPIRLDIEIIDLELIAKDLESVQKTIQKIEKIAKGG-------NQEAKQELEYLKKFEEIINPMIPLIRA : 182
tr|Q57NM7| : DIAGLVKGASKGEGLGNQFLTNIRETEAIGHVVRGFENDNIIHVAGK----VNPADLDVINTELAIADLDTCERAIHRVQIKAK---GGDK---DAKAELAAIE-KCLPHIAEAGMLISL : 181
tr|Q3Z0S1| : DIAGLVKGASKGEGLGNQFLTNIRETEAIGHVVRGFENDNIIHVSGK----VNPADLDIEVINTELAIADLDTCERAIHRVQIKAK---GGDK---DAKAELAVIE-KCLPQIENAGMLIAL : 181
tr|Q7VR78| : DVAGLIPGAYQGVGLGMQILNHIQTTQIFIHVVRGFDNDGILHIKNN----INPSRDIDIINIELIIFDIEQCEKHID-FLIKNR--MININ-----KDQNLFILEKCLNYINDGISLINI : 180
tr|Q492W3| : DIAGLIKGAAQGIGLGSKILNHIQTTKVICHVVRGFDDNQTVHVFNN----IDPCRDADIVNTELIIFDISQCEKSICSLQIKHK--LINEN-----IEQQLFVLKKCLDYIYDGVLLIKV : 181
tr|Q1LTH0| : DTAGLVKNASQGEGLGNQFLTKIREVEVIAHVVRGFEDETIIHINGN----IDPKSDININTNTELIIADIDLCEKAIIRVQIRIKS---NDQ----NAQLELVILKKCLKFISNFGMLIHL : 181
tr|B2IBV5| : DIAGLVRGASKGEGLGNQFLGNIRECDAIAHVVRGFEDSDITHVAGK----IAPIDLDIETIETEIMIADIESLEKRVGNIEKKAKG---GDKEA----KEILEIVNRCLVLIREGKPARLA : 182
tr|Q2GIK1| : DIAGLVRGASAGEGLGNKFLGHIREVDAIMHVLRGFEDEDIQHVNNA----VNPVELAEIVEMELIMADVESVKRRIMNVAIAAK----SDKTQ----HSQKEILEKILAAIEDGRPAISV : 181
tr|A7HVD6| : DIAGLVKGASKGEGLGNQFLATIREVDAVAYVLRGFVDDDITHVENR----IDPLSLAEIVETEIMIADIDSLEKRVVANEKKAKG---GDKEA----KLQVEVMNIALTYIRDGKPARLA : 182
tr|A9M6K2| : DIAGLVRGASKGEGLGNQFLANIREVDAIVHVLRGFEDDDITHVEGR----IDPVSLAQTVETEIMISDIESIERRIVQIRIRATG---KDKEA----LAILPVMEKALELIQNGQPVRLM : 182
tr|A5FZ49| : DIAGLVRGAAKGEGLGNQFLANIREVDAIIHVLRGFEDDDITHVEGH----IDPLRLDADTVETEIMIADIESIQKRVPNLQIKARG---NDRAA----AAELAIIEPLLKVIEDGRPAITA : 182
tr|A5V637| : DIAGLVRGASKGEGLGNQFLANIREVDAIVHVLRGFEDGDVTHVEGK----VDPIALDAETVETEIMISDIESLEKRVPAFAIKAAQ---GDKES----KAAASVLGQALDLIRDGKPAILT : 182
tr|B3EFF7| : DIAGLVRGASKGEGLGNQFLSHIREVDAIVHVVRGFDDSNIIHVEGK----IDPADLIATIDTEIMIADIDSMERRIEKLRINAR----KEKEL----LLQVEIAERIITGLSEGVPARAI : 180
tr|Q0C0W8| : DIAGLVEGASQGEGLGNKFLANIRETDAVIYVLRGFENDDITHVRGK----VDPIALFEIVETEIMIADIESLEKRKANIVIRRANT---GDKEA----KAQVAIMDLALNEIREGRPAFYA : 182
tr|Q0ARN9| : DIAGLVEGASKGEGLGNQFLATIRETDAIAYVLRGFDDDDITHVNNR----IDPLALLETVETEIMIADIESLEKRKSGLEIKAKS---GDKEA----KATLVIIDIALAQINDGKPAIFA : 182
