Testing Macromolecular Rate Theory

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Science (Research)

at

The University of Waikato

by

Kirsty Leigh Kraakman

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2017

Abstract

Enzymatic rate increase with respect to temperature has widely been described by transition-state theory. The experimentally observed rate decline above an optimum temperature (T_{opt}) for enzymes has previously been attributed to thermal denaturation, despite the known discrepancies with this rationalisation. Recently, a new model has been proposed to describe the temperature dependence of enzymatic rates: macromolecular rate theory (MMRT). This new theory incorporates heat capacity into the rate equation to provide a more robust thermodynamic description of rates, and account for the distinct curvature seen in biological temperature-rate profiles. The current study explores the effect of enzyme vibrational modes on heat capacity, and how alterations to the distribution of these modes, by heavy isotope substitution, affect the change in heat capacity along the reaction coordinate. The results presented show clear evidence for the role of vibrational mode frequency distributions in governing the curvature of temperature-rate profiles, and present a hypothesis for how this shift differs for the enzyme-substrate complex and the enzyme-transition state complex. The study also addresses the ability of MMRT to accurately model enzymatic rates over a wide temperature range for two different enzymes. As a result of this study, a new MMRT equation was generated to include a temperature dependence term for the heat capacity, a factor previously considered negligible. The data generated has shown that heat capacity is temperature dependent and this is a significant factor in accurately describing enzymatic rates with respect to temperature. In particular, for complex enzymatic reactions whose enthalpy distribution is significantly narrowed over the reaction co-ordinate. These findings have helped to develop MMRT to provide more accurate descriptions of a broader range of data. Additionally, the conclusions regarding vibrational mode distribution provide insight into the physical basis for heat capacity changes over the course of the enzymatic reaction. Investigation of the hypotheses generated from this research will offer further insight into the mechanistic contributions of vibrational modes and heat capacity in enzyme catalysis.

Acknowledgements

Thank you to Professor Emily Parker and the University of Canterbury for the use of their NanoDSC instrument. Special thanks to Gert-Jan Moggre and Effie Fan for their help in carrying out the DSC experiments for this work.

Thank you to the University of Waikato for the financial support I have received through the Postgraduate Taught Fees and Masters Research Scholarships. Also thank you to the Waikato Graduate Women's Trust Masters Fees Award scholarship for additional funding.

I would like to say a big thank you for the support and encouragement from the C.2.10 family. This journey would not have been possible without your kind words and wealth of knowledge. You are a truly wonderful and talented group of individuals who have made my experience one to be remembered.

Thank you to Judith for making the lab so great to work in. Life in C block would not be nearly as easy or enjoyable without you, your warm smile, and your macarons.

A very special thank-you to Erica Prentice, firstly for the cloning of both KSI and MalL - the results presented in this thesis would not have come about without your help. Mostly however, thank you for being not only a friend, but also a brilliant mentor. For your advice and guidance I am truly grateful - I owe a great deal of my sanity to you.

To my supervisor Vic Arcus. Thank you for your endless enthusiasm, and your passionate thermodynamic explanations; for teaching me the importance of optimism, and how to look sideways, backwards, and upside down at everything. The vivaciousness with which you approach science is inspiring.

To my family. Thank you for the consistent attempts to understand my '*emzymes*'. I cannot explain how much it means to me, for you I am truly blessed.

Finally to Henri. In the face of my somewhat irrational behaviour at times your continued faith in my abilities has been outstanding. Thank you for all that you do to encourage me and bring a smile to my face.

Table of Contents

List of Figures

List of Tables

List of Equations

List of abbreviations

AMP	Ampicillin
APS	Ammonium persulphate
BSA	Bovine serum albumin
DONM	Deoxynojirimycin
DSC	Differential scanning calorimetry
D-WT	Deuterium exchanged wild-type
EDTA	Ethylenediaminetetraacetic acid
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
KAN	Kanamycin
KSI	Ketosteroid isomerase (EC 5.3.3.1)
LB	Luria-Bertani
MalL	Isomaltase (EC 3.2.1.10)
MMRT	Macromolecular rate theory
p-NP	*para*-nitrophenol/ *p*-nitrophenol
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
TEMED	Tetramethylethylenediamine
Tris	Tris(hydroxymethyl)aminomethane
WT	Wild-type

1 Introduction

1.1 Temperature dependence of rates

Chemical rates are dependent on a number of factors and arguably, one of the most critical is temperature. The temperature dependence of chemical rates has been studied extensively since the 19th century, producing a plethora of equations and theories to describe the mechanisms involved with this relationship (Eyring, 1935; Eyring & Polanyi, 1931; Hobbs et al., 2013; La Mer, 1933; Peterson, et al., 2004a). Arrhenius first described the rate increase in product formation for simple chemical reactions in common solvents as an exponential function (Equation 1.1.) (Eyring, 1935). The increased kinetic energy possessed by reactants at higher temperatures increases the probability of successful collisions with enough energy to surpass the activation energy barrier (E_a) in an exponential fashion (Engel & Reid, 2006). This corresponds to an exponential increase in the rate of reaction (k) with increasing temperature. In 1935 Eyring introduced the idea of a transition state (TS, first termed the activated complex) (Eyring, 1935). The TS is an intermediate state existing between substrate binding and product formation at the highest point of the potential energy surface (Laidler & King, 1983; Schwartz & Schramm, 2009). This TS is associated with a much higher energy than substrate or product alone (Nelson, 2003) resulting from greatly reduced degrees of freedom by comparison (Eyring, 1935). The TS has partially broken and partially formed covalent bonds that behave unlike classical bonds to give no resisting force, resulting in a lowered $\Delta G^{\ddagger}$ and therefore an increase in reaction rate (Frey & Hegeman, 2006; Garrett & Grisham, 2010). Based on the existence of this transition state complex, Eyring and Polanyi provided a more accurate description of activation energy, expressing the energy barrier in terms of the change in Gibb's free energy between the TS and the reactants ($\Delta G^{\ddagger}$) (Eyring & Polyani, 1931). Since $\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger}$, transition state theory allows the energy barrier to include contributions from enthalpy (H), entropy (S), and temperature (T). Additionally, the Eyring-Polanyi equation (Equation 1.2) expresses pre-exponential factor (A) from Arrhenius' eponymous equation instead as $\kappa\left(\frac{k_B T}{h}\right)$, where κ is the transmission co-efficient, k_B is Boltzmann's constant, T is temperature, and h is Planck's constant (Frey & Hegeman, 2006). This modified Arrhenius form provides a more accurate thermodynamic description of observed

rates compared to the Arrhenius equation alone; however, there are notable discrepancies in the accuracy of this theory for predicting the behaviour of reactions beyond small molecules.

Equation 1.1. Arrhenius equation.

$$k = Ae^{\frac{-E_a}{RT}}$$

Equation 1.2. Transition state theory equation

$$k = \kappa\left(\frac{k_B T}{h}\right)e^{\left(\frac{-\Delta G^{\ddagger}}{RT}\right)} = \kappa\left(\frac{k_B T}{h}\right)e^{\left(\frac{\Delta S^{\ddagger}}{R}\right)}e^{\left(\frac{-\Delta H^{\ddagger}}{RT}\right)}$$

Biological catalysis is dependent on enzymes. These are macromolecular catalysts responsible for binding substrate in such a way as to produce favourable conditions for substrate turnover. Enzymes bind substrate in the active site though electrostatic, hydrogen bonds, van der Waals, and ionic interactions at the enzyme-substrate interface to form the enzyme-substrate complex (Creighton, 1993). Though enzymatic mechanisms are much debated (Hammes-Schiffer, 2002; Schwartz & Schramm, 2009; Sutcliffe & Scrutton, 2009; Tuñón, Laage, & Hynes, 2015), literature agrees on the formation of the transition state complex, a rigid, high-energy intermediate product tightly bound to the enzyme active site (Creighton, 1993). Following this, in the most general sense, enzyme catalysis follows the same principles as small molecule catalysis. This enables an argument for the use of basic Arrhenius-type functions such as transition state theory to describe the temperature dependence of enzymatic rates. When applied in practice to biological rates however, there are obvious discrepancies in the ability of transition state theory to fit data, resulting from the significant increase in complexity from small molecule catalysis to macromolecular enzyme catalysis. Biological systems involve large macromolecules and are subject to a number of influencing factors that are amplified when compared to chemical systems. The presence of complicating factors lends weight to the notion that traditional rate theory equations may not adequately describe the behaviour of enzyme-catalysed rates. Biological rate versus temperature data displays a distinct curvature, presenting an optimum temperature for enzyme activity, after which point a decline in rate is observed (Garrett & Grisham, 2010). The inadequacy of current rate theory to account for this negative curvature has been acknowledged thoroughly in the literature (Arcus et al., 2016; Buchanan et al., 1999; Corkrey et al., 2012; Daniel, Danson, & Eisenthal, 2001; Elias et al., 2014; Hobbs et al.,

2013; Peterson et al, 2004a), with rate decline currently being attributed to thermal denaturation of the enzyme (Garrett & Grisham, 2010). While this explanation provides a rationalization for empirical observations, it does not account for the known thermodynamic stability of many enzymes at temperatures above the optimum (Buchanan et al., 1999), or the loss of activity of psychrophilic enzymes well below their thermal inactivation temperature (Feller & Gerday, 2003). It is reasonable to attribute a portion of the rate decline to enzyme denaturation, however a more robust explanation of the negative curvature is required to fully account for the change in catalytic rate past T_{opt}.

1.2 **Alternative rate theories**

In 2001 Daniel *et al.* proposed an alternative to transition state theory, a proposal termed the "*equilibrium model*". This model describes the existence of an inactive folded form of the enzyme in rapid equilibrium with the folded active state. Increases in temperature shift the equilibrium position to favour the folded inactive state, causing a reduction in activity exclusive of irreversible denaturation (Daniel et al., 2001). This concept enables derivation of the curvature seen experimentally, since the concentration of the active enzyme (E_{act}) is decreased with increasing temperature. The authors concluded that thermal denaturation to the irreversibly inactive enzyme results from denaturation of the folded inactive population, with no contribution from the active folded form (Daniel et al., 2001). The equilibrium model introduced a new parameter, T_{eq}; the temperature at which ΔG_{eq} (ΔG associated with the transition of E_{act} to E_{inact}, occurs at T_{eq}) is equal to zero, meaning that the two folded states E_{act} and E_{inact} are of equal concentration (Peterson et al., 2004a). This parameter serves as the pseudo T_{opt}, since after T_{eq} equilibrium is shifted to favour E_{inact}, and a decline in overall rate is observed. Daniel and colleagues acknowledge that T_{opt} shows a slight shift to higher temperatures as assay length is decreased, and offer T_{eq} as a means of describing a true temperature optima, to replace the determination of T_{opt} based on experimental data (Daniel et al., 2001; Peterson et al., 2004a). The published work determines a value for T_{opt} based on the peak of the fitted curve as seen from the equilibrium model (Peterson et al., 2004a). The 2004 paper by Peterson *et al.* looked into modelling five monomeric enzymes with the equilibrium model using both experimental and computational data (Peterson et al., 2004a). Each data set generated a rate versus temperature experimental plot fitted with the

2-dimensional equation, and a computational 3-dimensional plot of rate versus time versus temperature (Peterson et al., 2004a). The paper concluded that all experimental data conforms to the equilibrium model. Although, the derived parameters showed significant disagreements between reported T_{eq} values and T_{opt} values for the enzymes β-lactamase and adenosine deaminase; 22 °C and 14 °C respectively (Peterson et al., 2004a). Based on the principles of equilibrium theory, these values should be in close agreement. T_{opt} should be approximately equivalent to T_{eq} (1-αT_{eq}), where α is so small that αT_{eq} is much less than one (Peterson et al., 2004a). In a follow up paper in 2009, Daniel *et al.* analysed over 50 data sets for 30 enzymes and claimed the equilibrium model described a universal pattern in enzyme catalysed rates, including enzymes ranging in structure and substrate complexity (Daniel et al., 2009). Parameters generated via the equilibrium model were tabulated for each of the 28 enzymes (Daniel et al., 2009). One limiting factor for the equilibrium model was the lack of physical evidence for the existence of the folded inactive state, E_{inact}. The rapid exchange associated with the E_{act}/E_{inact} equilibrium makes it very difficult to characterise the individual states despite a large amount of work seeking such a molecular description (Daniel & Danson, 2010; 2013).

In 2014, Elias *et al.* published an opinion paper reviewing the two-state Arrhenius-like model for enzyme catalysis (Elias et al., 2014). The authors noted that a two-state model for active folded-inactive denatured enzyme over a temperature gradient does not align with experimental observations based on current rate equations. It was suggested that Daniel *et al.* had overcome this obstacle by developing a three-state model (Daniel & Danson, 2013), referencing enzyme systems as ensembles. The flexibility of the enzyme and its ability to move through different conformations lends weight to this concept (Elias et al., 2014). Indeed, enzyme substrate interactions should also be considered conformational ensembles (Frauenfelder, Parak, & Young, 1988), showing reversible transitions between different substrate conformations with differences in binding and catalytic rate (Elias et al., 2014). With increasing temperature, an increase in the variability of the substrate is seen, as more non-specific binding can be achieved, an idea termed '*floppiness*' by Elias *et al.* (2014). The authors reason that increasing temperature shows rate increase relating to kinetic motion, but also increased floppiness. This increased floppiness results in curvature since

at higher thermal energy there is a reduced availability of sufficiently pre-organised substrate varieties (Elias et al., 2014). They argued that the flexibility of an enzyme is not an accurate measure of the flexibility of the active site, since backbone and side-chain conformational rigidity can be achieved without alteration to the active site (Elias et al., 2014). This effectively postulates that increased flexibility measured at higher thermal temperatures does not correspond to increased catalysis through flexible binding. A consequence of the proposed hypothesis is that distal conformational rigidity has limited effect on the active site, and therefore catalysis, and should not be considered important. However, it is well established that distal binding and conformational rigidity can affect the active site and catalytic activity via both allosteric and non-allosteric mechanisms. (Cui & Karplus, 2008; Motlagh et al., 2015; Raman et al., 2014). As with the equilibrium model, the most significant obstacle faced by Elias *et al.* is that no experimental evidence to support the floppiness hypothesis was presented, and may in fact be difficult to obtain.

1.3 **Macromolecular rate theory**

Macromolecular rate theory (MMRT) has been presented as a model to describe curvature of catalytic rates with temperature in the absence of denaturation. Derived from thermodynamics, MMRT answers questions left by current rate theory by incorporating heat capacity (C_p) into the rate equation; accounting for curvature and enzyme stability beyond T_{opt}. C_p is a fundamental thermodynamic parameter describing the change in internal energy of a system with a change in temperature (Tipler, 1998). Proteins are highly flexible macromolecules and are thus associated with high heat capacities (Arcus et al., 2016; Sturtevant, 1977). Consequently, protein C_p is an essential consideration in catalytic temperature dependence. Changes in C_p dramatically affect the thermodynamic description of the system when compared with models exclusive of C_p (Figure 1.1) MMRT produces a curvature reflective of empirical observations unlike Arrhenius-type functions providing a rate decline beyond T_{opt} prior to thermally induced denaturation (Figure 1.1). MMRT allows a more robust depiction of thermodynamic contributions to enzyme behaviour by providing a mathematically intricate description of the temperature dependence of $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ and, by definition, $\Delta G^{\ddagger}$.

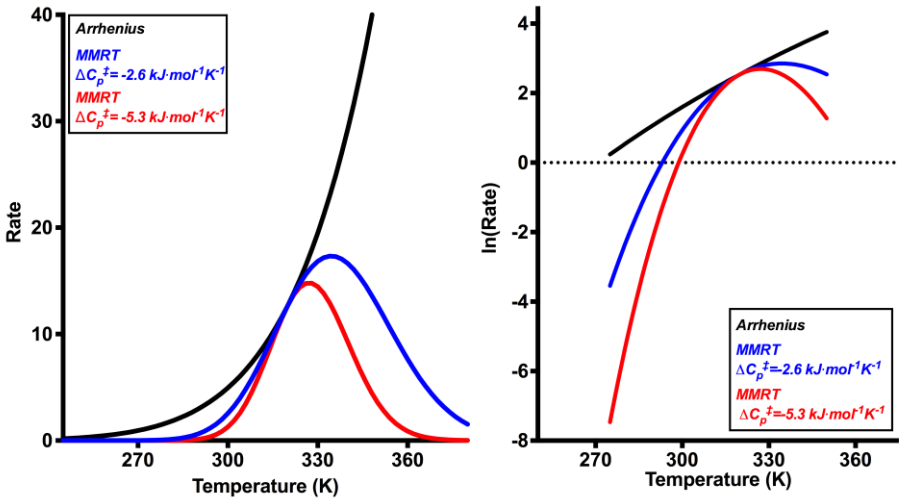

Figure 1.1. Arrhenius and MMRT functions to describe rate and ln(rate) versus temperature. Two different heat capacity values are shown for MMRT.

1.3.1 Equation derivation

The Eyring-Polanyi equation rests on the assumption that $\Delta G^{\ddagger}$ varies linearly with temperature according to the Gibb's equation (Equation 1.3.).

Equation 1.3. Gibb's free energy equation.

$$\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger}$$

This assumption requires both *H* and *S* to be independent of temperature. *H* and *S* cannot be entirely described in the absence of temperature, as both show significant variation (Prabhu & Sharp, 2005). The relationship between temperature and both enthalpy and entropy is characterised by heat capacity (Equation 1.4) (Creighton, 1993; Haynie, 2009b). The heat capacity at constant pressure, C_p, is used throughout as this is the parameter most relevant in biological settings.

Equation 1.4. Relationship of heat capacity, enthalpy, and entropy.

$$C_p = \left(\frac{\delta H}{\delta T}\right)_P = T\left(\frac{\delta S}{\delta T}\right)_P$$

This relationship can be visualised by plotting the temperature dependence of the difference between the enthalpy for the enzyme-substrate complex and the transition state ($\Delta H^{\ddagger}$) versus temperature (Figure 1.2). Integration of Equation 1.4 provides equations for the temperature dependence of the enthalpy and entropy, where T is the observed temperature, and T_0 is a suitable reference temperature (Figure 1.2). The entropic function is curved, by virtue of the logarithmic temperature term (Equation 1.6).

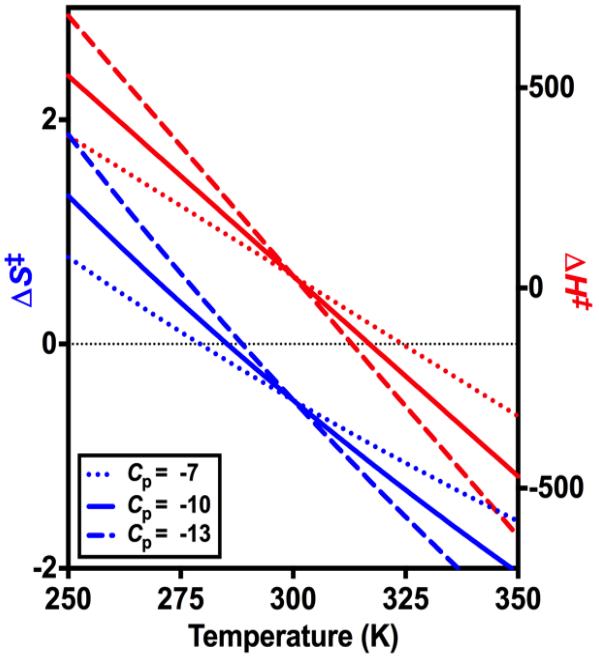

Figure 1.2. $\Delta H^{\ddagger}$ (right y-axis) and $\Delta S^{\ddagger}$ (left y-axis) versus temperature for ΔC_p values of -7, -10, and -13 kJ.mol^{-1}.K^{-1}

Equation 1.5. Enthalpy component to Gibb's free energy, including heat capacity.

$$H = H_0 + C_p{}^{\ddagger}(T - T_0)$$

Equation 1.6. Entropy component to Gibb's free energy, including heat capacity.

$$S = S_0 + C_{\mathrm{p}}{}^{\ddagger} ln\left(\frac{T}{T_0}\right)$$

$$= S_0 + C_{\mathrm{p}}{}^{\ddagger}(lnT - lnT_0)$$

Introducing these terms for heat capacity into the Gibbs equation (Equation 1.7):

Equation 1.7. Gibb's free energy component including modified enthalpy and entropy parameters.

$$\Delta G^{\ddagger} = \left[\Delta H^{\ddagger} + \Delta C_{\mathrm{p}}{}^{\ddagger}(T - T_0)\right] - T\left[\Delta S^{\ddagger} + \Delta C_{\mathrm{p}}{}^{\ddagger}(lnT - lnT_0)\right]$$

Transition state theory calls for a negative $\Delta G^{\ddagger}$ value in the exponential (Equation 1.8).

Equation 1.8. Negative Gibb's free energy component including modified enthalpy and entropy parameters.

$$-\Delta G^{\ddagger} = \left[-\Delta H^{\ddagger} - C_{\mathrm{p}}{}^{\ddagger}(T - T_0)\right] + T\left[\Delta S^{\ddagger} + \Delta C_{\mathrm{p}}{}^{\ddagger}(lnT - lnT_0)\right]$$

Incorporating this into the Eyring-Polyani equation in place of $-\Delta G^{\ddagger}$ (Equation 1.9):

Equation 1.9. Modified transition state theory equation including heat capacity.

$$k = \kappa\left(\frac{k_B T}{h}\right) e^{\frac{\left[-\Delta H^{\ddagger} - \Delta C_{\mathrm{p}}{}^{\ddagger}(T - T_0) + T\left[\Delta S^{\ddagger} + \Delta C_{\mathrm{p}}{}^{\ddagger}(lnT - lnT_0)\right]\right]}{RT}}$$

The transmission co-efficient (κ) describes the ratio between the molecules that proceed from the transition state to form product, and those that return to the reactants. For simplicity, we assume that κ is equal to one, and all transition state is converted to product (Equation 1.10).

Equation 1.10. MMRT equation.

$$k = \frac{k_B T}{h} e^{\left[\frac{-\Delta H^{\ddagger} - \Delta C_{\mathrm{p}}{}^{\ddagger}(T - T_0)}{RT}\right] + \left[\frac{\Delta S^{\ddagger} + \Delta C_{\mathrm{p}}{}^{\ddagger}(lnT - lnT_0)}{R}\right]}$$

The above equation provides the MMRT function to fit rate versus temperature data for enzyme activity. The MMRT equation can also be expressed in its log form (Equation 1.11).

Equation 1.11. Natural log form of MMRT equation.

$$lnk = ln\frac{k_B T}{h} - \left[\frac{\Delta H^{\ddagger} + \Delta C_{\mathrm{p}}{}^{\ddagger}(T - T_0)}{RT}\right] + \left[\frac{\Delta S^{\ddagger} + \Delta C_{\mathrm{p}}{}^{\ddagger}(lnT - lnT_0)}{R}\right]$$

1.4 **Protein heat capacity**

Heat capacity describes the change in internal energy with respect to temperature, formally defined as the amount of energy required to raise the temperature of one mole of a system by one degree at constant pressure (Equation 1.12) (Prabhu & Sharp, 2005; Tipler, 1998).

Equation 1.12. Formal definition of heat capacity.

$$C_{\mathrm{p}} = \frac{dU}{dT}$$

As a consequence, heat capacity measures the ability of a system to absorb thermal energy into partitioned energy modes (Haynie, 2009a). The Dulong-Petit law approximates that any solid system has a theoretical maximum heat capacity equal to 3R per atom (Tipler, 1998), however most systems plateau prior to reaching this maximum, and often only achieve a maximum specific heat capacity between 1.5R-3R per atom (Blundell & Blundell, 2006). Liquid water has a large heat capacity of ~75 $J.mol^{-1}K^{-1}$(Tipler, 1998), a value approximately equal to the theoretical maximum of 74.8 $J.mol^{-1}K^{-1}$ (3R per atom), however, ice has a specific heat capacity of ~37 $J.mol^{-1}K^{-1}$ (Tipler, 1998), equal to ~1.5R per atom. Proteins are large macromolecules associated with large heat capacities (Arcus et al., 2016), usually reaching ~1.5R per atom as a maximum (Makhatadze, 1998). A representative protein of 65 kDa has a heat capacity in water of ~140 $kJ.mol^{-1}K^{-1}$(Makhatadze, 1998). When compared with small molecule catalysis, this large heat capacity that cannot be ignored in thermodynamic calculations. It is predicted that at 298.15 K, each amino acid in the polypeptide chain contributes an additional ~185 $J.mol^{-1}K^{-1}$ to the heat capacity of the system (Makhatadze, 1998), thus increases in heat capacity show a strong correlation with increasing protein size (Arcus et al., 2016). A 2016 paper by Arcus *et al*. measured catalytic rate enhancement for a number of enzyme reactions using enzymes with catalytic domains ranging in molecular weight from ~5-75 kDa (Arcus et al., 2016). These findings implicate increasing heat capacity with increasing size and complexity of proteins.

Although C_p is known to be a temperature dependent quantity, this temperature dependence generally only has effects when extrapolated beyond the physiological temperature range (Creighton, 1993), although Makhatadze reports C_p values that are slightly temperature dependent for several proteins. In contrast, direct measurements of C_p have determined a largely temperature independent relationship (Liu et al., 2008). $\Delta C_p^{\ddagger}$ is a statistical thermodynamic quantity describing the difference in heat capacity of ensembles in the ground state and those at the transition state complex (Arcus et al., 2016). Changes to this protein heat capacity can be measured indirectly by generation of rate versus temperature data fitted using MMRT (Equation 10) allowing direct comparisons across multiple systems. $\Delta C_p^{\ddagger}$ cannot be directly measured under assay conditions, thus heat capacity information must be generated using model fits to experimental data.

Hobbs *et al.* (2013) determined that an increased $\Delta C_p^{\ddagger}$ value corresponds to decreased flexibility through mutagenesis experiments (Hobbs et al., 2013). A series of single residue mutants (V200S, G202P, V200T, and V200A) of the model enzyme MalL (glycoside hydrolase, EC 3.2.1.10) were generated and thermodynamically characterised. The crystal structures were solved for WT, V200S, and G202P to determine structural differences in flexibility. All three structures were very similar, displaying minimal RMSD values. Molecular dynamics simulations however, showed significant differences in flexibility. The V200S mutant introduced two additional hydrogen bonds, increasing the rigidity of the enzyme in the ground state. This decrease in flexibility is indicative of a decrease in heat capacity in the ground state. The flexibility of the enzyme transition state however is assumed to be unchanged, resulting in a difference in the $\Delta C_p^{\ddagger}$value between V200S and WT. This finding is seen in both molecular dynamic simulations and experimental observations. The G202P mutant showed no additional structural components and no changes in ground state flexibility, but had a significantly increased $\Delta C_p^{\ddagger}$value and reduced k_{cat} values compared to the WT. It was concluded from molecular dynamics simulations that this was the result of an increase in flexibility at the transition state (Hobbs et al., 2013). Flexibility is decreased at the binding of the transition state, a high-energy interaction with tight bonding interactions (Garrett & Grisham, 2010; Schramm, 2005). Compared with the ground state, the transition state enzyme complex has a significantly reduced number of available modes, thus a reduced number of conformations available, decreasing the overall flexibility of the enzyme complex, and increasing the value of $\Delta C_p^{\ddagger}$ (Hobbs et al., 2013).

1.4.1 Thermodynamics of protein heat capacity

Applied to enzyme thermodynamics, heat capacity describes the temperature dependence of enthalpy and entropy (Prabhu & Sharp, 2005), accounting for the change in these parameters between an observed temperature (T), and a reference temperature (T_0). Enthalpy and entropy represent the fundamental components of the Gibb's free energy, describing heat energy and energy of disorder respectively. In simple chemical systems, heat capacity is relatively small, and the temperature dependence of these parameters is linear with respect to temperature. In biological systems, heat capacity becomes a significant factor due to the large $\Delta C_p^{\ddagger}$ between the ground state and transition state, and dictates alteration of the Gibb's equation

to include a heat capacity term. Including ΔC_p into the entropy and enthalpy terms in the Gibb's free energy equation, results in curvature of the $\Delta G^{\ddagger}$ term. Plotting separate functions for $\Delta H^{\ddagger}$ and $-T\Delta S^{\ddagger}$ vs. temperature (Figure 1.3) visually demonstrates the competing enthalpic and entropic contributions to $\Delta G^{\ddagger}$. Significant deviation is seen from the linear relationship usually depicted in the absence of heat capacity.

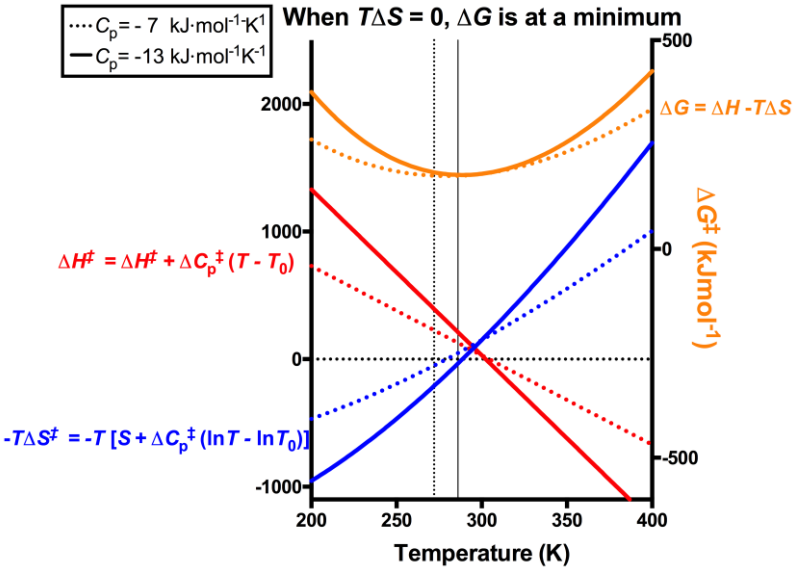

Figure 1.3. Contributions of entropy (ΔS) and enthalpy (ΔH) to ΔG with temperature. Two different values for C_p are shown.

Taking this a step further and plotting the entropic and enthalpic contributions to the natural log of the rate (that is, the entropic and enthalpic contributions to the $\frac{\Delta G^{\ddagger}}{RT}$ term of the Eyring-Polyani equation) allows functional derivation of curvature seen (Figure 1.4). As ΔC_p increases, the slope of the entropic and enthalpic terms also increases, resulting in a $\Delta G^{\ddagger}$ term reflecting the curvature observed experimentally. The $-\Delta G^{\ddagger}$ term is dominated by enthalpy at temperatures below the optimum, above which point entropy becomes the

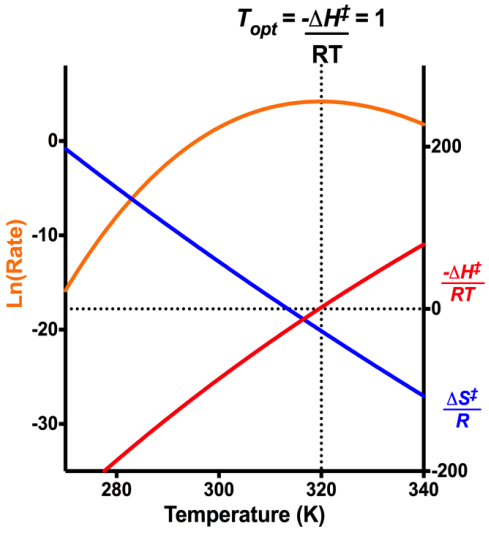

Figure 1.4. Individual contributions of $\frac{-\Delta H^{\ddagger}}{RT}$ and $\frac{\Delta S^{\ddagger}}{RT}$ to rate on right y-axis, ln(rate) as modelled by MMRT is seen on left y-axis

dominating term, and T_{opt} is seen as the temperature at which $\frac{-\Delta H^{\ddagger}}{RT}$ is equal to one (Figure 1.4).

