11-2-2020

Salt Dependence of Thermodynamic Stability of a Cold-Active DNA Polymerase I Fragment

Xinji Zhu

SALT DEPENDENCE OF THERMODYNAMIC STABILITY OF A COLD-ACTIVE DNA POLYMERASE I FRAGMENT

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Biological Sciences

by
Xinji Zhu
B.S., East China University of Science and Technology, 2012
December 2020

ACKNOWLEDGMENTS

As many long-term research projects, my dissertation demanded a lot of energy and time. During this process, many people have contributed to my dissertation and I am grateful for their support and time.

I would like to express my great appreciation to Dr. Vince LiCata, my major supervisor, for his patient guidance, enthusiastic encouragement, and valuable critiques of my research work. Without his support, this dissertation would not have been possible. I would also like to thank my committee members, Dr. Naohiro Kato and Dr. John Battista, who give me a lot of advice and support in my research.

I would also extend my thanks to the members of the LiCata Lab: Dr. Chin-Chi Liu, Dr. Hiromi Brown, Dr. Jaycob Warfel, Tod Baker, Wendy Hansing, Erin LeBoeuf, and Katelyn Jackson for their technical and emotional support.

I would like to thank the financial, academic and educational support of the Louisiana State University Department of Biological Sciences and its staff; Dr. Ted Gauthier and Dr. Rafael Cueto for their technical support in data collection; Dr. Bill Wischusen, Dr. Jane Reiland and Dr. Christopher Gregg for their teaching guidance.

Finally, I would especially thank my family. My wife, Yuli, is extremely supportive of me and has made extraordinary sacrifices to help me get his point. Also, I would like to thank my mother and father for their continued support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

[Salt]	Salt concentration
ΔG	Gibbs free energy change
ΔCp	Heat capacity change
ΔH	Enthalpy change
ΔS	Entropy change
AS	Ammonium sulfate
CD	Circular Dichroism
CHT	Ceramic Hydroxyapatite
DLS	Dynamic Light Scattering
HIC	Hydrophobic interaction chromatography
IDRs	Intrinsically disordered regions
Krel	The relaxation folding rate constant
Klenow	Large fragment of *Escherichia coli* PolI
Klenpin	Large fragment of *P. ingrahamii* PolI
Klentaq	A large fragment of *Thermus aquaticus* Pol I
PEI	Polyethyleneimine
PCR	Polymerase Chain Reaction
PI	Isoelectric point
PolI	DNA Polymerase I
PTS	Protein thermal shift assay
P. ingrahamii	*Psychromonas ingrahamii*
Tm	The temperature at the middle point of transition

ABSTRACT

P. ingrahamii is a halo-psychrophilic bacterium isolated from Arctic sea ice off the northern coast of Alaska that has been demonstrated to be able to grow down to -12°C. We have cloned and purified the large fragment of the cold-active DNA polymerase I from *P. ingrahamii*, named Klenpin. The objective of this project is to directly compare the thermodynamic stability of Klenpin, and the salt dependence of that stability, with Klenow and Klentaq; two homologous polymerases from a mesophile (*E. coli*) and a thermophile (*Thermus aquaticus*).

We first examined the effects of salts on the thermal stability(Tm) of Klenpin and Klenow across the Hofmeister series, which generally rank ions from stabilizing to destabilizing. Significantly different trends were observed on the melting temperature changes (ΔTm values) for Klenpin versus Klenow, even in chaotropic salts such as guanidine hydrochloride (GdnHCl) at low concentrations. Klenow polymerase responded to Hofmeister stabilizing salts and destabilizing salts as expected, while Klenpin was stabilized by all salts, even those that normally destabilize proteins.

We further examined the salt effect on the structure of Klenpin and Klenow using CD spectra, Trp fluorescence quenching, and dynamic light scattering. The results show that salt alters the structure of Klenpin to a more compact conformation, whereas no significant influence is observed in Klenow. Not only do salts affect the structure of Klenpin, but also they alter the unfolding process of Klenpin. In chemical denaturation, an intermediate is observed in the presence of salt, whereas Klenpin folds following a two-state model without salt. An increase in free energy of unfolding of Klenpin is also seen upon the addition of NaCl, and this improvement is mainly from the additional unfolding transition introduced due to the intermediate.

It is possible that non-specific screening of unfavorable electrostatic interactions on the surface of Klenpin maybe be responsible for many of the observed slats effects and so we conducted several computational comparisons to test this hypothesis. Unlike those typically found for archaebacterial halophilic proteins, no significant difference is observed in the comparison of amino acid preference between Klenpin and Klenow. But the electrostatic surface potential maps of Klenpin, Klenow, and Klentaq show larger clusters of acidic and basic residues on the Klenpin surface, supporting our hypothesis that salts modify Klenpin through the nonspecific screening of unfavorable electrostatic interactions on the surface.

CHAPTER 1. GENERAL INTRODUCTION

1.1 Life Adaptation in Extreme Environments

An extreme environment is a habitat characterized by harsh environmental conditions, such as temperatures of -20 °C or 122°C, high salinity, high radiation, high pressure, and pH levels of 2 or 11 [293]. Organisms with the ability to thrive in extreme environments are known as extremophiles. Nearly all extremophiles are microorganisms (archaea, bacteria, and eukaryotes), and their existence began approximately three billion years ago [26, 27], when the atmosphere was rich in greenhouse gases and the water from volcanism was acidic and hot [28]. In order to maintain structural and functional integrity, extremophiles generally have two strategies with which to adapt to these otherwise uninhabitable conditions, either by removing the extremes or inducing protective mechanisms for survival. For example, molar concentrations of KCl ("salt-in" strategy) are accumulated in *Haloarcular marismortui* in order to keep its cytoplasm isosmotic with a high salinity environment (3.4-3.9M NaCl) [29]. Other halophiles, such as *Halorhodospira halophila*, can synthesize organic osmolytes ("osmolyte" strategy), such as glycine betaine, sucrose, and trimethylamine N-oxide, to maintain proper osmotic pressure in the cytoplasm [29]. Proteins from halophiles may share similar structural adaptation. In the presence of molar salt concentration halophilic proteins contain a surface with low hydrophobicity to prevent aggregation and loss of function [35].

In this project, the Klenpin equivalent large fragment domain of DNA Polymerase I from *P. ingrahamii* was studied in order to better understand the protein folding and the stability of proteins in sea ice, where *P. ingrahamii* was originally isolated. The protein is referred to as Klenpin. Sea ice is primarily characterized by low temperature and high salinity, and therefore both the temperature dependence and the salt dependence of Klenpin were compared to Klenow and

Klentaq, two homologous proteins from *E.coli* (mesophile) and *Thermus aquaticus* (thermophile). In addition, amino acid composition comparison among Klenpin, Klenow, and Klentaq was investigated in order to explore the potential structural explanation for results from experiments.

Psychrophiles are defined as organisms that can live and thrive at low temperatures. In Moria's definition of a psychrophile, it has an optimal growth temperature at 15°C or below and cannot survive above 20°C [99]. Currently, *Planococcus halocryophilus*, an arctic permafrost bacterium, has the lowest growth temperature at -15°C, with a generation time of 1200 hours in lab conditions [100]. Compared to their homologous mesophilic enzymes, psychrophilic enzymes often have higher catalytic efficiency at low temperatures, as well as lower thermal stability and higher flexibility in their structures [94].

Extreme low temperatures can induce cytoplasmic ice crystals, which can cause cellular damage and osmotic imbalance [109]. One compensatory strategy is to synthesize compatible solutes (glycine betaine, TMAO (Trimethylamine-N-oxide), and sucrose) in order to lower the freezing point and maintain osmotic pressure [110]. Some psychrophilic organisms produce antifreeze proteins that control ice crystal growth [210]. Sea ice is characterized by the high salinity of the brine, which affects pH, dissolved inorganic nutrients, and gas [102]. Psychrophiles within sea ice are localized in hypersaline pockets and veins, which also contain dissolved organic matter [102]. Two genes in *P. ingrahamii* are responsible for the synthesis of glycine betaine which potentially balance the osmotic pressure *in vivo* [18]. Glycine betaine is utilized to investigate the effect on the stability of Klenpin and compared with that from salt dependence.

In addition to these physiological methods, changes in the composition of amino acid residues in a protein is also an essential strategy. Reduced proline content was observed to diminish the negative effect of proline isomerization on the folding rate in many psychrophilic enzymes [127].

Also, the lower content of electrostatic amino acids and increased polar residues (particularly serine) are responsible for the reduced thermostability and increased flexibility of cold-adapted proteins [38,39]. Moreover, enzymes from psychrophiles, with lower stability compared with its mesophilic and thermophilic counterparts, often contains lower helical content [135]. As a result, aspartic acid, serine, and threonine are often preferred in psychrophiles, since aspartic acid and serine are helix breakers, and threonine is helix indifferent [135]. With the high ability to form salt-bridge, glutamic acid also has a lower proportion in psychrophiles [42]. The composition of smaller amino acids is increased in the beta-sheets of psychrophilic proteins, because they are unable to form the long-range interaction that would increase the stability [44, 45].

Halophiles, which are also distributed in all three domains (Archaea, Bacteria, and Eukarya), are salt-loving organisms that require salt for growth [216]. Based on the salt concentration for optimal growth, halophiles are classified into three groups: slight, moderate, and extreme halophiles [217]. Slight halophiles only require 1.7% to 4.8% (w/v) salt for growth, whereas extreme halophiles prefer 20% to 30% (w/v) salt [217].

Halophiles primarily have two strategies with which to maintain osmotic balance in the cytoplasm: the salt-in strategy [218] and the salt-out strategy [219]. In the salt-in strategy, a high concentration of potassium chloride is accumulated in the cytoplasm in order to balance the extreme osmotic pressure. This mechanism is primarily observed in some archaea and anaerobic bacteria [218]. In the salt-out strategy, a high concentration of small soluble organic compounds was produced or transported in the cytoplasm in order to maintain the high osmotic pressure [219].

For halophiles using the salt-in strategy, many studies have found a high ratio of acidic to basic residues, especially at the protein surface [206, 221]. A stabilization model was proposed in order to explain the relationship between the acidic amino acid and salt binding [223]. The model

suggested that acidic residues on the protein surface can form a network with hydrated salt ions in order to stabilize the proteins. Klenpin was used as a model protein with which to investigate *P. ingrahamii's* adaptation to sea ice.

1.2 Protein Folding in Psychrophiles and Halophiles

When a long polypeptide (50 residues or more) comes off of the ribosome, it spontaneously starts to fold into a specific three-dimensional structure. The vital driving force for protein folding is the hydrophobic effect, resulting in the buried hydrophobic core. Other interactions and processes, such as van der Walls, H-bonds, salt bridges, and prolyl isomerization, also contribute to this process [85]. The protein folding rate was believed to be adversely affected by low temperatures. First, the hydrophobic effect depends on the temperature as it becomes weaker at lower temperature [150]. Furthermore, psychrophilic proteins tend to expose more non-polar residues to the solvent in order to disfavor a compact conformation [86]. As a result, the hydrophobic effect in psychrophilic proteins is perturbed, resulting in a lower folding rate compared to the homologous proteins from mesophiles and thermophiles.

The folding of cold shock proteins from *Thermotoga maritima* (hyperthermophilic bacterium, Tm-Csp), *Bacillus caldolyticus* (thermophilic bacterium, Bc-Csp), and *Bacillus subtilis* (mesophilic bacterium, Bs-Csp) have been investigated in previous studies [87]. The results show that the refolding rates of all three of these proteins are similar, whereas Tm-Csp has the lowest unfolding rate. However, such comparisons are currently lacking between psychrophilic proteins and mesophilic proteins [295]. Instead, Cipolla et al. compared the folding and unfolding rates of a psychrophilic alpha-amylase and its stabilized mutants through site-directed mutagenesis [89]. Interestingly, the folding rates constants (k_{fold}^{H2O}) of these proteins are very close, whereas the unfolding rate constants (k_{unfold}^{H2O}) of its stabilized mutants are 3.5 and 13.4 times slower with

respect to the wild alpha-amylase [89]. It suggests that the improved stability in both mutants are determined by the decreased unfolding rate rather than the folding rate [89].

During protein folding, prolyl isomerization is often a rate-limiting step because it involves the rotation of the peptide bond, which has a partial double bond character [90]. Since this process requires high activation energy, it is significantly correlated with the temperature. Therefore, one can show that prolyl isomerization is very slow in psychrophiles, whereas it is faster in hyperthermophiles [91]. Peptidyl-prolyl cis/trans isomerases (PPiases) accelerate the isomerization process. By comparing the genomes of a series of extremophiles, it has been found that there is a higher number of PPiase genes in psychrophiles than in hyperthermophiles. Psychrophiles can either improve the PPiase specific activity or increase the enzyme concentration in order to maintain a required prolyl isomerization rate [92,93]. Also, reduced proline content is often observed to improve structural flexibility in many cold-adapted proteins [94]. From the genomics of *P. ingrahamii,* five genes (919, 1049, 1080, 1469, 1619) are found to be responsible for the synthesis of Peptidyl prolyl isomerases, which may play a critical role in protein folding at low temperature [18].

It is still not fully clear how salt stabilizes a protein, but studies of halophilic enzymes may provide useful models with which to solve this problem. Previous studies on several halophilic enzymes, such as *Haloarcula marismortui* malate dehydrogenases and *Halobacterium salinarum* ferredoxins, have shown that their structure and functional properties require a molar concentration of salt, and that the removal of salt causes the unfolding of proteins [259, 260]. As noted earlier, many halophilic proteins are characterized by an abundance of acidic residues, especially at the protein surface, leading to a high negative surface electrostatic potential [261]. As a result, strongly

repulsive electrostatic interactions among acidic residues cause the destabilization of halophilic proteins in the absence of salt [261, 262].

For halophilic organisms with a salt-out strategy, a significant decrease of apolar surface in the proteins can be observed [35]. Most osmolytes synthesized by halophiles contain charged groups such as glycine betaine. The solubility and stability of halophilic proteins are affected by molar osmolyte concentration due to the interaction between the charged sites of the osmolytes and the oppositely charged area on the protein surface [264].

1.3 Thermodynamics and Kinetics in Protein folding

The thermodynamic stability of a protein is quantitatively defined as the difference in Gibbs free energy between the folded and unfolded states: $\Delta G_u = G_{fold} - G_{unfold} = -\mathrm{RT}\ln K_u$, where R is the gas constant, T is the temperature in Kelvins, and K_u is the equilibrium constant for unfolding.

1.3.1. Temperature

For mesophilic organisms, the denaturation of the protein upon heating seems to be a natural phenomenon. Any process induced by increasing temperature should process with heat absorption, and thus with an increase in enthalpy and entropy, which often cause protein destabilization [95]. Therefore, the decreasing temperature should lead to the stabilization of the protein. Nevertheless, it is predicted from Gibbs-Helmholtz relationships (Figure 1.1) that a protein can also be unfolded when cooling from room temperature to lower values, which is called cold denaturation. As a result, the difference between thermal and cold denaturation temperatures defines the temperature range in which the protein is thermodynamically stable. We also compared the temperature dependence of Klenpin and its homologous proteins (Klenow and Klentaq) to investigate how the environment temperature affects protein stability.

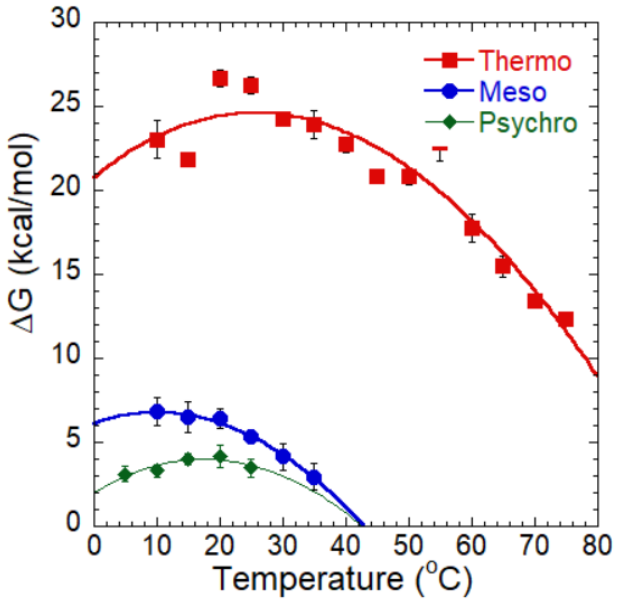

Figure 1.1.Gibbs free energy of unfolding, or conformational stability, of homologous extremophilic proteins.

1.3.2. pH

The pH value of the solution also plays a crucial role in terms of protein stability. A change in the pH value of the solvent determines the ionization state of the titratable amino acid on the protein, and thus directly affects the electrostatic interaction. For example, the stability of *Pyrococcus furiosus* methyl aminopeptidase decreased at low pH, where the acidic residues are protonated, and favorable ionic interactions are disrupted [299]. *Psychromonas ingrahamii* lives in neutral habitat with optimal pH range from 6.5 to 7.4 [306], whereas *Thermus aquaticus* lives at a higher pH condition (7.5~8.0) [115] and *E. coli* grows in a broad pH range (4.4~10.0) [113]. As a result, we also compared the effect of pH on the stability of Klenpin, Klenow and Klentaq.

1.3.3. Protein denaturant

Protein can also be denatured in the presence of molecules known as denaturants, such as GdnHCl and urea. The detailed molecular mechanism is still not entirely clear. Jorge Almarza et al propose that urea direct interact with protonated histidine for protein inactivation followed by hydrogen bond formation with polar residues, and the breaking of hydrophobic collapse [107]. For GdnHCl, it's general believed that the denaturing effect is due to its favorable interaction with the polar parts of proteins and that the non-polar side chains have no or little favorable interaction with guanidine hydrochloride [92]. Interestingly, the estimate of protein stability from urea or GdnHCl may provide different results depending on the contribution of electrostatic interactions [240].

Several studies also show that low concentration GdnHCl (0-0.3M) may increase the stability of some proteins [106, 109]. Bhuyan et al. designed the association and dissociation reactions of ferrocytochrome c and CO in order to investigate changes in conformation entropy of proteins in the presence of low concentrations of GdnHCl [281]. The results have shown that the rate coefficients decrease as the protein is stiffened in low GdnHCl concentration (<2.1M), suggesting that protein-denaturant interactions lower the conformational entropy of the protein [281]. In the presence of high GdnHCl concentration, the protein stiffening is suppressed, and the rate coefficients begin to increase (>2.1M) [281]. Another possible model proposes that specific cation binding stabilizes the protein with $GdnH^+$ [109]. In this project, GdnHCl and urea are both utilized in order to compare the Gibbs free energy upon unfolding between Klenpin and Klenow.

Salt effect & Hofmeister series

Salt effect on protein stability and solubility has been extensively investigated by Franz Hofmeister, who created the Hofmeister series by classifying different anions and cations based on salting-in and salting-out effects [104]. The mechanism of interaction between Hofmeister ions

and the protein is not well understood, and it is controversial whether this interaction is ion-specific or whether it is nonspecific. A series of Hofmeister ions are tested to investigate how the ionic strength affect the stability of Klenpin and Klenow, and whether this effect is ion-specific or not.

P. ingrahamii can live in concentrations of NaCl up to 2.6M [3], and it is reasonable to speculate that Klenpin is also a halophilic or halotolerant enzyme. Previous genomic and structural analysis results have shown that the enzymes from halophiles usually contain more acidic residues (asp and glu) and less basic residues, especially lys, when compared with their homologous mesophilic proteins [34]. Also, a decrease in the hydrophobic content in halophilic enzymes is often observed [35]. Surprisingly, no significant difference was observed when comparing the amino acid composition between Klenpin and Klenow. Although Klenpin has fewer basic residues (82, 13.4%) than Klenow (90, 15.9%), it also contains less acidic residues (83, 13.5%) than Klenow (89, 14.7%). As a result, Klenpin has much fewer charged residues than Klenow. And with regards to hydrophobicity, both Klenpin and Klenow contain around 300 hydrophobic residues. Therefore, it is interesting to compare the salt effect on the stability of Klenpin and Klenow. Protein thermal shift assay and circular dichroism spectroscopy were used in order to investigate the effect of five anions ($SCN^-, HCO_3^-, SO_4^{-2}, HPO_4^{-2}, Cl^-$) and four cations ($Na^+, K^+, Mg^{2+}, GdnH^+$) on the melting temperatures (Tm) of Klenpin and Klenow in different salt concentrations.

Salt effect on halophilic proteins

As mentioned in previous sections, halophiles have two methods by which to maintain proper osmotic pressure in the cytoplasm: increasing internal ion (K^+, Cl^-) concentrations (salt-in strategy**)** or the accumulation of osmolytes (salt-out strategy) [35]. For halophiles obeying the salt-in strategy, protein folding occurs in the presence of molar concentrations of potassium chloride.

For example, malate dehydrogenase from *halophilic malate* is inactivated and unfolded at NaCl concentrations below 2M [144].

Classical electrostatic stabilization was proposed as the essential factor for the halophilic adaptation of protein [290]. For mesophilic proteins, a high concentration of salt enhances the protein aggregation known as salting out. It also interferes with electrostatic interaction because of charge screening and decreases the natural hydration of proteins [206]. Extremely halophilic protein, however, can maintain its structure and function only in saturate salt conditions (>2M) [144]. One possible mechanism is that a high acidic/basic residue ratio, especially on the protein surface, diminishes the deleterious effect of salt by promoting non-specific electrostatic interactions with salts [146, 292].

An abnormal stabilization curve is observed for halophilic malate dehydrogenase from *Haloarcula marismortui* by Mg^{+2} [302]. At low $MgCl_2$ (<0.5M), the stability of the protein increases with an increasing $MgCl_2$ concentration, whereas at higher concentrations (0.8-1.5M $MgCl_2$), the stability of the protein decreases with an increasing $MgCl_2$ concentration [302]. The results show that magnesium stabilizes the protein at low concentrations, similar to monovalent cations (like Na^+ or K^+), whereas it destabilizes the protein at high concentrations, like guanidinium hydrochloride [143].

1.3.4. Osmolytes

Osmolytes are small organic compounds that maintain cell volume and affect protein stability. For halophiles, the accumulation of osmolytes (glycerol, glycine-betaine, and trimethylamine-N-oxide) is a common strategy with which to exclude salts from their cytoplasm and balance osmotic pressure [5]. Proteins in such halophiles lack the features of amino acid composition, which is often observed in salt-in halophiles [55]. In dilute aqueous solution, the thermodynamic forces for

protein folding and stability include hydrogen bonding, electrostatic, van der Waals, and hydrophobic interactions [53]. The cell environment is usually more complex, and specific small organic molecules (osmolytes) are accumulated in order to protect the intracellular macromolecules from denaturing environmental stresses.

Qu et al. found that the stabilization of proteins by osmolytes arises from the destabilization of the unfolded state by raising its Gibbs free energy [6]. The Stokes radius of reduced and carboxamidated ribonuclease A for the transfer of the protein from water to 1M osmolytes and the transfer Gibbs free energy changes are shown in Figure 1.2 [6]. The results indicate that osmolytes stabilize proteins by raising the chemical potential of the denatured ensemble. Furthermore, the crowded cell environment in the presence of osmolytes decreases the accessible volume, and therefore favors the more compact folded state [60]. Two genes in *P. ingrahamii* are responsible for the synthesis of glycine-betaine in order to balance the osmotic pressure *in vivo* [18]. Therefore, a series of osmolytes (glycine-betaine, TMAO, and sucrose) are utilized in order to investigate the osmolyte effect on the stability of Klenpin.

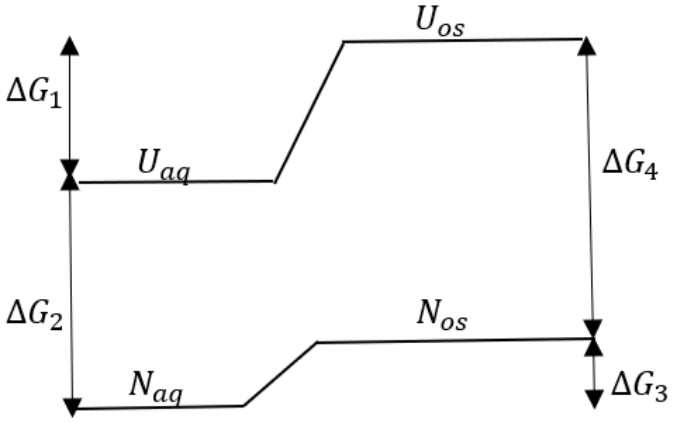

Figure 1.2.Gibbs energy diagram. ΔG_1: the increase in the chemical potential of the denatured ensemble ΔG_2: the unfolding Gibbs energy difference between a native and unfolded state in the aqueous buffer; ΔG_3: the increase in the chemical potential of the native ensemble; ΔG_4: the unfolding Gibbs energy difference in the presence of osmolytes (This figure is adapted from chart 1 of reference 6).

1.3.5. Protein folding kinetics

The folding mechanism for a protein consists of two parts: thermodynamics and kinetics. In the previous sections, we described how environmental factors influence the stability of the native state. Here, we ask: How does a polypeptide fold to the native state kinetically?

Protein folding kinetics for two-state proteins

The simplest model with which to describe protein folding is a two-state system in which only the unfolded state (U) and native state (N) are populated on the folding pathway [62]. It has been demonstrated that many monomeric proteins fold through simple two-state kinetics, with folding rates from microseconds to seconds [116-117].

Folding kinetics is highly sensitive to environmental conditions such as temperature, salt, and pH. For some proteins, the folding kinetics can be switched from multi-step to single-step when changing pH or temperature [64]. In some exceptional cases, both multi-step and single-step exist simultaneously [137-138]. Kiefhaber et al found that the major part of hen egg lysozyme molecules (86%) refold through a slow kinetic pathway with partially folded states, whereas 14% of the molecules fold faster by a two-state process[137]

The unfolding/folding kinetics for the two-state model is very simple, and can be analyzed by a single jump experiment with fluorescence or the circular dichroism method. Typical unfolding kinetics for a two-state protein are shown in Figure 1.3 [139], following single exponential kinetics.

1.3.6. Transition State

For a chemical reaction, the transition state is defined as the state corresponding to the highest potential energy, along with this reaction coordinate [304]. For protein folding/unfolding, the transition state represents the maximum free energy point on the folding/unfolding pathway. In a refolding experiment, a protein incubated in high denaturing conditions (high temperature, high

concentration of denaturant, or extreme pH) is diluted to its native state. The unfolding experiment is initiated by a condition shift from native to highly denaturing. These processes are monitored by a spectroscopic technique, or circular dichroism (CD) [119]. Information on the transition state can be used to determine folding/unfolding rate constants.[140-141].

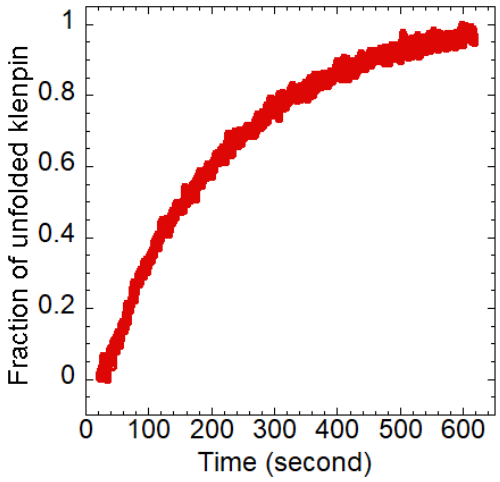

Figure 1.3.Single-mixing refolding kinetics of Klenpin by circular dichroism.

1.4 DNA Polymerase I

All living cells are faced with different tasks necessary for keeping the genome intact to be functional in a complex environment, to divide at the right time, and to die when it is appropriate. DNA polymerase (DNA pol) is therefore required to keep the genetic information intact during DNA synthesis, DNA repair, and bypassing lesions processes. Based on their crystal structures and amino acid sequences, DNA polymerases can be classified into several families (A, B, C, D, X, and Y).

DNA pol I is a member of family A polymerases and is the most abundant polymerase in *E. coli*. It is a single polypeptide enzyme with three activities: a template-directed DNA polymerase, a 3'-5' exonuclease, and a 5'-3' exonuclease. DNA Pol I is mainly involved in two cellular processes: DNA replication and DNA repair. In DNA replication, it removes RNA primers from the lagging

strand through its 5'-3' exonuclease activity and replaces the RNA primers with DNA using its polymerase activity. After that, DNA ligase seals the nicks [331].

DNA pol I is mostly involved in DNA repair, including nucleotide excision repair (NER) and base excision repair (BER). In nucleotide excision repair (NER), the damaged or mismatched base within a string of nucleotides will be cut out of the duplex DNA by a nuclease. Then DNA pol I can initiate synthesis from 3' OH at the single-strand break (nick) or gap [329]. Finally, DNA ligase seals the remaining nick to repair DNA.

DNA pol I was first discovered in *E. coli* cells by Arthur Kornberg and his colleagues in 1956 [47]. In contrast, Taq polymerase was originally isolated from *T. aquaticus*, a thermophile which thrives at high temperature up to 80 °C [334]. The removal of any potential 5'-3' exonuclease activity in *E. coli* pol I and Taq polymerase yield the so-called large fragment domains (Klenow and Klentaq), which retain the original polymerase activity.

The X-ray crystal structures of Klenow and Klentaq both show a highly similar architecture like a human's right hand, containing three domains: finger domain, palm domain, and thumb domain (Figure 1.4) [305].

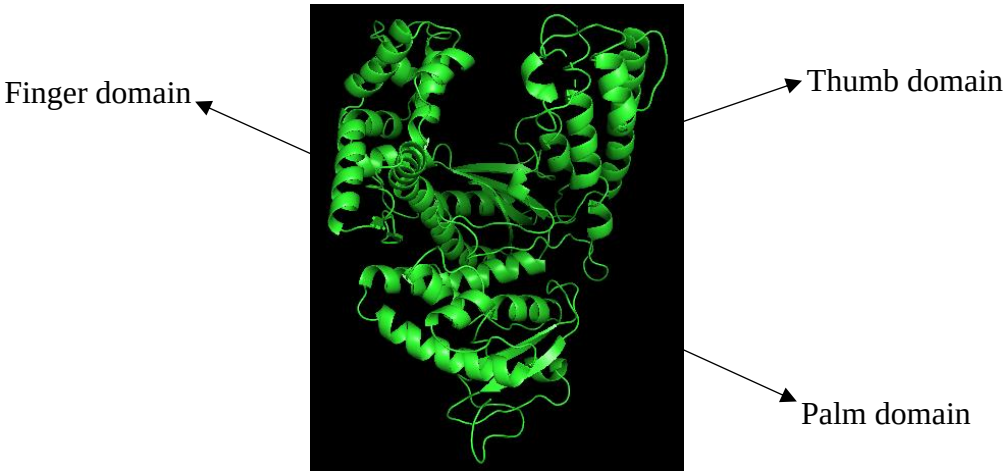

Figure 1.4. 3D structure of DNA polymerase I from *Thermus aquaticus*(5KTQ) by PyMOL.

During DNA replication, the finger domain can interact with the template and position the dNTP at the paired template base; the palm domain is responsible for the catalytic reaction of the transfer of the phosphoryl group (polymerase activity); and the thumb domain can interact with double-strand DNA and places the 3'-OH of the primer near the active site located in the palm domain[48].

1.4.1. DNA polymerase activity

The general DNA polymerase nucleotide incorporation reaction is shown in Figure 1.5. The reaction starts with the binding of the primer-template DNA (step 1) and dNTP (step 2) to the palm domain, which is in the "open" conformation. A change in the conformation is induced through the binding of the dNTP (step 3). A new phosphodiester bond is then formed via nucleophilic attack of 3'-OH of the terminal primer nucleotide to the 5' carbon of the dNTP (step 4). Another change in the conformation is then triggered (step 5) and causes the release of pyrophosphate (step 6). After that, the whole process will either be repeated (step 7-1) or terminated if the new DNA is released(step 7-2) [340]. Processivity is defined as the average number of nucleotides incorporated by a DNA polymerase on a single association/disassociation event. In other words, it measures the ability of DNA polymerase to continuously incorporate multiple nucleotides. Taq polymerase has a much higher processivity (50-80 bases) [338] than *E. coli* pol I does (< 30 bases) [337].

$$E + DNA_n \overset{1}{\leftrightarrow} EDNA_n + dNTP \overset{2}{\leftrightarrow} EDNA_n dNTP \overset{3}{\leftrightarrow} E_* DNA_n dNTP$$

$$\updownarrow 7\text{-}1 \qquad\qquad \updownarrow 4$$

$$E + DNA_{n+1} \overset{7\text{-}2}{\leftrightarrow} EDNA_{n+1} + PP_i \overset{6}{\leftrightarrow} EDNA_{n+1} PP_i \overset{5}{\leftrightarrow} E_* DNA_{n+1} dNTP$$

Figure 1.5. An overview of DNA polymerization reaction (Adapted from Reference 340). E is the enzyme, DNA_n is the DNA with n nucleotides, and E_* represent a conformational change in the enzyme. PP_i is a pyrophosphate group.

1.4.2. Exonuclease Activity of DNA pol I

The 3'-5' exonuclease activity or "proofreading" function allows the enzyme to check each nucleotide during DNA synthesis and excise mismatched nucleotides in the 3'-5' direction. Klenow 3'-5' exonuclease employs two divalent metal ions as cofactors. One metal ion is involved in the association with incoming dNTP, whereas the second ion is used to activate 3'-OH for the nucleophilic attack. This structural conservation of the proofreading domain is consistent within many DNA polymerases [339]. Enzymes with this conservation are grouped as DEDD exonuclease family, named after the related four amino acid residues in this process (aspartate(D), glutamate(E), and two additional aspartates(DD)). For Klenow. the four residues are Asp355, Glu357, Asp424, and Asp501 [340]. Klentaq, however, does not have the metal-binding sites within its exonuclease domain and therefore lacks the proofreading activity [336].

DNA polymerase I from *P. ingrahamii*

In 2003, a new organism called *P. ingrahamii* was isolated from sea ice near Point Barrow, Alaska [3]. It has a recorded growth temperature down to -12°C, with a generation time of 240h, and no growth is observed at 15°C. *P. ingrahamii* grows well at NaCl of 1-10% [306] and at near-neutral pH values (6.5-7.4) [3]. As a result, *P. ingrahamii* is moderately halophilic and strictly psychrophilic [305].

The large fragment domain of DNA polymerase I from *P. ingrahamii* was cloned along with the removal of any potential 5'-3' exonuclease activity [7]. Its sequence shares a high identity with the DNA polymerases from *E. coli* K-12 and *Thermus aquaticus* (66% and 44%, respectively) (Table 1.1). Like Klenow and Klentaq, the overall structure of Klenpin contains three domains: a polymerization domain and a 3'nuclease domain, the 5'-3' exonuclease domain. For the DNA

binding analysis, the activity of the 5'-3' exonuclease domain is removed by the D434A mutation. The protein is named Klenpin, by analogy with Klenow and Klentaq polymerases.

Table 1.1. Pairwise sequence alignment among Klenpin, Klenow, and Klentaq.

Identity	Klenpin	Klenow	Klentaq
Klenpin		66%	44%
Klenow	66%		47%
Klentaq	44%	47%	

The thermodynamics of Klenow and Klentaq, two homologous proteins from the mesophile (*E. coli*) and the thermophile (*Thermus aquaticus*), have been studied previously at the LiCata laboratory [162, 203]. Richard et al. have proved that salt effects on the stability of Klenow are highly consistent with the Hofmeister series [10]. Also, the cationic stabilization of Klenow can potentially be explained by reducing the repulsion on the protein surface [10]. With the thermodynamic information on Klenpin, stability comparisons among Klenpin, Klenow, and Klentaq are possible. Previous studies have shown that Klenow and Klentaq fold as a two-state, fully reversible model [204, 205]. As a result, the folding and unfolding kinetics of Klenpin are also studied by CD in order to characterize its folding pathway.

1.5 Intrinsically Disordered Protein (IDPs)

It has long been believed that a well-defined 3D structure is the prerequisite for protein function. About twenty years ago, significant amounts of evidence showed that many proteins are fully or partially intrinsically disordered (IDPs) under native states, or functional conditions [176-178]. It was demonstrated that IDPs are very common in nature. Results from 50 different predictors indicate that 15-45% of eukaryotic proteins and 10-35% of prokaryotic proteins have one or more significantly disordered regions, which is more than 30 residues in length [179].

The conformational stability of a typical ordered/globular protein results from a set of non-covalent interactions, such as hydrophobic interactions, hydrogen bonds, van der Waals interactions, and electrostatic interactions. From this traditional view, IDPs/IDRs are structurally unstable, with a low steepness of unfolding transition induced by denaturants or high temperatures.

Although lacking stable structures, IDPs are more likely to be stable in various extreme conditions due to their lack of ordered structures. In contrast, unfolding is very common for ordered proteins under extreme conditions, leading to the loss of biological function. Extended IDPs are known to possess high stability and are resistant to high temperatures or extreme pH values, and therefore maintain their functionality under these extreme environments [189]. A freezing-induced loss-of-function model proved that IDPs are more resistant to cold treatment when compared to globular proteins [192]. Conversely, prothymosin α, an IDP, is still active after having been boiled for a few days [193].

Several studies have shown that IDPs/IDRs are very common in many proteomes, and its level may contribute to both the evolution and adaptation to the environment [200-202]. Transcription factors are known to harbor a significant amount of of function-related intrinsic disorder [180]. The number of transcription factors in hyperthermophiles and psychrophiles is much less than that in mesophiles, suggesting reduced protein disorder in hyperthermophiles and psychrophiles [180]. Low levels of disorder content in hyperthermophiles or psychrophiles were also confirmed by Shen's group [181]. Halophiles, which thrive in high salinity habitats, were found to contain higher disorder compared to the mesophilic counterparts [182].

In general, IDPs/IDRs may play a significant role in an organism's adaptation to challenging environments. Therefore, IDR comparison between *E. coli* and *P. ingrahamii* was performed in order to explore potential composition bias and its effect on the stability of Klenpin.

CHAPTER 2. MATERIALS AND METHODS

2.1 Materials

2.1.1 Buffers and other solutions

Cell lysis buffer: 50mM Tris-HCl, 1mM DTT, 1mM K-EDTA, lysozyme (2mg/g pellet), 1mM PMSF, 10% glycerol, pH 7.5.

Potassium phosphate buffer for denaturation assays: 10mM KPO_4 at pH7.0, 1mM DTT.

Dialysis buffer for protein storage: 50mM Tris-Hcl, 1mM DTT, 50% glycerol, pH7.0

10X SDS running buffer/L: 30.0g Tris base, 144.0g Glycine, 10g SDS.

100ml Stain solution: 0.25 g Coomassie Brilliant Blue R250, 10 ml Glacial acetic acid，90 ml MeOH: H_2O (1: 1 v/v)

Destain solution: Destain solution is the same as the staining solution except for Coomassie R250.

2.1.2 Cell culture medium and *E. coli* strains

Culture medium and plates**:**

Solids (see Table 2.1) for culture media were suspended in about 800 deionized water, and the pH was adjusted to 7.0. Deionized water was added to a total volume of 1L. The medium was then transferred to a flask with aluminum foil and autoclave tape and autoclaved at 121 ℃ for 20 mins. The LB medium was cooled to room temperature, and antibiotics were added if necessary. LB Medium with antibiotics was used as soon as possible since the antibiotics will degrade over time.

For LB plate preparation, additional Agar-agar was added in the LB medium before autoclave. The bench was sterilized with 70% EtOH, and the Bunsen burner was light before depositing the LB medium. 100ml medium is enough for four plates. The LB medium was cooled to room temperature, and antibiotics were added if necessary. The recommended final concentrations for

Ampicillin (AMP) and Kanamycin are 100ug/ml and 50ug/ml, respectively. All LB plates were labeled at the edge of the bottom. Agar was slowly poured until the plate is covered and the volume of agar was around half of the plate. After pouring, the plates were cooled for at least 30 mins or until the agar is solidified. The LB plates were stored at 4°C for future use.

Cloning strains: DH5α (NEB, Ipswich, MA)

Expression strains: BL21(DE3) (NEB, Ipswich, MA). It contains pET24a/D434A-Klenpin plasmid. The position 434 was altered from aspartic acid to alanine prevent any potential 3' to 5' exonuclease activity. The same strategy has been applied to other large fragments of DNA Pol I, such as Klenow [203].

Table 2.1. Composition of bacterial growth media (Reagents were obtained from VWR, Fisher, and Sigma-Aldrich).

Media	Component	Amount g/L
Luria broth (LB)	Tryptone	10
	Yeast extract	5
	NaCl	10
	Agar-agar (plate only)	15
Super Optimal Broth (SOB)	Tryptone	20
	Yeast extract	5
	NaCl	10
	KCL	0.186

2.2 Methods

2.2.1 Cloning

Isolation of genomic DNA from *P. ingrahamii*

P. ingrahamii was generously provided by Dr. Brent Christner. 1ml of *P. ingrahamii* was cultured in Marine Broth [120] around 5 ºC until the O.D. was close to 0.9 (~ 10days). The culture was then spun down, and the pellet was dissolved in 100ul dH2O and then heated at 90 ºC for 15 minutes. After centrifuging, the supernatant containing the genomic DNA was collected. The

genomic DNA was used as a template in the Polymerase Chain Reaction (PCR) to amplify the gene. More details about the cloning work were described in previous studies [7].