tr|B1WLN7| : DIAGLVKGASKGEGLGNQFLATIREVDAVAHVVRGFEDSDITHVEGR----IAPLEDIETIETEIMIADIESLEKRVDNLSIKAKG---GDKDS----KEQLDIVERALKLIRDGKPAIFL : 182
tr|A6UBR8| : DIAGLVRGASKGEGLGNQFLANIREVDAVVHVLRGFEDEDITHVEGR----IHPVELADTIETEIMIADIESIERRADQTRIRATG---KDKES----MAQLPIMEASLKLIQEGKPVITL : 182
tr|B4S416| : DIAGLVRGASKGEGLGNQFLSHIREVDAIVHVVRGFDDGDVIHVHGK----VDPADLISTIETEIMIADIESMEKRVDKLRINAR----KEKEL----LPLVEIAEKVIKGLSDGIPARAL : 180
tr|Q31IN2| : DIAGLVEGASKGEGLGNKFLANIRETDAIVQVVRGFENDDIVHVAGK----VDPLSLEIEINMELIIADVESLEKAVQKVIEVAK---SGDK---EALAKQALYE-KVLKEVEGGKLVILV : 181
tr|Q8PC60| : DIAGLVAGAASGEGLGNKFLAHIREVDAITHVVRGFEHADIIHVNNK----VDPIALDIETIDTELAIADIDSVEKAINRAERISAK---GGDK---EAAARKPVIA-KLQAALADGKAGRSA : 181
tr|Q5GWR7| : DIAGLVAGAASGEGLGNKFLAHIREVDAITHVVRGFDNDDVIHVNNK----VDPIALDIETIDTELAIADIDSVEKAINRAERISAK---GGDK---DAAARKPVIA-KLQAALADGKSGRAA : 181
tr|Q87A25| : DIAGLVAGAASGEGLGNKFLAHIREVDAIAHVVRGFEHGEIIHVAGK----VDPLALDIETIDTELAIADIDTVEKACNRIMEAAK---GGGG---EAIARKSVIE-RLQVAINEGKSARSA : 181
tr|A1VSN2| : DIAGLVAGASTGEGLGNKFLAHIRETDAIVNVVRGFEDDNVIHVAGK----VDPIDLDIEVIQTEICIADIAAVEKAIHRVGILAR---SG-D---KDSAKLVAILEKCQAAINEAKPVITL : 181
tr|Q73HU8| : DIAGLVKGASKGEGLGNKFLSHIREVDAIVHLLRGFTDDDISHIHSK----IDPISLAEVVEMELIIADIDSIEKRIPQLEIKAKQ---GDKEL----KRQLEIMQEVLATIKLGKPARSL : 182
tr|Q2GHV2| : DIAGLVSGASQGEGLGNKFLSHIREVDAIIHVLRGFGDQNISHVNQT----VDPISLDAEIVEMELIIADIESLKRRIPATEKAVK----ANKEP----RKKLDTILEVLSVIEAGNLAKSA : 181
tr|Q6AQ62| : DIAGLVKGASQGEGLGNKFLGHIRQVEAIVHVVRGFEDDNIVHVDGS----VDPVRIREVITMELVISDIDTVEKRIKRTQSQAK---SGDK---AFKAQAVCIE-RLYDIIDGGKVAIVY : 181
tr|Q2S6K6| : DIAGLVEGAAEGEGLGNQFLAQIREVDAIIHVVRGFEDEEVAHVAGS----VDPTRDVEVINTELILKDIETVEKRIDTVKTAAK---KGDQ---EAAEELEVFE-RLHDHISEGAWARSF : 217
tr|Q6MJR2| : DIAGIVKGASQGEGLGNQFLSHIRQTDAIVHVVRGFDDPNIIHVAGS----VEPIRLDIEIINTELILADIDSVEKRIQRTEKQAKN--STDKKL----KLDVEVAKKIKEAIGKGLPAIAV : 182
tr|Q5HUM9| : DIAGLVKGASKGEGLGNKFLSNIRETEVILHIVRGFDEENITHVEGG----VDPLRLDVEIINTELIADIEQLSKKIEKLTKEAKAN-------QKGAKESLEIANSLLDHINKGLAASSY : 182
tr|A5I1T1| : DIAGLVKGASKGEGLGNKFLSHIREVEAIVHVVRGFEDSNIVHVDGS----VDPIRLDIETISLELIFSDIELMQRRIEKNSKLAKSG--------NKEAKTEEAIMKRIVAHIEKGKPVITL : 179