1.4.2 $\Delta C_p{}^{\ddagger}$ in enzyme catalysis

The tight binding of the transition state significantly lowers the ΔG to increase the rate of the enzyme in order for the reaction to proceed (Garrett & Grisham, 2010). Following the logic prescribed by MMRT, the C_p at the transition state ($C_p{}^{\ddagger}$) is expected to be lower than the C_p of the enzyme-substrate complex (Hobbs et al., 2013). As a result, the change in heat capacity between the transition state complex and the enzyme-substrate complex ($\Delta C_p{}^{\ddagger}$) is negative. Previous measurements for $\Delta C_p{}^{\ddagger}$have determined this to be the case (Arcus et al., 2016; Hobbs et al., 2013; Liu et al., 2008; Schipper et al., 2014; Sturtevant, 1977). This provides a curvature matching that seen experimentally, where an increase in $\Delta C_p{}^{\ddagger}$(ie: $\Delta C_p{}^{\ddagger}$ becomes less negative) corresponds to reduced curvature and vice versa.

1.5 Role of vibrational modes in macromolecular heat capacity

The total internal energy of a system is composed of electronic, rotational, vibrational, and translational modes (Engel & Reid, 2006). Electronic modes are high energy, and for systems solvated in water at temperatures found within the biological temperature range (-20-100°C) electronic transitions are high-energy and therefore rare. This is a result of the large energy difference between the ground state and higher states, and consequently, these modes do not contribute to the heat capacity of folded proteins. Rotational modes describe the energy associated with rotation of the entire system in one of three dimensions: X, Y, and Z (Engel & Reid, 2006). For a single molecule or protein, the total number of available rotational modes is three, as the entire system of a protein is allowed to rotate in three angles. Translational modes comprise movement associated with kinetic energy and momentum, and occur in the X, Y, and Z Cartesian dimensions. Similar to rotational movement, the entire protein system can translocate only in these three axes, resulting in a total of three translational modes. Vibrational modes also occur in the X, Y, and Z dimensions but a total of three modes exist for each atom making up the protein structure. The total number of modes for a protein is therefore three times the total number of atoms making up the system

(3N) (Engel & Reid, 2006). The total number of vibrational degrees of freedom in a protein are then 3N-6 (3 translational modes + 3 rotational modes) (Xie, van der Meer, & Austin, 2001). For example, a small hypothetical 10 kDa protein containing 92 randomly generated amino acids using average amino acid composition (RandSeq, ExPasy) has a total of 1,437 atoms (ProtParam, ExPasy) (Gasteiger et al., 2005). This equates to 4,311 total modes and 4,305 vibrational modes. The total vibrational modes of the protein then comprise 99.9% of the total contributing modes to internal energy, rendering the translational and rotational modes negligible in heat capacity considerations. This percentage will only increase as protein size increases, allowing the system to be treated as purely vibrational for ease of concept.

Since heat capacity refers to the change in internal energy with temperature, vibrational modes provide the foremost contribution to protein heat capacity. These specific vibrations describe alterations to the structure by stretching, bending, oscillating, and torsional motions (Xie et al., 2001). These movements can occur between single atoms; such as the stretching observed between oxygen and hydrogen in a hydrogen bond; or movements of whole groups of atoms such as methyl groups or entire secondary structures including conformational changes and ligand binding movements (Balog, Smith, & Perahia, 2006; Hammes-Schiffer, 2002; Perticaroli et al., 2015; Whitmire et al., 2003; Xie et al., 2001). Single atom vibrations are associated with high frequency and therefore high-energy modes (Xie et al., 2001). For example, a typical O-H stretching vibration occurs at a frequency of ~3500-3700 cm^{-1} (A I Vogel, 1989; Xie et al., 2001). These modes are high frequency and high-energy, but produce only very small energy changes. Increasing size and molecular complexity allows the system to occupy lower frequency modes (0-2000 cm^{-1}) correlating with the increase in heat capacity associated with an increase in molecular weight size (Arcus et al., 2016). These lower frequency vibrational modes allow a greater variation in the free energy landscape than small molecules, by increasing the number of conformational sub-states available in equilibrium (Jones et al., 2016). The dominant contribution to heat capacity from vibrational modes is produced by the low frequency vibrational modes (0-200 cm^{-1}) (Turton et al., 2014). These low frequency modes correspond to vibrational movements of large clusters of atoms such as those involved in flexible protein movements, and produce a broad

absorption peak in the far infrared spectrum indicative of a continuum of conformational ensembles (Xie et al., 2001). Studies using molecular dynamics (Moritsugu, Miyashita, & Kidera, 2000) and density function theory (Xie et al., 2001) have determined ~15% of vibrational modes in proteins to be < 200 cm^{-1} (Xie et al., 2001). The density of these low-frequency states is limited, since the collective vibration of a large group of atoms can occur in limited number of ways (Haynie, 2009a), however, these modes hold the greatest capacity to absorb energy, and therefore contribute a greater energy to heat capacity. The contribution of vibrational modes to heat capacity can be modelled using the harmonic oscillator equation (Equation 1.13) (Figure 1.5), if we assume a quasi-harmonic approximation (QHA).

Equation 1.13. Equation for C_p as a simple harmonic oscillator

$$C_p = \frac{(hc\upsilon)^2}{k_B T^2} \frac{e^{\left(\frac{hc\upsilon}{k_B T}\right)}}{\left[e^{\left(\frac{hc\upsilon}{k_B T}\right)} - 1\right]^2}$$

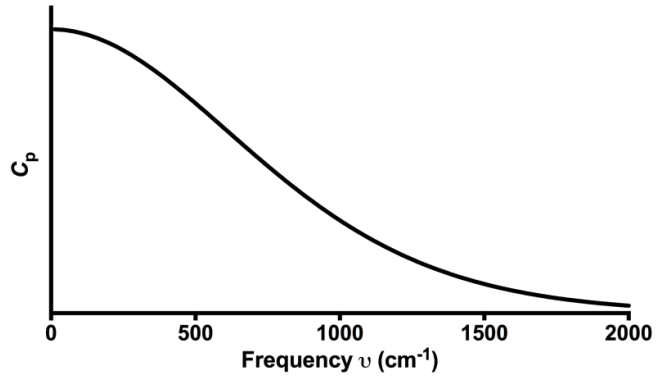

Figure 1.5. Contribution of vibrational frequencies up to 2000 cm^{-1} to heat capacity at 315 K.

1.6 **Vibrational mode hypothesis**

Low-frequency vibrational modes are expected to make the largest contribution to heat capacity (Figure 1.5). The population of vibrational modes within a protein is expected to dictate its flexibility, and therefore, its heat capacity. A 1993 paper by ben-Avraham assessed the vibrational spectra, up to 300 cm^{-1}, of five globular proteins to determine the density of vibrational states by normal mode computational analysis. Note that density of vibrational states [$g(\omega)$] is defined here as the number of modes per frequency range, divided by the total number of modes (ben-Avraham, 1993). Across all five proteins a universal curve was

observed. The size of the proteins ranged from 39-375 amino acid residues, indicating a vibrational mode distribution exclusive of size considerations. Figure 1.6 below shows the expected distribution of vibrational modes versus frequency as determined by ben-Avraham.

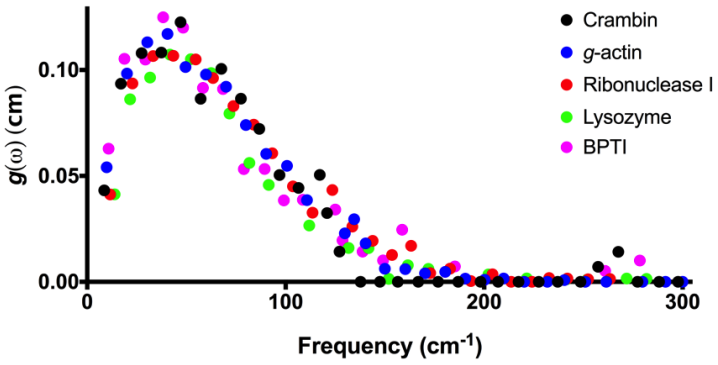

Figure 1.6. Density of vibrational modes versus frequency for crambin, g-actin, ribonuclease I, lysozyme, and BPTI. Taken from ben-Avraham (1993).

Further computational analysis of 135 proteins using enhanced normal mode analysis to include all atomic potentials was carried out by Na *et al.* in 2013. This analysis extended to 4000 cm^{-1} to give information on higher frequency, higher energy modes (Na, Song, & ben-Avraham, 2016). The study confirmed the universality of protein vibrational mode distributions across a wider range of protein sequences and frequencies, and provided detailed information about the type of vibrations populating each frequency. In 2000, Moritsugu *et al*. performed molecular dynamics simulations to elucidate the vibrational spectra of myoglobin. The results of this analysis generated a spectra consistent with the findings of Na *et al.* in 2016, however provided no detail on the exact movements corresponding to the peaks generated (Moritsugu et al., 2000). Figure 1.7 shows Figure 4 from Na *et al.* demonstrating that a universal curve is consistently represented across all proteins studied, and that at low frequencies the major contribution to vibrational modes comes from torsional and non-bonded movements, such as whole subunit or secondary structure motions. At high frequencies above 2000 cm^{-1}, the primary vibration contributing to vibrational modes is bond stretching (Na et al., 2016).

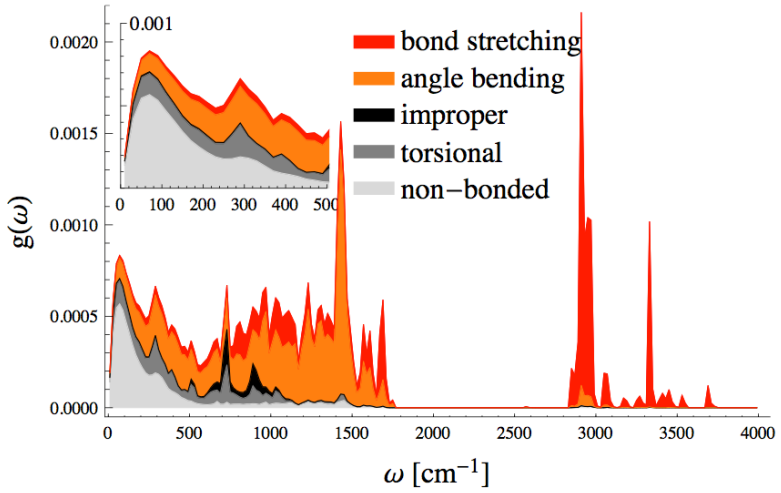

Figure 1.7. Vibrational mode distribution versus frequency for computational normal mode analysis of 135 proteins showing contributing vibrations. Taken from Na et al. (2016).

These frequency distributions describe the flexibility of the unbound protein (Balog et al., 2006). Ligand binding is well known to increase the stability and rigidity of the protein structure (Balog et al., 2006; Perticaroli et al., 2015), a fact which protein crystallographers readily exploit to assist in protein crystallisation (Arcus et al., 2016). Since vibrational modes provide the foremost contribution to protein heat capacity, it should follow that a change in heat capacity corresponding to a change in flexibility has an associated alteration in vibrational modes. It is then reasonable to hypothesise that on binding of a ligand, the distribution of vibrational modes will be altered to reflect the change in rigidity. Based on the equation for heat capacity as a simple harmonic oscillator (Equation 1.13), high flexibility and high heat capacity is associated with low-frequency vibrational modes (below 2000 cm^{-1}). As rigidity increases, the distribution of vibrational modes is expected to shift to higher frequency, while the number of vibrational modes available remains constant (3N-6). Turton *et al.* conducted femtosecond optical Kerr-effect (OKE) spectroscopy on lysozyme isolated from hen egg white (Turton et al., 2014). This type of spectroscopy measures orientational motions in the terahertz region including torsional and flexible motions of the protein backbone, and the liberations of amino acid side chains (Turton et al., 2014). OKE spectroscopy boasts a superior signal to noise ratio allowing discrete changes in flexibility, and therefore frequency distribution induced during biological processes, to be well defined. By obtaining the OKE spectra for lysozyme with and without the inhibitory ligand N,N',"-triacetylechitotriose bound, small but distinct alterations in intensity and

frequency were observed. On binding, a shift (blue) in frequency was seen, as well as an increase in intensity (Figure 1.8) (Turton et al., 2014). These results provide evidence for the potential frequency distribution shift expected on ligand binding, while confounding the hypothesis that the total number of modes remains unchanged. Turton *et al*. describe the alterations in intensity as resulting from increased polarizability of the delocalized oscillators upon binding of the inhibitor, however this is speculative. Further OKE spectroscopy to show reproducibility across a wider range of proteins with and without bound ligands would help to strengthen the arguments presented.

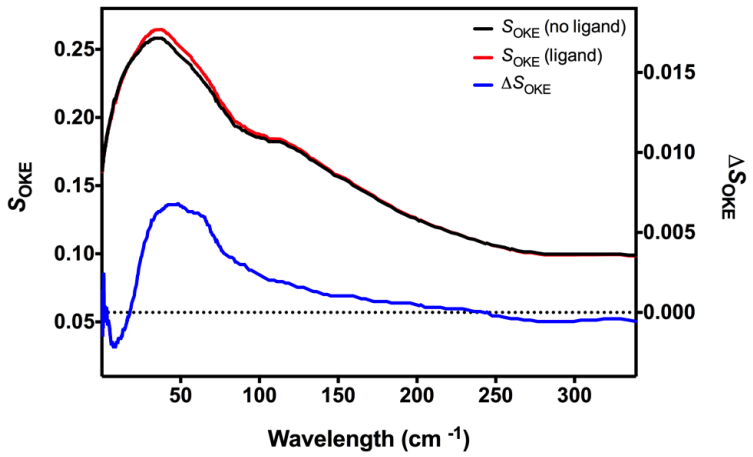

Figure 1.8. Spectra for lysozyme without and without inhibitory ligand bound showing alterations in frequency distribution. Change in spectra is shown in blue and defined as the sum of two Gaussian functions Taken from Turton et al (2014).

1.7 **Research Objectives**

Heat capacity is hypothesised to participate in catalysis (Arcus & Pudney, 2015; Austin, Roberson, & Mansky, 1989) by providing an energy reservoir that may or may not contribute to lowering the Gibb's free energy for the catalysed reaction. It follows that alterations to the distribution of vibrational modes might then alter the rate of catalysis. It is reasonable to postulate that manipulations of the vibrational mode energy within a system could result in changes to both rate and $\Delta C_p^{\ddagger}$, and that MMRT will accurately model each system. To investigate this hypothesis, a series of questions were proposed.

Does a change in vibrational mode distribution correspond to alterations in the overall rate? Does a change in vibrational mode distribution correspond to

alterations in $\Delta C_p^{\ddagger}$ for experimentally obtained rate versus temperature data? If this is the case, does MMRT sufficiently model these alterations in different enzyme systems? Can any improvements be made to better model these alterations?

1.7.1 Experimental design

It was postulated that discreet alterations to vibrational mode frequencies could be produced by slow exchange of freely exchangeable hydrogen atoms for deuterium atoms within a buffer solution. Purified enzymes suspended in buffer freely transfer exchangeable hydrogen atoms of amino acid residues with protons existing in solution. By making buffer solutions containing deuterated water (D_2O) rather than H_2O, it was expected that, over a period of time, the exchangeable hydrogen atoms in the enzyme would be replaced with deuterium atoms. Deuterium is a stable isotope of hydrogen with approximately 0.015% abundance, containing a neutron not present in common hydrogen (Sharp, 1990). The additional neutron present in deuterium results in mass increase from ~1 (^{1}H) to ~2 (^{2}H) (Sharp, 1990). The exchange of hydrogen for deuterium would therefore increase the overall mass of the protein without altering the number of electrons present. This should correspond to a change in the distribution of vibrational modes, resulting in alterations to heat capacity and possibly overall catalysis.

Experiments were devised based on the premise that the increase in exchanged deuterium atoms over the time course would correspond to an increase in rate. Alterations to rate could be correlated to increasing exchange and changes in vibrational mode distribution. A series of data were to be obtained for two enzymes, including data for wild-type (WT) and variants with increased isotopic mass; including deuterium exchanged and partially heavy isotope samples. By fitting data with Equation 1.11, differences in thermodynamic parameters could be visualised, to address how alterations to vibrational modes affected overall rate and $\Delta C_p^{\ddagger}$. Fitting data with the MMRT equation would model its efficacy in describing data for different systems and under different manipulations to energy modes. Particular attention was paid to goodness of fit (R^2) and error margins on data points and generated parameters.

1.1.1.1 Hydrogen-deuterium exchange

Exchange of covalently bound hydrogen atoms within the protein structure for hydrogen (or any of its isotopic forms) present in the solvent is a well-documented phenomenon (Creighton, 1993). Experimental data is commonly generated by exchanging hydrogen atoms involved in the amide backbone of the protein with isotopic forms due to the rapid and convenient exchange time of these particular bonds (Creighton, 1993). This exchange occurs via both acid and base catalysis and is thus, highly pH dependent. NMR experiments following the exchange of NMR active ^{1}H atoms for inactive ^{2}H atoms gives detailed information about the rate of exchange under different pH conditions. At pH 3.0, acid and base catalysed processes in water are approximately equal and the exchange rate is at a minimum. Above this pH, the rate of exchange increases approximately ten-fold per pH unit (Figure 1.9). Experimental NMR data shows slower exchange rates for hydrogens contained within the hydrophobic centre of folded proteins, particularly those involved in hydrogen bonding in β-sheets (Creighton, 1993). Based on this concept, rapid exchange of hydrogen atoms for deuterium could be used to manipulate kinetic behaviour of the enzyme by altering the vibrational frequencies of individual modes. H/D exchange in proteins has been used extensively to characterise local and global protein folding/unfolding and protein stability (Clarke et al., 1993).

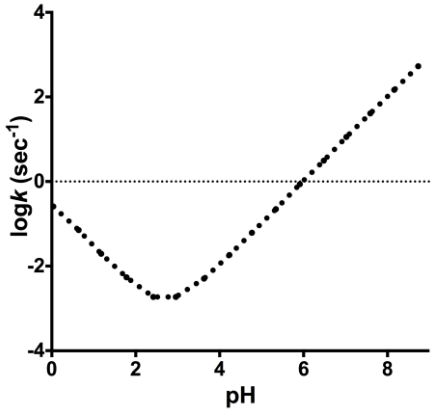

Figure 1.9. pH dependence of rate of exchange of hydrogen in a model amide group. Taken from Creighton (1993) (Figure 7.8).

1.1.1.2 Enzymes used in experimental design

The experimental design included the use of two model enzyme systems, ketosteroid isomerase (KSI) (EC 5.3.3.1) and isomaltase (MalL) (EC 3.2.1.10), each with a well-established reaction mechanism and assay protocol.

KSI is a small 14.9 kDa dimeric enzyme responsible for the catalysis of the two-step isomerisation of 3-oxo-Δ^5-steroids to their 3-oxo-Δ^4-conjugated derivatives (Pollack, 2004). During the reaction a C-H bond adjacent to a carbonyl group is broken, followed by a 1,3-proton transfer to the adjacent double bonded carbon group (C4 to C6) (Pollack, 2004) (Figure 1.10), effectively causing the double bond to migrate (Frey & Hegeman, 2006).

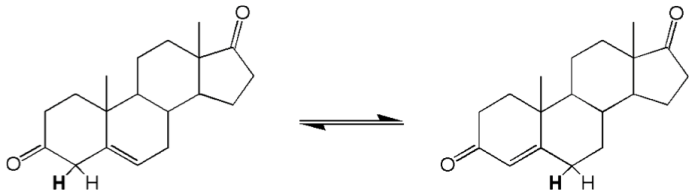

Figure 1.10. Reaction scheme for isomerisation of 3-oxo-Δ^5-steroid to 4-conjugate by ketosteroid isomerase. Taken from Pollack (2004).

MalL is a large 69 kDa α-glucosidase protein belonging to family 13 of the glycoside hydrolase group of proteins, and is responsible for hydrolysing the α 1-6 bond of isomaltose and some other oligosaccharides (" ENZYME entry 3.2.1.10," 2016.). The α1-6 bond is hydrolysed in a two-step retaining mechanism, proceeding via a covalent intermediate with two transition state species (Figure 1.11) (Frandsen, Palcic, & Svensson, 2002).

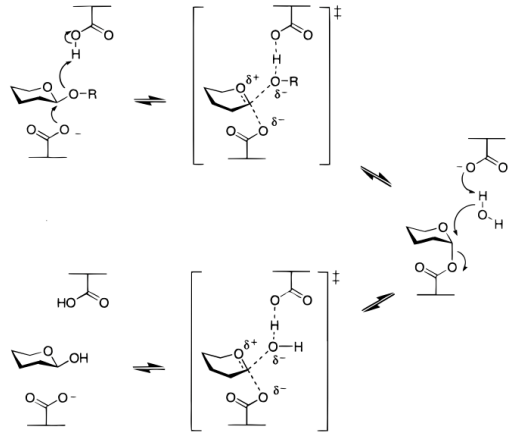

Figure 1.11. General mechanism for α-retaining glucosidases. Covalent intermediate is shown on the right hand side. Transition state species are indicated by ‡. Taken from Zechel and Withers (2000).

2 Materials and Methods

A considerable amount of method development was achieved preceding final protocols, the details of which are laid out in the following chapter. The methodology described here represents the final protocol for each experimental section. Details for media compositions can be found in Appendix I. Details for all buffers and reagents in this section can be found in Appendix II.

2.1 Protein expression

2.1.1 Ketosteroid isomerase expression

A glycerol stock of *Escherichia coli* strain BL21 containing the *Pseudomonas testeroni* ketosteroid isomerase (KSI) gene in pET-28b-pstI expression vector was provided by E. Prentice (University of Waikato) and plated onto an agar plate containing 0.05 mg/ml KAN. The plate was cultured overnight at 37 °C. A single colony was selected to create an 80% w/v glycerol stock for storage at -80 °C to be used in subsequent protein expression experiments.

2.1.1.1 LB media expression protocol

Glycerol stocks were used to grow 40 ml LB starter cultures with 0.5 mg/ml KAN overnight at 37 °C, shaking at 200 rpm. Starter cultures were used to inoculate 1 L of LB (1:100 inoculum) with 0.05 mg/ml KAN in a 2 L baffled flask, and grown at 37 °C while shaking at 200 rpm. OD was measured at 600 nm at various intervals using a ThermoSpectronic Helios spectrophotometer (ThermoFisher, USA) until log phase was reached (OD_{600} between 0.5 to 0.7). Protein expression was induced during log phase by the addition of IPTG to a final concentration of 1 mM, and culture was expressed overnight at 28 °C. Cells were isolated by centrifugation (4600 rpm, 20 min, 4 °C) and stored at -80 °C.

1.1.1.3 M9 minimal media expression protocol

Glycerol stocks were used to grow 5 ml LB starter cultures with 0.05 mg/ml KAN overnight at 37 °C, shaking at 200 rpm. Starter cultures were used to inoculate 25 ml M9 starter cultures (1:50 inoculum) with 0.05 mg/ml KAN in a 50 ml falcon tube, and grown at 37 °C while shaking at 200 rpm. M9 starter cultures were used to inoculate 250 ml of M9 with 0.05 mg/ml KAN in 1 L flask and returned to shaking incubator at 37 °C at 200 rpm. OD was measured at 600 nm at

various intervals using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA) until log phase was reached (OD_{600} between 0.5 to 0.7). Protein expression was induced during log phase by the addition of IPTG to a final concentration of 0.5 mM, and culture was expressed for approximately 60 hours at 28 °C. Cells were isolated by centrifugation (4600 rpm, 20 min, 4 °C) and stored at -80 °C.

2.1.2 Isomaltase expression

A glycerol stock of *E. coli* strain DH5α containing cloned *Bacillus subtilis* isomaltase (MalL) gene in pPROEX expression vector was provided by E. Prentice (University of Waikato).

2.1.2.1 LB media expression protocol

The glycerol stock was used to grow 40 ml LB starter cultures with 0.1 mg/ml AMP overnight at 37 °C, shaking at 200 rpm. Starter cultures were used to inoculate l L LB (1:100 inoculum) with 0.1 mg/ml AMP in a 2 L baffled flask, and grown at 37 °C while shaking at 200 rpm. OD was measured at 600 nm at various intervals using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA) until log phase was reached (OD_{600} between 0.5 to 0.7). Protein expression was induced during log phase by the addition of IPTG to a final concentration of 1 mM, and culture was expressed overnight at 18 °C. Cells were isolated by centrifugation (4600 rpm, 20 min, 4 °C) and stored at -80 °C.

2.1.2.2 M9 minimal media expression protocol

The glycerol stock was used to grow 5 ml LB starter cultures with 0.1 mg/ml AMP overnight at 37 °C, shaking at 200 rpm. Starter cultures were used to inoculate 25 ml M9 starter cultures (1:50 inoculum) with 0.1 mg/ml AMP in a 50 ml falcon tube, and grown at 37 °C while shaking at 200 rpm. M9 starter cultures were used to inoculate 250 ml of M9 with 0.1 mg/ml AMP in 1 L flask and returned to shaking incubator at 37 °C at 200 rpm. OD was measured at 600nm at various intervals using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA) until log phase was reached (OD_{600} between 0.5 to 0.7). Protein expression was induced during log phase by the addition of IPTG to a final concentration of 1 mM, and culture was expressed overnight at 28 °C. Cells were isolated by centrifugation (4600 rpm, 20 min, 4 °C) and stored at -80 °C.

2.2 SDS-PAGE Expression Gels

A 1 ml aquilot of cell culture was removed and transferred to a 1.5 ml Eppendorf tube prior to whole culture centrifugation. Samples were centrifuged (14,500 rpm, 3 mins) and supernatant discarded. Remaining cell pellet was re-suspended in lysis buffer prior to sonication on ice at level 4 for 5 bursts of 10 seconds with 30 second cooling intervals. Lysed cell solutions were centrifuged (14,500 rpm, 5 min) to pellet the insoluble fraction. Aquilots of 15 μL of both supernatant and pellet samples were combined with quenching buffer (Q4) (1:25 dilution) and incubated (95 °C, 5 min). A 10 μL aquilot of this solution was loaded onto a 12-16.5% SDS-PAGE gel alongside 10 μL 250-10 kDa Precision Plus Protein™ unstained protein ladder (Bio-Rad, USA). All gels were run at 15 mA for approximately 1.5 hrs in TG-SDS running buffer, prior to staining in Fairbanks A stain (30 min staining period). Gels were visualised using an Omega Lum™ G imager (Aplegen, USA) following an overnight destaining in Fairbanks D stain.

2.3 Protein purification

Purification of both proteins was achieved through immobilized metal affinity chromatography using a nickel HisTrap column (GE Life Sciences, NZ) to purify protein based on histidine tag affinity for nickel ions. Anionic purification was followed by size exclusion chromatography using a superdex 200 10/300 column (Amersham Biosciences, UK), to isolate the protein of interest based on size. Both purification steps were achieved using an ÄKTA purifier (GE Life Sciences, NZ) fast protein liquid chromatography system.

2.3.1 Nickel affinity chromatography purification

2.3.1.1 Cell preparation

Cell pellets were re-suspended in 20 ml lysis buffer and sonicated on ice at level 5 for 6 bursts of 15 seconds using fine sonication tip, allowing 30 second cooling intervals on ice. Sonicated cells were centrifuged (13,000 rpm, 20 minutes, 4 °C) to pellet cell fragments and leave expressed protein in the supernatant.

2.3.1.2 HisTrap column preparation

A 5 ml HisTrap column (GE Life Sciences, NZ) was prepared by addition of the following solutions at 1 ml/min:

- 10 ml MQ H_2O
- 10 ml EDTA
- 10 ml MQ H_2O
- 5 ml $NiCl_2$
- 10 ml MQ H_2O
- 10 ml Lysis buffer

Lysed cell supernatant was introduced to the equilibrated column at 1 ml/min through a series of filters (1.20 μm, 0.45 μm, 0.20 μm).

2.3.1.3 Purification protocol

The ÄKTA *Purifier* was prepared by manual pump and flow wash with filter sterilised MQ H_2O and appropriate buffers. HisTrap column was introduced at 1 ml/min (Lysis buffer) with alarm pressure set to 0.3 MPa. Following a stabilization period at 4% elution buffer, 100% lysis buffer was run at 1 ml/min for 10 ml prior to gradient introduction of elution buffer at 2% per ml. 2 ml fractions were collected from 10-60 ml (25 fractions, A1-15, B1-10) ending in 100% elution buffer. Fractions containing protein as given by peak in UV (280 nm) trace were retained for further purification. Column was removed and washed with 10 ml MQ H_2O, followed by 10 ml 20 % ethanol for storage.

2.3.2 Size exclusion purification

2.3.2.1 ÄKTA*Purifier* preparation

Pump and flow lines washed with filter sterilised MQ H_2O prior to introduction of a S200 10/300 column (Amersham Biosciences, UK) at 0.5 ml/min with alarm pressure set to 1.3 MPa. A 500 μL injection loop was connected and washed with 5 ml MQ H_2O. the size exclusion column was prepared by flow through of 25 ml MQ H_2O at 0.3 ml/min followed by equilibration with size exclusion buffer at 0.5 ml/min.

2.3.2.2 Purification protocol

Those fractions containing protein from nickel affitinity chromatography purification were pooled and concentrated to ≤20 mg/ml by centrifugation

(3300 rpm, 4 °C) using 5 kDa (KSI) or 10 kDa (MalL) molecular weight cut off concentrator [Amicon, USA (10 kDa); GE Life Sciences, NZ (5 kDa)]. Concentrated fractions were introduced to a 5 ml injection loop using 1 ml syringe and filter attachment (0.2 μm) following loop rinse with size exclusion buffer. A total of 25 ml size exclusion buffer was run through the column at 0.5 ml/min with alarm pressure set to 1.6 MPa. Fractions of 0.5 ml were collected from 7-25 ml (37 fractions, A1-15, B1-15, C1-7). Fractions containing protein as given by peak in UV (280 nm) trace were retained for use in downstream applications.

Following purification, the pump was washed with filter sterilized MQ H_2O. The column was flushed with 25 ml MQ H_2O (0.5 ml/min) followed by 25 ml 20% ethanol (0.3 ml/min) for storage before being removed from purification system. Pump and flow through were washed with MQ H_2O and 20% ethanol for storage.

2.3.3 Purification SDS-PAGE Gel

Aquilots of 15 μL of fractions corresponding to UV (280 nm) protein of interest peak and any additional protein peaks were retained for gel analysis. SDS-PAGE gel protocol was carried out as per previous expression gel (2.2) for these samples.

2.4 Protein concentration determination

2.4.1 Bradford assay (KSI)

Protein concentration was quantified based on the original dye-binding protocol determined by Bradford (1976) (Bradford, 1976). BSA solutions ranging in concentration from 0.0 mg/ml-0.7 mg/ml were made up in size exclusion buffer. Aquilots of 40 μl Bradford dye (Sigma, USA) and 5 μl standard BSA solutions were mixed in 96 well plate, prior to addition of 155 μl size exclusion buffer to each standard. Absorbance (595 nm) was read using ThermoScientific Multiskan Go plate reader (ThermoFisher, USA) to give a standard curve. Protocol was repeated for purified enzyme and enzyme dilutions in adjacent wells. Concentrations for samples with absorbance readings within the standard curve range were determined using the standard curve fit to a linear equation and appropriate dilution factor.