2.2.2 Protein purification

Protein overexpression in BL21(DE3)

BL21(DE3) containing pET24a/D434A-Klenpin plasmid was used as the expression cell line. The BL21 cells containing the recombinant plasmid are plated onto an LB plate with 30ug/ml kanamycin and incubated at 37 °C overnight. A single colony from the LB plate was inoculated in a 10 ml LB starter culture with 30ug/ml kanamycin and then cultured in an orbital shaker at 37 °C and 200 rpm overnight. After that, the 10ml culture was added into a 1L LB culture with 30ug/ml kanamycin and cultured at 37 °C and 200 rpm until O.D. 600 was close to 0.6. Protein overexpression is induced by adding 0.8M isopropylthio-ß-D-galactosidase (IPTG) to a final concentration of 1mM and grown in an orbital shaker at 16 °C and 100 rpm for 16 hours. The cell pellets are harvested at 15000 RPM for 40 mins at 4 °C and stored at -80 °C for future purification.

Polyacrylamide gel electrophoresis (SDS-PAGE)

The recipes used for stacking and resolving gels are listed in Table 2.2 and Table 2.3.

Table 2.2. Components for resolving gel

COMPONENT	Volume/ml
dH_2O	4.1
1.5M Tris-Hcl, pH8.8	2.5
20% w/v SDS	0.05
Acrylamide/Bis-Acrylamide 30%/0.8% w/v	2.3
10% w/v Ammonium persulfate (0.1g/ml)	0.05
TEMED	0.005
Total volume	10.005

Table 2.3. Components for stacking gel:

COMPONENT	Volume/ml
dH_2O	3.075
0.5M Tris-HCl, pH6.8	1.25
20% w/v SDS	0.025
Acrylamide/Bis-Acrylamide 30%/0.8% w/v	0.67
10% w/v Ammonium persulfate (0.1g/ml)	0.025
TEMED	0.005
Total volume	5.05

Polyacrylamide gel is prepared and run according to the protocol by Fanglian He [121]. The gel is stained with Coomassie and then destained in dH_2O for 1 hour. The results are analyzed using The Gel Doc XR+ system [288].

Polyethyleneimine (PEI) precipitation

PEI is the polymerized form of ethyleneimine. PEI is positively charged at neutral pH because it has a high pKa, around pH 10 [123]. Based on this property, PEI binds to negatively charged macromolecules, such as DNA and acidic proteins. The binding affinity depends on the concentration of salt. For example, a mildly acidic protein will bind to PEI at 0.1M NaCl, whereas it will elute from PEI at 0.3M NaCl [122]. Based on the analysis of Klenpin, it is a somewhat acidic protein with an estimated pI value at 5.0. As a result, high salt (0.3~0.4M NaCl) was used to elute the protein from PEI, whereas leaving DNA precipitated.

Ammonium sulfate (AS) precipitation

Proteins can be precipitated at high salt concentrations due to the salting-out effect. Although several salts can be used as precipitants, Ammonium sulfate is preferred because of its high solubility in water (up to 4.1M at 25°C) and its stabilization effect on protein structure [122]. Each protein has a different salting-out concentration range, which can be used to separate desired proteins from proteins that are not of interest.

Hydrophobic interaction chromatography (HIC)

Hydrophobic interaction chromatography (HIC) is a technique for separating proteins by their degree of hydrophobicity. The primary mechanism for this technique depends on the interaction between hydrophobic ligands on the media with hydrophobic amino acids exposed on the protein surface. Based on the amino acid composition of Klenpin, hydrophobic residues account for nearly 48% of the total residues. HIC was performed as the initial chromatographic step for the purification.

Ceramic Hydroxyapatite (CHT)

Hydroxyapatite $(Ca_5(PO_4)_3OH)_2$ is a form of calcium phosphate in which calcium doublets and phosphate triplets are arranged in a repeating geometric pattern [124]. Unlike other chromatographic methods, CHT is both the ligand and support matrix. CHT has two types of binding sites, negatively charged phosphate groups and positively charged calcium groups. As a result, proteins can be separated by cation exchange via the phosphate groups and metal affinity via the calcium atoms.

Protein concentration determination

Protein concentration is determined by a NanoDrop protein quantification system at A280nm with an extinction coefficient of 0.762 (https://web.expasy.org/protparam/) and a Bradford protein assay kit at A595nm.

Purification procedure

In general, Klenpin was purified by combining Hydrophobic interaction chromatography (HIC) and ceramic hydroxyapatite (CHT). The cell pellets were dissolved in lysis buffer (50 mM Tris-HCl, 2 mM K^+EDTA, 1 mM DTT, pH 7.5) and stirred at 4°C for 30 minutes. 10 mM PMSF and 10mg/g pellet lysozyme were added to the solution, which was sonicated with a Branson

Sonifier 250 for 6 x 30 sec with a 30 sec interval between each sonication. The supernatant was harvested at 15000 RPM for 40 mins at 4 °C. 10% polyethyleneimine (PEI) at pH 7.5 was slowly added to a final 0.2% to precipitate both Klenpin and DNA from other contaminating proteins. Klenpin was eluted by lysis buffer with 0.3M NaCl, whereas DNA remained precipitated. Ammonium sulfate was then gradually added to 44% saturation, and the mixture was stirred slowly for 1 hour at 4 °C and centrifuged at 15000 RPM for 40 mins at 4 °C. HIC was carried out using a Macro-prep Methyl HIC Support (BIO-RAD, Hercules, CA). Supernatant from Ammonium sulfate precipitation was loaded to the support containing 1.8M Ammonium sulfate, where Klenpin is eluted from the column and other proteins are bound according to the degree of hydrophobicity. Klenpin was located by denaturing SDS-PAGE and those fractions are then combined and dialyzed against dialysis buffer (10mM phosphate buffer,1mM DTT,10% glycerol, pH7.0) overnight. The partially purified Klenpin is loaded to Type I, 80 µM ceramic hydroxyapatite (Bio-Rad, Hercules, CA) column and eluted into 60 fractions using a 10-250 mM phosphate buffer gradient. The concentration of protein was estimated by the absorbance at 280nm and the purity of Klenpin was determined by denaturing SDS-PAGE. The fractions containing Klenpin were dialyzed against storage buffer (10mM Tris-HCl,10Mm DTT,50% glycerol, pH7.0) and stored at -20°C for future use.

2.2.3 Folding analysis

Circular dichroism (CD) measurement

Circular dichroism (CD) spectroscopy is a popular spectroscopic technique for studying the folding of proteins. Circular dichroism (CD) is the differential absorption of left-handed circularly polarized light(L-CPL) and right-handed circularly polarized light(R-CPL) in a sample that contains chiral chromophores [84]. Although CD spectroscopy is widely applied to study all types

of chiral molecules, such as 19 of 20 common amino acids, DNA and RNA, the primary application is to analyze the secondary structure of proteins which are sensitive to the environmental conditions (temperature, pH, ion strength and denaturants). The CD spectrum of a protein is not the sum of the CD spectra of individual amino acids and is highly affected by its 3D structure. As a result, changes in the CD spectra can represent the changes in the secondary structure of a protein. For example, proteins with a helical secondary structure possess two spectral minima at 222nm and 208nm. The addition of denaturants or osmolytes often affects the absorbance at the two peaks. As a result, protein unfolding at increasing temperature or denaturant concentration can be analyzed by CD spectroscopy. Also, this technique is very sensitive and fast and the spectra can be analyzed within minutes. In general, CD spectroscopy is a powerful tool to study dynamic changes in protein structure.

In this project, CD measurements were carried out using a Jasco Model J-815 circular dichroism spectrometer equipped with temperature control. All samples are measured from 25 °C to 64 °C for the thermal denaturation in 10mM phosphate buffer at pH7.0 at a protein concentration of 0.2~0.3mg/ml. Spectra are collected in the range of 200nm ~240 nm, and wavelength kinetics are collected at 218-222 nm. Free energy changes upon unfolding at different temperatures were determined via chemical denaturation with guanidine hydrochloride (GdnHCl) or urea. Denaturants, such as urea and GdnHCl, show high absorbance below 220nm and therefore the absorbance of the protein samples were analyzed at 218-222nm for chemical denaturation.

Equilibrium denaturation

For a chemical denaturation experiment, the following solutions are required: (A) 2.0ml protein solution (0.2mg/ml) in different concentrations of denaturant, (B) a solution containing denaturant only, and (C) a solution without denaturant or protein.

The spectrum of solution B and C in the appropriate cell between 250nm and 190nm was obtained as a baseline. Then the spectrum of Solution A was determined at the same wavelength range. The final CD signal for each denaturant concentration was obtained by subtracting the baseline signal from the CD signal in solution A.

The denaturation curves were fit by the linear extrapolation model [15]:

$$I = \frac{(I_N + m_N[D]) + (I_U + m_U[D])Exp(\frac{-\Delta G_U - m_G[D]}{RT})}{1 + Exp\left(\frac{-\Delta G_U - m_G[D]}{RT}\right)} \qquad (1)$$

I: observed CD signal

$I_N,\ I_U, m_N, m_U$: intercepts and slopes of the native and unfolded state baselines

m_G: slope in the transition region of the curve

ΔG_U: free energy upon unfolding

[D]: concentration of denaturant

T: temperature in Kelvin

R: gas constant

ΔG_U from each equilibrium denaturation measurement can be plotted as a function of temperature using the Gibbs-Helmholtz equation:

$$\Delta G_U(T) = \Delta H_m[(T_m - T)/\,T_m] \ - \Delta Cp[T_m - T\left(1 - \ln\left(\frac{T}{T_m}\right)\right)] \qquad (2)$$

T_m: Melting temperature

ΔH_m: Enthalpy change at melting temperature

ΔCp: heat capacity change

using Tm, ΔH_m and $\Delta Cp,\ \Delta H$ at each temperature can be calculated as follows:

$$\Delta H_T = \Delta H_m + \Delta Cp(T - T_m) \qquad (3)$$

TΔS values are then determined using $\Delta G = \Delta H - T\Delta S$.

Three-state model

For a three-state denaturation (N⇔I⇔U), it can be considered as a combination of two two-state curves as shown in Figure 2.1. Therefore, a three-state model can be written as follows:

$$F\ (N \Leftrightarrow I \Leftrightarrow U) = f_1 * F\ (N \Leftrightarrow I) + (1-f_1) * F\ (I \Leftrightarrow U)$$

F (N⇔I): Transition from Native state to Intermediate state.

F (I⇔U): Transition from Intermediate state to the unfolded state.

f_1: Fraction of F (N⇔I).

F (N⇔I) and F (I⇔U) can be fitted by linear extrapolation model, and therefore the three-state model is shown:

$$I = f_1 \frac{(I_N + m_N[D]) + (I_{I1} + m_{I1}[D]) Exp(\frac{-\Delta G_I - m_{G1}[D]}{RT})}{1 + Exp(\frac{-\Delta G_I - m_{G1}[D]}{RT})} + (1-f_1) \frac{(I_{I2} + m_{I2}[D]) + (I_U + m_U[D]) Exp(\frac{-\Delta G_U - m_{G2}[D]}{RT})}{1 + Exp(\frac{-\Delta G_U - m_{G2}[D]}{RT})} \quad (4)$$

I: observed CD signal

$I_N,\ I_{I1},\ I_{I2}\ I_U, m_N, m_{I1}, m_{I2}, m_U$: intercepts and slopes of the native, intermediate, and unfolded state baselines.

m_{G1}, m_{G2}: slopes in the transition state

$\Delta G_I, \Delta G_U$: Gibbs free energy upon unfolding

[D]: concentration of denaturant

f_1: the fraction of the overall signal change that is due to the N⇔I transition

T: temperature

R: gas constant

Data from three-state denaturation was analyzed by Mathematica (Champaign, Illinois) [236].

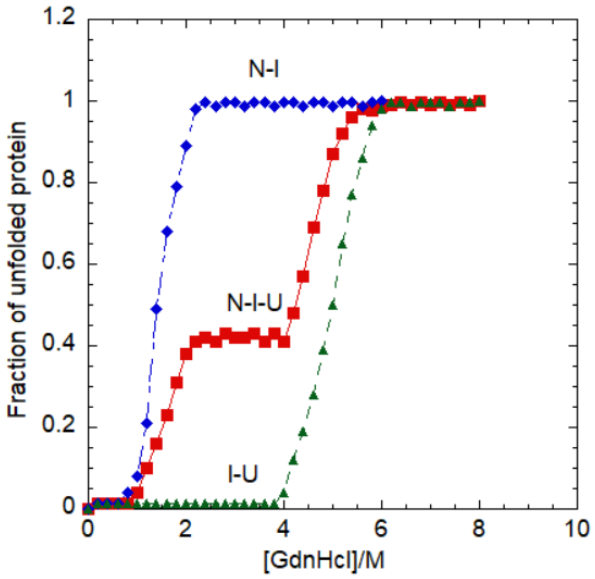

Figure 2.1. An idealized denaturation curve of the three-state model: Native to intermediate (N⇔I), intermediate to denatured (I⇔U).(The figure is adapted from reference 70)

Folding Kinetic Assays

Rate constant and equilibrium constant

For a two-state process, the transition between native state (N) and unfolded state (U) is given by the scheme:

$$U \underset{k_u}{\overset{k_f}{\rightleftharpoons}} N \quad (5)$$

$$\text{d[U]/dt} = -k_f[U] + k_u[N] \quad (6)$$

$$\text{d[N]/dt} = +k_f[U] - k_u[N] \quad (7)$$

where [U] and [N] are the instantaneous concentrations of the unfolded and folded conformations, and k_f and k_u are the folding and unfolding rate constants, respectively, if we denote $[U]_0$ and $[N]_0$ as the initial concentration. In folding experiments, fully unfolded protein is dissolved in a solution without denaturants, and the change of conformation is recorded as a function of time. As a result, $[N]_0 = 0$, and [N] is given by

$$[\mathrm{N}] = [U]_0 - [U] \quad (8)$$

using equation (8), equation (6) reduces to a first-order differential equation:

$$\mathrm{d}[\mathrm{U}]/\mathrm{dt} = -(k_f + k_u)[U] + [U]_0 \quad (9)$$

the solution of which is

$$\frac{[\mathrm{U}]}{[U]_0} = \frac{k_u}{k_f + k_u} + \frac{k_f}{k_f + k_u} \exp\left[-\left(k_f + k_u\right)t\right] \quad (10)$$

the decrease in [U] (or increase in [N]) is single exponential with an relaxation rate constant

$$k_{rel} = k_f + k_u \quad (11)$$

the equilibrium (t→∞) concentration of folded and unfolded conformations ($[N]_\infty, [U]_\infty$) follow the equation:

$$[N]_\infty k_u = [U]_\infty k_f \quad (12)$$

The equilibrium constant (K_{eq}) is defined by

$$K_{eq} = \frac{[N]_\infty}{[U]_\infty} = \frac{k_f}{k_u} \quad (13)$$

Finally, the Gibbs free energy of unfolding is given by:

$$\Delta G_{un} = -RTlnK_{eq} \quad (14)$$

where R and T are the gas constant and the absolute temperature in kelvins, respectively.

Single-jump folding experiments

30ul Klenpin (2.4mg/ml) was dissolved in 90ul 7.5M GdnHCl and pre-incubated at 20 °C for at least 1h. The final GdnHCl concentration is about 5.6M, which was high enough to fully denature Klenpin based on equilibrium denaturation data. 1.88ml 10 mM phosphate buffer with different concentrations of GdnHCl was equilibrated at the same temperature and stirred with a magnetic stirring bar. The previous 120ul Klenpin with 5.6M GdnHCl was added to the buffer, and the folding data was collected at 220nm with an interval of 50 msec. The dead time between

adding the protein sample and closing the cover was about 1-2 sec. The previous steps were repeated in different concentrations of GdnHCl (0.4-1.2M).

Single-jump unfolding experiments

1.97ml 10mM phosphate buffer with different concentrations of GdnHCl (1.0M-3.0M) was equilibrated at the same temperature and stirred with a magnetic stirring bar. 30ul Klenpin (2.4mg/ml) was added to the buffer, and the unfolding data was collected at 220nm with an interval of 50 msec.

Data analysis for folding kinetics experiments

Folding kinetics in the pre-transition and transition regions could be described by a single exponential function [114]:

$$F(t) = F(\propto)(1 - \exp(-k_{app}t)) \qquad (15)$$

k_{rel}: relaxation rate constant of folding and unfolding process

Unfolding kinetics in the post-transition region could be described by a single exponential

Chevron plot for the two-state model

A chevron plot is a way to represent protein folding kinetic data by plotting the logarithm of the relaxation rate constant (k_{rel}) as a function of the denaturant concentration [136].For a two-state model, the dependence of k_{app} on denaturant concentration in both folding and unfolding is linear and described as following [61]:

$$lnk_{app} = lnk(H_2O) + m_k[GdnHCl] \qquad (16)$$

$k(H_2O)$ is the rate constant of folding and unfolding in the absence of denaturant, and the minimum point in the V-shape plot is close to the unfolding midpoint Cm. An idealized chevron plot for a two-state protein is shown in Figure 2.2 [126]:

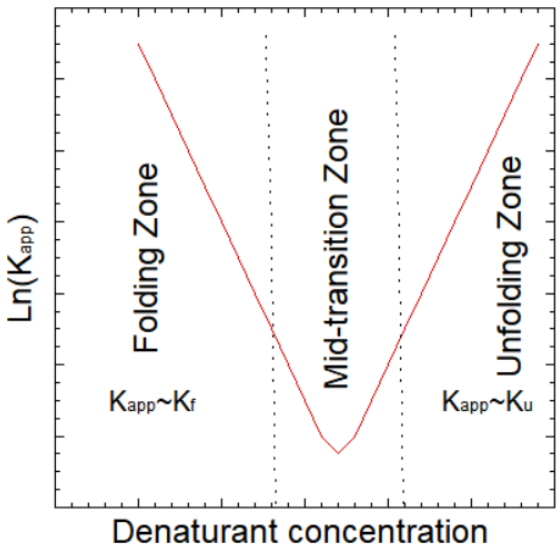

Figure 2.2. An idealized chevron plot showing the relaxation rate constant vs the denaturant concentration. The closed circles represent the folding process when unfolded protein refolds from a denaturation condition to a native or mid-transition condition. The open circles correspond to the unfolding process when a native state protein is denatured to mid-transition or unfolded state. The middle part is called the zone of thermodynamic equilibrium where is folded state is as stable as the unfolded state. (The figure is adapted from figure 9a of reference 126)

As shown in Figure 2.2, the left part corresponds to the folding process when the relaxation rate constant is approximately equal to the folding rate constant, whereas the relaxation rate constant for the denaturation process shown on the right side is approximately equal to the unfolding rate constant. In general, the natural logarithm of the rate constants for both folding ($\log k_U$) and unfolding ($\log k_F$) show a linear relationship with denaturant concentration ([D]), and the linear extrapolations to zero of the folding and unfolding arms of the plot indicate the rates without denaturant ($k_{f,H_2O}, k_{u,,H_2O}$).

The log-linear denaturant relationship between the microscopic and relaxation rate constants is given by

$$\log k_f = \log k_{f,H_2O} + m_f[D]/2.3RT \qquad (17)$$

$$\log k_u = \log k_{u,H_2O} + m_u[D]/2.3RT \qquad (18)$$

$$\log (k_f + k_u) = \log \left(k_{f,H_2O} * 10^{\frac{m_f[D]}{2.3RT}} + k_{u,H_2O} * 10^{\frac{m_u[D]}{2.3RT}} \right) \qquad (19)$$

Equation (19) can be fitted to a chevron plot to get the kinetic m-values [142]: $m_u(\frac{\delta \log k_u}{\delta[D]})$, $m_f(\frac{\delta \log k_f}{\delta[D]})$. A two-state model can be tested by evaluating the equality:

$$m_u + m_f = m_{eq} \quad (20)$$

Where m_{eq} is the m-value from the equilibrium experiment.

For a fully reversible folding process, the rate constant is independent of initial denaturant concentration. With the same final denaturant concentration, therefore, the relaxation rate constants from folding or unfolding experiments are the same. This phenomenon can be easily tested in the transition zone (Figure 2.2).

ΔG (H2O) based on kinetics data

The thermodynamic stability of a protein is defined as the Gibbs free energy change upon unfolding:

$$\Delta G\ (H2O) = -RT\ln K_{eq} \quad (21)$$

R: gas constant

T: the absolute temperature in kelvins.

K_{eq}: equilibrium constant

$$K_{eq}=[U]_{eq}/[N]_{eq}=k_u/k_f \quad (22)$$

combine equation (21) and (22),

$$\Delta G\ (H2O) = -RT\ln(\frac{k_u}{k_f}) \quad (23)$$

Protein Thermal Shift Assay (PTS) using a real-time PCR machine

The thermodynamic stability of a protein may be affected by a lot of environmental factors (temperature, pH, salt, and denaturant). To obtain the optimal buffer condition that favors

maximum stability of proteins, it usually costs a lot of protein samples and time. Therefore, the protein thermal shift assay (PTS) has been developed as a fast and convenient screening method for this purpose. A specific dye, such as SYPRO orange, shows an increase in fluorescence upon binding exposed hydrophobic residues of the unfolded protein. Also, different from other dyes that have been used for PTS, such as 1,8-ANS (1-anilinonaphthalene-8-sulfonate) and 2,6-TNS (naphthalene-6-sulfonic acid), SYPRO orange's fluorescence properties (λ_{ex} 470 nm /λ_{em} 570 nm) are compatible with settings in real-time PCR machines and allow its adaptation to this assay [111].

The protein thermal shift kit was purchased from Thermo Fisher (Waltham, MA, Cat# 4461146). The kit contains the protein thermal shift Dye(5000X). The buffer used in PTS experiment is 10mM phosphate buffer. The ViiA 7 Real-Time PCR System (Thermo Fisher, Waltham, MA) was used for the experiments. Each reaction sample contains 2.5ul 16X dye, 2ul protein(0.2mg/ml), and 15.5ul phosphate buffer at pH7.0. Three control samples contain only dye and buffer. All samples were mixed by pipetting up and down at least five times and added to a 96-well plate. A MicroAmp Optical Adhesive Film was used to seal the plate, which was spun for 1min at 1000rpm to get rid of the bubbles.

Melting temperature and enthalpy changes can be obtained by fitting a two-state model. All samples are measured from 15°C to 95°C in 10mM phosphate buffer at a concentration of 0.2mg/ml.

Data analysis

Unfolding transitions are fitted to a two-state model according to the equation by Kaleidagraph software (Synergy, Reading, PA):

$$I = \frac{I_{0,N} + I_{0,U}.\times e^{\frac{-\Delta H}{1.987*(T+273)}*(1-\frac{T+273}{T_m})}}{1 + e^{\frac{-\Delta H}{1.987*(T+273)}*(1-\frac{T+273}{T_m})}} \qquad (24)$$

I: observed signal intensity;

$I_{0,N}$: signal intensity at native state;

$I_{0,U}$: signal intensity at unfolded state;

ΔH: the change of enthalpy upon unfolding;

Tm: melting temperature where half of the proteins are denatured;

pH dependence of Secondary structure

The secondary structure of Klenpin, Klenow and Klentaq was tested by circular dichroism in different pH buffers: glycine-HCl (pH 2.0–3.0), sodium acetate (pH 4.0–5.0), sodium phosphate (pH 6.0–8.0), glycine-NaOH (pH 9.0–10.0) and phosphate-NaOH (pH 11.0–12.0) [215]. A stock protein solution was added to the appropriate buffer to a final concentration of protein 0.2mg/ml and incubated for 2h at room temperature (25 °C). All data was collected at 220nm.

Dynamic light scattering (DLS)

Dynamic light scattering (DLS) is a technique for measuring time-dependent fluctuations in the scattering intensity due to the random Brownian motion of particles [211]. The particle size can be calculated from the translational diffusion coefficient by Stokes-Einstein equation [212]. The hydrodynamic size of particles is then obtained based on the intensity distribution (Figure 2.3). The overall average hydrodynamic size of particles (Z-average) is used for the size comparison .

DLS is an excellent method to analyze protein folding or other processes in which proteins undergo conformational changes. DLS can thus be used to characterize the size of proteins in different conditions.

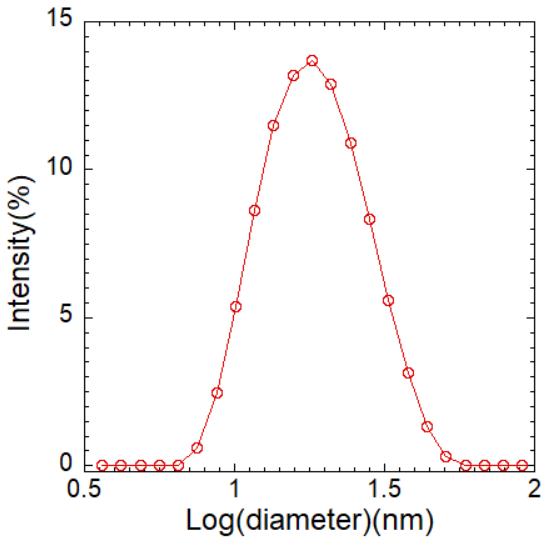

Figure 2.3. The intensity distribution of Klenpin in the absence of salt by DLS.

Tryptophan fluorescence quenching

Tryptophan is known to emit intrinsic fluorescence after absorbing light of a specific wavelength [278]. The fluorescence maximum peak and intensity of tryptophan is strongly affected by any change of the microenvironment [278]. The excited fluorophore can react with other molecules to form excited product, which decays by a non-radiative pathway [283]. This process is called fluorescence quenching. In the native state of proteins, hydrophobic residues, such as tryptophan (Trp), valine (Val) and leucine (Leu), tend to be buried in the interior of the protein. Under the unfolded state, these residues are exposed to the solvent molecules and easily quenched. Therefore, the protein folding process can be studied by analyzing tryptophan accessibility using fluorescence quenching.

In this project, the quencher, acrylamide, was used to investigate the salt dependence of the quenching rate of Klenpin and Klenow. The data were then analyzed using a modified Stern–Volmer equation [280]:

$$\frac{I_0}{I_0 - I} = 1 + 1/(K_{sv}[Q]) \quad (25)$$

I_0: the intensity of fluorescence without a quencher

I: the intensity of fluorescence with a quencher

K_{sv}: a parameter that senses the accessibility of the quencher to the fluorophore

[Q]: the concentration of quencher.

2.2.4 Amino acid composition

Student's T-test

Student's T-test is an excellent tool to test hypotheses about the mean of a sample drawn from a large population and therefore used to compare the amino acid composition between the homologous proteins from *P. ingrahamii* and *E. coli*. The two proteomes sequences were downloaded from the NCBI server in the FASTA format. One hundred pairs of homologous proteins were randomly selected with an average identity of 0.573. Sequence alignment for each pair of homologous proteins was performed by BLAST. The t-values are calculated as follows:

$$T = \frac{\bar{X}}{s/\sqrt{n}} \quad (26)$$

where $\bar{X}$ is the mean of the difference in individual amino acids. S is the standard deviation of the sample, and n is the sample size. The T-test was performed by the SAS program. The null hypothesis in the T-test is that there is no significant difference in the amino acid composition from *P. ingrahamii* and *E. coli.* A P value of 0.05 was used as the cut-off point to reject the null hypothesis.

Acidity and Hydrophobicity of proteomes

The acidity nature and Hydrophobicity of proteomes of halophiles and non-halophiles were compared by Isoelectric Point (pI), and Grand Average of Hydropathy (GRAVY) score distribution. Sequence information of *P. ingrahamii,* 'salt-in' halophiles (*Halobacterium salinarum, Haloarcula marismortui, Natronomonas pharaonic, Halobacterium sp. DL1*) and non-halophiles (*E. coli* K12, *Thermus thermophilus, Deinococcus radiodurans, Pseudomonas*

aeruginosa ATCC) are obtained from NCBI (https://www.ncbi.nlm.nih.gov). All pI distribution data is from a web server: http://isoelectricpointdb.org [232]. GRAVY score is calculated through a web server (http://www.gravy-calculator.de) [235].

Intrinsic disorder regions prediction

Four different predictors (PONDR VLXT, IUPred2a, IsUnstruct, and DisProt) were used to compare the level of intrinsically disordered regions (IDRs). With an accuracy of around 70%, PONDR VLXT integrated three feed-forward neural network predictors: VL1, XN, and XC [195, 196]. IUPred server can predict the IDRs by estimating total pairwise interresidue interaction with an assumption that IDRs are unable to form enough stabilizing interresidue interactions [197]. IsUnstruct can predict IDRs using the Ising model with 70%-85% accuracy [185]. DisProt (VL3H) uses an ensemble of neural networks as the predictor model [199].

Also, IDRs were compared based on the two classes of amino acids: disorder-promoting residues (Pro, Arg, Gly, Gln, Ser, Glu, Lys, and Ala) and order-promoting residues (Cys, Trp, Tyr, Phe Ile, Leu, Val, and Asn) [187, 188].

2.2.5 Homology modeling by Swiss-Model

Swiss-Model is an automated system for modelling the three-dimensional structure of a protein from its amino acid sequence [312]. For a give target protein, suitable templates are identified from a library of experimentally determined structures. A three-dimensional model is then generated based on the sequence alignment between the target protein and the template. The quality of model is evaluated by assessment tools. The three-dimensional structure of Klenpin is predicted from the web server: https://swissmodel.expasy.org/.

2.2.6 Electrostatic surface potential by PyMOL

The PyMOL software is applied to run APBS and visualize resulting electrostatic potentials. The APBS plugin tool is an interface for running APBS (Adaptive Poisson-Boltzmann Solver) electrostatics calculations. The input parameters (such as temperature (310K), ionic charge (+1 and -1), radius (2.0 for +1, 1.8 for -1), and ionic concentration (0.15M)) for all the calculations are the same. Surface charges are considered as continuous "patches" (blue patches for positive charges and red patches for negative charges) (Figure 2.4)

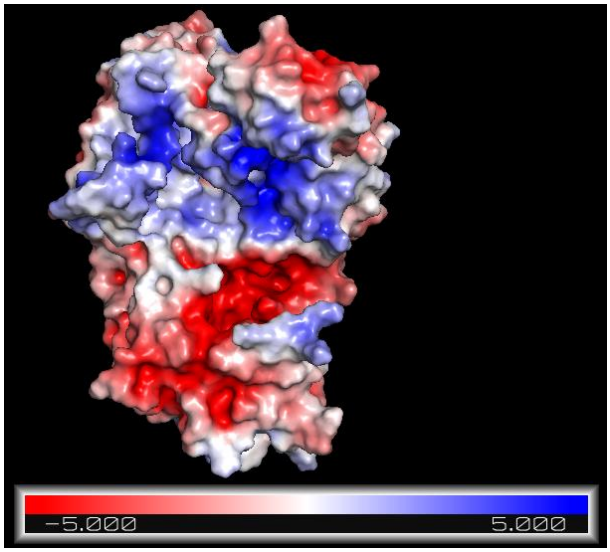

Figure 2.4. The electrostatic surface potential of Klenpin(Swiss-Model). Patches of positive (+) and negative (−) potential are shown, respectively, in blue and red. Neutral patches are white.

2.2.7 Investigation of reversibility of Klenpin folding by Protein refolding column

Protein refolding column (Profoldin, MA) is applied to analyze the reversibility of Klenpin unfolding by urea. Spin-column protein folding screen kit (Catalog #: SFC01-10) is first used to find the optimal folding condition. Then a large-scale preparative protein folding column set (SFC03) is applied for the refolding experiments. The details for the protocol can be found online:(https://drive.google.com/drive/folders/1rmdNdxMZYdABZjwKgoXd510oPZrOan_f).

CHAPTER 3. SALT DEPENDENCE OF THE THERMAL STABILITY OF KLENPIN POLYMERASE

3.1 General Characterization of Klenpin Polymerase

3.1.1 Protein purification

Klenpin expression was induced by IPTG at 16 °C and mainly purified by combining hydrophobic interaction chromatography and Ceramic hydroxyapatite chromatography (see Methods 2.2.2).

Hydrophobic interaction chromatography (HIC)

Purification results after HIC step are shown in Figure 3.1. The primary band in lanes 2-10 corresponds to Klenpin, which has an estimated molecular weight of 68.6kDa. Comparison of lanes 5-10 with lane 4 show some contaminating proteins and partial purification of Klenpin.

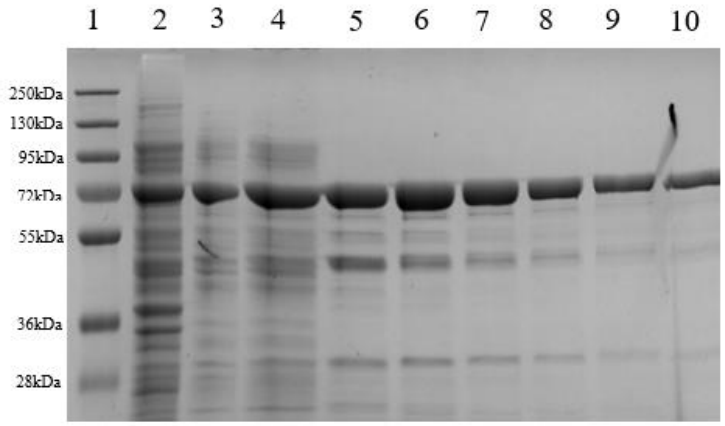

Figure 3.1. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) results after hydrophobic interaction chromatography (HIC). Lane 1: Protein standard. Lane 2: Sample after E.coli sonication. Lane 3: Sample after Polyethyleneimine (PEI) precipitation. Lane 4: Loaded sample for HIC. Lane 5-10: Samples eluted from HIC.

Ceramic hydroxyapatite (CHT) chromatography

Purification results after CHT are shown in Figure 3.2. Most of the contaminating proteins are removed. The purity of Klenpin in lanes 2-8 is above 98%, as determined by ImageJ (LOCI, University of Wisconsin). The samples (~100ml) with high absorbance at 280nm were collected

and concentrated to about 20ml. After dialysis against the storage buffer, the final concentration of Klenpin was determined by measuring absorbance at 280 nm with an extinction coefficient of 0.762 (see Methods 2.2.2).

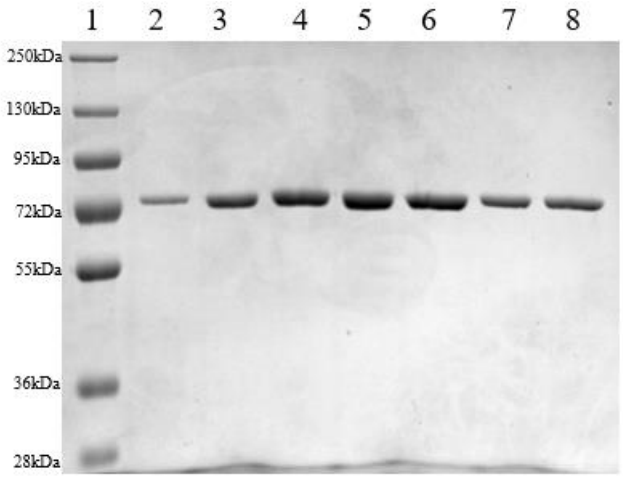

Figure 3.2. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) results after ceramic hydroxyapatite (CHT) chromatography. Lane 1: Protein standard. Lanes 2-8: Samples eluted from CHT.

3.1.2 Protein Sequencing and Identification by Mass Spectrometry (MS)

100ng purified protein was sent to the proteomics facility at the University of Texas at Austin for protein sequence analysis. The sequence alignment results are shown in Figure 3.3. The purified protein sample was digested to peptides with trypsin. Each isolated peptide was ionized by a beam of electrons and accelerated by an electric field. The ionized peptides were then shot into a magnetic field and split into spectra according to their mass and electrical charge. For each MS spectrum, the software Scaffold (Proteome Software, Portland, OR) was used to find the best match from a database of all the proteins from *P. ingrahamii* (3774 proteins). The whole sequence of *P. ingrahamii* Pol 1 has the highest coverage of about 62%, which is consistent with the fact that the purified enzyme is the large fragment domain of full-length DNA pol I. The percent coverage is calculated by dividing the number of amino acids in all found peptides by the total

number of amino acids in the entire protein sequence [277]. The coverage between the results from mass spectrometry and large fragment domain of the sequence is about 94.92%, which should be 100% in theory. Trypsin cleaves peptide chains mainly at the carboxyl side of lysine and arginine, resulting in some peptides that are too short, too long or too hydrophobic [307]. These peptides are poorly ionized or fragmented and cannot be identified by Mass Spectrometry. As a result, the sequence coverage is almost never 100%.

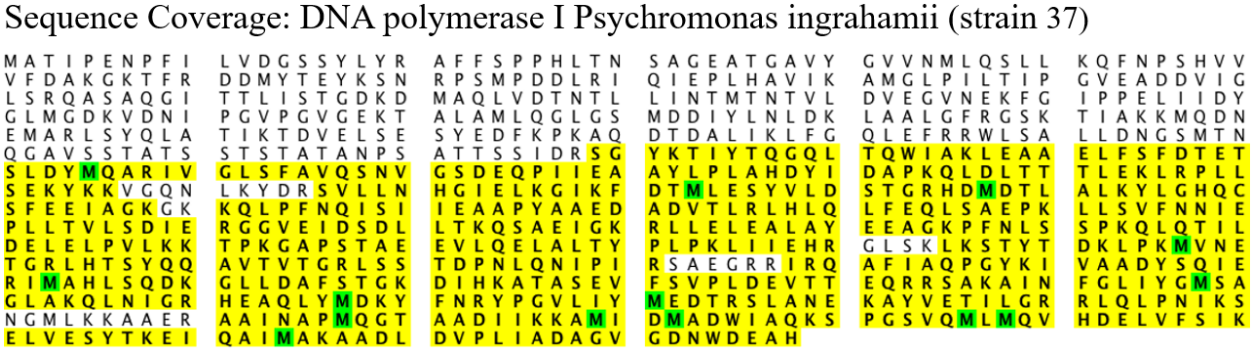

Figure 3.3. Sequence alignment between the results from mass spectrometry and P. ingrahamii DNA polymerase I (Pol I) from the database. Detected peptides are highlighted in yellow.

3.1.3 Spectra of Klenpin by Circular Dichroism (CD)

A comparison of the secondary structures of Klenpin, Klenow, and Klentaq was performed by CD (Figure 3.4). The CD spectra of α-helices have two negative bands at 208nm and 222nm, whereas β-sheets have one negative band at 218nm [8]. The secondary structure of Klenpin, Klenow and Klentaq was estimated by the program BeStSel [151]. Quantitatively, about 33.7% helical content was estimated, whereas Klenow and Klentaq have 50% and 61.5%, respectively (Figure 3.5). As will be discussed in Section 4.3.1, however, the helical content and secondary structure of Klenpin change with added salt. Lower helical content is often observed in psychrophilic proteins [135]. Spectra of Klenpin by CD in the native state (20°C) and denatured state (64°C) are shown in Figure 3.6. At 64°C, there is still a dominant-negative band (-13.7 mdeg)

at 224nm. One possible explanation is that denatured Klenpin forms aggregation, which caused an increase in negative ellipticity.

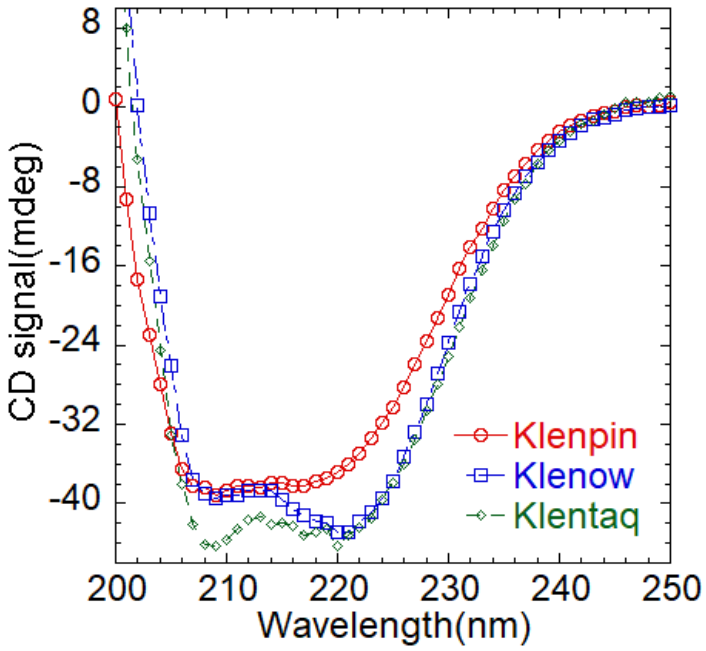

Figure 3.4. CD spectra of Klenpin, Klenow and Klentaq in 10mM phosphate buffer with 50mM NaCl at pH 7.0: Klenpin (empty circles), Klenow (empty squares), and Klentaq (empty diamonds).

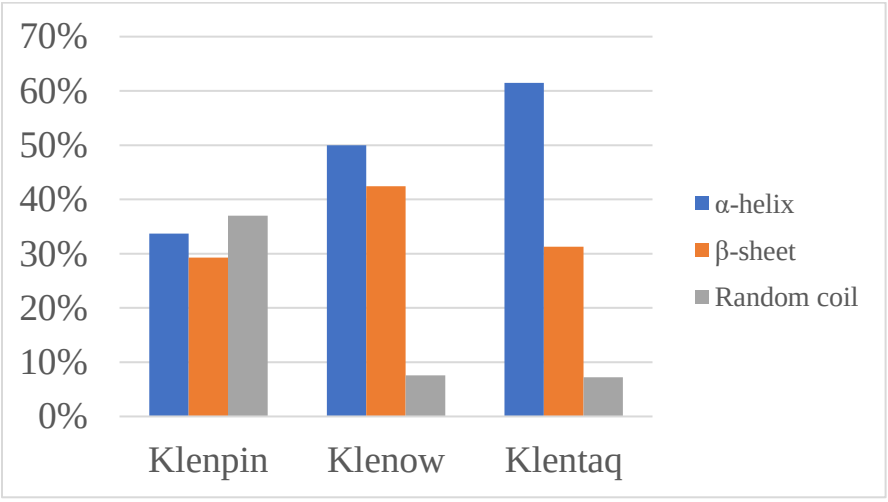

Figure 3.5. Composition of secondary structures (α-helix, β-sheet, and random coil) of Klenpin, Klenow, Klentaq by BeStSel.