tr|A4XK71| : DIAGLVKGASKGEGLGNKFLSHIREVDAIVHVVRGFDDSEIVHVEGS----VDPKRDIETINLELIFADIEVLERRIEKTRILAKN--------SKEAAHELEILEKIYSTIESGRMAITL : 178
tr|Q606I4| : DIAGLVAGASKGEGLGNQFLAHIRETDAIAHVVRGFENDDIVHVAGK----VDPLSDIEVIGVELIIADVATVEKAIQRAQIAAK---SGNK---EEIERVRVIE-LVSKGINSGTPVIAI : 181
tr|B7XSQ8| : DIAGLVKGASTGEGLGNKFLANIREVSLIVHVVRGFEEKKVIHINDD----INPEKDITTINVEICISDIETVQKSIQKQEINVK---STDKKISESSKAIVSILKRLEDHISNMNPAIEF : 185
tr|A9EY40| : DIAGLVRGASTGEGLGNKFLSHIREVDAILHVVRGFTLADVVHVDGG----VDPLRDIDTIETEIMIADIGSLEGMIERTRIQAK----GDKEL----LPKIAIMEKCLEAIQAGKPVIGL : 180
tr|B0TA60| : DIAGLVRGASKGEGLGNKFLSHIREVDAVIHVVRGFADPDVTHVDGK----VSPLRDIDTIELELIISDIEAVERRIDRVQIMLK---SGDK---KAQQELAVIQ-RLRDAFEDGKPAISI : 181
tr|B7DUK8| : DIAGLVKGASQGEGLGNRFLAHIREVDAIVHVVRGFENSDITHVAGT----VDPLRLDIDINVELIIADIETVAKRIDRARINMK---SGDK---RFKIEVEAIE-RIQAAIESERPARSV : 181
tr|Q81JI1| : DIAGIVKGASKGEGLGNKFLSHIRQVDAICQVVRGFEDENITHVSGK----VDPIDIETINLELIIADIESVDKRIERVAILAR---QKDK---EAVYEHEIIV-RLKEAFEEGKPAITV : 181
tr|A4ITW1| : DIAGIVKGASKGEGLGNKFLSHIRQVDAICQVVRGFNDENITHVAGK----VDPLADIETINLELIFADIETVDKRIERVGIIAK---QKDK---EAAFEYDILL-RLKETIEAGKPAITV : 181




* 260 * 280 * 300 * 320 * 340 * 360
sp|P0ABU2| : DL---SAE-EKAAIRYLSPLTLKFTMYIANVNED------GFENNPYLDQVREIAAKEGSV-VVPVCAAVEADIAELD-DEERDEFVQELGLEEP--GINRVIFAGYKLLNLQTYFTACVK : 288
tr|Q11SY2| : DL---PQEDK-AVIDDLHLLTIKFVIYAANVDEK---SLP--DGNEYVERIRKEVEKENAQ-IVIVCAAIEAQIQELTDPEEKAMFIQEYGLTDS--GINRLIFASYALLDLIIYFTACVK : 290
tr|A0L5U9| : DL---SKE-EQATIAELFPLINKFVLYLQNVSESLVAAAQTPGSHALVDQVREHIKGEGAE-VVAVCGSIEAEIAELE-GDEKAAFIEDLGLHES--GIDRLIQAAYRLLGLQTYFTACVK : 294
tr|A1K3G4| : DL---SRE-ELASIKQFCLITQKFVLYAANVAED------GFENNPHLDAVRAHIAAEGAE-VVAICAAIEAEIADID-DADKKDFIESMGLEEP--GIDRLIFAGYKLLGLQTYFTACVK : 288
tr|Q0AGY4| : PL---SDE-EREIIKSLCLLTIKFAMYVGNVLED------GFENNPHLDRVREYIAKENAP-VVAICAKIEAELADID-EADKKEFIADLGLEEP--GINRLIFTGYDLLGLQTYFTACVK : 288
tr|Q3A316| : GL---SYE-DQRLISDLHLITAKFVLYVANVAED-----DLEGEHPHVARVREYIAEEGAE-VVAICGKIESEIAELE-GDEKQEFIQELGLGSS--GIDRMIQAGYQLLGLITYFTACVQ : 289
tr|B4AU92| : EM---NED-ETKWIKQTPLLTSKFVLYIANVNED------GFENNPLLDKVVEYIKAENSN-VVPVCAAMEQEISQLE-ADEKLEFIADMGLSET--GIDRVIKAGYALLNLHIYLTACVK : 288