2.4.2 Nanodrop (MalL)

Absorbance (280 nm) of 2 μl purified protein was read using a ThermoScientific Nanodrop 2000 (ThermoFisher, USA). The instrument was blanked with size exclusion buffer. Concentration was determined in accordance with Beer-Lambert law(Equation 2.1), by division of absorbance reading by extinction co-efficient (ε) for MalL, (2.213 $M^{-1}cm^{-1}$), as determined from the protein sequence in ProtPram (Gasteiger et al., 2005).

Equation 2.1. Beer-Lambert equation

$$A = \varepsilon cl$$

2.5 Enzyme parameter assays

All assays were conducted at ambient temperature (22 °C). Note that for all KSI assays the substrate (19-Nor-androst-5(10)-ene-3,17-dione) was dissolved in 100% methanol at 10 mM before being diluted to give final concentration. Final methanol concentration in assay reaction was required to be at least as low as 7.5%.

2.5.1 Michaelis-Menten kinetic analysis

Michaelis-Menten plots were generated for both KSI and MalL to determine the optimum final concentration of substrate in exchange assays to ensure enzyme saturation. Reaction was initiated by addition of 5 μL enzyme [0.196 mg/ml (KSI); 0.360 mg/mL (MalL)] into cuvette containing 295 μL assay buffer and 100 μL substrate [19-Nor-androst-5(10)-ene-3,17-dione (KSI); *p*-nitrophenol α-D-glucopyranoside (MalL)]. Absorbance [248.0 nm (KSI); 405 nm (MalL)] was read continuously over a 20 s time period at 0.1265 s intervals using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA). Assays were conducted for various substrate concentrations (0.005, 0.010, 0.0250, 0.050, 0.100, 0.250, 0.500, 1.00, 2.50 mM) in triplicate. Rate (ΔAbsorbance/second) was determined for each substrate concentration over the first 10 s of the reaction by linear regression analysis. Rates were converted to k_{obs} (s^{-1}) (substrate turnover) values using Equation 2.1 and enzyme concentration, and graphed against substrate concentration to generate Michaelis-Menten curves. These curves were used to determine values for V_{max} and K_m (in mM) using Michaelis-Menten non-linear regression fit in GraphPad Prism version 6.0g for Mac (GraphPad Software,

La Jolla California USA). Substrate concentration for future experiments then required to be at least two times the concentration of K_m, ideally as high as ten times.

2.5.2 pH profiles for MalL and KSI

Activity (rate) was determined in various buffers ranging from pH 3.5-13.3 by spectrophotometric assay. Rate values were converted to k_{obs} (s^{-1}) values using Equation 2.1 and enzyme concentration. In the case of MalL this required correcting each value for the appropriate molar extinction co-efficient for p-nitrophenol at 405 nm as previously determined. A graph of k_{obs} versus pH was generated to give an appropriate, reliable pH curve for activity (Figure 4.8&Figure 4.9).

2.5.2.1 KSI pH profile assay protocol

The reaction was initiated by addition of 5 μL KSI (0.115 mg/ml) into 500 μl quartz cuvette containing 375 μL buffer and 20 μL 19-Nor-androst-5(10)-ene-3,17-dione substrate (10 mM). Final substrate concentration in assay was 0.5 mM. Formation of the product Δ^4-androstene-3,17-dione(Frey & Hegeman, 2006) was followed at 248 nm over 20 s at 0.125 s intervals using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA). Rate (ΔAbsorbance/second) was determined for the initial 10 s of the reaction at each pH condition using linear regression analysis. These values were corrected to k_{obs} (s^{-1}) and graphed to give a pH profile.

2.5.2.2 MalL pH profile assay protocol

The reaction was initiated by addition of 5 μL MalL (0.36 mg/ml) into 500 μl quartz cuvette containing 363 μL buffer and 32 μL *p*-nitrophenol α-D-glucopyranoside substrate (25 mM). Final substrate concentration in assay was 2 mM. Formation of the product *p*-nitrophenol was followed at 405 nm over 20 s at 0.125 s intervals using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA). Rate (ΔAbsorbance/second) was determined for the initial 10 s of the reaction at each pH condition using linear regression analysis. These values were corrected to k_{obs} (s^{-1}) using Equation 2.1 and enzymes concentration, and pH corrected ε values, prior to being graphed to give a pH profile.

2.6 Isotope labelled and unlabelled temperature profiles fitted using MMRT

Temperature profiles were determined for KSI and MalL variants of both standard (^{12}C, ^{14}N) and heavy (^{13}C, ^{15}N) isotope labelled protein variants. An additional T_{opt} assay was carried out for the V200S mutant form of MalL (undeuterated buffer only).

2.6.1 KSI temperature profiles

For KSI, data was gathered for deuterium-exchanged samples alongside non-deuterium exchanged controls for both $^{12}C^{14}N$ and $^{13}C^{15}N$ enzyme. Each sample was diluted into an exchange solution of either deuterated H_2O (D_2O) or undeuterated H_2O (control).

2.6.1.1 KSI exchange solution preparation

Purified KSI ($^{12}C^{14}N$, 1.10 mg/ml; $^{13}C^{15}N$ 5.19mg/ml) was transferred to 1.5 ml Eppendorf tubes containing either deuterated or undeuterated exchange buffer (pH 7.0) to give final concentrations between 0.35-0.05 mg/ml. Deuterated and undeuterated exchange solutions were left to exchange overnight (15 hours, 4 °C) before transfer to ice.

2.6.2 KSI temperature profile assay protocol

Assays were carried out in triplicate for temperatures ranging from 288 K-325 K for both deuterium exchanged and non-deuterium exchanged samples. The reaction was initiated by addition of 5 µl exchanged enzyme to the sample cuvette containing 395 µl of 19-Nor-androst-5(10)-ene-3,17-dione (0.5 mM) heated or cooled to the appropriate temperature using an Eppendorf cyrothermomixer (Eppendorf, Germany). Accurate initial temperature measurements were achieved using a thermocouple (Omega, USA), and consistent temperature control of the spectrophotometer was achieved using a ThermoSpectronic single cell Peltier (Life Sciences, USA). Absorbance (248 nm) was recorded over a 20 s time period in 0.125 s intervals using a temperature controlled ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA), and a final temperature was recorded immediately following the assay period using the thermocouple (Omega, USA). Assay temperature was taken to be the average final temperature across all three

replicates for each sample set. Rate (ΔAbsorbance/second) was determined for the first 10 s of the reaction using a linear regression analysis. Rates were converted to k_{cat} (s^{-1}) values using Equation 2.1 ans enzymes concentration. The natural logs (ln) of these values were graphed against temperature to generate temperature profiles. These profiles were fit with the natural log form of the temperature independent MMRT equation (Equation 1.11) by non-linear regression analysis in GraphPad Prism version 6.0g for Mac (GraphPad Software, La Jolla California USA). This analysis generated values for entropy, enthalpy, and change in heat capacity.

2.6.3 MalL temperature profiles

For MalL, data was gathered for $^{12}C^{14}N$ WT, $^{13}C^{15}N$ WT, and a standard isotope ($^{12}C^{14}N$) V200S mutant.

2.6.4 MalL temperature profile assay protocol

Assays were carried out in triplicate for temperatures ranging from 278 K-330 K. The reaction was initiated by addition of 5 µl enzyme to sample cuvette containing 395 µl of *p*-nitrophenol α-D-glucopyranoside (2 mM) heated or cooled to the appropriate temperature using an Eppendorf cyrothermomixer (Eppendorf, USA). Accurate initial temperature measurements were achieved using a thermocouple (Omega, USA), and consistent temperature control of spectrophotometer was achieved using ThermoSpectronic single cell Peltier (Life Sciences, USA). Absorbance (405 nm) was recorded over a 20 s time period at 0.125 s intervals using a temperature controlled ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA), and a final temperature was recorded using the thermocouple (Omega, USA) immediately following the assay period. Assay temperature was taken to be the average final temperature across all three replicates for each sample set. Rate (ΔAbsorbance/second) was determined for the first 10 s of the reaction using a linear regression analysis. Rates were converted to k_{cat} (s^{-1}) values using Equation 2.1 and enzyme concentration. The natural logs (ln) of these values were and graphed against temperature to generate temperature profiles. These profiles were fit with the natural log form of temperature dependent MMRT equation (Equation 4.4) by non-linear regression analysis in GraphPad Prism version 6.0g for Mac (GraphPad Software, La Jolla California

USA). This analysis generated values for entropy, enthalpy, change in heat capacity, and heat capacity temperature dependence.

2.7 **Differential Scanning Calorimetry**

Differential scanning calorimetry (DSC) using a NanoDSC (TA instruments, USA) was carried out for WT (0.944 mg/ml) and WT (0.546 mg/ml) bound in a 1:1 molar ratio to deoxynojirimycin (DONM), a transition-state analogue, to determine differences in the temperature dependence of heat capacity.

All samples were contained in pH 7.0 size exclusion buffer. Samples and reference buffers were degassed for 20 min prior to the experiments. Both sample and reference cells were cleaned thoroughly with H_2O, 1% decon detergent solution and H_2O again, prior to rinsing with the appropriate buffer before each experimental run. Baseline scans containing the appropriate buffer in both the reference and the sample cell were run overnight to sensitize the instrument and generate a stable, flat baseline. Samples were run against their identical buffer as a reference. Temperature was increased from 5-90 °C in increments of 1 °C per minute, under 3 atmospheres of constant pressure. Raw heat data (μJ/s) was analysed using Nanoanlalyze software version 3.6 (TA instruments, USA) to subtract baseline and convert to molar heat capacity. T_m was defined at the temperature where maximum heat capacity was reached. Slope was determined for 300–310 K for each sample by linear regression analysis in GraphPad Prism version 6.0g for Mac (GraphPad Software, La Jolla California USA).

3 Method Development

Details for the media used in this section can be found in Appendix I. Details for all buffers, reagents, and gels mentioned in this section can be found in Appendix II.

3.1 Protein expression

3.1.1 KSI Expression in M9 Media

Glycerol stocks of recombinant *Escherichia Coli* strain BL21 containing the *Pseudomonas testeroni* KSI gene were used to grow 5 ml LB starter cultures with 0.05 mg/ml KAN overnight at 37 °C, shaking at 200 rpm. These starter cultures were used to inoculate 25 ml M9 (1:10 inoculum) starter cultures with 0.05 mg/ml KAN in a 50 ml falcon tube, and grown at 37 °C while shaking at 200 rpm. M9 starter cultures were used to inoculate 100-250 ml of M9 (1:50 inoculum) with 0.05 mg/ml KAN and returned to shaking incubator at 37 °C at 200 rpm. OD was measured at 600 nm at various intervals using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA) until log phase was reached (OD_{600} between 0.5 to 0.7). Protein expression was induced during log phase by the addition of IPTG to a final concentration of 0.5 mM, and the culture was expressed overnight at 28 °C, shaking at 200 rpm. Initial M9 expression conditions resulted in insoluble protein expression. SDS-PAGE gel analysis of pellet and soluble fractions after sonication indicated that majority of the expressed KSI is insoluble and aggregated (Figure 3.1). Consequently, a protocol for

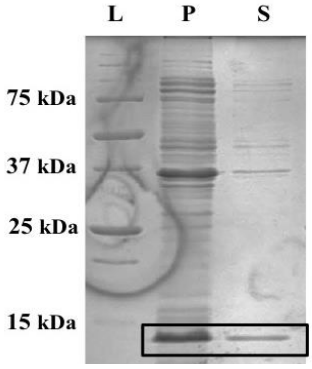

Figure 3.1. SDS-PAGE gel showing initial KSI expression in M9 minimal media. L= Precision Plus Protein™ standards unstained ladder (Bio-Rad); P= insoluble pellet fraction; S= soluble fraction. Products of expected size for KSI (14.9 kDa) are circled.

protein unfolding, isolation, and refolding was attempted (Hutchinson et al., 2006), in addition to a number of growth trials aimed at assessing optimal conditions for soluble expression.

3.1.1.1 Unfolding and refolding activity trials

To assess the viability of an unfold-refold protocol in producing workable amounts of active protein, trials were conducted using purified KSI. A protocol was modified from Hutchinson *et al*. (2006) (Hutchinson et al., 2006). A 1 ml aquilot of KSI (1.229 mg/ml, pH 5.0) was diluted into 1 ml unfolding buffer (50 mM Tris, 8 M urea, 3 mM EDTA, pH 9.0) and kept at either ambient temperature (22 °C) or 30 °C. Unfolding was followed by continuous assay (248 nm) using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA) to track unfolding (22 °C, 4 hours; 35 mins, 30 °C). The assay protocol followed was the same as used in the Michaelis-Menten assays (2.5.1). k_{obs} was graphed against time to generate unfolding curves (Figure 3.2A&C).

Once activity had plateaued at almost zero, the enzyme-unfolding buffer solution was diluted 1:10 into refolding buffer (50 mM KPO_4, pH 7.0) to give a final urea concentration of 0.8 M urea, and stored at 4 °C.

Refolding was followed by activity measurements at 4 °C using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA) at various intervals (20-180 min). The assay protocol followed was the same as used in the Michaelis-Menten assays (2.5.1). k_{obs} was graphed against time to generate refolding curves (Figure 3.2B&D). Continuous assays for the native folded purified protein were also conducted using the same protocol. Rates for undisturbed purified KSI were compared with unfolded and refolded samples to determine maximum refolding percentage as a function of activity (Figure 3.3).

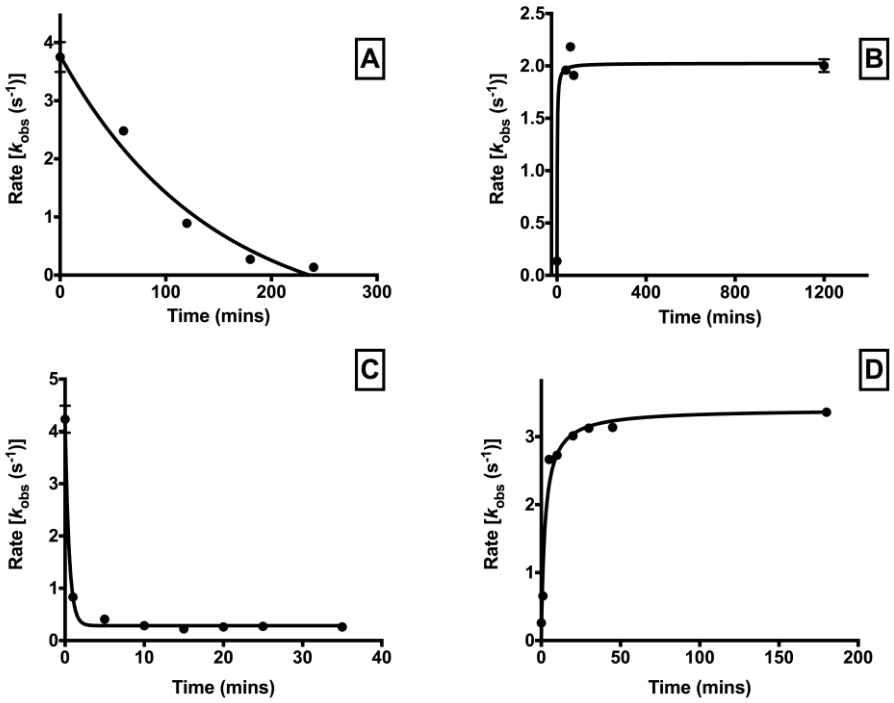

Figure 3.2. Rate versus time for unfolding and refolding of purified KSI. A Shows the unfolding rate at 22℃; B shows the corresponding refolding rate at 4℃. C shows the unfolding rate at 30℃; D shows the corresponding refolding rate at 4℃ .

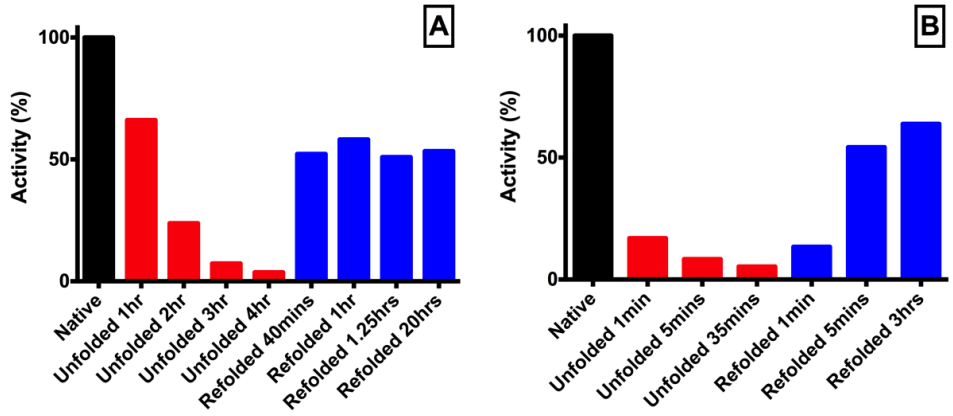

Figure 3.3. Enzyme activity as a percentage of native purified KSI activity for unfolding protocol at 22℃ (A) and 30℃ (B). A corresponds to A and B in figure above. B corresponds to C and D. Refolding occurred at 4℃ in both experiments.

Unfolding was 4-fold faster at 30 °C than 22 °C. Protein refolding at 4 °C yielded between 45% and 64% activity recovery, showing maximum activity recovering following unfolding at 30 °C. This finding was deemed significant enough to consider refolding KSI from an insoluble expression pellet grown in heavy M9

minimal media using previously established purification methods from the insoluble fraction (Cumming, 2012). Previous KSI expression in M9 media had presented significant insoluble expression, a total of 60% of which would be enough to perform the required experiments when refolded.

3.1.1.2 KSI isolation from insoluble pellet using nickel sepharose

Following the demonstration of effective refolding from unfolded purified KSI, an unfold-purify protocol was attempted using the insoluble pellet from an M9 minimal media expression trial. The protocol was based on that of Hutchinson *et al*. (2006) (Hutchinson et al., 2006). Pellet was re-suspended in 20 ml lysis buffer and sonicated on ice at level 5 for 6 bursts of 15 s, using fine sonication tip, allowing 30 s cooling intervals on ice. The sample was centrifuged (13,000 rpm, 20 min) and the supernatant was reserved for running on an SDS-PAGE gel. The pellet was re-suspended in 30 ml unfolding buffer (50 mM Tris, 8 M urea, 3 mM EDTA, pH 9.0) and incubated for 2 hours (37 °C, 200 rpm). Excess EDTA was bound by the addition of 6 mM $CaCl_2$. Nickel sepharose beads (20 µl) were washed with 1 ml wash buffer (8 M urea, 50 mM Tris, pH 9.0). A 1.5 ml aquilot of the pellet sample was loaded onto the washed beads and incubated for 15 min (37 °C, 350 rpm). The tube was centrifuged briefly to pull beads to the bottom of the tube. Supernatant was removed and reserved for running on an SDS-PAGE gel. This process was repeated five times to load a total 7.5 ml of pellet sample onto the beads. The beads were then washed 3 times with wash buffer, reserving each wash step to run on an SDS-PAGE gel. Two additional samples were retained for running on the gel. The foam layer that formed above the bead sample during the shaking and pipetting of the nickel beads, and a sediment layer, presumably of undissolved cell debris which was found at the bottom of the unfolded pellet sample.

3.1.1.2.1 Purification SDS-PAGE gel

A 15 µl aquilot of each retained sample was combined with Q4 buffer (Appendix II) (1:4 dilution) and boiled (95 °C, 5 min). A 10 µl sample of each of these steps was loaded onto a 12% SDS-PAGE gel and run at 15 mA for approximately two hours before a staining (one hour) and destaining (overnight)

process was carried out in Fairbanks A and Fairbanks D stain respectively (Appendix II).

The results of the gel (Figure 3.4) indicate that the nickel bead sample contained a significant amount of other protein and cellular debris. It is unlikely that this protein was bound to the nickel beads, but rather as the nickel beads were centrifuged to form a pellet, any undissolved insoluble fragments have remained as part of this sample. These results conclude that the pellet did not successfully dissolve in the extraction buffer, and it is unclear whether the KSI present in the nickel bead sample has been solubilised or remains aggregated. Consequently, it was decided that growth trials would be conducted in various M9 conditions to assess soluble expression levels before pellet refolding was revisited as a viable method for acquiring active KSI expressed in M9 minimal media.

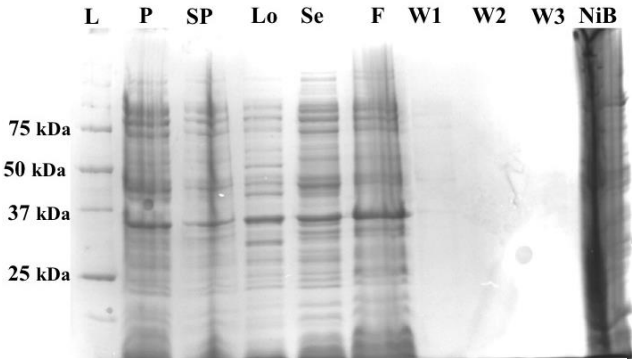

Figure 3.4. SDS-PAGE gel showing results of nickel pull-down. L= Precision Plus Protein™ standards unstained ladder (Bio-Rad); P= insoluble pellet fraction; SP= sonicated pellet; Lo= load; Se= Sediment layer found at bottom of Eppendorf tube; F= foam layer formed above sample; W= wash step; NiB= nickel beads.

3.1.1.3 Growth trials

A number of growth trials were conducted to determine whether or not any soluble KSI expression could be achieved by providing ideal conditions. This included testing various additives, temperatures, inoculums, and expression times in assessing growth rates and expression (Table 10, Appendix III).

3.1.1.3.1 Growth Curves

Growth curves of each culture were obtained through spectrophotometric monitoring at 600 nm using a ThermoSpectronic Heλios spectrophotometer (ThermoScientific, USA) to give optical density (OD_{600}) readings at various time points (Figure 3.5). Growth curves generated data about additive and temperature

effect on growth rate, and determined time of induction. Cultures were induced when OD_{600} reached 0.5-0.7 (log phase).

3.1.1.3.2 SDS-PAGE gels to assess protein expression

Following induction with 0.5 mM IPTG and an expression period, 1 ml of each cell culture was removed and reserved for an SDS-PAGE gel according to protocol in Section 2.2.. Remaining cells were harvested by centrifugation (4,600 rpm, 40 min) to provide cell pellets. Cell pellets were stored at -80°C.

1.1.1.3.1 Additives

The active form of thiamine (thiamine diphosphate) is an essential co-factor for all organisms (Makarchikov et al., 2003). Increased availability of co-factors is postulated to increase growth rates and potentially aid in soluble expression (Sambrook & Russell, 2001). Addition of 0.01% thiamine HCl had no significant effect on culture growth (Figure 3.5A&B), however soluble expression was marginally increased (Figure 3.5B, lanes 2-3 comparative to Figure 3.5A, lane 3). As a result, 0.01% thiamine HCl was continually added to the remaining M9 growth trials.

It was postulated that an increase in glucose availability in the media might reduce cellular stress and increase soluble protein expression. An additional 50% glucose was included in one culture. The high glucose culture showed no significant increase in soluble protein expression (Figure 3.5B), and excess glucose was omitted from further growth trials.

Literature suggests that iron supplementation is required for appropriate growth in minimal media as it is a critical co-factor for a number of enzymes (Benov & Fridovich, 1998). Additionally, a B vitamin mix was added to increase the media nutrient source. Both additives increased the growth rate (Figure 3.5C) and in combination, had a positive effect on soluble KSI expression (Figure 3.5C). Both additives were included in successive M9 growth trials.

The incorporation of a high molar excess of NaCl is postulated to increase soluble expression of by changing the ionic strength of solution (Chopra et al., 1994),

inducing cellular stress to activate transcription of chaperonin proteins. Chaperonin induction aids in post-transcriptional protein folding and prevents aggregation by binding hydrophobic residues (Baneyx, 1999). The addition of 250 mM and 500 mM NaCl significantly slowed culture growth (Figure 3.5C). A significant increase in soluble expression was noted (Figure 3.5C) and an additional protein expression band was seen at approximately 32 kDa, suggesting heat shock protein expression (Figure 3.5C). It was deduced that 250 mM additional NaCl provided a beneficial increase in protein solubility without slowing culture growth to an unreasonable rate.

Trials were conducted at a lower temperature (18°C), to increase soluble protein expression (Baneyx, 1999; Chopra et al., 1994). This drop in temperature could not be quantified as advantageous as the control culture (28°C) did not grow properly.

To assess the impact of culture size and aeration, a large-scale 250 ml culture was grown in identical conditions to the previous successful 100 ml culture (Figure 3.5D). The culture grew approximately 18% slower, providing a well-defined timeline for large-scale growth. Expression of these cultures was lower than for the previous trial (Figure 3.5D), prompting a final trial to re-assess the effect of expression temperature before proceeding to heavy isotope labelling.

Small and large-scale cultures for the determined media were grown and induced as per previous trials. Two cultures (100 ml, 250 ml) were expressed at 18 °C while the two remaining cultures were expressed at 28 °C. A similar level of expression was seen for cultures expressed at 28 °C, as for 18 °C, indicating no additional solubilisation of KSI at 18°C (Figure 3.5E). For this reason the heavy ($^{13}C^{15}N$) cultures were expressed at 28 °C for ease of experimental methods.

Original expression media used in these trials can be found in Appendix I. Summary of additives and culture conditions for growth trials can be found in Appendix III. Final conditions for KSI expression are laid out in Section 1.1.1.3. An initial MalL M9 expression using the same conditions as established for KSI

yielded sufficient soluble expression for further experiments. Final M9 minimal media expression protocol for MalL can be found in the methods and materials section (2.1.2.2).

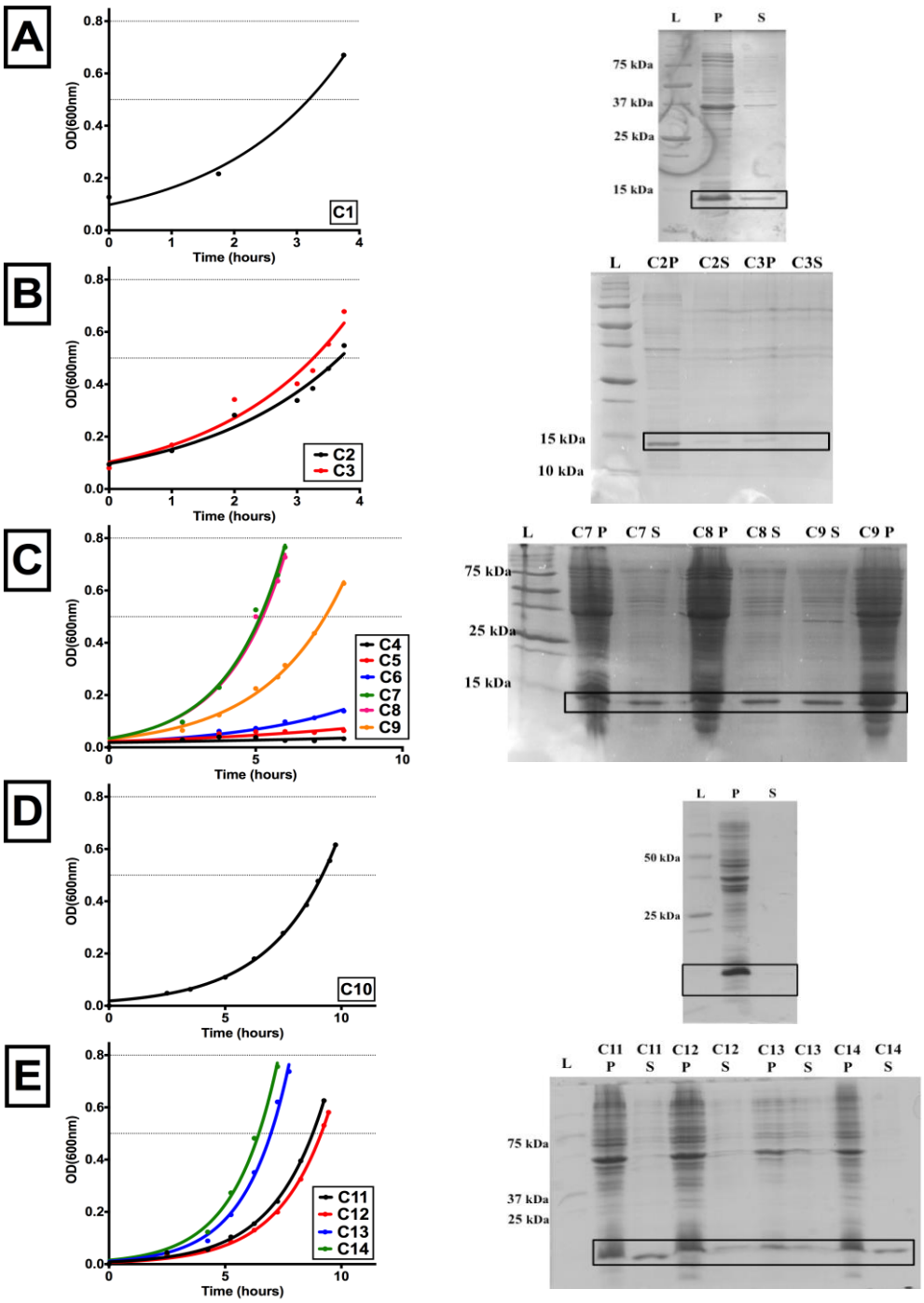

Figure 3.5. Growth curves and corresponding SDS-PAGE gels for all KSI minimal media M9 trials. Dotted lines show appropriate OD range for induction. Products of expected size for KSI (14.9 kDa) are circled. For full culture details see Table 10, Appendix III. In all cases, L= Precision Plus Protein™ standards unstained ladder (Bio-Rad); P= insoluble pellet fraction; S soluble fraction. Culture number is denoted as C'X'. A: Culture 1. No additives. B: Cultures 2 and 3. Thiamine (2,3) and additional glucose (3); C: Cultures 4-9. Thiamine (all), 250 mM NaCl (5,9), 500 mM NaCl (6), 10 µm $FeCl_3$ (7,9), B vits (8,9). D: Culture 10. Large scale (250 ml) C9. E: Cultures 11-14. C9 additives. 250 ml (11), 100 ml (12-14). 28℃ expression (11,12) 18℃ expression (13,14).

3.2 Determination of the pH dependence of *p*-nitrophenol molar extinction co-efficient

The molar extinction co-efficient (ε) of *p*-nitrophenol at 405 nm is known to be pH dependent (Beg, 1984; Biggs, 1954; Hriscu et al., 2013). In order to correctly calculate MalL k_{obs} and k_{cat} values for product turnover of enzymatic reactions, the correct ε (product) value is required. Since the pH of the assay was different to previously conducted assays in this laboratory, the correct *p*-nitrophenol ε needed to be determined.

3.2.1 Generation of *p*-nitrophenol

p-nitrophenol was created by adding 1.25% volume MalL (8.46 mg/ml) to a 2 mM solution of *p*-nitrophenol α-D-glucopyranoside and allowing the reaction to proceed for approximately 10 min at ambient temperature (22 °C). A 400 µl sample of the reaction mixture was assayed continuously using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA) at 405 nm to ensure enzymatic rate had depleted entirely over the 10 min period due to substrate availability. The *p*-nitrophenol-MalL mixture was centrifuged (3,300 rpm, 15 min) in a 10,000 MWCO concentrator to remove the enzyme.