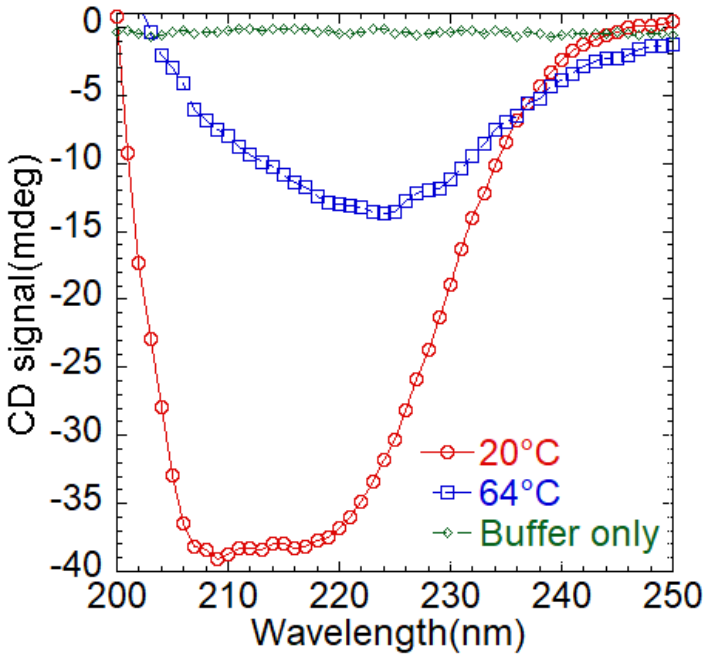

Figure 3.6. Spectra of Klenpin in 10mM phosphate buffer with 50mM NaCl at pH7.0 from 200nm to 240nm at two different temperatures: 20°C (empty circles), 64°C (empty squares), and buffer only (empty diamonds).

3.2 Effect of Salt on the Thermal Stability of Klenpin and Klenow

3.2.1 Salt dependence of thermal stability (ΔTm) of Klenpin and Klenow

As will be discussed in Chapter 4, initial chemical denaturations of Klenpin suggested an unusually low folding free energy, and extrapolation of that data to physiological temperatures for *P. ingrahamii* suggested near complete instability of the protein. For these and other reasons discussed throughout this dissertation, we hypothesized that the high salt in the natural sea ice environment of *P. ingrahamii* might be involved in stabilizing the protein. To test this hypothesis, a series of cations and anions from the Hofmeister series were tested to compare the salt dependence of thermal stability between Klenpin and Klenow. The salt effect on the melting temperature (Tm) of Klenpin and Klenow was tested by protein thermal shift (PTS) assay. Tm and ΔHm were obtained by fitting to a two-state unfolding model (Figure 3.7). Tm and ΔHm from PTS were also compared with that from CD and similar results were observed (Figure 3.8).

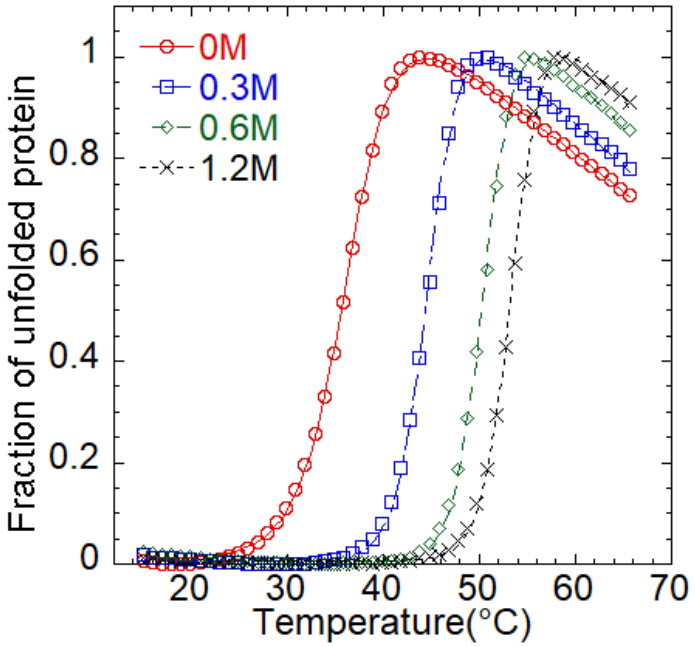

Figure 3.7. Thermal denaturation of Klenpin using the protein thermal shift assay at different NaCl concentrations (lines show the fits to a two-state unfolding equation as described in Methods 2.2.3): 0M(circles); 0.3M(squares); 0.6M(diamonds); 1.2M(crosses).

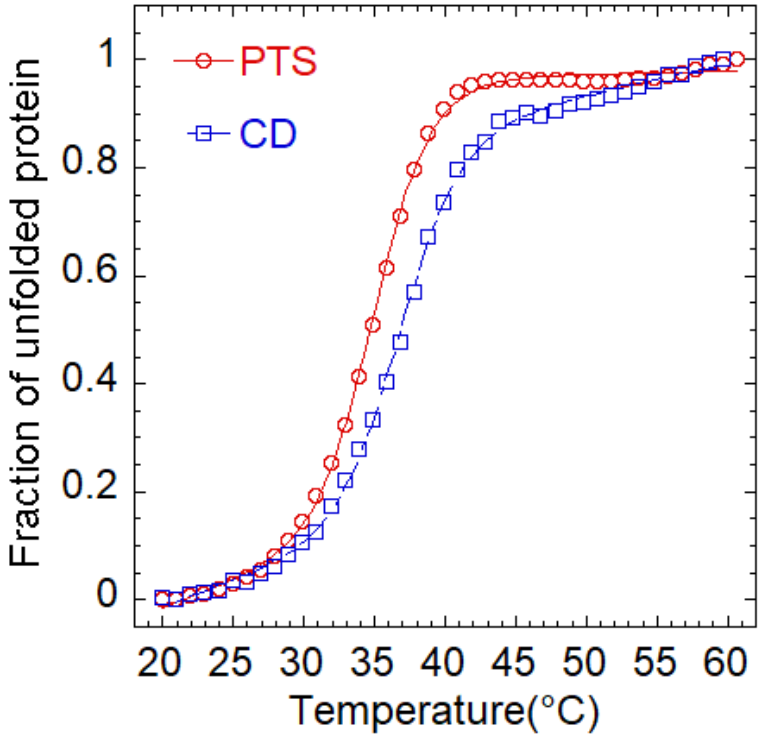

Figure 3.8. Thermal denaturation of Klenpin by PTS (circles) and CD (squares) in 10mM phosphate buffer at pH7.0. Tms from PTS and CD are 35 °C and 36.7 °C, respectively.

The Hofmeister series was used for the ion selection since it generally classifies ions in order of their ability to stabilize proteins [104]. Based on the Hofmeister series, the chlorides of four cations (Na^+, K^+, Mg^{2+}, $\mathrm{GdnH^+}$) and the sodium salts of four anions ($\mathrm{SCN^-}, \mathrm{HCO_3^-}, \mathrm{SO_4^{-2}}, \mathrm{HPO_4^{-2}}$) were selected for testing [104]. The salt concentration ranges tested in PTS assays mainly depends on the solubility limits. For GdnHCl and NaSCN, only low concentration (<0.3M) was tested since high concentration of GdnHCl or NaSCN already denature Klenpin and therefore the unfolding curve and melting temperature are not available.

At each salt concentration, the Tm was measured and compared to the Tm in the absence of salt by calculating $\Delta\mathrm{Tm}$ value:

$$\Delta\mathrm{Tm}=\mathrm{Tm\ (salt+10mM\ phosphate)} - \mathrm{Tm\ (10mM\ phosphate)}$$

If the ΔTm is greater than 0, the salt increases stability. Otherwise, it destabilizes the protein.

Monovalent cations ($\mathrm{Na^+}, \mathrm{K^+}$) and divalent anions ($\mathrm{SO_4^{-2}}, \mathrm{HPO_4^{-2}}$) were tested first using PTS (Figure 3.9). For NaCl, ΔTm for Klenpin increases sharply up to 0.6M and then become less dependent on NaCl concentration, whereas a linear dependence on salt concentrations was observed for Klenow. Similar trends were also observed in other salts. In addition, ΔTm for Klenpin was above 0 in all Na^+ ion conditions and the average maximal ΔTm is 16.5°C. By comparison, ΔTm for Klenow was also higher than 0 for all ion conditions with an average ΔTm 3.9°C over the same salt concentration range.

In general, both Klenpin and Klenow were stabilized by divalent anions and monovalent cations, which is consistent with the Hofmeister series. However, ions have a significantly greater stabilization effect on Klenpin than that on Klenow.

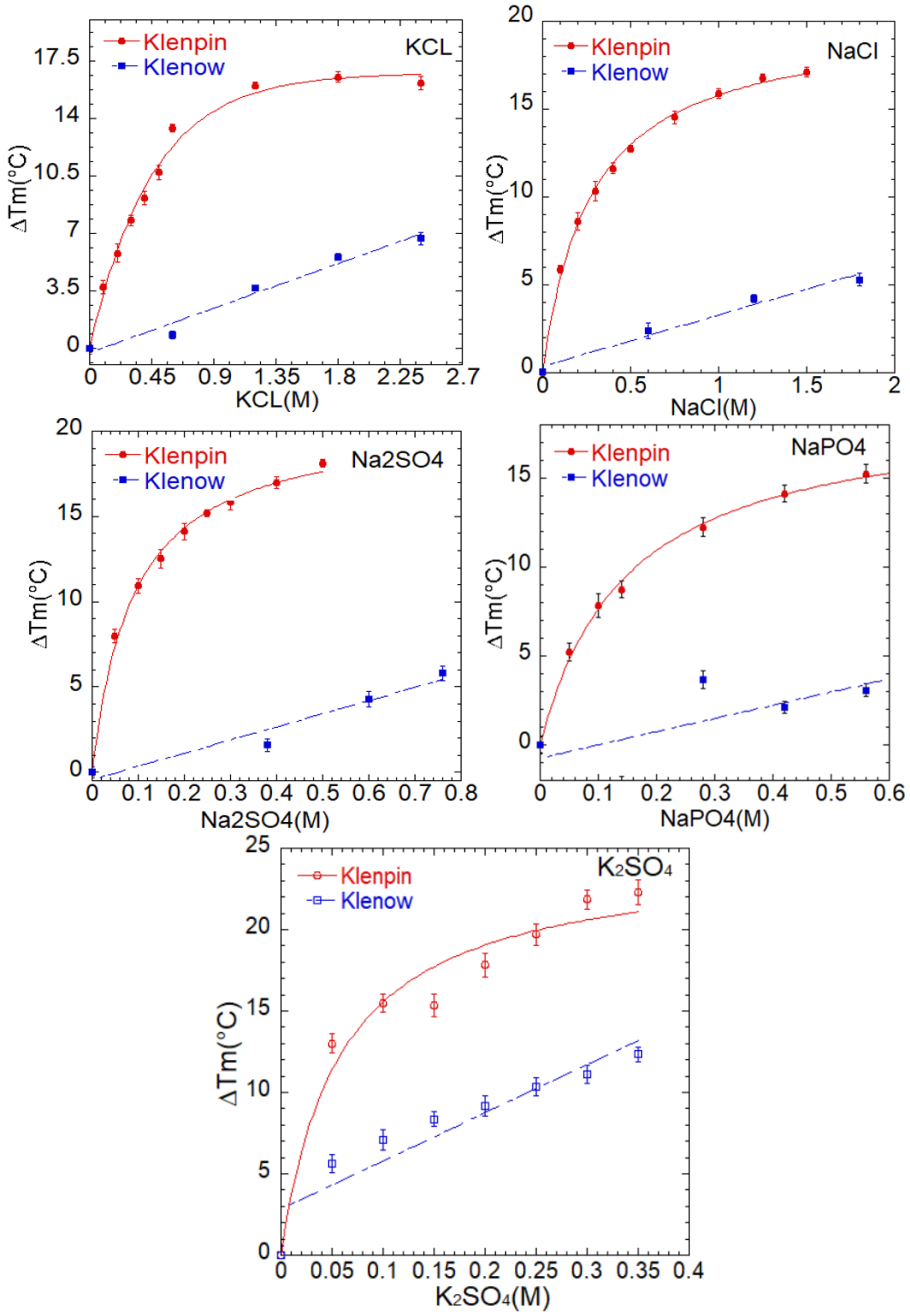

Figure 3.9. The effect of Hofmeister stabilizers ($SO4^{2-}, HPO4^{2-}$, Na^{+}, K^{+}) on ΔTm of Klenpin (circles) and Klenow (squares).

The effect of SCN^{-}, Mg^{2+} and $GdnH^{+}$ on the stability of Klenpin and Klenow were then investigated (Figure 3.10). In general, the effect of the salt on the stability of Klenow was consistent with the Hofmeister series with a negative ΔTm for these salts. For Klenpin, however, the effect of Mg^{2+} showed a bell-shaped curve as Na^{+}. Mg^{2+} stabilizes Klenpin (ΔTm>0) at

all concentration below 2M, whereas it destabilizes Klenpin at high concentration (>2M). This unexpected stabilization has been reported in some halophilic proteins. Tannous et al found that the thermal stability (Tm) of *Halobacterium* species NRC-1 seems to be significantly stabilized by salts (NaCl and $MgCl_2$), which they suggest is due to electrostatic repulsions on the protein surface of *Halobacterium* NRC-1 [310].

HCO_3^- is usually considered stabilizer in the Hofmeister series [169]. However, it caused a negative ΔTm for Klenow and a positive ΔTm for Klenpin, which is same as the destabilizers (SCN^-, Mg^{2+} and $GdnH^+$). One possible explanation is that in the presence of bicarbonate at neutral pH, protein tends to form high charged states (unfolded state) during thermal denaturation which causes the conversion of bicarbonate to CO2 [166].

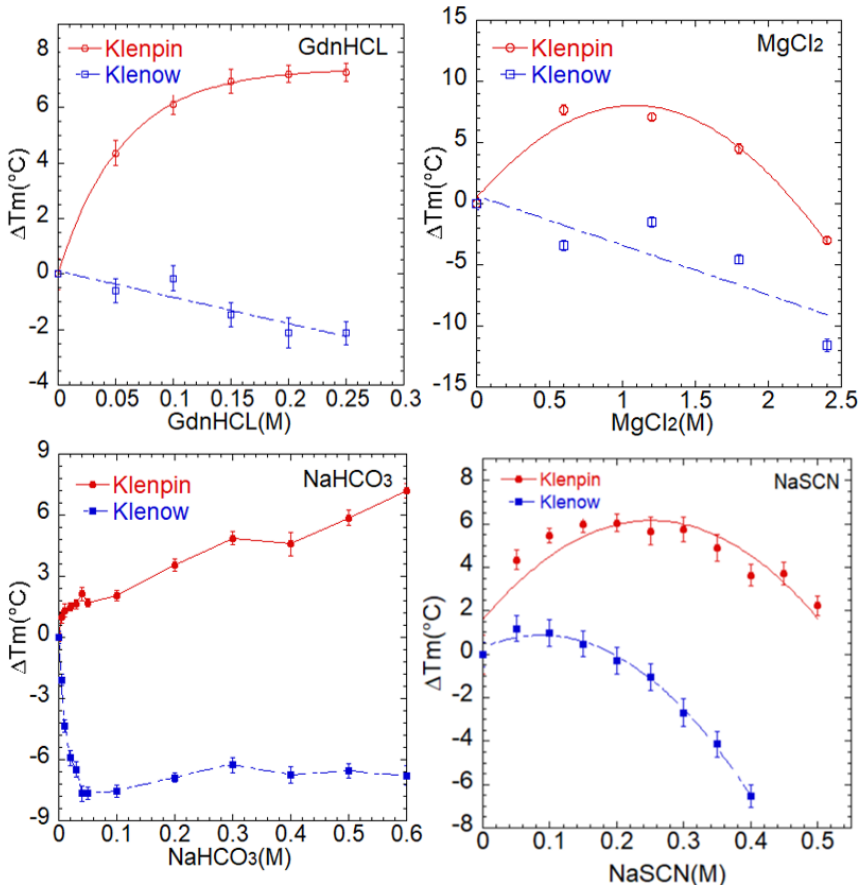

Figure 3.10. The effect of Hofmeister destabilizers ($HCO3^-, SCN^-, Mg^{2+}$) and $GdnH^+$ on ΔTm of Klenpin (circles) and Klenow (squares).

To further investigate the specificity of ion types for stabilization effect on Klenpin, salt concentrations were also converted to ionic strength (I) [289],

$$\mathrm{I} = \frac{1}{2}\sum_{i=1}^{n} c_i z_i^2,$$

Where c_i is the molar concentration of ion i, z_i is the charge number of the ion. The changes in melting temperature as a function of ionic strength are reported in Figure 3.11 and maximal Tm shifts and K_{50} ion dependency values are shown in Table 3.1. For Hofmeister Ions generally considered stabilizers ($Na^+, K^+, SO_4^{-2}, HPO_4^{-2}/H_2PO4^-$), there was a maximal thermal shift around 20°C (18.27~21.09°C) and an apparent binding affinity around 280mM (257-300mM). For Hofmeister Ions generally considered destabilizers (SCN^-, Mg^{2+} and $GdnH^+$), there was a maximal thermal shift less than 10°C (5.40~9.36°C) and an apparent binding affinity less than 80mM (35~80mM). Also, while the Hoffmeister "stabilizers" and "destabilizers" do behave as two separate groups (albeit, both stabilizing the protein in the case of Klenpin), within each separate group there are no additional specific ion effects.

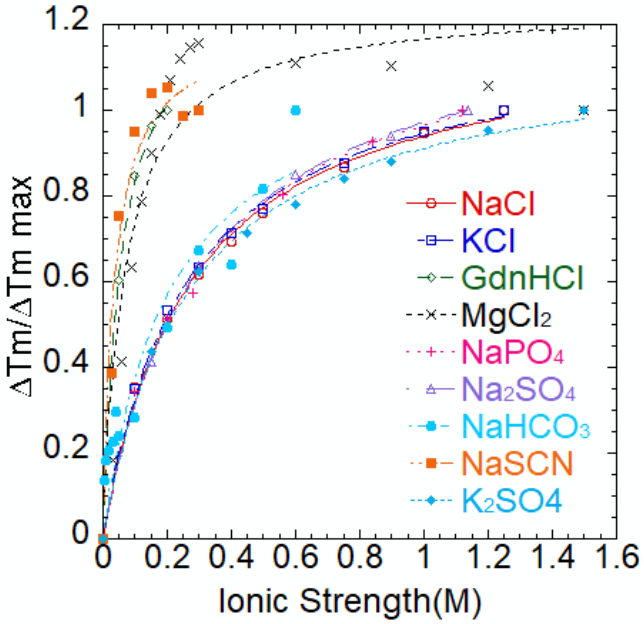

Figure 3.11. Incremental variation of the melting temperature ($\Delta Tm/\Delta Tm_{max}$) for the higher temperature thermal transition of Klenpin as a function of ionic strength of different salts.

Table 3.1. Summary of binding affinity (K_{50}) and maximal Tm shift of Klenpin in different salts.

	K_{50}(mM)	ΔTm_{max} (°C)(fitted plateaus value)
NaCL	273±17	20.07±0.42
Na_2SO_4	300±10	18.27±0.21
Na_nPO_4	298±25	19.07±0.58
KCl	257±11	19.46±0.27
K_2SO_4	264±16	21.09±0.38
NaSCN	35±9	6.88±0.42
GdnHCl	55±5	9.36±0.28
$NaHCO_3$	80±3	5.40±1.30
$MgCl_2$	68±19	7.94±0.50

In general, all tested ions, even the Hoffmeister destabilizers ($HCO_3^-, SCN^-, Mg^{2+}, GdnH^+$), will stabilize Klenpin, whereas salt effect on Klenow follows a typical Hoffmeister effect. Furthermore, the curve shapes for Klenpin resemble binding curves and it fall into two groups: low apparent affinity strong stabilizers ($Na^+, K^+, SO_4^{-2}, HPO_4^{-2}/H_2PO4^-$) and higher apparent affinity moderate stabilizers ($HCO_3^-, SCN^-, Mg^{2+}, GdnH^+$), suggesting that there is some ion "group-based specificity" in the effect on Klenpin.

3.2.2 Stabilization effect by low concentration of GdnHCl

GdnHCl and urea are the most common chemical denaturants used in the study of protein folding and stability. We then compared the effect of GdnHCl and urea on the thermal stability of Klenpin and Klenow (Figure 3.12). For urea, the melting points of both Klenow and Klenpin decreased as the denaturant concentration increased. However, GdnHCl showed an unexpected increase in the melting point for Klenpin at low denaturant concentration up to 0.3M, whereas it destabilized Klenow. This stabilization effect has also been observed in other proteins [106,155,156,281]. A clear understanding of this stabilization effect is lacking because even the destabilizing molecular mechanism for these denaturants is still a subject of debate. The leading

theories are that the denaturants either directly interact with proteins [157,158] or that they affect the surface hydration of proteins [159,160].

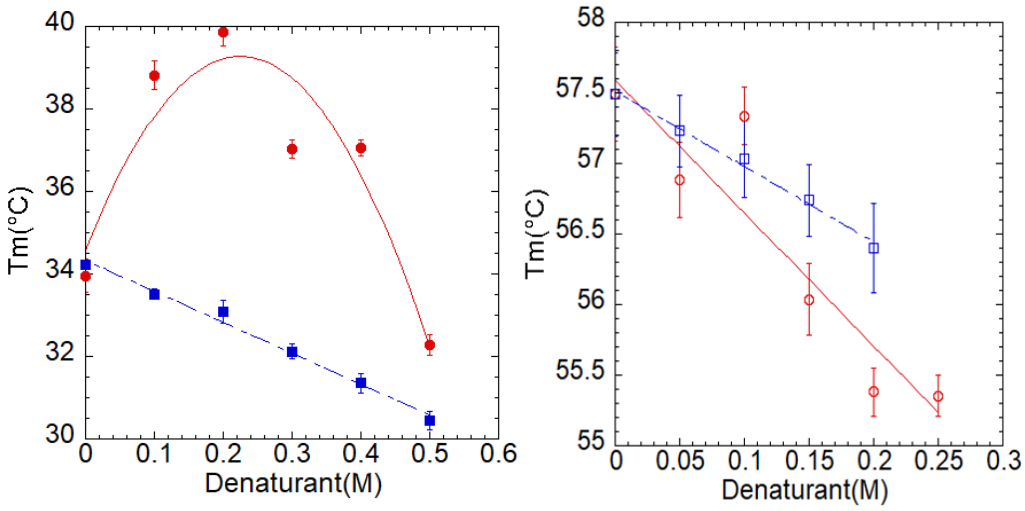

Figure 3.12. Effect of GdnHCl(squares) and urea(circles) on the melting temperature (Tm) of Klenpin(left) and Klenow(right). Error bars for Klenpin and Klenow data are similar but due to significant differences in the y-axis (Tm), the error bars for Klenpin are mostly smaller than the data symbols.

The difference between GdnHCl and urea is that the former exists as a $GdnH^+$ and Cl^- and therefore the stabilization effect may arise from an ionic interaction between Klenpin and $GdnH^+$ or Cl^-. One model proposes that $GdnH^+$-mediated cross-linking cause protein stiffening by forming hydrogen bonds and van der Waals bonds with nonspecific parts of the protein, which reduce the local entropy and increase protein stability [155]. However, $GdnH^+$ or Cl^- can also interact with specific/non-specific binding sites through electrostatic effects, which are also known to stabilize a protein [106].

To test if the stabilization effect of GdnHCl on Klenpin is mainly based on entropic effects or electrostatic effects, we compared this stabilization effect in the presence of different NaCl concentrations (0mM,250mM,500mM) (Figure 3.13). The previous results suggest that NaCl stabilizes Klenpin through electrostatic effects. If GdnHCl stabilize Klenpin through a different mechanism (entropic effects), then the stabilization effects from NaCl and GdnHCl should be

additive (ΔTm (NaCl+ GdnHCl) ~ΔTm (NaCl)+ ΔTm (GdnHCl)). Otherwise, the stabilization of GdnHCl on Klenpin will be offset by adding NaCl. For example, in the presence of both NaCl (500mM) and GdnHCl (200mM), ΔTm (NaCl+GdnHCl) is -4.8°C, whereas ΔTm (NaCl(500mM))+ ΔTm (GdnHCl (200mM)) is 18.7 °C. As a result, the stabilization of GdnHCl on Klenpin is offset in the presence of NaCl. Other salt combination shows similar trend (Figure 3.13). In the presence of NaCl, Klenpin is destabilized by GdnHCl with a negative ΔTm, which is consistent with the Hofmeister series. It indicates that the stabilization effects of GdnHCl is offset with the addition of NaCl, implying that Klenpin is stabilized by NaCl and GdnHCl through the same interactions (electrostatic effects). Another salt (NaSCN) shows a similar trend.

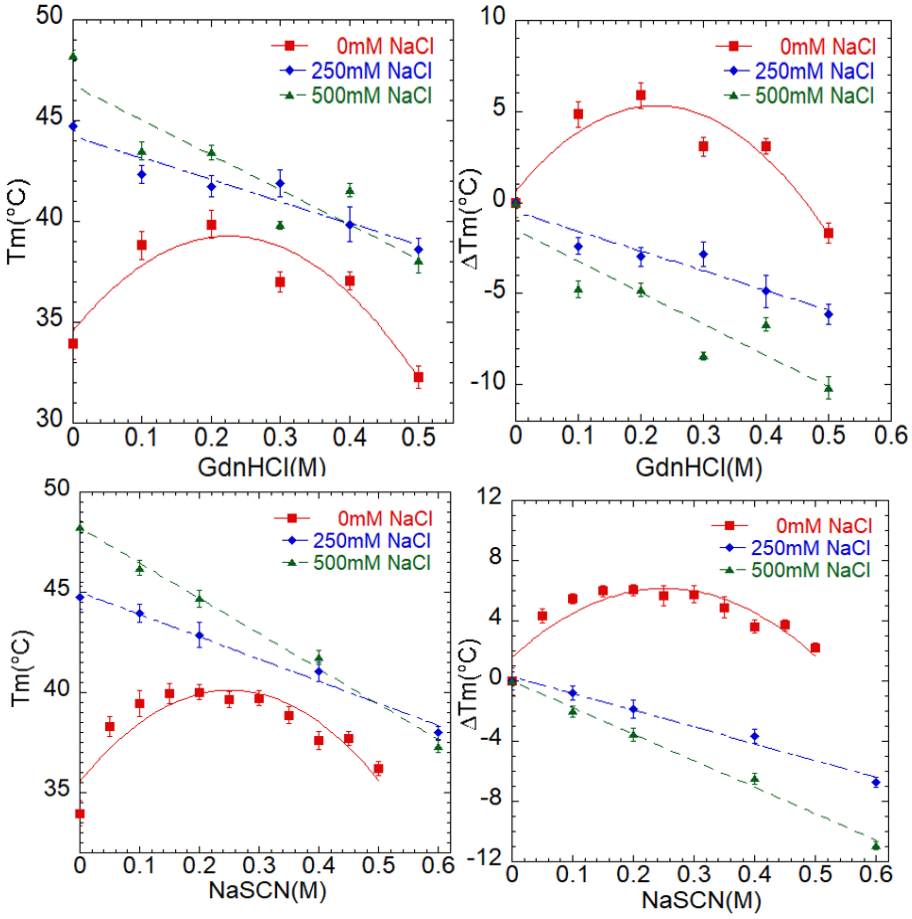

Figure 3.13. Effect of destabilizing ions ($GdnH^+$) on stability (Tm and ΔTm) of Klenpin in the presence and absence of NaCl: 0mM(square), 250mM (diamond), 500mM(triangle).

3.2.3 Osmolyte effect on stability

Osmolyte adaptation is another strategy used by halophiles to maintain osmotic pressure in the cytoplasm. Two genes (gene numbers 2071 and 2072 in the annotations in reference 18) are present in *P. ingrahamii* for enzymes that produce the osmolyte called glycine betaine [18]. Because of this we examined the effect of glycine betaine on the stability of Klenpin and Klenow by PTS and the data was fit by a two-state model (Figure 3.14). The results show that glycine betaine has a slightly enhanced stabilization effect on Klenpin relative to Klenow (Figure 3.15), but relative to the different effects of salt on the two proteins, the osmolyte effect is minor at best (Figure 3.16). while glycine betaine certainly stabilizes Klenpin, there appears to be no significant specific osmolyte effect on the protein-it responds to osmolyte nearly identically to its homologue from *E.coli*. *P. ingrahamii* may use glycine betaine in halophilic adaptation, but it does not appear to be a reason for the unique adaptation mechanism for the organism.

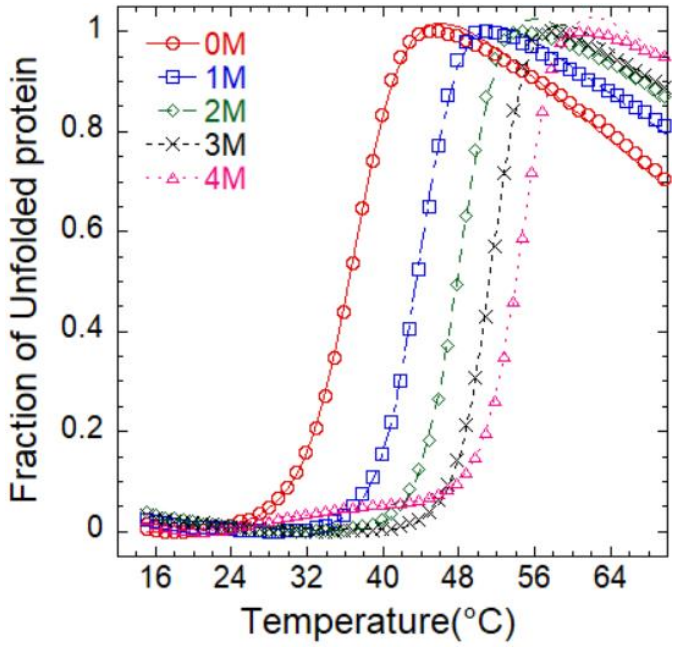

Figure 3.14. Thermal denaturation of Klenpin using the protein thermal shift assay at different NaCl concentrations (lines show the fits to a two-state unfolding equation as described in Methods 2.2.3): 0M(circles); 1M(squares); 2M(diamonds); 3M(crosses); 4M(triangles)

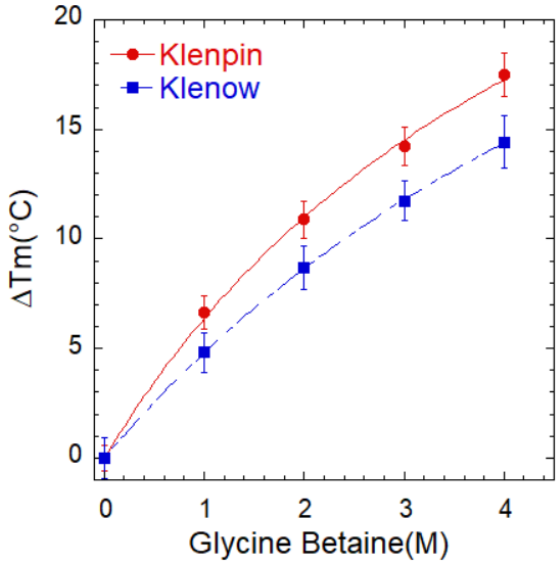

Figure 3.15. Effect of glycine betaine on stability (ΔTm) of Klenpin (circle) and Klenow (square) in 10mM phosphate buffer at pH7.0.

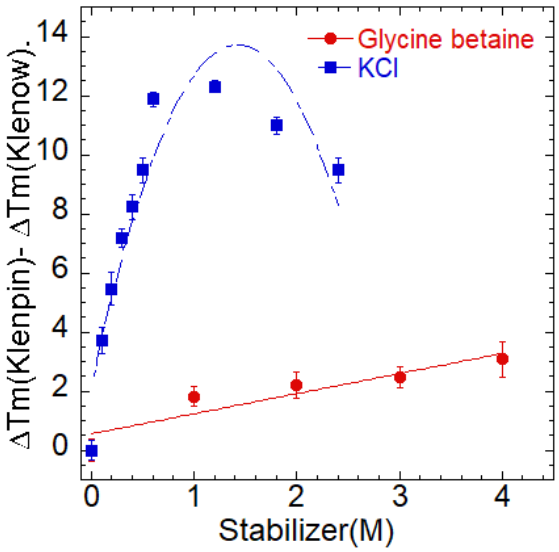

Figure 3.16. A comparison of the difference in ΔTm (ΔΔTm) between Klenpin and Klenow by glycine betaine(circles) or KCl(squares). ΔΔTm=ΔTm(Klenpin)- ΔTm(Klenow).

3.3 pH Effects on Polymerase Secondary Structures

For some halophilic proteins, a high acidic/basic amino acid ratio is often observed on the protein surface. It has been demonstrated that the high negative charge density can maintain repulsive protein-protein interactions in high salt concentration [313]. Another model proposes

that malate dehydrogenase from extreme halophile (*Halobacterium marismortui*) is stabilized in high salt environment through the formation of co-operative hydrate bonds between the protein and hydrated salt ions [223]. Charges on the protein surface are modulated by difference in pH conditions. Therefore, the pH dependence of the overall secondary structures of Klenpin, Klenow, and Klentaq were examined in different pH buffers by CD (Figure 3.17, Methods 2.2.3). Klentaq exhibited secondary structural stability over a wide range of pH values (3-12) in which CD spectroscopy did not show significant changes. With the further reduction of pH to 2, Klentaq showed disruption of secondary structure. Klenow, however, was only stable over a narrow range and rapidly lost secondary structure at low or high pH [10]. Klenpin had a secondary structure pH dependence similar to Klenow.

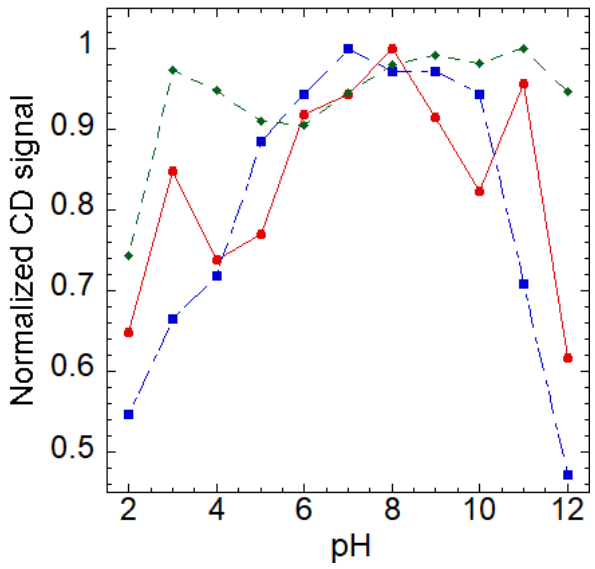

Figure 3.17. pH (2-12) dependence of CD spectra of Klenpin (circles), Klenow (squares) and Klentaq(diamonds). The spectra were recorded at 220nm.

The melting temperatures (Tms) of Klenpin and Klenow were also analyzed at different pH values (Figure 3.18). The raw data were fit using a two-state model (Figure 3.19). At pH above 4.8, the thermal stability of Klenpin was not affected with a narrow Tm range from 31.3°C to

32.2°C, whereas the Tm range for Klenow is from 43.6°C to 53.2°C. At pH4.6, a dramatically decreased Tm was observed for both Klenpin and Klenow.

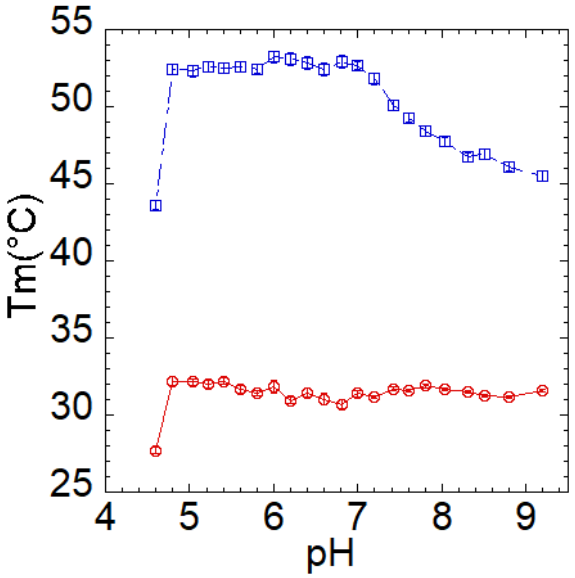

Figure 3.18. Melting temperature points of Klenpin(circles) and Klenow(squares) in different 10mM pH buffers: sodium acetate (pH 4.0–5.0), sodium phosphate (pH 6.0–8.0), glycine-NaOH and (pH 9.0–10.0).

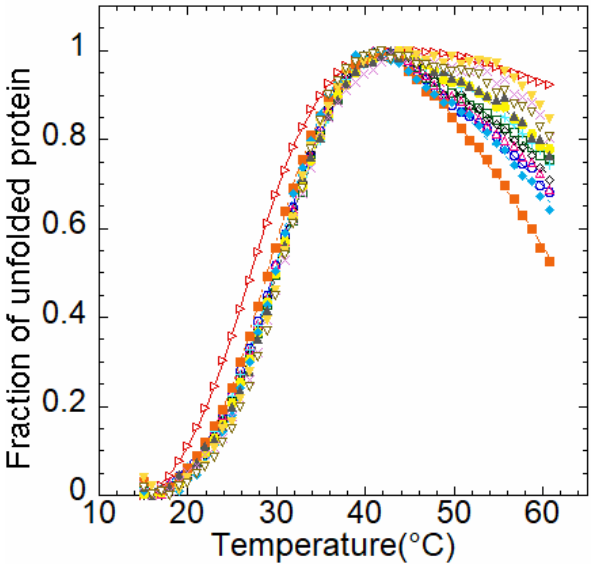

Figure 3.19. Thermal denaturation of Klenpin at different pH:4.6 (open triangles pointing right),4.8(open circles), 5.0(open squares), 5.2(open diamonds), 5.4(open triangles pointing up), 5.6(plus signal), 5.8(filled circles), 6.0(filled squares), 6.2(filled diamonds), 6.4(filled triangles pointing up), 6.6(crosses), 6.8(filled triangles pointing down), 7.0(empty triangles pointing down).

The previous results show that salts have an unexpected stabilization effect on thermal stability of Klenpin. We then investigated the pH dependence of thermal stability of Klenpin at different NaCl concentrations (0mM, 500mM, 1000mM) (Figure 3.20). At pH above 4.8, the melting temperatures of Klenpin are independent of pH conditions for all three NaCl concentrations. At pH 4.6, NaCl has much less effect on the thermal stability (Tm) of Klenpin.

Thermal stability of Klenpin (Tm) was sharply decreased at pH 4.6. We then test the salt effect on the thermal stability of Klenpin at pH 4.6 and compare the results with that at pH 7.0 (Figure 3.21). At pH 4.6, ΔTm for Klenpin linearly increases by adding NaCl, whereas a hyperbola shape curve was observed at pH 7.0. The CD spectra of Klenpin (Figure 3.17) also shows a reduction at pH 4 or 5, suggesting the disruption of secondary structure at low pH.

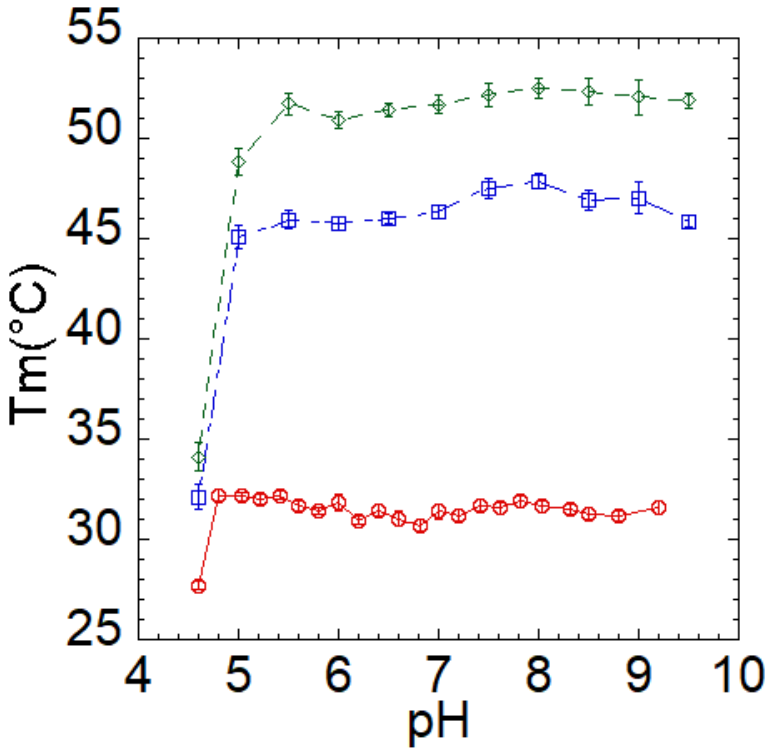

Figure 3.20. pH dependence of thermal stability (Tm) of Klenpin at different NaCl concentrations: 0mM (circles), 500mM (squares), and 1000mM (diamonds).