tr|A5EX41| : PL---TAE-ERLAVKELFLLTMKFMMYVANVDED------CAPDNPHLVAVQEKIAETDAV-VVTICAAIEAELAQLE-PDEQLELIAEYGIAEP--GINRVIQAGYRLLDLGIYFTACVQ : 288
tr|A9ZG94| : NL---TEE-EEILIRPYQLLTSKFVLYVANVAED------GFTDNSYLSQVYEFIAKEKAK-VVPVCAAIESEIADIS-ASDQQDFIQSLGLEEP--GIHRVIQAGYELLGLITFFTAGPK : 288
tr|Q74FE5| : QV---ADE-ELPWIRDLHLLTDKFVLYVANVAED-----DLDGAHPAVAKVRELAEREGAR-VVTICGRIEAEIAELD-GDEKRVFIEEMGIAES--GIDRLIFMGYELLGLITYFTACKK : 289
tr|Q73QY8| : DL---TD-DERAAIYDTHLITMKFCMYVGNVDETG----AQ-DSNPYIAAVKKIAESEGAD-TVVICGKFDAELADIESEEERLSFIAEVGIEES--GISQLAEAAYHLIGLRTFFTAGED : 293
tr|B1Y7U8| : DF---SKE-ELVIIKPLCLITAKFAMFVGNVDET------GFENNPFLDRIKEFIAKQNAP-VVAICAKTDAELAELE-DDDRAMFIAEMGQDEP--GINRLIFAAFKLLGLQTYFTACVK : 289
tr|B7YL90| : EF---TKE-EQPLVKSFTLITAKFAMFVGNVADD------GFENNPYLDRIREYIAKQGAP-VVAICAKIDAELAELD-DEDKKMFIAEIGQDEP--GINRLIFAAYKLLGLQTYFTACVK : 288
tr|A1WF44| : DL---AKE-EQPLIKPLCLITAKFTMFVANVGED------GFRNNPFLERINEYIASQNAP-VVAICAKMDAEMAELS-DADRQLFIAEMGQSEP--GIARLIFGAFKLLGLQTYFTACVK : 289
tr|B7WUA4| : AV---SKE-DAPLIKPFCLITAKFAMFVGNVSED------GFENNPLLDRIKEYIEAQGAP-VVAICAKIDAEMAELE-DEDRDMFIQELGLEEP--GINRLIFAGFKLLGLQTYFTAGEK : 289
tr|A9BUS0| : PV---SKE-DAPLIKQFCLITAKFAMFVGNVSED------GFENNPLLDSIKAYIEAQGAP-VVAICAKIDAEMAELS-DEDRDMFIQELGLEEP--GINRLIFAGFGLLGLQTYFTACVK : 289
tr|A1TKJ7| : PV---SKE-DAPLIKQFCLITAKFAMFVGNVADD------GFENNPLLDRIKEYIAAQNAP-VVAICAKIESEMSELS-DEDRDMFIAEMGQDEP--GINRLIFAGFRLLGLQTYFTACVK : 289
tr|A7CHM5| : NL---TDD-EWATIKPLCLITAKFTMYVANVRED------GFENNPHLDAVRNYIAQTNSP-VVAVCAAIEAEIADID-DADKAEFIADIGMEEP--GINRVIFAAFKLLGLQTYFTACVK : 289
tr|A3NZQ6| : DL---SDD-ERALIKPFCLITAKFAMYVANVKED------GFENNPHLDAVRRYIESEKSP-VVAVCAAIEAEIADID-DADKEAFIADMGMDEP--GIDRVIFAGFKLLGLQTYFTACVK : 289
tr|B6C3Y2| : VL---DPE-EWKEIQSLHLLTSKFVLYIANVAED------GFENNPWLKAIEEFIARERAE-VVPVCAAIEAELSQID-DAEREEFIAELGIEEP--GINRVIFAGYKLLGLQTFFTAGPK : 287
tr|B5JU04| : GF---DAD-QWVQIKDHAEITAKFMLYVANVNDD------GFDNNPHLDAVRKQIEKEGAECVVPVCAAMEAEIALID-DEEKVEFIADMGLEEP--GINRVIFAGYDLLGLQTFFTACVK : 289
tr|Q0ABZ6| : AL---DEA-ARALIADLHLITIKFIMYIANVAED------GFEDNPLLDQVRELARAEAAE-VVPVCAAIEAELAALD-DEDKAEFIAEYGLEEP--GINRVIRCGYRLLGLQTFFTACKK : 288