3.2.2 Determination of *p*-nitrophenol concentration

Dilutions of 0, 0.5, 1, 1.5, 2, 2.5 % *p*-nitrophenol were made up using a range of buffers of varying pH (pH 3.5-13.3). Absorbance readings were taken for each solution using a ThermoSpectronic Heλios spectrophotometer (ThermoScientific, USA) at 348 nm. 348 nm represents the isobestic point for *p*-nitrophenol absorbance (wavelength at which ε is fixed)(Hriscu et al., 2013) Absorbance readings were averaged over the range of pH buffers and concentration of each dilution was calculated according to Beer-Lambert law Equation 2.1), using the literature ε for 348 nm (5.4 $mM^{-1}cm^{-1}$)(Hriscu et al., 2013).

3.2.3 Determination of pH standard curves

Absorbance readings (348 nm, 405 nm) were determined for each *p*-nitrophenol concentration at various pHs (pH 3.5-13.3) using a ThermoSpectronic Heλios spectrophotometer (ThermoFisher, USA). Standard curves of absorbance versus

concentration were generated in GraphPad Prism version 6.0g for Mac (GraphPad Software, La Jolla California USA) for each pH, and fit using the linear regression software to determine the slope (εl) for each standard curve (Figure 3.6&Figure 3.7). Values for εl are summarised in Table 1.

Table 1. Slope values for absorbance versus p-nitrophenol concentration at 348 and 405 nm

pH	**εl (348 nm)**	**εl (405 nm)**
3.5	2173 ± 66	-2 ± 36
4.0	2298 ± 37	78 ± 22
5.0	2506 ± 92	316 ± 48
5.5	2230 ± 84	253 ± 36
6.0	2403 ± 159	868 ± 82
6.5	2605 ± 110	2072 ± 85
7.0	2601 ± 88	3751 ± 108
7.5	2457 ± 72	5401 ± 138
8.0	2541 ± 120	6635 ± 262
8.5	2512 ± 122	7269 ± 252
9.0	2348 ± 79	7300 ± 224
11.0	2410 ± 116	7665 ± 198
13.3	2640 ± 171	7814 ± 256

The slope for absorbance vs. *p*-nitrophenol concentration remains relatively stable between ~2200-2650 $M^{-1}cm^{-1}$ across all pH values at 348 nm, displaying no clear or significant linear trend (Figure 3.8). 348 nm is the isobestic point for *p*-nitrophenol (Hriscu et al., 2013), meaning that the molar extinction co-efficient, ε, remains constant over the entire pH range. The slope of the standard curve represents εl; the extinction co-efficient multiplied by the path length (l), which in this case is a constant 0.5 cm. εl at 405 nm however, shows a definite trend as pH increases. At pH 3.5 the slope is negative; we can assume this to be equal to zero, as the linear regression software determines the slope as non-significantly deviant from zero. As pH increases a steady slope is seen until pH 6.0, where the slope of the line increases dramatically. This dramatic increase is seen until pH 8.5. Increases in pH above 8.5 show increasing slope but more slowly than previously, demonstrating a sigmoidal relationship.

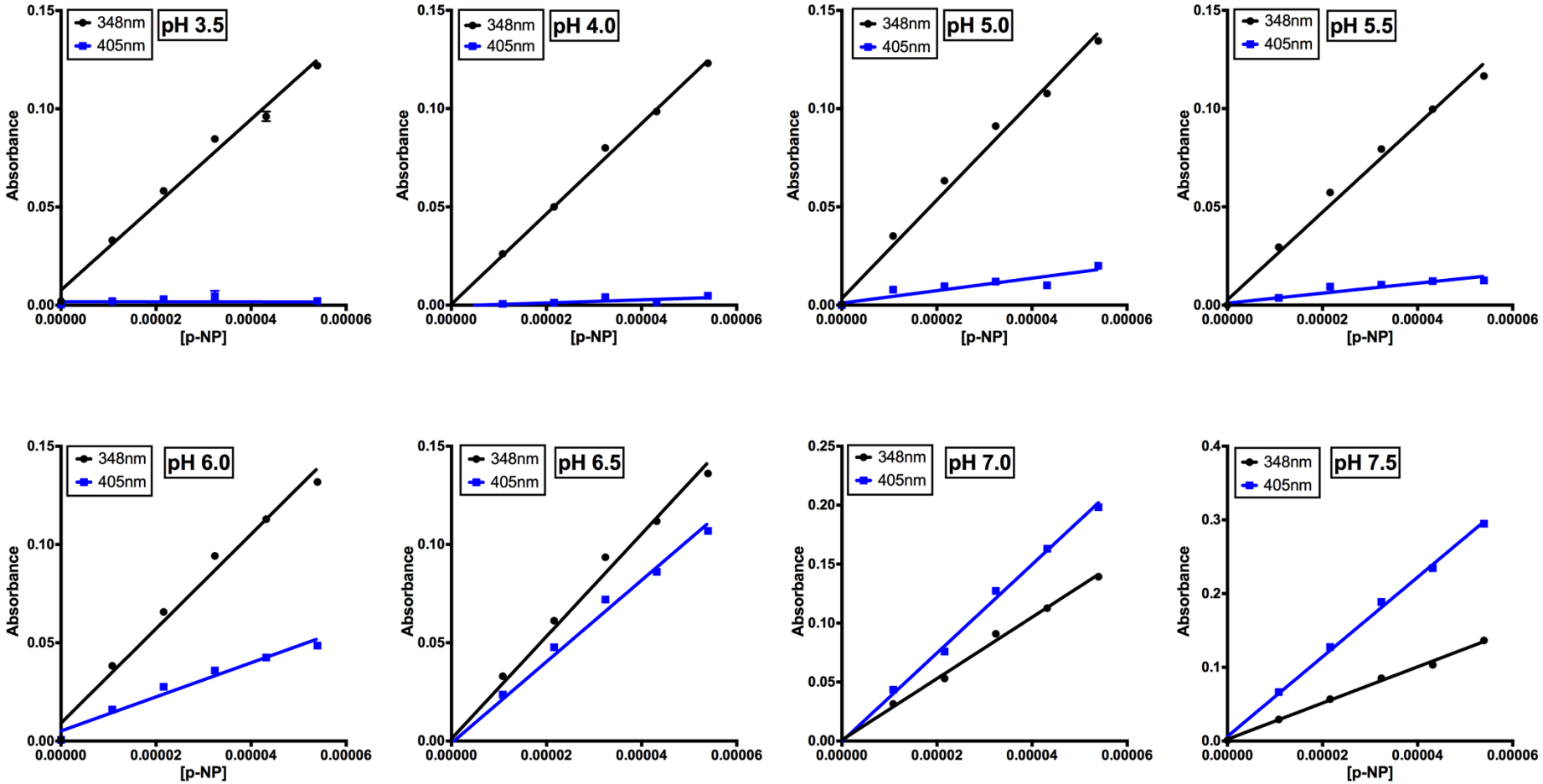

Figure 3.6. Standard pH versus p-nitrophenol concentration curves for pH 3.5-7.0.

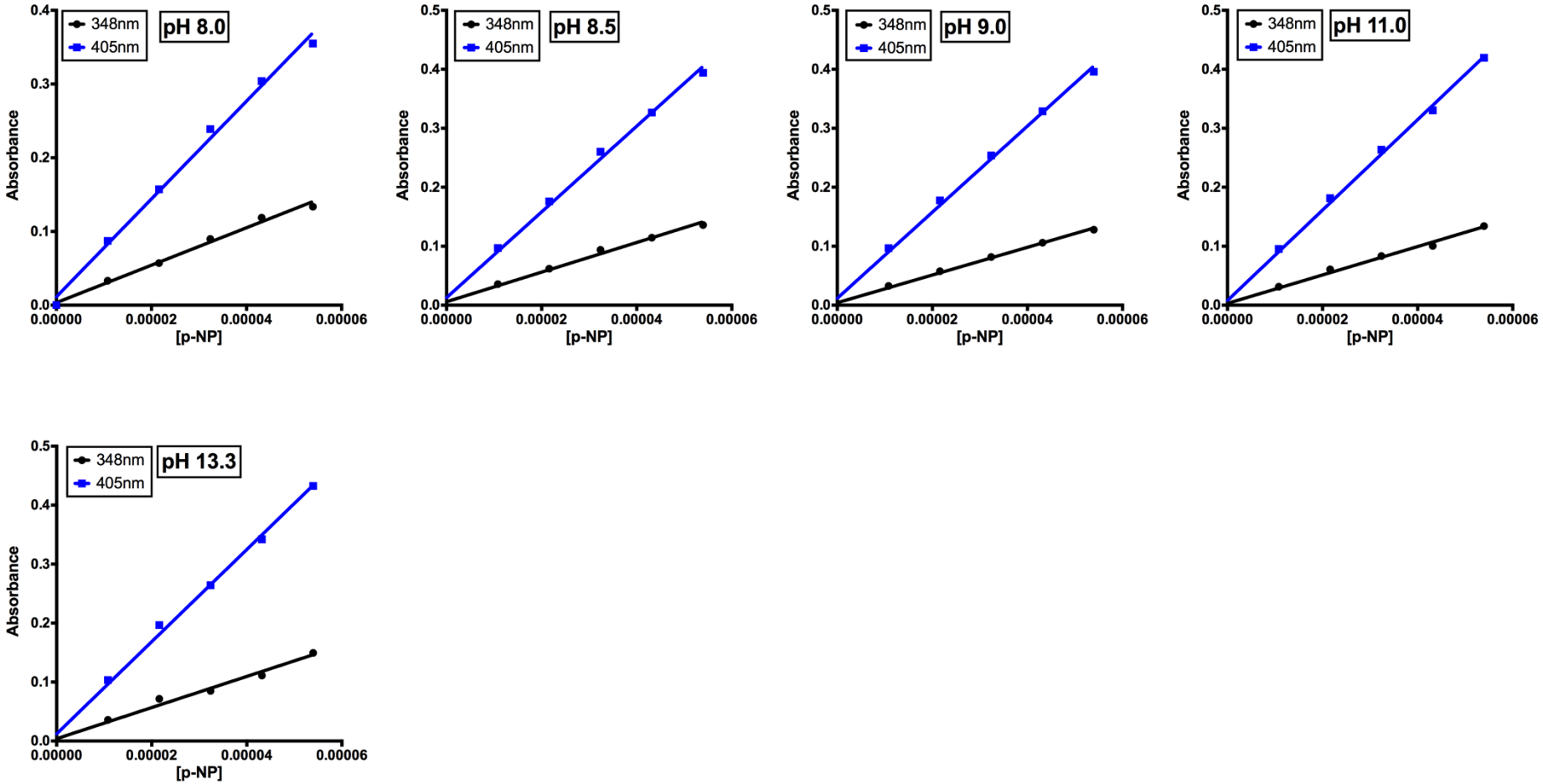

Figure 3.7. Standard pH versus p-nitrophenol concentration curves for pH 8.0-13.3

3.2.4 Determination of ε over the given pH range

The slope of each standard curve given by the linear regression analysis gave a value for εl, (according to Equation 2.1) where l is the path length in cm. The path length used in these experiments was 0.5 cm. Division of the slope of each standard curve by l gave values for ε in $mM^{-1}cm^{-1}$ (Table 2). Values for ε were graphed against pH to generate a curve displaying the pH variance of ε at 405 nm, and the variance in absorbance readings the isobestic point (348 nm) (Figure 3.8). Linear and non-linear regression analysis was conducted using GraphPad Prism version 6.0g for Mac (GraphPad Software, La Jolla California USA) for values of ε at 348 nm and 405 nm respectively. The relationship established for each generated an equation for calculation of values at a given pH (Equation 3.1).

At the isobestic point (348 nm), very little variance is seen in the values for ε over the pH range, allowing a linear relationship to be established with a slope determined as not significantly deviant from zero by linear regression (Figure 3.8). At 405 nm, a sigmoidal relationship is seen for ε versus pH (Figure 3.8), that is, in the pH range 6.0-8.5 any small change in pH can have a significant effect on the value of ε. This indicates that k_{cat} calculations will be systematically incorrect for pH values outside of the linear portion of the sigmoidal curve. ε values cited in literature are generally between pH 13.0 and 14.0, thus the use of these literature values here is inappropriate.

Equation 3.1. Molar extinction co-efficient equation

$$\varepsilon = 33.57 + \frac{15301.43}{[0.8325(7.037 - pH)]}$$

The R^2 value for the fit of this equation was 0.9991, giving very good confidence in its ability to calculate ε at a given pH. The ε value for *p*-nitrophenol used in k_{cat} calculations for all assays conducted at pH 7.0 was therefore taken as 7413 $M^{-1}cm^{-1}$, as determined using the equation above. Values of ε for the pH range 3-14 as determined by Equation 3.1 are summarised in Table 3. For pH activity characterisation curves all raw rate values were corrected for differences in ε to give a reliable k_{obs} versus pH curve.

Table 2. Experimentally determined molar extinction co-efficient of p-nitrophenol at 348 and 405 nm

pH	ε (348 nm) ($M^{-1}cm^{-1}$)	ε (405 nm) ($M^{-1}cm^{-1}$)
3.5	4346	~0
4.0	4596	156
5.0	5012	633
5.5	4460	506
6.0	4806	1736
6.5	5210	4144
7.0	5202	7502
7.5	4914	10802
8.0	5082	13270
8.5	5024	14538
9.0	4696	14600
11.0	4820	15330
13.3	5280	15628
Slope	0.03	Sigmoidal

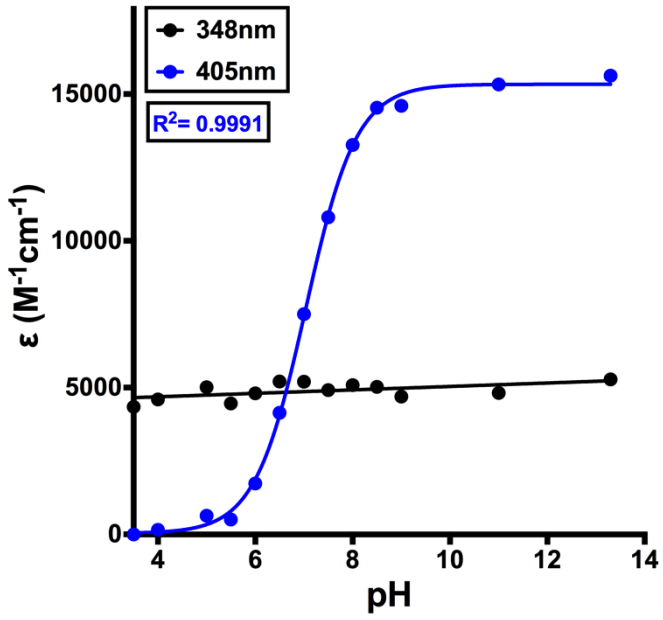

Figure 3.8. Molar extinction co-efficient of p-nitrophenol versus pH at 348 nm and 405 nm.

Table 3. Molar extinction co-efficient values of p-nitrophenol at 405 nm as determined by Equation 3.1.

pH	ε (405 nm) ($M^{-1}cm^{-1}$)
3.0	40
4.0	79
5.0	336
5.5	797
6.0	1877
6.5	4061
7.0	7413
7.5	10873
8.0	13249
8.5	14462
9.0	14988
10.0	15283
11.0	15327
12.0	15334
12.0	15335
14.0	15335

3.3 Deuterium exchange assay for KSI and MalL

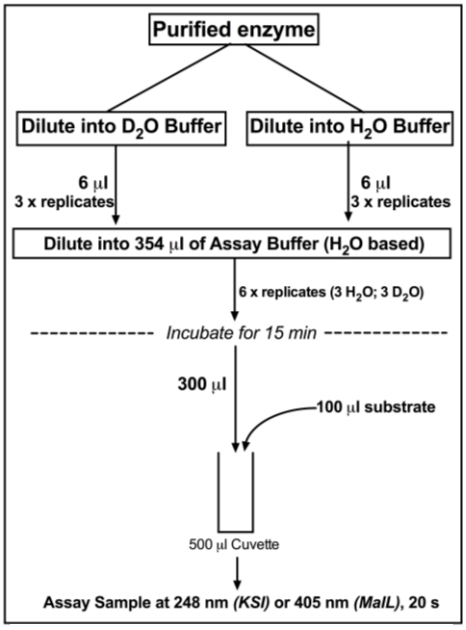

Figure 3.9. Workflow diagram for deuterium exchange assay protocol.

The initial method for determining differences in catalytic rate due to vibrational mode alterations was a forwards deuterium exchange assay compared to non-exchanged controls. This experiment required a time dependent deuterium exchange (or relevant hydrogen control) of the purified enzyme sample, followed by a fixed time reverse exchange prior to performing the assay. All steps were temperature controlled. The workflow for these experiments can be seen in Figure 3.9. The

objective of this experiment was to visualise any alterations in catalytic rate correlated to the increasing replacement of exchangeable hydrogens with deuterium atoms over a time course.

To accurately determine differences in rate between the H_2O and D_2O samples, exchange buffers of the exact same composition and pH were required, requiring extreme care in weighing of buffer components and pH measurements. Initially, pH 5.0 citrate buffers were used. This required the use of a pH correction for the measurement of D_2O using an H_2O calibrated pH meter (Equation 3.2). This correction was obtained from Krężel and Bal (2004). R_{pH} is the '*real*' pH of the D_2O solution, and R_{pH*} is the observed pH seen on the pH meter. Following the shift to pH 7.0 phosphate buffers, the exact weight of each buffer component was calculated to give the same pH to avoid pH measurement and correction.

Equation 3.2. D2O pH correction equation.

$$R_{pH} = 0.929R_{pH^*} + 0.42$$

A 10 μL aquilot of purified enzyme was transferred into 1.5 ml Eppendorf tube containing 290 μL of either H_2O or D_2O exchange buffer in triplicate to give three deuterium and three hydrogen (control) exchange enzyme stock solutions. These solutions were left to exchange at 18°C. A 6 μL aquilot of this solution was transferred into a new Eppendorf containing 354 μL of assay buffer at various time intervals (0, 1, 2, 3, 4, 5, 6, 7, 8 hrs). Each sample was incubated in an Eppendorf Thermomixer (Eppendorf, Germany) at 30 °C for 15 min following this final exchange step. Following incubation, 300 μL of sample was transferred to a 500 μL quartz cuvette. A 100 μL aquilot of substrate [30 °C; 0.5 mM 19-Nor-androst-5(10)-ene-3,17-dione (KSI); 2 mM *p*-nitrophenol α-D-glucopyranoside (MalL)] was introduced to the cuvette and absorbance [248 nm (KSI); 405 nm (MalL)] over 20 s was measured at 0.125 s intervals using a temperature controlled (30 °C) ThermoSpectronic Heλios spectrophotometer (ThermoScientific, USA). Temperature control of spectrophotometer was achieved using ThermoSpectronic single cell Peltier (Life Sciences, USA). Rate of activity [k_{cat} (s^{-1})] was determined for each sample replicate over a 10 s period using a linear regression analysis, Beer-Lambert law (Equation 2.1) and enzyme concentration. These values were graphed against time. Slope deviation from zero was determined for both deuterium exchanged and non-deuterium exchanged

samples using the linear regression tool in GraphPad Prism version 6.0g for Mac (GraphPad, La Jolla, California).

The initial conditions used in this protocol were modified significantly over the course of the experiments. Variations included substrate concentration, substrate solvent, enzyme concentration, exchange and assay temperatures, incubation period, and pH. These differences are summarised in Table of Appendix III.

Each exchange assay generated a rate versus time plot (Figure 3.13). KSI "exchange 5" is not shown as only three time points were gathered and this was not considered insufficient for a complete experiment.

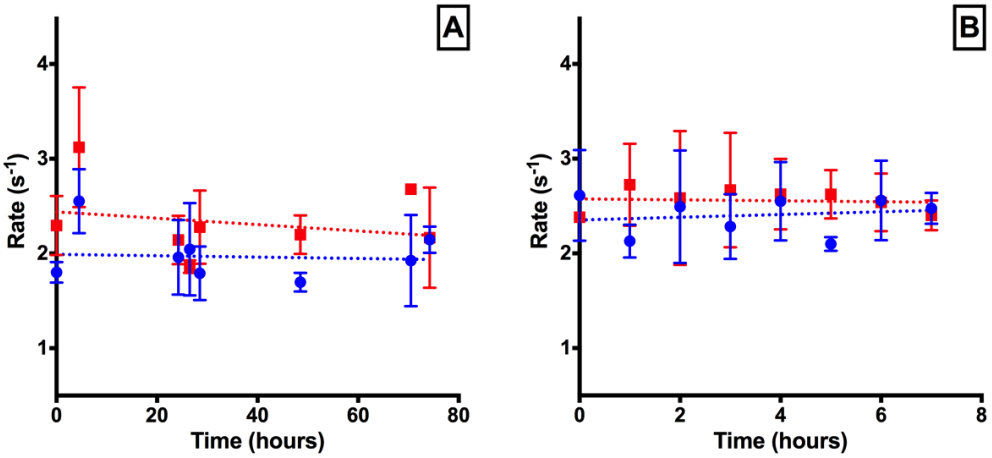

Figure 3.10. Rate [$k_{cat}(s^{-1})$] vs. time for deuterium exchanged (red) and non-deuterium exchanged (blue) KSI (purification 1) using 1.5 mM [S], final [MeOH] 25%. A: Exchange 1; B: Exchange 2.

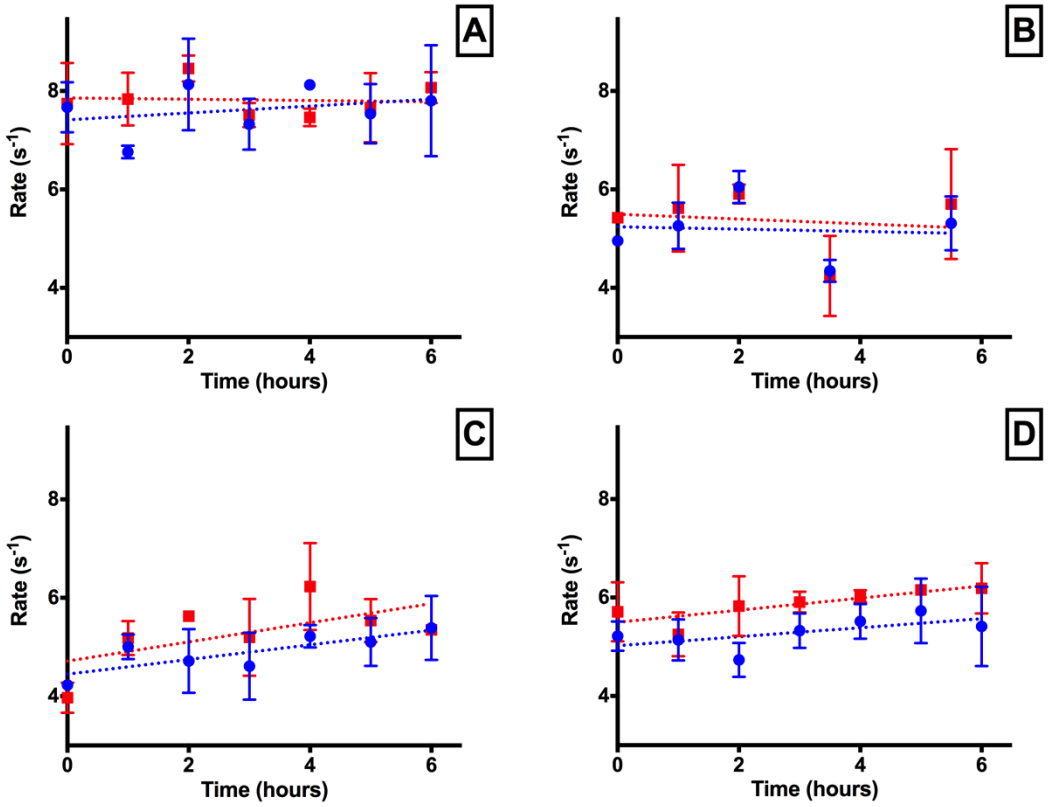

Figure 3.11. Rate [$k_{cat}(s^{-1})$] vs. time for deuterium exchanged (red) and non-deuterium exchanged (blue) KSI (purification 1) using 0.5mM [S], final [MeOH] 5%. A: Exchange 3; B: Exchange 4; C: Exchange 6; D: Exchange 7.

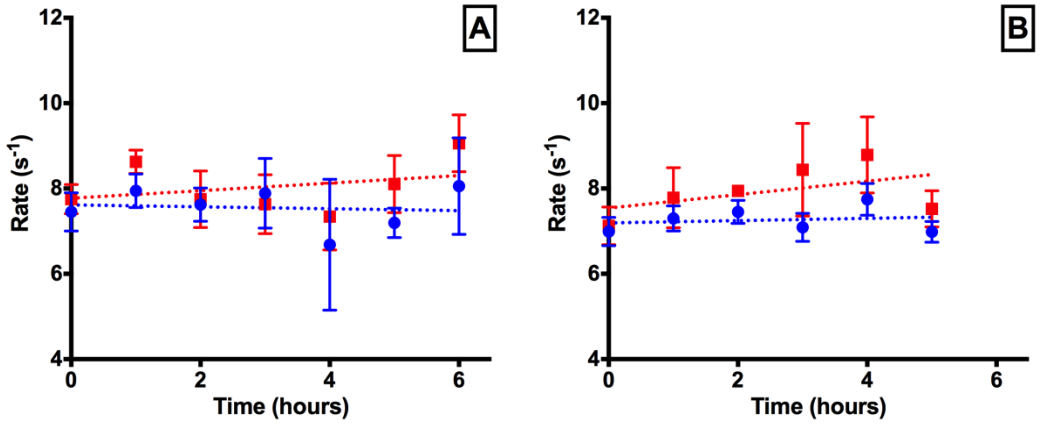

Figure 3.12. Rate [$k_{cat}(s^{-1})$] vs. time for deuterium exchanged (red) and non-deuterium exchanged (blue) KSI (purification 2) using 0.5mM [S], final [MeOH] 5%. A: Exchange 8; B:Exchange 9.

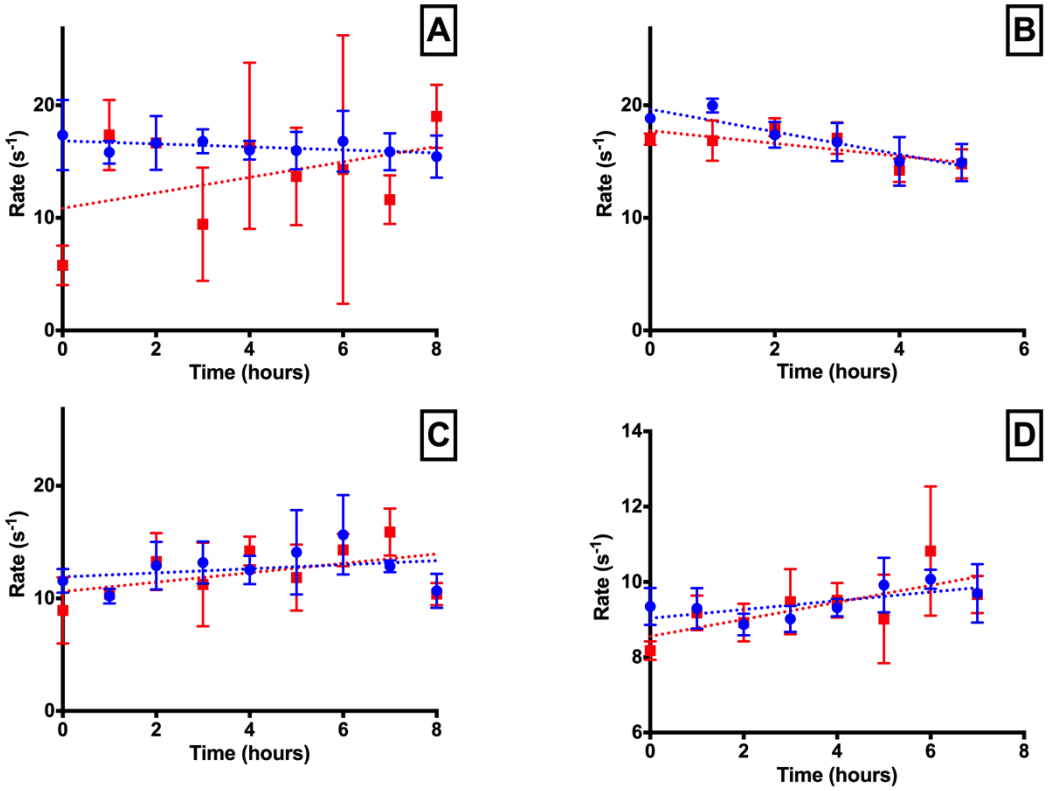

Figure 3.13. Rate [$k_{cat}(s^{-1})$] vs. time for deuterium exchanged (red) and non-deuterium exchanged (blue) KSI (purification 3, A, B, C ; Purification 4, D) using 0.5mM [S], final [MeOH] 5%. A: Exchange 10; B: Exchange 11; C: Exchange 12; D: Exchange 13.

The inconsistency in these results meant that no significant conclusions could be drawn regarding differences in thermodynamic parameters and catalysis between deuterium-exchanged and non-deuterium-exchanged samples. As a result of this, a new experimental protocol for visualising any possible differences had to be determined. It was postulated that temperature profiles of deuterium and non-deuterium exchanged samples of MalL and KSI would provide sufficient data. The exchange period would be extended to provide ample time for a significant exchange to occur prior to rate determination, and the sample would remain undisturbed during the exchange. The exchanged solution would be assayed directly, preventing error introduction by a second exchange step. This method would also allow all replicates to originate from a single sample, reducing the error associated with multiple isolated replicates as in the forwards exchange assay. The overall error within data for temperature profiles was therefore estimated to be much lower. Temperature profiles effectively produced more reliable and consistent data, by producing more data points to allow small differences to be elucidated outside of error range. The protocol for these temperature profiles can be found in Section 2.6.

4 Results

4.1 Protein expression in LB and M9 media

Successful protein expression in LB media was achieved using well established protocols for growth and expression of KSI and MalL in BL21 and DH5α *Escherichia coli* strains respectively (2.1.1.1, 2.1.2.1). Soluble expression in M9 minimal media was achieved using methods established and described in section 3.1.1 to produce $^{13}C^{15}N$ labelled enzyme. Expression was confirmed by SDS-PAGE expression gels for both proteins (Figure 4.1). Representative growth curves were generated for KSI and MalL grown in M9 media (Figure 4.2). Protein products of the expected size are circled.

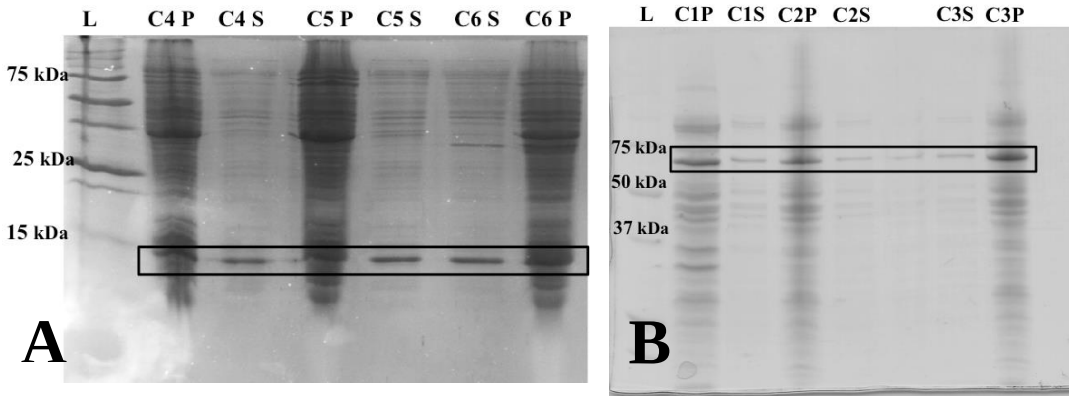

Figure 4.1. SDS-PAGE expression gels showing successful soluble expression of KSI (A) and MalL (B) in M9 minimal media. Culture number is indicated as C1-C6. L=Precision Plus Protein™ standards unstained ladder (Bio-Rad); P= insoluble pellet fraction; S= soluble fraction. Boxes circle products of ~14.9 kDa in gel A, and products of ~ 69 kDa in gel B, the expected size for KSI and MalL respectively.