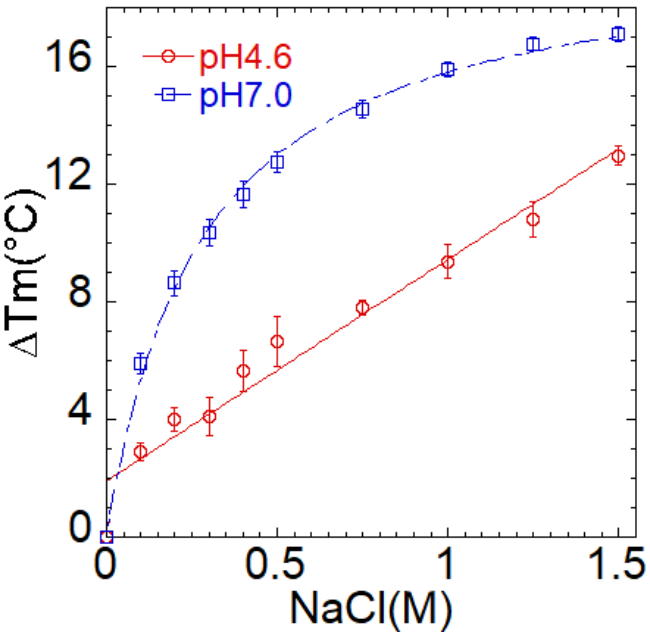

Figure 3.21. Salt dependence of thermal stability of Klenpin at different pH conditions: pH7.0(10mM phosphate buffer), pH4.6(10mM sodium acetate buffer).

3.4 Summary

In general, salts have significantly different effect on the thermal stability of Klenpin and Klenow. With NaCl, ΔTm for Klenpin increases sharply up to 0.6M and then becomes less dependent on salt concentration, whereas a linear dependence of ionic strength on the stability of Klenow was observed. For Klenpin, this trend is consistent with the nonspecific screening of unfavourable electrostatic intramolecular interactions. For Klenow, the trend is consistent with the order of the Hofmeister series ($SO_4^{-2} > HPO_4^{-2}/H_2PO4^- > Cl^- > SCN^-$) [104]. For Klenpin, Hofmeister ions generally considered stabilizers ($Na^+, K^+, SO_4^{-2}, HPO_4^{-2}/H_2PO4^-$) showed a similar stabilization effect with a maximal thermal shift around 20°C (18.27~21.09°C) and an apparent binding affinity around 280mM (257-300mM). Within this group of ions, the salt induced increases in thermal stability of Klenpin (Tm) were independent of the identity of the anion and cation, implying that salts probably improve the thermal stability of Klenpin by the nonspecific

screening of unfavorable electrostatic interactions rather than specific ion binding. The electrostatic screening is weak binding/attractive charge-charge interactions that are independent of ion species, which may still be interacting with specific charged residues on Klenpin surface. For specific ion binding, however, ion is bound tightly to a specific site on the protein and often participate in different types of enzymatic activities of the macromolecules.

Two common denaturants (GdnHCl and urea) are investigated to compare their effect on the thermal stability of Klenpin and Klenow. Klenow is destabilized by both GdnHCl and urea, whereas Klenpin was stabilized by GdnHCl and destabilized by urea. It implies that the stabilization effect of GdnHCl on Klenpin may relate to its ionic strength since that the main difference between GdnHCl and urea. With the addition of NaCl (0~500mM), Klenpin is destabilized by GdnHCl with a linear trend, indicating that the stabilization effect of GdnHCl (<0.4M) on Klenpin is offset by the addition of NaCl. It suggests that GdnHCl and NaCl stabilize Klenpin through the same method, which is nonspecific screening of unfavorable electrostatic interactions.

CHAPTER 4. EFFECTS OF SALT ON THE FOLDING FREE ENERGY AND STRUCTURE OF KLENPIN DNA POLYMERASE

4.1 Thermodynamic Stability of Klenpin by Chemical Denaturation

The thermodynamic stability of Klenpin was studied by chemical denaturation at different temperatures. ΔG_{H2O} values among Klenpin, Klenow and Klentaq were compared.

4.1.1 Chemical denaturation of Klenpin by GdnHCl in the absence of salt

Chemical denaturation of Klenpin was performed by incubation in different GdnHCl concentrations for at least one hour followed by analysis by CD (see Methods 2.2.3). Denaturation curves at five different temperatures (5°C, 10°C, 15°C, 20°C, 25°C) were determined (Figure 4.1 and Table 4.1). In general, Klenpin was found to have a relatively low ΔG_{H2O} (3.11-4.16 kcal/mol) and the midpoint of the transition varies from about 0.8 to 1.0M GdnHCl.

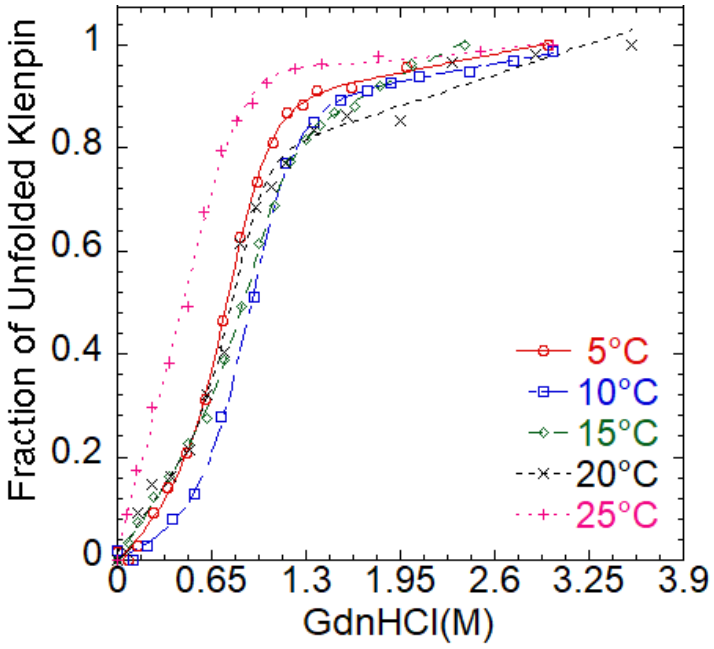

Figure 4.1. Chemical denaturation of Klenpin in GdnHCl by equilibrium measurement in 10mM phosphate, pH7.0, 1mM DTT at different temperature: 5°C (empty circles), 10°C (empty squares), 15°C (empty diamonds), 20°C (the letter x), 25°C(crosses). Lines are fit using the linear extrapolation model, as described in methods 2.2.3.

Table 4.1. Summary of ΔG_{H2O} at different temperatures.

Temperature(°C)	ΔG_{H2O}(kcal/mol)	$D_{0.5}$(M)
5	3.1±0.5	0.8±0.1
10	3.3±0.4	0.9±0.1
15	4.0±0.3	1.0±0.2
20	4.2±0.6	0.9±0.1
25	3.5±0.5	1.0±0.2

4.1.2 ΔG_{H2O} of Klenpin, Klenow and Klentaq

We compared the thermodynamic stability of Klenpin with Klenow and Klentaq, two homologous proteins from *E. coli* (a mesophile), and *Thermus aquaticus* (a thermophile), respectively. The thermodynamic data for Klenow and Klentaq are from previous studies in our lab [162]. Sequence alignment results showed that Klenpin has higher identity with Klenow (66%) than with Klentaq (44%). The comparison among the three organisms is shown in Table 4.2. *P. ingrahamii* has the lowest optional growth temperature (5°C), whereas *Thermus aquaticus* thrives at much higher temperature (70°C). Also, *P. ingrahamii* lives in a high salinity environment up to 2M NaCl. The stability curve (ΔG vs temperature) for the three proteins is shown in Figure 4.1. In general, Klentaq has the highest stability, whereas Klenpin is the most unstable one. All three curves have a similar peak near 20°C, indicating that they all reach the highest stability around 20°C. Similar patterns have been seen with other proteins and one hypothesis for this behavior is that the hydrophobic effect, the main force for protein stability, is optimal at that temperature [94]. Klenpin has the smallest temperature range in which the protein stays folded (-6°C to 43°C).

Table 4.2. Comparison of Klenpin, Klenow, and Klentaq

	Origin	$T_{optimal}$	Salinity tolerance
Klenpin	*P. ingrahamii*	5°C [3]	2M [3]
Klenow	*E. coli*	37°C	0.7M [213]
Klentaq	*T. aquaticus*	70°C	0.07M[214]

In Figure 4.2, each stability curve contains two X intercepts, the melting points for thermal denaturation (right) and cold denaturation (left), respectively. Both Klenpin and Klenow have a similar thermal melting point (~43°C), whereas Klenpin has a higher cold denaturation temperature (~-6°C) than Klenow (~-21°C). This indicates that Klenpin has a similar stability to Klenow at high temperatures, but surprisingly is less stable below 0°C. These results are unexpected because *P. ingrahamii* is a psychrophile and can survive at -12°C; therefore, DNA polymerase I in *P. ingrahamii* should be stable at -6°C. *P. ingrahamii* was isolated from a high salinity environment, and has been shown to have the genes for production of the natural osmolyte called glycine betaine [18].Both salts and osmolytes have the potential to stabilize proteins [285,286]. In Chapter 3, the effect of both salts and osmolytes on the thermal stability (Tm) of Klenpin and Klenow have been investigated and the results from Chapter 3 strongly suggest why these ΔG's are so low. Tm and ΔG are related but independent stability parameters and this chapter will explore the salt effects on the ΔG.

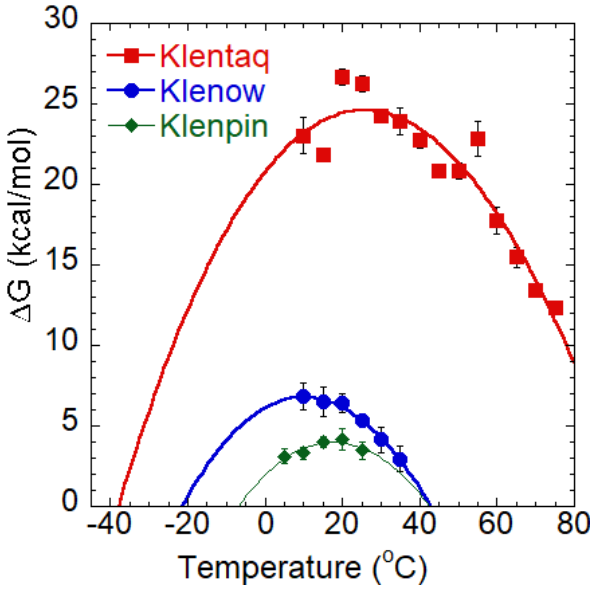

Figure 4.2. Stability curves for denaturation of Klenpin (diamonds), Klenow (circles), and Klentaq (squares) as a function of temperature. The data for Klenow and Klentaq is from previous studies in our lab [162]. The data points were fitted to the Gibbs Helmholtz equation (see Methods 2.2.3).

4.1.3 Temperature dependence of thermodynamic parameters in the absence of salt

Fits to the stability curves (ΔG versus T) can be used to determine the enthalpy, entropy, and heat capacity of folding of the proteins. The calculated folding enthalpies for Klenpin, Klenow, and Klentaq are shown as a function of temperature in the absence of added salt in Figure 3.9. ΔH_{fold} for Klenow is more favorable than that for Klenpin or Klentaq. At low temperature (< 15°C), Klenpin and Klentaq have similar ΔH_{fold}(Figure 4.3). Therefore, the big difference in ΔG_{fold} between Klenpin and Klentaq at low temperature (< 15°C) is probably due to $T\Delta S$.

Next we compared the calculated folding enthalpies, entropies, and Gibbs free energies for Klenpin, Klenow, and Klentaq at 20°C (Table 4.3, Figure 4.4), where all the proteins are close to their maximal stability. By comparing Klenpin and Klenow, the difference in ΔH_{fold} is about 29 kcal/mol, whereas the difference in $T\Delta S_{fold}$ is 26.8kcal/mol. For Klenpin and Klentaq, the difference in ΔH_{fold} is about 8.7kcal/mol, whereas the difference in $T\Delta S_{fold}$ is about 31.2kca/mol. As a result, the stability difference (ΔG_{fold}) between Klenpin and Klenow are mainly related to both ΔH_{fold} change and $T\Delta S_{fold}$, whereas the difference between Klenpin and Klentaq is mainly due to $T\Delta S_{fold}$.

Curvature of the stability plot reflects the difference in heat capacity (ΔCp) between the native and the unfolded states [162]. Klenpin has a similar ΔCp compared with Klenow (3.77 kcal/molK vs 3.70 kcal/molK). It indicates that the instability of Klenpin at cold temperatures is not due to the changes in ΔCp, but probably because of the lower ΔG_{fold}.

Table 4.3. Thermodynamic Stability comparison among Klenpin, Klenow, and Klentaq.

	ΔH_m	$\Delta H_{fold\ at\ 20\,°C}$	Tm(°C)	ΔCp	$\Delta G_{fold\ at\ 20°C}$(kcal/mol))
Klenpin	-98.8±15	-14.2±1.3	42.6±2.4	3.77±0.9	-4.2±0.6
Klenow	-128.3±12	-43.2±3.1	43.0±1.2	3.70±0.8	-6.4±0.6
Klentaq	-255.2±26	-5.5±0.8	94.1±3.9	3.37±0.7	-26.7±0.5

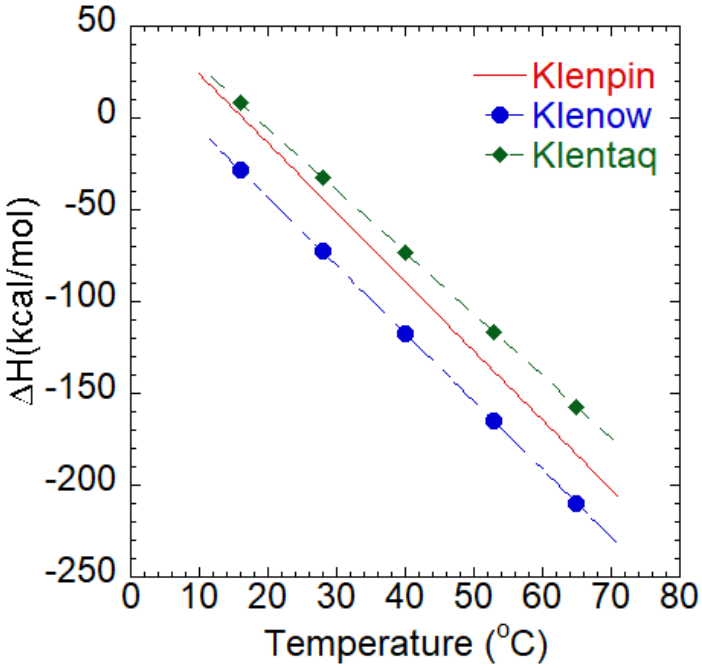

Figure 4.3. Temperature dependence of ΔH_{fold} during protein folding for Klenpin, Klenow, and Klentaq. The data of Klenow and Klentaq was from previous studies in our lab [162]

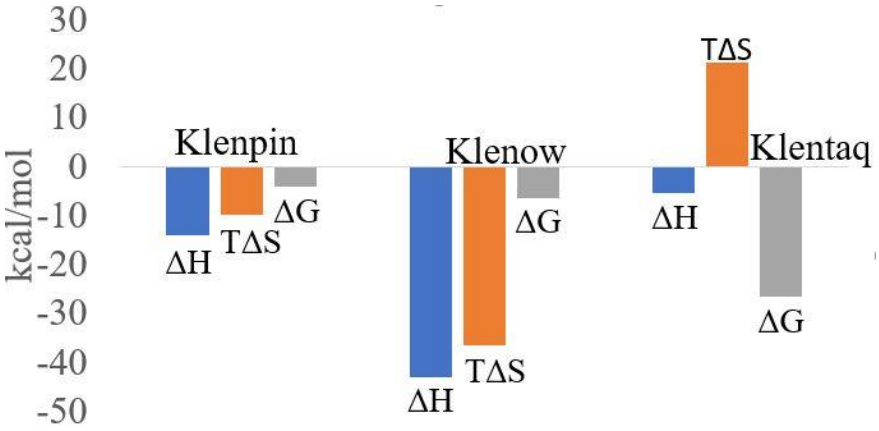

Figure 4.4. $\Delta H_{fold}, T\Delta S_{fold}$ and ΔG_{fold} values for the folding of Klenpin, Klenow, and Klentaq at 20°C (data from Figure 3.12 and other publications [162]).

4.1.4 Summary

In general, all three proteins (Klenpin, Klenow, and Klentaq) have the highest stability around room temperature (~20°C). According to Gibbs-Helmholtz analysis, the thermal unfolding of Klenpin and Klenow occurs at a similar temperature (Tm ~43°C), whereas Klentaq unfolds around

90°C. Georlette et al show that the psychrophilic DNA ligase (*P. haloplanktis*) has the lowest calorimetric enthalpy (ΔH, 46.1kcal/mol) compared with its mesophilic (*E.coli*, 153.1kcal/mol) and thermophilic (*T.scotoductus*, 413.24kcal/mol) counterparts[245], whereas Klenpin has slightly less ΔH_m than Klenow (98.8 vs 128.3kcal/mol). Interestingly, Klenpin has a significantly higher cold denaturation point (~ -6°C) than Klenow (~ -21°C). However, the environmental temperature range for *P. ingrahamii* (-12°C ~ 20°C) extends past the cold denaturation temperature for Klenpin (~-6°C) (Figure 3.13). The unexpected cold instability may come from the difference in conditions between *in vitro* (low salinity) and *in vivo* environment (high salinity) as showed in Chapter 3 and explored further below (section 4.3).

4.2 Chemical Denaturation of Klenpin: Reversibility and Kinetics

In the previous section, we directly compared the thermodynamic stability (ΔG) of Klenpin, Klenow, and Klentaq in equilibrium unfolding experiments. For chemical denaturation to be considered thermodynamically rigorous, two practical factors need to be considered: reversibility and establishment of equilibrium.

4.2.1 Reversibility of Klenpin folding and unfolding

To examine the reversibility of Klenpin folding, we first analyzed the CD spectra of three protein samples: Klenpin in its native state, Klenpin after denaturation, and denatured Klenpin dialyzed back into phosphate buffer (Figure 4.5).

In its native state (circle), the negative band at 222nm represents the α-helix and β-sheet contents of Klenpin. After denaturation, the band disappears (squares) as expected, indicating that Klenpin was unfolded. After dialysis (diamonds) to remove denaturant, the spectrum overlays the data in its native state, indicating the Klenpin reversibly regains its original secondary structure. High concentrations of GdnHCl cause significant background noise at lower wavelengths

(<215nm) (Figure 4.5), thus CD signals only at higher wavelengths were recorded for those experiments.

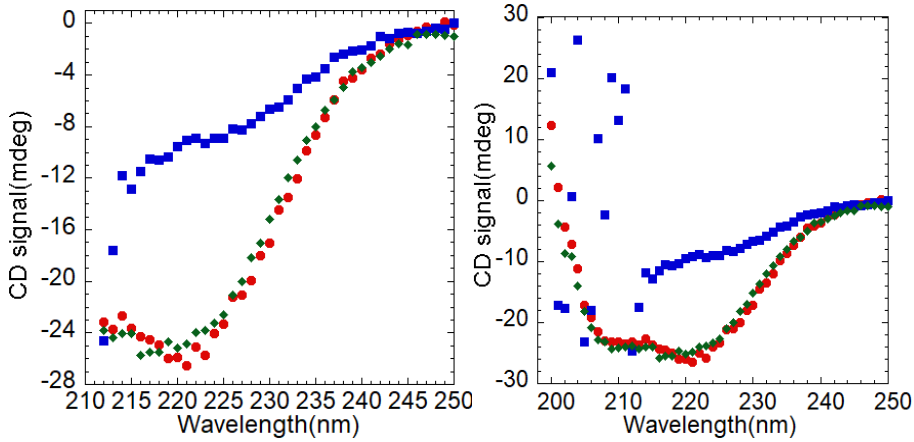

Figure 4.5. Comparing the CD spectra of three Klenpin samples at 20°C: native state (circles), denatured in 3M GdnHCl(squares), and dialyzed against phosphate buffer without GdnHCl(diamonds). Left: partial wavelength range; Right: full wavelength range.

4.2.2 Kinetics of Approach to Equilibrium

The reversibility of protein folding was also validated by directly comparing the forward and reverse folding relaxation rate constants, which should be the same when the system reaches equilibrium. The midpoint of the transition varies from about 0.8 to 1.0M GdnHCl. Therefore, folding and unfolding kinetics were examined by 1) concentration jumps up from 0M to GdnHCl concentrations between 0.6M and 1.55M GdnHCl, and 2) dilution jumps from 5M to GdnHCl concentrations between 0.4 and 1.2M.

Native Klenpin was added to 1.2M GdnHCl (0M → 1.2M) (Figure 4.6, left). By comparison, denatured Klenpin was diluted into 1.2M GdnHCl (5M → 1.2M) (Figure 4.6, left). In both directions, it takes about 8-10 mins to reach equilibrium. The relaxation rate constants (k_{rel}) were calculated by fitting the plot with a first-order equation (Figure 4.7, see Method 2.2.3). The results show that $log k_{rel}$ for folding and unfolding are the same (~-2.12 s^{-1}) when the final GdnHCl concentration is 1.2M (Table 4.4), indicating that Klenpin folding follows the same single

exponential process. Also, the folding direction of Klenpin was investigated at low GdnHCl concentration (0.6M) by either adding native Klenpin into 0.6M GdnHCl (0M → 0.6M) or diluting denatured Klenpin into 0.6M GdnHCl (5M → 1.2M) (Figure 4.6). Similar results were also observed and both directions have a $log k_{rel}$ around -2.5 s^{-1} (Table 4.4). In general, both refolding and unfolding processes of Klenpin are reversible and the relaxation rate constants only depend on the final concentration of the denaturant.

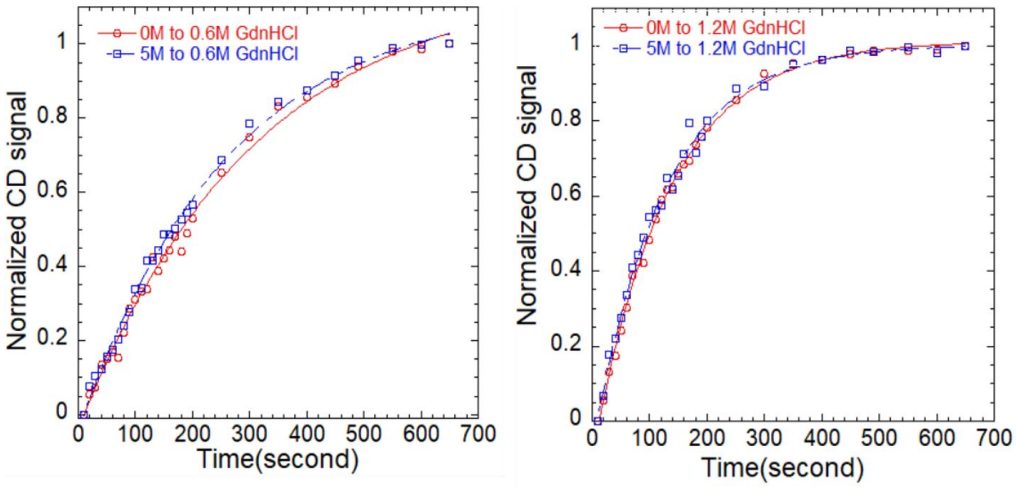

Figure 4.6. Unfolding (1.2M GdnHCl, left) and refolding (0.6M GdnHCl, right) of Klenpin as a function of time by CD at 20°C.

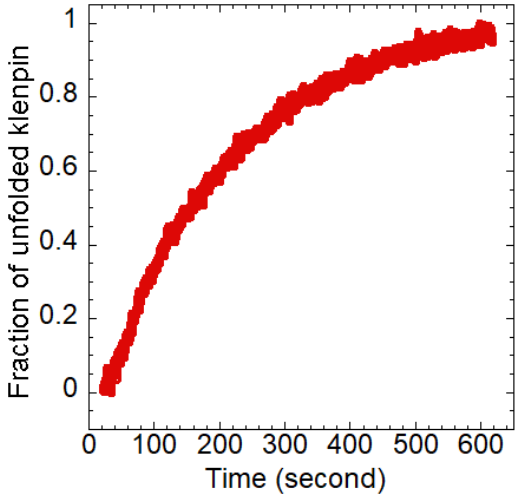

Figure 4.7. Representative unfolding of Klenpin by rapidly mixing folded Klenpin with 1.1M GdnHCl at 20°C. The signal was recorded by CD.

4.2.3 Chevron plot of Klenpin folding

Unfolding and refolding Klenpin kinetics at different GdnHCl concentrations were investigated at 20°C by CD spectroscopy and the relaxation rate constants (k_{rel}) were calculated by fitting the plot with a first-order equation (see Method 2.2.3) [194] (Figure 4.6, Table 4.4).By plotting the logarithm of the relaxation rate constants (k_{rel}) as a function of the denaturant concentrations, we obtain a V shaped plot (a Chevron plot) to represent the folding and unfolding processes of Klenpin (Table 3.4, Figure 4.8). $\log k_{rel}$ values are linearly fitted as a function of GdnHCl concentrations. These rate constants can be extrapolated back to zero molar denaturant to give the folding rate constant in the absence of denaturant ($logk_{fold}^{H_2O}$) and the unfolding rate constant in the absence of denaturant ($logk_{unfold}^{H_2O}$) (Figure 4.8). These values can then be used to determine the thermodynamic stability of Klenpin in the absence of denaturant ($\Delta G = RT\log\left(\frac{k_{unfold}^{H_2O}}{k_{fold}^{H_2O}}\right) = 1.987 * 298.15 * 5.9$. ΔG_{unfold} calculated from the Chevron plot is 3.5±0.4kcal/mol, which is within error of the ΔG_{unfold} (4.2±0.6kcal/mol) from equilibrium denaturation experiments (Table 4.1). We also fit the data for $logk_{unfold}^{H_2O}$ using both folding and unfolding rate constants, the new $logk_{unfold}^{H_2O}$ (-5.4) and ΔG (3.3±0.4kcal/mol)are both close to the original $logk_{unfold}^{H_2O}$ (-5.8) and ΔG (3.5±0.4kcal/mol). Moreover, the denaturation midpoint (0.8M GdnHCl) at which both the folded and unfolded states are equally populated is close to the $D_{1/2}$ (0.93M GdnHCl) at 20°C from equilibrium denaturation experiments (Table 4.1). These experimental consistencies suggest that Klenpin folding/unfolding follows a two-state model in the absence of salt [256]. In addition, the linearity of the folding and unfolding arms of the Chevron plot, the symmetric shape of the Chevron plot, and the overlap of the relaxation constants in the up and down jumps all indicate two-state behavior.

Table 4.4. The relaxation rate constants ($log\mathrm{K_{rel}}$) of Klenpin folding/unfolding at different GdnHCl concentrations.

GdnHCl (M)	Folding (s^{-1})	Unfolding (s^{-1})
0.40	-1.88±0.04	
0.50	-2.11±0.05	
0.60	-2.47±0.07	-2.50±0.06
0.70	-2.80±0.07	
0.80	-2.92±0.09	
0.90	-2.60±0.06	
1.00	-2.48±0.06	-2.56±0.10
1.10		-2.30±0.08
1.20	-2.12±0.09	-2.24±0.07
1.30		-1.71±0.10
1.40		-1.62±0.10
1.50		-1.23±0.09
1.55		-0.90±0.09

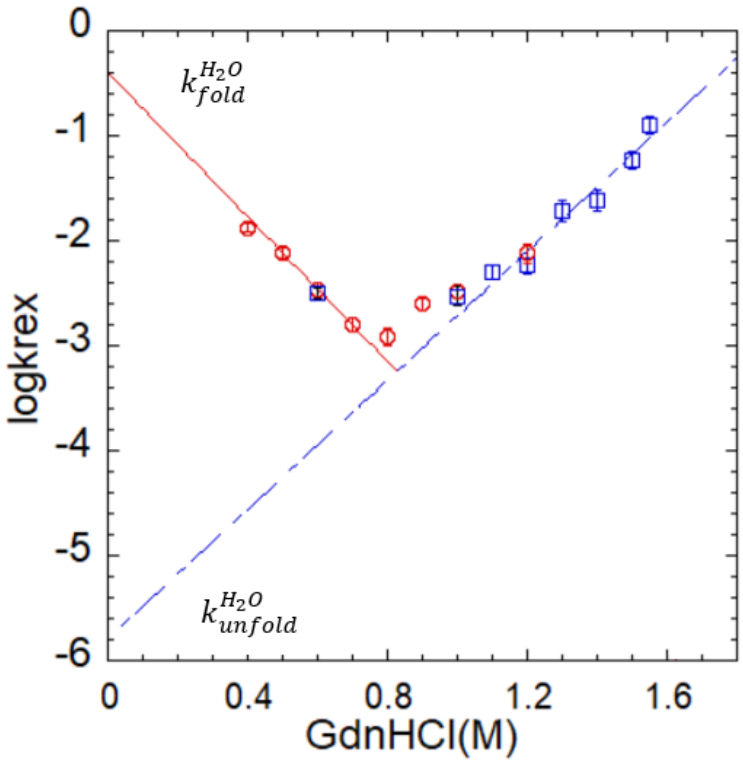

Figure 4.8. The natural logarithm of the relaxation rate constants is plotted as a function of denaturant concentration (Chevron Plot). Each dot represents a separate kinetic experiment and the error bars represent the standard deviation of square rate constants from three replicates.

4.3 Effect of Salt on Chemical Denaturation of Klenpin and Klenow

Results from protein thermal shift assay (3.5.1) have shown that ions have a much greater stabilization effect on Klenpin than that on Klenow. To directly compare the effect of salt on the thermodynamic stability of Klenpin and Klenow, we compared ΔG_{unfold} in the presence of salt. Both urea and guanidine are used as denaturants herein for comparison. For GdnHCl, only one NaCl concentration (500mM) was tested since GdnHCl is also a salt and will cause overlapping effects on Klenpin. For urea, a series of NaCl concentrations (100mM,250mM,500mM) were investigated. Klentaq is too stable to be denatured by urea and therefore was not used in this assay.

4.3.1 Reversibility of Klenpin folding/unfolding by urea in the absence of salt

The reversibility of Klenpin denaturation by GdnHCl has been discussed in section 4.2. For urea denaturation, Protein folding column (ProFoldin, MA) is used to investigate the reversibility of Klenpin folding/unfolding. Denatured Klenpin (in 8M urea) is loaded to the protein folding column(see methods 2.2.7) and then dialyzed against 10mM phosphate buffer at pH7.0 overnight. The CD spectrum from dialyzed sample mainly overlays the data in its native state, indicating the Klenpin reversibly regains its original secondary structure (Figure 4.9). Furthermore, both native sample and refolded sample have a similar ΔG_U (12.2 kcal/mol vs 11.8 kcal/mol), supporting its reversibility feature.

Urea denaturation of Klenpin is then investigated at different temperatures (5°C,10°C, 15°C, 20°C, 25°C) and the stability curved is compared with that from GdnHCl denaturation (Figure 4.10, Table 4.5). Unexpectedly, urea denaturation of Klenpin has much higher ΔG_U (12.2±2.6 kcal/mol at 20°C) than that from GdnHCl denaturation (4.2±0.6 kcal/mol at 20°C). For Klenow, however, ΔG_U from urea denaturation (5.2±0.7 kcal/mol at 20°C) is similar to that from GdnHCl denaturation (6.4±0.6 kcal/mol at 20°C). For the Tm estimate from stability curve (Table 4.50),

the Tm from GdnHCl method (42.6±2.4 °C) is closer to that from PTS experiments. The mechanism is not clear. Monera and et al suggest that GdnHCl and urea denaturations can give vastly different results of protein stability, depending on how important electrostatic interactions are to the protein since the ionic nature of GdnHCl masks electrostatic interactions in the proteins [240]. Previous results in Chapter 3 also show that salts have different effect on the thermal stability of Klenpin and Klenow. One possible explanation is that Klenpin and Klenow may have different surface charge distribution or overall charge (discussion in Chapter 5) and salts probably improve the thermal stability of Klenpin by the nonspecific screening of unfavorable electrostatic interactions, whereas a general Hofmeister effect is the primary salt effect on Klenow.

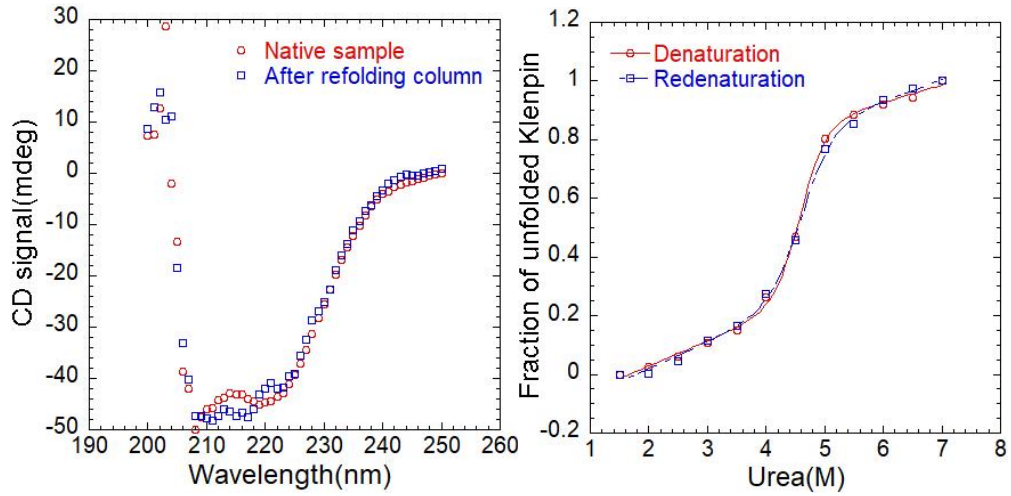

Figure 4.9. Left: CD spectra comparison of two Klenpin samples at 20°C: native state (circles) and washed through refolding column(squares). Right: Comparison of urea denaturation between Klenpin in native state (circles) and Klenpin washed through refolding column(squares)

Table 4.5. Thermodynamic Stability comparison between GdnHCl and urea denaturation.

Klenpin	ΔH_m (kcal/mol)	Tm(°C)	ΔCp (kcal/molK)	$\Delta G_{fold\ at\ 20°C}$(kcal/mol))
GdnHCl	-98.8±15	42.6±2.4	3.77±0.9	-4.2±0.6
Urea	-176.3±31	61.3±10.6	4.10±1.2	-12.2±2.6

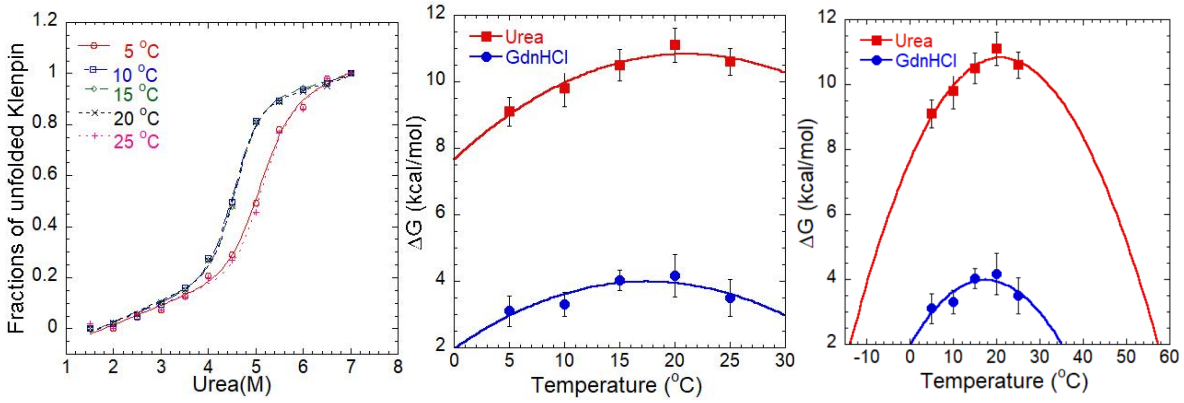

Figure 4.10. Chemical denaturation of Klenpin in GdnHCl by equilibrium measurement at different temperature: 5°C (empty circles), 10°C (empty squares), 15°C (empty diamonds), 20°C (the letter x), 25°C(crosses). Lines are fit using the linear extrapolation model, as described in methods 2.2.3.

4.3.2 Chemical denaturation of Klenpin by urea or GdnHCl with salt.

In previous sections, it has been shown that Klenpin follows a two-state model in the absence of NaCl, whereas in this section a possible intermediate is shown to be observed in the presence of 500mM NaCl for both urea and GdnHCl denaturation (Figure 4.11). Klenow, however, follows a two-state model in both GdnHCl and urea denaturation in the presence of added salt. Salts are extensively applied in experiments to investigate protein folding process since ions can influence the probability of observing a protein conformation through their interaction with protein surface [326]. Ohgushi et al found that several acid-denatured proteins, such as β-lactamase and apomyoglobin, form a compact intermediate of protein in the presence of salts [327]. The addition of salts shields the unfavorable charge repulsion between positive charges groups, which are protonated at low pH [328]. As a result, the folding process happen while native structure is not formed because of the distinct ionization of the titratable groups [328]. The shift of Klenpin folding from two-state to three-state may also be due to the salt shielding on its surface since the results from thermal stability comparison in Chapter 3 also suggest this hypothesis.

The quantitively analyze the salt effect on the conformational stability of Klenpin and Klenow. Urea denaturation of Klenpin and Klenow was tested in different NaCl concentrations (Figure 4.12) and the summary of thermodynamic parameters (ΔG_U and $D_{1/2}$) are shown in Table 4.6, Table 4.7. GdnHCl denaturation of Klenpin was only tested at 500mM NaCl concentration(Figure 4.13 and Table 4.8).

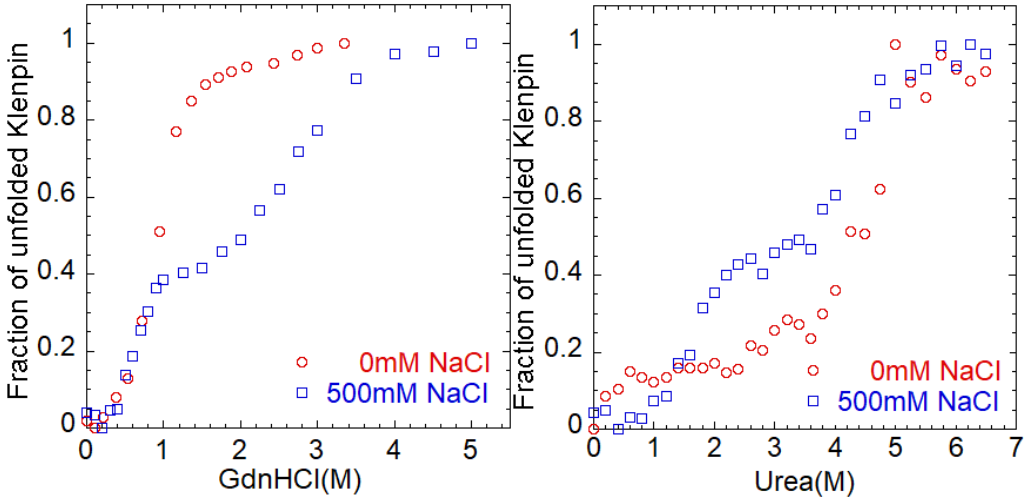

Figure 4.11. Chemical denaturation of Klenpin by GdnHCl (left) and urea(right) in 10mM phosphate buffer at 20 °C in the presence of NaCl: 0mM (circles), 500mM (squares).

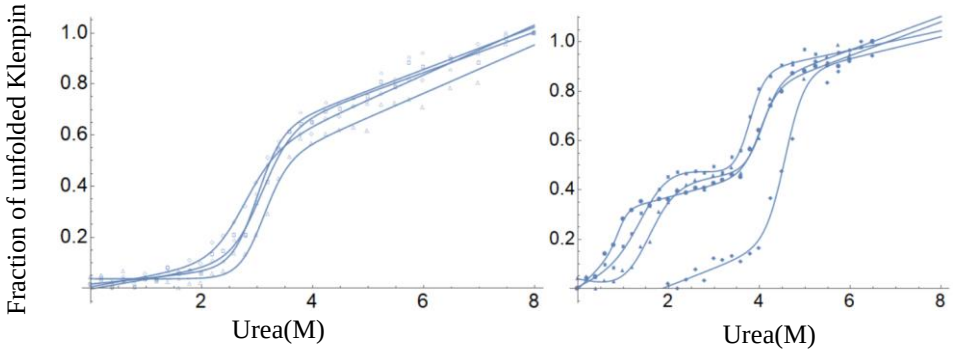

Figure 4.12. Urea denaturation of Klenpin (filled) and Klenow (empty) in the presence of NaCl: 0mM (diamond), 100mM (square),250mM (circle), 500mM (triangle). Note:the data fit Klenpin in the absence of salt is partial (2M urea to 6.5M urea) because Klenpin looks partially unfolded as shown in Figure 3.21.

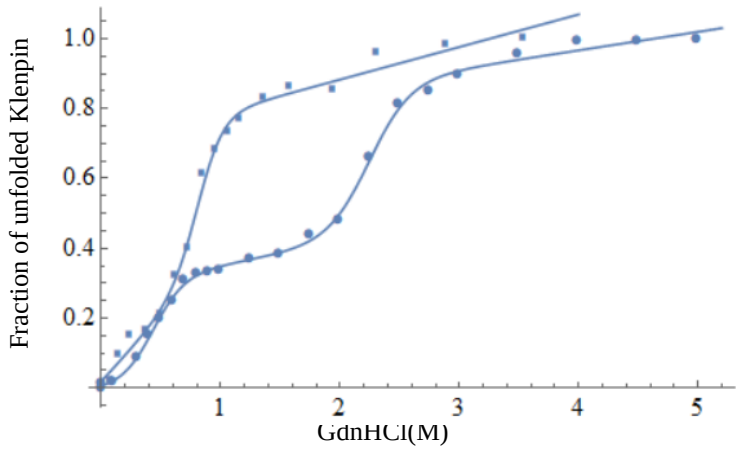

Figure 4.13. GdnHCl denaturation of Klenpin in the presence of NaCl: 0mM (squares), 500mM(circles).