tr|A1WVQ0| : GF---SEE-ECRQIQGYHLITLKFLMYIANVAED------GFSDNPLLADVQALIESEGAQ-VVPVCAAIEAELAVID-DEERDEFIAEYGLEEP--GINRVIFAGYDLLGLQTFFTACEK : 288
tr|B8KRB4| : MV---NDT-EAELIKTLHLLTAKFIMYIANVAED------GFEENPHLTAVETLIAEEGAQ-VVAICNKLDEIIEIE-EDEKLEFIETGMEEP--GINRVIFAAYQLLDLQTYFTAGPK : 289
tr|B3PJN4| : AL---STE-EKTLIKEINLLTLKFTMYIANVAED------GFENNPHLDVVRNIAEENSV-VVPICNKLDADIAEIE-GDDKKEFIAEMGMEEP--GINRVIFAGYQLLGLHTYFTACVK : 288
tr|Q0VS85| : EL---TDD-ERKAIKSLNLLTLKFTMYIANVDEG------GFEDNALLDAVRDLAEKENSS-VVPICAKIESEIAEID-DEEKADFIADLGVEEP--GINRVIFAGFDLLGLQTYFTACVK : 288
tr|Q47Y86| : EL---SKE-EKAAVDYLNELTLKFTMYIANVNDD------GFENNPYLDEVQKIADAEGAI-VVAVCAEIESELSEVD-DEDRVEFMADLGIEEP--GINRVINAGYSLLHLQTYFTACVK : 288
tr|A2PD89| : NL---TKE-ELLAVGYLNLLTLKFTMYIANVNEN------GFENNQYLDAVVEFIAKENAS-VVAVCAAIDAELSEIE-EEERDEFIADLGIEEP--GINRVIFSGYDLINLQTYFTACVK : 288
tr|A0KN05| : KL---DKD-ELAAVSHLNELTLKFTMYIANVAED------GFENNPYLDKVREIAAENAV-VVVVCCAIEADIAEID-DADRTEFMADLGIEEP--GINRVIFSGYELLKLQTYFTACVK : 288
tr|Q3IK87| : EL---TKE-EKAAISSINELTIKFTMYIANVAED------GFENNPYLDVVRAIAAEDAI-VVPVCAAIESELSEID-EEDKLEFMADLGIEEP--GINLVIRGGYELLQLQTYFTACVK : 288
tr|Q0HXT3| : EL---SKE-ELEAVAYLNELTLKFTMYIANVAED------GFENNPHLDAVRAIAEKENAI-VVSVCAAIESEIAEVE-AEERDEFMADLGIEEP--GIDRVIFAGYQLLSLQTYFTACVK : 288
sp|Q9CP90| : DL---DKD-ELQAIKGYNELTLKFTMYIANVNED------GFENNPYLDRVREIAEKEGAV-VVPVCAAIESEIAEID-DDEKIEFIQDLGIEEP--GINRVIFAGYKLLNLQTYFTACVK : 288
tr|Q6LNA8| : EL---SKE-ELAAIEYLHELTVKFTMYIANVSED------GFENNPFLDMVFEYIAKENAI-VVPICAAIESEIAEIE-DADREEFIADMGIEEP--GINRVIFSGYELLTLQTYFTACVK : 288
tr|A1STD5| : EL---TVE-EKAAVAYLNELTGKFTMYIANVSED------GFENNPYLDIVREIAEKEGSV-VVAVCASIESEIAEID-EEDKKEFIEGIQEEA--GINRVINSGYQLLHLQTYFTACVK : 288
tr|A8GD95| : DL---SDE-DRAAIRYLSELTLKFTMYIANVNED------GYENNPYLDKVREIAAEGSV-VVAVCAAVESDIAEIE-DEDRAEFMAELGIEEP--GINRVIFAGYELLNLQTYFTACVK : 288
tr|Q7CIA0| : DL---TAE-DKAAIRYLSELTLKFTMYIANVNED------GFENNPYLDQVRAIAAEGSV-VVAVCAAVESDIAEVE-DEDRAEFMAELGIEEP--GINRVIFAGYELLNLQTYFTACVK : 288
tr|A6TAP8| : DL---TKE-EKEAIRYLSELTLKFTMYIANVNED------GFENNPYLDQVRAIAEQEGSV-VVPVCAAVEADIAEID-DEERDEFMAELGIEEP--GINRVIFAGYALLNLQTYFTACVK : 288