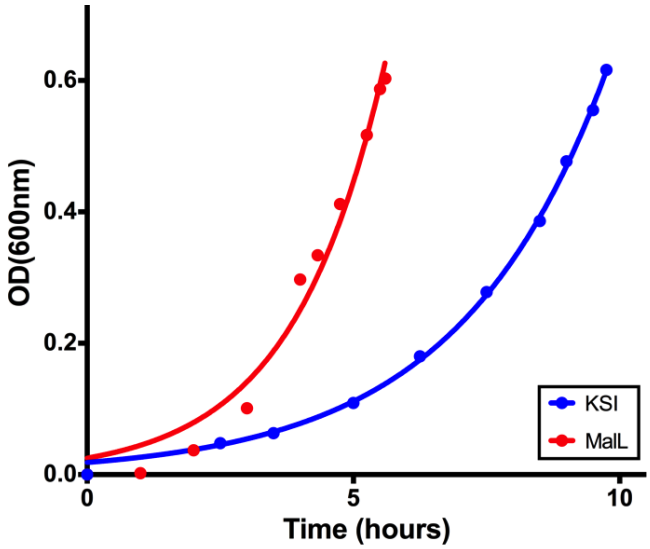

Figure 4.2. Representative growth rate data for E. Coli expressing KSI and MalL in M9 minimal media

Proteins were purified from soluble intracellular fraction by immobilised metal affinity chromatography (IMACS) using nickel affinity for the histidine tag, followed by a size exclusion purification to remove any residual contamination.

4.1.1 IMACS nickel affinity chromatography purification

KSI typically eluted at 0.43 M imidazole while MalL eluted at approximately 0.3 M imidazole. Typically, protein from 20-36 ml (KSI) and 18-28 ml (MalL) elution volume was collected and concentrated for further purification. Representative traces are shown (Figure 4.3A &C).

4.1.2 Size exclusion purification

Typical size exclusion traces for KSI and MalL are shown (Figure 4.3B &D). KSI was typically collected at 14-16 ml elution volume, while MalL typically eluted between 13-16 ml. A smaller peak was seen consistently at 12-13.5 ml. When run on a gel, this peak was slightly smaller, however was confirmed as MalL. These fractions were not collected due to potential degradation.

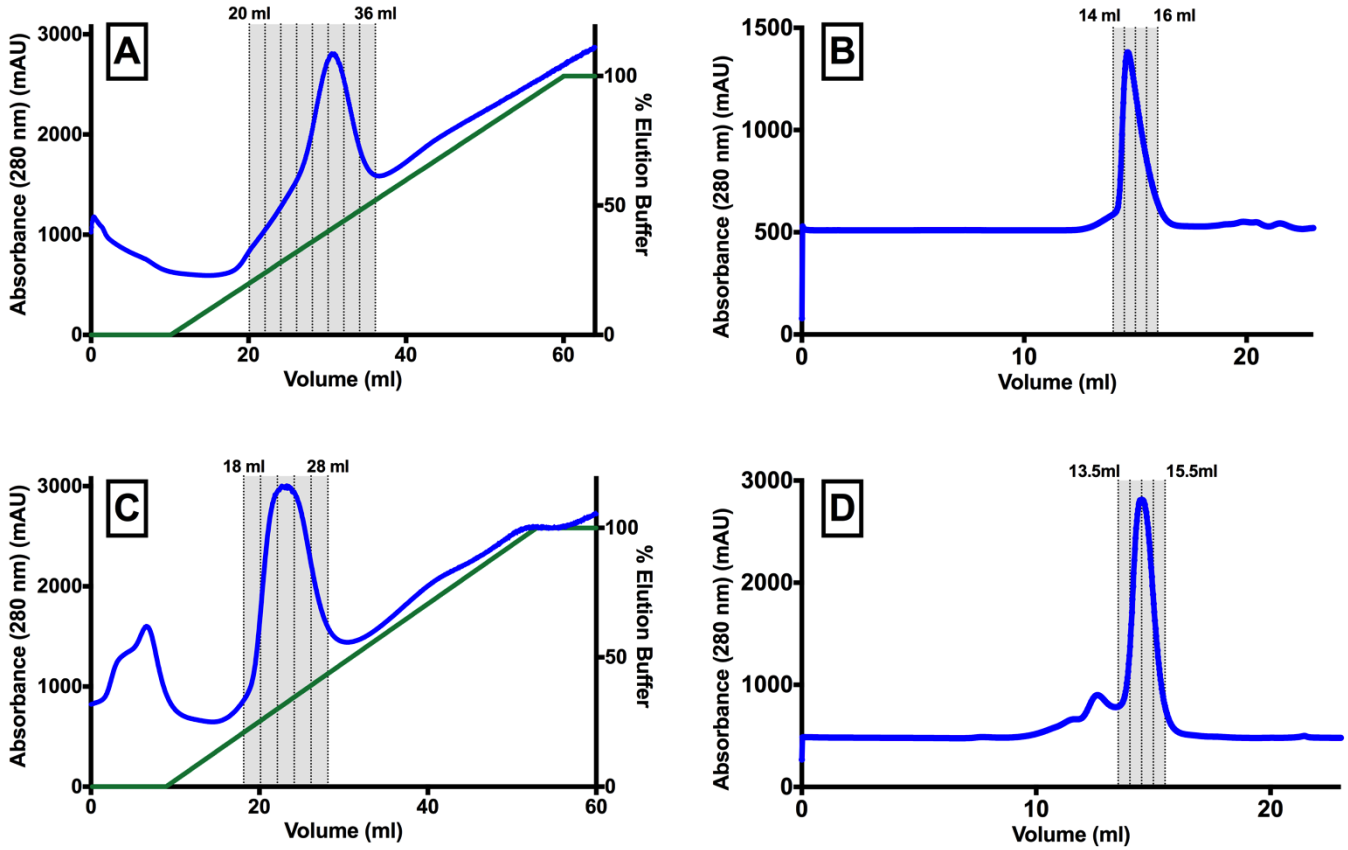

Figure 4.3. Typical purification traces for KSI and MalL showing collected fractions shaded in grey. A:KSI nickel affinity chromatography purification trace; B: KSI size exclusion trace; C: MalL nickel affinity chromatography purification trace; D: MalL size exclusion trace.

4.1.3 Purification SDS-PAGE Gel

Successful purifications of heavy ($^{13}C^{15}N$) KSI and MalL and those MalL WT and V200S purifications used in DSC experiments were confirmed by SDS-PAGE gel (Figure 4.5).

4.2 Protein concentration determination

4.2.1 Bradford assay (KSI)

Figure 4.4 shows a typical Bradford curve for determination of the concentration of KSI. When sample concentration fell outside of the standard curve range, samples were diluted so concentrations could be interpolated. The standard curve equation (Equation 4.1) generated from BSA standards by linear regression analysis has an R^2 value of 0.9753.

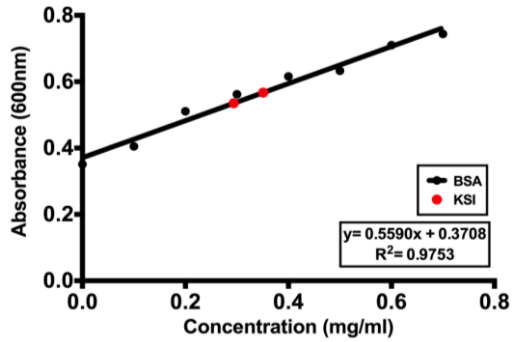

Figure 4.4. Representative Bradford curve. Prism version 6.0g for Mac software (GraphPad, La Jolla California USA) linear regression tool generates equation used in concentration determination. Shows duplicates for dilution of $^{13}C^{15}N$ KSI used in T_{opt} assays.

Equation 4.1. Standard curve equation for Bradford curve of $^{13}C^{15}N$ KSI purification.

$$y = 0.5590x + 0.3708$$

For example, the equation determines this KSI dilution, once duplicates are accounted for, to be 0.32 mg/ml.

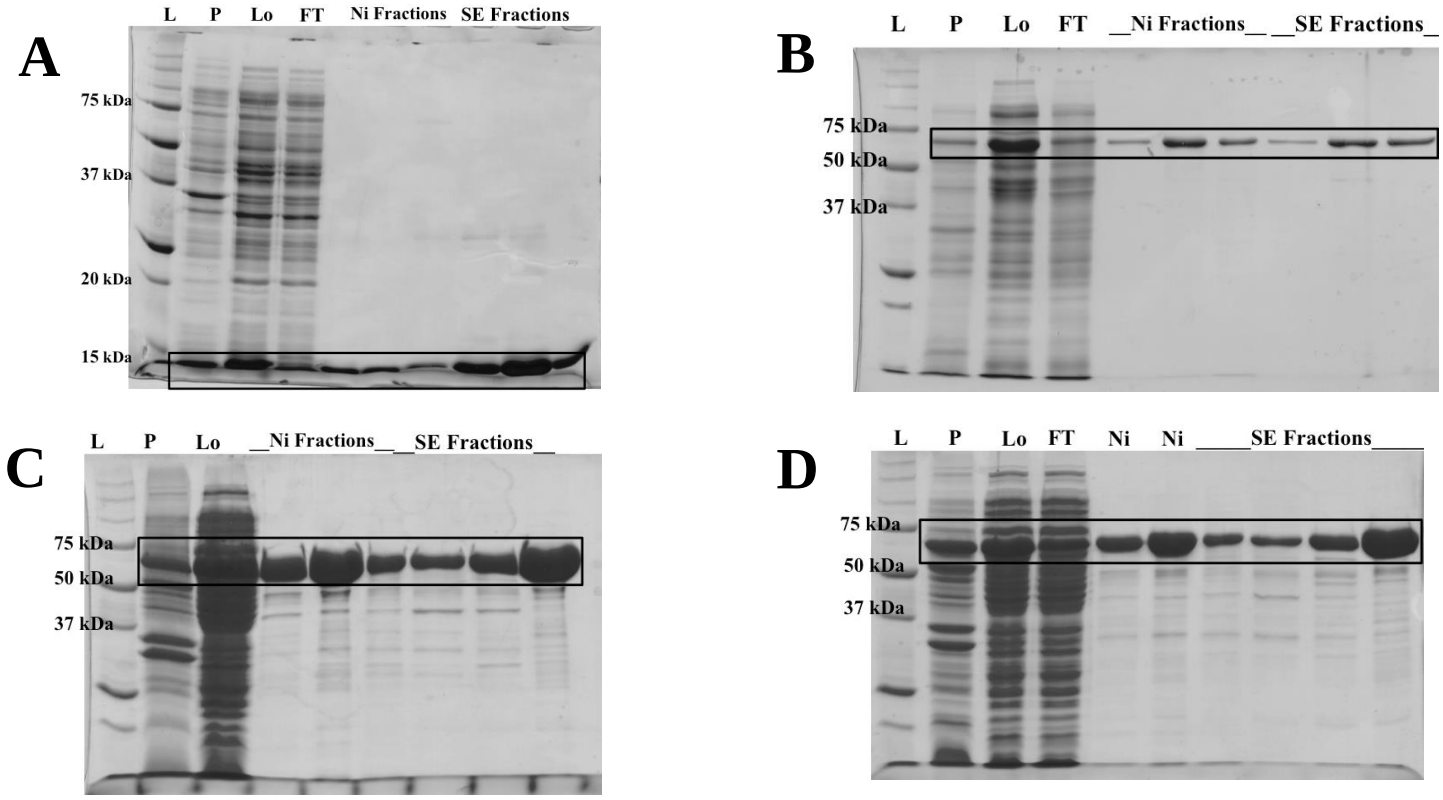

Figure 4.5. SDS-PAGE purification gels. A: $^{13}C^{15}N^{5}$ KSI purification; B: $^{13}C^{15}N$ MalL purification; C: Purification of WT MalL for use in DSC experiments; D: Purification of V200S MalL for use in DSC experiments. In all Cases, L= Precision Plus Protein™ standards unstained ladder (Bio-Rad), P= insoluble pellet fraction, Lo= load onto nickel chromatography column, FT= flow through. Nickel and size exclusion purification fractions are labelled. Products of expected sizes are circled. ~14.9 kDa for KSI; and ~69 kDa for MalL.

4.3 Enzyme parameter assays

4.3.1 Michaelis-Menten kinetic analysis

Michaelis-Menten plots were generated for KSI and MalL (Figure 4.6&Figure 4.7) to determine V_{max} and K_m values as determined by non-linear regression Michaelis-Menten fit in GraphPad Prism version 6.0g for Mac (GraphPad Software, La Jolla California USA) (

Table 4).

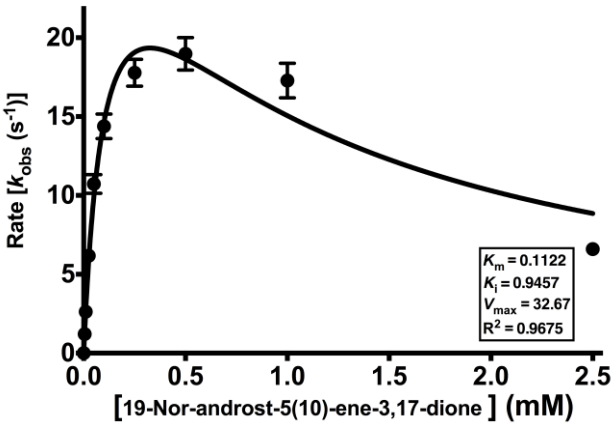

Figure 4.6. Michaelis-Menten plot incorporating substrate inhibition of rate versus substrate concentration for KSI at pH 7.0

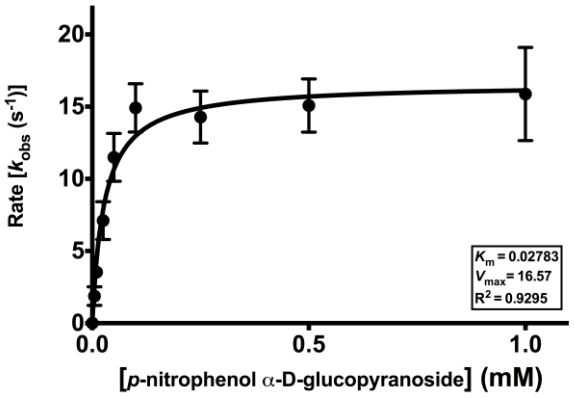

Figure 4.7. Michaelis-Menten plot of rate versus substrate concentration for MalL at pH 7.0

Table 4. Kinetic data generated from Michaelis-Menten non-linear regression fit of k_{obs} versus substrate concentration

Enzyme	**V_{max} (s^{-1})**	**K_m (mM)**	**K_i (mM)**
KSI	32.67	0.11220	0.9457
MalL	16.57	0.02783	-

K_m defines the substrate concentration at which half of the enzyme is saturated. The K_m value for MalL is low, ~0.03 mM, indicating that at pH 7.0, MalL is binding the substrate moderately tightly. It was determined from these data that future experiments should use a substrate concentration of 2 mM, almost 100 times greater than the value of K_m. At this concentration it is highly unlikely that the substrate will become limiting, even in different conditions such as those used in the pH optima assays. The Michaelis-Menten plot for KSI shows significant substrate inhibition, and has thus been fit with the appropriate inhibition curve. The K_m for KSI has been determined as ~0.1 mM at pH 7.0, however the strong inhibition above 0.75 mM substrate indicates that the substrate in excess will hinder the reaction. The inhibitory constant, K_i, generated from the curve is 0.9 mM, describing the concentration at which half maximum inhibition is seen. The substrate concentration with the highest activity was 0.5 mM, and the rate remains linear over a 20 s time period as required for future assays. It was concluded that 0.5 mM was the appropriate substrate concentration for subsequent assays. This concentration is ~5 times greater than K_m so should not become limiting over the course of the reaction, and well below the value for K_i.

4.3.2 pH activity assay

The pH optimum for KSI was determined to be ~8.0 (Figure 4.8), however the error in this data point was much larger than for other pH values. Deuterium exchange assays required a slow exchange where easily measurable rates could still be achieved, in order to visualise the slow exchange of the sample over time, but still obtain reasonable data. A lower pH slows the exchange of hydrogen for deuterium, thus the pH of buffers used for these experiments was altered to be pH 5.0 based on this curve (Figure 4.8). When forward exchange assays were determined as inappropriate to assess the hypothesis, the pH of further experiments was shifted to pH 7.0 to be working near the optimum for KSI where measured rates showed little error in pH characterisation.

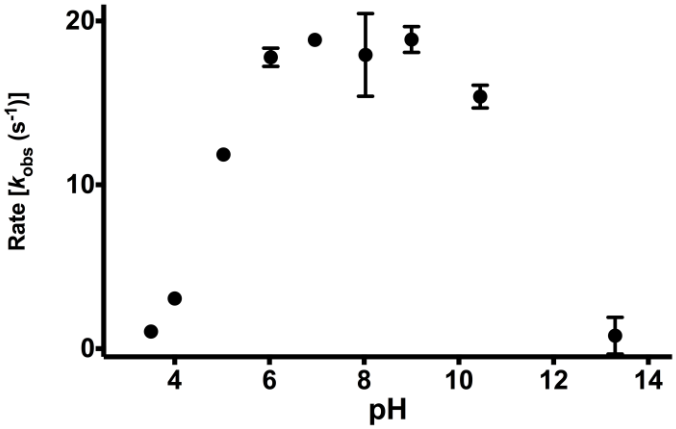

Figure 4.8. Rate [k_{obs} (s^{-1})] versus pH for KSI. [Substrate]:0.5 mM, [KSI]:0.196 mg/ml. Error bars represent the spread of three replicates.

Two curves were generated for MalL: raw absorbance and corrected activity rate versus pH data (Figure 4.9). Corrected data has been converted to k_{obs} (s^{-1}) values using the appropriate correction for *p*-nitrophenol at each pH. This correction shifts the observed pH optima from ~8.25 to ~7.53. As the pH moves outside of the optimum range, rate determination becomes more difficult as raw data become more skewed by either an extremely high or an extremely low molar extinction co-efficient. For this reason, pH 7.0 was determined as an appropriate pH for following assays. pH 7.0 shows good enzyme activity and also gives good measurable raw data output to reduce the error between replicates.

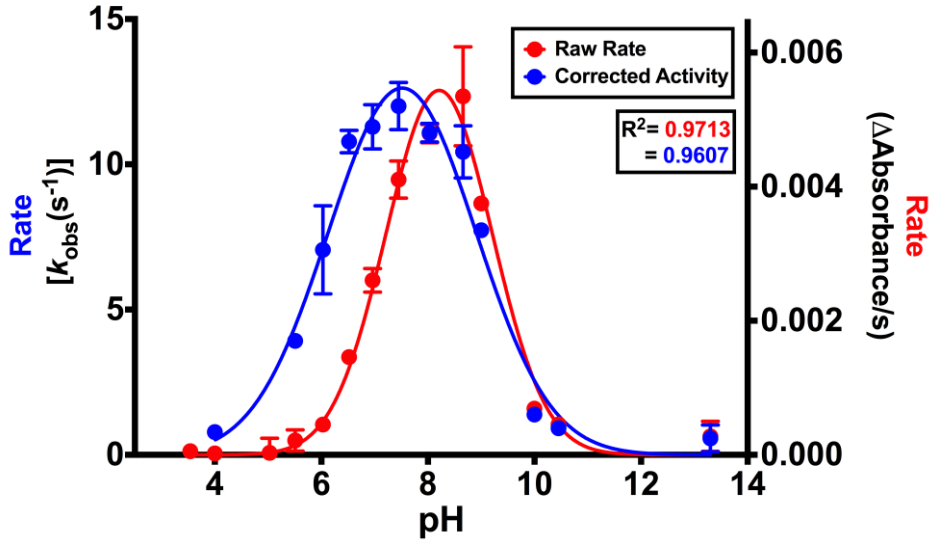

Figure 4.9. Rate versus pH data for MalL. Red data set shows raw absorbance per second rates where blue data set has been corrected for pH differences in molar extinction co-efficient to give k_{obs} values.

4.4 Isotope labelled and unlabelled temperature profiles fit with MMRT

4.4.1 KSI isotope labelled and unlabelled temperature profiles

The MMRT equation was fit to ln(rate) versus temperature data (Figure 4.10) to generate values for the change in heat capacity (ΔC_p) (Table 5). T_0 is 315 K in all cases. Goodness of fit (R^2) values are shown .

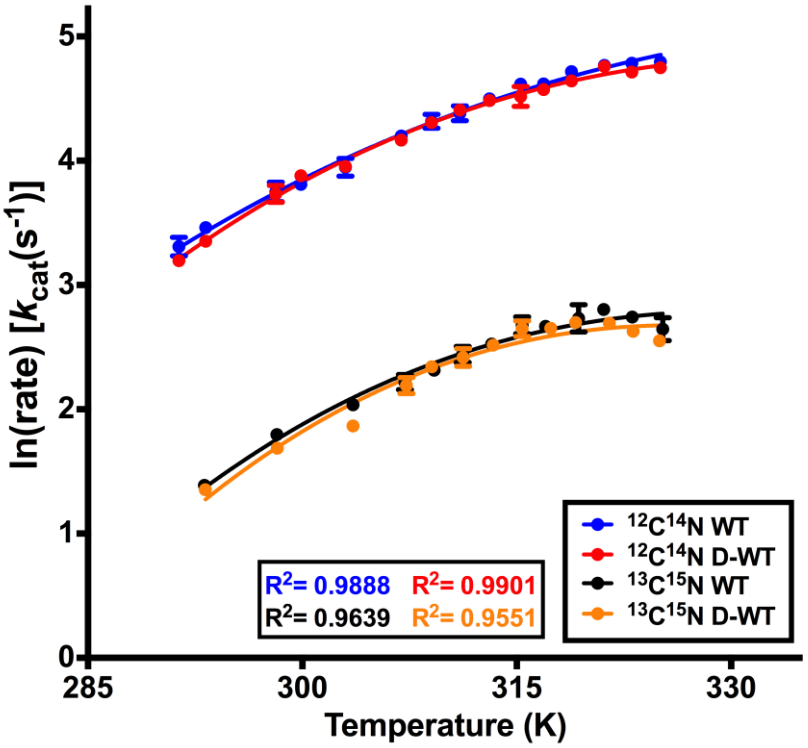

Figure 4.10. ln(rate)[$k_{cat}(s^{-1})$] versus temperature for deuterium exchanged (D-WT) and non-deuterium exchanged WT ($^{12}C^{14}N$), and partially heavy ($^{13}C^{15}N$) KSI.

Table 5. $\Delta C_p^{\ddagger}$ values for deuterium and non-deuterium exchanged WT and $^{13}C^{15}N$ -WT KSI.

KSI	$\Delta C_p^{\ddagger}$ (J.mol^{-1}K^{-1})
WT	- 860 ± 100
D-WT	- 1300 ± 100
WT $^{13}C^{15}N$	- 1700 ± 200
D-WT $^{13}C^{15}N$	- 2000 ± 300

There is a significant decrease in rate when comparing the $^{13}C^{15}N$ enzyme data to the WT $^{12}C^{14}N$ data. This indicates that decreasing vibrational mode frequencies

causes a decrease in rate. The change in heat capacity ($\Delta C_p^{\ddagger}$) decreases and becomes more negative as the enzyme gains mass. A more negative value is seen for $\Delta C_p^{\ddagger}$ for the $^{13}C^{15}N$ enzyme than for the $^{12}C^{14}N$ WT. A decrease in total value to a more negative $\Delta C_p^{\ddagger}$is seen for deuterium-exchanged samples over the WT. For WT $^{12}C^{14}N$ KSI, these differences are outside of error range, demonstrating significance. The difference between WT and $^{13}C^{15}N$ KSI is also outside of error range, however the deuterium exchanged $^{13}C^{15}N$ $\Delta C_p^{\ddagger}$ value is within error for the non-deuterium exchanged sample. WT and partially heavy variants can be visualised in separate graphs for ease in Figure 4.11 as A and B respectively.

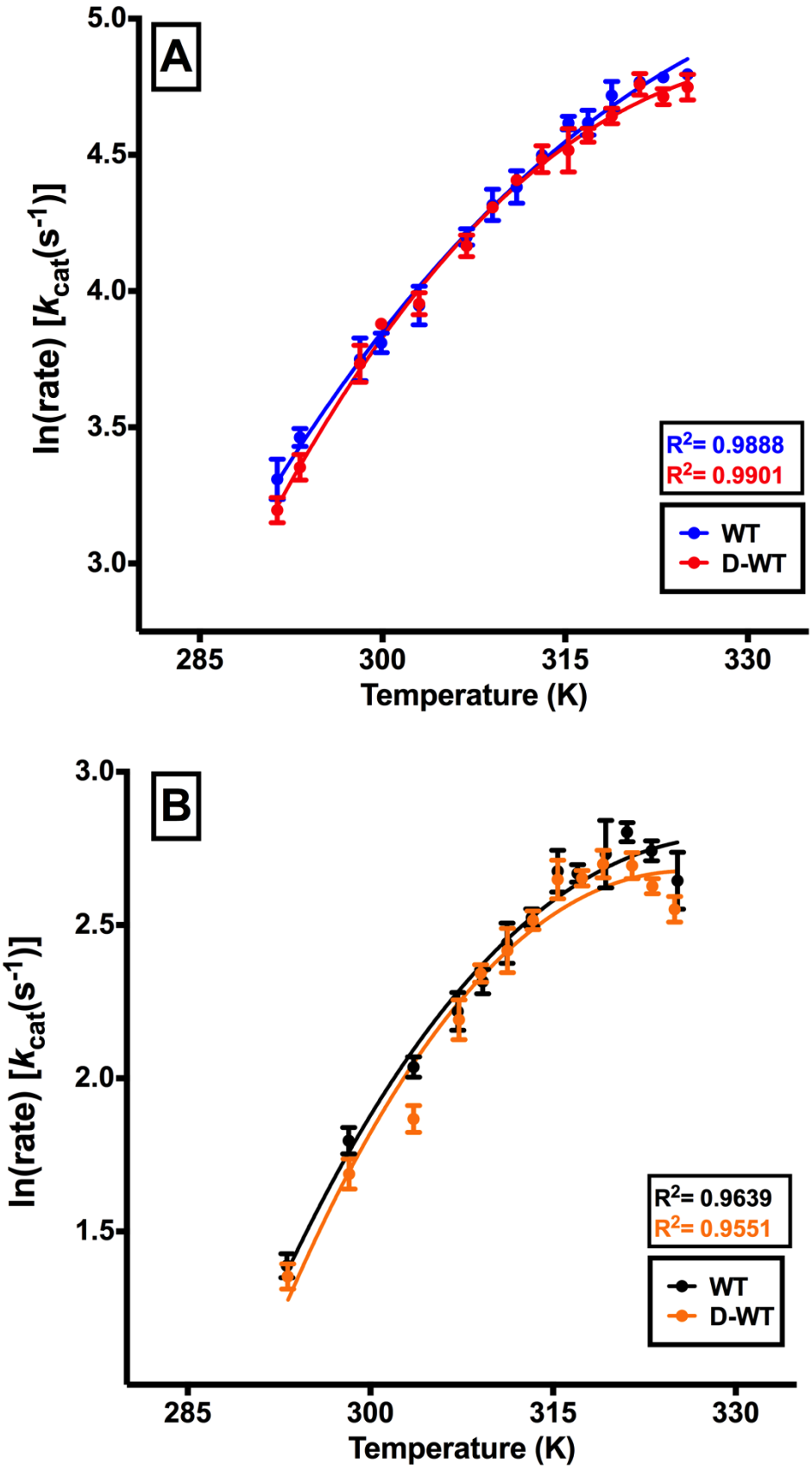

Figure 4.11. ln(rate) [k_{cat} (s^{-1})] versus temperature for KSI. A: WT and deuterium-exchanged (D-WT) for native $^{12}C^{14}N$ KSI.; B: WT and D-WT for partially heavy $^{13}C^{15}N$ KSI.

4.4.2 MalL isotope labelled and unlabelled temperature profiles

MalL data was originally generated for the temperature range 299-327 K for the WT enzyme and was fit with MMRT (T_0 = 315 K) (Figure 4.12A), based on the protocol by Hobbs *et al.* (2013). Following analysis of the data it was determined that the temperature range should be extended to reach 285 K. The R^2 value was ~0.96, however there was a non-random distribution of residuals about the line of fit, suggesting that MMRT is not providing an accurate model over wide temperature ranges in this case (Figure 4.15B). A full data set was gathered over 285-327 K and here, a more pronounced non-random distribution of residuals (from the fit of MMRT to the data) is evident (T_0 = 315 K) (Figure 4.12A&B).

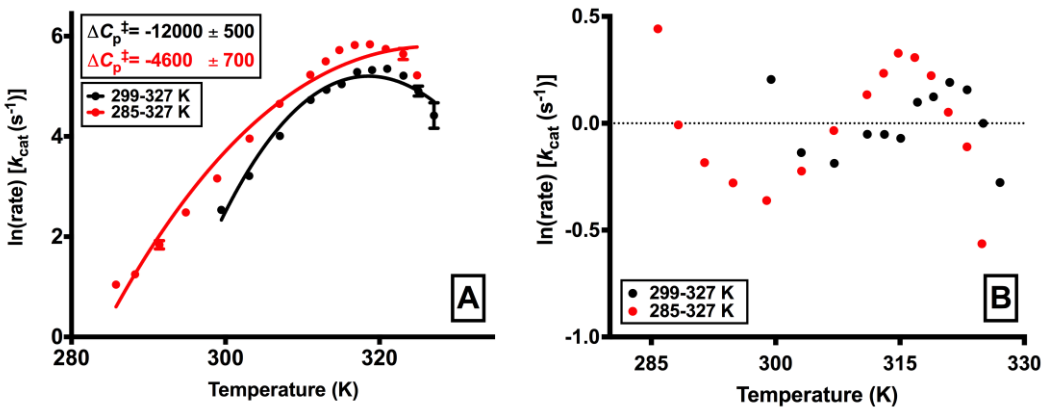

Figure 4.12. Deviation analysis of MMRT for MalL data. A: ln(rate) versus temperature data for WT MalL over narrow (black) and wide (red) temperature range showing deviation of data points from fit. B Residuals plot of MMRT fit for half scale and full scale WT MalL temperature profiles.

The two data sets provide similar curvature over the shared temperature range, however, the skewed fit generates significantly different values for $\Delta C_p^{\ddagger}$ (Figure 4.12A). These values differ by a factor of three indicating an inappropriate model for determining accurate thermodynamic parameters. It was hypothesised that the systematic difference in curvature was the result of temperature dependence of the heat capacity term in MMRT. Consequently, further assays using MalL variants were fit using a newly developed temperature dependent MMRT equation as described in the following section.

4.4.3 Temperature dependence of heat capacity in MalL thermal profiles

The original equation works well to model experimental data generated from small enzymes involved in simple reactions, however assumes the temperature dependence of heat capacity to be negligible. Heat capacity is temperature dependent parameter itself (Prabhu & Sharp, 2005; Tipler, 1998), a notion usually excluded from mathematical considerations due to its limited effect (Liu et al., 2008; Prabhu & Sharp, 2005). When temperature dependence becomes a factor, curvature is altered, particularly when addressing low temperatures where rates begin to approach zero. The MalL WT temperature profiles over a large temperature range appear to display some evidence of this temperature dependence. As a result of these data, a new equation was developed by Professor Vic Arcus to incorporate this function into the MMRT equation. It was decided that by comparing fits of temperature dependent and independent MMRT to MalL data, this hypothesis could be accurately tested.