Table 4.6. Summary of thermodynamic parameters of Klenow in urea denaturation at 20°C.

NaCl(mM)	ΔG(kcal/mol)	$D_{0.5}$(N<=>U) (M)
0	5.2±0.7	2.7±0.2
100	6.4±0.9	3.1±0.2
250	7.6±1.4	3.0±0.2
500	7.9±1.6	3.1±0.1

Table 4.7. Summary of ΔG(kcal/mol) values and $D_{0.5}$(M) of Klenpin in urea denaturation in the presence of NaCl (mM) at different temperatures (°C)

Temperature	NaCl	ΔG1	ΔG2	ΔΔG	$D_{0.5}$(N<=>I)	$D_{0.5}$(I<=>U)
20	0	na	12.2±2.3		na	4.6±0.1
20	100	2.7±1.2	12.7±2.7	3.2±0.7	0.9±0.1	4.0±0.1
20	250	4.6±2.8	11.7±2.3	4.1±1.0	1.5±0.2	3.8±0.2
20	500	3.0±1.3	12.2±3.2	3±0.7	1.5±0.2	4.1±0.2
10	0	na	8.9±1.38		na	4.5±0.1
15	0	na	9.9±1.47		na	4.6±0.1
25	0	na	10.7±3.1		na	5.1±0.1
30	0	na	5.0±1.3		na	4.9±0.2

Table 4.8. Summary of ΔG(kcal/mol) values and $D_{0.5}$(M) of Klenpin in GdnHCl denaturation in the presence of NaCl (mM).

NaCl(mM)	ΔG1	ΔG2	$D_{0.5}$(N<=>I) (M)	$D_{0.5}$(I<=>U) (M)
0	na	4.2±0.6	na	0.9±0.1
500	4.1±0.5	4.7±0.7	0.5±0.1	2.4±0.2

For Klenow, ΔG_U urea denaturation in different salt concentrations is 5.2~7.9 kcal/mol. In low ion strength condition (0mM~ 250mM), salt improves the thermodynamic stability of Klenow by 2.4 kcal/mol, whereas there is no significant stabilization effect from 250mM to 500mM (+0.3kcal/mol). The midpoint of the transition ($D_{1/2}$) is around 3.0 M urea in all salt conditions.

For Klenpin, salt improves the thermodynamic stability of Klenpin by 4.10 kcal/mol ($\Delta G_U(250mM\ NaCl)$- $\Delta G_U(0mM\ NaCl)$) in low ion strength condition (0mM~ 250mM), whereas there is not significant increase in ΔG_U above 250mM. The addition of ΔG_U in the presence of salt is mainly due the new transition (ΔG_1) since all ΔG_2 values are very similar (Table 4.7). Therefore, a more compact structure of Klenpin maybe formed after the addition of salt (salt shielding), which not only increase the thermal stability of Klenpin (Chapter 3), but also improve its conformational stability. $D_{1/2}$ from native state to intermediate state is increased by 0.6M as salt concentration increases from 100mM to 250mM. $D_{1/2}$ from the intermediate state to the unfolded state are all around 4.0M. For GdnHCl denaturation, a significant increase of ΔG_U of Klenpin is observed by the addition of 500mM NaCl (Table 4.8). Also, the increase of ΔG_U mainly comes from the new transition ($\Delta G_1 \sim \Delta\Delta G$) (Table 4.8).

4.3.3 Summary

Salt effect on the thermodynamic stability of Klenpin and Klenow are further compared by urea and guanidine denaturation. In the absence of salt, GdnHCl and urea denaturation of Klenpin has different estimates of ΔG_U. The mechanism is not clear. One possible explanation is that the ionic nature of GdnHCl masks electrostatic interactions in the proteins, which give vastly different results of protein stability [240]. In the presence of salt, the unfolding of Klenpin behave as a three-state model in both GdnHCl and urea denaturation. An increase in ΔG_U is observed mainly from the new transition (ΔG_1).

4.4 Structural Effect of Salt on Klenpin

4.4.1 Salt Effect on the Secondary Structure of Klenpin

Previous studies have shown that salt is a prerequisite for the function and structure of halophilic proteins [206-209]. Most of enzymes from extreme halophiles were found to be stable only in the presence of at least 1M salt [206]. At low salt concentrations, the cell envelopes of the halobacteria are slowly disintegrated to fragments, and some membrane-bound enzymes are inactivated [208]. It has been long assumed that archaea of the family Halobacteriaceae are the only active aerobic heterotrophs at or near NaCl saturation (extreme halophilies) [209]. One interesting exception is *Salinibacter ruber*, which belongs to extremely halophilic bacteria and is no less halophilic than the archaeal halophiles: no growth was observed below 15% NaCl, and for optimal growth concentrations between 20% and 30% NaCl are required [65]. A reasonable hypothesis for the unique salt dependence was that high salt concentration (high ionic strength) can protect the protein structures through the shielding of negative charges on their surfaces [234]. This hypothesis is similar to that from our previous data (salt dependence on Tm and ΔG in section 3.2 and 4.3). However, *Psychromonas ingrahamii* has a much lower NaCl requirement (1%-10%). As a result, salt effect on the secondary structure of Klenpin will be investigated in this section.

Klenpin was cloned from *Psychromonas ingrahamii*, which was isolated from sea ice in Alaska, so it is of interest to investigate the effect of salt on the secondary structure of Klenpin. The CD spectra of Klenpin were therefore analyzed in different NaCl concentrations. The result was then compared with that of Klenow and Klentaq, two homologous proteins from *E. coli* and *Thermus aquaticus* (Figure 4.14). In general, all CD spectra have two negative bands at 208 nm and 220 nm, representing α-helix and β-sheet content. NaCl significantly promoted an increase in negative ellipticity of Klenpin, especially in the low concentration range from 0 to 100mM,

whereas the CD spectra of Klenow and Klentaq show no significant difference. By plotting CD signal as a function of NaCl concentrations, an increase in negative ellipticity was observed in Klenpin (circle), whereas Klenow (square) and Klentaq (diamond) show much less increase. The results suggested that Klenpin is partially or differently folded at low salt concentration, and it may be that some salt is required for its complete folding. This characteristic is different from enzymes of some extreme halophiles (most of them are archaea)[207], which have completely distorted ellipticity pattern.

In section 3.1.3, a lower helical content was observed in Klenpin (33.7%) compared with Klenow (50%) and Klentaq (61.5%) in low salt condition (50mM NaCl). At high NaCl concentration (500mM NaCl), an increase in helical content (53.2%) was detected by BeStSel (Figure 4.15), suggesting that the enzyme adopted it proper conformation by the formation of the helix structure. The loss of helical structure after the removal of salt is seen in many halophilic proteins [191,195,198]. Several computational methods are also applied to investigate the helical difference among Klenpin, Klenow, and Klentaq in Chapter 5.

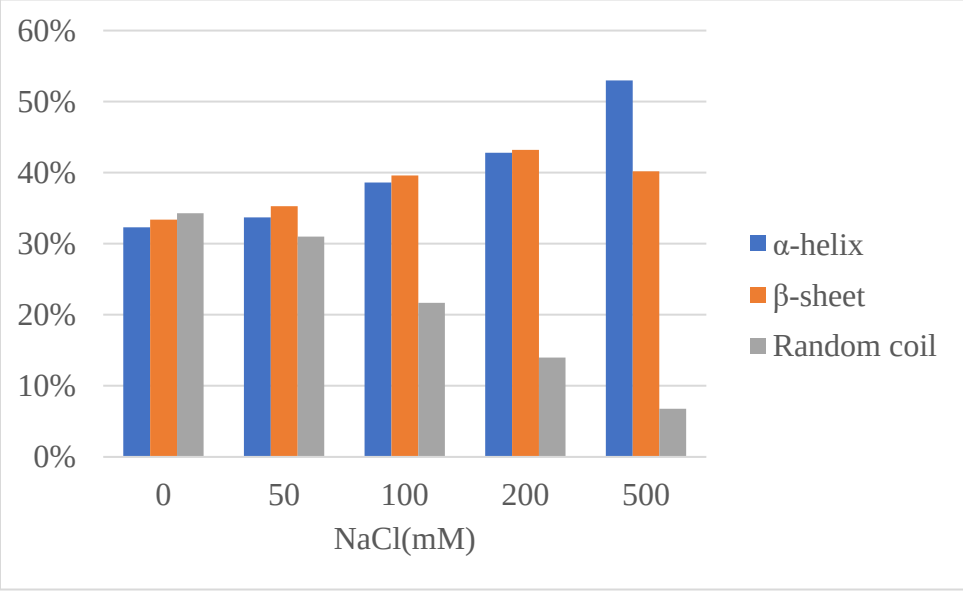

Figure 4.14. Prediction of the secondary structure composition of Klenpin by BeStSel.

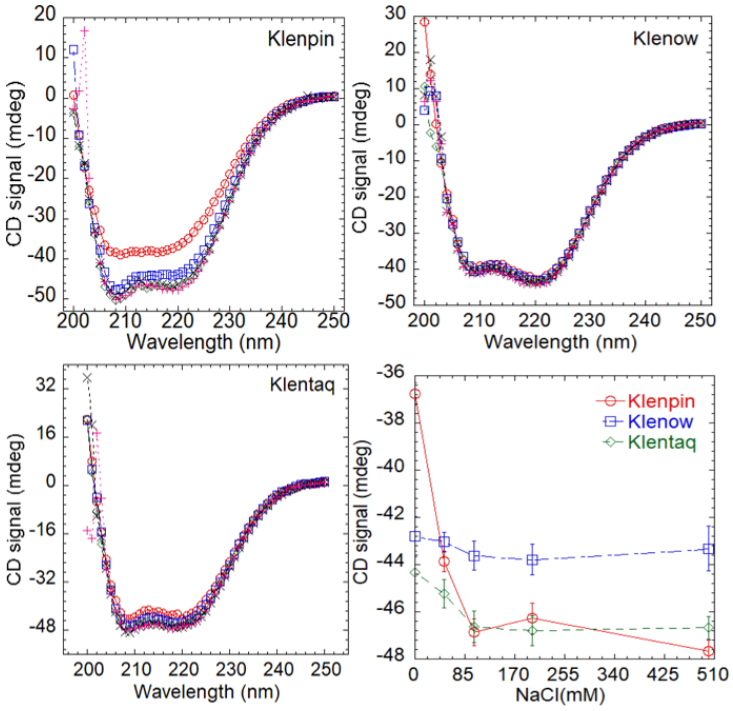

Figure 4.15. Effect of NaCl on Spectra of Klenpin, Klenow and Klentaq at 20°C: 0mM(circles), 50mM(squares), 100mM(diamonds), 200mM(crosses) and 500mM(triangles). Ellipticity comparison among Klenpin(circles), Klenow(squares), and Klentaq(diamonds) at 220nm in different concentrations of NaCl:0mM, 50mM, 100mM, 200mM, and 500mM.

4.4.2 Salt dependence of protein size by dynamic light scattering (DLS)

The salt dependence of the structure of Klenpin was further investigated by dynamic light scattering. In DLS experiments, the intensity fluctuations from the scattered light are measured over time (1 ns~1 ms). The hydrodynamic size of particles is then obtained based on the intensity distribution. The overall average hydrodynamic size of particles (Z-average) is used for the size comparison . Hydrodynamic sizes of Klenpin, Klenow, and Klentaq were investigated in different NaCl conditions (0mM-1000mM) (Figure 4.16, Table 4.9). The size of Klenpin significantly decreases as NaCl concentration increases, especially at low concentrations (<250mM). As comparison, there is no significant changes in the size of Klenow and Klentaq as adding salt. Interestingly, the hydrodynamic size of Klenpin is always larger than Klenow and Klentaq, which have a similar size. One possible explanation is the formation of dimer for Klenpin at high salt.

One good example is bovine β-lactoglobulin, which is a monomeric native structure in the absence of salt and the addition of NaCl stabilized the dimer [134].

Results from DLS also implies that Klenpin is not fully folded in the absence of salt, whereas Klenow and Klentaq do not require salt for protein folding. It is also interesting to note that even in the presence of high salt concentration, Klenpin does not compact to the same hydrodynamic size as Klenow and Klentaq. One possible explanation is that Klenpin, like many psychrophilic proteins, tend to expose more non-polar residues to the solvent in order to disfavor a compact conformation [86]. In other words, a psychrophilic protein is a flexible molecule with many opening and closing motions, whereas thermophilic protein behaves as a compact molecules undergoing few internal motions [245].

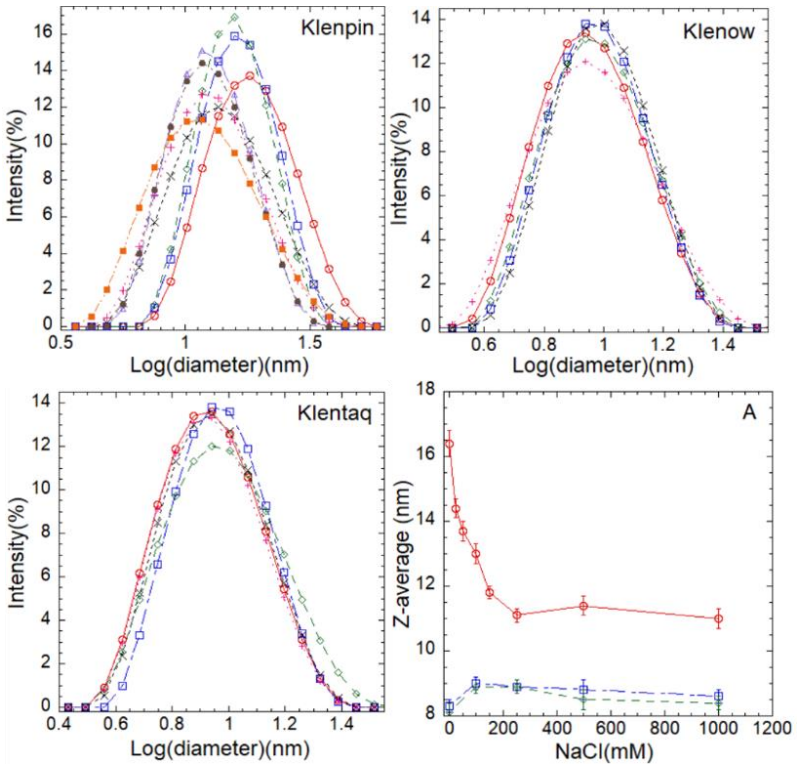

Figure 4.16. Size distribution of Klenpin, Klenow and Klentaq in different NaCl concentrations by dynamic light scattering at 20°C:0mM(circles), 25mM(squares), 50mM(diamonds), 100mM(crosses), 150mM(pluses), 250mM(triangles), 500mM(filled circles), and 1000mM(filled squares). A: Salt effect on Z-average of Klenpin(circles), Klenow(squares) and Klentaq(diamonds).

Table 4.9. Summary of Z-average of Klenpin, Klenow and Klentaq at different NaCl concentrations.

	Z-average(nm)					
NaCl(mM)	Klenpin	PDI	Klentaq	PDI	Klenow	PDI
0	16.4±0.4	0.21±0.02	8.1±0.3	0.15±0.01	8.3±0.2	0.19±0.02
25	14.4±0.3	0.17±0.01				
50	13.7±0.3	0.17±0.01				
100	13.0±0.3	0.19±0.01	9±0.2	0.21±0.02	8.9±0.2	0.20±0.02
150	11.8±0.2	0.22±0.02				
250	11.1±0.2	0.17±0.01	8.9±0.2	0.18±0.01	8.9±0.2	0.21±0.02
500	11.4±0.3	0.18±0.02	8.5±0.3	0.19±0.01	8.8±0.1	0.19±0.03
1000	11.0±0.3	0.20±0.02	8.4±0.2	0.19±0.02	8.6±0.2	0.19±0.02

4.4.3 Tryptophan (Trp) fluorescence quenching

Results from CD spectra and DLS both imply that the structure of Klenpin differs in the presence and absence of salt. One possible explanation for this difference is that Klenpin may be partially unfolded in the absence of salt. The Trp fluorescence quenching method was also applied to further investigate this hypothesis. Trp residues tend to be buried in the interior of the protein. In the unfolded or partially unfolded state, however, Trp residues are exposed to the solvent environment and easily quenched after binding to acrylamide or iodide. The quenching data for protein--acrylamide system were initially analyzed using the Stern–Volmer equation,

$$I_0/I = 1 + K_{sv}[Q],$$

Where I_0 is the fluorescence intensity in the absence of the quencher; I is the fluorescence intensity in the presence of the quencher [Q]; K_{sv} is the Stern–Volmer quenching constant.

In the Stern–Volmer equation, K_{sv} is a parameter that quantifies the quencher accessibility to the fluorophore. Therefore, we compared the K_{sv} for Klenpin and Klenow in different NaCl concentrations (Figure 4.17, Table 4.10). The Stern-Volmer plot of I_0/I vs [Q] is found to be

nonlinear (Figure 4.18) and this is characteristic of multi-fluorophore proteins with unequal accessibility of fluorophores to quenching [309]. A modified Stern–Volmer equation was applied to such systems: $I_0/(I_0 - I) = 1 + 1/(K_{sv}[Q])$.

Table 4.10. Summary of K_{sv} values for Klenpin and Klenow at different NaCl concentrations.

NaCl(mM)	K_{sv}(Klenpin)	ΔK_{sv}	K_{sv}(Klenow)	ΔK_{sv}
0	9.32±0.1	0	10.39±0.05	0
62.5	8.52±0.02	0.8	10.00±0.02	0.39
125	7.98±0.02	1.34	10.07±0.07	0.32
250	7.86±0.06	1.46	10.11±0.02	0.28
500	7.75±0.03	1.57	10.00±0.02	0.39

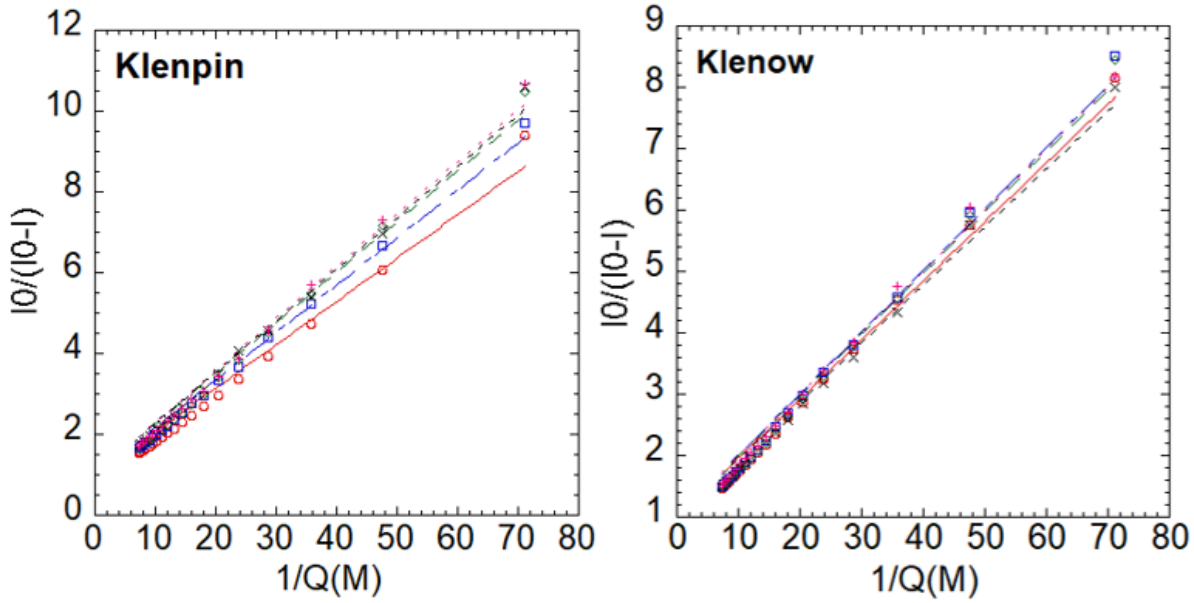

Figure 4.17. Tryptophan fluorescence quenching comparison between Klenpin (left) and Klenow (right) by acrylamide at different NaCl concentrations at 25°C:0mM(circle), 62.5mM(square), 125mM(diamond), 250mM(triangle), 500mM(cross).

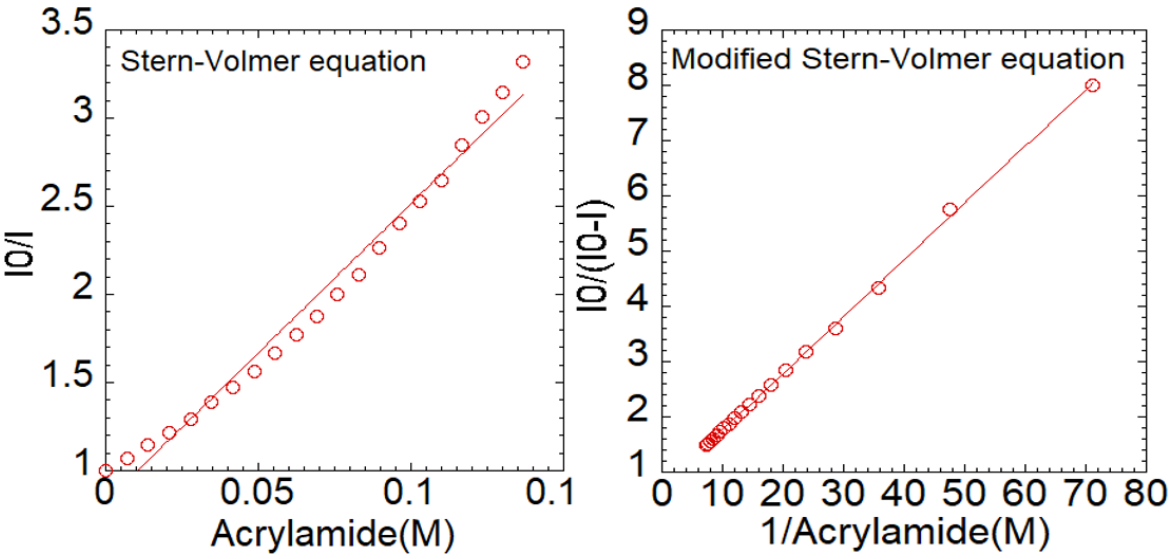

Figure 4.18. Tryptophan fluorescence quenching of Klenow in 10mM phosphate buffer with 250mM NaCl at pH 7.0 at 25°C by acrylamide. The data were analyzed by Stern-Volmer equation (left) and modified Stern-Volmer equation (right). $\lambda_{excitation} = 280nm, \lambda_{emission} = 340nm$.

Klenpin has three Trp residues, whereas Klenow has five Trp residues. Solvent accessibility of Trp residues in Klenpin and Klenow was compared by STRIDE [323], an algorithm for the assignment of protein secondary structure [311] (Table 4.11). The structure information of Klenow is from the protein data bank (PDB ID: 2kfn), whereas swiss model method is used to predict the structure of Klenpin. Swiss-Model is an automated system for modelling the three-dimensional protein structure of a protein based on its amino acid sequence [312]. As noted in Table 4.10, not only does Klenow have more Trp residues than Klenpin (5 vs 3), but also Trp residues in Klenow have more solvent accessible area than that in Klenpin (155.7 $Å^2$ vs 49.8 $Å^2$). Therefore, Klenpin has higher K_{sv} values than Klenow at all NaCl concentrations (Table 4.11).

Each condition was analyzed three times and the summary of K_{sv} values is shown in Table 4.9. A decrease in K_{sv} represents the decrease in solvent accessibility of Trp. For Klenpin, K_{sv} values are significantly decreasing (9.32→7.75) by adding NaCl, whereas the changes in K_{sv} are

much smaller for Klenow (10.39→10.00) (Table 4.11). It suggests that NaCl significantly lowers Trp accessibility of Klenpin, whereas it has much less effect on Klenow.

Also, for Klenpin, the changes of K_{sv} (ΔK_{sv}) at low NaCl concentrations (0mM-125mM) (0→1.34) are much more significant than that at high NaCl concentrations (125mM-500mM) (1.43→1.57) (Table 4.11, Figure 4.19). It suggests that Klenpin has a more compact and possibly less dynamic structure at high NaCl concentrations compared with that at low NaCl concentrations, which is also consistent with previous results.

Table 4.11. Solvent accessibility of Tryptophan in Klenpin and Klenow by STRIDE.

Protein	Position	Solvent accessible area (Å^2)
Klenow (PDB ID:1kfd)	19	18.6
	142	60.1
	190	45.3
	498	28.3
	556	3.4
Klenpin (swiss model)	18	18.6
	550	28.2
	611	3.0

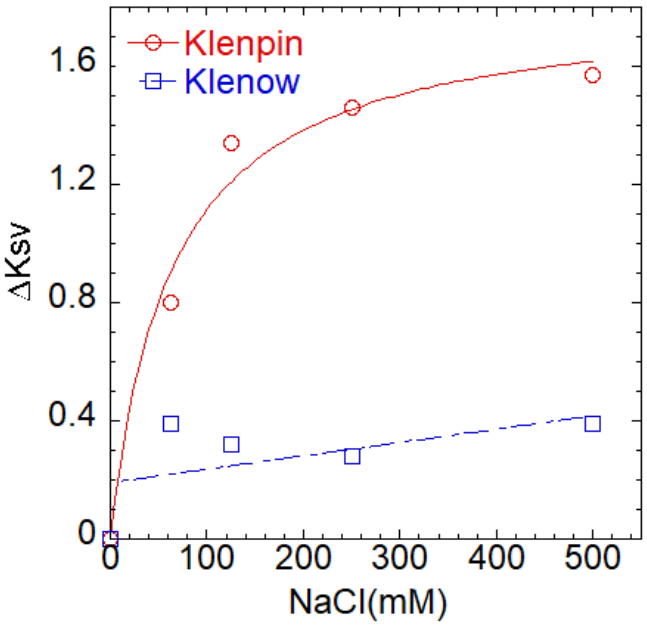

Figure 4.19. Changes in Stern–Volmer accessibility (ΔK_{sv}) for Klenpin (circles) and Klenow (squares) at different concentrations of NaCl.

4.4.4 Summary

Results in both Chapter 3 and 4 of this dissertation demonstrate that salt strongly increase thermal stability of Klenpin. By comparing the CD spectra of Klenpin, Klenow, and Klentaq at different salt concentrations, a significant increase in the negative ellipticity of Klenpin is observed as the NaCl concentration increases, whereas only small changes is seen in Klenow and Klentaq, implying that the secondary structure of Klenpin was partially disrupted in the absence of salt. Salt effect on the structure of Klenpin is further investigated by DLS and Trp quenching assays. Based on the results from DLS, Klenpin has a larger hydrodynamic diameter than Klenow and Klentaq in the absence of salt. By adding NaCl, the size of Klenpin significantly decrease, whereas Klenow and Klentaq were not affected. The salt dependence was also observed in Trp quenching experiments and solvent accessibility for Klenpin was decreased in high NaCl concentration. All these features strongly imply that Klenpin is partially or differently folded in the absence of salt, and a low concentration of salt is required for a complete Klenpin folding. This characteristic is different from enzymes of some extreme halophiles (most of them are archaea)[207], which have completely distorted ellipticity pattern. Therefore, several computational methods are applied in Chapter 5 to compare Klenpin with its two homologous proteins (Klenow and Klentaq) and enzymes from extreme halophiles.

CHAPTER 5. COMPUTATIONAL BASED INVESTIGATION OF THE STABILITY OF KLENPIN POLYMERASE

In the previous chapters, it was shown that Klenpin has lower stability (ΔG_{H2O}) than Klenow and Klentaq in the absence of salt except as measured by urea denaturation. With the addition of salt, a much greater stabilization effect (ΔTm) is observed on Klenpin than that on Klenow. Several computational approaches are applied to explore a possible explanation through three perspectives: electrostatic surface, amino acid composition, and intrinsic disorder regions (IDRs).

One possible hypothesis is that the strong salt stabilization on Klenpin is mainly from the nonspecific screening of unfavourable electrostatic interactions on the protein surface. So electrostatic surface calculations for Klenpin, Klenow and Klentaq are compared by PyMOL to see if there is any difference in the distribution of charges. A crystal structure for Klenpin is not available for now and so the three-dimensional structure of Klenpin is predicted from Swiss-Model based on its sequence.

Also, the amino acid preference may influence the stability of Klenpin. *P. ingrahamii* was isolated from sea ice and is probably a facultative psychrohalophile. Previous genomic and structural analysis shows that halophilic enzymes usually contain more acidic residues (asp and glu) and less basic residues, especially lys, when compared with their homologous mesophilic proteins [34]. Therefore, A comprehensive comparison of global amino acid preferences among the three homologous proteins was investigated. We also compare 100 pairs of homologous proteins from *P. ingrahamii* and *E.coli* by Student's t-test (see Methods 2.2.4).

Some studies reported that intrinsic disorder regions (IDRs) are often resistant to boiling temperatures or cold environments [189,190,192,193], and disorder regions (IDRs) may play a role in *P. ingrahamii*'s adaptation. The previous results show that the structure of Klenpin is more

flexible and larger in the absence of salt than that for Klenow. Our hypothesis is that Klenpin may have more disordered structures than Klenow in the absence of salt. The IDRs of the three proteins are predicted from their sequences and then compared to test this hypothesis.

5.1 Comparison of the Electrostatic Potential of Klenpin, Klenow, and Klentaq

Results in Chapter 3 and 4 indicate that salts from the Hofmeister series have a much stronger stabilization effect on Klenpin than Klenow. One possible explanation is that there is much stronger electrostatic repulsion on the surface of Klenpin than that on Klenow. Therefore, the charge distribution (electrostatic potential) of Klenpin and Klenow are calculated and compared to explore the potential pattern difference.

The three-dimensional structure of Klenpin is predicted by Swiss-Model (See Methods 2.2.5) since no experimental structure is available (Figure 5.1). The electrostatic surface potential of Klenpin, Klenow, and Klentaq is calculated and compared by The APBS (Adaptive Poisson-Boltzmann Solver) tool in PyMOL (See Methods 2.2.6, Figure 5.2). All input parameters (such as temperature, ionic charge, ionic concentration, and radius of ion species) are the same. Surface charges are considered as continuous "patches" (blue patches for positive charges and red patches for negative charges). Figure 5.2 shows that there are more blue and red patches on the surface of Klenpin than that on the surface of Klenow and Klentaq. The results are consistent with the hypothesis from salt experiments that the electrostatic repulsion on the surface of Klenpin is much stronger than that on Klenow. In addition to the difference in their charge distribution, it's also interesting to note that there is no gap between fingers domain and thumb domain in the electrostatic surface potential of Klenpin.

However, any conclusions from modeling structures of halophilic proteins should be used with caution. For example, a structure-based model of dihydrofolate reductase from *Halobacterium*

volcanii suggests that its low-salt instability is due to charge repulsion of the overrepresented acidic residues[220], whereas such repulsion was not found in its solved crystal structure [171]. Therefore, a crystal structure of Klenpin is needed to eventually clarify our hypothesis in the future.

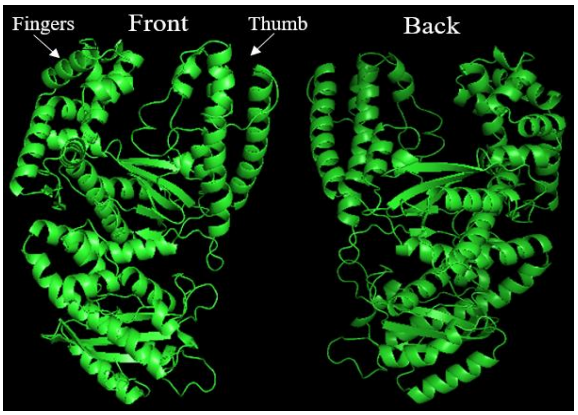

Figure 5.1. Three dimensional structures of Klenpin(Swiss-Model).

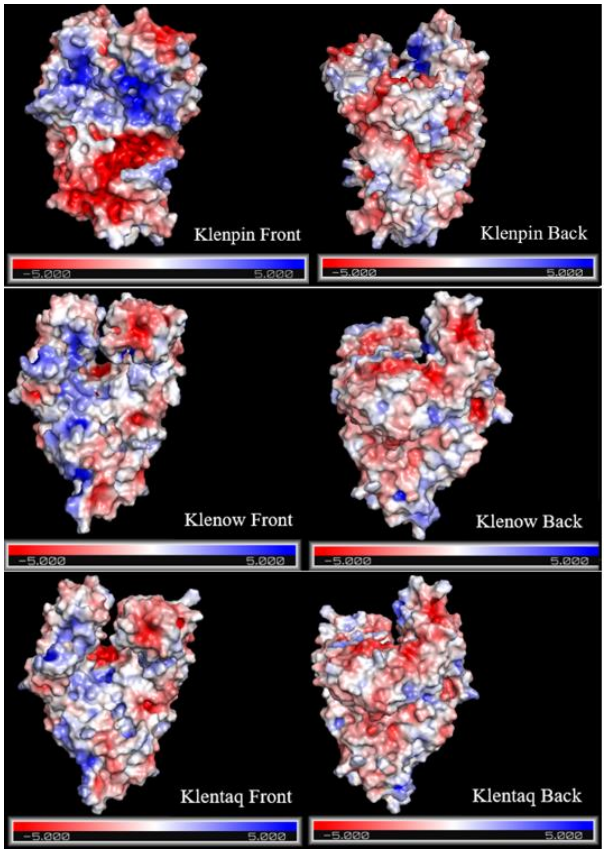

Figure 5.2. Comparison of the electrostatic surface potential of Klenpin(Swiss-Model), Klenow(1kfd), and Klentaq(1ktq) by PyMOL. Patches of positive (+) and negative (−) potential are shown, respectively, in blue and red. Neutral patches are white.

5.2 Amino Acid Composition Comparison among Klenpin, Klenow, and Klentaq

The individual amino acid frequency of the three homologous proteins from *P. ingrahamii* (psychrophile), *E.coli* (mesophile), and *Thermus aquaticus* (thermophile) are first compared (Figure 5.3) and several residues showed a significant difference among the compositions of Klenpin, Klenow, and Klentaq. The overall Amino acid compositions of Klenpin, Klenow and Klentaq are shown in Table 5.1.

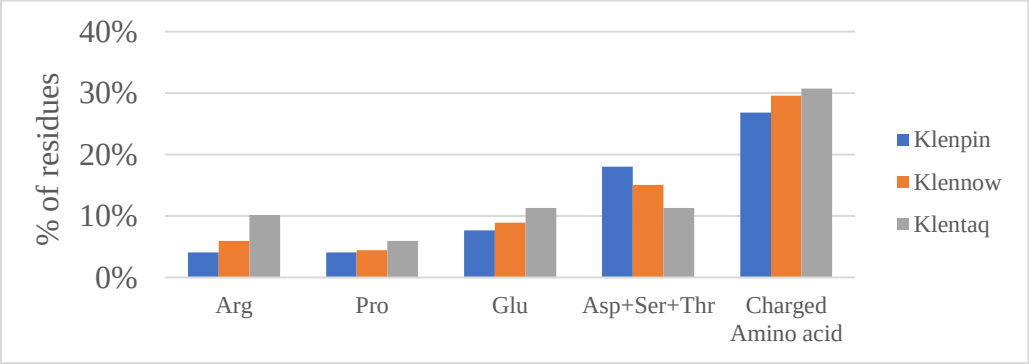

Figure 5.3. Amino acid composition comparison showing a systematic trend from Klenpin to Klenow to Klentaq.

Table 5.1. Amino acid composition of Klenpin, Klenow and Klentaq

Amino acid	Klenpin	Klenow	Klentaq
Ala	9.4%	10.2%	11.7%
Arg	4.1%	6.0%	10.2%
Asn	2.9%	3.8%	1.9%
Asp	5.9%	5.8%	3.7%
Cys	0.2%	0.2%	0.0%
Gln	5.4%	4.1%	2.2%
Glu	7.6%	8.9%	11.3%
Gly	5.5%	5.5%	6.7%
His	2.0%	2.6%	1.9%
Ile	7.2%	6.4%	3.0%
Leu	12.5%	11.4%	13.7%
Lys	7.3%	6.3%	3.7%
Met	2.4%	2.3%	2.4%
Phe	2.6%	2.6%	3.3%
Pro	4.1%	4.5%	5.9%
Ser	6.7%	4.5%	3.7%
Thr	5.5%	4.8%	3.9%
Trp	0.5%	0.8%	1.7%
Tyr	3.9%	3.5%	2.6%
Val	4.4%	5.8%	6.7%

The percentage of arginine in Klenpin is less than half of that in Klentaq. Arginine residues can form multiple hydrogen bonds to enhance the stability of a protein and is often reduced in psychrophilic proteins [153]. Also, Klenpin has the lowest number of proline residues compared with Klenow and Klentaq. During protein folding, proline isomerization is considered to be a rate-limiting step [154]. Therefore, many cold-adapted proteins tend to possess low proline content [127]. Moreover, Klenpin contains the lowest amount of glutamic acid, which can form salt-bridges to enhance protein stability. Therefore, a lower content of glutamic acid is often observed in cold-adapted proteins [135]. In addition, Klenpin has a much higher lysine-to-arginine ratio (1.8) than Klenow (1.05) and Klentaq (0.32). Siddiqui et al have shown that replacing lysine with arginine in the cold-adapted α-amylase generates mesophilic-like characteristics [170]. One possible mechanism is that the arginine facilitates more ionic (two salt bridges and five H bond) interactions than lysine through the guanidino group [316]. Also, Klenpin has many residues that inhibit the formation of helix structure, such as aspartic acid, serine, and threonine. One possible explanation is that high helical content is avoided in psychrophilic proteins to maintain high flexibility [135]. Surprisingly, the number of charged residues in Klenpin is slightly less than that in Klenow or Klentaq, while electrostatic surface potential comparison suggests higher exposure of charged residues in Klenpin, greater clustering of similar charges, and possible higher juxtapositions of charged residues and hydrophobic residues (since lower local dielectric constants enhance charge effects in proteins). One example is Dodecin from *Halobacterium salinarum*, which contain much higher surface charges than that from *Mycobacterium tuberculosis* (144 vs 30), whereas no difference is observed in the overall number of charged residues (20 vs 21) [143]. The high number of acidic residues on halophilic protein surface is thought to promote the tight binding of the hydration shell to the proteins [19]. In general, Klenpin is a typical cold-adapted

protein that has a significantly higher proportion of residues that contribute to higher protein flexibility and a lower proportion of residues that improve stability.

Due to its high salinity tolerance (2.6M NaCl), *P. ingrahamii* is also a halophilic or halotolerant microorganism. Proteins from extremely halophilic bacteria (most of them are archae) often have a higher ratio of acidic to basic residues and lower hydrophobic content compared with their homologous mesophilic proteins [34,35]. However, the ratio of acidic to basic residues is close to 1 for Klenpin, Klenow, and Klentaq (Table 5.2). Also, Klenpin has a similar number of hydrophobic residues (265) compared with that for Klenow and Klentaq (Table 5.2). In general, there is no difference between Klenow and Klenpin in either ratio of acidic to basic residues or the hydrophobic content, so while the amino acid content of Klenpin follows patterns associated with other psychrophilic proteins, it does not follow the patterns associated with other halophilic proteins, suggesting that *P. ingrahamii* may use a different saline adaptation mechanism since sea ice is a different and potentially more variable environment than other hypersaline environments.

Table 5.2. Comparison of charged and hydrophobic residues in Klenpin, Klenow, and Klentaq

	Hydrophobic residues	Acidic residues /basic residues	% of charged residues
Klenpin	265	1.01	26.83%
Klenow	267	0.99	29.59%
Klentaq	261	0.95	30.74%

5.3 Acidic Signature in DNA Polymerase I of *P. ingrahamii*

In this section, we further compare the proteomes of salt-in halophilic bacteria (*Halorhodospira halophila*, *Salinibacter ruber*), salt-out halophilic bacteria (*Bacillus subtilis*, *Halobacillus halophilus*), sea ice bacteria (*Psychromonas ingrahamii*, *Colwellia psychrerythraea*), and non-halophilic bacteria (*E. coli, Thermus aquaticus*). Also, the electrostatic surface potential maps of DNA pol I from those organisms are investigated to explore the potential differences. In Table 5.3, the acidic/basic residues is defined as the ratio of acidic residues (aspartic acid, glutamic

acid) to basic residues (arginine, histidine, lysine), whereas the acidic/basic proteins is defined as the ratio of acidic proteins (pI<7) to basic proteins (pI>7). The distribution of the isoelectric point (pI) of proteins in a proteome is also used as a simple indication of the acidic nature of an organism .

In general, the proteomes from salt-in halophiles have higher acidic/basic proteins and lower average proteome isoelectric point. For the 8 Pol I polymerases examined here, seven have a low acidic/basic residue ratio, close to 1 (0.95-1.19), the exception being *Salinibacter ruber* with a ratio of 1.82. Also, the electrostatic surface potential of DNA pol I from *Salinibacter ruber* show an extremely high acidic residues distribution, which is not observed in another salt-in halophile (*Halorhodospira halophila*) (Figure 5.4). It suggests that not all proteins in salt-in halophiles follows the same adaptation method to hypersaline environment.

There is no significant difference in acidic/basic residues DNA pol I among salt-out halophiles, sea ice bacteria and non-halophiles (Table 5.3). By comparing the electrostatic surface potential maps, however, there is more charges (both positive and negative) on the protein surface of salt-out halophiles and sea ice bacteria that that on non-halophiles. It suggests that the saline adaptation of DNA pol I may occur through the modification of surface charges instead of the overall charges.