tr|Q9Z8L7| : EL--TPEQI--VALKPYPELTMKFMFYIANVDESS--LPDM--DNDYVAAVREVIAKENSK-VVPICVRIEEEIVSIP-IEERLEFIMSLGIEKS--GIHRLVFAAYDTLGLISYFTTCPQ : 290







```
tr|B4U7U1| : -ESRAWAIAKDTKAQKAAGKIHSDIERGFIAAEVINYDDYIKVKSMTKAKEQGLIRLEGKDYIVKDGDIIYFRFNV---- : 368
tr|Q57NM7| : -EVRAWTIPVGATAPQAAGKIHTDFEKGFIRAQTIAFDDFITYKGEQGAKEAGKMRAEGKDYIVKDGDIMNFLFNV---- : 363
tr|Q3Z0S1| : -EVRAWTIPVGATAPQAAGKIHTDFEKGFIRAQTISFEDFITYKGEQGAKEAGKMRAEGKDYIVKDGDVMNFLFNV---- : 363
tr|Q7VR78| : -MTRSWIYESETIALDAAYNIHSDFKKGFIRAQIIKFNDFIANHGESGAKKSGRIRYEGKNYLIEEGDIIKFLFKN---- : 348
tr|Q492W3| : -MTRAWIYVVGMTALEAAKKVHSDFKKGFIRVQIIKFDDFIFYRGELGAKKAGKICYAGKDYCVKDGDIIKFLFQI---- : 358
tr|Q1LTH0| : -EVRAWTIPVGTTALQAANKIHTDFAKGFIRAQTISFNDFITFKGEKGAKKAGKIHLEGREYIVKDGDVINFLFNV---- : 356
tr|B2IBV5| : -EARAWTIPAGTKGPQAAGVIHTDFEKGFIRAETIAYPDYVTHQGETGAREAGKFRLEGKDYVVADGDIIHFRFAT---- : 365
tr|Q2GIK1| : -EARAWSIMRNTPADKAAGTIHTDFEKGFIKAETISFEDYIQCGGEAECKKAGKVRFEGRDYIVQDGDIMHFRVNK---- : 363
tr|A7HVD6| : -ETRAWTITRGTRAPAAAGVIHTDFEKGFIRAETIAYDDYIALNGEQGAKDAGKMRLEGKEYVVQDGDVMHFRFAN---- : 369
tr|A9M6K2| : -ETRAWTIKRGTKAPAAAGVIHTDFERGFIRAQTIAYEDYVKYNGELGAKEAGKARDEGKEYIVHDGDVMLFRFNT---- : 367
tr|A5FZ49| : -ETRAWTITRGTAAPAAAGVIHGDFERGFIACETIAYDDYVACNGEAGAREAGKLRIEGRNYIVQDGDVLLFRFNV---- : 364
tr|A5V637| : -EARAWTVEKGSKAPQAAGVIHSDFERGFIRAETMAYADYVALGGEAAAREAGKLRSEGKDYVVHDGDIMLFRFNV---- : 366
tr|B3EFF7| : -EVHAWTIRKGAAAPEAAGAIHSDFEKGFIRAEVMSYRDLIELGSEQKVKEAGKLRSEGRDYIVRDGDVIVFRFNV---- : 363
tr|Q0C0W8| : -ESRAWTIRKGWTAPKAAGVIHGDFEKGFIRAETIAFEDYVACGGEVGAKEAGKMRSEGKEYIVQDGDVMLFRFNV---- : 365
tr|Q0ARN9| : -EARAWTIREGWTAPQAAGVIHGDFEKGFIRAETIAYDDYVACGGEAPAKEAGKMRAEGKEYIVKDGDVLHFRFNV---- : 365
tr|B1WLN7| : -EARAWTITKGTKAPAAAGVIHTDFEKGFIRAETIAYDDYVTLGGEAGARDAGRLRLEGKEYVVADGDVMHFRFNT---- : 365
tr|A6UBR8| : -ETRAWTIPRGTKAPQAAGVIHSDFERGFIRANTIAYDDYIAYGGENGAKEAGKARDEGKEYVVQDGDVIHFRFNT---- : 367
tr|B4S416| : -EVRAWTIRKGAAAPEAAAAIHTDFEKGFIRAEVIAYNDMITYGSEQKAKEAGKMRSEGKEYIVKDGDVIVFRFNV---- : 363
tr|Q31IN2| : -EVRAWTVKKGATAPQAAGVIHTDFEKGFIRAEVTAYDDFIEYKGDAGAKAAGKQRLEGKEYIVKDGDVMHFRFNV---- : 363