A heat capacity versus temperature plot assumes a linear relationship (Prabhu & Sharp, 2005), the slope of which denotes the temperature dependence. Heat capacity can thus be mathematically defined as Equation 4.2

Equation 4.2. Temperature dependence of heat capacity.

$$\Delta C_p = A(T - T_0) + B$$

Where A is the temperature dependence of C_p, with B providing an observed reference C_p at T_0. The temperature dependence of heat capacity becomes important in more complex systems as it allows for differences in curvature between low and high temperature. By incorporating this temperature dependent term into the MMRT equation, a robust model for complex protein catalysis is generated (Equation 4.3).

Equation 4.3. Temperature dependent MMRT.

$$k = \frac{k_B T}{h} e^{\left[\frac{-\Delta H^{\ddagger}_{To} - (A(T-T_0)+B)(T-T_0)}{RT}\right] + \left[\frac{\Delta S^{\ddagger}_{To} + (A(T-T_0)+B)(lnT - lnT_0)}{R}\right]}$$

Equation 4.4. Natural log form of temperature dependent MMRT.

$$lnk = ln\frac{k_B T}{h} - \left[\frac{\Delta H^{\ddagger}_{T_0} + (A(T - T_0) + B)(T - T_0)}{RT}\right] + \left[\frac{\Delta S^{\ddagger}_{T_0} + (A(T - T_0) + B)(lnT - lnT_0)}{R}\right]$$

Equation 4.4 was fit to the previous MalL data set (Figure 4.13) along with the original, temperature independent MMRT equation; and residual plots were generated for both fits (Figure 4.13).

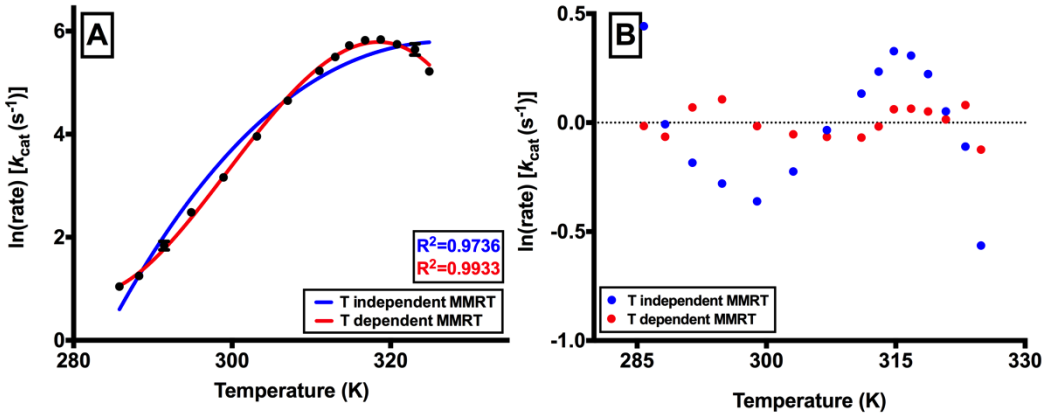

Figure 4.13. A:Ln Rate versus temperature data fit with temperature independent (blue) and dependent (red) MMRT fit to data. B: Residuals plot for both fits.

It is clear that the temperature dependent equation provides a much more accurate fit than the original temperature independent model. Not only is the fit visibly more accurate, but the residual plot shows much less absolute deviation about the fit line. It must be noted that the distribution of points is still non-random for the temperature dependent model. Taken together with the reduced absolute value of the deviations, this suggests that the temperature dependent MMRT model is much better than the previous model, but still demonstrates some inaccuracy for modelling the data. It is expected that this is the result of the temperature dependence of heat capacity being a complex non-linear parameter. A linear function has been used to generate the new MMRT equation, as a curved function would further increase the complexity and provide too many parameters to justify. Additionally, an F test was run for the two models GraphPad Prism version 6.0g for Mac (GraphPad Software, La Jolla California USA) to give a comparison of fit.

The F ratio was determined to be 467.6, and a P value of <0.0001 was given for the analysis. Taken together, these statistics allow rejection of the null hypothesis that temperature independent MMRT is the appropriate curve. It was thus determined based on the differences in fit between the two equations that the temperature dependence of heat capacity was not negligible for MalL. Consequently the remaining MalL assays for the V200S mutant and heavy $^{13}C^{15}N$ WT were fit using the temperature dependent MMRT equation.

4.4.3.1 Function predictions from temperature dependent MMRT

The wide temperature range was repeated experimentally for WT MalL in undeuterated buffer to confirm the pattern (Figure 4.14A). This data set was also fit with the temperature dependent MMRT equation. The function was extended to generate a curve for the temperature range 260-330 K to determine rate predictions beyond the previous experimental range (Figure 4.14B).

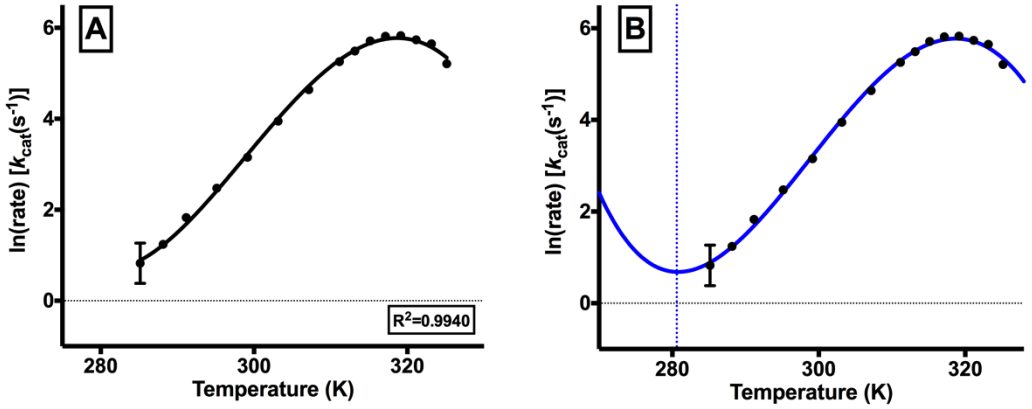

Figure 4.14.MalL wide temperature scale repeated run. A: Data fit with Equation 4.4. B: Function determined from fit in A extended to 260- 330 K temperature range.

Interestingly, the new function predicted a minimum rate would be reached at 280.6 K (dotted line, Figure 4.14), below which an increase in rate would be seen. Further experimental data was collected for temperatures reaching 278 K to assess this possibility (Figure 4.15). No increase in rate was seen below 280 K. The new extended data set was also fit with Equation 4.4, generating another excellent fit. As the temperature approaches 273 K, the rates approach zero as expected thermodynamically. Lower temperature measurements were unable to be made due to instrument limitations, however it is expected that cold denaturation processes would likely have an effect at temperatures beyond those observed experimentally.

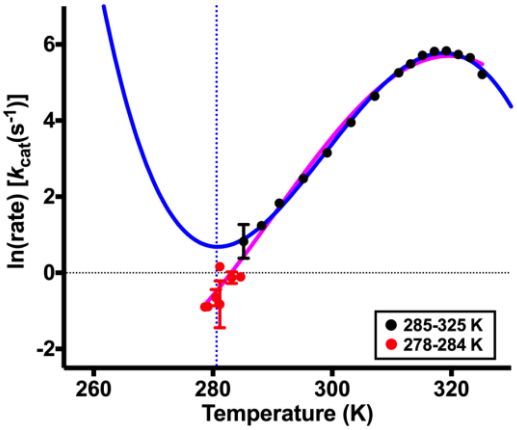

Figure 4.15. MalL extended function data set extended to 278 K. Additional data points gathered for 278-284 K are seen in red. Extended function from smaller data set seen in blue. Temperature dependent MMRT fit seen in pink.

4.4.3.2 Temperature dependent MMRT fit to existing data sets for additional enzyme data

To increase confidence in the experimental conclusions, and to broaden this explanation beyond MalL, a rate versus temperature data set containing four different enzymes was obtained from Peterson *et al*. (2004)(Peterson et al., 2004b), and fit with temperature dependent MMRT (Figure 4.16). A residual plot for all sets was also generated (Figure 4.17).

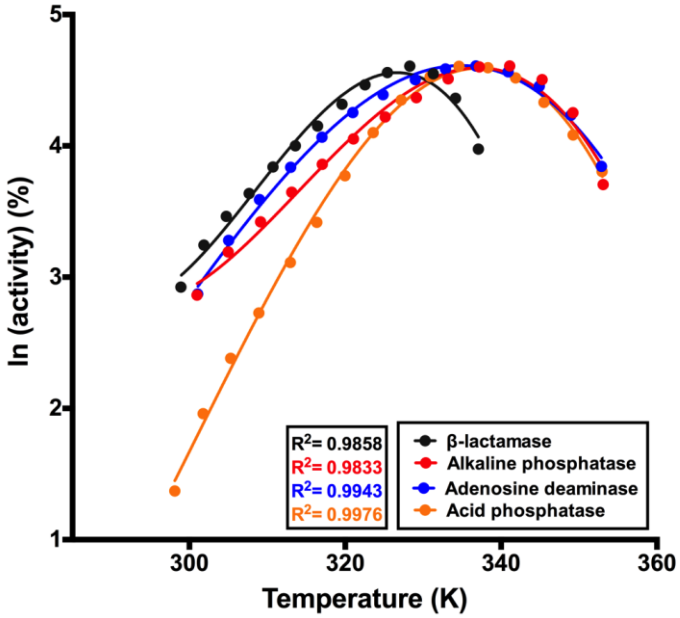

Figure 4.16. ln(activity) (%) versus temperature showing fit with temperature dependent MMRT. Data from Peterson et al. (2004).

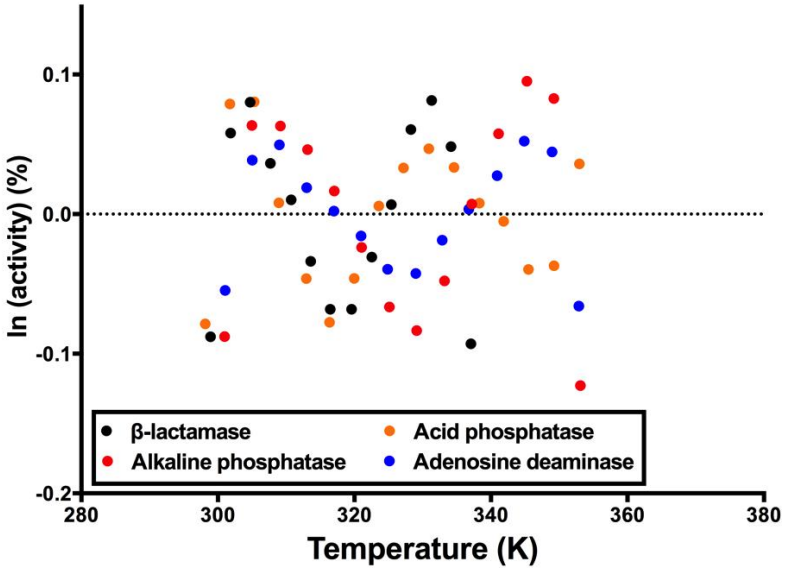

Figure 4.17. Residuals plot for temperature dependent MMRT fit from Peterson et al. (2004).

The fit for each of these four data sets is excellent, with each displaying an R^2 value above 0.98. The residuals plot shows little deviation for each data set, with most deviations falling within 0.1%. The spread of these points around the centre zero line is non-random, suggesting some inaccuracy in the ability of the model to fit the data presumably for the same reasons discussed for the MalL data (non-linear heat capacity temperature dependence). The small absolute values of these residual errors however, indicates temperature dependent MMRT is providing a good fit for the data.

4.4.3.3 Temperature dependent MMRT fit to MalL experimental data

Ln(rate) versus temperature data for MalL variants was fit using Equation 4.4 (Figure 4.18), generating values for both $\Delta C_p^{\ddagger}$, and the temperature dependence (slope) of $\Delta C_p^{\ddagger}$ at T_0 (Table 6).

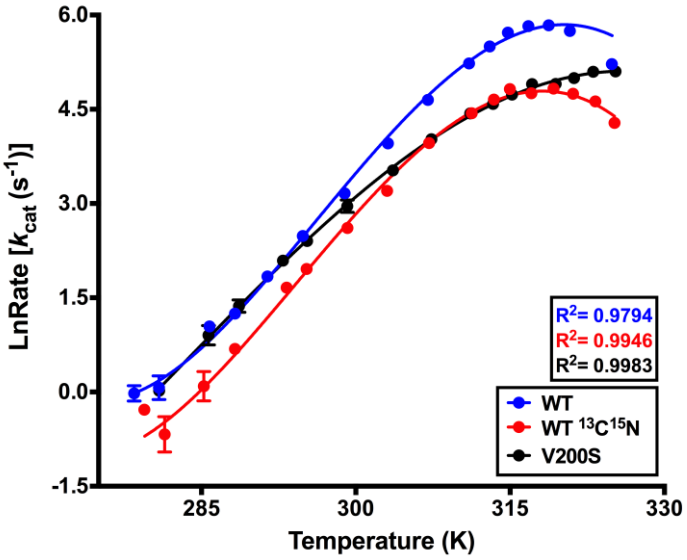

Figure 4.18. ln(k_{cat}) versus temperature for WT, partially heavy ($^{13}C^{15}N$) WT and V200S variants of MalL

Table 6. $\Delta C_p^{\ddagger}$ (at T_0) and temperature dependence of $\Delta C_p^{\ddagger}$ (slope) values for MalL variants seen in Figure 4.18

MalL	$\Delta C_p^{\ddagger}$ (J.mol^{-1}K^{-1})(at T_0)	T-dependence $\Delta C_p^{\ddagger}$ (J.mol^{-1}K^{-2})
WT	-9930 ± 1000	-196 ± 30
WT $^{13}C^{15}N$	-11000 ± 600	-196 ± 20
V200S	-4280 ± 300	- 54 ± 8

As for KSI, there is a significant decrease in rate when comparing the $^{13}C^{15}N$ enzyme data to the WT $^{12}C^{14}N$ data, and the change in heat capacity for the transitions state ($\Delta C_p^{\ddagger}$) decreases and as the enzyme becomes heavier. Standard errors of fit for $\Delta C_p^{\ddagger}$values for MalL WT variants overlap, meaning more work is required to determine exact differences; however the data still conform to the general trend identified in the KSI data. V200S MalL shows a significantly reduced $\Delta C_p^{\ddagger}$ value compared to WT, concurrent with data from Hobbs *et al.* (2013).

KSI has not been fit with temperature dependent heat capacity MMRT as there is little evidence to warrant a temperature dependent $\Delta C_p^{\ddagger}$, likely due to the small size of the molecule and the simplicity of the reaction it catalyzes (Figure 1.). It

should be noted however, that the data does start to shown some evidence of curvature possibly indicative of temperature dependence in the partially heavy forms (Figure 4.10&Figure 4.11B). In order to determine whether this is true temperature dependence, a wider temperature range would need to be included in the analysis to properly define differences, and DSC should be carried out to confirm an absolute trend. For the purposes of this study however, it was more beneficial to treat all systems of KSI with temperature independent MMRT to allow a direct comparison of results. These differences do however indicate that temperature dependence may be correlated with enzyme size.

MalL data has been fit with the the temperature dependent $\Delta C_p^{\ddagger}$ equation, as the data displays differences in curvature between low and high temperatures. This new equation generates a slope value equal to the temperature dependence of $\Delta C_p^{\ddagger}$. This value is the same for WT and partially heavy MalL, indicating that while $\Delta C_p^{\ddagger}$ may alter with changes to vibrational mode energy, the specific temperature dependency of $\Delta C_p^{\ddagger}$ remains constant for the enzyme structure. The slope value for V200S MalL is much smaller than for WT. This indicates a smaller difference in the temperature dependence of the heat capacity between the enzyme transition state complex and the enzyme-substrate complex than for WT.

4.5 Direct evidence of temperature dependent heat capacity

Following the change in model from temperature independent to temperature dependent MMRT, it was decided that thermodynamic evidence of heat capacity temperature dependence should be identified. Differential scanning calorimetry (DSC) of MalL WT and WT + transition state analogue could generate direct heat capacity data over a given temperature range, allowing an accurate determination of heat capacity temperature dependence in the ground and transition states.

DSC was carried out to determine alterations to heat capacity with respect to temperature, giving a slope equal to temperature dependence. Raw data was given as heat energy in μJ/s. Accounting for enzyme concentration, molecular weight and partial specific volume generated an output of molar heat capacity versus temperature for duplicate samples of WT MalL, as well as a single data set for WT with the transition state analogue deoxynojirimycin (DONM) added to

solution at a 1:1 molar ratio (Figure 4.19). Melting temperature (T_m) is shown as a dotted line for WT at approximately 319 K. T_m represents the temperature at which half of the protein sample is denatured. Information following this peak is not relevant in this study as protein is unfolded and aggregated.

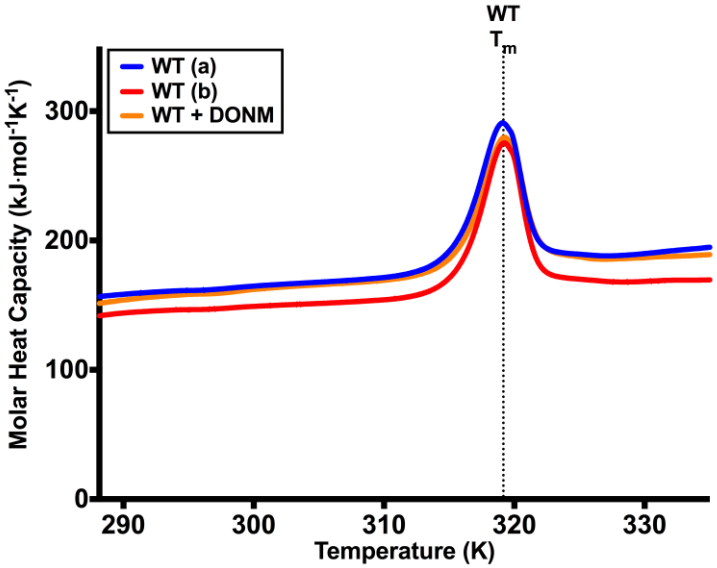

Figure 4.19. DSC data for WT and WT + DONM samples over 288-335 K. T_m values shown by dotted lines.

WT in solution with DONM shows no significant difference to WT only samples, likely due to relatively weak binding in the micromolar range. It was concluded that this data set was unable to provide information on differences in temperature dependence for the transition state versus the unbound enzyme, and for this reason WT + DONM was not considered in further analysis. The data indicates definite temperature dependence for WT, confirming the existence of heat capacity temperature dependence in MalL. A linear regression analysis of the data was conducted in the linear region between 300-310 K for WT in GraphPad Prism version 6.0g for Mac (GraphPad Software, La Jolla California USA) (Table 7).

Table 7. Linear regression results over 300 K-310 K for WT and V200S MalL DSC data

MalL	T. dependence of C_p (kJmol^{-1}K^{-2})	R^2	Average (kJmol^{-1}K^{-2})
WT (a)	0.63 ± 0.002	0.9960	
WT (b)	0.48 ± 0.001	0.9945	0.56

The average slope value is 0.56 kJ.mol^{-1}K^{-2}. To determine whether this value was appropriate, data from Makhatdze (1998)(Makhatadze, 1998), was obtained and

corrected for molecular weight to give a general range for protein heat capacity temperature dependence (

Table 8). The data set contained heat capacity values for a number of amino acids and proteins at various temperatures, however data used in this analysis only included proteins where three or more temperature points were presented. This allowed a linear relationship to be established to give a temperature dependence of heat capacity as the slope. A total of six proteins were included. Most values were determined by DSC, however some data points have been determined by flow calorimetry (FC) or drop calorimetry (DC) (Makhatadze, 1998)

Table 8. Temperature dependence of heat capacity for various proteins as determined by DSC, FC or DC.

Protein	**Molecular weight (Da)**	**T. dependence C_p ($J.mol^{-1}K^{-2}Da^{-1}$)**
BPTI	6565	0.0059
Tendamistat	7960	0.0045
Cytochrome c	12300	0.0069
Barnase	12365	0.0083
Lysozyme	14314	0.0062
α-chymotrypsin	25400	0.0099
MalL WT	69472	0.0080

The six additional proteins from Makhatdze (1998) give a temperature dependence of heat capacity in the range of ~0.005-0.01 $Jmol^{-1}K^{-2}Da^{-1}$. WT MalL is well within this range. Additionally, to determine whether the absolute value for heat capacity generated for WT MalL was appropriate, a theoretical value for total C_p was generated using values for the heat capacity of solid amino acid side chains at 298.15 K including a 15 $Jmol^{-1}K^{-1}$ contribution from each amino acid backbone (Makhatadze, 1998), coupled with the amino acid sequence for WT MalL (Obtained from PDB, 4M56) (Table 9). This generated a total C_p value of ~114 $kJmol^{-1}K^{-1}$, approximately 36% lower to the 155 $kJmol^{-1}K^{-1}$ obtained by averaging the C_p values of MalL generated in DSC experiments.

Table 9. Estimated theoretical heat capacity contributions from solid amino acid heat capacity values

Amino Acid	**Number**	**C_p (J.mol^{-1}K^{-1})**
Ala	24	3294.24
Arg	28	7726.88
Asn	29	6464.1
Asp	47	7998.93
Cys	2	553.84
Gln	24	4780.32
Phe	23	5014.23
Pro	23	3821.91
Ser	34	5119.04
Val	33	6066.06
Glu	44	8362.64
Gly	35	3997
His	14	3703.42
Ile	33	6708.24
Leu	35	7558.6
Lys	40	10156.4
Met	23	7015.92
Thr	18	2921.4
Trp	17	4303.55
Tyr	35	8100.4
Total	**561**	**113667.12**

5 Discussion

5.1 Isotope labelled and unlabelled temperature profiles fit with MMRT

Both KSI and MalL demonstrate a decreasing $\Delta C_p^{\ddagger}$ to become more negative as the overall mass of the enzyme is increased through isotopic labelling. This finding suggests that altered vibrational mode distribution is having an effect on the heat capacity of the system. When the mass of the atoms contained within the enzyme is increased, the vibrations occurring are at lower frequencies (Engel & Reid, 2006). Since low frequency vibrational modes provide the largest contribution to heat capacity, this lowering of total vibrational frequencies alters the total contribution to heat capacity. A consequence of the vibrational mode hypothesis postulates that a decrease in $\Delta C_p^{\ddagger}$, as evidenced by increased curvature in the temperature profile, should correspond to a shift of the enzyme-substrate complex distribution to lower frequencies. Recent data from Professor Vic Arcus (2016), has demonstrated this shift using molecular dynamics simulations of unlabelled versus fully labelled (^{2}H, ^{13}C, ^{15}N) enzyme vibrational mode distributions (Figure 5.1) (Unpublished data). A shift in the distribution of states to lower frequencies is expected to increase the absolute heat capacity value.

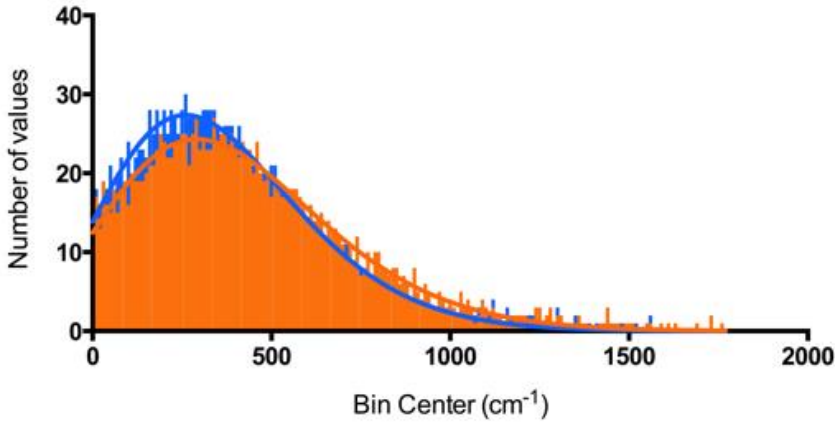

Figure 5.1. Vibrational mode distribution for deuterated (blue) and undeuterated (orange) rigid enzyme system as determined by molecular dynamics simulations (Arcus, Unpublished,).

A more negative $\Delta C_p^{\ddagger}$ indicates a greater difference in heat capacity between the enzyme-substrate complex (ground state) and the transition state complex (transition state). The data implies that the shift in vibrational mode distribution is less pronounced for the transition state than for the ground state, suggesting that

the increase in mass affects the C_p of the ground state more significantly than that of the transition state. The effect of this is to imply the following as a potential model for vibrational modes distribution alterations (Figure 5.2).

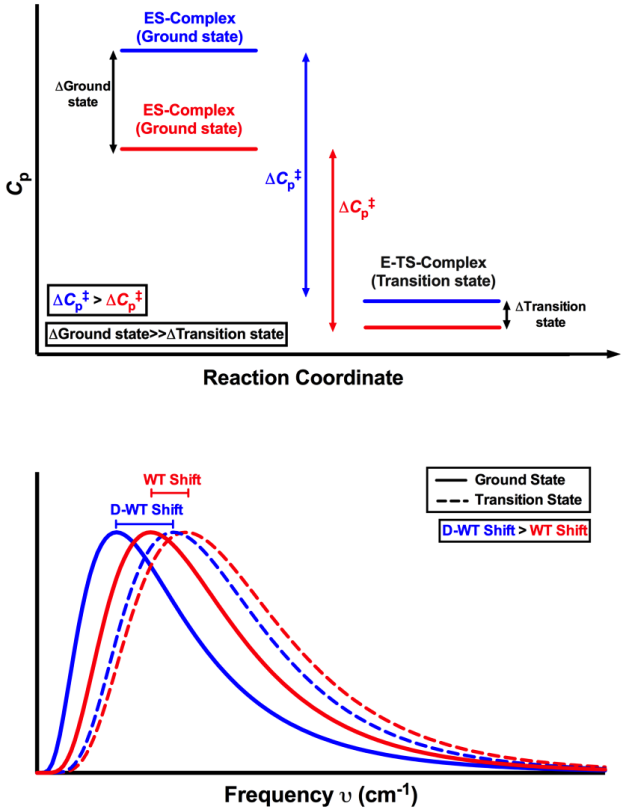

Figure 5.2 Schematic of the changes to $\Delta C_p^{\ddagger}$ and vibrational mode distribution between ground state and transition state for deuterated (blue) and undeuterated (red) enzyme. Proposed model.

5.1.1.1 KSI temperature profiles

The total distribution of vibrational modes is shifted to significantly lower frequencies for the $^{13}C^{15}N$ KSI samples when compared with the WT, as evidenced by the significant drop in overall rate from the WT to the $^{13}C^{15}N$ samples. The data implies that the distribution of the deuterium exchanged WT (D-WT) for both $^{12}C^{14}N$ and $^{13}C^{15}N$ is shifted slightly in comparison to the WT (Figure 5.2). These alterations are outside of standard error range (of fit) for all except the WT and D-WT of $^{13}C^{15}N$, lending confidence to these conclusions. Further experimental work to compare these rates with those from KSI expressed from *E. coli* grown in media containing heavy nitrogen, carbon and hydrogen sources ($^{2}H^{13}C^{15}N$) would provide additional support to establish the trend if this hypothesis is correct. The current data only works to establish a trend for partially heavy systems comparative to the WT '*light*' system. By adding D_2O to the

growth media as well as heavy glucose and heavy ammonium, a fully heavy system could be generated to give a more complete data set.

It should be noted that as the mass of the enzyme is increased, the R^2 values for the fit of data with MMRT are decreased. The fit still provides an accurate description of the data, however extended analysis would assist in determining whether a data skew is present. The data for the $^{13}C^{15}N$ variants appears to display differences in curvature not seen in the WT samples (Figure 4.11), reflecting the temperature dependent shape seen in the MalL data sets. When fit with the temperature dependent MMRT equation, the fit appears much better for these two data sets, generating R^2 values of 0.9739 and 0.9839 for non-exchanged, and deuterium exchanged samples respectively (Figure 5.3).

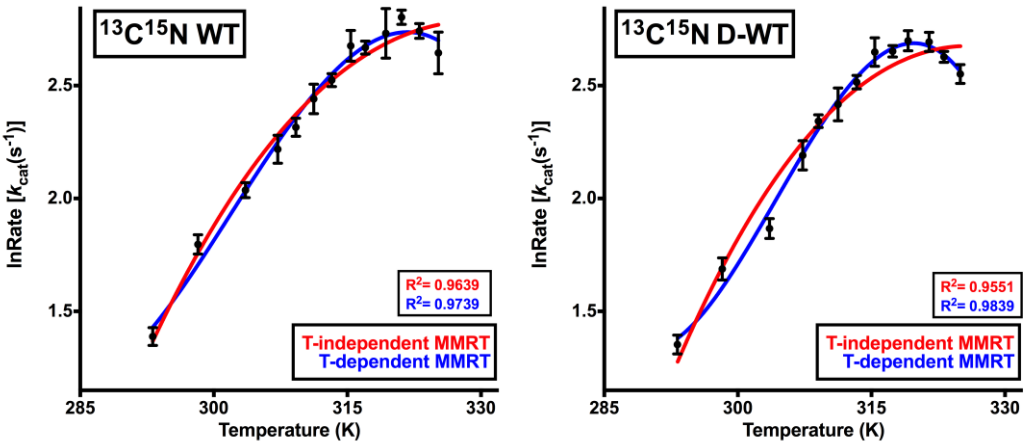

Figure 5.3.Temperature independent (red) and dependent (blue) MMRT fits for partially heavy KSI WT and deuterium exchanged variants.

Based on these data it is reasonable to speculate that over a greater temperature range KSI would more clearly display some heat capacity temperature dependence as is seen for MalL. The difference in curvature seen in these experiments could be argued to be the result of thermal denaturation as a consequence of reduced enzyme stability due to introduced heavy isotopes. However this seems unlikely, since the temperature at which half of the enzyme solution is denatured (T_m) for WT KSI has been experimentally determined using circular dichroism spectroscopy at 222 nm to be 338.45 K ± 0.3 (Cha et al., 2013); more than 10 degrees above the maximum assay temperature used in these experiments.

5.1.1.2 MalL temperature profiles

MalL displays the same general trends for changes in $\Delta C_p^{\ddagger}$ as KSI, decreasing as mass is increased (Table 6). Increased mass by incorporation of heavy carbon and nitrogen isotopes has slowed vibrations to shift vibrational mode distribution to lower frequencies, resulting in the same decrease in rate and decreased $\Delta C_p^{\ddagger}$ value. The overall decrease in rate is not as substantial as for KSI, and the reason for this discrepancy is unknown. Experimental rate data gathered for the MalL mutant V200S, gave information on the thermodynamic parameters for a rigidified version of the enzyme by comparison. A rigid system is expected to present with a vibrational mode distribution at higher frequencies (Figure 5.4) when compared to WT, essentially mimicking the event of a ligand or transition state binding. As expected, V200S has an increased $\Delta C_p^{\ddagger}$value comparative to the WT and displays decreased curvature about T_{opt}. Hobbs *et al.* (2013) previously determined that V200S shows differences in $\Delta C_p^{\ddagger}$ to WT resulting from changes to ground state exclusively, meaning that the ground state for V200S is much closer to the transition state complex in terms of structure and flexibility than the WT. It is reasonable to assume then, that as the enzyme gains mass without altering the total number of vibrational modes present, the $\Delta C_p^{\ddagger}$value will move further away from that determined for V200S. As with KSI, the MalL data would be more conclusive if supported by a fully heavy isotope variant. Rate versus temperature for $^{2}H^{13}C^{15}N$ MalL exists (Prentice, Unpublished data), however only for a high temperature range of 307-327 K, and at pH 6.5, lending little relevance to the data collected in this study.