Table 5.3. Comparison of the acidic features among 'salt-in' halophiles, moderate halophiles, sea-ice bacteria, and non-halophiles

	Organism	Average pI	Acidic/basic proteins	Acidic/basic residues DNA pol *I*[1]
Extreme halophiles	*Salinibacter ruber*	5.74	4.60	1.82
	Halorhodospira halophila	6.04	3.60	1.18
Moderate halophiles	*Bacillus subtilis*	6.52	1.76	1.17
	Halobacillus halophilus	6.15	2.58	1.15
Sea ice bacteria	*Psychromonas ingrahamii*	6.52	1.70	1.09
	Colwellia psychrerythraea	6.43	1.81	1.19
Non-halophiles	*E. coli(K12)*	6.60	1.77	1.05
	Thermus aquaticus	7.34	1.04	0.90

[1] DNA pol I for *E.coli* and *thermus aquaticus* are the full length, whereas Klenow and Klentaq are large fragments of DNA pol I.

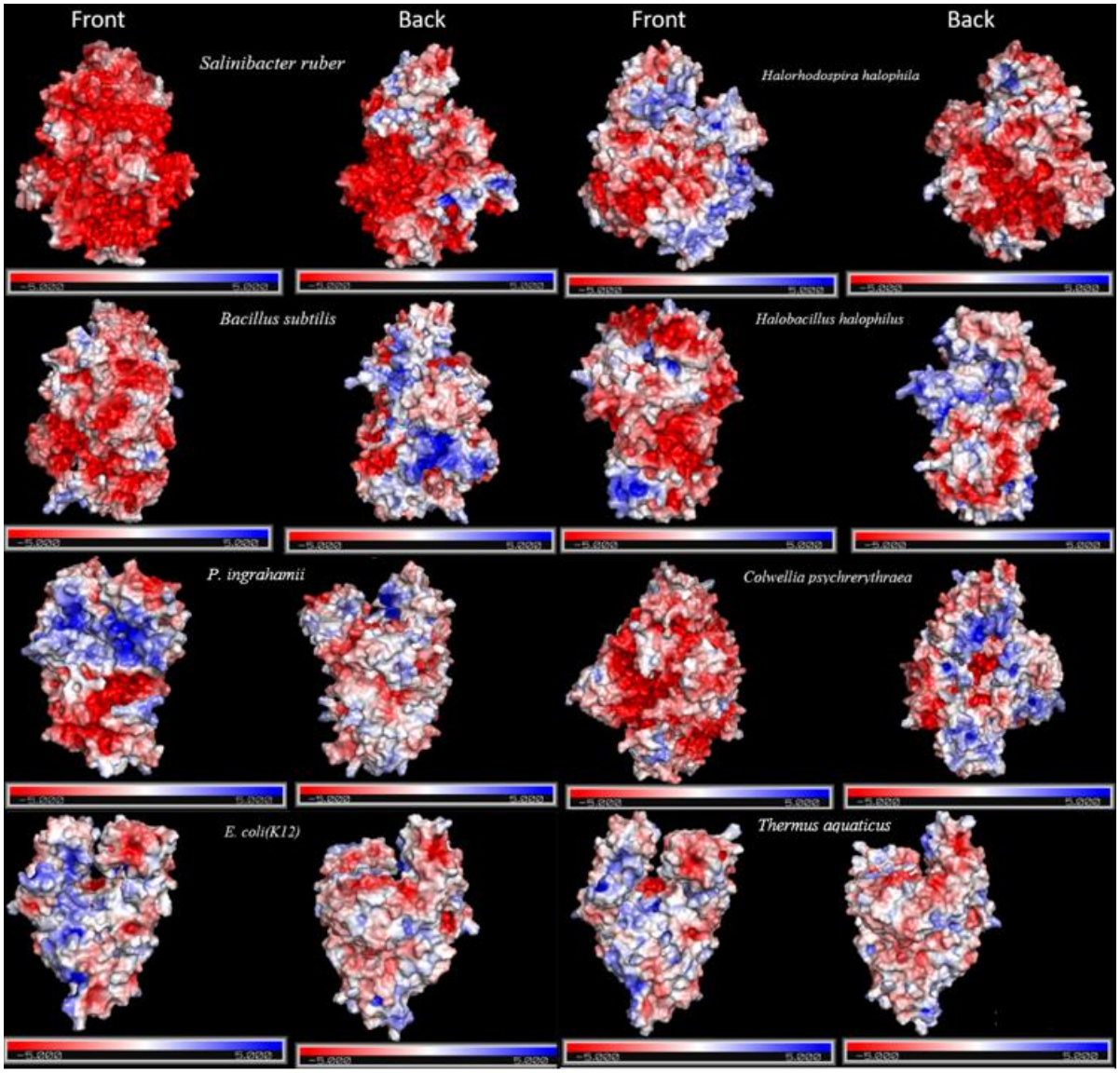

Figure 5.4. Comparison of the electrostatic surface potential of DNA polymerase I from salt-in halophiles (*Halorhodospira halophila*, *Salinibacter ruber*), salt-out halophilic bacteria (*Bacillus subtilis*, *Halobacillus halophilus*), sea ice bacteria (*Psychromonas ingrahamii*, *Colwellia psychrerythraea*), and non-halophilic bacteria (*E. coli*, *Thermus aquaticus*) by PyMOL. Patches of positive (+) and negative (−) potential are shown, respectively, in blue and red. Neutral patches are white. The three-dimensional structures of Klenow and Klentaq are from crystal structure data (1kfd and 1ktq), all other structures are models generated with Swiss-Model.

5.4 Comparison of Helix Content Through Computational Method

In previous section (3.6.1), the salt dependence of helix content among Klenpin, Klenow and Klentaq by CD spectra was investigated, and the results suggest that salt significantly improves the helical content of Klenpin (33.7%→ 53.2%) and similar to the helix content of Klenow(50%)

and Klentaq(61.5%). Therefore, the low helical content in Klenpin maybe correlated with salt rather than cold adaptation. We further compared the helical content among Klenpin, Klenow and Klentaq based on their sequences or three-dimensional structures.

For sequence-based prediction, six methods (PSIPRED[317], CFSSP[318], GOR IV[319], PredictProtein[320], SCRATCH[321], RaptorX-Property[322]), whereas two methods (STRIDE[323] and DSSP[324]) are applied to three-dimensional structures-based investigation (Table 5.4).

Table 5.4. Comparison of helical content among Klenpin, Klenow, and Klentaq by computational methods.

	Klenpin	Klenow	Klentaq
Method		Helical content	
PSIPRED	52.70%	54.40%	51.90%
CFSSP	54.30%	58.30%	54.80%
GOR IV	52.20%	52.23%	56.30%
PredictProtein	50.08%	53.55%	57.78%
SCRATCH	53.17%	52.89%	54.26%
RaptorX-Property	49.02%	49.13%	55.20%
STRIDE	50.57%	51.08%	53.02%
DSSP	52.70%	50.41%	52.83%

In general, the helical contents from computational methods are similar to those from CD spectra analysis in the presence of salt, suggesting that the three-dimensional structure of Klenpin from Swiss-Model is for the high salt structure. Also, the results from all methods predict that there is no difference in the helical contents among Klenpin, Klenow, and Klentaq, and It is notable that all of these programs predict a helical content similar to that found for Klenpin at high salt rather than at low salt. Therefore, a higher percentage in aspartic acid, serine, and threonine may not correlated with the avoidance for helical content in Klenpin, which is proposed by many studies [135,314].

5.5 Amino Acid Composition Preferences from *P. ingrahamii, E. coli,* and *T. aquaticus*

Next, we compared 100 pairs of homologous proteins randomly selected from *P.ingrahamii* and *E.coli* by subtracting the frequency of individual amino acid (the percentage of each amino acid in each protein) in *E. coli* from that in *P. ingrahamii*. The frequency difference was then analyzed by Student's t-test to see if we can find similar trends. For this project, Student's t-test (See Methods 2.2.4) is used to test the null hypothesis that there is no difference in the amino acid composition between *P.ingrahamii* and *E.coli*. Based on the T-distribution, a large t-value represent a low p-value. A small p-value (<0.05) or high t-value (>1.96 or <-1.96) indicates that the null hypothesis should be rejected and there is a significant difference in the frequency of individual amino acid between the two organisms. A positive t value means that the frequency of specific amino acid in *P. ingrahamii* is higher than that in *E. coli*, whereas a negative t value indicates an opposite trend. The results were shown in Table 5.5.

In general, some amino acids, such as Asn, Gln, Ile, Lys, Phe, and Ser, are significantly preferred in *P. ingrahamii*. On the other hand, amino acid residues, such as Ala, Arg, Glu, Gly, Pro, Trp, and Val, are less favoured in *P. ingrahamii*.

A significant decrease in Arg content (t-value= -11.98, p-value<0.001) was observed in *P. ingrahamii*. Compared with other amino acids, Arg has the highest number of hydrogen donor/acceptor atoms and therefore a decrease in Arg can significantly reduce the structural flexibility and increase the stability of psychrophilic proteins [163, 164]. A higher lysine (t-value=8.93, p-value <0.001) content in *P. ingrahamii* than that in *E.coli* is consistent with the previous comparison among Klenpin, Klenow and Klentaq.

A lower proline content (t-value= -7.30, p-value<0.001) is present in *P. ingrahamii*, which is often observed in cold-adapted proteins to improve the protein folding rate since proline

isomerization is a rate-limiting step [154]. Also, there are 12 peptidylprolyl isomerases in *P. ingrahamii*, whereas 8 peptidylprolyl isomerases are known in *E.coli* [325]. The two strategies may be both applied to maintain a necessary folding rate in *P. ingrahamii* at extreme low temperature.

Table 5.5. The Difference in individual amino acids of 100 pairs of homologous proteins. Positive values of t indicate an increased frequency in *P.ingrahamii* and negative values indicate a decreased frequency in *P.ingrahamii*.

Amino acids	SD	Minimum difference	Maximum difference	p-value	t-value
Ala	0.02	-0.04	0.03	< **0.001**	-5.54
Arg	0.01	-0.04	0.02	< **0.001**	-11.98
Asn	0.01	-0.02	0.04	< **0.001**	5.58
Asp	0.01	-0.02	0.04	0.46	0.75
Cys	0.01	-0.02	0.02	0.25	1.16
Gln	0.01	-0.02	0.04	< **0.001**	4.06
Glu	0.01	-0.04	0.02	< **0.001**	-3.08
Gly	0.01	-0.03	0.01	< **0.001**	-6.19
His	0.01	-0.02	0.02	0.46	-0.74
Ile	0.01	-0.02	0.05	< **0.001**	8.45
Leu	0.01	-0.03	0.03	0.73	0.34
Lys	0.01	-0.03	0.05	< **0.001**	8.93
Met	0.01	-0.02	0.03	0.13	-1.52
Phe	0.01	-0.02	0.04	< **0.001**	3.68
Pro	0.01	-0.04	0.01	< **0.001**	-7.30
Ser	0.01	-0.03	0.05	< **0.001**	6.30
Thr	0.01	-0.05	0.03	0.33	-0.98
Trp	0	-0.02	0.01	< **0.001**	-4.21
Tyr	0.01	-0.02	0.02	0.11	1.63
Val	0.01	-0.04	0.02	< **0.001**	-5.18

Significant compositional differences as indicated by t-test parameter are shown in color. Minimum difference ,maximum difference, and standard deviations (SD) represent the range and variability of difference for individual amino acid among the comparison of 100 pairs of homologous proteins.

Asparagine (t-value=5.58, p-value<0.001) are present in a higher percentage in *P. ingrahamii* than that in *E.coli*. An asparagine excess has also been seen other psychrophiles, such as *P. haloplanktis* [315]. One possible explanation for this bias is that the process of deamidation in

which asparagine is converted to aspartate is sensitive to high temperature [18] and *P. ingrahamii* grows exponentially at extremely low temperature (-12 °C).

Although serine, as a helix breaker, is present in a higher percentage in *P. ingrahamii* and maybe partly responsible for the reduced thermostability and increased flexibility of cold-adapted proteins. The contents of other helix breaker, such aspartic acid (t-value=0.75, p-value=0.46) and threonine (t-value=-0.98, p-value=0.33), show no difference between *P. ingrahamii* and *E.coli*. Metpally et al reported that in psychrophiles, serine, aspartic acid, threonine and alanine are overrepresented in the coil regions, whereas glutamic acid and leucine are underrepresented in the helical regions [135]. Therefore, we compared the distribution of those amino acids in coil regions and helical regions for Klenpin and Klenow by PSIPRED (Table 5.6).

Table 5.6. Distribution of S, D, T, A, E, L in coli regions and helical regions between Klenpin and Klenow by PSIPRED.

			Klenpin			
PSIPRED	Ser(S)	Asp(D)	Thr(T)	Ala(A)	Glu(E)	Leu(L)
Coil regions	25	16	14	13	8	16
Helical regions	12	15	15	36	28	48
Coil/helical regions	2.08	1.07	0.93	0.36	0.29	0.33
			Klenow			
PSIPRED	Ser(S)	Asp(D)	Thr(T)	Ala(A)	Glu(E)	Leu(L)
Coil regions	16	16	17	14	9	15
Helical regions	8	14	10	41	31	44
Coil/helical regions	2.00	1.14	1.70	0.34	0.29	0.34

In general, there is no difference in the distribution (coil/helical regions) for serine, aspartic acid, alanine, glutamic acid and leucine between Klenpin and Klenow. For threonine, more residues are in helical regions for Klenow than that in Klenpin, which is even opposite with the results from Metpally's analysis, suggesting that threonine in Klenpin may not act as a helix breaker. Instead, Lanyi et al suggest that low content of hydrophobic amino acids in halophilic

proteins is often offset by a higher content of borderline hydrophilic–hydrophobic residues such as threonine [206].

Therefore, not only Klenpin has similar helical contents compared with Klenow (Table 5.4), but also there is no difference in the distribution of helix related amino acids (Table 5.6).

5.6 Intrinsic Disordered Regions (IDRs)

Intrinsic disorder (ID) is also an essential feature of many protein sequences. Previous studies have shown that IDPs are still active after incubating at both boiling and freezing temperatures [189,192]. Although Klenpin is a well-ordered protein, intrinsic disorder regions (IDRs) were observed using several predictors. Therefore, it is of interest to test the relationship between the level of IDRs and the conformation stability of Klenpin.

We first compared the intrinsically disordered regions among Klenpin, Klenow, and Klentaq using four different predictors (PONDR, IUPred2A, IsUnstruct, and DISpro) [183-186] (Table 5.7). Although each predictor has a different prediction about the level of intrinsic disorder regions, the trend is very similar. Klenpin always has lower intrinsic disorder than Klenow and Klentaq. Whether Klenow or Klentaq has the next highest percentage of disorder differs among the different predictors.

Moreover, intrinsic disorder regions of 100 pairs of homologous proteins from *P. ingrahamii* and *E. coli* were predicted by Pondr, and the data is analyzed by t-test. As mentioned in previous section, A small p-value (<0.05) indicates a significant difference between the two organisms. A positive t value means that the frequency of specific amino acid in *P. ingrahamii* is higher than that in *E. coli*. In Table 5.8, proteins from *P. ingrahamii* contain a lower percentage of disordered residues than those from *E. coli*, as shown by a negative t-value (-7.52).

In general, *P. ingrahamii* contains a lower level of intrinsically disordered regions than *E.coli*. Therefore, the loss of structure of Klenpin in the absence of salt is probably not due to intrinsic disordered regions (IDRs).

Table 5.7. Comparison of the intrinsic disorder among Klenpin, Klenow, and Klentaq.

	Predictor	Predicted	Residues disordered	Overall percentage
	Pondr[1]			
Klenpin		615	133	21.63%
Klenow		605	171	28.26%
Klentaq		540	266	49.26%
	IUPred2A[2]			
Klenpin		615	19	3.09%
Klenow		605	46	7.60%
Klentaq		540	37	6.85%
	IsUnstruct[3]			
Klenpin		615	69	11.22%
Klenow		605	93	15.37%
Klentaq		540	66	12.22%
	DISpro[4]			
Klenpin		615	98	15.93%
Klenow		605	108	17.85%
Klentaq		540	181	33.52%

1 Predictor Of Naturally Disordered Regions: the first tool designed specifically for prediction of protein disorder based on artificial neural networks

2 a combined web interface that allows to identify disordered protein regions using IUPred2 and disordered binding regions using ANCHOR2

3 prediction of the residue status to be ordered or disordered in the protein chain by a method based on the Ising model

4 a member of the SCRATCH suite of protein data mining tools

Table 5.8. Intrinsically disordered region comparison between *P. ingrahamii* and *E.coli.* based on intrinsic disorder regions of 100 pairs of homologous proteins.

	Mean	Minimum	Maximum	P-value	T-value
P. ingrahamii	0.20	0.01	0.53	<0.001	20.91
E.coli	0.26	0.04	0.62	<0.001	24.29
Difference	-0.06	-0.29	0.22	<0.001	-7.52

5.7 Summary

Both the average proteome isoelectric points and acidic/basic proteins show no difference between *P. ingrahamii* and *E.coli*, whereas the comparison of the electrostatic surface potential indicates that there are more surface charge patches on Klenpin than that on Klenow. It suggests that the saline adaptation of DNA pol I may occur through the modification of surface charge

patterns instead of the overall charges, supporting the hypothesis that the electrostatic repulsion on the surface of Klenpin is much stronger than that on Klenow. Also, there is no excess of acidic amino acid (acidic/basic residue ~1) in Klenpin, suggesting that *P. ingrahamii* does not belong to salt-in halophiles. The amino acid composition comparisons between *P. ingrahamii* and *E.coli* indicate a preference for Asn, Gln, Ile, Lys, Phe, and Ser and an avoidance of Ala, Arg, Glu, Gly, Pro, Trp, and Val. All these features indicate that *P. ingrahamii* is a typical psychrophile. Predication of IDRs from several computational methods find a lower a lower intrinsic disorder in Klenpin than that in Klenow or Klentaq, suggesting the flexible and large structure of Klenpin in the absence of salt probably does not arise from intrinsic disordered regions (IDRs).

CHAPTER 6. DISCUSSION AND CONCLUSION

6.1 Discussion

Organisms with the ability to thrive in extreme environments (temperatures of -20 ℃ or 122°C, high salinity, high radiation, high pressure, and pH levels of 2 or 11) are known as extremophiles [293]. Most of current studies on the adaptation mechanism usually focus on one of these extremes and a more complex environment (such as sea ice) has been less studied, especially at a molecular level. Sea ice is characterized as a dynamic environment with seasonal changes in both temperature and salinity and requires organisms to have many adaptations to survive or thrive [238]. Previous studies have shown that high concentration of extracellular polymeric substances (EPS) act as cryo- and osmo-protectants against high salinity [300,301]. Some sea ice organisms also produce ice-binding or anti-freeze proteins to prevent the formation of ice crystals [297].

In this project, we mainly focus on the thermodynamic characteristics of Klenpin, a large fragment of DNA Polymerase I from a sea ice bacterium called *P. ingrahamii*. With low temperature (-2°C to -30°C) and high salinity (15% or 2.6M NaCl) [18], organisms from sea ice are probably facultative psychrohalophiles because sea ice also melts and these organisms must survive such seasonal thaw-freeze cycles. Both the temperature and salt dependence of Klenpin were compared with Klenow and Klentaq, two homologous proteins from *E.coli* (mesophile) and *Thermus aquaticus* (thermophile). DNA polymerase I is a key enzyme required to keep the genetic information during DNA synthesis, DNA repair, and DNA recombination. Little is known regarding halophilic and psychrophilic adaptation of the DNA processing machinery, and protein-DNA interaction are both salt and temperature sensitive [263,275]. The present dissertation reports the first folding and stability study of a DNA polymerase I with halophilic and

psychrophilic features, and the main objective is to study the effect of salt and temperature on the thermodynamic stability of Klenpin by comparing it with Klenow and Klentaq.

Ions from the Hofmeister series were first tested for their effect on the melting temperatures of Klenpin and Klenow , and the results show significant differences. Klenow is stabilized by salting out ions ($Na^+, K^+, SO_4^{-2}, HPO_4^{-2}/H_2PO4^-$), whereas it is destabilized by salting in ions (SCN^-, Mg^{2+} and $GdnH^+$), which is consistent with the general Hofmeister preferential hydration effect. Klenpin , however, is stabilized by all test ions, at low concentrations (<0.2M). Many factors can cause an increase in protein stability when adding salt. At first, high salt concentration can increase the exposure of buried hydrophobic residues, improving the apparent hydrophobic effect [253]. Secondly, it has been shown that the screening of unfavorable electrostatic intramolecular interactions would increase the native-state stability [254]. In addition, specific anion binding to the Lys residues located on the surface leads to a stabilizing effect [255].

Within the same salt group (salting in ions or salting out ions), the salt effect on Klenpin is independent of the identity of the salt cation and anion. Therefore, specific ion binding on the protein is not the possible mechanism for Klenpin stabilization. In addition, a linear increase in stability of Klenow (ΔTm) is present as salt concentration increases, whereas a bell-shaped curve is observed for Klenpin. It is interesting to note that Klenpin is even stabilized by low concentration of GdnHCl (<0.4M), whereas Klenow was destabilized as expected. This rules out the possibility that salt stabilization of Klenpin is due to apparent hydrophobic effect ,as weakly hydrated GdnHCl would disrupt hydrophobic interactions [257]. All these features imply that there is much higher repulsive electrostatic interaction on the surface of Klenpin than on Klenow, and these unfavorable electrostatic interactions can be screened by salt.

The importance of electrostatic interaction on proteins is often a debated topic. Electrostatic interaction can affect the stability of protein through attraction of opposite charges and repulsion of charges with equal sign. The net contribution of these interactions is highly context dependent. In some cases, the stability of proteins is sensitive to the modification of charges on the surface [291,298], whereas, in other cases, little effect has been observed on the overall stability by changing surface charges [287].

To test our hypothesis that salts mainly stabilize Klenpin through the screening of unfavorable electrostatic interactions. We first compare the electrostatic surface potential among Klenpin, Klenow, and Klentaq. The results show that there are more red patches (negative charges) and blue patches (positive charges) on Klenpin than those on Klenow and Klentaq. This type of strong electrostatic repulsion is often observed in some halophilic enzymes, which have unusually high charge densities, due to the large numbers of acidic residues [262,279]. The abundance of acidic residues in halophilic proteins is often required to prevent aggregation in high salinity environment [262].

The comparison of amino acid composition among Klenpin, Klenow, and Klentaq, however, shows a similar acidic/basic residue ratios of around 1.0. The large number of acidic residues (acidic signature) may only exist on the surface of Klenpin, whereas the overall number of acidic residues is similar to basic residues.

We further investigate the possible acidic signature of the whole proteome and the DNA pol I among sea ice bacteria, salt-in halophiles, salt-out halophiles, and non-halophiles. The electrostatic surface potential maps of the DNA pol I are also compared.

The proteomes of salt-in halophiles (*Salinibacter ruber* and *Halobacterium sp. DL1*) show a high acidic/basic proteins (4.6 and 11.8). For the DNA pol I, however, only *Salinibacter ruber* has

a high acidic/basic residues and extremely high number of acidic residues (red patches) on its surface, whereas the DNA pol I from *Halobacterium sp. DL1* has similar acidic/basic residues and charge distribution to non-halophiles. It may imply that not all proteins in salt-in halophiles follows the same adaptation method to hypersaline environment.

For *P. ingrahamii*, neither the whole proteome or the DNA pol I shows any acidic signature, which is consistent with salt-out halophiles or non-halophiles and yet in contrast with the experimental finding that salt significantly stabilizes Klenpin, suggesting that a high acidic/basic residue ratio is not the only mechanism for making proteomes halotolerant. The electrostatic surface potential maps comparison of the DNA pol I from *P. ingrahamii*, *E. coli*, and *Thermus aquaticus* show a higher number of acidic residues (red patches) on its surface. All these clues may suggest that the adaptation of proteins in *P. ingrahamii* is through the modification of surface charge rather than the overall charge.

Salt effects on the thermodynamic stability of Klenpin and Klenow are further compared by urea and GdnHCl denaturations. First, the unfolding of Klenpin shifts from two-state to three-state in the presence of salt for both urea and GdnHCl denaturations. One possible explanation is that a more compact structure is formed by the addition of salt, which is supported by the structure data in CD spectra, DLS, and Trp quenching. Another evidence is that the increase in ΔG by the addition of salt is mainly from the new transition (ΔG1~ ΔΔG), which is barely affected by salt concentration above 250mM. Although Tm and ΔG are quantitatively different and non-linearly related indicators of a protein's stability, our results from Tm and ΔG comparison both show that Klenpin are enhanced by salt. In addition, Klenpin seems to have higher conformation stability than Klenow in the absence of salt in urea (Δ G(Klenpin) > Δ G(Klenow)), especially at low salt concentration.

In the absence of salt, the strong repulsive interactions not only destabilize halophilic proteins, but also affect their structures. Salt is required in the folding of some halophilic proteins [207,259], while the mechanism by salts stabilize a protein is still debatable. It has been reported that some halophilic enzymes are fully or partially unfolded in the absence of salt [191,195,198,206-209]. Complete distortion of the ellipticity pattern was observed in enzymes from some extremely halophiles [207]. Therefore, salt dependence of the CD spectra of Klenpin, Klenow, and Klentaq is investigated. An increase in the ellipticity of Klenpin was observed as the NaCl concentration increase, whereas small affect was seen on Klenow and Klentaq. The analysis of the secondary structures by BeStSel further suggest that salt significantly improves the helical content of Klenpin (33.7%→ 53.2%) and similar to the helix content of Klenow (50%) and Klentaq (61.5%).

Low helical contents were often observed in some psychrophilic proteins to maintain high flexibility at low temperature [135]. Proteome comparison between psychrophilic and mesophilic bacteria shows that there is a significant preference for aspartic acid, glycine, and threonine, as aspartic acid, glycine are helix breakers and threonine is a helix indifferent [135]. The amino acid composition of Klenpin, Klenow, and Klentaq, however, show no difference in the proportion of those amino acids. Furthermore, several computational methods are also applied to investigate the helical contents of the three homologous proteins based their sequences or three-dimensional structures. The results also show no difference in the helical contents among Klenpin, Klenow, and Klentaq (Table 5.4). Therefore, the lower helical content of Klenpin than Klenow and Klentaq in the absence of salt by BeStSel (Figure 4.15) is more salt related rather than cold adaptation.

The loss of structures of Klenpin in low salt concentration is further confirmed by dynamic light scattering (DLS) and tryptophan (Trp) fluorescence quenching. Both the size of Klenpin (from DLS) and the quenching constant (from Trp fluorescence quenching) significantly decreases

as NaCl concentration increases, whereas only small changes were observed in Klenow and Klentaq. It suggests that Klenpin has a more compact and possibly less dynamic structure at high NaCl concentrations compared with that at low NaCl concentrations. The structural work in this dissertation also suggests that Klenpin might form a dimer at high salt.

Salt also affects the unfolding pathway for Klenpin. In the absence of salt, the unfolding of Klenpin and Klenow by adding GdnHCl follows a two-state model. The two-state unfolding model of Klenow has been discussed previously [276]. Agreement of the Gibbs free energy between Kinetics unfolding experiments and equilibrium unfolding experiments is used as an indicator of two-state unfolding for Klenpin. In the absence of salt, however, an intermediate was observed in both GdnHCl and urea denaturation of Klenpin, whereas unfolding of Klenow still follows a simple reversible two-state model at high salt. Also, with a higher $\Delta\Delta G$, salt have a much stronger stabilization effect on Klenpin than that on Klenow, which is consistent with salt effect on Tm between Klenpin and Klenow. The mechanism for the formation of the salt induced intermediate is not clear, but it is clearly involved in the higher salt stabilization on Klenpin.

Little is known about how halophilic DNA-binding proteins work in the cell. DNA binding requires positively charged residues to interact with the negative charges on DNA phosphate groups and it's interesting to investigate how those proteins function at elevated salt concentrations. For mesophilic proteins, the interaction with DNA shows a decreased affinity with increasing salt concentration as ion release becomes less favorable [333]. For some halophilic DNA-binding proteins, however, a tighter binding with DNA with increasing salt concentrations is observed. Bergqvist et al proposed that this unexpected behavior may be due to the sequestration of cations into the protein-DNA complex by the overrepresented negative charges [332].

The salt dependence of DNA binding has not yet been studied for Klenpin, but several possibilities can be hypothesized. If the binding affinity of Klenpin is much less sensitive to the salt than Klenow and Klentaq, which have weaker binding as salt concentration increases, it suggests that Klenpin has some unique features in the structure which can maintain necessary binding affinity in the presence of high salinity. If Klenpin has a similar salt dependence of DNA binding as Klenow and Klentaq do, it suggests that the salinity in *P. ingrahamii* is similar to that in *E coli. or Thermus aquaticus* and the accumulation of osmolyte, rather than salt, is applied by the *P. ingrahamii* to balance osmotic pressure. Also, there could be a reversal of salt linkage in the Klenpin binding as found in the protein-DNA interactions in some halophilic bacteria [332]. The possible reversal of salt linkage for Klenpin binding in the presence of salt (NaCl or KCl) may suggest that high salt concentration could induce a conformational change in Klenpin, which may allow it to maintain high binding affinity in both low and high salinity environment.

The whole organism's requirement for NaCl was investigated by comparing the growth of *P. ingrahamii* on CLED (cystine–lactose–electrolyte-deficient) agar [306]. The results showed that no growth was observed in the absence of 0% NaCl. It grew well at 1-12% NaCl (0.17M to 2M). although details of the relative growth rate across this salt range were not described. Only weak growth was seen at 20% NaCl (~3.4M). These results are consistent with our ΔTm data (Figure 3.9) that the increases of ΔTm reach the plateau around 2M NaCl.

6.2 Conclusion

These studies provide insight into the temperature and salt dependence of the thermodynamic stability of Klenpin, a large fragment of DNA polymerase I with halophilic and psychrophilic features. By most measures other than urea denaturation, the overall conformational stability of Klenpin is slightly lower than Klenow, and a similar melting point for the two polymerases was

observed in the absence of salt. The unfolding of Klenpin follows a simple reversible two-state model in both GdnHCl and urea denaturation, whereas an intermediate was observed in presence of salt. The concentration of salts around 500mM is required for complete Klenpin folding, and salt dependence of ΔTm of Klenpin appears to possible be due to nonspecific screening of unfavorable electrostatic interactions on the protein surface, whereas a general Hofmeister effect is the primary salt effect on Klenow. No signature features were observed in proteome when comparing *P. ingrahamii* with other halophiles, whereas large number of acidic residues are observed on the surface of Klenpin, indicating a strong repulsive electrostatic interaction. The crystal structure is still needed to further study the specific mechanism for its salt dependence.

APPENDIX A. SUMMARY OF MELTING TEMPERATURES(TMS) OF KLENPIN AND KLENOW

Table 1 Summary of Tms of Klenpin at different salts or osmolyte.

NaCl(mM)	Tm(°C)	GdnHCl(mM)	Tm(°C)	K2SO4(mM)	Tm(°C)
0	35	50	39.36	50	48.02
100	40.9	100	41.11	100	50.5
200	43.63	150	41.94	150	50.36
300	45.33	200	42.21	200	52.83
400	46.64	250	42.27	250	54.7
500	47.75	MgCl2(mM)	Tm(°C)	300	56.86
750	49.54	600	42.69	350	57.29
1000	50.9	1200	42.12	NaPO4(mM)	Tm(°C)
1250	51.77	1800	39.51	50	40.23
1500	52.11	2400	32.02	100	42.83
KCl(mM)	Tm(°C)	NaHCO3(mM)	Tm(°C)	140	43.75
100	38.73	5	35.99	280	47.25
200	40.78	10	36.32	420	49.13
300	42.8	20	36.48	560	50.24
400	44.16	30	36.63	NaSCN(mM)	Tm(°C)
500	45.7	40	37.13	50	39.33
600	48.41	50	36.72	100	40.46
1200	50.99	100	37.05	150	40.98
1800	51.53	200	38.55	200	41.05
2400	51.15	300	39.86	250	40.67
Na2SO4(mM)	Tm(°C)	400	39.6	300	40.75
50	42.98	500	40.87	350	39.9
100	45.94	600	42.2	400	38.63
150	47.53	Glycine betaine(mM)	Tm(°C)	450	38.72
200	49.12	1000	41.63	500	37.23
250	50.2	2000	45.89		
300	50.82	3000	49.21		
400	51.97	4000	52.48		
500	53.1				

Table 2 Summary of Tms of Klenow at different salts or osmolyte.

NaCl(mM)	Tm(°C)	GdnHCl(mM)	Tm(°C)	NaSCN(mM)	Tm(°C)
0	57.5	50	56.895	50	58.68
600	59.9	100	57.345	100	58.47
1200	61.7	150	56.045	150	57.97
1800	62.8	200	55.39	200	57.2
KCl(mM)	Tm(°C)	250	55.364	250	56.46
600	58.3	MgCl2(mM)	Tm(°C)	300	54.81
1200	61.18	600	54.1	350	53.37
1800	63.05	1200	55.98	400	50.95
2400	64.18	1800	52.96	K2SO4(mM)	Tm(°C)
Na2SO4(mM)	Tm(°C)	2400	45.91	50	63.14
380	59.09	NaHCO3(mM)	Tm(°C)	100	64.58
600	61.8	5	55.42	150	65.85
760	63.3	10	53.16	200	66.67
NaPO4(mM)	Tm(°C)	20	51.58	250	67.88
140	55.35	30	51.02	300	68.6
280	61.17	40	49.83	350	69.83
420	59.6	50	49.85		
560	60.57	100	49.96		
Glycine betaine(mM)	Tm(°C)	200	50.62		
1000	62.31	300	51.24		
2000	66.19	400	50.75		
3000	69.23	500	50.95		
4000	71.91	600	50.721		

APPENDIX B. RESEARCH SUMMARY OF GREEN FLUORESCENT PROTEIN

B.1. Background

Green fluorescent protein (GFP) is a protein composed of 236 amino acids which exhibits fluorescence under the light range from blue to ultraviolet. Osamu Shimomura first purified GFP from *Aequorea victoria* in 1960s[1]. The successful expression of wtGFP in heterologous cells of *E. coli* and *C. elegans* was first published in *Science* by the lab of Martin Chalfie in 1994[2]. Since then, green fluorescent protein (GFP) has been increasingly used for *in vivo* imaging, intracellular targeting studies, and many other applications because of its remarkable self-catalyzed chromophore formation.

The overall structure of GFP contains a beta-barrel of eleven strands, surrounding a central helix. The barrel ends are capped by short distorted helical segments, which protect the chromophore from solvents [3]. Upon unfolding, the chromophore is exposed to the solvent and therefore loses all fluorescence. As a result, the loss of fluorescence can be used to analyze the unfolding of GFP. Wild type (wt) GFP has some drawbacks, such as low folding efficiency at 37°C and low quantum yield [4], which have severely limited its applications. Mutagenesis strategies has been applied to produce different variants of GFP, such as GFPuv, GFPmut2 and superfolder GFP, which have better folding efficiency and higher quantum yield [5,6].

My research focuses on developing a new method to measure protein unfolding in living cells. The first part of my project is the comparison of the folding and unfolding of green fluorescent protein (GFP) *in vivo* and *in vitro*. Green fluorescent protein exhibits bright green fluorescence when exposed to ultraviolet light. By using this property, it's easy to quantitatively analyze the unfolding fraction of GFP by testing its fluorescence loss. During protein unfolding in cells, the protein is surrounded by a very complex environment, which includes the nucleus,

organelles, and many other types of molecules. Also, a lot of diseases, such as cancer, diabetes, and liver disorders, are related to the accumulation of unfolded protein in the cell. For these reasons, understanding how protein folding *in vivo* differs from *in vitro* is important.

Since many GFP variants are a highly thermally stable proteins with a melting points above 70 °C, we chose *Thermus aquaticus*, which can resist temperatures up to 85 °C, as a model bacterium for studying protein folding *in vivo*. The next step in this project is the expression of GFP in *Thermus aquaticus*. After searching the literature, we found that a cloning vector called pMK18 can be used to express different variants of GFP in *Thermus aquaticus*[7]. Another potential method is using cell penetrating pepetides(CPPs) which are short peptides facilitating cellular uptake of various molecular cargo[8].

In these experiments, we plan to get Tm, ΔH, and ΔCp values for *in vivo* protein folding at different temperatures, and then these values can be used to calculate ΔG by using the Gibbs-Helmholtz equation.

Research Activities

We have successfully purified GFP from *E. coli* and compared the thermal denaturation of two variants of GFP, GFPuv and GFPmut2 using both fluorescence and circular dichroism (CD) *in vitro*.

B.2. Methods and Materials

Two-state model

Unfolding transitions are fit to a two-state model of protein unfolding according to the equation:

$$I = \frac{I_{0,N} + I_{0,U}.\times e^{\frac{-\Delta H}{1.987*(T+273)}*(1-\frac{T+273}{T_m})}}{1 + e^{\frac{-\Delta H}{1.987*(T+273)}*(1-\frac{T+273}{T_m})}}$$

I: observed signal intensity;

$I_{0,N}$: signal intensity of the native state of the protein;

$I_{0,U}$: signal intensity of the native state of the protein;

ΔH: the change of enthalpy upon unfolding;

Tm: melting temperature where half of the proteins are denatured;

Protein Expression and Extraction

The *E. coli* DH5α cells containing plasmid pKen-GFPmut2 were obtained from Addgene and were cultured in Luria-Bertani (LB) broth containing ampicillin (100 µg/mL). The culture was incubated in an orbital shaker at 37°C and 200 rpm. Protein expression was induced by isopropylthio-ß-D-galactosidase(IPTG) at a final concentration of 0.5mM when cell density is 0.6. The culture was further incubated for 16 hours for the overexpression of GFP.

$(NH_4)_2SO_4$ was gradually added to the clarified supernatant at 12g/100ml (w/v) and the mixture was stirred slowly for 1 hour at 4 °C and then was immediately heated at 70°C for 7 minutes[9]. Precipitated proteins were removed after the heat step by centrifugation for 20 min at 8000g [9].

Purification

We first purified two variants of green fluorescent protein, GFPuv and GFPmut2 by using hydrophobic interaction chromatography (HIC) and size exclusion chromatography (SEC).

Hydrophobic Interaction Chromatography (HIC)

Hydrophobic interaction chromatography (HIC) (BioRad Macro-Prep Methyl HCI Support) is a technique for separating proteins by their degree of hydrophobicity. The main mechanism for

this technique depends on the interaction between hydrophobic ligands on media with hydrophobic amino acids exposed on the protein surface. Green fluorescent protein is extremely hydrophobic compared with other proteins in bacterium and some of its amino acids are exposed to the surface, therefore we can use HIC as an initial step for the purification.

Size exclusion chromatography (SEC)

Size exclusion chromatography (SEC) is a method in which the molecules are separated by their size. Small molecules can pass through the pores of the packing material in stationary phase then elute later while large molecules can only pass through the interparticle volume and elute earlier.

Fluorescence Measurement

Fluorescence emission spectra and single-wavelength kinetic traces were collected on an ISS PX01 Koala photon counting fluorometer . The protein fluorescence was excited at 460nm (GFPmut2) or 370nm (GFPuv), using 2-nm excitation slits. The emission was acquired at 509nm (GFPuv) or 505nm (GFPmut2). The experiment was carried in 50mM phosphate buffer at a protein concentration of 8ng/ml.

Circular Dichroism (CD) Measurements

CD measurements were carried out using a Jasco Model J-815 circular dichroism spectrometer equipped with temperature control. All the samples were measured from 25 °C to 90 °C for thermal denaturation in 50mM phosphate buffer at the protein concentration of 50ng/ml. Spectra were collected in the range 210–260 nm, and single-wavelength kinetics were collected at 220 nm.

B.3. Results

Protein purification

Hydrophobic interaction chromatography was carried out using a Macro-prep Methyl HIC Support (BIO-RAD). Proteins were bound to the support by loading in 50 mM Phosphate buffer, containing 2M $(NH_4)_2SO_4$ and then eluted from the column according to the degree of hydrophobicity by reducing $(NH_4)_2SO_4$ concentration from 2M to 0M. Elution of protein was monitored at the absorption maximum of GFP[9]. The results are shown in Figure 1.

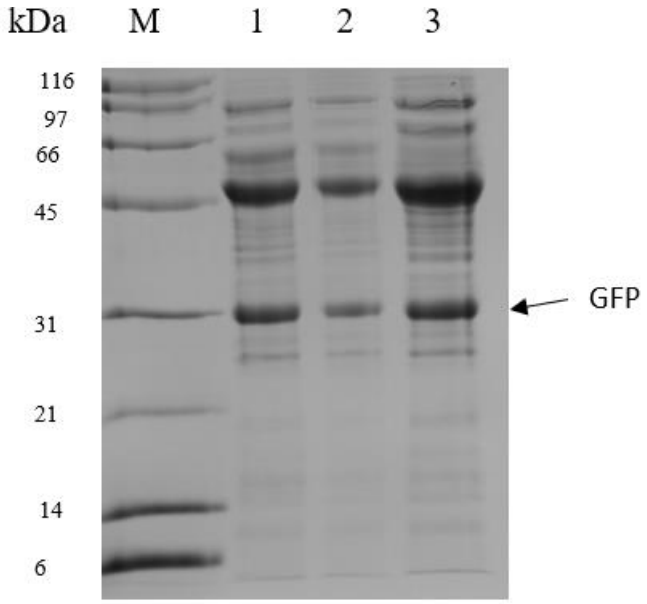

Figure B.1. SDS-PAGE analyses of partially purified GFP after HIC. Lane M: molecular mass markers in kDa, lane 1, 2, 3: purified GFP at different concentrations (lane2<lane1<lane3). The arrow shows the location of GFP

Size exclusion chromatography (SEC) was carried out using Sephacryl 300-HR Size Exclusion Media (SIGMA). Protein from HIC was loaded on the column, which was equilibrated with 20mM Tris and 1M ammonium sulfate, and then eluted with 50mM phosphate buffer (pH7.5). The result is shown in Figure 2.