tr|Q8PC60| : -EVRAWTVKAGSTAPQAAAVIHTDFEKGFIRAETIAFDDFIKYRGESGARDAGRLRLEGKEYRVKEGDVLHFRFNV---- : 363
tr|Q5GWR7| : -EVRAWTVRAGSTAPQAAAVIHTDFEKGFIRAETIGFDDFIKYRGEAGARDAGRLRLEGKEYRVQEGDVLHFRFNV---- : 363
tr|Q87A25| : -EVRAWTVKSGATAPQAAAVIHTDFEKGFIRAETIAFDDFIKYKGEAGARDAGRMRLEGKEYRVQEGDVLHFRFNV---- : 363
tr|A1VSN2| : -EVRAWTIHVGDTGPQAAGVIHTDFEKGYIRAQTIAFDDFIAFKGEQGAKDAGKMRAEGKEYVVKDGDVMNFLFSS---- : 363
tr|Q73HU8| : -EARAWPVKIGSTADKAAGVIHTDFEKGFIKAETISFADYIKYGSESACKDAGKIRFEGRDYIVQDGDIMHFRFNV---- : 364
tr|Q2GHV2| : -EARAWPIKRFSTASSAAGVIHTDFEKGFIKAELISFDDYIKYNGEAKCKEAGKVRFEGRDYIVQDGDIIHFRHNK---- : 362
tr|Q6AQ62| : -ETRAWTIPVGAKAPEAAGKIHTDFEHGFIRAEVIGYDDFVKCQGEAKAKEQGLMRVEGKEYVVADGDCMHFRFNV---- : 365
tr|Q2S6K6| : -GAYAWAIEEGTRAPQAAGEVHSDFEQGFIKAETIHFSDFDEAGSEAAAREAGLLRTEGKDYVVEDGDVMEFRFNV---- : 400
tr|Q6MJR2| : -EVRAWTIRANTKAPQAAGVIHSDFERGFIKAETYHCEDLFKYKSEQAIKDAGKYRIEGKEYVVKDGDVIFFKFNV---- : 366
tr|Q5HUM9| : -EVRSWTIKKGWKAPKAASVIHNDFEKGFIKAEVISYEDYIQYKGENGAKEAGKLRLEGKDYIVLDGDVMHFRFNV---- : 367
tr|A5I1T1| : -EVRAWTIVKGTKAPKAAGKIHTDIEKGFIRAEVISYDKLIECGSEAHAKEKGFFRLEGKEYVMQDGDIVNFRFNV---- : 365
tr|A4XK71| : -EVRAWTIKKGTKAPQAAGKIHSDFERGFIRAEVVPFDVLVECGGFAQAKEKGLVRSEGKDYVMQDGDVVIFRFNV---- : 362
tr|Q606I4| : -EVRAWTIPAGATAPQAAGVIHSDFEKGFIRAEVIAYEDFVACQGEQGAKDAGKWRLEGKDYVMRDGDIVHFRFNV---- : 363
tr|B7XSQ8| : -EVRAWTFKQGIKAPKAAGIIHSDFERGFIKAEVYSCEDLFKFQSVQKLKEKGLIRIEGKEYLVRDGDVIFFKFNV---- : 368
tr|A9EY40| : -EIRAWTIPIGCTAPQAAGVIHSDFEKGFIRAEVYKVDDLVEHRTEAAIRAAGKLRVEGKTYVFQDADVAHFLFNV---- : 363
tr|B0TA60| : -EVRAWTIPRGTKAPQAAGVIHTDFERGFIRAEVTAYDDLVACGSQTAAREKGLQRLEGKEYVMKDGDVVHFRFNV---- : 366
tr|B7DUK8| : -EVRAWTIRKGTKAPQAAGVIHSDFERGFIRAEVVAYDDLVAAGSFQAAREQGKIRLEGKDYVVQDGDVCYFRFNV---- : 366
tr|Q81JI1| : -EVRAWTFKQGMKAPQCAGVIHTDFERGFIRAETVSYEDLMTNGSMTAAKEAGKVRLEGKEYIVKDGDVMHFRFNV---- : 366
tr|A4ITW1| : -GVRAWTFRKGMKAPQCAGIIHSDFERGFIRAETVSYADLVAAGSMAAAREAGKVRLEGKDYEVQDGDIMHFRFNV---- : 366
```

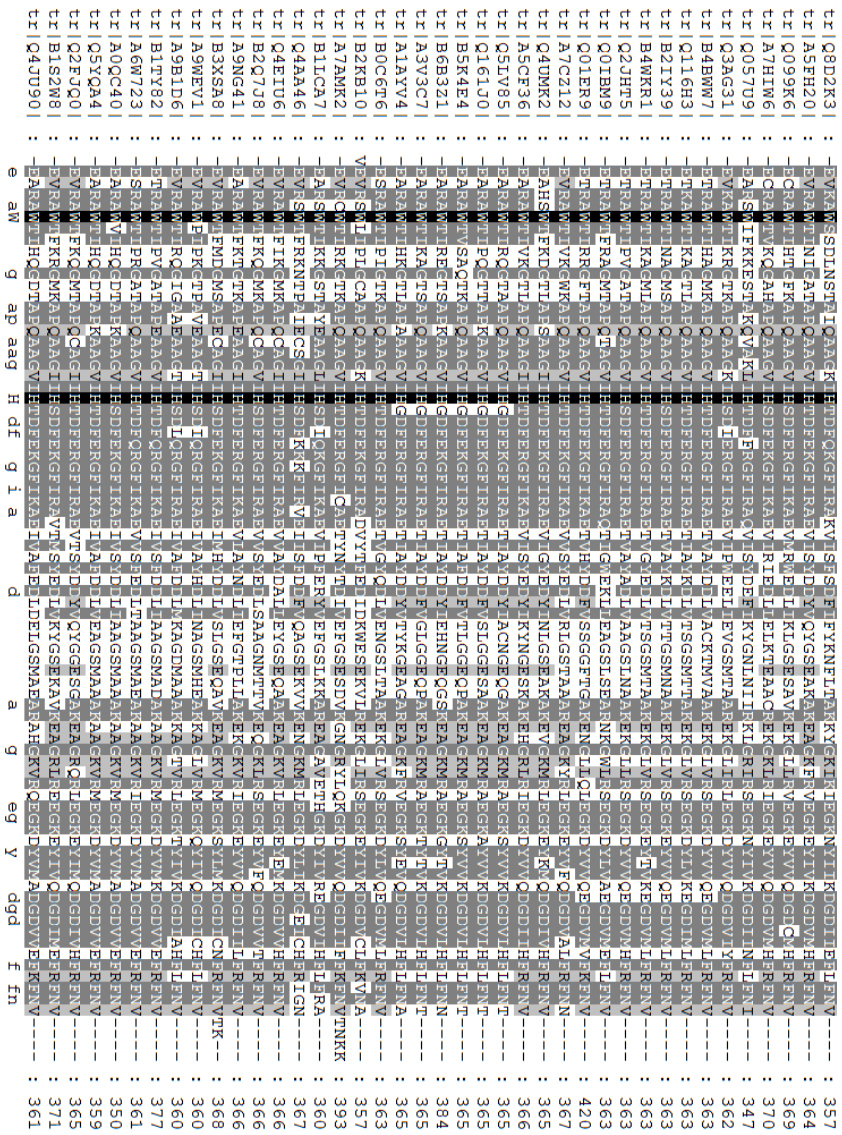


Fig. A.1 Primary sequence alignment of YchF within bacteria Black indicates 100% conservation, dark grey 80 to 100% conservation, light grey 60 to 80% and white is less than 60% conserved.