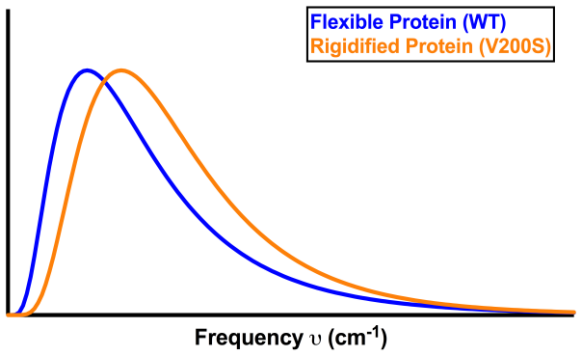

Figure 5.4. Schematic showing flexible versus rigid protein vibrational mode distribution as expected for WT and V200S MalL.

5.2 Temperature dependence of heat capacity

5.2.1 Effect of temperature dependent heat capacity on rate and thermodynamic parameters

The temperature dependence of heat capacity was previously assumed to be non-essential in MMRT considerations (Arcus & Pudney, 2015), however the data generated over the course of this study provides evidence to the contrary. Hobbs *et al.* (2013) used the temperature independent equation to generate results for MalL with very good fits, however, this study conducted assays over a much narrower temperature range (310-330 K) than the current study. At higher temperatures, the temperature dependence of heat capacity is not as apparent, since the curvature about T_{opt} is similar. At lower temperatures, curvature becomes much shallower, skewing the fit and providing an inaccurate description of experimental data (Figure 4.12). By incorporating a simple linear function to account for the temperature dependence of heat capacity as a first approximation (Prabhu & Sharp, 2005), the model becomes more robust in describing large temperature ranges and thermodynamic behaviour. Previous measurements of protein heat capacity have determined a linear temperature dependence, lending confidence to the use of a simple function to describe the temperature dependence of C_p (Makhatadze, 1998; Privalov & Makhatadze, 1990). The incorporation of heat capacity temperature dependence into the MMRT equation (Equation 4.3) gives significantly different functions for entropic and enthalpic contributions to rate when compared with temperature exclusive functions (Figure 5.5A). At temperature between 300 K-330 K, the functions show similar curvature, where the temperature dependent C_p appears mostly linear. Outside of this temperature range however, the functions show significant differences. Consequently, the function for Gibb's free-energy used in the MMRT equation is also significantly altered (Figure 5.5B). At higher temperatures (>300 K) the temperature dependent and independent functions for the Gibb's free energy are very similar as governed by the enthalpy and entropy functions. Combined, these factors account for the goodness of fit of the Hobbs *et al.* (2013) data for 300 K-330 K. At lower temperatures however, significant divergence is seen between the Gibb's free energy functions, as they begin to show opposing curvature (Figure 5.5B).

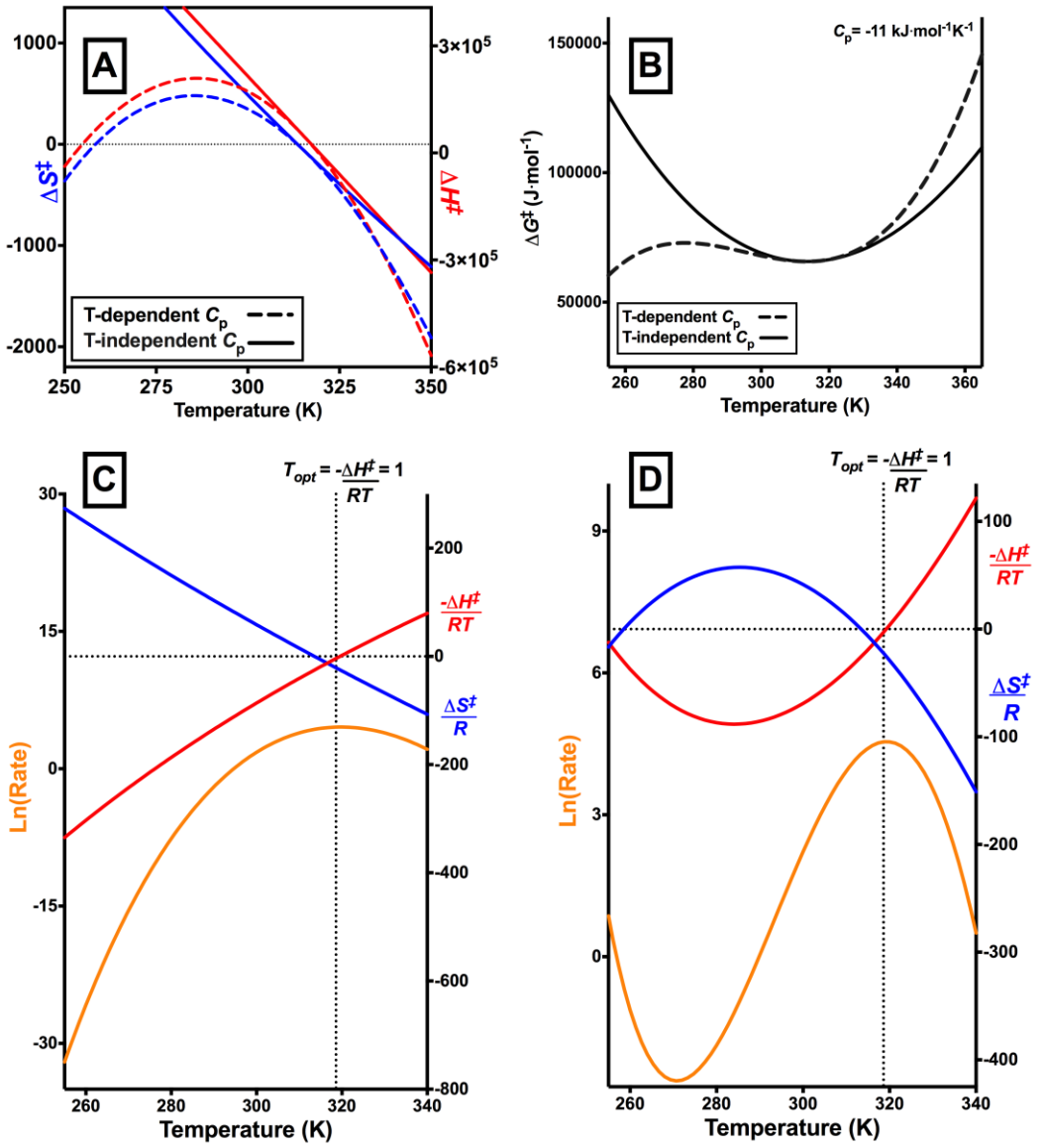

Figure 5.5. Effect of temperature independent and dependent heat capacity functions on enthalpy and entropy (A), Gibb's free-energy (B) and ln(rate) (C and D).

The alterations to thermodynamic parameters resulting from the incorporation of a temperature dependent heat capacity governs a shift in the rate function to significantly alter curvature in rates (Figure 5.5C&D). The T_{opt} remains unchanged, since the $\frac{-\Delta H^{\ddagger}}{RT}$ contribution is still equal to one at the same temperature. The entropic and enthalpic contributions also intersect at the same temperature, meaning that above T_{opt}, the governing contribution of each parameter is similar to the temperature independent model (Figure 5.5C&D). Below the intersection of these two terms is where significant alterations are seen, as there is a second intersection for the temperature dependent C_p rate function prior to the predicted

rate minimum. The overall result is an '*S-shaped*' function for rates with significantly differing curvature than the previous MMRT equation (Figure 5.5D).

The rate versus temperature MalL data shows that while an increase in mass by incorporation of heavy isotopes alters the change in heat capacity ($\Delta C_p^{\ddagger}$) by a shift in vibrational mode distribution, this shift has no effect on the temperature dependence of $\Delta C_p^{\ddagger}$. The $^{13}C^{15}N$ MalL had a $\Delta C_p^{\ddagger}$ slope identical to the WT, suggesting that the temperature dependence of $\Delta C_p^{\ddagger}$ for MalL is exclusive of vibrational considerations, or that these vibrational alterations are not significant enough to affect the temperature dependence. This may not provide an accurate description of all systems however, and it is frivolous to extrapolate this finding as universal. More data for a number of enzymes is required before conclusions can be drawn regarding vibrational mode effect on heat capacity temperature dependence.

5.2.2 Predicted rate increase at low temperatures using temperature dependent MMRT

The new temperature dependent MMRT equation predicts an increase in rates at very low temperatures when the function is extended beyond preliminary experimental results (Figure 4.14). Further experimental data contradicts this extrapolation, and increased data at low temperatures alters the function to predict this rate increase at a lower temperature than previously (Figure 5.6). This suggests that the predicted increase in rate following the minimum is an asymptotic artefact of the function, rather than a description of a real phenomenon.

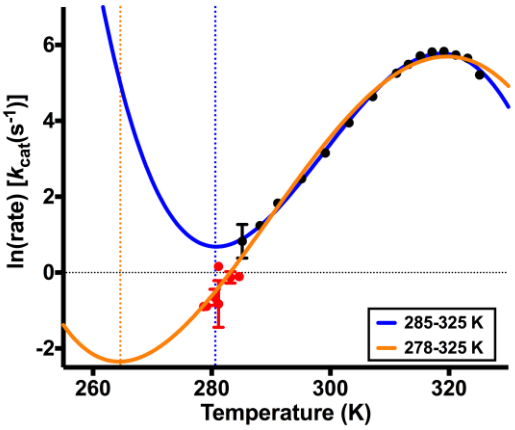

Figure 5.6. MalL wide temperature range data and extended functions. Data points in black show 285-325 K range. Blue line is the extended function of the temperature dependent MMRT for these data points. Predicted minimum is shown as blue dotted line. Red data points show additional lower temperature assays. Yellow line shows extended function for fit of black and red points together. Predicted minimum is shown as yellow dotted line.

5.3 Direct evidence of temperature dependent heat capacity

Identification of the temperature dependence of heat capacity over a biologically relevant temperature range for WT MalL provides further support for the new MMRT model (Equation 4.3 and Equation 4.4). The temperature dependence value generated falls well within range of previously determined protein heat capacity temperature dependence values (Makhatadze, 1998). The theoretical absolute value for C_p generated for MalL WT at 298.15 K from amino acid heat capacities is 36% smaller than the experimental value of ~155 $kJmol^{-1}K^{-1}$ obtained from the DSC experiments (Figure 4.19). The difference is considered to be small enough that the values are of the same magnitude, since the theoretical value is an estimate. The DSC values from Makhatdze *et al.* are for solid amino acids, meaning that solvation effects have not been accounted for. The DSC for MalL WT was conducted in buffer. Although the heat capacity was measured relative to a reference, it is likely that amino acid side chains when solvated contribute a higher heat capacity than when in solid form. An earlier paper by Makhatdze and Privalov (1990) looked at the heat capacity of aqueous amino acid side chains to address the hydration effect (Privalov & Makhatadze, 1990). By taking these heat capacity values and a general backbone contribution term of 15 $Jmol^{-1}K^{-1}$, an estimate for fully solvated MalL was determined to be ~138 $kJmol^{-1}K^{-1}$. This

value is 12% smaller than the experimentally determined value, significantly closer than the previous estimate. It is unlikely however, that all amino acid side chains will be subject to solvation effects. Those hydrophobic residues packed tightly in the centre of the structure will likely contribute heat capacity values similar to the solid values, whereas hydrophilic residues on the outside of the structure will likely contribute solvated heat capacities to the system.

The DSC data provides strong physical evidence for the inclusion of heat capacity temperature dependence in enzyme rate considerations. While it is unclear exactly how this temperature dependence changes between the apo-enzyme state and the transition state, it seems likely that differences here will have a role in explaining thermodynamic phenomena. Based on thermodynamic principles, it is expected that the temperature dependence of a more rigid system, such as the enzyme-transition state complex, will have a steeper temperature dependence than a more flexible system (Blundell & Blundell, 2006). A shift of the vibrational mode distribution to higher frequencies would indicate a steeper slope value for the transition state, resulting in an increase in $\Delta C_p^{\ddagger}$at higher temperatures, as the difference between the slope of the heat capacity at the ground state and the transition state becomes wider. This however is opposed to experimental observations, where the difference in heat capacity becomes greater (more negative) as temperature increases. It is possible that some other contributing factor is influencing this difference, or that the temperature dependence does not remain linear for the entire temperature range considered. It is also possible that the simple explanation offered here is naïve, and that the vibrational distribution exhibits only a partial shift. Further experiments to properly elucidate the nature of heat capacity temperature dependence at the vibrational mode level are required to reconcile experimental data with thermodynamic predictions.

5.4 Thermodynamic justification for temperature dependent heat capacity

From a thermodynamic basis, the relationship between heat capacity and temperature is much more complex than the simple linear relationship established for the MMRT equation (Blundell & Blundell, 2006). Using simple approximations, and an N-level system model for heat capacity at constant volume (C_v) versus temperature, allows derivation of the temperature dependence

of heat capacity at constant pressure (C_p). We can approximate that $C_v \approx C_p$ based on the formal definition of C_p-C_v (Equation 5.1).

Equation 5.1. Heat capacity at constant pressure minus heat capacity at constant volume in a liquid system.

$$C_p - C_v = VT\frac{\beta^2}{\kappa}$$

Where β equals the thermal expansion co-efficient of the system, and κ equals the compressibility of the system. In water at ambient temperature and pressure, β is approximately $207\epsilon^{-6}$per degree ("Properties of common liquid materials," 2016), and κ is $46\epsilon^{6}$ per atmosphere (4.5 Pa^{-1}) (Millero, Curry, & Drost-Hansen, 1969). Thus, $\frac{\beta^2}{\kappa} = \frac{207\epsilon^{-12}}{4.5} = 4.6\epsilon^{-11}$. The volume ($V$) of the system under biological conditions remains unchanged, giving $C_p - C_v \approx T \times 4.6\epsilon^{-11}$. The difference between C_p and C_v is so small that C_v can be assumed to be approximately equal to C_p for systems in aqueous solvents.

By treating the system as an N-level system (Equation 5.2), a visual representation of absolute C_p with respect to temperature can be generated (Figure 5.7). As temperature increases from absolute zero, there is a short lag phase, where insufficient energy is present to move the system beyond the ground state to higher energy modes. Following this, there is a steep linear increase in temperature dependence of heat capacity, as energy requirements for transitions are reached, and higher energy modes become populated. At a given temperature, heat capacity begins to plateau, reaching a maximum when all of the vibrational modes are populated. Beyond this maximum, the function begins to display a negative, curved slope downward, the steepness of which is dictated by the number of states in the system. This decline in heat capacity is known as the Schottky anomaly (Blundell & Blundell, 2006) (Figure 5.7).

Equation 5.2. N-level system equation. Derived from partition function (Z).

$$C_p = \frac{(hcv)^2}{k_B T^2}\left[\frac{e^{\frac{-hcv}{k_B T}}}{\left(e^{\frac{-hcv}{k_B T}} - 1\right)^2} - \frac{N^2 e^{\frac{-N^2 hcv}{k_B T}}}{\left(e^{\frac{-Nhcv}{k_B T}} - 1\right)^2}\right]$$

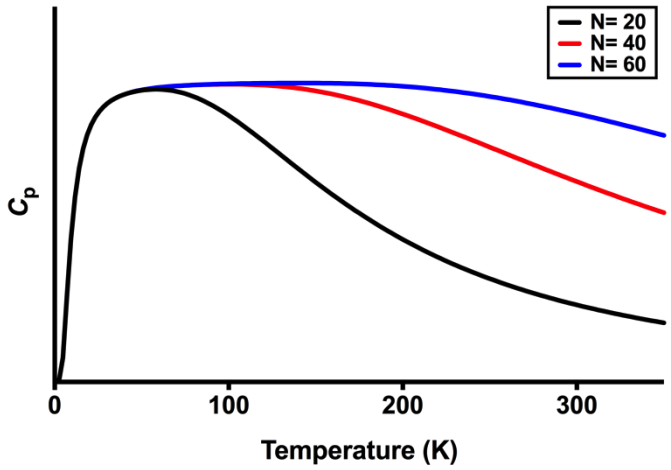

Figure 5.7. Absolute heat capacity (C_p) versus temperature (K) for 20, 40, and 60 state systems showing Schottky anomaly above maximum C_p.

From this definition, it is clear that as fewer states are available in the system, heat capacity diverges at a lower temperature, and with a more negative slope.

Vibrational mode availability is altered on binding of the ligand, and is further restricted by binding of the transition state. Molecular dynamics simulations of MalL with substrate and transition states bound, have revealed a 10% narrowing in the enthalpy distribution on binding of the transition state (Prentice, Unpublished data) (Figure 5.8). This provides evidence for the narrowing of the free-energy landscape at the transition state. Changes to the free energy landscape are reflected in the heat capacity of the system, since heat capacity is the mean enthalpy deviation squared (variance), over RT^2 (Equation 5.3) If the enthalpy distribution is narrowed, the variance is reduced, having a direct effect on the absolute value of the heat capacity (Equation 4).

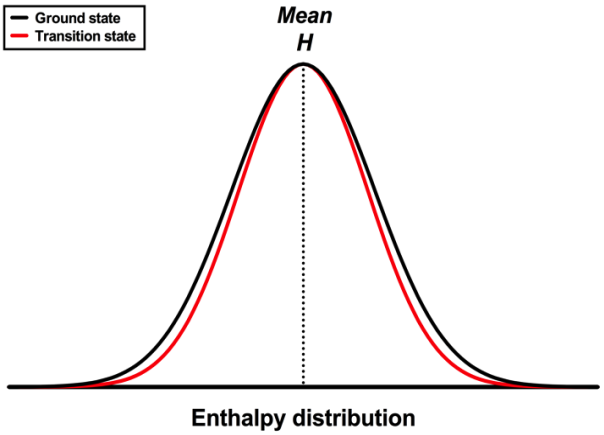

Figure 5.8. Enthalpy distribution for the enzyme-substrate complex (ground state, black) and the enzyme-transition state complex (red).

Equation 5.3. Heat capacity in terms of enthalpy change.

$$C_p = \frac{< \delta H^2 >}{RT^2}$$

It is reasonable then to assume that narrowing of enthalpy distribution at the transition state, corresponds to a reduction in the total heat capacity of the system. This results in a lower number of available vibrational modes due to reduced energy availability for vibrational transitions. Using an N-level system to describe the difference between the ground and transition states of the enzyme provides a simple model for the difference in heat capacity at biologically relevant temperatures (Figure 5.9).

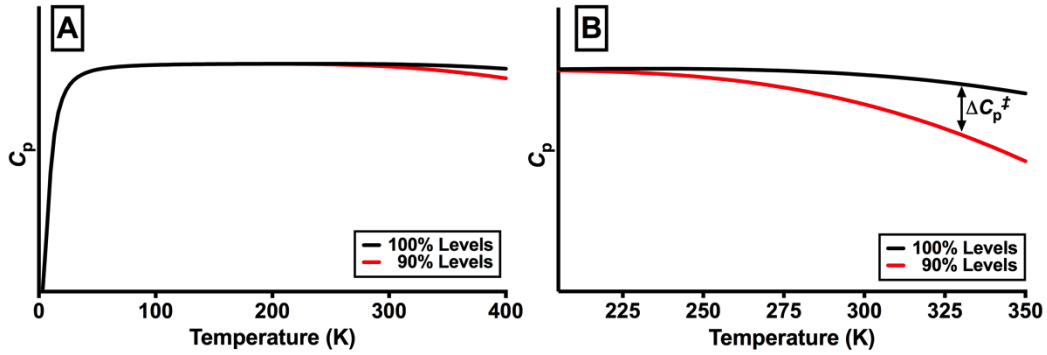

Figure 5.9. N-level system showing 10% reduction in states. B shows figure A zoomed to biologically relevant temperatures.

On binding of the transition state, the energy availability for vibrational transitions is restricted. Consequently, the slope of C_p is more negative for the transition state complex, demonstrating increased temperature dependence compared with the enzyme-substrate complex. The difference between these two values predicts a negative value for the change in temperature dependence between the two states. This rationalization agrees with experimental data generated from rate versus temperature assays, where negative values are seen for the temperature dependence of $\Delta C_p^{\ddagger}$ of MalL variants. This justification does not however agree with the DSC data generated for WT MalL. The data determines an absolute C_p slope that is positive at biologically relevant temperatures, contradicting the model above. The reason for this discrepancy is unknown and should be addressed in the future. It is unlikely that the DSC data generated for MalL is incorrect, as it agrees with data compiled in a review by Makhatdze (1997), combining a number of amino acid and protein C_p values determined from independent studies.

5.4.1.1 Evidence for the role of enthalpy distribution in heat capacity temperature dependency

The free-energy landscape associated with reaction coordinate and folding is governed by the number of sub-states available in equilibrium, which is governed by the enthalpy distribution (Blundell & Blundell, 2006). Changes to the enthalpy distribution are therefore reflected in the nature of the free energy landscape. Narrowing of the free-energy landscape has previously been associated with differences in heat capacity change across a temperature range. Oliveberg *et al.* (1995) looked at differences in protein folding activation enthalpies for two enzymes, barnase and chemotrypsin inhibitor 2 (CI2). CI2 is a 64 residue protein (Otzen et al., 1994) of ~7.5 kDa (Roesler & Rao, 1999). Barnase has 110 amino acids (Otzen et al., 1994) and is ~12.4 kDa (Makhatadze, 1998). The data generated looked at the temperature dependence of the rate constant for refolding (k_{obs}). The temperature at which $\frac{\Delta G^{\ddagger}}{RT}$ is at a minimum dictates the maximum for the k_{obs}, consistent with rate versus temperature data for enzyme catalysis (Oliveberg, Tan, & Fersht, 1995). It was determined that for CI2, the curvature was equal either side of this optimum, similar to that seen in the temperature independent model for MMRT. CI2 folds via a simple two-step kinetic mechanism, proceeding from a fully unfolded state to the folded state via a single high-energy transition state (Otzen et al., 1994) (Figure 5.10). The free-energy landscape associated with this folding mechanism is simple, since the reaction co-ordinate follows a one-step path with the rate-determining step being the rearrangement of the collapsed unfolded state (Otzen et al., 1994). This landscape reflects a small enthalpy distribution, owing to the limited number of sub-states existing for the pathway. There is only a single narrowing of the enthalpy distribution at the transition state, and therefore a larger (less negative) $\Delta C_p^{\ddagger}$ between the unfolded and folded state. Barnase is a larger enzyme, and folds by a multi-state, modular mechanism, proceeding via an intermediate folded step, resulting in two high-energy transition states (Oliveberg et al., 1995; Otzen et al., 1994). The result of this modular mechanism is the folding of smaller modules of the protein that then associate, which provides the rate determining step (Otzen et al., 1994) (Figure 5.10). Barnase has a larger number of possible sub-states, thus has a wider enthalpy distribution. The folding of barnase therefore produces a much more rugged free-energy landscape than for CI2, and a smaller (more

negative) $\Delta C_p^{\ddagger}$ reflecting two separate narrowing events of the enthalpy distribution.

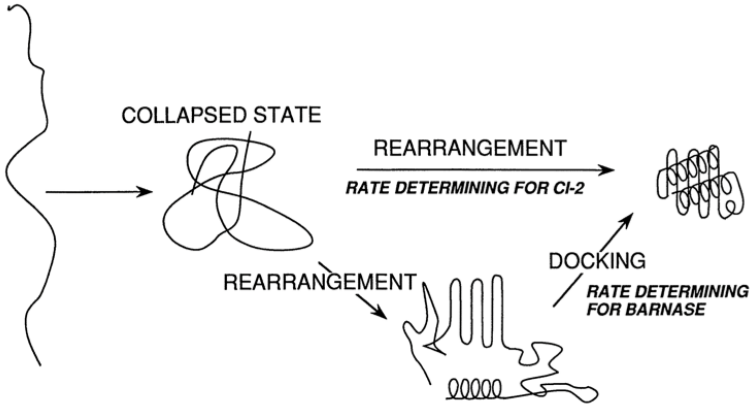

Figure 5.10. Schematic for folding pathways of CI2 and barnase. Retrieved from Otzen et al. (1994).

The difference in complexity between the two folding models has an effect on the temperature dependence of the rate constant for folding (Oliveberg et al., 1995). CI2 data shows a consistent curvature prior to and after the maximum, whereas the data generated for folding of barnase shows a distinct difference in curvature above and below the maximum. These data sets reflect differences seen for KSI and MalL data respectively, where the larger enzyme and more complex reaction scheme shows evidence of temperature dependent heat capacity. Consequently, both data sets from Oliveberg *et al.* (1995) were obtained and fit with temperature independent MMRT. The data set for barnase was also fit with temperature dependent MMRT (Figure 5.11).

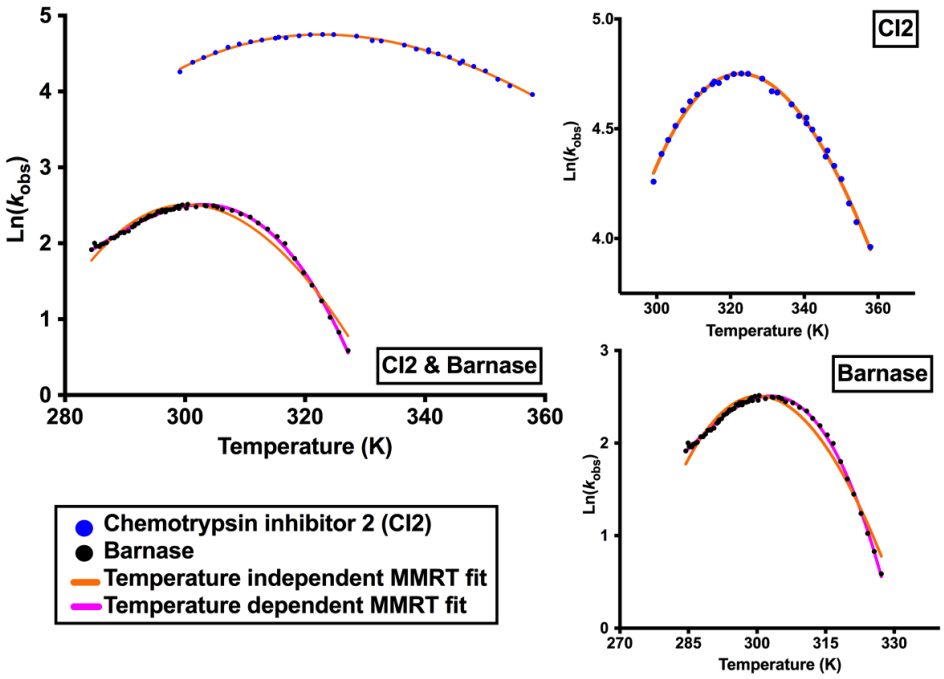

Figure 5.11. Temperature independent and dependent MMRT fits for CI2 and barnase rate constant for folding (k_{obs}) versus temperature. Data from Oliveberg et al. (1995).

The temperature independent MMRT model provides an excellent fit for the CI2 data, giving at R^2 value of 0.9942, however clearly does not provide an accurate fit for the barnase data set. The temperature dependent MMRT model provides an excellent fit for barnase, generating an R^2 value of 0.9974. The absence of temperature dependence from CI2 data supports the concept of increased process complexity being important in heat capacity considerations. As the enthalpy distribution between the ground and transition states undergoes a more substantial restriction, it is expected that heat capacity temperature dependence will become a greater factor. Unpublished data from Pudney *et al.* (2016) suggests the free-energy landscape of MalL is linked to catalytic rate based on a thorough thermodynamic analysis. Linking this data to temperature dependent heat capacity considerations will be critical in gaining a more advanced understanding of this correlation.

5.5 Future work

In future, temperature profiles should be obtained for fully heavy variants of both KSI and MalL at pH 7.0 to support the trends in $\Delta C_p^{\ddagger}$ established in these experiments. Enzymes expressed from *E. coli* grown in media with heavy carbon (C) and nitrogen (N) sources have been used in these experiments, however the

generation of data from enzymes with various combinations ($^{2}H^{13}C$, $^{2}H^{15}N$, $^{2}H^{13}C^{15}N$) of these isotopes incorporated, additional to those used in the current study, would give a robust data set to confirm the trend identified here. A number of other MalL mutants were produced by Hobbs *et al.*, and these should also be included in further analysis. If possible, KSI data should be extended to a wider temperature range to determine whether or not heat capacity temperature dependence becomes apparent.

DSC data should be gathered for MalL V200S and KSI to determine the differences in heat capacity temperature dependence by direct measure. This would provide further support for the conclusions reached through the course of this study, and assist in visualising differences in heat capacity temperature dependence between rigid and flexible systems. Further experiments using enzymes with and without transition state analogues that are known to bind tightly would provide insight into differences in the temperature dependence of heat capacity between the apo-enzyme and the transition state. Paired with temperature profiles modelled by temperature dependent MMRT, these data would enable the development of a more detailed description of experimental phenomena. One potential enzyme for use in this application would be human 5'methylthioadenosine phosphorylase (MTAP) (Guan et al., 2013). MTAP is a metabolic enzyme responsible for the breakdown of 5'methylthioadenosine, and has a series of transition state analogues that have been developed for potential as anticancer molecules, two of which include MT-ImmA and MT-DADMe-ImmA (Guan et al., 2013). These transition state analogues have been well characterised in terms of mode of action, geometry, and affinity, making this system ideal for future DSC work. Unfortunately, there is no evidence in the literature for a spectrophotometric assay to allow rate versus temperature plots to be generated easily. MalL also has known transition state analogues (Hansen & Bols, 2000) that could be used in this application, and the established activity assay would allow a full data set to be obtained. The transition state analogue used in the DSC experiments completed here, DONM, did not bind effectively to show a difference in temperature dependence thus the binding affinity of any analogues used would need to be determined to generate an accurate experimental design.

The ability of MMRT to describe the temperature dependence of folding for small, simple molecules such as CI2 and barnase extends the model to another

experimental protocol. The concept of enthalpy distribution changes determining the temperature dependence of heat capacity should be explored further. KSI is a small enzyme, shown to refold following a urea denaturation (Unfolding and refolding activity trials 3.1.1.1). The experimental protocol employed by Oliveberg *et al.* (1995) could be conducted for KSI to gather similar data sets, and modelled with MMRT. KSI is however a dimeric enzyme, which may provide complication. Kinetic analysis of the temperature dependence of a number of small refolding enzymes combined with computational analysis of their enthalpy distributions would assist in investigating this hypothesis further.

The possibility of using protein infrared spectroscopy to determine discreet differences in vibrational spectra between the apo-enzyme and enzyme with a transition state analogue bound should also be explored. Detailed information on the differences between vibrational motions occurring in each system would shed light on which movements are involved in experimental phenomena. Femtosecond optical kerr-like spectroscopy as conducted by Turton *et al*. could also be a useful technique for gathering this information. Coupled with molecular dynamics simulations, this information would assist in developing a physical explanation for the differences in thermodynamic parameters seen in temperature profiles, and allow a detailed description of vibrational mode distribution between the apo-enzyme and the transition state to be obtained. This would allow differences in vibrational mode distribution more complex than a simple distribution shift to be visualised should they be important.