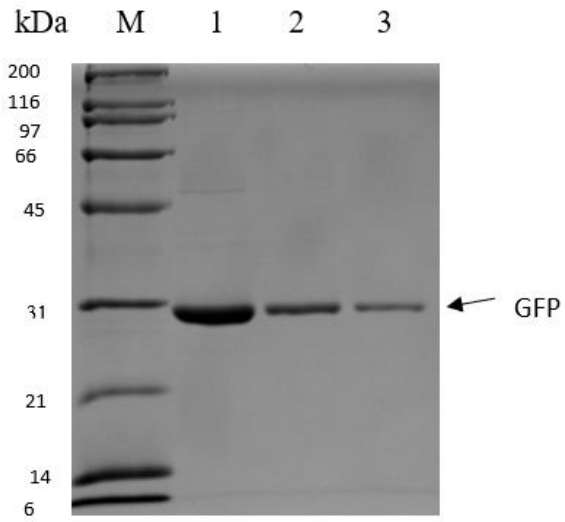

Figure B.2. SDS-PAGE analyses of partially purified GFP after SEC. Lane M: molecular mass markers in kDa, lanes 1, 2, 3: purified GFP in different amounts. The arrow shows the location of GFP

Unfolding of GFP by Fluorescence and CD

Unfolding of the two GFP variants by fluorescence and CD was carried out by increasing the temperature from 25 ℃ to 90 °C in 50 mM phosphate buffer (pH 6.2) and the results were fit to a two-state model. The difference in melting temperature for GFPuv between two methods is 10.3°C while the difference in Tm for GFPmut2 is only 1.4℃ (Figure 3). Also, the two fitted lines for GFPmut2 are almost overlapped and therefore suggest t that GFPmut2 is a better candidate than GFPuv for unfolding studies. However, the equilibrium state of unfolding can take a long time to reach as a previous study showed [10] and therefore this data cannot be used to calculate the free energy of unfolding at present.

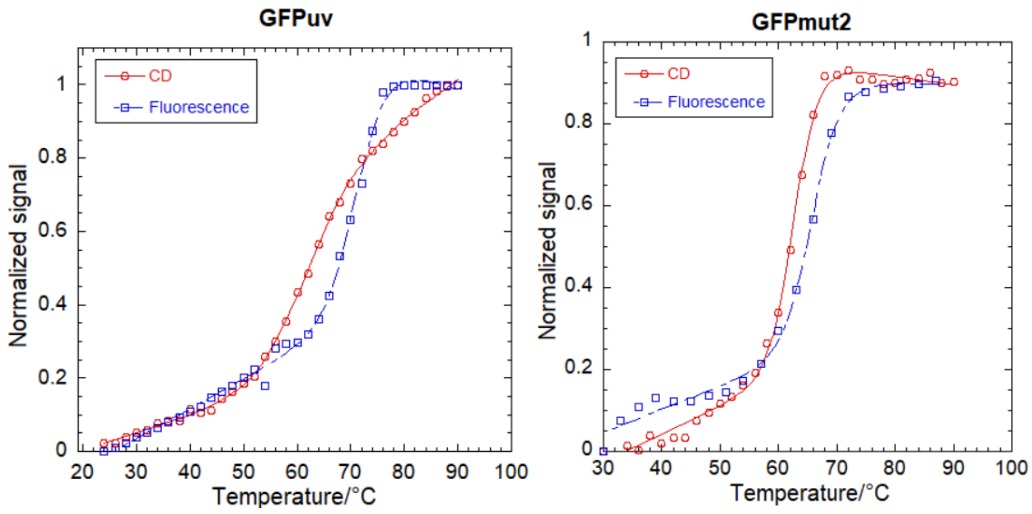

Figure B.3. Unfolding of GFPuv (left) and GFPmut2 (right) by CD (circle) or fluorescence (square) from 24°C to 90°C.

pH sensitivity

pH sensitivity of GFPuv and GFPmut2 was analyzed in phosphate buffer at different pH value from pH 5.5 to 8 by both fluorescence and CD. The results show that the secondary structure was relatively resistant to pH for both GFPuv and GFPmut2 since the CD signal only shifts a little when the pH changes (Figure5). The fluorescence, however, is sensitive to pH change especially for GFPmut2, which lost more than 80% fluorescence when pH dropped from 7.5 to 5.5 (Figure 4). Comparing the different pH sensitivity between two methods, we conclude that the loss of fluorescence is not due to the protein unfolding but probably because of the protonation-de-protonation of residues at or near the chromophore [11].

Also, unexpected increase of CD signal during the thermal denaturation of GFPmut2 suggests that aggregation of unfolded GFP happen at high pH (>6.5). (Figure6)

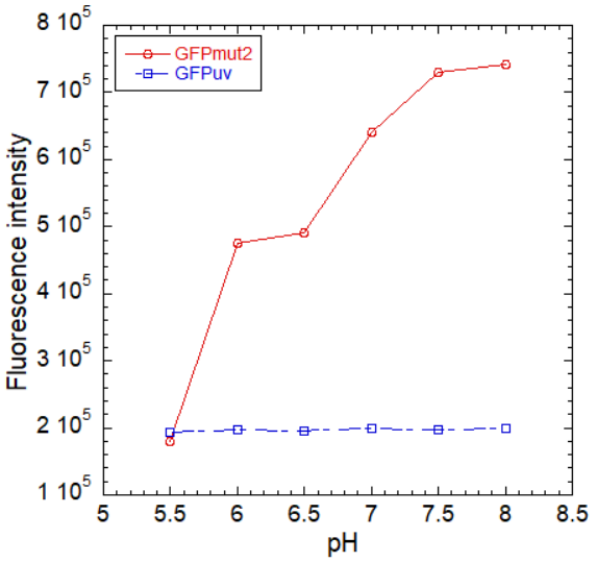

Figure B.4. Fluorescence intensity of GFPuv(square) and GFPmut2(circle) in 50mM phosphate buffer at different pH.

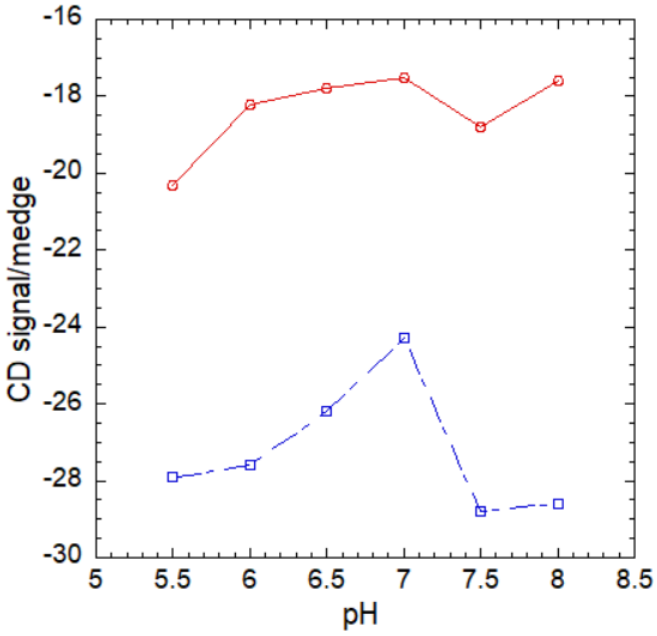

Figure B.5. Circular dichroism (CD) signal of GFPuv(square) and GFPmut2(circle) in 50mM phosphate buffer at different pH.

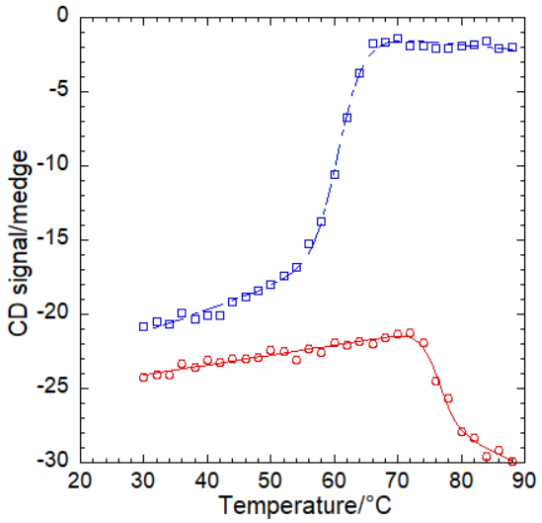

Figure B.6. Aggregation of denatured GFPmut2 at high pH (>6.5). Unexpected increase of CD signal at pH 6.5 (circle) is observed when the temperature is above 70°C while the CD signal is close to 0 at 90°C for pH 5.5(square)

Hysteresis in the unfolding of GFP

In the previous studies, a lag phase before reaching the equilibrium state has been observed [12]. Unfortunately, we also observed this hysteresis in our unfolding experiments. The CD signal gradually decreased when GFP was incubated at 60°C for 1 hour, suggesting that it did not reach equilibrium during this time.

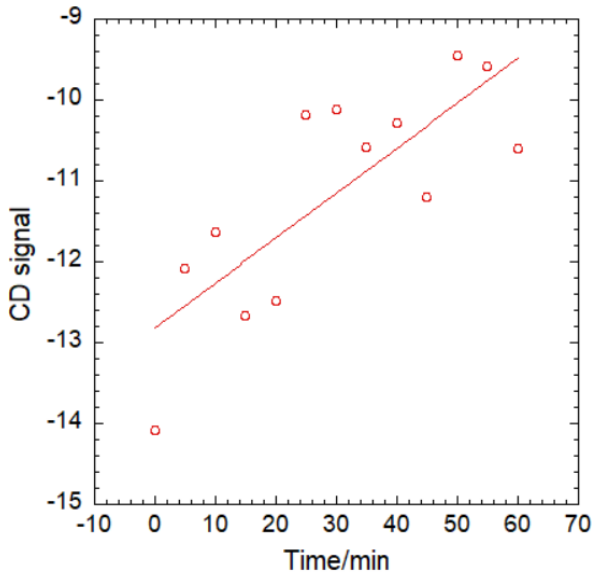

Figure B.7. Decrease of CD signal for GFP incubated at 60°C in 50mM phosphate buffer (pH 6.0).

Reversible Photobleaching of GFPmut2

Photobleaching, the photochemical alteration of a fluorophore molecule, usually causes the permanent loss of fluorescence. Although photobleaching occurs when GFPmut2 is exposed to ambient light, it has the unusual property of being reversible photobleaching[13] that the fluorescence slowly recovers in the dark and recovery can even be further accelerated by additional excitation light. In stage 1 of Figure 8, nearly 40% of fluorescence was lost under ambient light before GFP was incubated under excitation light at 470nm. The fluorescence nearly recovered to 100% after 30-40 min. In stage 2, fluorescence was quickly lost when incubated under ambient light within 15 min. In stage 3: after partial photobleaching, GFP was incubated in the dark and fluorescence slowly recovered. (Figure 8)

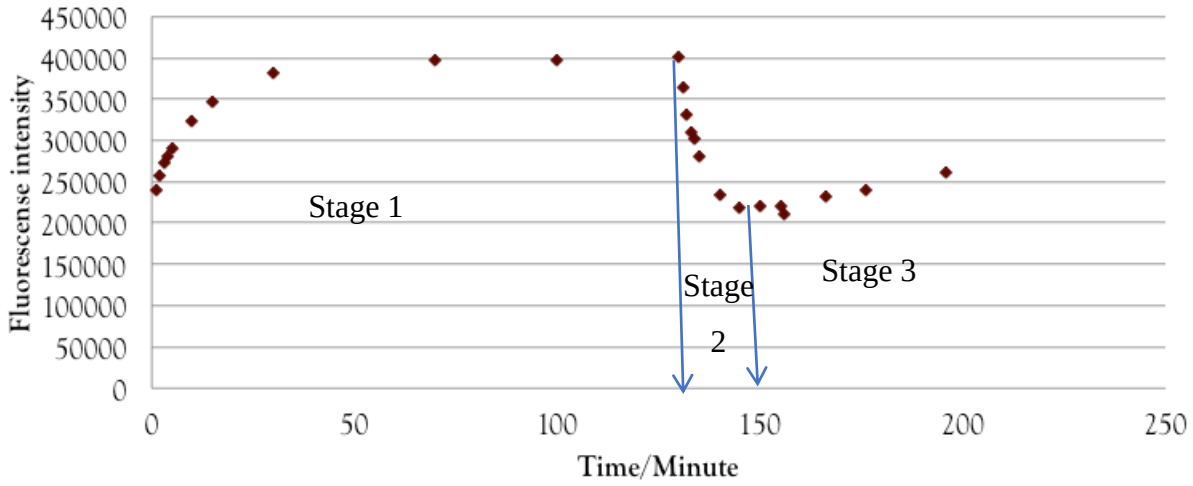

Figure B.8. Reversible photobleaching of GFPmut2. Stage 1: Fast fluorescence recovery with excitation light after photobleaching by ambient light; Stage 2: photobleaching by ambient light; Stage 3: Slow fluorescence recovery in the dark.

B.4. Conclusions

We have investigated the unfolding of two variants of GFP, GFPuv and GFPmut2 *in vitro*. The results indicate that GFPmut2 shows a better consistency between fluorescence and CD results, which suggest it may be a better model protein for *in vivo* study. However, hysteresis and a lag

phase before reaching equilibrium state have been observed during the unfolding of both GFPuv and GFPmut2, which make it difficult to fit the data to a two-state unfolding model. In addition, GFPmut2 has high pH sensitivity, where high pH (pH>6.5) can cause protein aggregation, whereas lost of fluorescence will happen at low pH (pH<5.5).

B.5. Future plans

The next step for my project is to study the unfolding of GFPmut2 in *Thermus aquaticus.* As mentioned previous, pMK18 is a potential vector for expression of GFP in *Thermus aquaticus.* [7]Another potential method is to use cell penetrating pepetides(CPPs).[8]

A main problem for my project right now is the hysteresis in the unfolding of GFP and therefore we are unable to calculate ΔH and ΔG by fitting data to a two-state model. We aim to find regions of the unfolding landscape where the proteins unfold in a reversible two-state manner by examining the unfolding of the two GFPs under different solution conditions. Also, we will examine other GFP variants to find a better behaved variant for *in vivo* study.

B.6. Reference

[1] Shimomura O, Johnson FH, Saiga Y. "Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea". Journal of Cellular and Comparative Physiology (1962), 223–39.

[2] Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC . "Green fluorescent protein as a marker for gene expression". Science 263 (1994), 80.

[3] S. James Remington. "Green fluorescent protein: A perspective". Protein Science (2011), 1509-1519.

[4] Peternel S, Grdadolnik J, Gaberc-Porekar V, Komel R. "Engineering inclusion bodies for non denaturing extraction of functional proteins. "Microb Cell Fact (2008), 34.

[5]Brendan P. Cormack , Raphael H. Valdivia and Stanley Falkow. "FACS-optimized mutants of the green fluorescent protein (GFP) ". Gene(1993), 33-38.

[6] Jean-Denis Pedelacq, Stephanie Cabantous, Timothy Tran, Thomas C Terwilliger & Geoffrey S Waldo, "Engineering and characterization of a superfolder green fluorescent protein." Nature Biotechnology(2006), 79-89.

[7] Cava F, de Pedro MA, Blas-Galindo E, Waldo GS, Westblade LF, Berenguer J. "Expression and use of superfolder green fluorescent protein at high temperatures in vivo: a tool to study extreme thermophile biology". Environmental Microbiology (2008), 605–613

[8] Gunaratna Kuttuva Rajarao, Natalia Nekhotiaeva, Liam Good. "Peptide-mediated delivery of green fuorescent protein into yeasts and bacteria". FEMS Microbiology Letters(2002), 267–272.
[9] Shelley R. McRae, Christopher L. Brown, Gillian R. Bushell. "Rapid puriWcation of EGFP, EYFP, and ECFP with high yield and purity". Elsevier(2005), 121-127.

[10] Shang-Te Danny Hsu, Georg Blaser and Sophie E. Jackson. "The folding, stability and conformational dynamics of b-barrel fluorescent proteins". Chemical Society Reviews (2009), 2951–2965.

[11] Malea Kneen, Javier Farinas, Yuxin Li, and A. S. Verkman. "Green Fluorescent Protein as a Noninvasive Intracellular pH Indicator." Biophysical Journal(1998), 1591-1599.

[12] Shang-Te Danny Hsu, Georg Blaser and Sophie E. Jackson. "The folding, stability and conformational dynamics of b-barrel fluorescent proteins." Chemical Society Review (2009), 2951–2965

[13] Daniel Sinnecker, Philipp Voigt, Nicole Hellwig, and Michael Schaefer. "Reversible Photobleaching of Enhanced Green Fluorescent Proteins." Biochemistry (2005), 7085-7094.

REFERENCES

[1] De Maayer, P., Anderson, D., Cary, C., & Cowan, D. A. (2014). Some like it cold: understanding the survival strategies of psychrophiles. EMBO reports 15(5), 508-517.

[2] Margesin, R., & Miteva, V. (2011). Diversity and ecology of psychrophilic microorganisms. Research in microbiology, 162(3), 346-361.

[3] Breezee, J., Cady, N., & Staley, J. T. (2004). Subfreezing growth of the sea ice bacterium "P. ingrahamii". Microbial Ecology, 47(3), 300-304.

[4] Von Hippel, P. H., & Wong, K. Y. (1965). On the Conformational Stability of Globular Proteins, the effects of various electrolytes and nonelectrolytes on the thermal ribonuclease transition. Journal of Biological Chemistry, 240(10), 3909-3923.

[5] Ventosa, A., & Ventosa, A. (2004). Halophilic microorganisms (p. 349pp). Berlin; Springer.

[6] Qu, Y., Bolen, C. L., & Bolen, D. W. (1998). Osmolyte-driven contraction of a random coil protein. Proceedings of the National Academy of Sciences, 95(16), 9268-9273.

[7] Baker, John Tod, (2018) "Cloning, Purification, and Preliminary DNA-binding and Unfolding Results for the DNA Polymerase I from the Psychrophile P. ingrahamii". *LSU Master's Theses*. 4693

[8] Greenfield, N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. Nature protocols, 1(6), 2876.

[9] Feller, G., & Gerday, C. (1997). Psychrophilic enzymes: molecular basis of cold adaptation. Cellular and Molecular Life Sciences CMLS, 53(10), 830-841.

[10] Richard, A. J., Liu, C. C., Klinger, A. L., Todd, M. J., Mezzasalma, T. M., & LiCata, V. J. (2006). Thermal stability landscape for Klenow DNA polymerase as a function of pH and salt concentration. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1764(10), 1546-1552.

[11] Bye, J. W., & Falconer, R. J. (2013). Thermal stability of lysozyme as a function of ion concentration: a reappraisal of the relationship between the Hofmeister series and protein stability. Protein Science, 22(11), 1563-1570.

[12] Canchi, D. R., & García, A. E. (2011). Backbone and side-chain contributions in protein denaturation by urea. Biophysical journal, 100(6), 1526-1533.

[13] Gounot, A. M. (1976). Effects of temperature on the growth of psychrophilic bacteria from glaciers. Canadian journal of microbiology, 22(6), 839-846.

[14] Takai, K., Nakamura, K., Toki, T., Tsunogai, U., Miyazaki, M., Miyazaki, J., ... & Horikoshi, K. (2008). Cell proliferation at 122 C and isotopically heavy CH4 production by a hyperthermophilic methanogen under high-pressure cultivation. Proceedings of the National Academy of Sciences, 105(31), 10949-10954.

[15] Santoro, M. M., & Bolen, D. W. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl. alpha.-chymotrypsin using different denaturants. Biochemistry, 27(21), 8063-8068.

[18] Riley, M., Staley, J. T., Danchin, A., Wang, T. Z., Brettin, T. S., Hauser, L. J., ... & Thompson, L. S. (2008). Genomics of an extreme psychrophile, P. ingrahamii. BMC genomics, 9(1), 210.

[19] Madern, D., Ebel, C., & Zaccai, G. (2000). Halophilic adaptation of enzymes. Extremophiles, 4(2), 91-98.

[20] Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. Science, 181(4096), 223-230.

[21] Levinthal, C. (1968). Are there pathways for protein folding? Journal de chimie physique, 65, 44-45.

[22] Stelea, S., Pancoska, P., & Keiderling, T. A. (1999). Intermediates in the thermal unfolding of ribonuclease A. In Spectroscopy of biological molecules: New directions (pp. 65-66). Springer, Dordrecht.

[23] Jackson, S. E., & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. Biochemistry, 30(43), 10428-10435.

[24] Huang, G. S., & Oas, T. G. (1995). Structure and Stability of Monomeric. lambda. Repressor: NMR Evidence for Two-State Folding. Biochemistry, 34(12), 3884-3892.

[25] Dill, K. A., & Chan, H. S. (1997). From Levinthal to pathways to funnels. Nature structural biology, 4(1), 10.

[26] Schopf, J. W., Kudryavtsev, A. B., Czaja, A. D., & Tripathi, A. B. (2007). Evidence of Archean life: stromatolites and microfossils. Precambrian Research, 158(3-4), 141-155.

[27] Schopf, J. W. (2006). Fossil evidence of Archaean life. Philosophical Transactions of the Royal Society B: Biological Sciences, 361(1470), 869-885.

[28] Mars Exploration Rover Launches - Press kit" NASA. June 2003. Retrieved 14 July 2009.

[29] Roberts, M. F. (2005). Organic compatible solutes of halotolerant and halophilic microorganisms. Saline systems, 1(1), 5.

[30] Creighton, T. E. (1993). Proteins: structures and molecular properties. Macmillan.

[31] Oakenfull, D., & Fenwick, D. E. (1979). Hydrophobic interaction in aqueous organic mixed solvents. Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases, 75, 636-645.

[32] Tadeo, X., López-Méndez, B., Castaño, D., Trigueros, T., & Millet, O. (2009). Protein stabilization and the Hofmeister effect: the role of hydrophobic solvation. Biophysical journal, 97(9), 2595-2603.

[33] Collins, K. D. (2004). Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. Methods, 34(3), 300-311.

[34] Fukuchi, S., Yoshimune, K., Wakayama, M., Moriguchi, M., & Nishikawa, K. (2003). Unique amino acid composition of proteins in halophilic bacteria. Journal of molecular biology, 327(2), 347-357.

[35] Siglioccolo, A., Paiardini, A., Piscitelli, M., & Pascarella, S. (2011). Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. BMC structural biology, 11(1), 50.

[36] Marguet, E., & Forterre, P. (1998). Protection of DNA by salts against thermodegradation at temperatures typical for hyperthermophiles. Extremophiles, 2(2), 115-122.

[37] Seckbach, J. (Ed.). (2013). Enigmatic microorganisms and life in extreme environments (Vol. 1). Springer Science & Business Media.

[38] Methé, B. A., Nelson, K. E., Deming, J. W., Momen, B., Melamud, E., Zhang, X., ... & Brinkac, L. M. (2005). The psychrophilic lifestyle as revealed by the genome sequence of Colwellia psychrerythraea 34H through genomic and proteomic analyses. Proceedings of the National Academy of Sciences, 102(31), 10913-10918.

[39] Sælensminde, G., Halskau, Ø., & Jonassen, I. (2009). Amino acid contacts in proteins adapted to different temperatures: hydrophobic interactions and surface charges play a key role. Extremophiles, 13(1), 11.

[40] Lewis, P. N., Kotelchuck, D., & Scheraga, H. A. (1970). Helix probability profiles of denatured proteins and their correlation with native structures. Proceedings of the National Academy of Sciences, 65(4), 810-815.

[41] Gosink, J. J., & Staley, J. T. (1995). Biodiversity of gas vacuolate bacteria from Antarctic sea ice and water. Appl. Environ. Microbiol., 61(9), 3486-3489.

[42] Gromiha, M. M., Oobatake, M., & Sarai, A. (1999). Important amino acid properties for enhanced thermostability from mesophilic to thermophilic proteins. Biophysical chemistry, 82(1), 51-67.

[43] Johns, G. C., & Somero, G. N. (2004). Evolutionary convergence in adaptation of proteins to temperature: A4-lactate dehydrogenases of Pacific damselfishes (Chromis spp.). Molecular biology and evolution, 21(2), 314-320.

[44] Pack, S. P., & Yoo, Y. J. (2004). Protein thermostability: structure-based difference of amino acid between thermophilic and mesophilic proteins. Journal of Biotechnology, 111(3), 269-277.

[45] Panasik Jr, N., Brenchley, J. E., & Farber, G. K. (2000). Distributions of structural features contributing to thermostability in mesophilic and thermophilic α/β barrel glycosyl hydrolases. Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology, 1543(1), 189-201.

[46] Yang, L. L., Tang, S. K., Huang, Y., & Zhi, X. Y. (2015). Low temperature adaptation is not the opposite process of high temperature adaptation in terms of changes in amino acid composition. *Genome biology and evolution*, *7*(12), 3426-3433.

[47] Lehman, I. R., Bessman, M. J., Simms, E. S., & Kornberg, A. (1958). Enzymatic synthesis of deoxyribonucleic acid I. Preparation of substrates and partial purification of an enzyme from Escherichia coli. Journal of Biological Chemistry, 233(1), 163-170.

[48] Loeb, L. A., & Monnat Jr, R. J. (2008). DNA polymerases and human disease. Nature Reviews Genetics, 9(8), 594.

[49] Myers, J. K., & Pace, C. N. (1996). Hydrogen bonding stabilizes globular proteins. Biophysical journal, 71(4), 2033-2039.

[50] Pace, C. N., Horn, G., Hebert, E. J., Bechert, J., Shaw, K., Urbanikova, L., ... & Sevcik, J. (2001). Tyrosine hydrogen bonds make a large contribution to protein stability. Journal of molecular biology, 312(2), 393-404.

[51] Shi, Z., Krantz, B. A., Kallenbach, N., & Sosnick, T. R. (2002). Contribution of hydrogen bonding to protein stability estimated from isotope effects. Biochemistry, 41(7), 2120-2129.

[52] Cao, Z., & Bowie, J. U. (2014). An energetic scale for equilibrium H/D fractionation factors illuminates hydrogen bond free energies in proteins. Protein Science, 23(5), 566-575.

[53] Dill, K. A. (1990). Dominant forces in protein folding. Biochemistry, 29(31), 7133-7155.

[54] Edelhoch, H., & Osborne Jr, J. C. (1976). The thermodynamic basis of the stability of proteins, nucleic acids, and membranes. In Advances in protein chemistry (Vol. 30, pp. 183-250). Academic Press.

[55] Harding, T., Brown, M. W., Simpson, A. G., & Roger, A. J. (2016). Osmoadaptative strategy and its molecular signature in obligately halophilic heterotrophic protists. Genome biology and evolution, 8(7), 2241-2258.

[56] Arakawa, T., & Timasheff, S. N. (1985). The stabilization of proteins by osmolytes. Biophysical journal, 47(3), 411-414.

[57] Burg, M. B., & Ferraris, J. D. (2008). Intracellular organic osmolytes: function and regulation. Journal of Biological Chemistry, 283(12), 7309-7313.

[58] Auton, M., Rösgen, J., Sinev, M., Holthauzen, L. M. F., & Bolen, D. W. (2011). Osmolyte effects on protein stability and solubility: A balancing act between backbone and side-chains. Biophysical chemistry, 159(1), 90-99.

[59] Gerday, C. (2008). Psychrophiles: from biodiversity to biotechnology (p. 462). R. Margesin, F. Schinner, & J. C. Marx (Eds.). Berlin: Springer.

[60] Linhananta, A., Hadizadeh, S., & Plotkin, S. S. (2011). An effective solvent theory connecting the underlying mechanisms of osmolytes and denaturants for protein stability. Biophysical journal, 100(2), 459-468.

[61] Tanford, C. (1970). Theoretical models for the mechanism ofdenaturation. Adv. Protein Chem, 24, 2-97.

[62] Jackson, S. E. (1998). How do small single-domain proteins fold? Folding and Design, 3(4), R81-R91.

[63] Solomon, E. I., & Lever, A. B. P. (Eds.). (1999). Inorganic Electronic Structure and Spectroscopy, Applications and Case Studies (Vol. 2). Wiley-Interscience.

[64] Dalby, P. A., Clarke, J., Johnson, C. M., & Fersht, A. R. (1998). Folding intermediates of wild-type and mutants of barnase. II. Correlation of changes in equilibrium amide exchange kinetics with the population of the folding intermediate. Journal of molecular biology, 276(3), 647-656.

[65] Antón, J., Oren, A., Benlloch, S., Rodríguez-Valera, F., Amann, R., & Rosselló-Mora, R. (2002). Salinibacter ruber gen. nov., sp. nov., a novel, extremely halophilic member of the Bacteria from saltern crystallizer ponds. International journal of systematic and evolutionary microbiology, 52(2), 485-491.

[66] Morita, R. Y. (1975). Psychrophilic bacteria. Bacteriological reviews, 39(2), 144.

[67] Mykytczuk, N. C., Foote, S. J., Omelon, C. R., Southam, G., Greer, C. W., & Whyte, L. G. (2013). Bacterial growth at− 15 C; molecular insights from the permafrost bacterium Planococcus halocryophilus Or1. The ISME journal, 7(6), 1211.

[68] Price, P. B., & Sowers, T. (2004). Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. Proceedings of the National Academy of Sciences, 101(13), 4631-4636.

[69] Mock, T., & Thomas, D. N. (2005). Recent advances in sea-ice microbiology. Environmental Microbiology, 7(5), 605-619.

[70] BinSheng, Y. (2011). The Stability of a Three-State Unfolding Protein. Thermodynamics: Physical Chemistry of Aqueous Systems, 365.

[71] Rabus, R., Ruepp, A., Frickey, T., Rattei, T., Fartmann, B., Stark, M., ... & Amann, J. (2004). The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium from permanently cold Arctic sediments. Environmental microbiology, 6(9), 887-902.

[72] Ayala-del-Río, H. L., Chain, P. S., Grzymski, J. J., Ponder, M. A., Ivanova, N., Bergholz, P. W., ... & Rodrigues, D. (2010). The genome sequence of Psychrobacter arcticus 273-4, a psychroactive Siberian permafrost bacterium, reveals mechanisms for adaptation to low-temperature growth. Appl. Environ. Microbiol., 76(7), 2304-2312.

[73] Lauro, F. M., Tran, K., Vezzi, A., Vitulo, N., Valle, G., & Bartlett, D. H. (2008). Large-scale transposon mutagenesis of Photobacterium profundum SS9 reveals new genetic loci important for growth at low temperature and high pressure. Journal of bacteriology, 190(5), 1699-1709.

[74] Math, R. K., Jin, H. M., Kim, J. M., Hahn, Y., Park, W., Madsen, E. L., & Jeon, C. O. (2012). Comparative genomics reveals adaptation by Alteromonas sp. SN2 to marine tidal-flat conditions: cold tolerance and aromatic hydrocarbon metabolism. PLoS One, 7(4), e35784.

[75] D'Amico, S., Collins, T., Marx, J. C., Feller, G., & Gerday, C. (2006). Psychrophilic microorganisms: challenges for life. EMBO reports, 7(4), 385-389.

[76] Klähn, S., & Hagemann, M. (2011). Compatible solute biosynthesis in cyanobacteria. Environmental microbiology, 13(3), 551-562.

[77] Casanueva, A., Tuffin, M., Cary, C., & Cowan, D. A. (2010). Molecular adaptations to psychrophily: the impact of 'omic'technologies. Trends in microbiology, 18(8), 374-381.

[78] Celik, Y., Drori, R., Pertaya-Braun, N., Altan, A., Barton, T., Bar-Dolev, M., ... & Braslavsky, I. (2013). Microfluidic experiments reveal that antifreeze proteins bound to ice crystals suffice to prevent their growth. Proceedings of the National Academy of Sciences, 110(4), 1309-1314.

[79] Gerday, C., Aittaleb, M., Bentahir, M., Chessa, J. P., Claverie, P., Collins, T., ... & Hoyoux, A. (2000). Cold-adapted enzymes: from fundamentals to biotechnology. Trends in biotechnology, 18(3), 103-107.

[80] Sellek, G. A., & Chaudhuri, J. B. (1999). Biocatalysis in organic media using enzymes from extremophiles. Enzyme and Microbial Technology, 25(6), 471-482.

[81] Rina, M., Pozidis, C., Mavromatis, K., Tzanodaskalaki, M., Kokkinidis, M., & Bouriotis, V. (2000). Alkaline phosphatase from the Antarctic strain TAB5: properties and psychrophilic adaptations. European journal of biochemistry, 267(4), 1230-1238.

[82] Struvay, C., & Feller, G. (2012). Optimization to low temperature activity in psychrophilic enzymes. International journal of molecular sciences, 13(9), 11643-11665.

[83] Chan, C. K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A., & Hofrichter, J. (1997). Submillisecond protein folding kinetics studied by ultrarapid mixing. Proceedings of the National Academy of Sciences, 94(5), 1779-1784.

[84] Šmidlehner, T., Piantanida, I., & Pescitelli, G. (2018). Polarization spectroscopy methods in the determination of interactions of small molecules with nucleic acids–tutorial. Beilstein journal of organic chemistry, 14(1), 84-105.

[85] Jenkins, J., & Pickersgill, R. (2001). The architecture of parallel β-helices and related folds. Progress in biophysics and molecular biology, 77(2), 111-175.

[86] Feller, G. (2010). Protein stability and enzyme activity at extreme biological temperatures. Journal of Physics: Condensed Matter, 22(32), 323101.

[87] Perl, D., Welker, C., Schindler, T., Schröder, K., Marahiel, M. A., Jaenicke, R., & Schmid, F. X. (1998). Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. Nature structural biology, 5(3), 229-235.

[88] Okada, J., Okamoto, T., Mukaiyama, A., Tadokoro, T., You, D. J., Chon, H., ... & Kanaya, S. (2010). Evolution and thermodynamics of the slow unfolding of hyperstable monomeric proteins. BMC evolutionary biology, 10(1), 207.

[89] Cipolla, A., D'Amico, S., Barumandzadeh, R., Matagne, A., & Feller, G. (2011). Stepwise adaptations to low temperature as revealed by multiple mutants of psychrophilic α-amylase from Antarctic bacterium. Journal of Biological Chemistry, 286(44), 38348-38355.

[90] Baldwin, R. L. (2008). The search for folding intermediates and the mechanism of protein folding. Annu. Rev. Biophys., 37, 1-21.

[91] Schmid, F. X., Mayr, L. M., Mucke, M., & Schonbrunner, E. R. (1993). Prolyl isomerases: role in protein folding. In Advances in protein chemistry (Vol. 44, pp. 25-66). Academic Press.

[92] A Qasim, M., & Taha, M. (2013). Investigation of the mechanism of protein denaturation by guanidine hydrochloride-induced dissociation of inhibitor-protease complexes. Protein and peptide letters, 20(2), 187-191.

[93] Piette, F., Struvay, C., & Feller, G. (2011). The protein folding challenge in psychrophiles: facts and current issues. Environmental microbiology, 13(8), 1924-1933.

[94] Feller, G., & Gerday, C. (2003). Psychrophilic enzymes: hot topics in cold adaptation. Nature reviews microbiology, 1(3), 200.

[95] Privalov, P. L. (1990). Cold denaturation of protein. Critical reviews in biochemistry and molecular biology, 25(4), 281-306.

[96] Cipolla, A., Delbrassine, F., Da Lage, J. L., & Feller, G. (2012). Temperature adaptations in psychrophilic, mesophilic and thermophilic chloride-dependent alpha-amylases. Biochimie, 94(9), 1943-1950.

[97] Feller, G., d'Amic, D., & Gerday, C. (1999). Thermodynamic stability of a cold-active α-amylase from the Antarctic bacterium Alteromonas haloplanctis. Biochemistry, 38(14), 4613-4619.

[98] Linderstrøm-Lang, K. (1924). On the ionization of proteins. CR Trav. Lab. Carlsberg, 15(7), 1-29.

[99] Makki, F. (1996). Thermal inactivation kinetics of lysozyme and preservative effect beer (Doctoral dissertation, University of British Columbia).

[100] Cueto, M., Dorta, M. J., Munguía, O., & Llabrés, M. (2003). New approach to stability assessment of protein solution formulations by differential scanning calorimetry. International journal of pharmaceutics, 252(1-2), 159-166.

[101] Salvi, G., De Los Rios, P., & Vendruscolo, M. (2005). Effective interactions between chaotropic agents and proteins. Proteins: Structure, Function, and Bioinformatics, 61(3), 492-499.

[102] Hedwig, G. R., Lilley, T. H., & Linsdell, H. (1991). Calorimetric and volumetric studies of the interactions of some amides in water and in 6 mol dm–3 aqueous guanidinium chloride. Journal of the Chemical Society, Faraday Transactions, 87(18), 2975-2982.

[103] Rahimpour, F., & Pirdashti, M. (2007). The effect of guanidine hydrochloride on phase diagram of PEG-phosphate aqueous two-phase system. World Academy of Science, Engineering and Technology, 1(5), 29.

[104] Baldwin, R. L. (1996). How Hofmeister ion interactions affect protein stability. Biophysical journal, 71(4), 2056-2063.

[105] Harrison, R. (Ed.). (1993). Protein purification process engineering (Vol. 18). CRC Press.

[106] Mayr, L. M., & Schmid, F. X. (1993). Stabilization of a protein by guanidinium chloride. Biochemistry, 32(31), 7994-7998.

[107] Almarza, J., Rincon, L., Bahsas, A., & Brito, F. (2009). Molecular mechanism for the denaturation of proteins by urea. Biochemistry, 48(32), 7608-7613.

[108] D'Amico, S., Marx, J. C., Gerday, C., & Feller, G. (2003). Activity-stability relationships in extremophilic enzymes. Journal of Biological Chemistry, 278(10), 7891-7896.

[109] Smith, J. S., & Scholtz, J. M. (1996). Guanidine hydrochloride unfolding of peptide helices: separation of denaturant and salt effects. Biochemistry, 35(22), 7292-7297.

[110] Bai, Y., Sosnick, T. R., Mayne, L., & Englander, S. W. (1995). Protein folding intermediates: native-state hydrogen exchange. Science, 269(5221), 192-197.

[111] Kauzmann, W. (1959). Some factors in the interpretation of protein denaturation. In Advances in protein chemistry (Vol. 14, pp. 1-63). Academic Press.

[112] Javid, N., Vogtt, K., Krywka, C., Tolan, M., & Winter, R. (2007). Protein–protein interactions in complex cosolvent solutions. ChemPhysChem, 8(5), 679-689.

[113] Gullian-Klanian, M., & Sánchez-Solis, M. J. (2018). Growth kinetics of *Escherichia coli* O157: H7 on the epicarp of fresh vegetables and fruits. brazilian journal of microbiology, 49(1), 104-111.

[114] Shastry, M. C. R., Agashe, V. R., & Udgaonkar, J. B. (1994). Quantitative analysis of the kinetics of denaturation and renaturation of barstar in the folding transition zone. Protein Science, 3(9), 1409-1417.

[115] Brock, T. D., & Freeze, H. (1969). Thermus aquaticus gen. n. and sp. n., a nonsporulating extreme thermophile. Journal of bacteriology, 98(1), 289-297.

[116] Burton, R. E., Huang, G. S., Daugherty, M. A., Fullbright, P. W., & Oas, T. G. (1996). Microsecond protein folding through a compact transition state. Journal of molecular biology, 263(2), 311-322.

[117] Kragelund, B. B., Robinson, C. V., Knudsen, J., Dobson, C. M., & Poulsen, F. M. (1995). Folding of a four-helix bundle: studies of acyl-coenzyme A binding protein. Biochemistry, 34(21), 7217-7224.

[118] Weikl, T. R. (2008). Transition states in protein folding kinetics: Modeling Φ-values of small β-sheet proteins. Biophysical journal, 94(3), 929-937.

[119] Plaxco, K. W., & Dobson, C. M. (1996). Time-resolved biophysical methods in the study of protein folding. Current opinion in structural biology, 6(5), 630-636.

[120] Belkin, S., & Jannasch, H. W. (1985). A new extremely thermophilic, sulfur-reducing heterotrophic, marine bacterium. Archives of microbiology, 141(3), 181-186.

[121] He, F. (2011). Laemmli-SDS-PAGE. Bio-101: e80. DOI: 10.21769/BioProtoc.80.

[122] Burgess, R. R. (2009). Protein precipitation techniques. In Methods in enzymology (Vol. 463, pp. 331-342). Academic Press.

[123] ZILLIG, W., ZECHEL, K., & HALBWACHS, H. J. (1970). A new method of large scale preparation of highly purified DNA-dependent RNA-polymerase from E. coli. Hoppe-Seyler´ s Zeitschrift für physiologische Chemie, 351(1), 221-224.

[124] Yuan, H., Kurashina, K., de Bruijn, J. D., Li, Y., De Groot, K., & Zhang, X. (1999). A preliminary study on osteoinduction of two kinds of calcium phosphate ceramics. Biomaterials, 20(19), 1799-1806.

[125] Luo, J., Wu, C., Xu, T., & Wu, Y. (2011). Diffusion dialysis-concept, principle and applications. Journal of Membrane Science, 366(1-2), 1-16.

[126] Finkelstein, A. V., & Galzitskaya, O. V. (2004). Physics of protein folding. Physics of Life reviews, 1(1), 23-56.

[127] Siddiqui, K. S., & Cavicchioli, R. (2006). Cold-adapted enzymes. Annu. Rev. Biochem., 75, 403-433.

[128] Hatta, T. (1987). Le Châtelier principle. Institute of Social and Economic Research, Osaka University.

[129] Heremans, K. (1982). High pressure effects on proteins and other biomolecules. Annual review of biophysics and bioengineering, 11(1), 1-21.

[130] Mozhaev, V. V., Heremans, K., Frank, J., Masson, P., & Balny, C. (1996). High pressure effects on protein structure and function. Proteins: Structure, Function, and Bioinformatics, 24(1), 81-91.