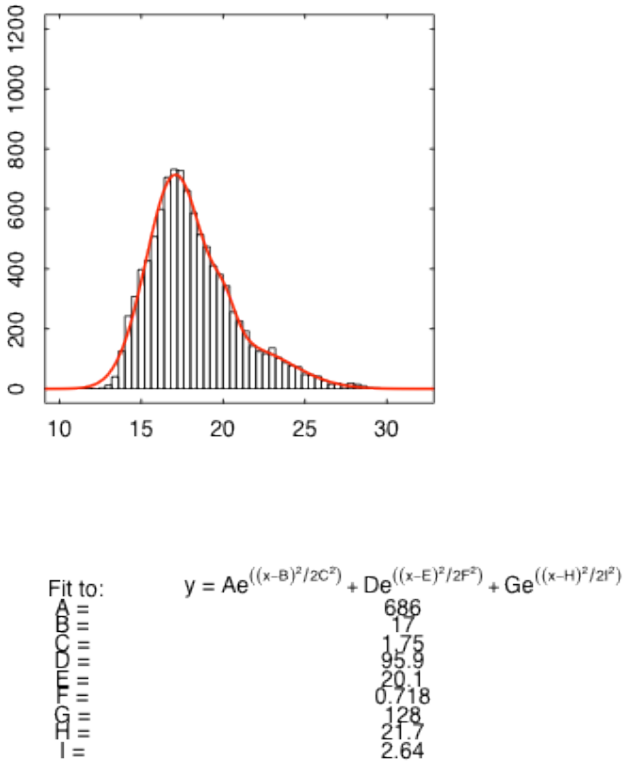
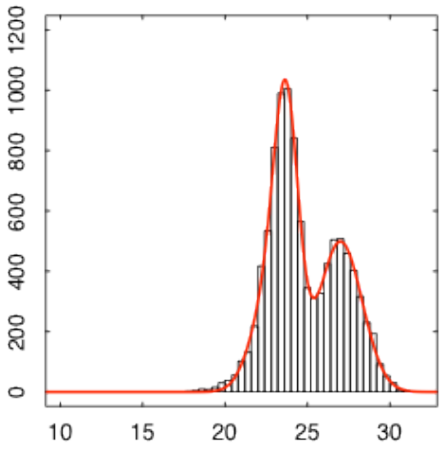


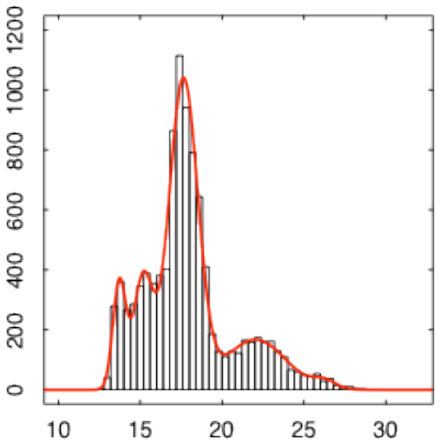
Figure A.2 Distance between α carbons of His 114 and Ser 16 during 50 ns simulations of *apo* YchF.



Fit to: $y = Ae^{((x-B)^2/2C^2)} + De^{((x-E)^2/2F^2)} + Ge^{((x-H)^2/2I^2)}$

A =	308
B =	22.7
C =	1.1
D =	805
E =	23.7
F =	0.733
G =	498
H =	27
I =	1.31

Figure A.3 Distance between α carbons of His 114 and Ser 16 during 50 ns simulations of YchF•ADP•Mg^{2+}.



$$\text{Fit to: } Ae^{((x-B)^2/2C^2)} + De^{((x-E)^2/2F^2)} + Ge^{((x-H)^2/2I^2)} + Je^{((x-K)^2/2L^2)} + Me^{((x-N)^2/2O^2)}$$

A =	353
B =	13.7
C =	0.395
D =	373
E =	15.2
F =	0.595
G =	1030
H =	17.6
I =	0.88
J =	167
K =	22
L =	1.86
M =	26.8
N =	26.2
O =	0.696

Figure A.4 Distance between α carbons of His 114 and Ser 16 during 50 ns simulations of YchF•ATP•Mg^{2+}.