6 Conclusion

The results presented here confirmed that alterations to vibrational mode frequency and distribution through heavy atom substitutions are reflected by changes to values of $\Delta C_{\mathrm{p}}{}^{\ddagger}$, although the exact molecular nature of these alterations cannot be confirmed. The data suggest that there are differences in the extent of these alterations between the enzyme-substrate complex and the enzyme-transition state complex. This will not be substantiated until further support is generated by molecular dynamics simulations and physical evidence obtained through advanced infrared techniques using transition state analogues. Future work to establish the trends identified here will help to better elucidate the physical alterations responsible for experimental phenomena. It is clear from the data presented here that vibrational mode distribution can successfully be shifted to lower frequency by increasing the mass of the enzyme to slow vibrations. Coupled with molecular dynamics simulations and data from Turton *et al.* (2015), where ligand and inhibitor binding exhibits a shift to higher vibrational frequencies, experimental observations can be reconciled to expected behaviour identified through physics concepts.

The second conclusion from these data is the identification of temperature dependent heat capacity as an integral factor in MMRT considerations. For more complex enzymatic reactions with greater enthalpic restrictions over the reaction coordinate, heat capacity temperature dependence significantly alters the MMRT model for ln(rate) over a wider temperature range than previously characterised. The development of a modified MMRT equation in response to this finding provides a more robust description of thermodynamic behaviour than the previous equation, and offers a more efficient model for rate versus temperature data moving forward. DSC data indicates absolute temperature dependence for heat capacity in WT MalL significantly deviant from zero. Combined with the data generated in temperature profiles for MalL, these data postulate a temperature dependent $\Delta C_p{}^{\ddagger}$ will also be present, although further experimental work is required to confirm this. Future DSC experiments to determine how heat capacity temperature dependence is altered with system flexibility and complexity will also be beneficial in characterising the nature of this temperature dependence.

References

A I Vogel, B. S. F. E. A. (1989). Vogel's Textbook of Practical Organic Chemistry, (5 ed., pp. 1–1540). New York: John Wiley & Sons, Inc.

Arcus, V. L., & Pudney, C. R. (2015). Change in heat capacity accurately predicts vibrational coupling in enzyme catalyzed reactions. *FEBS Letters*, *589*(17), 2200–2206. http://doi.org/10.1016/j.febslet.2015.06.042

Arcus, V. L., Prentice, E. J., Hobbs, J. K., Mulholland, A. J., Van der Kamp, M. W., Pudney, C. R., et al. (2016). On the Temperature Dependence of Enzyme-Catalyzed Rates. *Biochemistry*, acs.biochem.5b01094–8. http://doi.org/10.1021/acs.biochem.5b01094

Austin, R. H., Roberson, M. W., & Mansky, P. (1989). Far-infrared perturbation of reaction rates in myoglobin at low temperatures. *Physical Review Letters*, *62*(16), 1912–1915. http://doi.org/10.1103/PhysRevLett.62.1912

Balog, E., Smith, J. C., & Perahia, D. (2006). Conformational heterogeneity and low-frequency vibrational modes of proteins. *Physical Chemistry Chemical Physics*, *8*(47), 5543–6. http://doi.org/10.1039/b610075a

Baneyx, F. (1999). Recombinant protein expression in *Echerichia coli*, (10), 411–421.

Beg, A. E. (1984). The effect of pH various additives on extinction coefficients for p-nitrophenol. J. Chem. Soc. Pak.

ben-Avraham, D. (1993). Vibrational normal-mode spectrum of globular proteins. *Physical Review B*, *47*(21), 14559–14560. http://doi.org/10.1103/PhysRevB.47.14559

Benov, L., & Fridovich, I. (1998). Growth in iron-enriched medium partially compensates Escherichia coli for the lack of manganese and iron superoxide dismutase. *Journal of Biological Chemistry*, *273*(17), 10313–10316.

Biggs, A. I. (1954). A spectrophotometric determination of the dissociation constants of p-nitrophenol and papaverine. *Transactions of the Faraday Society*, *50*(0), 800–802. http://doi.org/10.1039/TF9545000800

Blundell, S. J., & Blundell, K. M. (2006). Concepts in Thermal Physics. Oxford University Press. http://doi.org/10.1093/acprof:oso/9780199562091.001.0001

Bradford, M. (1976). A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Biochemistry*, *72*(1-2), 248–254. http://doi.org/10.1006/abio.1976.9999

Buchanan, C. L., Connaris, H., Danson, M. J., Reeve, C. D., & Hough, D. W. (1999). An extremely thermostable aldolase from Sulfolobus solfataricus with specificity for non-phosphorylated substrates. *The Biochemical Journal*, *343* (3), 563–570.

Cha, H. J., Jang, D. S., Kim, Y.-G., Hong, B. H., Woo, J.-S., Kim, K.-T., & Choi, K. Y. (2013). Rescue of deleterious mutations by the compensatory Y30F mutation in ketosteroid isomerase. *Molecules and Cells*, *36*(1), 39–46. http://doi.org/10.1007/s10059-013-0013-1

Chopra, A. K., Brasier, A. R., Das, M., Xin-Jing, X., & Petersoir, J. W. (1994). Improved synthesis of Salmonella typhimurium enterotoxin using gene fusion expression systems. *Gene*, *144*(1), 81–85. http://doi.org/10.1016/0378-1119(94)90207-0

Clarke, J., Hounslow, A. M., Bycroft, M., & Fersht, A. R. (1993). Local breathing and global unfolding in hydrogen exchange of barnase and its relationship to protein folding pathways. *Proceedings of the National Academy of Sciences*, *90*(21), 9837–9841.

Corkrey, R., Olley, J., Ratkowsky, D., McMeekin, T., & Ross, T. (2012). Universality of Thermodynamic Constants Governing Biological Growth Rates. *PLoS ONE*, *7*(2), e32003–8. http://doi.org/10.1371/journal.pone.0032003

Creighton, T. E. (1993). Proteins: structures and molecular properties (2nd ed.). W. H. Freeman and Company.

Cui, Q., & Karplus, M. (2008). Allostery and cooperativity revisited. *Protein Science*, *17*(8), 1295–1307. http://doi.org/10.1110/ps.03259908

Cumming, M. H. (2012). *Structural and enzymatic characterisation of nucleoside*

triphosphate diphosphohydrolases from Trifolium repens and Dolichos biflorus. (V. L. Arcus & N. Roberts, Eds.). University of Waikato.
Daniel, R. M., & Danson, M. J. (2010). A new understanding of how temperature affects the catalytic activity of enzymes. *Trends in Biochemical Sciences*, *35*(10), 584–591. http://doi.org/10.1016/j.tibs.2010.05.001
Daniel, R. M., & Danson, M. J. (2013). Temperature and the catalytic activity of enzymes: A fresh understanding. *FEBS Letters*, *587*(17), 2738–2743. http://doi.org/10.1016/j.febslet.2013.06.027
Daniel, R. M., Danson, M. J., & Eisenthal, R. (2001). The temperature optima of enzymes: a new perspective on an old phenomenon. *Trends in Biochemical Sciences*, *26*(4), 223–225.
Daniel, R. M., Peterson, M. E., Danson, M. J., Price, N. C., Kelly, S. M., Monk, C. R., et al. (2009). The molecular basis of the effect of temperature on enzyme activity. *The Biochemical Journal*, *425*(2), 353–360. http://doi.org/10.1042/BJ20091254
Elias, M., Wieczorek, G., Rosenne, S., & Tawfik, D. S. (2014). The universality of enzymatic rate– temperature dependency. *Trends in Biochemical Sciences*, *39*(1), 1–7. http://doi.org/10.1016/j.tibs.2013.11.001
Engel, T. R., & Reid, P. (2006). Physical Chemistry (2nd ed.). United States of America: Pearson Education.
ENZYME entry 3.2.1.10.. ENZYME entry 3.2.1.10. Retrieved November 2, 2016, from http://enzyme.expasy.org/EC/3.2.1.10
Eyring, H. (1935). The Activated Complex in Chemical Reactions. *The Journal of Chemical Physics*, *3*(2), 107–10. http://doi.org/10.1063/1.1749604
Eyring, H., & Polanyi, M. (1931). Uber einfache Gasreaktionen, Z. physik. Chem., B12: 279 (1931).: The Energy of Activation for Bimolecular Reactions Involving Hydrogen and the …. J. Am. Chem. Soc.
Feller, G., & Gerday, C. (2003). Psychrophilic enzymes: hot topics in cold adaptation. *Nature Reviews Microbiology*, *1*(3), 200–208. http://doi.org/10.1038/nrmicro773
Frandsen, T. P., Palcic, M. M., & Svensson, B. (2002). Substrate recognition by three family 13 yeast alpha-glucosidases. *European Journal of Biochemistry*, *269*(2), 728–734.
Frauenfelder, H., Parak, F., & Young, R. D. (1988). Conformational substates in proteins. *Annual Review of Biophysics and Biophysical Chemistry*, *17*(1), 451–479. http://doi.org/10.1146/annurev.bb.17.060188.002315
Frey, P. A., & Hegeman, A. D. (2006). Enzymatic Reaction Mechanisms. New York: Oxford University Press.
Garrett, R., & Grisham, C. M. (2010). Biochemistry (4 ed.). Boston: Mary Finch.
Gasteiger, E., Hoogland, C., Gattiker, A., & Duvaud, S. (2005). Protein identification and analysis tools on the ExPASy server. *The Proteomics Protocols Handbook* (pp. 571–607). Totowa, NJ: Humana Press. http://doi.org/10.1385/1-59259-890-0:571
Guan, R., Tyler, P. C., Evans, G. B., & Schramm, V. L. (2013). Thermodynamic Analysis of Transition-State Features in Picomolar Inhibitors of Human 5′-Methylthioadenosine Phosphorylase. *Biochemistry*, *52*(46), 8313–8322. http://doi.org/10.1021/bi401188w
Hammes-Schiffer, S. (2002). Impact of Enzyme Motion on Activity †. *Biochemistry*, *41*(45), 13335–13343. http://doi.org/10.1021/bi0267137
Hansen, S. U., & Bols, M. (2000). Azafagomine hydrazones: an argument against a "Flat" transition state in glycoside cleavage. *Journal of the Chemical Society, Perkin Transactions 2*, (4), 665–667. http://doi.org/10.1039/a909648e
Haynie, D. T. (2009a). Biological Thermodynamics. Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9780511754784
Haynie, D. T. (2009b). Glossary. In *Biological Thermodynamics* (pp. 363–372). Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9780511754784.015
Hobbs, J. K., Jiao, W., Easter, A. D., Parker, E. J., Schipper, L. A., & Arcus, V. L. (2013). Change in Heat Capacity for Enzyme Catalysis Determines Temperature Dependence of Enzyme Catalyzed Rates. *ACS Chemical Biology*, *8*(11), 2388–2393.

http://doi.org/10.1021/cb4005029
Hriscu, M., Chiş, L., Toşa, M., & Irimie, F. D. (2013). pH-Profiling of thermoactive lipases and esterases: Caveats and further notes. *European Journal of Lipid Science and Technology*, *115*(5), 571–575. http://doi.org/10.1002/ejlt.201200305
Hutchinson, M. H., Morreale, G., Middelberg, A. P. J., & Chase, H. A. (2006). Production of enzymatically active ketosteroid isomerase following insoluble expression inEscherichia coli. *Biotechnology and Bioengineering*, *95*(4), 724–733. http://doi.org/10.1002/bit.21043
Jones, H. B. L., Weels, S. A., Prentice, E. J., Kwok, A., Arcus, V. L., & Pudney, C. R. (2016). A complete thermodynamic analysis of enzyme turnover links the free energy landscape to enzyme catalysis. *Unpublished*.
Krężel, A., & Bal, W. (2004). A formula for correlating pKa values determined in D2O and H2O. *Journal of Inorganic Biochemistry*, *98*(1), 161–166. http://doi.org/10.1016/j.jinorgbio.2003.10.001
La Mer, V. K. (1933). Chemical Kinetics. The Temperature Dependence of the Energy of Activation. The Entropy and Free Energy of Activation. *The Journal of Chemical Physics*, *1*(5), 289–296. http://doi.org/10.1063/1.1749291
Laidler, K. J., & King, M. C. (1983). Development of transition-state theory. *The Journal of Physical Chemistry*, *87*(15), 2657–2664. http://doi.org/10.1021/j100238a002
Liu, C.-C., Richard, A. J., Datta, K., & LiCata, V. J. (2008). Prevalence of Temperature-Dependent Heat Capacity Changes in Protein-DNA Interactions. *Biophysical Journal*, *94*(8), 3258–3265. http://doi.org/10.1529/biophysj.107.117697
Makarchikov, A. F., Lakaye, B., Gulyai, I. E., Czerniecki, J., Coumans, B., Wins, P., et al. (2003). Thiamine triphosphate and thiamine triphosphatase activities: from bacteria to mammals. *Cellular and Molecular Life Sciences (CMLS)*, *60*(7), 1477–1488. http://doi.org/10.1007/s00018-003-3098-4
Makhatadze, G. I. (1998). Heat capacities of amino acids, peptides and proteins. *Biophysical Chemistry*, *71*(2-3), 133–156. http://doi.org/10.1016/S0301-4622(98)00095-7
Millero, F. J., Curry, R. W., & Drost-Hansen, W. (1969). Isothermal compressibility of water at various temperatures. *Journal of Chemical & Engineering Data*, *14*(4), 422–425. http://doi.org/10.1021/je60043a018
Moritsugu, K., Miyashita, O., & Kidera, A. (2000). Vibrational energy transfer in a protein molecule. *Physical Review Letters*, *85*(18), 3970–3973. http://doi.org/10.1103/PhysRevLett.85.3970
Motlagh, H. N., Wrabl, J. O., Li, J., & Hilser, V. J. (2015). The ensemble nature of allostery. *Nature*, *508*(7496), 331–339. http://doi.org/10.1038/nature13001
Na, H., Song, G., & ben-Avraham, D. (2016). Universality of vibrational spectra of globular proteins. *Physical Biology*, *13*(1), 016008–42. http://doi.org/10.1088/1478-3975/13/1/016008
Nelson, P. (2003). Biological Physics. Macmillan Higher Education.
Oliveberg, M., Tan, Y. J., & Fersht, A. R. (1995). Negative activation enthalpies in the kinetics of protein folding. *Proceedings of the National Academy of Sciences*, *92*(19), 8926–8929.
Otzen, D. E., Itzhaki, L. S., elMasry, N. F., Jackson, S. E., & Fersht, A. R. (1994). Structure of the transition state for the folding/unfolding of the barley chymotrypsin inhibitor 2 and its implications for mechanisms of protein folding. *Proceedings of the National Academy of Sciences*, *91*(22), 10422–10425.
Perticaroli, S., Russo, D., Paolantoni, M., Gonzalez, M. A., Sassi, P., Nickels, J. D., et al. (2015). Painting biological low-frequency vibrational modes from small peptides to proteins. *Physical Chemistry Chemical Physics*, *17*, 11423–11431. http://doi.org/10.1039/C4CP05388E
Peterson, M. E., Eisenthal, R., Danson, M. J., Spence, A., & Daniel, R. M. (2004a). A new intrinsic thermal parameter for enzymes reveals true temperature optima. *Journal of Biological Chemistry*, *279*(20), 20717–20722. http://doi.org/10.1074/jbc.M309143200
Peterson, M. E., Eisenthal, R., Danson, M. J., Spence, A., & Daniel, R. M. (2004b). A

new intrinsic thermal parameter for enzymes reveals true temperature optima. *Journal of Biological Chemistry*, *279*(20), 20717–20722. http://doi.org/10.1074/jbc.M309143200

Pollack, R. M. (2004). Enzymatic mechanisms for catalysis of enolization: ketosteroid isomerase. *Bioorganic Chemistry*, *32*(5), 341–353. http://doi.org/10.1016/j.bioorg.2004.06.005

Prabhu, N. V., & Sharp, K. A. (2005). HEAT CAPACITY IN PROTEINS. *Annual Review of Physical Chemistry*, *56*(1), 521–548. http://doi.org/10.1146/annurev.physchem.56.092503.141202

Prentice, E. J. (Unpublished.). *The Temperature Dependence of Biological Rates*. (V. L. Arcus & J. Lane, Eds.).

Privalov, P. L., & Makhatadze, G. I. (1990). Heat capacity of proteins. II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: protein unfolding effects. *Journal of Molecular Biology*, *213*(2), 385–391. http://doi.org/10.1016/S0022-2836(05)80198-6

Properties of common liquid materials. (2016, December). Properties of common liquid materials. Retrieved December 9, 2016, from http://www.efunda.com/materials/common_matl/Common_Matl.cfm?MatlPhase=Liquid&MatlProp=Thermal

Raman, S., Taylor, N., Genuth, N., Fields, S., & Church, G. M. (2014). Engineering allostery. *Trends in Genetics*, *30*(12), 521–528. http://doi.org/10.1016/j.tig.2014.09.004

Roesler, K. R., & Rao, A. G. (1999). Conformation and stability of barley chymotrypsin inhibitor-2 (CI-2) mutants containing multiple lysine substitutions. *Protein Engineering*, *12*(11), 967–973.

Sambrook, J., & Russell, D. W. (2001). Molecular cloning (3rd ed., Vol. 3). Cold Spring Habour Laboratory Press.

Schipper, L. A., Hobbs, J. K., Rutledge, S., & Arcus, V. L. (2014). Thermodynamic theory explains the temperature optima of soil microbial processes and high Q10values at low temperatures. *Global Change Biology*, *20*(11), 3578–3586. http://doi.org/10.1111/gcb.12596

Schramm, V. L. (2005). Enzymatic transition states: thermodynamics, dynamics and analogue design. *Archives of Biochemistry and Biophysics*, *433*(1), 13–26. http://doi.org/10.1016/j.abb.2004.08.035

Schwartz, S. D., & Schramm, V. L. (2009). Enzymatic transition states and dynamic motion in barrier crossing. *Nature Chemical Biology*, *5*(8), 551–558. http://doi.org/10.1038/nchembio.202

Sharp, D. W. A. (1990). The Penguin Dictionary of Chemistry. Penguin (Non-Classics).

Sturtevant, J. M. (1977). Heat capacity and entropy changes in processes involving proteins. *Proceedings of the National Academy of Sciences*, *74*(6), 2236–2240. http://doi.org/10.1073/pnas.74.6.2236

Sutcliffe, M. J., & Scrutton, N. S. (2009). A new conceptual framework for enzyme catalysis. *European Journal of Biochemistry*, *269*(13), 3096–3102. http://doi.org/10.1046/j.1432-1033.2002.03020.x

Tipler, P. A. (1998). Physics for Scientists And Engineers Volume 1 + Volume 2. W H Freeman & Company.

Tuñón, I., Laage, D., & Hynes, J. T. (2015). Are there dynamical effects in enzyme catalysis? Some thoughts concerning the enzymatic chemical step. *Archives of Biochemistry and Biophysics*, *582*, 42–55. http://doi.org/10.1016/j.abb.2015.06.004

Turton, D. A., Senn, H. M., Harwood, T., Lapthorn, A. J., Ellis, E. M., & Wynne, K. (2014). Terahertz underdamped vibrational motion governs protein-ligand binding in solution. *Nature Communications*, *5*, 1–6. http://doi.org/10.1038/ncomms4999

Whitmire, S. E., Wolpert, D., Markelz, A. G., Hillebrecht, J. R., Galan, J., & Birge, R. R. (2003). Protein Flexibility and Conformational State: A Comparison of Collective Vibrational Modes of Wild-Type and D96N Bacteriorhodopsin. *Biophysical Journal*, *85*(2), 1269–1277. http://doi.org/10.1016/S0006-3495(03)74562-7

Xie, A., van der Meer, A. F. G., & Austin, R. H. (2001). Excited-State Lifetimes of Far-Infrared Collective Modes in Proteins. *Physical Review Letters*, *88*(1), 018102–4. http://doi.org/10.1103/PhysRevLett.88.018102

Appendix I: Expression media compositions

LB media

10 g/L tryptone

10 g/L NaCl

5 g/L yeast extract

Made up to volume in baffled flask using distilled H_2O. Autoclaved prior to use.

LB-glucose media

10 g/L tryptone

10 g/L NaCl

5 g/L yeast extract

2 g/L D-glucose

Made up to volume in baffled flask using distilled H_2O. Autoclaved prior to use.

Minimal media

1.1.1.4 5X M9 salts

33.89 g/ L Na_2HPO_4 (anhydrous)

15 g/L KH_2PO_4

2.5 g/L NaCl

5.0 g/L NH_4Cl

Made up to volume in Schott bottle and autoclaved prior to use.

1.1.1.5 Heavy 5X M9 salts

Made up as for 5X M9 salts replacing NH_4Cl with $^{15}NH_4Cl$.

1.1.1.6 BME vitamins

100 mg/L D-biotin

100 mg/L folic acid

100 mg/L nicotinamide

100 mg/L pyridoxal HCl

10 mg/L riboflavin

100 mg/L Thiamine HCl

50 mg/L cyanocobalamin

Filter sterilized prior to use.

1.1.1.7 M9 media

200 ml/L 5X M9 salts (autoclaved)

2 ml/L 1 M $MgSO_4$ (autoclaved)

20 ml/L 20% glucose solution (filter sterilized)

0.1 ml/L 1 M $CaCl_2$ (autoclaved)

0.01% 1 mg/ml thiamine (filter sterilized)

10 ml/L BME vitamins (filter sterilized)

10 μm $FeCl_3$ (filter sterilized)

250 mM NaCl (autoclaved)

Components added to sterile autoclaved flask containing 750 ml (or appropriate equivalent if preparing volumes outside 1 L) MQ H_2O. Made up to full volume with sterilized MQ H_2O.

1.1.1.8 Heavy M9 media

Made up as for M9 media replacing 5X M9 salts with heavy 5X M9 salts, and 20% glucose solution with 20% ^{13}Cglucose solution.

Appendix II: Buffer and gel compositions

Protein purification buffers

All buffers were filter sterilized prior to use. Adjusted to correct pH prior to making up to full volume with distilled H_2O.

Lysis Buffer

Enzyme	pH	Chemical composition
KSI	7.0	50 mM Tris 150 mM NaCl 20 mM imidazole
MalL	7.0	10 mM KCl 1 mM $MgCl_2$ 1 mM $CaCl_2$ 25 mM imidazole

Elution Buffer

Enzyme	pH	Chemical composition
KSI	7.0	50 mM Tris 150 mM NaCl 1 M imidazole
MalL	7.0	10 mM KCl 1 mM $MgCl_2$ 1 mM $CaCl_2$ 1 M imidazole

Size exclusion buffer

Enzyme	pH	Buffer Type	Chemical composition
KSI	7.0	Phosphate	50 mM $NaPO_4$
			150 mM NaCl
	5.0	Citrate	50 mM citrate
			150 mM NaCl
MalL	7.0	Phosphate	50 mM $NaPO_4$
			150 mM NaCl
	6.0	Phosphate	100 mM $NaPO_4$

Assay buffer

Enzyme	pH	Buffer Type	Chemical composition
KSI	7.0	Phosphate	50 mM $NaPO_4$
			150 mM NaCl
	5.0	Citrate	50 mM citrate
			150 mM NaCl
MalL	7.0	Phosphate	50 mM $NaPO_4$
			150 mM NaCl
	6.0	Phosphate	100 mM $NaPO_4$
			150 mM NaCl

pH parameter buffers

pH	Buffer Type	Chemical composition
3.5	Acetate	100 mM CH_3COONa
4.0		
5.0	Citrate	50 mM citrate
5.5		150 mM NaCl
6.0		
6.5	Phosphate	50 mM $NaPO_4$
7.0		150 mM NaCl
7.5		
8.0		
8.5	Glycine	50 mM glycine
10.0		300 mM NaCl
11.0	Sodium Hydroxide	0.1 M NaOH
13.5		

SDS-PAGE gel buffers

Resolving buffer

1.5 M Tris (pH 8.8)

Stacking buffer

1 M Tris (pH 6.8)

Quenching buffer (Q4)

250 mM Tris HCl (pH 6.8)

20 % glycerol (w/v)

4 % SDS (w/v)

10 % mercaptoethanol (w/v)

0.025 % bromophenol blue (w/v)

TG-SDS running buffer

25 mM Tris

250 mM glycine

0.1 % SDS (w/v)

Fairbanks A staining solution

0.05 % R250 coomassie blue dye (w/v)

25 % isopropanol (bulk grade) (w/v)

10 % acetic acid (analytical grade) (w/v)

Fairbanks D destaining solution

10 % acetic acid (analytical grade) (w/v)

SDS-PAGE gels

Gels were made in multiples of 5 using a stacked 5-gel cast. Each gel set was made up of a resolving gel and a stacking gel. Resolving layer was poured first and covered in 2 ml isopropyl alcohol (to prevent gel dehydration) and left to set for 40 minutes. Isopropyl alcohol layer was removed and the resolving gel rinsed thoroughly with MQ H_2O. Stacking gel was poured to a height of 2 cm above the resolving layer and 10 well combs inserted. Gels were left to set for 30 minutes prior to removal from cast and storage in damp paper towelling inside an airtight plastic bag in the fridge (4 °C).

Resolving gel

Component	Volume added (ml) for 12 % gel	Volume added (ml) for 16.5 % gel
Distilled H_2O	10.05	5.55
30 % acrylamide (w/v)	12.00	16.50
Resolving buffer	7.50	7.50
10 % SDS (w/v)	0.30	0.30
10 % APS (w/v)	0.15	0.15
TEMED	0.015	0.015

Stacking Gel

Component	Volume added (ml)
Distilled H_2O	8.500
30 % acrylamide (w/v)	2.125
Stacking buffer	1.600
10 % SDS (w/v)	0.125
10 % APS (w/v)	0.063
TEMED	0.0063

Buffers used in determination of ε (*p*-nitrophenol)

pH	Composition
3.5	100 mM CH_3COONa
4.0	
5.0	50 mM citrate
5.5	150 M NaCl
6.0	
6.5	50 mM $NaPO_4$
7.0	150 mM NaCl
7.5	
8.0	50 mM Tris
8.5	150 mM NaCl
9.0	20 mM imidazole
11.0	100 mM NaOH
13.3	

Appendix III: Experimental variations in method development

Table 10. Summarised conditions for all growth trials conducted in minimal media for KSI.

Culture number	Additives	Culture size	Inoculum dilution	Induction temperature	Expression time	Soluble expression?
1	None	250 ml	1:10	28°C	24 hours	Minimal
2	100 mg/L thiamine	250 ml	1:25	28°C	36 hours	Minimal
3	100 mg/L thiamine 50% additional glucose	250 ml	1:25	28°C	36 hours	
4	100 mg/L thiamine,	100 ml	1:50	18°C	60 hours	
5	100 mg/L thiamine, Addition 250 mM NaCl	100 ml	1:50	18°C	60 hours	
6	100 mg/L thiamine, additional 500 mM NaCl	100 ml	1:50	18°C	60 hours	
7	100 mg/L thiamine, 10 μm $FeCl_3$	100 ml	1:50	18°C	60 hours	50:50
8	100 mg/L thiamine, 1 ml B vitamin mix	100 ml	1:50	18°C	60 hours	50:50
9	100 mg/L thiamine, Additional 250 mM NaCl, 10 μm $FeCl_3$, 1 ml B vitamin mix	100 ml	1:50	18°C	60 hours	60:40
10	100 mg/L thiamine, Additional 250 mM NaCl, 10 μm $FeCl_3$, 1 ml B vitamin mix	250 ml	1:50	18°C	60 hours	Minimal

Table 10 Ctd. Summarised conditions for all growth trials conducted in minimal media for KSI.

Culture number	Additives	Culture size	Inoculum dilution	Induction temperature	Expression time	Soluble expression?
11	100 mg/L thiamine, Additional 250 mM NaCl, 10 μm $FeCl_3$, 1 ml B vitamin mix	250 ml	1:50	28°C	60 hours	60:40
12	100 mg/L thiamine, Additional 250 mM NaCl, 10 μm $FeCl_3$, 1 ml B vitamin mix	100 ml	1:50	28°C	60 hours	50:50
13	100 mg/L thiamine, Additional 100 mM NaCl, 10 μm $FeCl_3$, 1 ml B vitamin mix	100 ml	1:50	18°C	60 hours	50:50
14	100 mg/L thiamine, Additional 100 mM NaCl, 10 μm $FeCl_3$, 1 ml B vitamin mix. 20 min heat shock at 47°C	100 ml	1:50	18°C	60 hours	50:50

Table 11. Summarised conditions for deuterium exchange assays

Experiment number	Enzyme	[Enzyme] (mg/ml)	Substrate conditions (Final concentration, solvent)	Forward exchange T	Incubation T (Reverse exchange)	Incubation period (minutes)	Assay T	Buffer pH
1	KSI	0.00191	1.5 mM, in 100% methanol stored on ice during experiment	18°C	18°C	30	30°C	pH 7.0
2	KSI	0.00191	1.5 mM, in 100% methanol Stored on ice during experiment	18°C	18°C	30	30°C	pH 7.0
3	KSI	0.00191	0.5 mM, in 20% methanol Stored in thermomixer at 30°C during experiment	18°C	18°C	30	30°C	pH7.0
4	KSI	0.00191	0.5 mM, in 20% methanol Stored in thermomixer at 30°C during experiment	18°C	18°C	15	30°C	pH7.0 size exclusion buffer, pH5.0 exchange and assay buffers
5	KSI	0.00191	0.5 mM, in 20% methanol Stored in thermomixer at 30°C during experiment	18°C	18°C	15	30°C	pH7.0 size exclusion buffer, pH5.0 exchange and assay buffers
6	KSI	0.00191	0.5 mM, in 20% methanol Stored in thermomixer at 30°C during experiment	18°C	30°C	15	30°C	pH7.0 size exclusion buffer, pH5.0 exchange and assay buffers

Table 11 Ctd. Summarised conditions for deuterium exchange assays

Experiment number	Enzyme	[Enzyme] (mg/ml)	Substrate conditions (Final concentration, solvent)	Forward exchange T	Incubation T (Reverse exchange)	Incubation period (minutes)	Assay T	Buffer pH
7	KSI	0.00191	0.5 mM, in 20% methanol Stored in thermomixer at 30°C during experiment	18°C	30°C	15	30°C	pH7.0 size exclusion buffer, pH5.0 exchange and assay buffers
8	KSI	0.001165	0.5 mM, in 20% methanol Stored in thermomixer at 30°C during experiment Substrate from freezer stocks	18°C	30°C	10	30°C	pH5.0
9	KSI	0.002329	0.5 mM, in 20% methanol Stored in thermomixer at 30°C during experiment Substrate from freezer stocks	18°C	30°C	15	30°C	pH5.0
10	KSI	0.0006828	0.5 mM, in 20% methanol Stored in thermomixer at 30°C during experiment. Substrate from freezer stocks.	18°C	30°C	15	30°C	pH5.0
11	KSI	0.0006828	0.5 mM, in 28% methanol Stored in thermomixer at 30°C during experiment	18°C	30°C	15	30°C	pH5.0
12	KSI	0.0006828	0.5 mM, in 28% methanol Stored in thermomixer at 30°C during experiment	18°C	30°C	15	30°C	pH5.0

Table 11 Ctd. Summarised conditions for deuterium exchange assays

Experiment number	Enzyme	[Enzyme] (mg/ml)	Substrate conditions (Final concentration, solvent)	Forward exchange T	Incubation T (Reverse exchange)	Incubation period (minutes)	Assay T	Buffer pH
13	KSI	0.003233	0.5 mM in 28% methanol Stored in thermomixer at 30°C during experiment	18°C	30°C	15	30°C	pH5.0
1	MalL	0.001733	2 mM in assay buffer	18°C	30°C	15	30°C	pH6.0
2	MalL	0.0014	2 mM in assay buffer	18°C	30°C	15	30°C	pH6.0 size exclusion buffer, pH7.0 exchange and assay buffers
3	MalL	0.003621	2 mM in assay buffer	22°C	22°C	10	22°C	pH7.0