[131] Chen, J. H., Chi, M. C., Lin, M. G., Lin, L. L., & Wang, T. F. (2015). Beneficial effect of sugar osmolytes on the refolding of guanidine hydrochloride-denatured trehalose-6-phosphate hydrolase from Bacillus licheniformis. BioMed research international, 2015.

[132] Sharma, G. S., & Singh, L. R. (2017). Polyols have unique ability to refold protein as compared to other osmolyte types. Biochemistry (Moscow), 82(4), 465-473.

[133] Boltzmann, Ludwig (1868). "Studies on the balance of living force between moving material points". Wiener Berichte. 58: 517–560.

[133] Ohgushi, M., & Wada, A. (1983). 'Molten-globule state': a compact form of globular proteins with mobile side-chains. FEBS letters, 164(1), 21-24.

[134] Sakurai, K., Oobatake, M., & Goto, Y. (2001). Salt-dependent monomer–dimer equilibrium of bovine β-lactoglobulin at pH 3. Protein Science, 10(11), 2325-2335.

[135] Metpally, R. P. R., & Reddy, B. V. B. (2009). Comparative proteome analysis of psychrophilic versus mesophilic bacterial species: Insights into the molecular basis of cold adaptation of proteins. BMC genomics, 10(1), 11.

[136] Matthews, C. R. (1987). [26] Effect of point mutations of the folding of globular proteins. In Methods in enzymology (Vol. 154, pp. 498-511). Academic Press.

[137] Kiefhaber, T. (1995). Kinetic traps in lysozyme folding. Proceedings of the National Academy of Sciences, 92(20), 9029-9033.

[138] Thirumalai, D., & Hyeon, C. (2005). RNA and protein folding: common themes and variations. Biochemistry, 44(13), 4957-4970.

[139] Gruber, T., & Balbach, J. (2015). Protein folding mechanism of the dimeric amphiphysinII/Bin1 N-BAR domain. PloS one, 10(9), e0136922.

[140] Fersht, A. R. (1995). Characterizing transition states in protein folding: an essential step in the puzzle. Current opinion in structural biology, 5(1), 79-84.

[141] Dobson, C. M., & Karplus, M. (1999). The fundamentals of protein folding: bringing together theory and experiment. Current opinion in structural biology, 9(1), 92-101.

[142] Matouschek, A., Otzen, D. E., Itzhaki, L. S., Jackson, S. E., & Fersht, A. R. (1995). Movement of the position of the transition state in protein folding. Biochemistry, 34(41), 13656-13662.

[143] Bieger, B., Essen, L. O., & Oesterhelt, D. (2003). Crystal structure of halophilic dodecin: a novel, dodecameric flavin binding protein from Halobacterium salinarum. Structure, 11(4), 375-385.

[144] PUNDAK, S., ALONI, H., & EISENBERG, H. (1981). Structure and activity of malate dehydrogenase from the extreme halophilic bacteria of the Dead Sea: 2. Inactivation, dissociation and unfolding at NaCl concentrations below 2 M. Salt, salt concentration and temperature dependence of enzyme stability. European journal of biochemistry, 118(3), 471-477.

[145] Bonneté, F., Madern, D., & Zaccai, G. (1994). Stability against denaturation mechanisms in halophilic malate dehydrogenase" adapt" to solvent conditions. Journal of molecular biology, 244(4), 436-447.

[146] Frolow, F., Harell, M., Sussman, J. L., Mevarech, M., & Shoham, M. (1996). Insights into protein adaptation to a saturated salt environment from the crystal structure of a halophilic 2Fe-2S ferredoxin. Nature structural biology, 3(5), 452-458.

[147] Robinson, D. R., & Jencks, W. P. (1965). The effect of concentrated salt solutions on the activity coefficient of acetyltetraglycine ethyl ester. Journal of the American Chemical Society, 87(11), 2470-2479.

[148] Nandi, P. K., & Robinson, D. R. (1972). Effects of salts on the free energy of the peptide group. Journal of the American Chemical Society, 94(4), 1299-1308.

[149] Von Hippel, P. H., Peticolas, V., Schack, L., & Karlson, L. (1973). Model studies on the effects of neutral salts on the conformational stability of biological macromolecules. I. Ion binding to polyacrylamide and polystyrene columns. Biochemistry, 12(7), 1256-1264.

[150] PL, P. (1990). Cold denaturation of protein. Crit. Rev. Biochem. Mol. Biol., 25, 281-305.

[151] Micsonai, A., Wien, F., Bulyáki, É., Kun, J., Moussong, É., Lee, Y. H., ... & Kardos, J. (2018). BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. Nucleic acids research, 46(W1), W315-W322.

[152] Thompson, G. S. (2005). The role of perturbation on biochemical systems: salt, pH, and microgravity effects on protein-ligand and small molecule interactions.

[153] Aghajari, N., Feller, G., Gerday, C., & Haser, R. (1998). Structures of the psychrophilic Alteromonas haloplanctis α-amylase give insights into cold adaptation at a molecular level. Structure, 6(12), 1503-1516.

[154] Wedemeyer, W. J., Welker, E., & Scheraga, H. A. (2002). Proline cis− trans isomerization and protein folding. Biochemistry, 41(50), 14637-14644.

[155] Bhuyan, A. K. (2002). Protein stabilization by urea and guanidine hydrochloride. Biochemistry, 41(45), 13386-13394.

[156] Zarrine-Afsar, A., Mittermaier, A., Kay, L. E., & Davidson, A. R. (2006). Protein stabilization by specific binding of guanidinium to a functional arginine-binding surface on an SH3 domain. Protein science, 15(1), 162-170.

[157] Pike, A. C., & Acharya, K. R. (1994). A structural basis for the interaction of urea with lysozyme. Protein Science, 3(4), 706-710.

[158] Dunbar, J., Yennawar, H. P., Banerjee, S., Luo, J., & Farber, G. K. (1997). The effect of denaturants on protein structure. Protein Science, 6(8), 1727-1733.

[159] Timasheff, S. N. (1992). Water as ligand: preferential binding and exclusion of denaturants in protein unfolding. Biochemistry, 31(41), 9857-9864.

[160] Breslow, R., & Guo, T. (1990). Surface tension measurements show that chaotropic salting-in denaturants are not just water-structure breakers. Proceedings of the National Academy of Sciences, 87(1), 167-169.

[161] Hagihara, Y., Aimoto, S., Fink, A. L., & Goto, Y. (1993). Guanidine hydrochloride-induced folding of proteins.

[162] Liu, C. C., & LiCata, V. J. (2014). The stability of Taq DNA polymerase results from a reduced entropic folding penalty; identification of other thermophilic proteins with similar folding thermodynamics. Proteins: Structure, Function, and Bioinformatics, 82(5), 785-793.

[163] Huston, A. L., Haeggström, J. Z., & Feller, G. (2008). Cold adaptation of enzymes: Structural, kinetic and microcalorimetric characterizations of an aminopeptidase from the Arctic psychrophile Colwellia psychrerythraea and of human leukotriene A4 hydrolase. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1784(11), 1865-1872.

[164] Michaux, C., Massant, J., Kerff, F., Frère, J. M., Docquier, J. D., Vandenberghe, I., ... & Van Beeumen, J. (2008). Crystal structure of a cold-adapted class C β-lactamase. The FEBS journal, 275(8), 1687-1697.

[165] Zhao, J. S., Deng, Y., Manno, D., & Hawari, J. (2010). Shewanella spp. genomic evolution for a cold marine lifestyle and in-situ explosive biodegradation. PloS one, 5(2), e9109.

[166] Hedges, J. B., Vahidi, S., Yue, X., & Konermann, L. (2013). Effects of ammonium bicarbonate on the electrospray mass spectra of proteins: evidence for bubble-induced unfolding. Analytical chemistry, 85(13), 6469-6476.

[167] Georlette, D., Blaise, V., Collins, T., D'Amico, S., Gratia, E., Hoyoux, A., ... & Gerday, C. (2004). Some like it cold: biocatalysis at low temperatures. FEMS microbiology reviews, 28(1), 25-42.

[168] Klein, W., Weber, M. H., & Marahiel, M. A. (1999). Cold shock response of Bacillus subtilis: isoleucine-dependent switch in the fatty acid branching pattern for membrane adaptation to low temperatures. Journal of bacteriology, 181(17), 5341-5349.

[169] Hang, Y., & Cremer, P. S. (2006). Interactions between macromolecules and ions: the Hofmeister series. Current opinion in chemical biology, 10(6), 658-663.

[170] Siddiqui, K. S., Poljak, A., Guilhaus, M., De Francisci, D., Curmi, P. M., Feller, G., ... & Cavicchioli, R. (2006). Role of lysine versus arginine in enzyme cold-adaptation: Modifying lysine to homo-arginine stabilizes the cold-adapted α-amylase from Pseudoalteramonas haloplanktis. Proteins: Structure, Function, and Bioinformatics, 64(2), 486-501.

[171] Pieper, U., Kapadia, G., Mevarech, M., & Herzberg, O. (1998). Structural features of halophilicity derived from the crystal structure of dihydrofolate reductase from the Dead Sea halophilic archaeon, Haloferax volcanii. Structure, 6(1), 75-88.

[172] Oren, A. (2002). Adaptation of halophilic archaea to life at high salt concentrations. In Salinity: Environment-Plants-Molecules (pp. 81-96). Springer, Dordrecht.

[173] Fields, P. A., & Somero, G. N. (1998). Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. Proceedings of the National Academy of Sciences, 95(19), 11476-11481.

[174] Detrich III, H. W. (1997). Microtubule assembly in cold-adapted organisms: functional properties and structural adaptations of tubulins from antarctic fishes. Comparative Biochemistry and Physiology Part A: Physiology, 118(3), 501-513.

[175] Gianese, G., Argos, P., & Pascarella, S. (2001). Structural adaptation of enzymes to low temperatures. Protein engineering, 14(3), 141-148.

[176] Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., ... & Ausio, J. (2001). Intrinsically disordered protein. Journal of molecular graphics and modelling, 19(1), 26-59.

[177] Uversky, V. N., Gillespie, J. R., & Fink, A. L. (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions?. Proteins: structure, function, and bioinformatics, 41(3), 415-427.

[178] Wright, P. E., & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. Journal of molecular biology, 293(2), 321-331.

[179] Tompa, P., & Fersht, A. (2009). Structure and function of intrinsically disordered proteins. Chapman and Hall/CRC.

[180] Burra, P. V., Kalmar, L., & Tompa, P. (2010). Reduction in structural disorder and functional complexity in the thermal adaptation of prokaryotes. PloS one, 5(8), e12069.

[181] Vicedo, E., Schlessinger, A., & Rost, B. (2015). Environmental pressure may change the composition protein disorder in prokaryotes. PloS one, 10(8), e0133990.

[182] Uversky, V. N. (2013). A decade and a half of protein intrinsic disorder: biology still waits for physics. Protein Science, 22(6), 693-724.

[183] Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J., & Dunker, A. K. (2001). Sequence complexity of disordered protein. Proteins: Structure, Function, and Bioinformatics, 42(1), 38-48.

[184] Dosztányi, Z. (2018). Prediction of protein disorder based on IUPred. Protein Science, 27(1), 331-340.

[185] Lobanov, M. Y., & Galzitskaya, O. V. (2011). The Ising model for prediction of disordered residues from protein sequence alone. Physical biology, 8(3), 035004.

[186] Cheng, J., Sweredoski, M. J., & Baldi, P. (2005). Accurate prediction of protein disordered regions by mining protein structure data. Data mining and knowledge discovery, 11(3), 213-222.

[187] Radivojac, P., Iakoucheva, L. M., Oldfield, C. J., Obradovic, Z., Uversky, V. N., & Dunker, A. K. (2007). Intrinsic disorder and functional proteomics. Biophysical journal, 92(5), 1439-1456.

[188] Campen, A., Williams, R. M., Brown, C. J., Meng, J., Uversky, V. N., & Dunker, A. K. (2008). TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. Protein and peptide letters, 15(9), 956-963.

[189] Tompa, P. (2002). Intrinsically unstructured proteins. Trends in biochemical sciences, 27(10), 527-533.

[190] Kalthoff, C. (2003). A novel strategy for the purification of recombinantly expressed unstructured protein domains. Journal of Chromatography B, 786(1-2), 247-254.

[191] Rao, L., Zhao, X., Pan, F., Li, Y., Xue, Y., Ma, Y., & Lu, J. R. (2009). Solution behavior and activity of a halophilic esterase under high salt concentration. PloS one, 4(9).

[192] Tantos, A., Friedrich, P., & Tompa, P. (2009). Cold stability of intrinsically disordered proteins. FEBS letters, 583(2), 465-469.

[193] Uversky, V. N. (2013). Unusual biophysics of intrinsically disordered proteins. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1834(5), 932-951.

[194] Greenfield, N. J. (2006). Analysis of the kinetics of folding of proteins and peptides using circular dichroism. Nature protocols, 1(6), 2891.

[195] Karan, R., & Khare, S. K. (2011). Stability of haloalkaliphilic Geomicrobium sp. protease modulated by salt. Biochemistry (Moscow), 76(6), 686.

[196] Li, X., Romero, P., Rani, M., Dunker, A. K., & Obradovic, Z. (1999). Predicting protein disorder for N-, C-and internal regions. Genome Informatics, 10, 30-40.

[197] Dosztányi, Z., Csizmok, V., Tompa, P., & Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics, 21(16), 3433-3434.

[198] Larsen, R. W., Yang, J., Hou, S., Helms, M. K., Jameson, D. M., & Alam, M. (1999). Spectroscopic characterization of two soluble transducers from the Archaeon Halobacterium salinarum. Journal of protein chemistry, 18(3), 269-275.

[199] Peng, K., Vucetic, S., Radivojac, P., Brown, C. J., Dunker, A. K., & Obradovic, Z. (2005). Optimizing long intrinsic disorder predictors with protein evolutionary information. Journal of bioinformatics and computational biology, 3(01), 35-60.

[200] Uversky, V. N. (2009). The mysterious unfoldome: structureless, underappreciated, yet vital part of any given proteome. BioMed Research International, 2010.

[201] Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., & Jones, D. T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. Journal of molecular biology, 337(3), 635-645.

[202] Xue, B., Dunker, A. K., & Uversky, V. N. (2012). Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. Journal of Biomolecular Structure and Dynamics, 30(2), 137-149.

[203] Datta, K., Wowor, A. J., Richard, A. J., & LiCata, V. J. (2006). Temperature dependence and thermodynamics of Klenow polymerase binding to primed-template DNA. Biophysical journal, 90(5), 1739-1751.

[204] Schoeffler, A. J., Joubert, A. M., Peng, F., Khan, F., Liu, C. C., & LiCata, V. J. (2004). Extreme free energy of stabilization of Taq DNA polymerase. Proteins: Structure, Function, and Bioinformatics, 54(4), 616-621.

[205] Karantzeni, I., Carmen, R. U. I. Z., Chin-Chi, L. I. U., & LiCATA, V. J. (2003). Comparative thermal denaturation of Thermus aquaticus and Escherichia coli type 1 DNA polymerases. Biochemical Journal, 374(3), 785-792..

[206] Lanyi, J. K. (1974). Salt-dependent properties of proteins from extremely halophilic bacteria. Bacteriological reviews, 38(3), 272.

[207] PUNDAK, S., & EISENBERG, H. (1981). Structure and activity of malate dehydrogenase from the extreme halophilic bacteria of the Dead Sea: 1. Conformation and interaction with water and salt between 5 M and 1 M NaCl concentration. European journal of biochemistry, 118(3), 463-470.

[208] Kushner, D. J. (1964). Lysis and dissolution of cells and envelopes of an extremely halophilic bacterium. Journal of bacteriology, 87(5), 1147-1156.

[209] Oren, A. (2002). Molecular ecology of extremely halophilic Archaea and Bacteria. FEMS Microbiology Ecology, 39(1), 1-7.

[210] Niesen, F. H., Berglund, H., & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature protocols, 2(9), 2212.

[211] Chu, B. (1970). Laser light scattering. Annual review of physical chemistry, 21(1), 145-174.

[212] Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Annalen der physik, 322(8), 549-560.

[213] Kunin, C. M., Hua, T. H., Van Arsdale White, L., & Villarejo, M. (1992). Growth of Escherichia coli in human urine: role of salt tolerance and accumulation of glycine betaine. Journal of Infectious Diseases, 166(6), 1311-1315.

[214] Brumm, P. J., Monsma, S., Keough, B., Jasinovica, S., Ferguson, E., Schoenfeld, T., ... & Mead, D. A. (2015). Complete genome sequence of Thermus aquaticus Y51MC23. PloS one, 10(10), e0138674.

[215] Kishore, D., Kundu, S., & Kayastha, A. M. (2012). Thermal, chemical and pH induced denaturation of a multimeric β-galactosidase reveals multiple unfolding pathways. PLoS One, 7(11), e50380.

[216] Quillaguamán, J., Guzmán, H., Van-Thuoc, D., & Hatti-Kaul, R. (2010). Synthesis and production of polyhydroxyalkanoates by halophiles: current potential and future prospects. Applied Microbiology and Biotechnology, 85(6), 1687-1696.

[217] Ollivier, B., Caumette, P., Garcia, J. L., & Mah, R. A. (1994). Anaerobic bacteria from hypersaline environments. Microbiology and Molecular Biology Reviews, 58(1), 27-38.

[218] Oren, A. (2002). Diversity of halophilic microorganisms: environments, phylogeny, physiology, and applications. Journal of Industrial Microbiology and Biotechnology, 28(1), 56-63.

[219] Oren, A. (2008). Microbial life at high salt concentrations: phylogenetic and metabolic diversity. Saline systems, 4(1), 2.

[220] Böhm, G., & Jaenicke, R. (1994). A structure-based model for the halphilic adaptiation of dihydrofolate reductase from Halobacterium volcanii. Protein Engineering, Design and Selection, 7(2), 213-220.

[221] Rao, J. M., & Argos, P. (1981). Structural stability of halophilic proteins. Biochemistry, 20(23), 6536-6543.

[222] Vincent, J. F. (2012). Structural biomaterials. Princeton University Press.

[223] Zaccai, G., Cendrin, F., Haik, Y., Borochov, N., & Eisenberg, H. (1989). Stabilization of halophilic malate dehydrogenase. Journal of molecular biology, 208(3), 491-500.

[224] Chen, G. Q. (2009). A microbial polyhydroxyalkanoates (PHA) based bio-and materials industry. Chemical Society Reviews, 38(8), 2434-2446.

[225] Chen, G. Q. (2012). New challenges and opportunities for industrial biotechnology. Microbial cell factories, 11(1), 111.

[226] Kolp, S., Pietsch, M., Galinski, E. A., & Gütschow, M. (2006). Compatible solutes as protectants for zymogens against proteolysis. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1764(7), 1234-1242.

[227] Setati, M. E. (2010). Diversity and industrial potential of hydrolaseproducing halophilic/halotolerant eubacteria. African Journal of Biotechnology, 9(11), 1555-1560.

[228] Rodrigues, L., Banat, I. M., Teixeira, J., & Oliveira, R. (2006). Biosurfactants: potential applications in medicine. Journal of antimicrobial chemotherapy, 57(4), 609-618.

[229] Oren, A., Larimer, F., Richardson, P., Lapidus, A., & Csonka, L. N. (2005). How to be moderately halophilic with broad salt tolerance: clues from the genome of Chromohalobacter salexigens. Extremophiles, 9(4), 275-279.

[230] Karan, R., Capes, M. D., & DasSarma, S. (2012). Function and biotechnology of extremophilic enzymes in low water activity. Aquatic Biosystems, 8(1), 4.

[231] Oren, A. (2013). Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. Frontiers in microbiology, 4, 315.

[232] Kozlowski, L. P. (2016). Proteome-pI: proteome isoelectric point database. Nucleic acids research, 45(D1), D1112-D1116.

[233] Fedyukina, D. V., Jennaro, T. S., & Cavagnero, S. (2014). Charge segregation and low hydrophobicity are key features of ribosomal proteins from different organisms. Journal of Biological Chemistry, 289(10), 6740-6750.

[234] Baxter, R. M. (1959). An interpretation of the effects of salts on the lactic dehydrogenase of Halobacterium salinarium. Canadian Journal of Microbiology, 5(1), 47-57.

[235] Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. Journal of molecular biology, 157(1), 105-132.

[236] Wolfram Research, Inc., SystemModeler, Version 12.0, Champaign, IL (2019).

[237] "All About Sea Ice." National Snow and Ice Data Center. Accessed 1 February 2016.

[238] Ewert, M., & Deming, J. W. (2013). Sea ice microorganisms: environmental constraints and extracellular responses. Biology, 2(2), 603-628.

[239] Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; Protein Identification and Analysis Tools on the ExPASy Server;

[240] Monera, O. D., Kay, C. M., & Hodges, R. S. (1994). Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. Protein Science, 3(11), 1984-1991.

[241] Rashid, F., Sharma, S., & Bano, B. (2005). Comparison of guanidine hydrochloride (GdmHCl) and urea denaturation on inactivation and unfolding of human placental cystatin (HPC). The protein journal, 24(5), 283-292.

[242] Lim, W. K., Rösgen, J., & Englander, S. W. (2009). Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. Proceedings of the National Academy of Sciences, 106(8), 2595-2600.

[243] Povarova, O. I., Kuznetsova, I. M., & Turoverov, K. K. (2010). Differences in the pathways of proteins unfolding induced by urea and guanidine hydrochloride: molten globule state and aggregates. PLoS One, 5(11), e15035.

[244] Sinha, R., & Khare, S. K. (2014). Protective role of salt in catalysis and maintaining structure of halophilic proteins against denaturation. Frontiers in microbiology, 5, 165.

[245] Georlette, D., Damien, B., Blaise, V., Depiereux, E., Uversky, V. N., Gerday, C., & Feller, G. (2003). Structural and functional adaptations to extreme temperatures in psychrophilic, mesophilic, and thermophilic DNA ligases. Journal of Biological Chemistry, 278(39), 37015-37023.

[246] Kumar, S., & Nussinov, R. (2004). Experiment-guided thermodynamic simulations on reversible two-state proteins: implications for protein thermostability. Biophysical chemistry, 111(3), 235-246.

[247] Rees, D. C., & Robertson, A. D. (2001). Some thermodynamic implications for the thermostability of proteins. Protein Science, 10(6), 1187-1194.

[248] Kumar, S., Tsai, C. J., & Nussinov, R. (2002). Maximal stabilities of reversible two-state proteins. Biochemistry, 41(17), 5359-5374.

[249] Knapp, S., Karshikoff, A., Berndt, K. D., Christova, P., Atanasov, B., & Ladenstein, R. (1996). Thermal unfolding of the DNA-binding protein Sso7d from the hyperthermophile sulfolobus solfataricus. Journal of molecular biology, 264(5), 1132-1144.

[250] McCrary, B. S., Edmondson, S. P., & Shriver, J. W. (1996). Hyperthermophile protein folding thermodynamics: differential scanning calorimetry and chemical denaturation of Sac7d. Journal of molecular biology, 264(4), 784-805.

[251] Robic, S., Guzman-Casado, M., Sanchez-Ruiz, J. M., & Marqusee, S. (2003). Role of residual structure in the unfolded state of a thermophilic protein. Proceedings of the National Academy of Sciences, 100(20), 11345-11349.

[252] Makhatadze, G. I., & Privalov, P. L. (1995). Energetics of protein structure. In Advances in protein chemistry (Vol. 47, pp. 307-425). Academic Press.

[253] Leberman, R., & Soper, A. K. (1995). Effect of high salt concentrations on water structure. Nature, 378(6555), 364.

[254] Robinson, C. R., & Sligar, S. G. (1993). Electrostatic stabilization in four-helix bundle proteins. Protein Science, 2(5), 826-837.

[255] Kenar, K. T., Garcia-Moreno, B., & Freire, E. (1995). A calorimetric characterization of the salt dependence of the stability of the GCN4 leucine zipper. Protein Science, 4(9), 1934-1938.

[256] Roman, E. A., & González Flecha, F. L. (2014). Kinetics and thermodynamics of membrane protein folding. Biomolecules, 4(1), 354-373.

[257] Mason, P. E., Neilson, G. W., Dempsey, C. E., Barnes, A. C., & Cruickshank, J. M. (2003). The hydration structure of guanidinium and thiocyanate ions: implications for protein stability in aqueous solution. Proceedings of the National Academy of Sciences, 100(8), 4557-4561.

[258] Vladilo, G., & Hassanali, A. (2018). Hydrogen bonds and life in the universe. *Life*, *8*(1), 1.

[259] Mevarech, M., Eisenberg, H., & Neumann, E. (1977). Malate dehydrogenase isolated from extremely halophilic bacteria of the Dead Sea. 1. Purification and molecular characterization. Biochemistry, 16(17), 3781-3785.

[260] Gafni, A., & Werber, M. M. (1979). Ferredoxin from Halobacterium of the Dead Sea. Structural properties revealed by fluorescence techniques. Archives of biochemistry and biophysics, 196(1), 363-370.

[261] Müller-Santos, M., de Souza, E. M., Pedrosa, F. D. O., Mitchell, D. A., Longhi, S., Carrière, F., ... & Krieger, N. (2009). First evidence for the salt-dependent folding and activity of an esterase from the halophilic archaea Haloarcula marismortui. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 1791(8), 719-729.

[262] Elcock, A. H., & McCammon, J. A. (1998). Electrostatic contributions to the stability of halophilic proteins. Journal of molecular biology, 280(4), 731-748.

[263] Moraitis, M. I., Xu, H., & Matthews, K. S. (2001). Ion concentration and temperature dependence of DNA binding: comparison of PurR and LacI repressor proteins. Biochemistry, 40(27), 8109-8117.

[264] Street, T. O., Bolen, D. W., & Rose, G. D. (2006). A molecular mechanism for osmolyte-induced protein stability. Proceedings of the National Academy of Sciences, 103(38), 13997-14002.

[265] Bolen, D. W., & Baskakov, I. V. (2001). The osmophobic effect: natural selection of a thermodynamic force in protein folding. Journal of molecular biology, 310(5), 955-963.

[266] Cavicchioli, R., Siddiqui, K. S., Andrews, D., & Sowers, K. R. (2002). Low-temperature extremophiles and their applications. Current Opinion in Biotechnology, 13(3), 253-261.

[267] Perunov, N., & England, J. L. (2014). Quantitative theory of hydrophobic effect as a driving force of protein structure. Protein Science, 23(4), 387-399.

[268] Arunan, E., Desiraju, G. R., Klein, R. A., Sadlej, J., Scheiner, S., Alkorta, I., ... & Kjaergaard, H. G. (2011). Definition of the hydrogen bond (IUPAC Recommendations 2011). Pure and applied chemistry, 83(8), 1637-1641.

[269] Warshel, A., & Aqvist, J. (1991). Electrostatic energy and macromolecular function. Annual review of biophysics and biophysical chemistry, 20(1), 267-298.

[270] Arakawa, T., & Timasheff, S. N. (1984). The mechanism of action of Na glutamate, lysine HCl, and piperazine-N, N'-bis (2-ethanesulfonic acid) in the stabilization of tubulin and microtubule formation. Journal of Biological Chemistry, 259(8), 4979-4986.

[271] Saum, S. H., & Müller, V. (2008). Regulation of osmoadaptation in the moderate halophile Halobacillus halophilus: chloride, glutamate and switching osmolyte strategies. Saline systems, 4(1), 4.

[272] Cheng, X., Guinn, E. J., Buechel, E., Wong, R., Sengupta, R., Shkel, I. A., & Record Jr, M. T. (2016). Basis of protein stabilization by K glutamate: unfavorable interactions with carbon, oxygen groups. Biophysical journal, 111(9), 1854-1865.

[273] Leirmo, S., Harrison, C., Cayley, D. S., Burgess, R. R., & Record Jr, M. T. (1987). Replacement of potassium chloride by potassium glutamate dramatically enhances protein-DNA interactions in vitro. Biochemistry, 26(8), 2095-2101.

[274] Csonka, L. N., Ikeda, T. P., Fletcher, S. A., & Kustu, S. (1994). The accumulation of glutamate is necessary for optimal growth of Salmonella typhimurium in media of high osmolality but not induction of the proU operon. Journal of Bacteriology, 176(20), 6324-6333.

[275] Datta, K., & LiCata, V. J. (2003). Salt Dependence of DNA binding by Thermus aquaticusand Escherichia coli DNA Polymerases. Journal of Biological Chemistry, 278(8), 5694-5701.

[276] Karantzeni, I., Ruiz, C., Liu, C. C., & LiCATA, V. J. (2003). Comparative thermal denaturation of Thermus aquaticus and Escherichia coli type 1 DNA polymerases. Biochemical Journal, 374(3), 785-792.

[277] Abouelkhair, M. A., Bemis, D. A., Giannone, R. J., Frank, L. A., & Kania, S. A. (2018). Characterization of a leukocidin identified in Staphylococcus pseudintermedius. PloS one, 13(9), e0204450.

[278] Ghisaidoobe, A., & Chung, S. (2014). Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Förster resonance energy transfer techniques. International journal of molecular sciences, 15(12), 22518-22538.

[279] Arakawa, T., & Tokunaga, M. (2004). Electrostatic and hydrophobic interactions play a major role in the stability and refolding of halophilic proteins. Protein and peptide letters, 11(2), 125-132.

[280] Stern, O., & Volmer, M. (1919). Über die abklingungszeit der fluoreszenz.

[281] Kumar, R., Prabhu, N. P., Yadaiah, M., & Bhuyan, A. K. (2004). Protein stiffening and entropic stabilization in the subdenaturing limit of guanidine hydrochloride. Biophysical journal, 87(4), 2656-2662.

[282] Zheng, P., Cao, Y., Bu, T., Straus, S. K., & Li, H. (2011). Single molecule force spectroscopy reveals that electrostatic interactions affect the mechanical stability of proteins. Biophysical journal, 100(6), 1534-1541.

[283] Möller, M., & Denicola, A. (2002). Protein tryptophan accessibility studied by fluorescence quenching. Biochemistry and Molecular Biology Education, 30(3), 175-178.

[284] Thomas, D. N., & Dieckmann, G. S. (2002). Antarctic sea ice--a habitat for extremophiles. Science, 295(5555), 641-644.

[285] Tadeo, X., López-Méndez, B., Castaño, D., Trigueros, T., & Millet, O. (2009). Protein stabilization and the Hofmeister effect: the role of hydrophobic solvation. Biophysical journal, 97(9), 2595-2603.

[286] Arakawa, T., & Timasheff, S. N. (1985). The stabilization of proteins by osmolytes. Biophysical journal, 47(3), 411-414.

[287] Loladze, V. V., & Makhatadze, G. I. (2002). Removal of surface charge-charge interactions from ubiquitin leaves the protein folded and very stable. Protein Science, 11(1), 174-177.

[288] Safe, D. N. A. Molecular Imager® Gel Doc™ XR+ and ChemiDoc™ XRS+ Systems.

[289] IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997).

[290] Nayek, A., Gupta, P. S. S., Banerjee, S., Mondal, B., & Bandyopadhyay, A. K. (2014). Salt-bridge energetics in halophilic proteins. Plos one, 9(4).

[291] Akke, M., & Forsén, S. (1990). Protein stability and electrostatic interactions between solvent exposed charged side chains. Proteins: Structure, Function, and Bioinformatics, 8(1), 23-29.

[292] Tadeo, X., López-Méndez, B., Trigueros, T., Laín, A., Castaño, D., & Millet, O. (2009). Structural basis for the aminoacid composition of proteins from halophilic archea. PLoS biology, 7(12).

[293] Gómez F. (2011) Extreme Environment. In: Gargaud M. et al. (eds) Encyclopedia of Astrobiology. Springer, Berlin, Heidelberg

[294] Moyer, C. L., & Morita, R. Y. (2001). Psychrophiles and psychrotrophs. e LS.

[295] Feller, G. (2013). Psychrophilic enzymes: from folding to function and biotechnology. Scientifica, 2013.

[296] Ferguson, N., & Fersht, A. R. (2003). Early events in protein folding. Current opinion in structural biology, 13(1), 75-81.

[297] Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J. A., & Rapp, J. Z. (2015). Microbial ecology of the cryosphere: sea ice and glacial habitats. Nature Reviews Microbiology, 13(11), 677-690.

[298] Grimsley, G. R., Shaw, K. L., Fee, L. R., Alston, R. W., Huyghues-Despointes, B. M., Thurlkill, R. L., ... & Pace, C. N. (1999). Increasing protein stability by altering long-range coulombic interactions. Protein Science, 8(9), 1843-1849.

[299] Ogasahara, K., Lapshina, E. A., Sakai, M., Izu, Y., Tsunasawa, S., Kato, I., & Yutani, K. (1998). Electrostatic Stabilization in Methionine Aminopeptidase from Hyperthermophile Pyrococcus f uriosus. Biochemistry, 37(17), 5939-5946.

[300] Stibal, M., Hasan, F., Wadham, J. L., Sharp, M. J., & Anesio, A. M. (2012). Prokaryotic diversity in sediments beneath two polar glaciers with contrasting organic carbon substrates. Extremophiles, 16(2), 255-265.

[301] Krembs, C., Eicken, H., & Deming, J. W. (2011). Exopolymer alteration of physical properties of sea ice and implications for ice habitability and biogeochemistry in a warmer Arctic. Proceedings of the National Academy of Sciences, 108(9), 3653-3658.

[302] Madern, D., & Zaccai, G. (1997). Stabilisation of halophilic malate dehydrogenase from Haloarcula marismortui by divalent cations: effects of temperature, water isotope, cofactor and pH. European journal of biochemistry, 249(2), 607-611.

[303] Schonbrun, J., & Dill, K. A. (2003). Fast protein folding kinetics. Proceedings of the National Academy of Sciences, 100(22), 12678-12682.

[304] Solomons, T. W. Graham and Fryhle, Craig B.(2004). Organic Chemistry (8th ed.). John Wiley & Sons, Inc.

[305] Federley, R. G., & Romano, L. J. (2010). DNA polymerase: structural homology, conformational dynamics, and the effects of carcinogenic DNA adducts. Journal of nucleic acids, 2010.

[306] Auman, A. J., Breezee, J. L., Gosink, J. J., Kämpfer, P., & Staley, J. T. (2006). Psychromonas ingrahamii sp. nov., a novel gas vacuolate, psychrophilic bacterium isolated from Arctic polar sea ice. International Journal of Systematic and Evolutionary Microbiology, 56(5), 1001-1007.

[307] Savidor, A., Barzilay, R., Elinger, D., Yarden, Y., Lindzen, M., Gabashvili, A., ... & Levin, Y. (2017). Database-Independent Protein Sequencing (DiPS) Enables Full-Length de Novo Protein and Antibody Sequence Determination. Molecular & Cellular Proteomics, 16(6), 1151-1161.

[308] Lothrop, A. P., Torres, M. P., & Fuchs, S. M. (2013). Deciphering post-translational modification codes. FEBS letters, 587(8), 1247-1257.

[309] Maruthamuthu, M., & Selvakumar, G. (1995, February). Selective quenching of tryptophanyl fluorescence in bovine serum albumin by the iodide ion. In Proceedings of the Indian Academy of Sciences-Chemical Sciences (Vol. 107, No. 1, pp. 79-86). Springer India.

[310] Tannous, E., Yokoyama, K., You, D. J., Koga, Y., & Kanaya, S. (2012). A dual role of divalent metal ions in catalysis and folding of RNase H1 from extreme halophilic archaeon Halobacterium sp. NRC-1. *FEBS open bio*, *2*, 345-352.

[311] Frishman, D., & Argos, P. (1995). Knowledge-based protein secondary structure assignment. Proteins: Structure, Function, and Bioinformatics, 23(4), 566-579.

[312] Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., ... & Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic acids research*, *42*(W1), W252-W258.

[313] Ebel, C., Costenaro, L., Pascu, M., Faou, P., Kernel, B., Proust-De Martin, F., & Zaccai, G. (2002). Solvent interactions of halophilic malate dehydrogenase. *Biochemistry*, *41*(44), 13234-13244.

[314] Alimenti, C., Vallesi, A., Pedrini, B., Wüthrich, K., & Luporini, P. (2009). Molecular cold adaptation: Comparative analysis of two homologous families of psychrophilic and mesophilic signal proteins of the protozoan ciliate, Euplotes. IUBMB life, 61(8), 838-845.

[315] Médigue, C., Krin, E., Pascal, G., Barbe, V., Bernsel, A., Bertin, P. N., ... & Fang, G. (2005). Coping with cold: the genome of the versatile marine Antarctica bacterium Pseudoalteromonas haloplanktis TAC125. Genome research, 15(10), 1325-1335.

[316] Cupo, P., El-Deiry, W., Whitney, P. L., & Awad, W. M. (1980). Stabilization of proteins by guanidination. Journal of Biological Chemistry, 255(22), 10828-10833.

[317] McGuffin, L. J., Bryson, K., & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics*, *16*(4), 404-405.

[318] Ashok Kumar, T. (2013). CFSSP: Chou and Fasman Secondary Structure Prediction server. WIDE SPECTRUM: Research Journal. 1(9):15-19.

[319] Garnier, J., Gibrat, J. F., & Robson, B. (1996). GOR method for predicting protein secondary structure from amino acid sequence. In Methods in enzymology (Vol. 266, pp. 540-553). Academic Press.

[320] Yachdav, G., Kloppmann, E., Kajan, L., Hecht, M., Goldberg, T., Hamp, T., ... & Richter, L. (2014). PredictProtein—an open resource for online prediction of protein structural and functional features. Nucleic acids research, 42(W1), W337-W343.

[321] Cheng, J., Randall, A. Z., Sweredoski, M. J., & Baldi, P. (2005). SCRATCH: a protein structure and structural feature prediction server. Nucleic acids research, 33(suppl_2), W72-W76.

[322] Wang, S., Li, W., Liu, S., & Xu, J. (2016). RaptorX-Property: a web server for protein structure property prediction. Nucleic acids research, 44(W1), W430-W435.

[323] Heinig, M., & Frishman, D. (2004). STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic acids research, 32(suppl_2), W500-W502.

[324] Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers: Original Research on Biomolecules, 22(12), 2577-2637.

[325] Hottenrott, S., Schumann, T., Plückthun, A., Fischer, G., & Rahfeld, J. U. (1997). The Escherichia coli SlyD is a metal ion-regulated peptidyl-prolyl cis/trans-isomerase. Journal of Biological Chemistry, 272(25), 15697-15701.

[326] Maity, H., Muttathukattil, A. N., & Reddy, G. (2018). Salt effects on protein folding thermodynamics. The journal of physical chemistry letters, 9(17), 5063-5070.

[327] Ohgushi, M., & Wada, A. (1983). 'Molten-globule state': a compact form of globular proteins with mobile side-chains. FEBS letters, 164(1), 21-24.

[328] Goto, Y., Adachi, M., Muta, H., & So, M. (2018). Salt-induced formations of partially folded intermediates and amyloid fibrils suggests a common underlying mechanism. Biophysical Reviews, 10(2), 493-502.

[329] Cooper, P. K., & Hanawalt, P. C. (1972). Role of DNA polymerase I and the rec system in excision-repair in Escherichia coli. Proceedings of the National Academy of Sciences, 69(5), 1156-1160.

[330] Persky, N. S., & Lovett, S. T. (2008). Mechanisms of recombination: lessons from E. coli. Critical Reviews in Biochemistry and Molecular Biology, 43(6), 347-370.

[331] Cooper, P. K., and P. C. Hanawalt. 1972. Role of DNA polymerase I and the rec system in excision-repair in Escherichia coli. Proc Natl Acad Sci U S A 69:1156-1160.

[332] Bergqvist, S., Williams, M. A., O'Brien, R., & Ladbury, J. E. (2003). Halophilic adaptation of protein–DNA interactions.

[333] Record Jr, M. T., Zhang, W., & Anderson, C. F. (1998). Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. Advances in protein chemistry, 51, 281.

[334] Chien, A., Edgar, D. B., & Trela, J. M. (1976). Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. Journal of bacteriology, 127(3), 1550-1557.

[335] Steitz, T. A. (1998). A mechanism for all polymerases. Nature, 391(6664), 231-232.

[336] Wowor, A. J., Datta, K., Brown, H. S., Thompson, G. S., Ray, S., Grove, A., & LiCata, V. J. (2010). Thermodynamics of the DNA structural selectivity of the Pol I DNA polymerases from Escherichia coli and Thermus aquaticus. Biophysical journal, 98(12), 3015-3024.

[337] McClure, W. R., & Jovin, T. M. (1975). The steady state kinetic parameters and non-processivity of Escherichia coli deoxyribonucleic acid polymerase I. Journal of Biological Chemistry, 250(11), 4073-4080.

[338] Davidson, J. F., Fox, R., Harris, D. D., Lyons-Abbott, S., & Loeb, L. A. (2003). Insertion of the T3 DNA polymerase thioredoxin binding domain enhances the processivity and fidelity of Taq DNA polymerase. Nucleic acids research, 31(16), 4702-4709.

[339] Bernad, A., Blanco, L., Lázaro, J., Martin, G., & Salas, M. (1989). A conserved 3′→ 5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases. Cell, 59(1), 219-228.

[340] Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A., & Benkovic, S. J. (1987). Kinetic mechanism of DNA polymerase I (Klenow). Biochemistry, 26(25), 8410-8417.

VITA

Xinji Zhu was born in Ningbo, Zhejiang, China on October 31, 1989. In 2008, He attended East China University of Science and Technology, Shanghai, where he earned a Bachelor of Science Degree in Bioengineering through the Department of Bioengineering in May 2012. In September 2012, Xinji entered the laboratory of Dr. Vince J. LiCata in the Department of Biological Sciences at Louisiana State University and Agricultural and Mechanical College in Baton Rouge, Louisiana for doctoral study.